

EUDP Slutrapport

Project: Beslutningsstøttesoftware baseret på digitale tvillinger for energioptimering og drift af bygninger

List of Appendices

Appendices included:

1. Gantt Diagram as of September 2024
2. A report on the energy modeling requirements for Digital Twins applications.
3. A report on the data modeling requirements for Digital Twins applications.
4. A feedback report summarizing the key findings of the field interviews with owners and clients
5. Twin4Build Digital Twin Architecture
6. Digital Twin demonstration in OU44 building room
7. Published Article in HVAC Magsinet: Digitale tvillinger I byggeriet
8. Published Paper in Journal of Building Engineering: Component-level re commissioning of a newly retrofitted Danish healthcare building
9. Published Paper in the Proceedings of BSO-VI 2022- Sixth Building Simulation and Optimisation Virtual Conference: A Modular Thermal Space Coupling Approach for Indoor Temperature Forecasting Using Artificial Neural Networks
10. Published Paper in the Proceedings of Building Simulation Applications BSA 2022: A Fully Automated and Scalable Approach for Indoor Temperature Forecasting in Buildings Using Artificial Neural Networks
11. Published Paper in the Proceedings of IEEE ACTEA 2023 conference: A Digital Twin Platform for Energy Efficient and Smart Buildings Applications.
12. Published Paper in the Proceedings of Building Simulation Conference BS 2023: Adaptable and Scalable Energy Modeling of Ventilation Systems as Part of Building Digital Twins.

13. Published Paper in the Proceedings of Building Simulation Conference BS 2023: Implementation and demonstration of an automated energy modeling framework for scalable and adaptable building digital twins based on the SAREF ontology.
14. Published Paper in the Proceedings of Building Simulation Applications Conference BSA 2024: Data-driven digital twining of ventilation systems for performance optimization: A university building case study.
15. Published Paper in the Proceedings of the International Conference on Sustainable Energy Technologies (SET23): Using Digital Twins to investigate the need for active cooling systems in Danish public buildings.
16. Published Paper in Energy and Buildings Journal: An Ontology-Based Innovative Energy Modeling Framework for Scalable and Adaptable Building Digital Twins.
17. Published Paper in the Proceedings of the Sixth International Conference on Efficient Building Design: Systematic Commissioning and Performance Testing of Next Generation Building Management Systems.
18. Submitted Paper to the International Journal of Building Simulation: Development and Demonstration of a Digital Twin Platform Leveraging Ontologies and Data-driven Simulation Models.
19. Submitted Paper to Energy and building Journal: Automated Model Generation and Parameter Estimation of Building Energy Models Using an Ontology-based Framework.
20. Submitted Paper to Applied Energy Journal: Large-scale Field Demonstration of an Interoperable and Ontology-Based Energy Modeling Framework for Building Digital Twins.

Energy Modeling Requirements for Digital Twins Applications

Jakob Bjørnskov and Muhyiddine Jradi

Center for Energy Informatics, University of Southern Denmark

Abbreviation	Definition
HVAC	Heating, Ventilation, and Air Conditioning
DT	Digital twin
IFC	Industry Foundation classes
ISO	International Organization for Standardization
ANN	Artificial Neural Network

Introduction

In this specification, the fundamental requirements for a DT-integrated energy model are formulated. This includes both input requirements concerning the static design data for installed cooling, heating, and ventilation systems as well as dynamic time-series data collected on-site from sensors and energy meters. These requirements are highly dependent on what types of services the Digital Twin solution is targeting. Therefore, this specification has been developed such that the Digital Twin solution delivers the following three fundamental service blocks:

1. Collecting, integrating, and managing data generated in the building by various sources including sensors, meters, and IoT devices within a flexible, effective, and user-friendly open standard context information model.
2. Real-time performance monitoring and automated continuous commissioning and fault detection enabling smarter facility management, efficient building operation, timely identification of malfunctions, the anticipation of risks, and heading off problems before they occur.
3. Strategy and planning support over the whole building life cycle through informing optimal decisions, running 'What if?' scenarios in advance, predicting impacts, and testing different control and management patterns in a virtual environment.

While service block 1 is mostly related to the context information model, blocks 2 and 3 impose certain requirements to the energy model design. In recent years, successful implementations of continuous commissioning as listed in service block 2 have been reported numerous times with different modeling approaches. Typically, fault detection methods are divided into two categories. The first category is model-based methods, which rely on comparisons between the simulated behavior of a system with actual measurements from a system. The second category is model-free methods, which typically utilizes only operational data collected from a system to find patterns that distinguish faulty operation from normal operation. Although model-free methods are easier to implement, their effectiveness is also often very limited in comparison to model-based methods when applied on large-scale distributed systems such as HVAC systems [1]. Therefore, most efforts in fault detection and continuous commissioning have been through model-based methods, often with the use of white-box simulation tools [2, 3, 4, 5]. Over the years much effort has been invested in developing these energy simulation tools, e.g. EnergyPlus, DOE-2, and TRNSYS [6, 7, 8]. However, as is the case for the majority of large-scale white-box models, they require extensive information on the building and a substantial amount of resources and time to develop. Taking EnergyPlus as an example, a typical workflow consists of three time-consuming phases before an accurate energy model is obtained. First, the building geometry is defined with detailed geometry data from floor plans, cross-sections, 3D models, etc. Second, HVAC systems design data must be specified along with material and thermal properties of the envelope and occupancy, lighting, and equipment schedules of the spaces. Finally, to achieve acceptable prediction accuracy, the model must be calibrated using actual consumption data from the building, a process that often relies on a manual trial-and-error approach that requires domain-specific expertise from the modeler [9]. In addition to the high amount of resources and time that must be dedicated to the manufacturing of these high-resolution white-box models, high computational costs and execution time is also a well-known challenge, which makes them unsuitable for operational optimization purposes [10, 11].

As a result of these shortcomings of white-box models, data-driven modeling methods have recently gained increasing popularity for building modeling and simulation. Data-driven modeling methods are typically divided into black-box and grey-box methods. Black-box models are purely data-driven that generally assumes no domain-specific model-structure in finding appropriate functional relationships that maps system inputs to system outputs using data. Commonly used black-box models are Artificial Neural Networks (ANN) [12, 13], Support Vector Machines (SVM) [14], and Decision Trees [15].

Typically, black-box models are much lighter than white-box models in terms of computational complexity. Therefore, a popular approach to deal with the model complexity of white-box models is to train black-box *surrogate models* to mimic the behavior of white-box simulation engines. As an attempt to lower computational demands of exploring different building design scenarios, Geyer et al. [16] trained machine-learning surrogate models on synthetic EnergyPlus-generated data to represent parametrized building

components such as walls, floors, windows, etc. These components were then used to predict both the static and dynamic energy consumption response to certain design configurations such as heat transfer coefficients, window to wall ratios, surface areas, etc. As also reported by Geyer et al., these models can achieve very high prediction similarity while being much lighter and faster than their white-box counterparts [17]. However, while this approach addresses the problem of model complexity, it does not address the problem of modeling time. In fact, additional time must be allocated to the already time-consuming white-box modeling workflow. Furthermore, surrogate models can only be as accurate as the model it is trained to mimic, i.e. the assumptions and inaccuracies of the simulation engine are inherently built into the surrogate model. Therefore, surrogate models are mostly useful for building design problems where no operational data is yet available [18].

For existing buildings with available historical data, black-box models have been used extensively for both heating and electricity load forecasting [19, 20, 21], usually in the form of *monolithic* models. Within the building domain these types of models usually refer to whole-building models that do not incorporate any knowledge of the modeled building, besides historical weather and operational data such as heating or electricity consumption. While this offers convenient and fast implementation, there are also certain downsides to consider. If the building operation or design is changed, e.g. if system setpoints are changed or if the building is retrofitted, these models have to be re-trained on new data.

One of the often highlighted drawbacks of this approach is the lack of interpretability, i.e. the model structure and the obtained parameters have no direct physical meaning. This is opposed to the second data-driven category called grey-box models. These types of models are a mix of white-box and black-box models as they take advantage of physical principles to define a general model structure, while the unknown parameters of this model structure are estimated through parameter identification techniques using data. An example of grey-box models in buildings is the thermal Resistor Capacity (RC) model, a thermal analogue to electrical circuits [22, 23]. However, grey-box modeling is a general concept that can also be applied to individual building components such as heat pumps, recovery units, thermal storage tanks, cooling coils, etc. as demonstrated by Afram et al. [24].

Systems and components

To meet the defined service blocks from the energy model perspective, a modular and component-based modeling approach is proposed. The energy model is thus split into a set of building blocks that can be automatically assembled and adapted, based on the specific use-case.

This concept has been conceptualized in the context of a generic building energy system as shown in Figure 1. As seen, the concept can be divided into three main parts; the physical system, the digital twin, and a flow of data linking these two systems. The fundamental task of the digital twin is to mimic the behavior of the physical system as closely as possible, and it should thus consist of key components as also shown in Figure 1. Concerning service block 2, this approach will make it possible to continuously monitor and compare the performance of actual components in real-time with their digital counterparts in the digital twin environment. Furthermore, a highly modular and flexible model allows for the interchange of component models during the entire building lifecycle, which is necessary to maintain the Digital Twin as a close replica

of the actual building. Therefore, the component models should be accurate enough to meet the defined services to a satisfactory level, but simple enough to make the model development as robust and automated as possible.

When modeling buildings, it is common practice to take advantage of the major difference in physical time-constants that are associated with the HVAC system and the building envelope and spaces. The HVAC system is often governed by fast-moving dynamics that reaches equilibrium within seconds or minutes, while the large thermal inertia of the building envelope and spaces has very slow-moving dynamics. Therefore, it can often be justified to model the HVAC systems as quasi-steady-state systems, which decreases model complexity and lowers the computational burden considerably [25, 26]. This approach has also been adopted by the widely used building simulation software EnergyPlus [27].

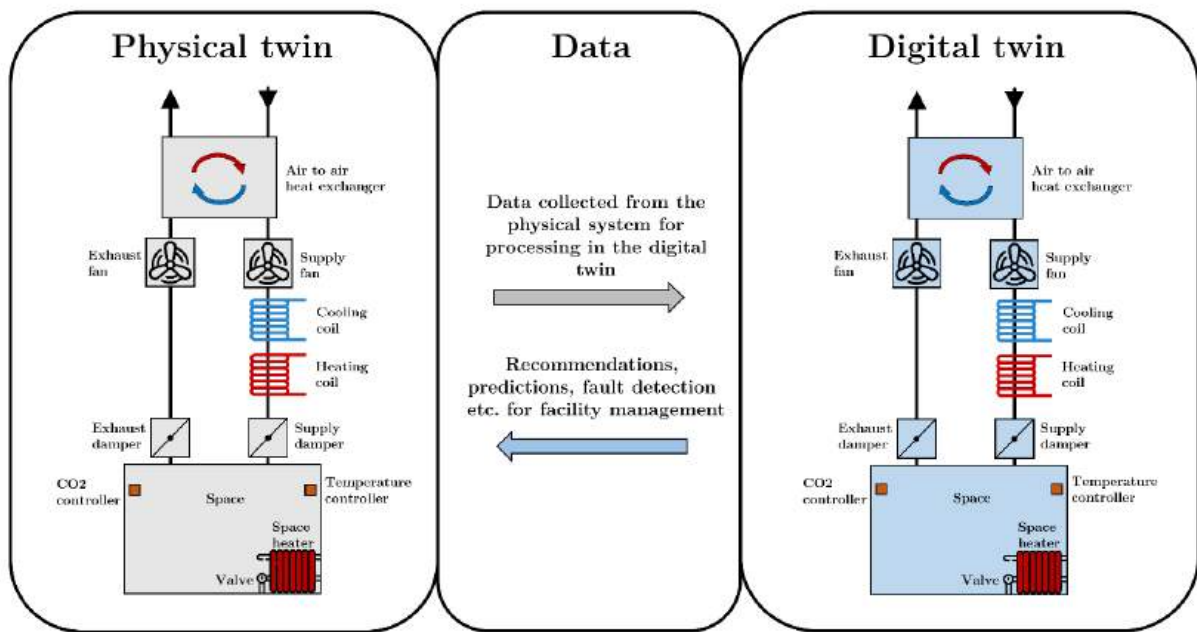


Figure 1: Digital Twin framework shown in the context of a generic building energy system with different components.

To comply with international standards and maximize the impact and applicability of this specification, it has been decided that this specification will follow the SAREF4BLDG extension to the SAREF ontology [28]. SAREF4BLDG is based on the ISO-published Industry Foundation Classes (IFC) [29] and has 72 classes and 179 object properties that represents typical appliances and devices in buildings. Hence, this ontology allows an almost direct mapping between IFC and SAREF for most of these objects within the building domain. Therefore, a review of the available IFC-files for the studied buildings will be conducted later in this document to investigate how much of the required information ifc-files can provide.

With basis in the components depicted in Figure 1, a description of the components and their properties adopted from the SAREF4BLDG ontology is given in Table 1. The superscripts *op* and *dp* stands for *object property* and *data property*, respectively. A complete list of available components and a description of the listed properties can be found in the SAREF4BLDG documentation [28].

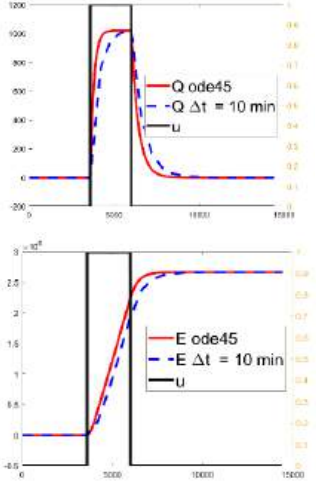
Table 1: Component description and related properties adopted from the SAREF4BLDG ontology [28].

Component	Description	Properties
Valve	<i>A valve is used in a building services piping distribution system to control or modulate the flow of the fluid.</i>	closeOffRating ^{op} flowCoefficient ^{op} size ^{op} testPressure ^{op} valveMechanism ^{dp} valveOperation ^{dp} valvePattern ^{dp} workingPressure ^{op}
Space heater	<i>Space heaters utilize a combination of radiation and/or natural convection using a heating source such as electricity, steam or hot water to heat a limited space or area. Examples of space heaters include radiators, convectors, baseboard and finned-tube heaters.</i>	bodyMass ^{op} energySource ^{dp} heatTransferDimension ^{dp} heatTransferMedium ^{dp} numberOfPanels ^{dp} numberOfSections ^{dp} outputCapacity ^{op} placementType ^{dp} temperatureClassification ^{dp} thermalEfficiency ^{op} thermalMassHeatCapacity ^{op}
Coil	<i>A coil is a device used to provide heat transfer between non-mixing media. A common example is a cooling coil, which utilizes a finned coil in which circulates chilled water, antifreeze, or refrigerant that is used to remove heat from air moving across the surface of the coil. A coil may be used either for heating or cooling purposes by placing a series of tubes (the coil) carrying a heating or cooling fluid into an airstream. The coil may be constructed from tubes bundled in a serpentine form or from finned tubes that give a extended heat transfer surface.</i>	airFlowRateMax ^{op} airFlowRateMin ^{op} nominalLatentCapacity ^{op} nominalSensibleCapacity ^{op} nominalUa ^{op} operationTemperatureMax ^{op} operationTemperatureMin ^{op} placementType ^{dp}
Air to air heat recovery	<i>An air-to-air heat recovery device employs a counter-flow heat exchanger between inbound and outbound air flow. It is typically used to transfer heat from warmer air in one chamber to cooler air in the second chamber (i.e., typically used to recover heat from the conditioned air being exhausted and the outside air being supplied to a building), resulting in energy</i>	hasDefrost ^{dp} heatTransferTypeEnum ^{dp} operationTemperatureMax ^{op} operationTemperatureMin ^{op} primaryAirFlowRateMax ^{op} primaryAirFlowRateMin ^{op} secondaryAirFlowRateMax ^{op} secondaryAirFlowRateMin ^{op}

	<i>savings from reduced heating (or cooling) requirements.</i>	
Fan	<i>A fan is a device which imparts mechanical work on a gas. A typical usage of a fan is to induce airflow in a building services air distribution system.</i>	capacityControlType ^{dp} motorDriveType ^{dp} nominalAirFlowRate ^{op} nominalPowerRate ^{op} nominalRotationSpeed ^{op} nominalStaticPressure ^{op} nominalTotalPressure ^{op} operationTemperatureMax ^{op} operationTemperatureMin ^{op} operationalRiterial ^{op}
Damper	<i>A damper typically participates in an HVAC duct distribution system and is used to control or modulate the flow of air.</i>	airFlowRateMax ^{op} bladeAction ^{dp} bladeEdge ^{dp} bladeShape ^{dp} bladeThickness ^{op} closeOffRating ^{op} faceArea ^{op} frameDepth ^{op} frameThickness ^{op} frameType ^{dp} leakageFullyClosed ^{op} nominalAirFlowRate ^{op} numberOfBlades ^{dp} openPressureDrop ^{op} operation ^{dp} operationTemperatureMax ^{op} operationTemperatureMin ^{op} orientation ^{dp} temperatureRating ^{op} workingPressureMax ^{op}
Controller	<i>A controller is a device that monitors inputs and controls outputs within a building automation system. A controller may be physical (having placement within a spatial structure) or logical (a software interface or aggregated within a programmable physical controller).</i>	

With basis in the presented components in Table 1, potential mathematical models have been identified in the literature to establish an input-output relationship for each component. The main selection criterias have been that the models should be validated, simple and robust to be suited for parameter estimation.

Table 2: Potential mathematical models of components and their associated inputs, outputs and parameters.

Component	Mathematical Model	Inputs	Outputs	Parameters
Valve [30]	$\tilde{u} = \frac{u}{\sqrt{u^2(1 - N_v) + N_v}}$ $\dot{m}_w = \tilde{u} \dot{m}_{w,max}$	u Valve position	$\dot{m}_w$ Water massflow through the valve	$N_v \in [0, 1]$ Valve authority $\dot{m}_{w,max}$ Maximum waterflow through the valve
Space heater [31] Adapted to discrete time domain for 1 lumped element with simplified convective term.	$K_1 = \frac{T_{w,in} \dot{m}_w c_{p,w} + UA T_z}{C_r} + \frac{T_{r,t-1}}{\Delta t}$ $K_2 = \frac{1}{\Delta t} + \frac{UA + \dot{m}_w c_{p,w}}{C_r}$ $T_{r,t} = \frac{K_1}{K_2}$ $\dot{Q}_r = \dot{m}_w c_{p,w} (T_{w,in} - T_{r,t})$ “The use of numerical integration in a long time simulation is a cause for some concern due to the potential build-up of truncation error over many time steps. In this case, the finite difference approximation is of low order that further aggravates the problem. However, the cyclic nature of building energy simulations should cause truncation errors to cancel over each daily cycle so that no net accumulation of error occurs, even over many days of simulation (Walton, 1990)” – Energy Plus Engineering reference – p. 8 equation 9 	$\dot{m}_w$ Water massflow through the space heater $T_{w,in}$ Inlet temperature of the water T_z Indoor temperature of the space in which the space heater is located	$\dot{Q}_r$ Heating output of the space heater	$c_{p,w}$ Specific heat capacity of water. UA Overall heat transfer coefficient for the space heater Δt Timestep constant C_r Heat capacity of the radiator
Heating coil	$\dot{Q}_{hc} = \begin{cases} \dot{m}_a c_{p,a} (T_{a,set} - T_{a,in}) & \text{if } T_{a,in} < T_{a,set} \\ 0 & \text{if } T_{a,in} \geq T_{a,set} \end{cases}$	$\dot{m}_a$ Air massflow through the heating coil $T_{a,set}$ Air temperature setpoint of the supply flow $T_{a,in}$ Inlet air temperature of the heating coil	$\dot{Q}_{hc}$ Heating output of the heating coil	$c_{p,a}$ Specific heat capacity of air

Cooling coil	$Q_{cc} = \begin{cases} \dot{m}_a c_{p,a} (T_{a,in} - T_{a,set}) & \text{if } T_{a,in} > T_{a,set} \\ 0 & \text{if } T_{a,in} \leq T_{a,set} \end{cases}$	$\dot{m}_a$ Air massflow through the cooling coil $T_{a,set}$ Air temperature setpoint of the supply flow $T_{a,in}$ Inlet air temperature of the cooling coil	$\dot{Q}_{cc}$ Cooling output of the cooling coil	$c_{p,a}$ Specific heat capacity of air
Air to air heat recovery [27]	$\varepsilon_{75\%} = \begin{cases} \varepsilon_{75\%,h} & \text{if } T_{a,sup,in} < T_{a,exh,in} \\ \varepsilon_{75\%,c} & \text{if } T_{a,sup,in} > T_{a,exh,in} \end{cases}$ $\varepsilon_{100\%} = \begin{cases} \varepsilon_{100\%,h} & \text{if } T_{a,sup,in} < T_{a,exh,in} \\ \varepsilon_{100\%,c} & \text{if } T_{a,sup,in} > T_{a,exh,in} \end{cases}$ $f_{flow} = \frac{\frac{1}{2}(\dot{m}_{a,sup} + \dot{m}_{a,exh})}{\dot{m}_{a,max}}$ $\varepsilon_{op} = \varepsilon_{75\%} + (\varepsilon_{100\%} - \varepsilon_{75\%}) \frac{f_{flow} - 0.75}{1 - 0.75}$ $\dot{C}_{sup} = \dot{m}_{a,sup} c_{p,a}$ $\dot{C}_{exh} = \dot{m}_{a,exh} c_{p,a}$ $\dot{C}_{min} = \min(\dot{C}_{sup}, \dot{C}_{exh})$ $T_{a,sup,out} = T_{a,sup,in} + \varepsilon_{op} (T_{a,exh,in} - T_{a,sup,in}) \frac{\dot{C}_{min}}{\dot{C}_{sup}}$	$\dot{m}_{a,sup}$ Supply air massflow through the heat recovery unit $\dot{m}_{a,exh}$ Exhaust air massflow through the heat recovery unit $T_{a,sup,in}$ Supply inlet air temperature of the heat recovery unit $T_{a,exh,in}$ Exhaust inlet air temperature of the heat recovery unit	$T_{a,sup,out}$ Supply outlet air temperature of the heat recovery unit	$\varepsilon_{75\%,h} \in [0, 1]$ Effectiveness of the heat recovery unit at 75% airflow in heating mode $\varepsilon_{75\%,c} \in [0, 1]$ Effectiveness of the heat recovery unit at 75% airflow in cooling mode $\varepsilon_{100\%,h} \in [0, 1]$ Effectiveness of the heat recovery unit at 100% airflow in heating mode $\varepsilon_{100\%,c} \in [0, 1]$ Effectiveness of the heat recovery unit at 100% airflow in cooling mode $\dot{m}_{a,max}$ Maximum air massflow on either the supply or exhaust side $c_{p,a}$ Specific heat capacity of air
Fan [27]	$f_{flow} = \frac{\dot{m}_a}{\dot{m}_{a,max}}$ $f_{pl} = c_1 + c_2 f_{flow} + c_3 f_{flow}^2 + c_4 f_{flow}^3$ $\dot{W}_{fan} = \frac{f_{pl} \dot{m}_{a,max} \Delta P_{max}}{\eta_{tot} \rho_a}$	$\dot{m}_a$ Air massflow through the fan	$\dot{W}_{fan}$ Electrical power consumed by the fan	$c_1 \in [0, 1]$ Power coefficient 1 $c_2 \in [0, 1]$ Power coefficient 2 $c_3 \in [0, 1]$ Power coefficient 3 $c_4 \in [0, 1]$ Power coefficient 4 $\dot{m}_{a,max}$ Maximum air massflow ΔP_{max} Pressure rise across the fan at maximum air flow $\eta_{tot} \in [0, 1]$ Total efficiency of the fan ρ_a Density of air

Damper [32]	$\dot{m}_a = ae^{bu} + c$	u Damper position	$\dot{m}_a$ Air massflow through the damper	$a \in \mathbb{R}$ $b \in \mathbb{R}$ $c \in \mathbb{R}$
Controller [33]	$e_t = y_{set} - y_t$ $u_t = u_{t-1} + K_p(e_t - e_{t-1}) + K_i\Delta t + \frac{K_d}{\Delta t}(e_t - 2e_{t-1} + e_{t-2})$	y_{set} Setpoint for the controlled variable y_t Value of the controlled variable	u_t Output signal	K_p Proportional coefficient K_i Integral coefficient K_d Derivative coefficient

The parameters for the models in Table 2 can either be given through design specifications or calculated through parameter estimation techniques, if time-series data is available for the inputs and outputs. In the case where input-output data can be continuously provided, parameter estimation can be performed on a regular interval to attain increasing accuracy on the different digital twin components. Additionally, each physical component can be monitored closely, by comparing its operation with its digital replica to provide fault detection and continuous commissioning capabilities on a component-level. If neither static data or operational input-output data is available, standard values must be assumed. As seen from Table 2, many of the parameters are very model specific and are therefore not part of the SAREF4BLDG ontology properties. Therefore, these parameters must either be added as extensions to the ontology or treated as internal parameters only accessible by the energy model.

System topology

To present requirements in relation to the system topology information that is required for a DT-integrated energy model, the SAREF4SYST ontology is considered. SAREF4SYST is aimed at providing a generic framework for representing the general topology of systems and how they interact [33]. Through SAREF4SYST, the system topology can be expressed through three generic classes, *System*, *Connection*, and *ConnectionPoint* and nine applicable properties as shown in Figure 2. As seen in the figure, a *System* can be a subsystem of other systems specified through the property *subSystemOf* or have subsystems itself specified through *hasSubSystem*. A *System* can be connected to other systems through the *Connection* and *ConnectionPoint* classes. The type of information exchanged between systems can be specified as a property of the *ConnectionPoint*.

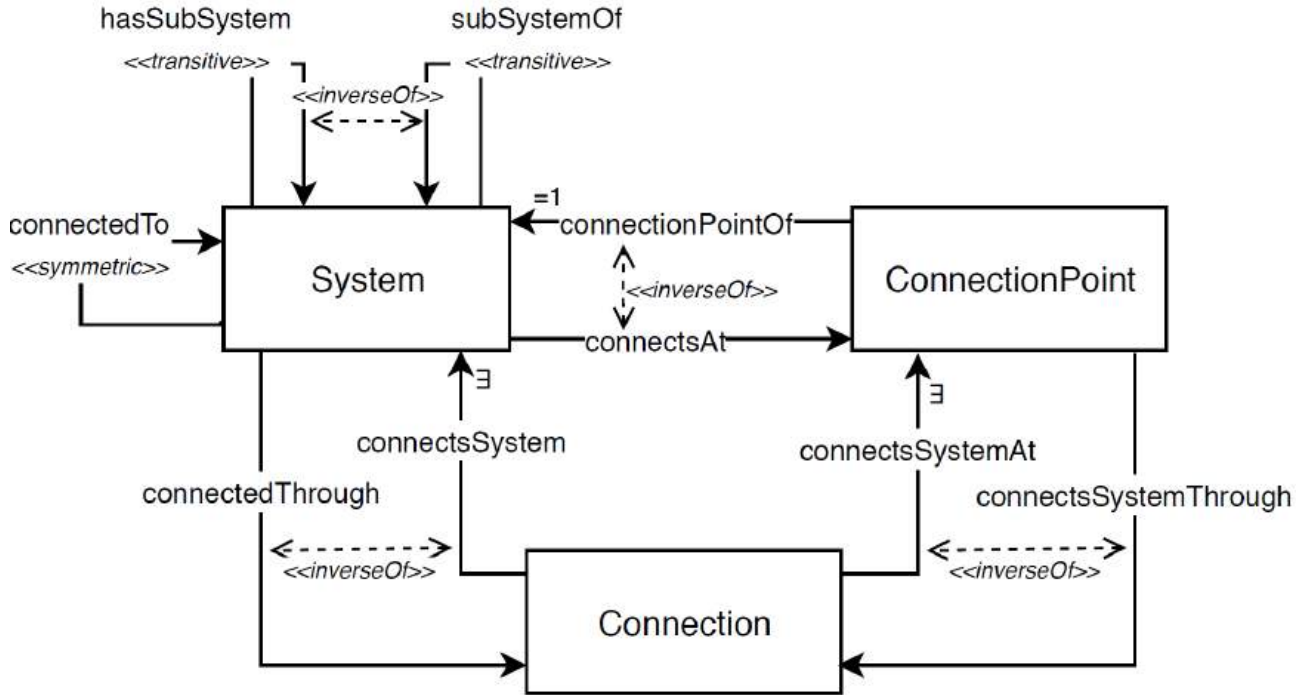


Figure 2: SAREF4SYST overview with the three classes and their nine applicable properties [33].

As an example on the use of these concepts, consider Figure 3. Here, various SAREF4BLDG components are shown which all inherit from the *SAREF4SYST:System* superclass (grey) and can thus interact and communicate through the SAREF4SYST ontology. It should be noted that the components shown under the classes *SAREF4BLDG:DistributionControlDevice* and *SAREF4BLDG:DistributionFlowDevice* are further categorized under different subclasses in the SAREF4BLDG ontology. However, these subclasses are omitted for clarity along with the *Connection* and *ConnectionPoint* instances from the SAREF4SYST ontology that are used to represent information flows between systems.

In Figure 3, two examples of a simple information flows between components is provided. The first information flow relates a space, an indoor temperature controller, a valve, and a space heater. From the space, the current indoor temperature is sent to an indoor temperature controller which responds by calculating an appropriate valve position to drive the indoor temperature towards a specified setpoint. This signal is sent to the valve component which outputs the current water massflow through the space heater. Finally, this massflow is sent to the space heater component for energy consumption calculations. For the second example, the space component sends the current CO₂-concentration to the CO₂-controller, which sends a new position signal to the damper. The damper component then calculates the current air flow and forwards this to the heat recovery unit.

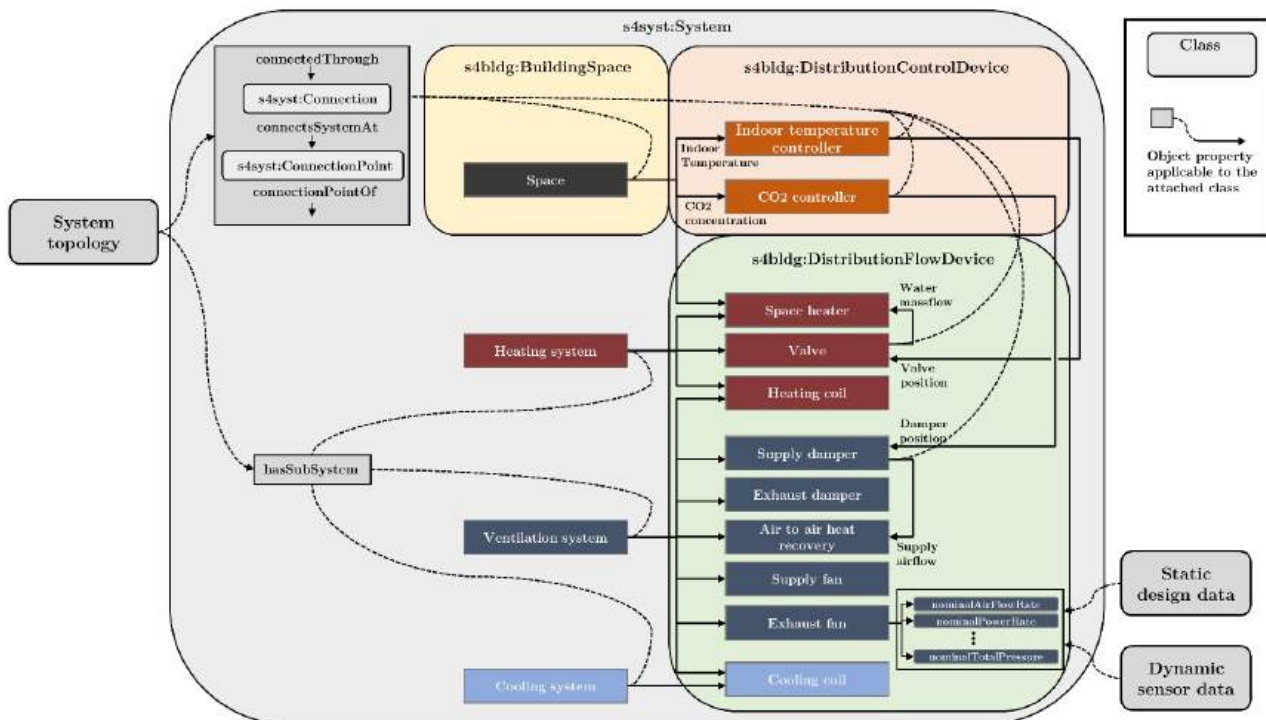


Figure 3: Data structure hierarchy showing examples of required relations and properties of different components.

In Figure 3, examples on the use of the *hasSubSystem* property is also shown. As seen, the different low-level components can be sorted under higher-level systems such as heating, ventilation, and cooling systems, allowing for performance monitoring and reporting on various levels of the building. By structuring the model in systems and sub-systems, the system performance monitoring can thus be customized and adapted based on the specific set of sensors and meters available in the building, ranging from high-resolution monitoring of individual components to a low-resolution whole-building monitoring of e.g. heating, and cooling consumption. Based on the included examples, it is evident that topology information on *which* components are connected and *how* they are connected is crucial. Furthermore, this information must be available in a standardized format.

Simulation framework

Assuming a data model is available, which follows the defined requirements to topology and design data, a simulation framework for the energy model can be outlined. As shown in Figure 4, the energy model can be divided into two general types of models that interact during simulation; the space models and the HVAC models.

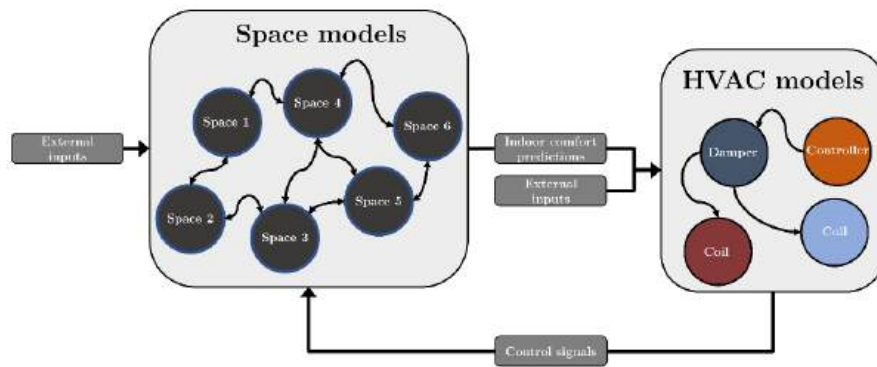


Figure 4: Overview of the proposed simulation framework for the energy model divided into two general types of models, space models and HVAC models.

The space models takes external inputs such as weather data and occupancy data along with Control signals from the HVAC system such as damper or valve positions. The drawn connections between spaces represent the necessary information flow to accurately model the interdependence of temperatures in adjacent spaces. However, the specific configuration of these connections will depend on the choice of space model. The space models will then predict indoor comfort indicators such as temperature and CO₂ concentration for the space they represent. This output is given as input to the HVAC model, which will actuate the different HVAC components to meet predefined setpoints given as external inputs. The specific components included in the HVAC model depends on the specific building modeled. However, based on the system shown in Figure 1, an example of a possible configuration of components and information flows is shown in Figure 5.

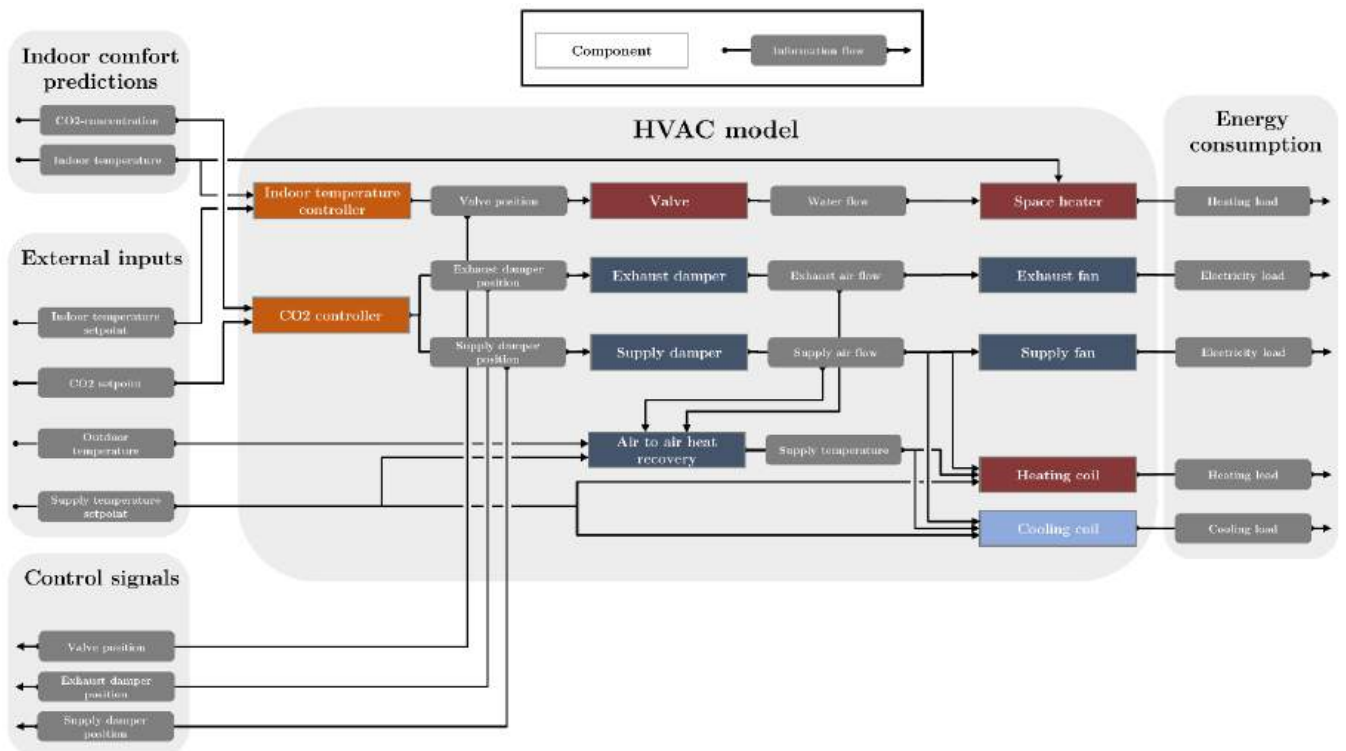


Figure 5: Component-based HVAC model with required inputs and outputs during simulation.

The configuration of Figure 5 is based on the required inputs, outputs and relations between components defined in Table 2. As shown, the temperature and CO₂ controllers each receives a setpoint and an actual value to output a control signal. These signals are fed to valves and dampers that outputs equivalent flows, which are further sent to space heaters, fans, coils etc. Additionally, the signals are also fed back to the space model, thereby closing the loop. Depending on the specific use-case, relevant operational data can be reported or used for performance monitoring, e.g. heating loads, electricity loads, cooling loads, waterflows, airflows, efficiencies, damper positions etc.

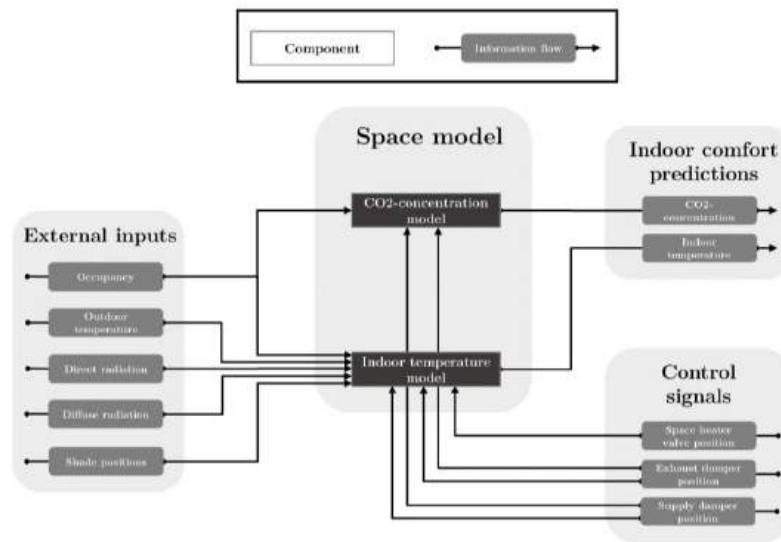


Figure 6: Component-based space model with required inputs and outputs during simulation.

Figure 6 shows an example of a space model with two submodels; an indoor temperature model and an indoor CO₂-concentration model. The model responsible for predicting the indoor temperature receives the control signals from the HVAC system in the form of valve and damper positions. In addition, external inputs such as weather data in the form of outdoor temperature and irradiation levels are also given as input together with occupancy. Because the available inputs for each room can differ considerably between cases, the temperature model is modeled as a black-box.

Sensors and meters

To develop reliable models that can predict the room temperature in buildings, preliminary requirements to sensors and meters in the ideal case have been outlined in Table 3. The list of sensors have divided into three categories; weather data, rooms, and components. For each category, several sensors are listed along with the expected measurement frequency.

Table 3: Requirements to sensor and meter data in the ideal case.

Type	Sensors/Meters	Measurement frequency
Weather data	- Outdoor temperature - Solar irradiance	<1 hour

Rooms	• Indoor temperature • Space heater valve positions or fluid massflow • Damper positions or air massflow • Shade positions • Occupancy count, CO₂-concentration, or Passive Infrared sensor (PIR)	<10 minutes
Components	• Inputs and outputs from Table 2 for all installed components in the building	<10 minutes

Table 4: Minimum requirements to sensor and meter data.

Type	Sensors/Meters	Measurement frequency
Weather data	• Outdoor temperature • Solar irradiance	<1 hour
Rooms	• Indoor temperature • Space heater valve positions or fluid massflow • Damper positions or air massflow	<30 minutes
Components	• Total heating consumption • Total cooling consumption • Total electricity consumption	<Daily

In Table 4 the minimum requirements are outlined. As seen, requirements to the available sensors and the resolution of the collected data are much lower here. Generally, when using data-driven approaches, higher amounts of data will yield better models and vice versa. This also means that there is a tipping point where too little data is available and data-driven methods are no longer reliable. This concept is shown in Figure 7 (in the context of fault detection), where the reliability of data-driven methods is compared to the reliability of physics-based methods, based on the amount of data available.

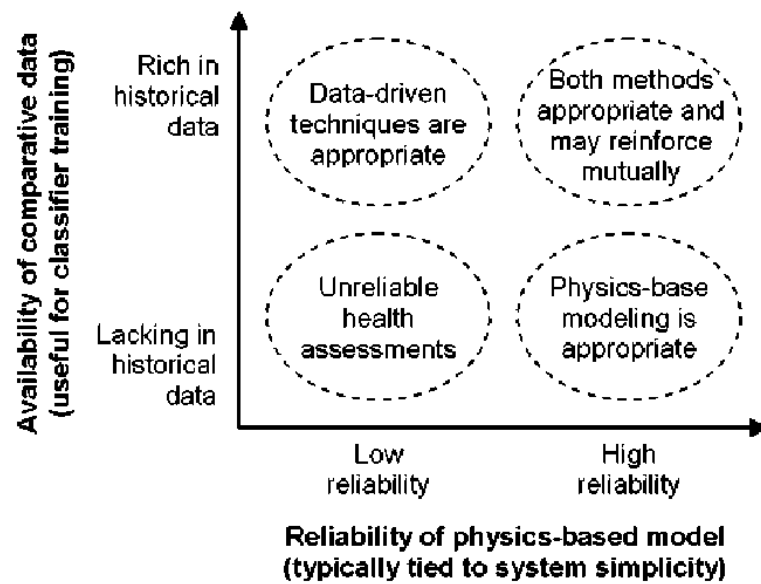


Figure 7: Reliability of data-driven methods compared to the reliability of physics-based methods based on the amount of data available [35].

Review of IFC files

A review has been conducted on the IFC-files for the case study buildings studied in the project. A visual representation of two case-study IFC-files is shown in Figure 8. The building seen on the left is a university teaching building located in Odense and the building seen on the right is a public library located in Aarhus.

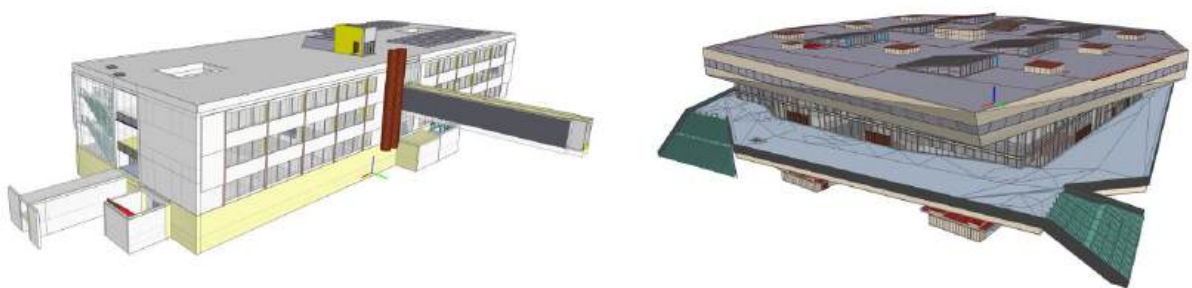


Figure 8: IFC architecture files for the teaching building in Odense (left) and the public library in Aarhus (right).

In this review both the quality and detail of the ifc-files is evaluated based on whether the required information is available and where it can be found. Specifically, ifc-files representing both architecture, heating, cooling and ventilation systems for each case study building has been investigated to find how well each component class presented in Table 1 is represented in relation to the defined properties. As already mentioned, the component properties defined in SAREF inherits directly from the native *IfcPropertySets* in the IFC standard with the purpose of enabling a direct mapping between IFC and SAREF. All native *IfcPropertySets* have the prefix “Pset_”, with the name of the property set following. However, in the IFC

standard, the user also has the option of attaching custom property sets to IFC entities. Even though this adds flexibility to IFC development, it also makes it harder to standardize and automate extraction of information from IFC. As shown in the following review, majority of the properties of the investigated components are represented through custom *IfcPropertySets*, which introduces certain challenges in extracting information from IFC autonomously. Furthermore, many of the investigated components were not found in the IFC files. Therefore, there is a need for either different data sources or more comprehensive IFC-files. If the first option is chosen, the different data sources could be managed e.g. through the Linked Data method.

Information is available in the expected format.	
Information is available, but is either represented through an unexpected <i>IfcEntity</i> , or is part of a custom <i>IfcPropertySet</i> .	
Information is not available.	

Space	Teaching building, Odense	Library building, Aarhus	Hospital, Bergen	Expected
IfcEntity	IfcSpace	IfcSpace	IfcSpace	IfcSpace
Valve	Teaching building, Odense	Library building, Aarhus	Hospital, Bergen	Expected
IfcEntity				IfcValve
Space heater	Teaching building, Odense	Library building, Aarhus	Hospital, Bergen	Expected
IfcEntity	IfcBuildingElementProxy	IfcBuildingElementProxy	IfcFlowTerminal	IfcSpaceHeater
bodyMass				Pset_SpaceHeaterTypeCommon
energySource				
heatTransferDimension				
heatTransferMedium	Custom IfcPropertySet			
numberOfPanels	Custom IfcPropertySet		Custom IfcPropertySet	
numberOfSections				
outputCapacity	Custom IfcPropertySet		Custom IfcPropertySet	

placementType				
temperatureClassification	Custom IfcPropertySet			
thermalEfficiency				
thermalMassHeatCapacity				
Heating coil	Teaching building, Odense	Library building, Aarhus	Hospital, Bergen	Ideal
IfcEntity				IfcCoil
Cooling coil	Teaching building, Odense	Library building, Aarhus	Hospital, Bergen	Ideal
IfcEntity				IfcCoil
Air to air heat recovery	Teaching building, Odense	Library building, Aarhus	Hospital, Bergen	Ideal
IfcEntity				IfcAirToAirHeatRecovery
Fan	Teaching building, Odense	Library building, Aarhus	Hospital, Bergen	Ideal
IfcEntity				IfcFan
Damper	Teaching building, Odense	Library building, Aarhus	Hospital, Bergen	Ideal
IfcEntity	IfcFlowTerminal		IfcFlowTerminal	IfcDamper
airFlowRateMax				Pset_DamperTypeCommon
bladeAction				
bladeEdge				
bladeShape				
bladeThickness				
closeOffRating				
faceArea				

frameDepth				
frameThickness				
frameType				
leakageFullyClosed				
nominalAirFlowRate	Custom IfcPropertySet		Custom IfcPropertySet	
numberOfBlades				
openPressureDrop				
operation				
operationTemperatureMax				
operationTemperatureMin				
orientation				

References

- [1] F. Xiao and S. Wang, "Progress and methodologies of lifecycle commissioning of HVAC systems to enhance building sustainability," *Renewable and Sustainable Energy Reviews*, vol. 13, pp. 1144-1149, 2009.
- [2] L. Wang, S. Greenberg, J. Fiegel, A. Rubalcava, S. Earni, X. P. a. R. Yin, S. Woodworth and J. Hernandez-Maldonado, "Monitoring-based HVAC commissioning of an existing office building for energy efficiency," *Applied Energy*, vol. 102, pp. 1382-1390, 2013.
- [3] Z. O'Neill, X. Pang, M. Shashanka, P. Haves and T. Bailey, "Model-based real-time whole building energy performance monitoring and diagnostics," *Journal of building performance simulation*, vol. 7, pp. 83-99, 2014.
- [4] E. Markoska, M. Jradi and B. N. Jørgensen, "Continuous Commissioning of Buildings: A Case Study of a Campus Building in Denmark," in *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 2016.
- [5] M. Jradi, K. Arendt, F. Sangogboye, C. Mattera, E. Markoska, M. Kjærgaard, C. Veje and B. Jørgensen, "ObepME: An Online Building Energy Performance Monitoring and Evaluation Tool to Reduce Energy Performance Gaps," *Energy and Buildings*, vol. 166, 2018.
- [6] D. Crawley, L. Lawrie, F. Winkelmann, W. Buhl, Y. Huang, C. Pedersen, R. Strand, R. Liesen, D. Fisher, M. Witte and J. Glazer, "EnergyPlus: creating a new-generation building energy simulation program," *Energy and Buildings*, vol. 33, pp. 319-331, 2001.
- [7] F. C. Winkelmann, B. E. Birdsall, W. F. Buhl, K. L. Ellington, A. E. Erdem, J. J. Hirsch and S. Gates, "DOE-2 supplement: Version 2.1E".
- [8] S. Klein and W. Beckman, "TRNSYS 16: A transient system simulation program: mathematical reference," *TRNSYS*, vol. 5, pp. 389-396, 2007.
- [9] A. Reddy, "Literature review on calibration of building energy simulation programs: Uses, problems, procedure, uncertainty, and tools," *ASHRAE Transactions*, vol. 112, pp. 226-240, 2006.
- [10] J. Reynolds, M. W. Ahmad and Y. Rezgui, "Holistic modelling techniques for the operational optimisation of multi-vector energy systems," *Energy and Buildings*, 2018.
- [11] Y. Chen, M. Guo, Z. Chen, Z. Chen and Y. Ji, "Physical energy and data-driven models in building energy prediction: A review," *Energy Reports*, vol. 8, pp. 2656-2671, 2022.
- [12] F. Mtibaa, K.-K. Nguyen, M. Azam, A. Papachristou, J.-S. Venne and M. Cheriet, "LSTM-based indoor air temperature prediction framework for HVAC systems in smart buildings," *Neural Computing and Applications*, 2020.

- [13] Z. Fang, N. Crimier, L. Scanu, A. Midelet, A. Alyafi and B. Delinchant, "Multi-zone indoor temperature prediction with LSTM-based sequence to sequence model," *Energy and Buildings*, 2021.
- [14] D. B., C. C. and L. S.E., "Applying support vector machines to predict building energy consumption in tropical region," *Energy and Buildings*, vol. 37, pp. 545-553, 2005.
- [15] Z. Yu, F. Haghighat, B. C. Fung and H. Yoshino, "A decision tree method for building energy demand modeling," *Energy and Buildings*, vol. 42, pp. 1637-1646, 2010.
- [16] P. Geyer and S. Singaravel, "Component-based machine learning for performance prediction in building design," *Applied Energy*, vol. 228, pp. 1439-1453, 2018.
- [17] P. Westermann and R. Evins, "Surrogate modelling for sustainable building design – A review," *Energy and Buildings*, vol. 198, pp. 170-186, 2019.
- [18] P. Westermann and R. Evins, "Surrogate modelling for sustainable building design – A review," *Energy and Buildings*, vol. 198, pp. 170-186, 2019.
- [19] S. Idowu, S. Saguna, C. Åhlund and O. Schelén, "Applied machine learning: Forecasting heat load in district heating system," *Energy and Buildings*, vol. 133, pp. 478-488, 2016.
- [20] Y. Guo, J. Wang, H. Chen, G. Li, J. Liu, C. Xu, R. Huang and Y. Huang, "Machine learning-based thermal response time ahead energy demand prediction for building heating systems," *Applied Energy*, vol. 221, pp. 16-27, 2018.
- [21] F. Pallonetto, C. Jin and E. Mangina, "Forecast electricity demand in commercial building with machine learning models to enable demand response programs," *Energy and AI*, vol. 7, pp. 100-121, 2022.
- [22] X. Yu, L. Georges, M. D. Knudsen, I. Sartori and L. Imsland, "Investigation of the Model Structure for Low-Order Grey-Box Modelling of Residential Buildings," *International Building Performance Simulation Association*, 2019.
- [23] H. Park, N. Martaj, M. Ruellan, R. Bennacer and E. Monmasson, "Modeling of a Building System and its Parameter Identification," *Journal of Electrical Engineering and Technology*, vol. 8, pp. 975-983, 2013.
- [24] A. Afram and F. Janabi-Sharifi, "Gray-box modeling and validation of residential HVAC system for control system design," *Applied Energy*, vol. 137, pp. 134-150, 2015.
- [25] P. Li, Z. D. O'Neill and J. E. Braun, "Development of Control-Oriented Models for Model Predictive Control in Buildings," *ASHRAE Transactions*, vol. 119, pp. 8-1Z,2Z,3Z,4Z,5Z,6Z,7Z,8Z, 2013.
- [26] F. Jorissen, M. Wetter and L. Helsen, "Simulation Speed Analysis and Improvements of Modelica Models for Building Energy Simulation," in *11th International Modelica Conference*, Versailles, 2015.
- [27] U.S. Department of Energy, "EnergyPlus Version 9.6.0 Documentation," 2021.
- [28] M. Poveda-Villalón and R. Garcia-Castro, "SAREF extension for building," 13 April 2020. [Online]. Available: <https://saref.etsi.org/saref4bldg/>. [Accessed 7 February 2022].

- [29] International Organization for Standardization, "ISO 16739-1:2018," 1 11 2018. [Online]. Available: <https://www.iso.org/standard/70303.html>. [Accessed 9 2 2022].
- [30] I. Zajic, T. Larkowski, M. Sumislawska, K. Burnham and D. Hill, "Modelling of an Air Handling Unit: A Hammerstein-bilinear Model Identification Approach," in *International Conference on Systems Engineering*, Las Vegas, 2011.
- [31] T. H. Pedersen, R. E. Hedegaard, K. F. Kristensen, B. Gadgaard and S. Petersen, "The effect of including hydronic radiator dynamics in model predictive control of space heating," *Energy and Buildings*, vol. 183, pp. 772-784, 2019.
- [32] G. Huang, "Model predictive control of VAV zone thermal systems concerning bi-linearity and gain nonlinearity," *Control Engineering Practice*, vol. 19, no. 7, pp. 700-710, 2011.
- [33] M. Lefrançois, "SAREF4SYST: an extension of SAREF for typology of systems and their inter-connections," 06 06 2019. [Online]. Available: <https://saref.etsi.org/saref4syst/>. [Accessed 20 04 2022].
- [34] A. Afram and F. Janabi-Sharifi, "Review of modeling methods for HVAC systems," *Applied Thermal Engineering*, no. 67, pp. 507-519, 2014.
- [35] M. Jafari, A. Ghofrani, E. Zaidan and A. Abulibdeh, "Improving Building Energy Footprint and Asset Performance Using Digital Twin Technology," in *Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction*, 2021.

Data Modeling Requirements for Digital Twins Applications

Project Acronym: Twin4Build

Project Full Title:

Grant Agreement:

Project Duration:

Deliverable D2.2 Requirements Specification

Work Package: **WP2**

Task(s):

Lead Beneficiary:

Due Date:

Submission Date:

Deliverable Status:

Deliverable Type:

Dissemination Level:

File Name:

xx.pdf

Authors List

Leading Author (Editor)				
Surname	First Name	Beneficiary	Contact email	
Damm	Peter		pda@kmd.dk	
Co-authors (in alphabetic order)				
No	Surname	First Name	Beneficiary	Contact email
1.	Bonne	Mads		mbq@kmd.dk
2.	Kołodziej	Bartłomiej		bkd@kmd.dk
3.		i		
4.				
5.				
Reviewers List				
List of Reviewers (in alphabetic order)				
No	Surname	First Name	Beneficiary	Contact email
1.	Hargaard	Thorsten		teo@kmd.dk

Table of Contents

Executive Summary	5
Introduction	5
Approach.....	6
Methodology	6
TWIN4BUILD requirements specification	7
Building domain Requirements (TWIN4BUILD-R-BLD).....	7
Energy Domain Requirements (TWIN4BUILD-R-ENR)	19
Common business requirements (TWIN4BUILD-R-BUS, T.B.D).....	19
Business Requirement Specification (BRS) document (T.B.D)	19
References	22

List of Definitions & Abbreviations

Abbreviation	Definition
API	Application Programming Interface
BIM	Building Information Model
EC	European Commission
EE	Energy efficiency
EU	European Union
IFC	International Foundation Classes
SDU	University of Southern Denmark
T.B.D.	To be determined
GDPR	General data protection regulation
IoT	Internet of Things
WP	Work Package

Executive Summary

Building Information Modelling (BIM) is one of the major drivers of building sector digitalization. In the current practice, BIM covers the minimum requirements set by guidelines and is put in a drawer after the building handover with no use for facility operation and management. As owners will spend over 80% of the facility life cycle cost during the operation phase, the project will develop a holistic Digital Twin platform allowing coordinated decision support over the whole building life cycle. The platform replaces the traditional static BIM with a living operational model combining flexible and adaptable dynamic energy models and real-time building data.

Employing the proposed Digital Twin, 1) owners will have higher energy efficient buildings from day one, 2) facility managers will have a tool for smarter facility management through advanced monitoring, automated continuous commissioning and timely maintenance intervention, 3) consultants and planners will have a solution aiding in effective strategy planning and optimal decision-making support through running 'What if?' scenarios in advance, predicting impacts and testing control and management regimes in a virtual environment.

The project is carried out as a collaborative project between four main partners. SDU Center for Energy Informatics at University of Southern Denmark, COWI A/S, KMD A/S and Aarhus Municipality. Demonstration of the solution will be done through case studies conducted on three buildings of different character; the SDU Live-Lab campus building OU44, Dokk1 on Aarhus harbor and a hospital building in Bergen Norway.

Introduction

This deliverable provides high-level requirements specification.

The following documents will further develop the requirements presented in this document:

- D2.1 Semantic interfaces – to shape the TWIN4BUILD Ontology;
- D5.1 functional and technical requirements specification – to identify main functional features;
- D5.1 architectural design – to shape the TWIN4BUILD architecture;
- D5.1 value-added services definition requirements– to define value –added services demonstrated in the project.
- D1.5 exploitation business plan – to include relevant business motivations and inhibitors in business plans.

The technical requirement specification and architectural design drives the whole TWIN4BUILsolution definition and subsequent implementation.

The following relevant deliverables and activities are used to extract requirements specification:

- D1.4 Project meetings and coordination between partners
- A2.3 Carrying out field interviews with potential clients, facility managers and building owners to assess the needs and demands

- D2.3 A feedback report summarizing the key findings of the field interviews with owners and clients

The requirements in this document are organised as follows:

- Chapter 1+2 provides a brief descriptive introduction and background to this report
- Chapter 3 describes the methodology adopted for gathering relevant issues
- Chapter 4 formulates the requirements for Twin4build- based on the methodology in chapter 3.

These requirements cover domains represented in Twin4Build, including common business requirements which were identified as relevant.

Approach

This chapter describes the approach taken to define the requirements specification as part of the over-all TWIN4BUILD requirement life cycle defined in WP2 “Building Digital Twin Platform Requirements”.

Methodology

Requirement management process facilitates elicitation of stakeholders’ expectations through requirements definition, technical specification and architecture design followed by the detailed design, implementation and validation phase. The requirements establish a consistent baseline for stakeholders’ expectation and reflect realistic real-world needs. The requirements provide basis for solution validation in the demonstration phase.

The TWIN4BUILD requirement specifications consolidate the following inputs from different activities and deliverables:

- Stakeholders drivers and barriers (part of the Business requirements not written yet) from A2.3;
- Operational and interface requirements from A2.3;
- IoT platforms, infrastructures, ontologies and relevant standards from D1.4;

Requirement identifiers and traceability

Every requirement of the requirements specification is identified by its unique identifier according to the following table. In this document we focus at the Building - and Energy Domain requirements.

Requirement group	Requirement identifier
Building domain requirements	TWIN4BUILD-R-BLD
Energy domain requirements	TWIN4BUILD-R-ENR
Common business requirements	TWIN4BUILD-R-BUS

Table 2: List of requirements groups

Requirement's template

The requirements will be included in chapter "TWIN4BUILD requirements specification" of this deliverable using the following requirement template:

TWIN4BUILD-REQ-00xx	<Requirement title>
<Requirement description>	

TWIN4BUILD requirements specification

The Requirements Specification describes TWIN4BUILD consortium's motivation for why the system is being developed, defines processes and policies/rules under which the system will be used and documents the high-level requirements from the stakeholder perspective.

The TWIN4BUILD business requirements specification is divided into 3 domains:

- Building domain
- Energy domain (see D2.1)
- Common business requirements. (T.B.D)

Building domain Requirements (TWIN4BUILD-R-BLD)

Digitalization enables smart buildings to establish a smooth interaction with their users and a proper understanding of the internal and external environment. The key in establishing this interaction is a large set of data. Smart buildings nowadays generate large amounts of data from all building services and at various levels and components of the built environment. The main added value here is that immediate access to data provide a comprehensive view on how a building is performing and enables owners and operators to manage assets and facilities more cost- and energy efficient.

However, data collected from all different sources and devices in the building needs to be organized, integrated, and communicated in relation to the building context. During the design and construction phases of a building, a large amount of information is produced such as 2D and 3D drawings, documents, and notes about almost everything in the facility, however maintaining and finding this information after construction is often very costly and time-consuming. In this context, BIM provides a standardized machine-readable information model developed specifically for each facility or building, regardless of the building age, type, size, or use. Such models comprise relevant data and information created and gathered on the facility starting by the geometry and orientation to the constructions and materials and the HVAC systems and electrical and thermal properties of various components.

But the BIMs for the built environment are static and non-interoperable descriptions based on International Foundation Classes that only covers minimum requirements set by guidelines and is often put in a drawer after the building handover with no use for facility operation and management.

A large block of initiatives, investigations and accumulated efforts have been devoted in the recent years to link and integrate various datasets in a standardized manner with BIM formats. This includes generating spatial data employing Geographic Information Systems (GIS) (1), integrating IoT devices (2), linking facility

management data with BIM information (3), and incorporating environmental performance data of infrastructure (4). But these efforts have been focusing on the construction industry and engineering of buildings. In this project we focus on buildings in a more holistic approach reaching outside the construction industry and frameworks and trying to relate to the building as a “living organism” being part of the greater society.

Dynamic building models

In the following chapters we explain why the data and the models that represent the buildings have to be dynamic, flexible, adapted in real-time, highly interoperable, interlinked and holistic to better understand and improve the construction, operation and usage of buildings and to fully represent the building in a digital context of the circular economy.

In future smart buildings, digitalization is going to establish new levels of efficiency, security and comfort. By having a contextualized and dynamic view on digital information, buildings will transition from being a static reactive element, to becoming a proactive key component with the ability to intelligently respond to the changes in the dynamic environment and connect efficiently with other buildings and infrastructures within smart communities and grids.

The project proposes combining and linking historical, real, and forecast data from buildings into a context information model to describe and understand the building context better over time and also dynamically in real time.

TWIN4BUILD-REQ-001

Digital twins represent building assets

To help understand and trust the data in a building context, data must be linked intelligently from decoupled systems, hence the project suggests representing building assets via Digital Twins and harmonized data standards. The Digital Twins are based on semantic knowledge graphs to ensure high quality, comparability, and enriched data, based on a holistic Digital Twin platform, allowing coordinated data-driven decision support over the whole building life cycle.

Digital Twins

Overall, a digital twin is a digital replica of physical assets, systems, devices or processes which can be used for various purposes. Thus, a digital twin is not just a digital model of a component or a system, it is a live operational model where real-time data is integrated with the virtual replica of the system to form a digital twin. Such integration allows continuous monitoring of the actual state of the system and inform decision making for future states and applications. The range of use of digital twins has broadened in the recent years and successful implementations have been reported in different industrial sectors on different levels including design, production, manufacturing, and maintenance. Such digital twin applications have demonstrated positive impacts especially when used in the design and production lines, assisting in the quick understanding the needs of customers, predicting weaknesses and faults in the models, and controlling and optimizing processes to respond to dynamic changes and alerting on maintenance needs in good time before failure occurrence (5).

From the conceptual perspective, a digital twin is made up of three main components:

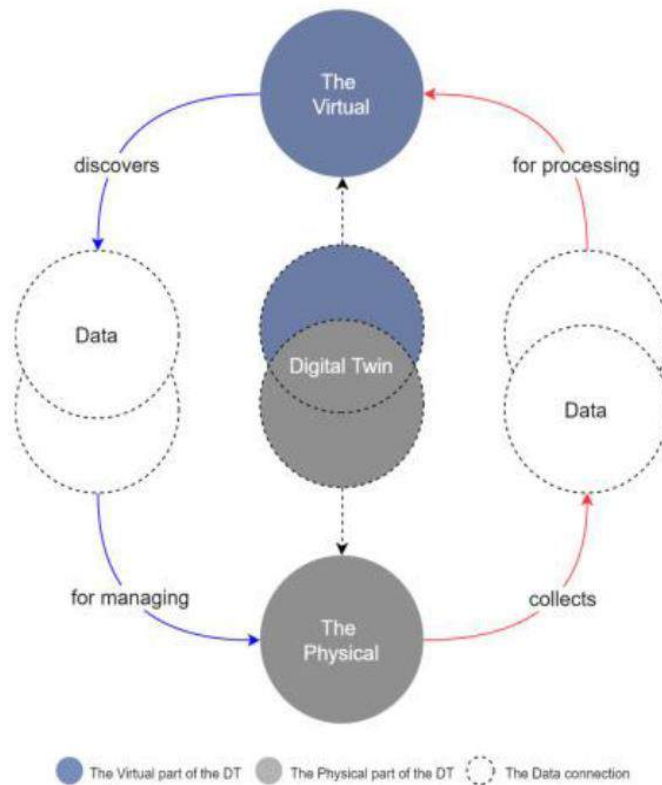


Figure 1. Main components of a standard digital twin (6)

1. Physical Component (real component)
2. Virtual Component
3. Data Connecting them

As shown in Figure 1, the connection and information carryover between the physical component and the virtual replica is provided by data on different levels and in different forms.

The live Digital Twin of a building
A building digital twin provides a comprehensive and holistic picture on the actual building in operation, bridging the physical and digital world. It combines a virtual building model with actual data collected from various meters and IoT devices to understand, analyze, manipulate, and optimize processes within a building and inform decision making throughout the life cycle. Thus, a digital twin is valuable at every single stage of the building value chain, including the

conception, design, build, commission, operate and maintain stages.

In this project, where we handle live change of building information, the focus is primarily on the operation phase where 2 levels of digital twins are relevant:

- **Data twin:** it provides a queryable representation of observed properties of the physical twin. The representation is limited to the past, current or predicted properties of that physical artifact.
- **Smart twin:** In addition to the current state of observed properties the digital twin includes a behavioural model of the physical twin. This model can be of any kind such as multiphysics model, machine learned model, etc. that represents the functional states within defined boundaries.

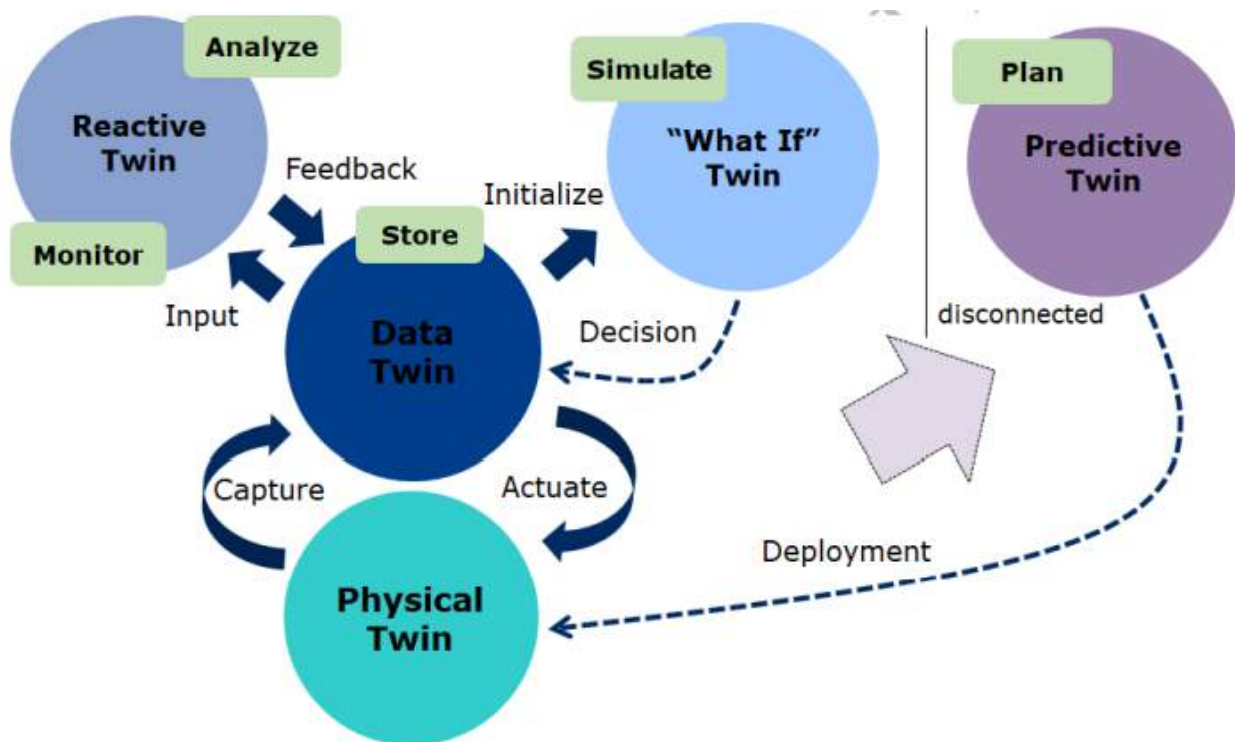


Figure 2 Digital Twins can be used to store, monitor, analyze, simulate and plan activities (6)

While the data twin captures and stores a passive model of the physical twin, which is updated with current sensor information - it is important to actively monitor and analyze the current situation - referred to as "reactive twin" in Figure 2 above. The figure also describes that several usage patterns are foreseen, including predicting behaviour of the system, detecting anomalies in operation or executing what-if scenarios based on the current state of the data twin. In addition, coupling to actuation can be envisioned to allow physical systems being driven by their twin in deployment.

Finally, as the result of the situation analysis and simulation, functionality is needed to automatically and reliably trigger actuations on the physical twin. All-in-all we consider this dynamic and holistic digital twin as being a "smart digital twin".

Smart Digital Twins as Building Information Models

A smart digital twin is regarded as the next evolution of the BIM model. It brings the static BIM to life by integrating actual data and dynamic information, resulting into a live dynamic digital twin that can serve the building throughout the life cycle; design, (de)commissioning and operation. The smart digital twin can utilize information and operational technologies such as construction data and building floor plans from BIM, real-time sensor data from the BMS, including data from smart HVAC systems and Internet of Things (IoT) devices and other building services sensors, as well as meter energy consumption and building usage data.

TWIN4BUILD-REQ-002

A Digital Twin replaces the BIM

The proposed holistic Smart Digital Twin replaces the traditional static Building Information Model (BIM) with a living operational model combining flexible and adaptable dynamic energy models and real-time data to provide the following services:

- Collecting, integrating, and managing building data within a flexible, effective and user friendly open standard context information model.
- Real-time building performance monitoring and automated continuous commissioning and fault detection enabling smarter operation and facility management along with performance optimization
- Strategy and planning support over the whole building life cycle through informing optimal decisions and running 'What if?' scenarios in advance.

On the overall building level, a digital twin also provides a platform for comparing the actual performance of a building to an expected dynamic baseline and thus serve as a basis for commissioning processes and fault and anomaly detection and diagnosis.

Having both static building information and dynamic data and predictions within the same platform allow the facility managers and building owners to localize malfunctions and errors in high precision and less time leading to an optimized predictive maintenance solution. At a system level, a digital twin allows operational optimization by making use of various data inputs from IoT devices to predict and inform control and management strategies.

Digital twins also improve the interactions capabilities with the real artefacts. Developing the appropriate Digital twin data model and putting data into context is a key ingredient for interoperability and bridging the gap between different building verticals, including the need for maintenance and to predict energy savings and CO2 reduction potential

TWIN4BUILD-REQ-003**Digital Twin applications are the foundation for building management**

The project takes a step forward to establish a leading role in terms of Digital Twin applications in the building sector and aims at:

- Laying the proper foundation for understanding, comparing, and sharing data in a building context and successful building digital twins
- Identifying optimal conditions in terms of proper building information modelling (BIM) design and building datasets and ontology requirements
- Designing a building Digital Twin framework and the associated services and strategies
- Developing a holistic Digital Twin platform for building applications
- Demonstration of the Digital Twin in case study buildings and assessment of the technical and socio-economic impacts.

Digital Twins classification and context

The connection between digital twins, representing the interactions and relationships, is also important. A core element of the solution is a Digital Twin data model of the target asset in a building. Building Digital Twins can be considered as connected entities that provide a network interface supporting direct interaction of applications with the entity/device.

Figure 3 gives an example of this for an apartment captured as a subsystem of a building.

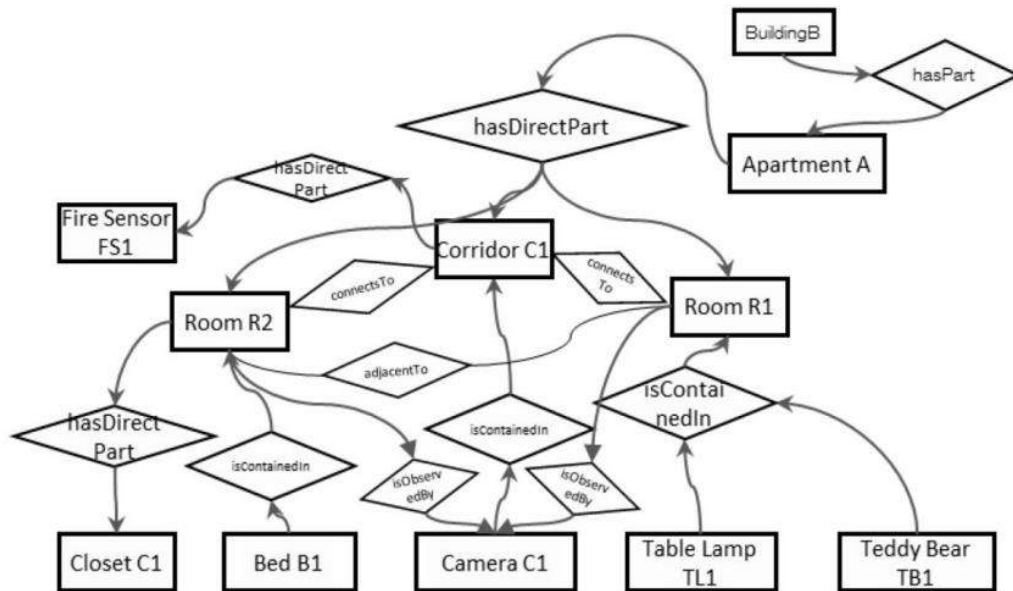


Figure 3 Digital Twins are connected in graphs representing the interactions and relationships (6)

The relationships between the nodes that make up self-contained systems like these may correspond to:

- **vertical top-down links** between the root system and its constituent parts (with type `hasPart/hasDirectPart`), when these parts are designed to be included in the overall system and this system cannot work if these parts are removed
- **vertical bottom-up links** (with type `isContainedIn`) between parts and systems that would not always be captured as such, like the sets of all things (furniture, appliances, etc.) contained inside a room. This type of relationship may also be used to capture an even more informal and purely contingent set-based location, without implying a “systemic” relationship.
- **transversal links** (with type `ConnectsTo`) to capture the way through which one may go from one room to another, or a room is near another (with type `AdjacentTo`)
- **transversal directed links** between a sensor and what it observes (with type `isObservedBy`), or an actuator and what it acts upon (with type `sActedOnBy`)

Context

Context is defined as the set of properties characterizing these nodes, together with the set of relationships, and the properties of these relationships. In this perspective, the primary data of one application may be the context of another, and vice versa. Context is, thus decentred and broadly defined, the graph itself.

TWIN4BUILD-REQ-004	The context of Digital Twins is represented as connected entities in a graph
--------------------	--

Digital Twins can be considered as connected entities that provide a network interface supporting direct interaction of applications with the entity/device. This network is represented as a graph that extend beyond active devices including the passive or legacy “stuff” that can be sensed by sensors or acted upon by actuators. The entities are well-characterized by their external relationships with other entities at the same level, or with the larger systems which they are part of.

The Digital Twin for buildings in a larger energy context

Since the building cannot be considered a fully isolated digital twin, we also must consider relationships to digital twins outside the building. As indicated in the figure below, previous research (7) show that a digital twin for a building in an energy context, need to relate to many external entities that all affect the building. This project will include these external digital twins to the extent that we can get access to the needed data.

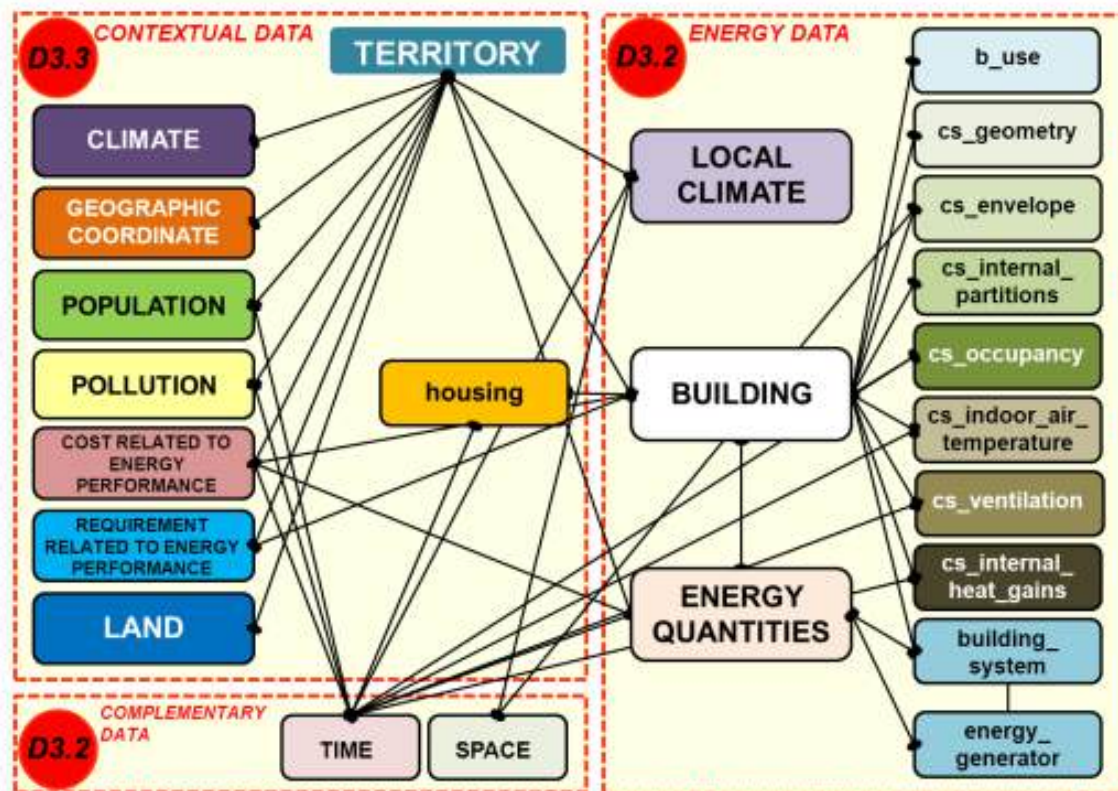


Figure 4 Relationships to digital twins outside the building must be considered to represent the building in “real life” (7)

Digital Twin lifecycle

To identify what are the expected functionalities and usages of digital twins, it is proposed to split the analysis over the different stages of a system lifecycle: Design stage, commissioning, and Operations.

As more data is collected and integrated within the digital twin, more opportunities for scenario planning and what-if analyses arise using various intelligence algorithms. Using organized data, a digital twin provides real-time understanding of how a building is performing, enabling immediate adjustment to optimize the performance at various levels, resulting in a more energy efficient, cost-effective, and sustainable smart buildings and facilities.

TWIN4BUILD-REQ-005

Digital Twins are used as input to the intelligent algorithms

The different intelligent algorithms examined in the project are:

- Predictive analytics: to provide a prediction of the system state in the future, based on the current state
- Outliers/anomaly detection: to raise alerts in case the system deviates from its anticipated state.
- What-if scenario: how will evolve the system if an action is taken from current state, modify some parameters in the model
- System optimisation: real time optimisation of the system by the DT.

Digital twin data models and ontology requirements

Standard ontologies are the building blocks which support the different digital twins coming together, creating a common understanding, shared values and language. More so, standards evolve constantly, and new ones are added as time passes.

TWIN4BUILD-REQ-006**Digital Twins are based on standard ontologies and data models**

In this project we use the latest research in data models and technical standards to ensure that Digital Twins are based on the most relevant standard ontologies and data models

Data Models

As described previously, Digital Twins help generating a holistic view on what is currently going on in a building and help to manage real-time data gathered from the different vertical systems by precisely model the entities, their interactions, and their context as experienced in both time and the physical space. Processing and analysis of this data, referred to as context information, will bring support to adoption of smart decisions or the smart automation of certain processes.

The connections between digital twins, representing the interactions and relationships, are also important in ensuring security and protection of personal data and the resilience of assets and systems. Standards and data exchange protocols will allow the ecosystem of connected digital twins to be created, including attributes related to the legal and regulatory contexts within which the assets operate, but at the core, the Digital Twin need a domain specific data model of the target assets in a building.

Developing the appropriate Digital twin data model and putting data into context is a key ingredient for interoperability and bridging the gap between different building verticals including the need for maintenance and to predict energy savings and CO2 reduction potential.

Data models ensure interoperability

The data model ensures a harmonized representation of the formats and semantics that will be used by applications both to consume and to publish data and serve as a language in which systems can talk to each other and hereby opening up data across solutions.

Data models should capture as much the complete context they are representing. This enables other applications to define what they need for their context and request the specific attributes they are interested in.

Several EU-funded pilots as well as supporting actions and partnerships have produced standards, mechanisms, services, guidelines, and tools that enable the interoperability of urban platforms. But choosing the right ontology for creating the data schema for the data model, is very important for achieving the previously mentioned semantic interoperability between the systems. To make the choice of ontologies easier, the Semantic Interoperability Expert Group of the Alliance for Internet of Things Innovation (AIOTI) (8) has created an Ontology Landscape that currently includes 30 ontologies from different application areas of IoT. The overview gives a first indication of which ontologies could be suitable candidates. As described previously this project focus on digital twins in a broader context than only buildings. Hence the chosen ontology must cover several domains, both the Building, the Energy, the City and Environment domain.

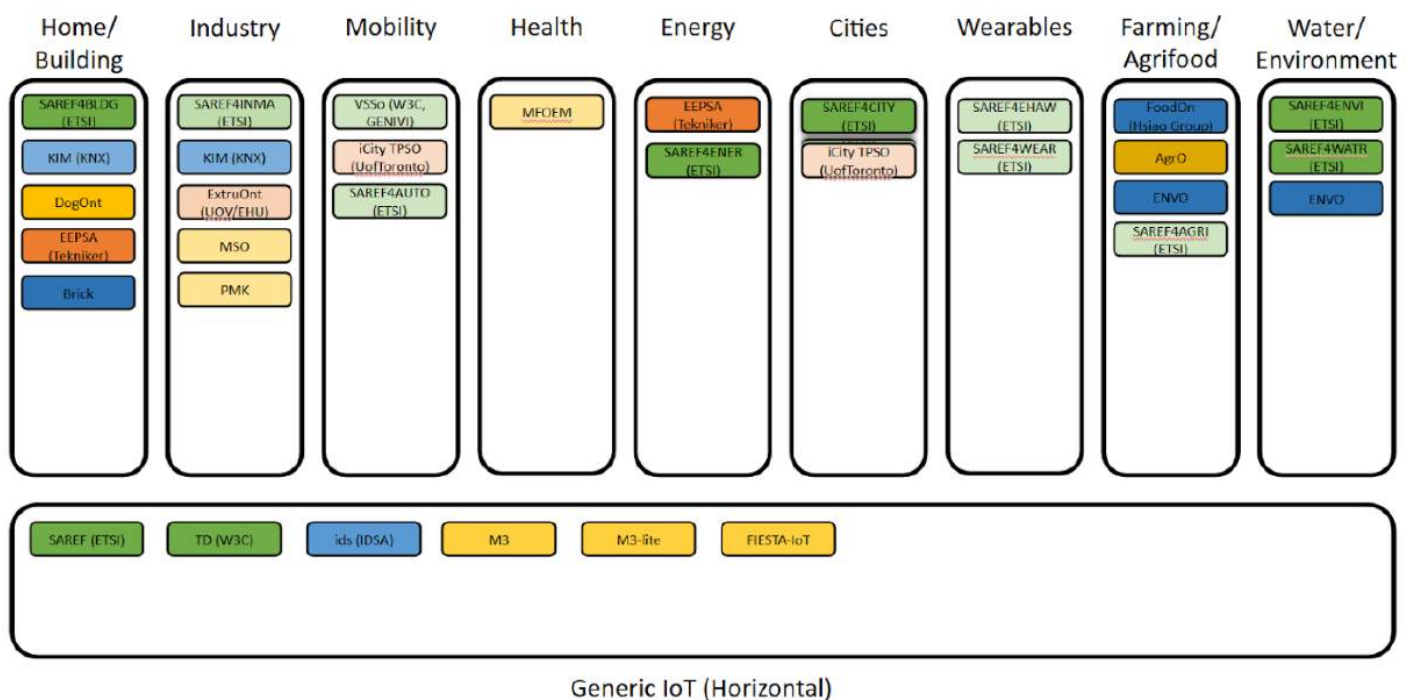


Figure 5 SAREF is the ontology covering the most domains (21)

As seen from Figure 5, the only ontology that cover the needed domains is SAREF, primarily due to its many extensions.

TWIN4BUILD-REQ-007

Digital Twin data models are based on SAREF

The project uses Smart Appliance/Anything REference (9) to represent Digital Twins in a data model: The SAREF ontology, an ETSI/OneM2M standard, is a shared model of consensus that helps match existing assets, such as standards, protocols, and data models. It consists of base ontology and extensions for the relevant sectors, including one for cities (SAREF4CITY). SAREF Ontologies are promoted by the European standards community and inherits from the Ontology-based Representation of IFC, which make it a perfect match to also represent the BIM.

Digital Twin Ecosystems, Platforms and Context Data Management

The key obstacles in getting broader value from existing building data collection, storage and management, are dissimilar and inconsistent standards and data models due to siloed efforts. As described earlier some of these problems can be solved by specifying requirements for digital twins, building on the standards for the domain, but to be able to easily exchange this standardized information a digital twin platform need to build on open set of APIs, but also on other platform components to reduce the cost in finding, selecting, and integrating the diverse building data sources.

The digital twin eco-system also needs to gather and manage context information, processing that information and informing external actors, enabling them to actuate and therefore alter or enrich the current context – this is the job for a “broker” as shown in Figure 7, including event-triggered service executions. The Service Executor has a similar role as the Context Source, but instead of providing context information, it enables the execution of services related to one or more entities (digital twins). The Service Execution Application is an application that creates service executions.

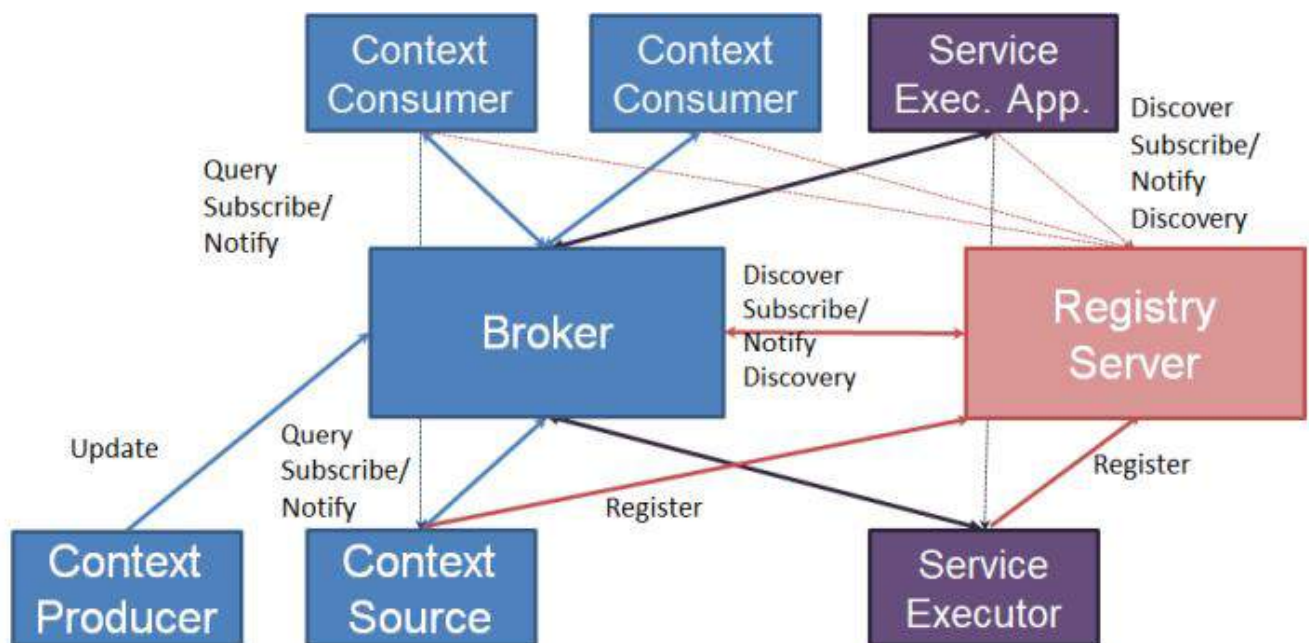


Figure 7 The digital platform should also include the needed infrastructure components surrounding a broker (6)

The digital platform should also include the needed infrastructure components surrounding a broker, for supporting processing, analysis and visualization of data or bringing support to data access control, publication or monetization via software applications on top of the platform. It should allow developers to source components from different suppliers and to reduce fragmentation, increase reusability and scalability. Developers can also manage their risk exposure to single-vendor solutions and increase the ability to commission multi-stage projects without locking into a single technology or vendor from the outset. This approach also increases the potential for innovation and to reveal new application opportunities (10).

Digital Twin platform

The Digital Twin Platform can never encompass every conceivable aspect of a building. To be able to support the before mentioned interoperability, the platform needs to establish a minimal set of interoperability requirements to link systems together when they need to.

The Open Agile Smart City community call these requirements for “Minimal Interoperability Mechanisms” (11) (12) (13), or MIMs for short. The MIMs are a set of common APIs (to access data), common data models (to structure data), and a common, but optional, data platform (to store and serve data). In addition, a reference architecture and a reference implementation complete the set of MIMs.

The MIMs help (14) to avoid vendor lock-in on the demand-side and ‘city lock-in’ on the supply-side through the use of open standards and open source components such as the ETSI OneM2M reference library (15), the SynchroniCity reference architecture (16) and the FIWARE reference architecture (17).

Architecture framework model

MIMs also ensure that the platform is reliable, durable, scalable and performant so that the city can build on the platform and guarantee privacy and security by design, making the platform trustworthy. The platform should allow to support formal, de-facto and emerging standards, making sure that it is future-proof and stable.

High level architecture

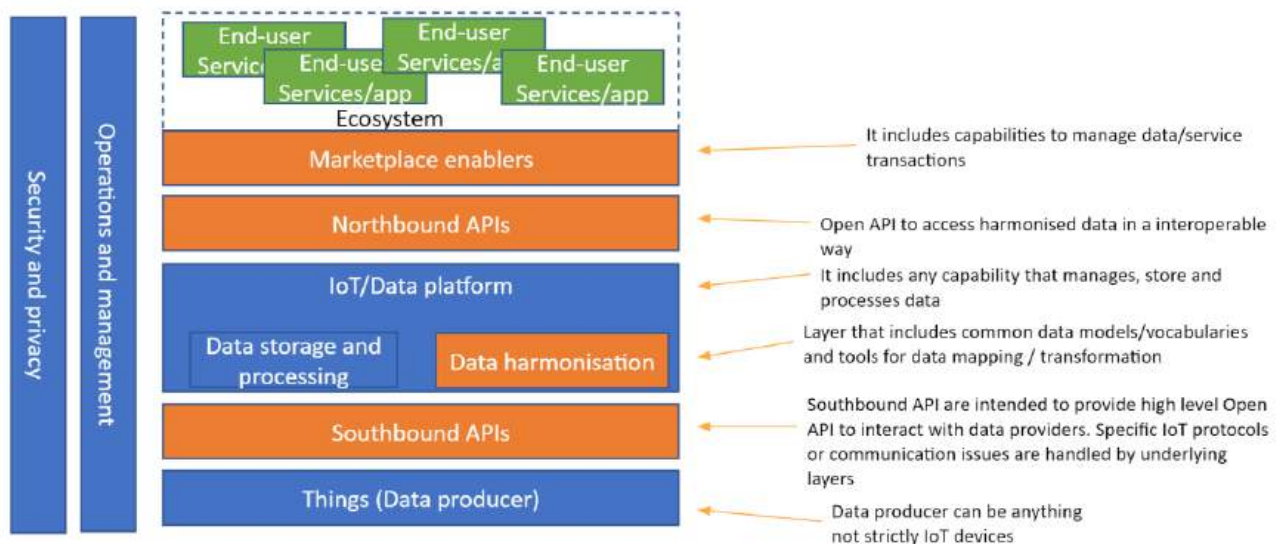


Figure 8 MIMs require a layered and capability-based common architectural model in different cities/domains (12)

TWIN4BUILD-REQ-008

MIMs as guiding principles

The project uses the MIMs as guiding principles to the extend that it fits the scope of the project and consider the following MIMs topics as common architectural design principles:

- A layered and capability-based approach to follow a common architectural model in different cities/domains.
- Based on open international standards (where available)

-
- Compliant with existing technical solutions (e.g. already present in the cities with many legacy systems) focusing on interoperable interfaces rather than component implementation.
 - Modular and scalable solutions for small and big cities: e.g. support different deployment scenarios and performance requirements.
 - Security and privacy by design.
 - Availability of reference implementations to foster and simplify the adoption in the cities.
 - The architecture modularity assures the possibility to implement any component with different/proprietary technologies.
 - Based on open API to enable both southbound/northbound interoperability.
 - Data harmonization and semantic interoperability through the adoption of common linked data models
-

Energy Domain Requirements (TWIN4BUILD-R-ENR)

This refers to the D2.1 document but could be included here to make it easier to read in the context of this document.

Common business requirements (TWIN4BUILD-R-BUS, T.B.D)

This section includes business requirements identified for cross-domain applicability. It builds upon the findings presented in the domain specific requirements.

The common business requirements are clustered in the following categories:

Usability requirements;

Implementation requirements;

Security requirements;

Privacy requirements;

Legal & Ethics requirements.

Business Requirement Specification (BRS) document (T.B.D)

This annex will be used in development of concrete business plans that will look to maximize the exploitation potential of identified assets results as well as address stakeholders' needs and expectations. Exploitation Strategy and Business Plan Development

1.1 Business purpose

Describe at the organisation level the reason and background for which the organisation is pursuing new business or changing the current business in order to fit a new management environment. In this context it should describe how the proposed system will contribute to meeting business objectives.

1.2 Business scope

Define the business domain under consideration

1.3 Business overview

Describe major internal divisions and external entities of the business domain concerned and how they are interrelated. A diagrammatic description is recommended.

1.4 Major Stakeholders

List the major stakeholders or the classes of stakeholders and describe how they will influence the organisation and business, or will be related to the development and operation of the system.

1.5 Business environment

Define external and internal environmental factors that should be taken into consideration in understanding the new or existing business and eliciting the stakeholder requirements for the system to be developed or changed. The environmental factors should include possible influences to the business and consequently the system from external conditions like market trends, laws and regulations, social responsibilities, and technology base.

1.6 Mission, Goals and Objectives

Describe the business results to be obtained through or by the proposed system.

1.7 Business model

Describe methods by which the business mission is expected to be achieved. The description should be concentrated on the methods supported by the system to be developed or changed with the items such as product and services, geographies and locales, distribution channels, business alliance and partnership, and finance and revenue model.

1.9 Business processes

Provide description of the procedures of business activities and possible system interfaces within the processes. The purpose of this information item is to represent how and in which context the system supports the business activities. In general, business processes make a hierarchical structure with decomposition and classification. Each business process should be uniquely named and numbered in the hierarchy. The description of the individual business process should be represented as a diagram representing a sequence of activities.

1.10 Business operational policies and rules

Describe logical propositions applied in conducting the business processes. The propositions will be conditions to start, branch and terminate the sequence of the business activities in the business processes; criteria for judgment in the business processes; or formula to evaluate a quantity, which will likely be addressed in functional requirements in the SyRS and SRS. The policies and rules shall be uniquely named and numbered, and shall be referenced in the description of the business processes.

1.11 Business operational constraints

Describe conditions to be imposed in conducting the business process. The conditions may be on a performance constraint (e.g., the process shall be finished within a day after the triggering event occurs), or may be from a management requisite such as 'every occurrence of the process shall be monitored and recorded'.

1.12 Business operation modes

Describe methods to conduct the business operation in an unsteady state, for example, a state when business operations might be extremely busy due to some intensive occurrence of events. An unsteady state of business operation includes a manual operation mode when the proposed system is not available due to some unexpected situation like an accident or natural disaster.

1.13 Business operational quality

Define the level of quality required for the business operation. For example, a business process may address required urgency with higher priority than the reliability of the business process.

1.14 Business structure

Identify and describe the structures in the business relevant to the system, such as organizational structure (divisions and departments), role and responsibility structures, geographic structures, and resource sharing structures. There may be a need to align the system functions to these structures, and to support future structural changes.

1.15 Preliminary operational concept

Describe the proposed system in a high-level manner, indicating the operational features that are to be provided without specifying design details. The following information should be included: d) Operational policies and constraints

Description of the proposed system

Modes of system operation

User classes and other involved personnel

Support environment

1.16 Preliminary operational scenarios

Describe examples of how users/operators/maintainers will interact with the system (context of use). The preliminary (upper-level) scenarios are described for an activity or a series of activities of business processes supported by the system. The scenarios should be uniquely named and numbered, and should be referenced in the description of the business processes in 6.1.9.

1.17 Other preliminary life-cycle concepts

Describe how the system of interest is to be acquired, deployed, supported and retired.

1.18 Project constraints

Describe constraints to performing the project within cost and schedule.

References

1. *Integration of BIM and GIS in sustainable built environment: A review and bibliometric analysis*,. **H. Wang, Y. Pan, X. Luo**. s.l. : Automation in Construction, 2019, Vols. 103 pp. 41-52.
 2. *A framework for integrating BIM and IoT through open standards*. **B. Dave, A. Buda, A. Nurminen, K. Främling**. s.l. : Automation in Construction, (2018), , Vols. 95 pp. 35-45.
 3. *BIM for facilities management: A framework and a common data environment using open standards*. **J. Patacas, N. Dawood, M. Kassem**. 103366, (2020) : Automation in Construction, 2020, Vol. 120.
 4. *Building information modeling (BIM) for transportation infra-structure – literature review, applications, challenges, and recommendations*. **A. Costin, A. Adibfar, H. Hu, S.S. Chen**. s.l. : Automation in Construction, 2018, Vols. 94 pp. 257-281.
 5. *Uses of the digital twins concept for en-ergy services, intelligent recommendation systems, and demand side management*. **A.E. Onile, R. Machlev, E. Petlenkov, Y. Levron, J. Belikov**. s.l. : A review, En-ergy Reports, 2021, Vols. 7 pp. 997-1015.
 6. ETSI, Context Information Management (CIM); DRAFT. [Online] https://docbox.etsi.org/ISG/CIM/Open/drafts/ETSI_GR_CIM-0017_Digital_Twin_DRAFT-for-PUBLIC-COMMENT-V002.pdf.
 7. SEMANCO. [Online] http://www.semanco-project.eu/index_htm_files/X-SEMANCO_D3.5_20140930.pdf.
 8. Ontology landscape. [Online] <https://aioti.eu/aioti-ontology-landscape-report/>.
 9. SAREF. [Online] <https://www.etsi.org/technologies/smart-appliances>.
 10. Benefits of OneM2M. [Online] <https://www.onem2m.org/using-onem2m/what-is-onem2m#value>.
 11. OASC MIMs. [Online] <https://oascities.org/wp-content/uploads/2019/06/OASC-MIMs.pdf>.
 12. Open Agile Smart Cities, MIMs. [Online] <https://oascities.org/wp-content/uploads/2019/06/OASC-MIMs.pdf>.
 13. ETSI Smart Appliances. [Online] <https://www.etsi.org/technologies/smart-appliances>.
 14. Declaration on joining forces to boost cities. [Online] https://living-in.eu/sites/default/files/files/declaration_english.pdf.
 15. Onem2m. [Online] www.onem2m.org.
 16. Synchronocity tech. [Online] <https://synchronicity-iot.eu/tech/>.
 17. Consolidated Report on Tech Specs v2. [Online] <https://living-in.eu/sites/default/files/files/Consolidated-Report-on-Tech-Specs-v2.pdf>.
 18. Consolidated Report on Tech Specs v2 (living-in.eu). [Online]
 19. C. Boje, A. Guerriero, S. Kubicki, Y. Rezgui, Towards a semantic construction digital twin: directions for future research, Automation in Construction, 114 (2020). [Online] <https://www.sciencedirect.com/science/article/pii/S2352484721000913>.
-

20. Context Information Management (CIM); Draft. [Online]

https://docbox.etsi.org/ISG/CIM/Open/drafts/ETSI_GR_CIM-0017_Digital_Twin_DRAFT-for-PUBLIC-COMMENT-V002.pdf.

21. AIOTI Ontology Landscape Report. [Online] <https://aioti.eu/aioti-ontology-landscape-report/>.

EUDP DIGITAL TWIN - WP2 INTERVIEWS

CONTENTS

1	Introduction and methodology	2
2	The situation right now (Tema 1)	4
3	Expectations and potential (Tema 2)	6
4	Identified challenges (Tema 3)	8
5	The digital twin (Tema 4)	10
6	Conclusion	11
7	Appendix:	12

PROJEKTNR.

DOKUMENTNR.

VERSION

UDGIVELSESDATO

BESKRIVELSE

UDARBEJDET

KONTROLLERET

GODKENDT

1 Introduction and methodology

In connection with this EUDP project, work package 2 defines that we must carry out field interviews with potential clients, facility managers and building owners to assess the needs and demands for digital twins (DT).

In this work package, we conducted interviews with six customers:

- 1) Aarhus kommune
- 2) Newsec
- 3) FSB
- 4) KEID
- 5) Bygningsstyrelsen
- 6) Helse Bergen (NO)

The interviews were conducted in February and March in 2022 and as far as possible participants from both COWI and KMD were present.

Each interview is held according to an Interview Guide which has served as a template for the 6 interviews. The 6 interviews are shown in the appendix.

In the following, a brief introduction is given to the customers.

Aarhus Kommune	Aarhus Municipality Technology and Environment administers and monitors approx. 1.500 properties in Aarhus municipality.
Newsec	Newsec is a large professional real estate administrator. Newsec manages 47 million m ² in the north. Many of the properties are owned by large pension companies which focus hard on returns.
FSB	FSB is a public housing organization with 13.000 homes in 80 housing departments in Copenhagen. FSB is the largest housing association in Copenhagen. The FSB is non-profit. The FSB is governed by an organization's board which consists of the tenants. The FSB is thus a political organization.
Københavns Ejendomme og Indkøb (KEID)	Copenhagen Properties and Purchasing (KEID) is a public developer, with a total of approximately 3.800 properties a total of 2.6 million m ² . KEID is a political organization.
Bygningsstyrelsen	The Danish Building Authority is a real estate company and developer for the state

	and the universities. The Danish Building Authority manages DKK 4.2 million. m ² properties.
Helse Bergen (NO)	Helse Bergen is a Norwegian university hospital. They manage about 450 000 m ² of property divided into 115 buildings.

2 The situation right now (Tema 1)

The following questions were answered:

"Where are you in the process of digitizing, including DT?"

Client/Building Owner	Note
Aarhus Kommune	BIM is used only for design models for building construction. Data collection and use of data today is very static and reactive and not proactive and takes place primarily from very simple data sources.
Newsec	Newsec does not work with BIM models nor digital twins. For Newsec's customers, returns are the alpha and omega. If energy savings or sustainable initiatives can be translated into a significant return, then it is interesting for Newsec's customers. If you can make a better economic model based on a digital twin, then the interest is great - if not, it is not interesting. Newsec follows the development but does not take the initiative themselves.
FSB	The FSB gradually started up in 2017 by establishing BIM models of all their properties. The FSB would like to make electronic inspections when moving in and moving out of dwellings and uses the BIM models to make bulk extracts of surfaces that must be renovated when moving out in order to be able to make correct estimates of costs and offers for craftsmen. The FSB is very interested in the potential.
Københavns Ejendomme og Indkøb (KEID)	KEID has a relatively low level of information in their BIM models but would like to be better at using the data contained in BIM. It is the ambition that all their buildings must have a 3D model. KEID wants to go where they have a digital engine, in which you can put a decision in, and see the consequences of it - also for managing activities. KEID knows about digital twins, and demands more knowledge about how much or how little data is needed to exploit the potential of a digital twin. KEID is working on various data projects, e.g. LeanHeat (intelligent control of district heating) and collaborates with HOFOR on

	consumption profiles and outsourcing of heat control to HOFOR (which presupposes detailed knowledge of the buildings - to make the consumption profiles).
Bygningstyrelsen	The Danish Building Authority are good at using BIM actively in the design not in operational contexts and simulations. The Danish Building Authority are having 3D models made for all buildings in their portfolio.
Helse Bergen (Norge)	The building which has been selected to have a DT established is in the construction phase The project is far ahead in terms of data in the project organization on construction. On the operating side, they are less mature. but they have set in motion a system so that they also mature in operation. Helse Bergen tries to streamline their data management (and processes) across construction and operations, including also implementing the same systems - however, it is perceived as very difficult. Helse Bergen tries to incorporate the ideas from construction into operation.

3 Expectations and potential (Tema 2)

The following questions were answered:

"What do you expect to be able to achieve by using DT?"

Client/Building Owner	Note
Aarhus Kommune	Aarhus Municipality has a desire to get a visual overview of the buildings they operate in their portfolio. The primary desire is to be able to achieve increased building performance and energy optimization of the entire portfolio. There is a desire that this should apply to both new and existing construction. The DT must obtain forecast on weather, expected number of users, show energy consumption, parameters such as temperatures, CO₂ but also be able to detect and estimate the number of users - all this down to room level. These parameters must be able to be used to regulate the building installations up to a desired / expected indoor climate with the least possible energy consumption. The DT should make it easy for operations managers, possibly replacements or new operations managers to see if the building is in optimal operation.
Newsec	Newsec is uncertain about where the development is going, in relation to the digital twins. Can they optimize Newsecs or their customers' business - can you get the digital twins to take over some human operating costs, then it will be very relevant. But Newsec has not discovered that potential yet. The value of the digital twin is not clear to them yet. DT are of interest to Newsec if the DT could predict how the indoor climate and energy consumption will be if the interior / floor plan changes and.
FSB	The DT must be able to provide explanations for the development in energy consumption - why are the changes happening? Automating the explanation so that it is not handheld. They must be able to provide more data to determine the direction. Does improvements the energy savings that had been assumed? What is the potential of that improvement? Are there other potentials for energy improvements?

	The DTs should make it easy to compare the performance of different buildings.
Københavns Ejendomme og Indkøb (KEID)	A DT should show the processes - and the connection between elements. Contribute to giving the building an operational strategy that strengthens decisions. A digital twin should provide more knowledge on how to operate the building in the best way. A digital twin must detect deviations in energy consumption that go beyond the norm and give an alarm, also in relation to multi-functional buildings. Be able to accumulate heat in a building. Drain heat from the district heating system when it is cheap and store it in the building constructions until it is needed.
Bygningsstyrelsen	The Danish Building Authority is very focused on BIM and not on DT yet. The construction projects are very fragmented - so it is difficult to get from BIM in the construction phase and implement it in operation. The Danish Building Authority has not dived into the potential of the DT but is open to the concept. Right now, the Danish Building Authority is in clarification phase. What are the needs when operating a building? It is not entirely clear to them what data is needed, in relation to the operating side. They lack knowledge and data about the interface between construction and operation.
Helse Bergen (Norge)	There is an expectation that the digital twin can help reduce energy consumption, intelligently monitor data, including troubleshooting. In addition, Helse Bergen is not fully aware of what DT can be used for. But there is clearly a potential for creating value. They are open to new ideas.

4 Identified challenges (Tema 3)

The following questions were answered:

"Are you ready to implement a DT? Do you expect any challenges?"

Client/Building Owner	Note
Aarhus Kommune	Aarhus Municipality is ready and looking forward to the challenge. It is the municipality's experience that the existing BMS (CTS) is too general, and more knowledge is needed.
Newsec	Newsec has 7-8 buildings where you could set up a digital twin - if the finances are made available. Newsec has HVAC systems on some larger hosted environments. In these environments, it would in theory be able to set up a DT. It is a challenge that they do not know the potential of a DT. The business case is very important.
FSB	The BIM models must be updated, and the level of detail must be increased, it is necessary to make a requirements specification for BIM and data. There will be restrictions on what data can be obtained from the residents, as it is not legal to register whether the residents are at home, how much they are in the home, etc. CTS on older buildings may need to be upgraded. The model must represent reality. Updating master data is a prerequisite. New interiors can be a challenge.
Københavns Ejendomme og Indkøb (KEID)	One thing is to present data - see the red spots and energy consumption - something else is to link data together and more clearly state the operating scenario - and implement the digital twin in operation. Great challenge in competencies in the operation - if the operation cannot understand the systems, then the systems become indifferent. The model must represent reality. Updating master data is a prerequisite.

	KEID needs knowledge about how much or how little data you need to succeed with a DT.
Byggningsstyrelsen	There must be some requirements for handing over data from construction to operation. It is difficult, but a prerequisite. The operation is not far in terms of using data smart. The Danish Building Authority builds for many different operators. They mostly use 2D drawings - and are not so much inside the 3D universe. The operators lack the competencies and / or IT that support it.
Helse Bergen (Norge)	Interaction between the project organization and the operating organization is a clear prerequisite for valuable use of the digital twin. Those who need the twin must understand it. Operation and maintenance are central. The operations department must feel ownership of data.

5 The digital twin (Tema 4)

The following questions were answered:

"Which types of properties will best be suited for DT?"

Client/Building Owner	Note
Aarhus Kommune	The primary desire is to be able to achieve increased building performance and energy optimization of the entire portfolio. There is a desire that this should apply to both new and existing construction.
Newsec	Commercial properties - the more complex the properties, the more the need for data (complex: cooling, combined ventilation, high demands on interaction) Housing will be less interesting (less technical). New properties will definitely be best. Non-existent buildings, unless a major renovation is undertaken.
FSB	FSB primarily has residential properties and DT will be interested in both new and existing properties. The FSB's old buildings are great sinners in relation to CO ₂ . They would have to look at both single buildings and on-portfolio level.
Københavns Ejendomme og Indkøb (KEID)	The primary desire is to be able to achieve increased building performance and energy optimization of the entire portfolio. There is a desire that this should apply to both new and existing construction. A digital twin will be highly relevant on buildings that have a very varied heat consumption, such as. children's institutions.
Bygningsstyrelsen	DT could be relevant on both existing buildings and new buildings.
Helse Bergen (Norge)	Especially the large central buildings.

6 Conclusion

Based on the interviews conducted, we can conclude the following:

The DT should give you a visual overview of all the buildings that are operated in a portfolio.

The primary desire is to be able to achieve increased building performance and energy optimization.

The DT should apply to both new and existing buildings.

The DT should be able to detect and estimate the number of users down to room level.

The detected parameters should be used to regulate the building installations to a desired / expected indoor climate with the least possible energy consumption.

The DT should make it easy for operations managers, replacements or new operations managers to see if the building is in optimal operation.

The DT should predict how the indoor climate and energy consumption will be if the interior / floor plan changes.

The DT should be able to provide explanations for the development in energy consumption - why are the changes happening?

The DT should be able to provide more data to determine the direction and provide a suggestion for improvement, and what the potential is of that improvement.

The DT should make it easy to compare the performance of different buildings.

The DT should show the processes - and the connection between elements and contribute to the operational strategy of the buildings.

The DT should provide more knowledge on how to best operate the building.

The DT should detect deviations in energy consumption that go beyond the norm and alert the users, perhaps even take action.

The DT should be a digital engine, in which you can put in a decision, and see the consequences of it - also for managing activities.

We expect that the DT can help reduce energy consumption, intelligently monitor data, and provide troubleshooting.

The DT should provide energy management by activation of historical data in new buildings.

7 Appendix:

- 1) Interview with Aarhus kommune
- 2) Interview with Newsec
- 3) Interview with FSB
- 4) Interview with KEID
- 5) Interview with Bygningsstyrelsen
- 6) Interview with Helse Bergen (NO)

Appendix 1 - Interview with Aarhus commune

Velkommen	
Bordrunde og rammesætning fra COWI/KMD. (5-10 min) Jakob Hovgaard Kaiser kjah@aarhus.dk Søren Stig Andersen – xxxx, xxxx@aarhus.dk Jesper Lund – xxxx, xxxx@aarhus.dk Thorsten Larsen Harggaard – TEO, TEO@kmd.dk Peter Christian Damm – PDA, PDA@kmd.dk Lars Gyldenholm – LGO, LGO@kmd.dk Niels Trelldal – NITL, NITL@COWI.COM Peter Tolstrup – PETE, PETP@cowi.com Lise Lund-Hansen – LILH, LILH@cowi.com Belinda Godsk – BMGO, bmgo@cowi.com	
Spørgsmål	Noter
Tema 1: Hvor står I? (10 minutter) Hvilke strategi-/initiativer har I gang i på digitalisering, herunder digitale tvillinger?	> BIM bruges i dag alene til projekteringsmodeller for bygningskonstruktion. > Teknik og miljø overvåger ca. 1500 ejendomme. Dataindsamling og brug af data er i dag meget statisk og reaktivt og ikke proaktivt og sker primært fra et generelt forbrug/eller retur temperatur osv. dvs. fra meget simple datakilder. Der er brug for flere data til at underbygge beslutningsprocesser med flere detaljerede data og man ønsker at bevæge sig væk fra et "hovedløst" driftsprincip. > KMD EnergyKey er en cloud-løsning til datahjemtagning og kvalitetskontrol samt analyse og beregning af energidata. Løsningen kan desuden anvendes som dataplatform til forbrugsafregning, driftsoptimering, energiledelse og -rådgivning. > KMD EnergyKey indsamler i dag eksisterende data. Der er altså tale om historisk data. KMD er en software leverandør der leverer en platform som kan integreres med alle mulige systemer og bygget som open source platform.

Tema 2: Forventninger og potentialet ved digitale tvillinger? (ca. 20 min) Jeres forventninger til brugen af digitale tvillinger?	> Aarhus kommune har et ønske om at få et visuelt overblik over de bygninger de drifter i deres portefølje. Det primære ønske er at kunne opnå øget bygningsperformance og energioptimering af hele porteføljen hvor man i kommunen har et ønske om, at det skal gælde både for nyt og eksisterende byggeri. > Modellen skal indhente f.eks. forecast på vejr og forventet antal brugere, kunne vise energiforbrug, parametre som temperaturer, CO₂ men også kunne detektere et estimat af antallet af brugere alt dette ned på rumniveau. Disse parametre skal kunne bruges til at regulere bygningsinstallationerne op imod et ønsket/forventet indeklima med mindst muligt energiforbrug. > Ift. indeklima vil det også være tilladeligt at reducere kvaliteten af indeklima hvis BEM kan forudsige at der ikke er brug for "godt" indeklima når der ikke er mennesker i bygningen/området og derved undgå overventilering. Dette reguleres ikke i dag. Tilsvarende kan den klassiske natsænkning af temperaturen udfordres af andre modeller som udnytter opmagasinering af energi i bygningsmassen. Generelt skal bygningernes arealer også udnyttes bedre.
Tema 3: Udfordringer ved implementering af digitale tvillinger (ca. 20 min) Praktiske udfordringer? Strategiske udfordringer?	> Det er vigtigt at der sikres energirigtig drift af alle typer af bygninger både ift. en individuel intuning for såvel nye som ældre, men også ift. afbrydelighed da kommunen allerede har aftaler med fjernvarme leverandører om afbrydelighed således at peaks i energibehovet flades ud. > Der er også erfaring med at den eksisterende BMS (CTS) er for generel og der er brug for mere viden, da der generelt kan tages 20% af det forudsagte energiforbrug ved commissioning af bygninger ift. teori/BIM. Commissioning er statisk og der er brug for at arbejde i flere dimensioner og komme ud over 2D modeller. > Selv om benchmark fra BIM kan bruges ved overlevering, er der i høj grad brug for BEM. > Kommunen har også et projekt. BART, ift. en åben og distribueret branchestandard for automatisk regeltjek af digitale byggeprojekter som vi kan undersøge om kan give nogle synergier med dette projekt.
Tema 4: Konkret på digitale tvillinger (ca. 20 min)	> Modellen skal gøre det let for driftsledere, evt. afløsere eller nye driftsledere at overskue om bygningen er i optimal drift. > Modellen skal kunne simulere forskellige scenarier og forudsige energiforbrug eller lignende ved ændring af forskellige parametre.

På hvilke type ejendomme vil det give mest værdi for jer at få en digital tvilling? (Overblik, drift, energi?)	> Man ønsker også at bringe BIM ind i energiregnskabet og f.eks., som nævnt, opmagasinere energi i bygningen og i højere grad at få et bygningsoverblik, om bygningen lever op til den oprindelige specifikation hvor også tiden/alderen inkluderes som parameter og hvor digitale tvillinger understøtter sammenligning af bygninger på tværs af bygningsmassen og med bedre forståelse af transformationen af bygninger, bygningsdynamik ift. årstider og energibalance ved opvarmning/køling/performancetab og prioriteringer af bygningsmassen.
Tema 5: Afslutning (10 min)	Tak for i dag. Der udveksles data for Dokken nu så det kan afklares hvad BIM-modellen indeholder. Aarhus kommune skal forsøge at stille følgende til rådighed: • Bim-model og GIS data. • Adgang til data der logges i dag. • Adgang til CTS-systemet, Det er ikke tanken at CTS reguleres på ventil niveau men i stedet, hvis det er muligt, ift. set-punkt ændringer osv. • Adgang til diverse IKT-sensorer, hvis det er muligt. > Aarhus kommune ønsker, at der efter at ovenstående bliver afleveret/givet adgang til, bliver udarbejdet en fælles afklaringsliste med spørgsmål fra KMD, COWI og SDU som der kan følges op på. > Vedr. Data og logninger så skal logninger gennemgås, så vi får logninger så ofte så muligt (kan være sat op til f.eks. timeværdier hvor man kan sættes ned til f.eks. 5 minutters for at give KMD bedst mulige grundlag ved at have en finmasket logning for 2022.

Appendix 2 - Interview with Newsec

Velkommen Bordrunde og rammesætning fra COWI/KMD. Hanne Berggren Winther, Director, Technical Management, Newsec Peter Horsager Pejtersen, Fagansvarlig for automatik, Newsec Peter Christian Damm, Applied Research Director, KMD Belinda Marianne Godsk, Sektionsleder VVS og Ventilation Bygninger Øst, COWI (5-10 min)	
Spørgsmål	Noter
Tema 1: Hvor står I? (10 minutter) Hvilke strategi-/initiativer har I gang i på digitalisering, herunder digitale tvillinger?	> Newsec er professionel ejendomsadministration. Vi sidder en masse forskellige fagligheder. 47 mio. etage kvadratmeter. > Enormt uforløst potentiale i at bruge data smart. > Newsec har porteføljer hvor der benyttes EnergiKey. Data er kun relevant ift. afrapportering eller hvis de kan forøge investorernes afkast. > Hanne kender ikke rigtigt til begrebet, Peter har arbejdet med det tidligere med en model som er udarbejdet hos MIT. Der er dog mandat til at arbejde med Machine Learning og kunstig intelligens. Vi skal have nogle kraftigere databaser. Vi er et stort dataware house. Så det er next step – men vi er ikke endnu overhovedet. > Newsec arbejder ikke med BIM-modeller, og dermed heller ikke digitale tvillinger. > Afkast er alfa og omega. Hvis energibesparelser eller bæredygtige tiltag kan omsættes til et betydeligt afkast så er det interessant for Newsecs kunder. Kan man lave en bedre økonomisk model på baggrund af den digitale tvilling? Så er interessen stor – ellers er den lig 0. > GVP aftryk /CO₂ aftryk – vi har års- og halvårsrapporter. (Peter, KMD, vi skal gerne bruge den digitale tvilling til at måle og afrapportere ift. CO₂ footprint).
Tema 2: Forventninger og potentialet ved digitale tvillinger?	> Vi er ikke helt enige om hvor udviklingen går hen, i forhold til de digitale tvillinger. Det kan optimere vores forretning – kan vi få computerne til at overtage nogle menneskelige driftsomkostninger

(ca. 20 min) Jeres forventninger til brugen af digitale tvillinger?	så vil det være meget relevant. Men "man" har ikke opdaget det potentiale endnu. Værdien af den digitale tvilling er ikke tydelig endnu. > Et godt indeklima er det absolut vigtigste, energiforbruget er sekundært. Det er vigtigere for lejerne at deres medarbejdere befinder sig godt end at de har et øget energiforbrug. De primære energiforbrug i kontorejendomme er strøm og køling og herefter varme > Digitale tvillinger er interessante ved ændring af lejemålsindretninger. Hvorledes bliver indeklimaet og energiforbruget hvis indretningen ændres – det vil lejere være interesseret i. > Data skal være realdata! Ikke bygge på antagelser. Hvis det er det, så rykker det ned på bevidsthedsskalaen. Troværdig rapportering ift. den reelle bygning. Reelle målinger fra bygningerne.
Tema 3: Udfordringer ved implementering af digitale tvillinger (ca. 20 min) Praktiske udfordringer? Strategiske udfordringer?	> Vi har en 7-8 bygninger hvor man godt kunne sætte en digital tvilling op – hvis økonomien bliver stillet til rådighed. HVAC anlæg på nogle store hostede miljøer. På disse miljøer ville vi i teorien godt kunne opsætte en DT. > Der skal være et helt klart økonomisk businesscase – en besparelse for vores kunder. > At få systematiseret data, og indsamlet den rigtige data. Manuelt data er ikke god. Så det skal være automatiseret data. Modellen bliver ikke bedre end den data der er. Der kan være mange usikkerheder og antagelser forbundet med data. > Afklaring: Hvis vi har en model over energiforbrug – så skal vi jo have styr på hvordan og hvornår der skal reageres på afvigelser i energiforbruget? Skal der reageres på en daglig/ugentlig/månedlig afvigelse? Skal det være en mail man reagere med? Ellers skal det være mere analytisk reaktion? > Vores kunder er bygherrer – dem der ejer ejendommene. Typisk de store pensionskasser. Det som de efterspørger er økonomi og myndighedskrav. > Hvis lejereren kan spare penge på forbrug, så kan ejendomsejerne sætte leje op – så her har de også et incitament. > Det er en udfordring at man ikke kender potentialet. Businesscasen er super vigtig. Men man skal kunne tale sproget, for at forstå potentialet. > Bæredygtighed betyder meget for markedsføringen. Det skal lyde rigtig fint. Det kan sælges.

	>
Tema 4: Konkret på digitale tvillinger (ca. 20 min)	> Energi: Akkumulerede energi overfor budget. Årsagen til variansen, så vil det jo være alle relevante data. Hvorfor er vi der hvor vi er. Det er et kæmpe spørgsmål. Et helt CTS system. Hvilken data måler vi på : > Centralvarme, volumen på gas (centralvarme på 8 bygninger), elforbrug. > Schneider servere indkøbt. Model-sammensmeltning. > Modellen skal ud over Hvac anlægget. Mere data end det. > Behovstyret trykregulering. Hvilke typer ejendomme vil passe sig bedst til DT? > Erhvervsejendommene – jo mere komplekse ejendommene, jo mere behov for data (kompleks: køle, kombineret ventilation, store krav til samspil) Boliger vil være mindre interessante (mindre tekniske tunge). > Nye ejendomme vil helt klart være bedst. Ikke så meget gamle bygninger – hvorfor skulle man lave en 3D model af noget eksisterende? Hvad er businesscase? Medmindre man skulle i gang med en ombygning. > Energiforbruget på kølingen er stort.
Tema 5: Afslutning (10 min)	Tak for i dag. Opfølgning?

Appendix 3 - Interview with FSB

Interviewguide:

Velkommen	
Bordrunde og rammesætning fra COWI/KMD. (5-10 min) <u>Deltagere:</u> FSB: Anders, BIM/IKT mand, Mikkel, teamchef, Mads (praktikant) COWI: Belinda, Niels, Lise KMD: Peter	
Spørgsmål	Noter
Tema 1: Hvor står I? (10 minutter) > Hvilke strategi-/initiativer har I gang i på digitalisering, herunder digitale tvillinger? > Hvordan arbejder I med at bringe jeres data i spil? > Hvad forstår I ved digitale tvillinger? > Er det et begreb I kun anvender ifm. BIM? > Hvilke behov har I? Hvilke udfordringer står I over for? > Jeres fokus-/indsatssområder? Hvor sætter I ind?	<u>Indflyvning ved FSB:</u> > Meget spændende dagsorden. Vi vil rigtig gerne sådan noget her – få mere ud af vore data. Vil gerne deltage i hvad der giver mening for os – bidrage med data mv. FSB er ikke i gang med digitale tvillinger, og har også relativt enkle BIM-modeller. <u>Kort om FSB:</u> > Vi er en almen boligorganisation med 13.000 boliger fordelt på 80 boligafdelinger i Kbh. Vi ejer alle selv, modsat andre organisationer som er administratorer af andres bygninger. Vi er den største i Kbh og har mange byggesager. Vi har også et strategisk samarbejde med COWI. Både eksisterende bygninger og nybyg (i Valby f.eks.). 350 medarbejdere. 120 herinde på Rådhuspladsen. Alle afdelinger i FSB er en selvstændig økonomisk enhed. > Vi er non-profit. Vi er ikke sat i verden for at tjene penge. Lejerne ejer os. Vi er styret af en organisations bestyrelse som består af lejerne. Vi er dermed en politisk org. Alt kommer oppe fra. Vi kigger på bundlinjen i form af huslejen for vores lejere, men vi kigger også på alle mulig andre sociale parametre. F.eks. bæredygtighed fylder meget. Organisationsbestyrelsen har nogle strategier. F.eks. er der krav til DGBN certificering ved nybyg, strateg på indkøbsaftaler mv. > Vi bliver p.t. ikke målt på CO₂. Men vi har effektiviseringsmål og andre typer mål. = indkøbsaftaler. Miljø og bæredygtighed betyder

	meget qua strategier for byggeri på indkøbsaftale ift. FNs verdensmål, men også krav fra politikere. Det kan sagtens være at det kommer i fremtiden. Så måske fint at tappe ind i det. <u>Indflyvning til projektet ved Niels/Peter:</u> > SDU (simuleringsmotor). Hvad burde en bygning bruge af energi? Afhængig af belastningen i bygningen. Den digitale tvilling er ikke en konkurrent til et CTS system, men kan man lave et styrings-system, hvorigennem man kan give sin varmemester eller CTS systemet mere viden, vejledning til at styre bedre. Hvor lidt BIM skal man bruge for at kunne komme længere? Hvor lidt skal der til? Hvad er behovet? Hvad er udfordringen? Oplagt med et energitudgangspunkt - men også bredere. Hvordan rammer vi skiven? Måske en energisk dynamisk model. > Brug af dynamisk data mere end statisk data (f.eks. energidata fra EnergyKey KMD). BIM er statisk data. Dynamisk data skal måles løbende – og samles sammen. Kan man f.eks. sætte kamera på brugerne – så der er data på hvem og hvordan man er i bygningen. Hvornår bruger folk bygningen – spændende data. Opvarmningscyklus. > Output: Vi skal lave en kravspek og så en pilot når løsningen er på plads.
Tema 2: Forventninger og potentialet ved digitale tvillinger? (ca. 20 min)	Energistyring og energisimulering: > Driftsafdelingen består af administratorer og fagspecialister – teamet (energiteknikere og miljøteknikere) har specielle værktøjer til at styre og monitorere energiforbrug. Vi bruger systemet

> Jeres forventninger til brugen af digitale tvillinger? > Hvad er de primære udfordringer som bygningsejer/forvalter ift. at sikre driftsoptimering og brugeroplevelse i jeres bygningsportefølje? > Hvad kan det give jer at bruge digitale tvillinger? > Hvad vil I kunne, som I ikke kan nu? > Hvordan arbejder I med energi-data:	MinEnergi, men er ved at få et nyt. Vi har gang i en forundersøgelse, hvor vi taler med leverandører om hvilke krav vi har. Der er blandt andet fjernaflæsning og visualisering af data. > I 2017 begyndte vi for første gang at se på energisimuleringer. Vi så ind i at lave simple 3D-modeller på alle vores boliger. Det kunne gøre det nemmere for os at lave udbud på de behandlinger der er i forbindelse med fraflytning. Hvordan laver vi sådan en model, og hvordan kan vi få værdi ud af modellen. Det er et stort arbejde – så det skal give os værdi. Vi vil gerne lave elektronisk syn ved indflytning og fraflytning. Det endte med at vi lavede modeller der kunne give os plantegninger over alle boligerne. Dem kan lejerne se på hjemmesiden, og når de får tildelt en bolig. Vi har meget fokus på nyttiggørelse af vores BIM-modeller. Men de kom ikke rigtig længere end plantegningsvisualisering. På de store renoveringssager over 20 mio., er det et krav at der skal være IKT og ift. vedligeholdelseefterslæb, som finansieres via renoveringer, er en brugbar BIM model vigtig. Vi har haft lange diskussioner om hvordan vi får lavet dem. Det giver super god mening – men er også svært. Hvis modellerne skal være omstændige – hvilket giver mening – trækker det mange ressourcer. Og vi har måske ikke haft kompetencerne til det. Vi har dyrket forskellige områder – energisimuleringer. Det kom der dog ikke så meget ud af. Teknologien udvikler sig. Vi er super interesserede i at deltage og udsøge. Men historisk set har vores modeller ikke været lige så gode som vi kunne tænke os. 60 afdelinger har tegnet modellerne og de er bygget ud fra eksisterende tegninger og kontrolmål på steder. Men nogle gange passer de ikke på virkeligheden. Det er jo frustrerende. Eseebase (digitaliseringsfirma under Årstiderne Arkitekter), og måske COWI, var lidt skuffede over niveauet for vores BIM-modeller. > Potentiale ved BIM. Især ved råderetssager – hvor der f.eks. skal sættes en væg op i lejligheden. Det kræver nye planer. Det er et stort vedligeholdelses arbejde. Vi skal have knowhow internt, for at kvalitetssikre. Det er først for et par år siden at vi har fået interne BIM-kompetencer. > FSB's BIM modeller: Vores modeller har ikke så højt informationsniveau, de er ret skrællet. Men har en facade, vinduer mv. > > Potentiale ved DT: Nedsættelse af effektregulering hos HOFOR. Vi mangler dokumentation til HOFOR ift. fjernvarme, hvilken tilkoblingskategori som bygningen skal ligge på. Stort besparelses potentiale. (Vi betaler for den store tv-pakke, men ser kun dr). Det er en stor kamp med HOFOR og vi har ikke knækket koden. Tænk hvis man kunne få data til det? Optimere vores kendskab til vores bygninger – og kan dokumentere det. Det er dog ikke helt klart hvad HOFOR skal have af dokumentation for at
--	--

	ændre kategori. Vi er meget optaget af nedsættelse af effektregulering. Ved at mindske ydelsen fra fjernvarme kan man spare mange penge. Høfor skal være enige om den energivurdering, men mangler data, så det er en stor slåskamp for os. HOFOR anerkender ikke ændringer. > Potentiale ved DT: Energimærke rapporter er ikke relevante som input men vil i høj grad være interessante som et output fra modellen. Som input er de alt for upræcise. Lovpligtigt at opdatere energimærke rapporter. Data i rapporten er svært (en lejer kan have 18 grader, en anden kan have 26 grader, det påvirker jo i høj grad forklaringerne på energiforbrug og denne nuance ses ikke i rapporten). En automatisk genereret energimærkerapport. Det har umiddelbart stor interesse hos FSB. > Potentiale ved DT: årsagsforklaring til udsving. Give forklaringer på udviklingen i tallene – hvorfor sker forandringerne? Automatisering af forklaringen, så den ikke er håndholdt. Det kunne give mere data til at sætte retningen. Er forbedringen som den burde være? Hvad er potentialet i den forbedring? Skal det bare være bedre end sidste år, eller er der andre forventninger? Har to bygninger samme potentiale? Data fra BIM-modellen skal ind sammen med energimålingen (EnergyKey data og anden dynamisk data). Vi har historisk data som kan hjælpe os med at identificere et potentiale. Data på forklaringen. Varmen stiger – men der var også mange flere mennesker hjemme – det kan data vise os. Stor interesse hos FSB. Der er ellers en opfattelse af, at man ikke kan få denne data fra systemer. Men hvis man kan det, vil vi spare utrolig meget tid. = god businesscase for os så. > Potentiale ved DT: benchmarking på tværs af bygninger/afdelinger. Detaljeringsgraderne i energimodellerne skal hjælpe til at to boliger er sammenlignelige. Fair sammenligning jo bedre vi forstår de enkelte lejligheder. Måske ikke altid fair at sammenligne to afdelinger -der skal flere detaljer i spil. Så de bedre kan sammenlignes. Måling op imod en potentiel besparelse. Hvor tæt på idealmålet kan vi få bygningerne? Eller som vi kan forvente af denne bygning. Lige nu kigges der ofte på en middelværdi af ejendomsmassen – og det er vel ikke helt fair. Der er brug for forbehold over for de enkelte forhold ved den enkelte lejlighed. Udligne nuancer. Mere korrekt grundlag for styring. Relevant data for FSB: > Vi er meget optaget af at reducere CO₂. Det er dog ikke et krav. Men det kan jo være at det kommer. > Interessant med dynamisk data: Hvad er potentialet? – dynamiske data. Men det statiske er også interessant for FSB. Hvad skal
--	---

der til for at modne data for at bruge det? Og hvad kræver det at vedligeholde det. Vi har 5-15 bygninger på hver af de 80 afdelinger. BIM modeller på det hele. Man ved godt at det ikke er helt opdateret. Så der er altså et stort optimeringspotentiale.

- > **Obs-punkt:** Hvad sker der når VVS'er river radiator ned? Det skal jo ind i modellen. Vi skal ikke sætte en mere kompliceret model op, end at den kan vedligeholdes.

Data vi gerne vil have:

- > **Grønt regnskab med Excelfil samt BIM-model for samme bygning.** = Ud fra den data kan vi så lave et case, hvor meget skulle der til for at vi kan give noget mere. Fremtidige BIM-modeller, hvad skal den have for at vi kan bruge det til noget ift. den digitale tvilling. Hvor meget skal der bygges på?
- > **Kan vi få en version af den allermest simple BIM-model I har, samt en version af den mest detaljerede BIM-model I har** (evt. på nybyg). Så vi får en forståelse for skalaen.
- > **Viden om hvilke data I har.** Og hvordan? Hvad er der for data? Hvad samler I op af energi data? CTS data? Open source. CTS – systemet, i hvor høj grad regulerer det sig selv? Intelligent styring ift. køl og varme? Er der snitflader til dem? Og har I overhovedet køl?

Hvilken data har FSB:

- > Al nybyg er på CTS. Vand, el, energi, generelt forbrugsdata.
- > **Spørgsmål: Har I Bi-målere på enkelte lejemål? Viden om forbrugsdata?**
- > **Vi får meget data gennem DGNB** – herigennem får vi systematisk data som er generede under projekteringen. Det viser sig dog i praksis at det teoretiske grundlag ikke er så præcist – det finder man ud af når det skal driftes (f.eks. et stort vindue som skal pudses, det indgår ikke altid i det teoretiske grundlag). Vi får meget data fra DGBB men vi følger ikke op på det.

	> Data fra de grønne regnskaber: Data fra 2001. Det er ikke data på time-niveau, men på månedlig aflæsning. Der er ikke automatisk fjernaflæsning. Det er dog fremtiden. Vi kan evt. kigge FSB's grønne strategi igennem – for at se på måltal og data. > CTS: Der er CTS anlæg på alle nybyg, og på nogle af de eksisterende. Denne data må vi gerne få, og se på. > Økonomidata. 30 års budget på en bygning, som det opdateres hvert år. Især konto 116 – planlagt periodiske vedligehold (PPV) – er interessant. Vedligehold til tag, vinduer. Det sparer beboere op til gennem henlæggelser. Drøm om at lave bæredygtige planer - bedre og mere bæredygtig langtidspanlægning? FSB mangler konceptet. Når der f.eks. skal udskifte et tag: her kunne det jo være spændende at se på data. Simulering på tilbagebetalingssplanen. Hvornår skal taget skiftes ud? Hvad er varme forbruget på de øverste lejligheder, end de lavere. Er der noget interessant her? Planer ift. hvem er mest bæredygtige, mangler koncept om hvad der kan måles på, DGNB måske? Kom med anbefalinger ift. vedligeholdelsesplan. > Sundhedstjek. Hver 5 år laver FSB en obligatorisk ydelse. Sundhedstjek hos alle lejere. Spørgeramme på 20 spørgsmål. Er der drypbakker på alle opvaskemaskiner – ellers er det miljø og energiteknikere der laver spørgsmål. Personale ind i samtlige boliger og svarer på alle 20 spørgsmål. Den data kommer tilbage til os. Den data har vi også. Ikke hardcore forbrugsdata, men boligens stand. > Vi har voldsomt meget data. Beboerserviceopgaver, gennemsyn af boliger. Men også registrering af pasning af varmecentral, beboerservice, reparationer mv. Vi kan trække data på meget. > Hvor hurtigt skal I kunne reagere? Hvor tit har I brug for output? De grønne regnskaber laves en gang om året. Tanken er at de skal bruges som dialogværktøj sammen med tilstandsvurdering på et årligt møde. Derudover ser vi en gang om mdr. i MinEnerg i om der er ting der stikker af. Vi kunne godt tænke os at der er mere automatisering. Tricker baseret. Sms hvis der sker noget – der vil vi gerne hen. Den digitale tvilling skal give intelligent overvågning. Vi skal væk fra, at der er et menneske der bliver tricket. Det er AI der skal alarmeres i det øjeblik når der sker noget unormalt. Der er ingen grund til at der er mennesker der skal holde øje. Det handler derefter om at AI skal adviseres de rette folk. Note fra Peter:
--	---

	> BIM-modellerne er ikke nok (siger SDU). Vi har brug for nogle digitale tvillinger som går ud over BIM. Med inkludering af alt det der også ligger uden om bygningerne og relationerne i bygningerne. Det er en teknisk mangel ved BIM – BIM indeholder ikke relationer mellem de enkelte enheder, f.eks. om hvordan ting hænger sammen i bygningen. F.eks. radiatoren som trækker vand fra andre steder i bygningen. Den data ligger ikke i en BIM-model. Der er derfor behov for mere omfattende modeller end BIM. IFC formatet er desværre ikke altid løbende blevet opdateret, så vi bruger BIM modeller og IFC-data som datakilder til den digitale tvilling, samt også anden relevant data (Energykeys).
Tema 3: Udfordringer ved implementering af digitale tvillinger (ca. 20 min) Praktiske udfordringer? Strategiske udfordringer?	> Kompetencer. Kun 1 mand sidder med BIM. [står også lidt i det ovenstående]
Tema 4: Konkret på digitale tvillinger (ca. 20 min) > På hvilke type ejendomme vil det give mest værdi for jer at få en digital tvilling?	Hvilke typer bygninger/porteføljer? > Nybyg eller eksisterende? – begge kunne være interessante. Vores gamle byggeri er en stor synder ift. CO₂. Så her vil CO₂ jo være centralt. Vi ville skulle kigge på både enkelt bygningsniveau, og på porteføljeniveau.

(Overblik, drift, energi?) > Hvilke krav er der til brugerfladen? > Hvad skal kunne overvåges? > El-forbrug, energiforbrug, vandforbrug, brugerbelastning, mv.? > Overvågning på hvilket niveau? > Bygning, etage, område, lokale?	> Vi kunne sagtens finde på at benchmarke afdelingerne op imod hinanden. Det gør i allerede i de grønne regnskaber. Kan det regnes om til CO₂? CO₂ benchmark? [står også lidt i det ovenstående]
Tema 5: Afslutning (10 min)	Tak for i dag. Opfølgning?

Appendix 4 - Interview with KEID

Velkommen	
Bordrunde og rammesætning fra COWI/KMD. (5-10 min) Deltagere: COWI: Belinda, Niels og Lise KMD: Peter KEID: Michael Ørsted, Anders Christian Lyngtorp og Sebastian Bruhn	
Spørgsmål	Noter
Tema 1: Hvor står I? (10 minutter) Hvilke strategi-/initia- tiver har I gang i på digi- talisering, herunder digi- tale tvillinger?	> (Niels Note: Performance i udbuddet. Kontrol op imod bygher- restandarder? Er det tænkt ind? Benchmarker. Store forventninger til de grønne afrapporteringer. Cirkulære økonomi skal kunne repræsenteres i den digitale tvilling.) Mere drift scenariet end anlæg. Brugen af digitale tvillinger skal kunne gøre energimodellen dynamisk. BIM-modeller i KK > Vi skal have mere viden om de forskellige aktiviteter. Vi har et forholdsvist lavt informationsniveau. Vi har sat ressourcer af til at alle bygninger skal have en 3D model – i REVIT. Det kommer vi til at bruge det næste års tid på. Det er i alt omkring 2.6 mia. kvadratmeter – alle modeller i BIM. > Vi er ikke gode nok til at bruge den data der er til stede i BIM. Vi har et ok IKT grundlag. Det kræver kvalitetssikring at bruge den data i den anden ende. Potentialet er der. Lille pilot med TMF – kunne deres by data og vores bygningsdata ind i en model. Ejerskab og renhold. Kunne det give os noget? Den digitale tvilling 1.0 > Vi vil gerne derhen hvor vi har en motor – hvor vi kan putte en beslutning ind. Og se konsekvens af det. Hvilke typer af udstyr der kan sætte ind. Herunder også styring af aktiviteter. Forudse tem- peratur i håndholdhallen, f.eks. Vi er forholdsvist umodne men vi vil gerne derhen. > Der er flere internationale standarder – vi skal finde den rigtige data model. Og have en intelligent styring af vores bygninger. Bedre udnyttelse af nogle standard systemer.

Tema 2: Forventninger og potentialet ved digitale tvillinger?

(ca. 20 min)

Jeres forventninger til brugen af digitale tvillinger?

Hvor ser I det største potentiale?

- > En rigtig tvilling skal også vise processerne – og forbindelserne mellem tvillingerne. Linket data.
- > Spare drift-folk eller energi? Svært at sige. Vi kender ikke de socioøkonomiske gevinster. Gaspriserne stiger – så det ville tale for energi-dagsordenen. Der er hurtige gevinster ved energidagsordenen. Økonomi er gevinsten. Hvad er det for en matrix der er rammesættende for hvilke trin vi skal sætte gang i.
- > På BIM-delen får vi value for money? Vi skal skrue op for objekttankegangen for vores modeller, for at blive bedre til at knytte data sammen.
- > Indtil vi er der hvor det hele er automatisk og som det skal være. der er der en kæmpe udfordring i at få driftspersonalet til at drifte deres ejendomme så det er optimalt. Give dem adgang til de systemer der skal til. Visualisering af data. Deres smartphone skal komme med vejledninger til pedellerne – i dag skal du gøre dette. Det er altså også en stor ting. At få værdi ud af systemerne, kræver stor tilpasning og kompetencer blandt driftspersonalet.
- > Bygningerne skal have en strategi – man skal ikke bare skrue op fordi det ønskes af brugerne. Der må være mere strategi bag. Hvordan drifter man bedst den bygning? Det kunne være fedt med viden om det. Implementering – dokumentation for om der er reel problemer. Er det faktisk koldt?
- > En censor pr. 100 m². Måling af CO₂, fugtighed og temperatur. Skal den være kablet eller trådløs. Nogle steder er det ok at de er kablet. Børn hiver dem ned. Så er det bedst med trådløs. GDPR problemer – krænkelse i privatlivet. Der skal være en vis mængde personer i rummet – for at komme uden om det.
- > Wi-Fi triangulering – enheder fra telefonerne. Router pr. kvadratmeter. Kan man lave en integration til indeklima?
- > Open source platform. Stor bekymring: hvordan får vi holdt det hele ved lige. AI skal indover for at detektere afvigelser, som træder uden for normen.
- > KEID skal lave løsninger som taler ind i KK's samlede plan. Derfor også vidensdeling til den private sektor.
- > KEID har et samarbejde med Cleaver – datastandere. Linked Data strategi. Vi skal være tydelige omkring data governance. Markedet skal også være klar. Vi skal koble op imod samarbejdspartnere og leverandører.

	> Internationale standarder for typologier og platforme. Er det så LinkedData der er den hellige gral. Blockchain? Vi ved ikke nok om det.
Tema 3: Udfordringer ved implementering af digitale tvillinger (ca. 20 min) Praktiske udfordringer? Strategiske udfordringer?	Udfordring: Vedligeholdelse af BIM-data > Det er et stort problem. Vi skal have accepteret at der er et vist Gab. Det er et stort projekt. En model som ikke repræsenterer virkeligheden. Det er et problem. > I vores servicekontrakter kan man måske godt sætte flere krav ind, men der kan desværre være langt til praksis – det er med at de skal vedligeholde data omkring ventilation. Anlæg er mærket op med en QR kode. Stamdata skal opdateres. Hvor meget skal modellen dække – det skal afklares (f.eks. sidder ventilationen på den ene eller anden side af bygningen – det skal være præcist for at man kan bruge data korrekt). CTS kode niveau.
Tema 4: Konkret på digitale tvillinger (ca. 20 min) På hvilke type ejendomme vil det give mest værdi for jer at få en digital tvilling? (Overblik, drift, energi?)	Hvilke data har I? > Miljømærker - her er en del data. Hvad er det for nogle bygningsdele der kan optimeres. Det er dog ikke kædet sammen med vedligeholdes efterløb. De data om planlagt vedligehold og energioptimeringsmuligheder skal kobles sammen med miljømærkning. Det kan give mening. Det overblik har vi ikke. Her har vi lavet en opgørelse over data - EERA data (Lise: ??). > Energiovervågningsdata. Beskrevet data i samme projekt. HOFOR-projekt. EnergyKey data pr. ejendom og på bygningsniveau. Hvilke bygninger bliver forsynet fra hvilke hovedmålere (vand, varme, el). Logget en gang i timen, vi modtager data pr døgn. 98 % af alle vores bygninger. Det har været et krav at hver varmepumpe kan trække data ud. > Styringsautomatik – 98 % på alle vores bygninger. Hvordan skal de styres. Lidt forskelligt hvilken data der kommer ind. Men ikke CO₂ målinger på rumplan. Ikke temperaturmålinger på alle vores arealer. Det arbejder vi på en løsning på at få det. Vi bruger noget styringsautomatik med detaljeret viden om hvordan energien bliver forbrugt på fjernvarme – Lean Heat. Her styrer vi opvarmning på vores ejendomme (ift. hvordan temperaturen er på forecast og hvordan situationen er på ejendommen). Planlagt og forecast. Så vi kan starte opvarmningen af vores ejendomme på et givent tidspunkt, til en bestemt temperatur. Og nedskalere på samme måde. Også ift. peak load. Lige nu har vi dette test-setup på 20 ejendomme og vil få det på 100 ejendomme inden et år. > Vi har i alt 3800 ejendomme.

	> Kunstig intelligens / Machine learning profiler: Forecast vejr data og bygningsdata. På selve fjernvarme enhed og indeklimate sensorer op. Krydse data så man sikrer at komfort opretholdes. Og hvordan reagerer bygningen. Spiral af forskellige datapunkter, hvor vi kan lave en profil – hvis vi skruer ned et sted hvordan påvirker det så andre steder (system: LeanHeat). LeanHeat er ikke kun afgrænset til Danfoss produkter. > Bygningens brug – hvor mange mennesker er det og relationer – er ikke med i LeanHeat. LeanHeat kigger ikke på indeklimadata – det vil vi gerne have med. Vi giver styringen til HOFOR, uden at kunne kontrollere indeklimate i ejendommen. Vi er meget interesseret i at få vores indeklimamålinger ind, også ift. CO₂. Så hvis vi har indeklimate målinger kan vi se hvordan arealerne bliver brugt. Og vi kan sikre os at der hverken er for koldt eller varmt. > Indeklima: Vi forestiller os at den effektivisering vi ser, også kan imødekomme den finansiering det kræver at få indeklimate med. Det kigger vi også på ift. ventilation. Vi skal have mere af det hvor systemet selv bestiller elementer der mangler – men det kræver at data er korrekt og afstemt. Ellers nytter det ikke. Det næste step er, især for kontor og andre bygninger, indeklimate skal individualiseres til bygningen. AI skal i høj grad til at regulere det. Det er et nyt projekt, som bygningsstyrelsen har kørt med Danfoss. "Den digitale fjernvarme mester". LeanHeat løsning. Ikke kun skrue op og ned, men gå længere ind. Men den teknologi har vi ikke i ejendommene. Det ville være nemt at lave – men vi skal have businesscasen. Hvorfor er det at det har en værdi? Hvorfor skal vi have fjernaflæsning? Hvorfor er det noget værd? Vi er en politisk institution. > Vi har indeklimamålere på 20 snart 100 ejendomme. Her kan vi se hvor meget folk lufte ud, og hvor de glemmer at lukke vinduerne. > CO₂ - hvor mange mennesker er der til stede. Bruger vi møderummene rigtigt? Det er meget potentiale. > Hvor meget data har vi brug for? KEID spørger om vi har en matrix? Hvis vi har de her type data, så kan vi lave det her. Hvis vi kommer længere ned i data, så kan vi lave det her. Niels: Hvor lidt BIM kan vi nøjes med? Sebastian: Matrix over datapunkter/kilder. Det kunne give os grundlaget for hvad vi skal og hvad værdien er. Det er nok det der har været svært for os. Hvad får vi ud af at tracke vandvær. Vi skal hele tiden tænke i businesscase og genomsigtighed. Balancen – projektet kan måske hjælpe os med at sige du behøver ikke at være på "guld-niveau". Hvilke bygninger:
--	---

	> Bygninger som har en fleksibel varme – som f.eks. børne institutioner. Bygninger under 3000 kvm. Det er også de mest sårbare. Her der bliver klaget mest over CO₂. > Grænseflade mellem CTS og LeanHeat. LeanHeat har adgang til DCL. Men det bliver en sammenkoblet integration. Danfoss har købt LeanHeat. Det er HOFOR vi udliciterer til. Vi laver profiler i CTS hvor vi bestemmer hvornår HOFOR må gå ind. Det er et meget følsomt område. Så her er målet at vi får opbygget en indeklimas overvågnings platform. > Censorer: Tommelfingerreglen er en censor pr 100 m². 80 % af arealet har indeklimamåling. Det er kravet. Vi forbereder en businesscase med håb om at vi kan rulle indeklima ud i alle bygninger. Også på de bygninger hvor vi ikke har CTS styring. > HOFOR samarbejde: HOFOR kan ikke slukke for varmen, men ned til 50 %. Det er et pilotprojekt. Men her vil det jo blive en aftale med HOFOR som skruer op og ned (ift. nogle forbrugskriterier (profiler som vi laver)). Hvilke krav er der til komforten ift. brugen af ejendommen. Lige nu har vi forbrugsprofiler på skolerne, her kan vi jo godt skrue ned for varmen i påsken. Tidsprogram. Men hvis vi skal gøre det i stor skal, så skal vi kende forbrugsprofilerne meget nøje på vores ejendomme. Hvis vi skal give HOFOR lov til det. Vi skal som ejendomsejer kende vores profiler. Det kan hjælpe so til at bruge strøm når det er grønt og skrue ned når det er sort. = Opmagasiner af varme i bygningen. Andre data? > LCA – total økonomisk på det. Det kunne også være. Måretning af hvad bygningerne skal bruges til. > Byrumstankegangen. Hvordan udnytter vi vores bygninger bedre? Udnyttet kapacitet. Multifunktionelle bygninger. > Vi kunne godt stille flere krav til vores bygninger. Og udnytte ejendommene. Vi mangler businesscasen: Hvad er det værd at identificere alle de bygninger der er for varme? Det burde kun være måske 19 grader på alle skoler (standard i Paris) – sådan en politik kunne vi godt have i dk. Men vi har ikke data. Det er bevidst at eleverne får højere karakter med godt indeklima. Der er så mange cross-overes – som vi skal have mere viden om! > Energi forbrug skal ind i den store infrastruktur. Der skal bygges så vi sparer i driften. En ting er at præsentere data – se de røde pletter – noget andet er at tydeligere angivelse af driftssceneriet. DGBP standarder – det kræver viden blandt dem der drifter bygningerne. Varmemestrene. Man lufter ikke ud – det klarer
--	---

	systemet. Det skal folk jo vide – og handle efter. Ellers er det lige meget med systemet. Kan man slippe styringen: Hvordan får man den bedste styring af bygningerne? > Vi kan kun sætte en DT op på de anlæg som kører optimalt. CTS systemer. Visuel rundering. Alle set punkter går igennem en gang hver halve år. Det behøver man ikke gøre manuelt (BSS) det kan man gøre automatisk. Det vil være en kæmpe besparelse for os. Optimering af driften pba. data. > I den offentlige sektor har vi svært ved at finde optimering i kroner og øre på lønkroner. Man sætter bare de ansatte til at lave noget andet, hvis der effektiviseres et sted.
Tema 5: Afslutning (10 min)	Tak for i dag. Opfølgning? > KEID vil gerne være med som sparring undervejs.

Appendix 5 - Interview with Bygningsstyrelsen

Velkommen	
Bordrunde og rammesætning fra COWI/KMD. (5-10 min) Mathias Hjorth Løvhøj fra Bygningsstyrelsen (ansvarlig for BIM og IKT) i Center for Byggeri. <i>Bemærk: Drift/energi er dermed ikke en del af dette interview.</i>	
Spørgsmål	Noter
Tema 1: Hvor står I? (10 minutter) Hvilke strategi-/initiativer har I gang i på digitalisering, herunder digitale tvillinger?	> Vi er længere i forhold til at bruge bygningsmodeller aktivt i projekteringen end vi er i forhold til at bruge bygningsmodeller i driftssammenhæng og simuleringer (IoT og sådan). > Ved at starte et projekt op for at gøre modellerne mere aktive ift. drift – men vi er helt i starten og vi har dermed ikke så mange erfaringer her. > Vi modtager nogle ret gode modeller når vi afslutter et byggeprojekt – modellerne er på et ret højt niveau. Vi skal dog have styrket vores IT-understøttelse (systemerne) og kompetencer, for at få modellerne til at lave i den videre sammenhæng. Og her er vi ikke så langt endnu. BIM på ejendomme: > Vi har netop ansat to folk til at digitalisere modellerne - til 3D modeller. På den måde er vi kommet i gang med den eksisterende masse. Lige nu er det med henblik på huslejeregulering – men det åbner for mange ting. Vinduer osv. Krav til bæredygtighed. Hvad forstår I ved Digitale Tvillinger? > Lige nu handler det meget om BIM. Byggeriet er meget fragmenteret – så det gør det måske lidt svært at komme fra BIM og over til driften i den digitale tvilling. > (Peter KDM: Men der er mere i det end at se på en 3D-tegning. Den digitale forståelse skal ind i driften.) Snitfladen mellem byggeri og drift:

	> Forretningsmodel, driften i Bygningsstyrelsen: Vi bygger og forvalter, men vi drifter dem ikke. Det er udliciteret. Leverandører der skifter hver 4 år. Det gør det hele bare lidt mere komplekst. > IKT aftale – skal censorer ind her. Det er et spørgsmål om pengekassen. Det er ikke byggeprojektet der skal finansiere en bedre drift. Så det bliver et praktisk problem, og en udfordring for implementeringen i driften. Der skal være nogle krav til overlevering af data fra byggeri til drift. Det er svært, men en forudsætning. > Driften er ikke langt i forhold til at bruge data smart – det er sådan i hele branchen. Bygningsstyrelsen bygger for mange forskellige driftsherrer. De bruger mest 2D tegninger – og er ikke så meget inde i 3D universet. Driftsherrerne har måske ikke kompetencerne eller IT der understøtter det. Det kræver platforme. > Den digitale tvilling kan jo også startes op i drift – den skal vel ikke nødvendigvis starte i byggeriet? Det skaber sikkert nogle andre udfordringer hvis den starter i driften. Det kræver jo at data er retvisende. Man tør ikke tage andres 3D modeller – man starter forfra. De sidder og tegner de samme vægge igen. Hensyn til risikoafdækning – men dobbeltarbejde. > Peter, KMD: Kan det være en motivation i overlevering af bygningen at den drifter som den skal? og at man kan påvise dette? (gennem DT) – mere end bare en statisk dokumenteret energimærkning. Mathias: Der laves test på driften – men det har ikke så meget at gøre med bygningsmodellen, mere anlægget. Det vil helt klart være relevant at have masser af driftsdata med i modellen (CTS/ventilation/køling/brugen).
Tema 2: Forventninger og potentialet ved digitale tvillinger? (ca. 20 min) Jeres forventninger til brugen af digitale tvillinger?	Forventninger: > Lige nu er vi i en behovsaflklaringsfase. Hvilke behov er der når man drifter en bygning. Vi ved ikke helt hvad vi har brug for i drift. Snitfladen mellem byggeri og drift. > Det er brugeren af bygningerne som kan få noget ud af den digitale tvilling. Visning på dashboard. Hvor meget strøm bruger bygningen i dag – nudging af forbrugeren. Viden om energiforbruget. Brugerfladen har spændende iagttagelser. Få bygningen til at performe bedre, billigere og mere bæredygtigt. Ikke varme op om aftenen – det har man styr på. Men det kan også være i løbet af dagen, eller ferier. > Et reduceret energiforbrug vil stadig være i Bygningsstyrelsens interesse – også selvom driften udliciteres. Vedligeholdelse af systemerne – det er det der bliver udliciteret. Det er stadig vores ejendomme. Det vil derfor stadig være til vores interesse.

	> Vi har 4.200.000 km² ejendomme som vi forvalter. Også nogle gamle ejendomme – og vi kan ikke renovere det hele på én gang. > Der kommer helt sikkert nogle krav i BR23 – som vi skal forholde os til. Mængderne i bygningerne bliver især relevante. Lige nu udbyder vi ikke på mængderne. Der er meget stor forskel på hvordan man tegner og hvilke mængder det bliver, og så det reelle materiale forbrug. Så det er en udbudstekniks beslutning at vi ikke har mængderne med i udbuddet, det er ikke en bæredygtighedsbeslutning. Har I krav til CO₂ reduktion (fra direktion/politikere)? > Det er jeg ikke bekendt med. Vi er en politisk styret organisation. Det kommer oppefra hvis det kommer. Tid og kvalitet er meget fokus på der hvor jeg sidder (byggeteknisk kvalitet / holdbarhed / undgå fejl i udførelsen – ikke kvalitet ift. bæredygtighed). IFC: > (Peter, KMD: IFC er ikke beskrivende nok. Eller dynamisk nok. IFC er den åbne standard – men detaljeringsgraden er der ikke. Og ikke dynamisk opdateret. IFC er for gammel en standard. Det er et stort problem – men der er heldigvis andre standarder. SmartCities / LinkedData. Dataelementer er sammensat i relationer. I BIM er der ingen relation mellem systemer og enheder).
Tema 3: Udfordringer ved implementering af digitale tvillinger (ca. 20 min)	BIM i byggeriet: > Format fra projekteringer – de nyeste IFC formater er ikke opdateret. 70-80 % af den danske byggebranche bruger Revit. Så hvis AutoDesk ikke kan eksportere til det format – så er det urealistisk at implementere. Der skal være en mapning mellem standarder. > (Peter, KMD: BIM-BEM -> det er målet. Alt fra IFC mappes over i den nye standard. Vidensgraf vi bygger op. Og de værktøjer kan også leveres videre. Vi kommer ikke til at ændre på byggebranchens måde at gøre det på.) Software forudsætninger (udfordring fra Mathias): > Der er både udveksling mellem forskellige typer software. Det skifter, og skal vi lave en mapping eller noget teknisk snart. Det kan man ikke selv gøre i denne branche – det har vi ikke kompetencerne til – og det har entreprenørerne heller ikke. Vi er derfor afhængige af underleverandører. Noget andet er udbudsloven – vi kan komme til at lempe nogle krav så flere kan byde ind. Det er vi nødt til. Der er flere der skal kunne byde. Ellers skal man dele det

	for meget op. Det gør det lidt svært at lave disse datamæssige manøvrer – sådan er det at være en offentlig bygherre. Vi er underlagt nogle regler – og det påvirker vores udbud og dermed også vores opgaveløsning. > F.eks. kan det hele gå i stå hvis entreprenøren/samarbejdspartneren ikke har 365/Sharepoint. Det er det svære i det digitale. At skifte mellem platforme. Vi kan ikke få en platform der kan tale hele. Det vigtige er at lave specialiserede værktøjer der kan tale ind i de andre. > (Peter, KMD: Vi skal fastholde standarder og systemer, i en ny platform - til brug for AI og ML. Vi arbejder med datagrundlaget. Det her kalder vi digital tvilling. Det er et godt begreb for at få folk til at forstå sammenhængen. 1:1 sammenhæng mellem det du ser i verden helst i realtid.) > Data skal ligge det sted hvor det bliver skabt – og så hentes det ind i platformen. Dataspace – Datafabric. Blande de data man har brug for. > Data er født af mennesker – og dermed tit ustruktureret. Hedder det en door eller en dør. Vi lægger os op ad internationale standarder – door i mappningen.
Tema 4: Konkret på digitale tvillinger (ca. 20 min)	<i>Her mangler lidt da vi i dette interview ikke dækker energi/driftssiden.</i> > DT kunne være relevant på både eksisterende byggerier – og nybyg. > Bygningsstyrelsen er meget interesseret i at følge med i den. Observatør. Invitere ind så de kan følge med.
Tema 5: Afslutning (10 min)	Tak for i dag. Opfølgning?

Appendix 6 - Interview with Helse Bergen (NO)

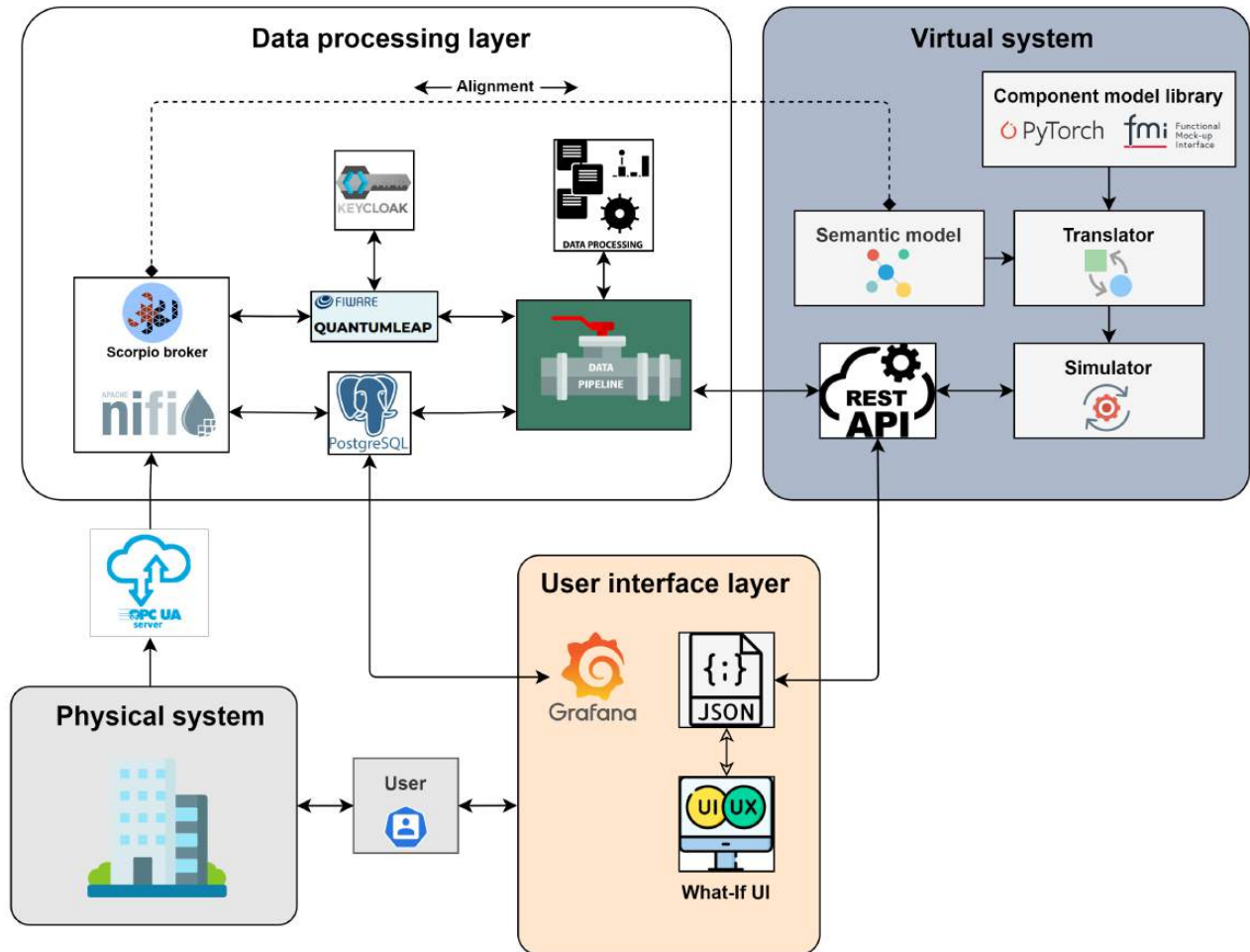
Velkommen	
Bordrunde og rammesætning fra COWI/KMD. (5-10 min) Deltagere: Torgeir, Kyrre, Lise fra COWI Kristian Branseth og Tord Monsen, Bjørn Blomvik fra Bergen Helse.	
Spørgsmål	Noter
Tema 1: Hvor står I? (10 minutter) Hvilke strategi-/initia- tiver har I gang i på digi- talisering, herunder digi- tale tvillinger?	> Vi er startet på nul i vores projektet, og har bygget på undervejs. I udgangspunktet skal vi bruge samme systemer i driften som i byggeriet. Vi jobber for at hvordan vi skal håndtere data og bruge vores kompetencer. Hvilke processer skal vi igennem når vi skal i gang med ombygning? Alle de problemstillinger vi arbejder med i byggeriet – forsøger vi at tænke ind i driften. > Vi har forskellige systemer – det er svært at koble dem. Det er gammeldags systemer. > Vi forsøger at koble projektkomponenter ind i de systemer vi har. Det lader sig at gøre – men det kræver beskrivelser/arbejdsgange. Metoderne skal være tydelige. Men vi ser at de i driften bliver vant det. Modstand er ikke et så stor længere. Især på Bygningsdelen. > Meget med data – handler om at konkretiserer – og det er lidt svært. Afgrænse data. Vi har mange data – det er en lang proces at jobbe med data. Som er bra. Men det kommer fra projekterne – og skal ske på projekterne. > Vi behøver en samkøring fra byggeri til drift. Det er svært. > Ikke krav til driften at bruge BIM-model? Nej – ikke krav. De får det som de får. Hvorfor gør vi det? Vi ser på alt der er muligt. Hvordan forstår dere den digital tvilling? > Et levende billede af bygningen. Akkurat nu. F.eks. temperaturmålinger. I dag bliver alt logget. <i>Hvor ofte?</i> – Hver 15 minutter. Det er forskelligt. Hver 5 sekund på nogle ting som er mere interessante. Typisk energi data. > Bruge data, hvor er potentialet: F.eks. på fejlmeldinger/fejlkoder. Overvågning. Bruge data til at finde fejlen. Det er et stort potentiale.

	> Maskinlæring og AI. Det skal vi mere og mere ind i. Men vi starter i driften. Det er en defineret ambition. > En enkel digital tvilling fra projekt til projekt – bruge data fra en gammel bygning til en ny bygning. Store bygninger vi er i gang med. Vi implementerer metodikken i driften af alle bygningerne. Vi bruger samme krav i nye projekter. IoT på udstyr – det er spændende. > Forbedringskultur – vi bruger erfaringer fra tidligere projekter.
Tema 2: Forventninger og potentialet ved digitale tvillinger? (ca. 20 min) Jeres forventninger til brugen af digitale tvillinger? Hvad er de primære udfordringer som bygningsejer/forvalter ift. at sikre driftsoptimering og brugeroplevelse i jeres bygningsportefølje? Hvad kan det give jer at bruge digitale tvillinger? Hvad vil I kunne, som I ikke kan nu? Hvordan arbejder I med energi-data:	> Forudsætningerne er udfordringen. At data logges. At selve objektet findes i en database. Højere temperatur i et rum – det er viden. Det kræver at vi får defineret hvad der er en fejl for driften (høj temperatur kan være en fejl, men er det ikke nødvendigvis). Viden om hvor mange mennesker der er i et rum. Relevant data: > Statisk kontrol. Alt skal være smart. Udstyr? Mobilt udstyr. Stort behov for – sambrug. Data om brugen af udstyr, kan optimere brugen af udstyret. > Censor på taget. Medicinkøleskab. Det er alt for mange muligheder. Mærkningssystem. > Prioritering af data. Der er mange muligheder. Rapporteringskrav? > Det gør vi i forvejen. Kan en digital tvilling give jer mere præcis afrapportering? Materialebrug? Det gør vi i forvejen? Drift er det største CO₂ forbrug – varer og maskiner. Mere end de materialer der bruges i byggeriet. Det er ikke et behov at materialer er en del af den digitale tvilling. Mærkning: Usystematisk kortlægning af mærkning - dobbeltmærkning i modellen – ikke fysisk.
Tema 3: Udfordringer ved implementering af digitale tvillinger (ca. 20 min) Praktiske udfordringer? Strategiske udfordringer?	Vedligeholdelse af data: > Det vil kræve nogle nye systemer og opfølgninger. Vedligeholdelse af data/opdateret. Hvis ikke, så fungerer det ikke. > Driftsafdelingen skal føle ejerskab til data – for at data er rigtigt. Nu har vi kontrol på mange ting – og så skal vi insistere på at der er måde hvorpå modellerne skal implementeres. Vil vi være i stand til at tage vare på det?

	> Uanset kompetencer – så er det omstillingen der er den største udfordring. At der er procedurer der gør at man skal bruge den. > Hvilke systemer skal vi bruge? Det skal være tydeligt. Og drifts-model. Vi konterer det som et projekt. Det er den eneste måde vi har data på i et system. > Drift er ikke helt klar. Vi har et team for digitalt samarbejde. Et nyt mandat. Bygningsdata skal håndteres på en "Glasblok-metode". Vi har store ambitioner i projektet (Schneider samarbejde (?).) > Drift og vedligehold er super centralt. > Hvis man skrifter nøgler eller andet – så skal den data jo ind i databasen. Det er udfordringen. Det skal der være gode systemer og måde på. For at den kommer op og køre, så er der mange ting der skal fungere. Hvor er I om 10 år? > Driften opdateres. Den måde bygningen skal driftes på. StreamBIM. Al data overføres og metoden bruges. Læse data. Vi skal ikke begrænse os. Vi trænger alle data.
Tema 4: Konkret på digitale tvillinger (ca. 20 min) På hvilke type ejendomme vil det give mest værdi for jer at få en digital tvilling? (Overblik, drift, energi?) Hvilke krav er der til brugerfladen? Hvad skal kunne overvåges?	> Alle tekniske anlæg. Ikke kun energi. > Især de store central bygninger. > Der er mange muligheder. Drift til klinisk virksomhed. Planlægge og logistik løsninger. Det kunne være helt oplagt. Strategi – vi jobber strategisk med digital planlægning af bygningerne. > Helhedsperspektiv: Vi har gennemført et projekt på infrastruktur. Her har vi samlet det hele – intern håndtering af overblikket. Dem der skal bruge tvillingen, skal forstå den. > Vigtigt at der systematiseres på tværs af bygherrerådgivere (COWI, Multikoncern, mv). > Tegninger skal være rigtige før vi skal bygge. > Entrepriseformen kan være afgørende for hvor meget der arbejdes med digitale tvillinger (krav ved f.eks. totalentreprise).
Tema 5: Afslutning (10 min)	Tak for i dag. Opfølgning?

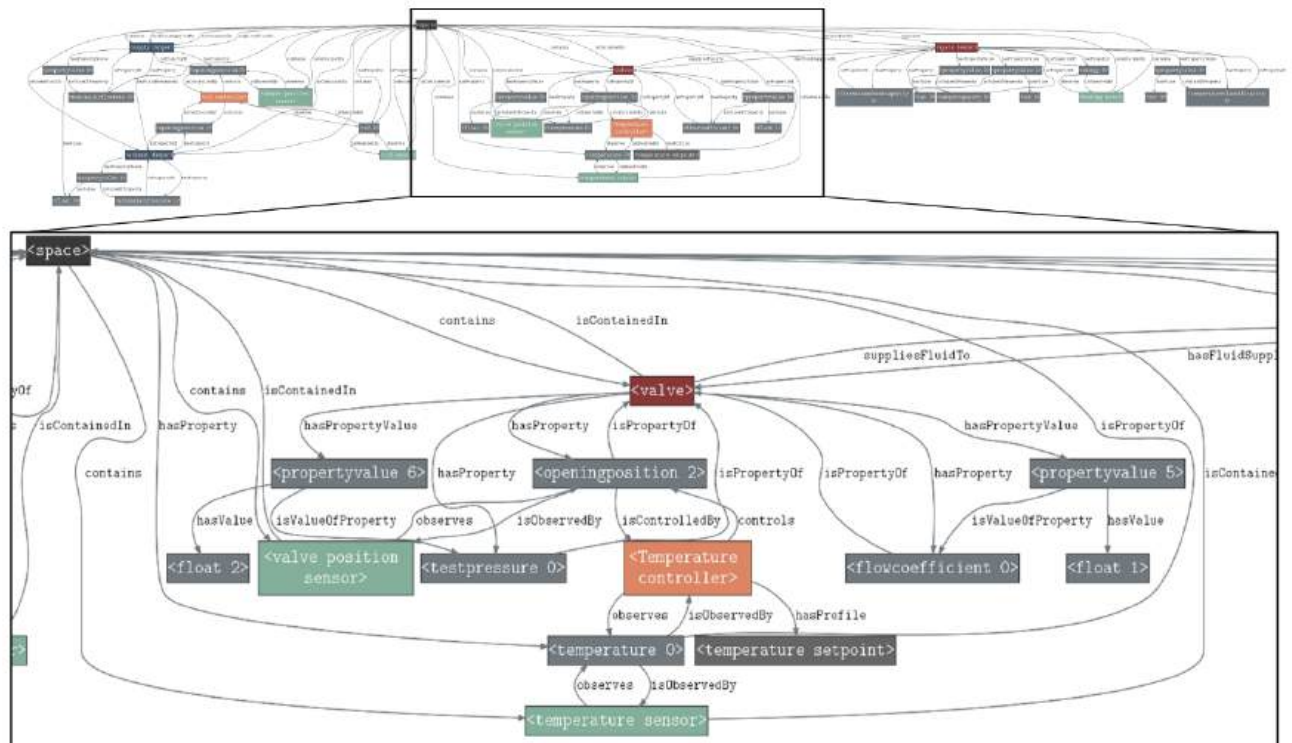
	Rufus – populært system i Norge til at holde styr på udstyr.
--	--

Architecture of the Twin4Build Digital Twin platform and the information flow between high-level components.

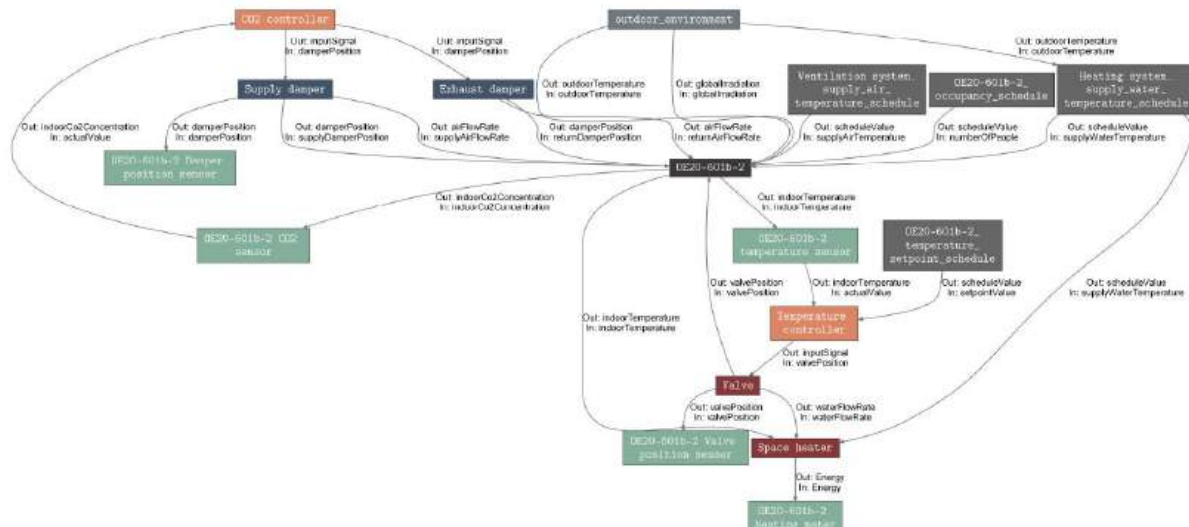


Twin4Build Solution Demonstration in OU44 Room

- A. Semantic model of the classroom case study representing the system topology and equipment design Information of the physical system. The model is based on concepts from the SAREF ontology extended with SAREF4BLDG, SAREF4SYST, and FSO:



B. Graph of the generated simulation model for the classroom case study showing the information flow and logical connections between component models. The graph has been automatically generated using the opensource Python package available on GitHub Bjørnskov (2023):



- C. Dashboard in the Grafana interface showcasing real-time performance monitoring of the classroom. (a) Two days of normal operation; (b) An anomaly detected, most likely caused by an opened window:



(a)



(b)

Digitale tvillinger i byggeriet

AF BELINDA MARIANNE GODSK

Sektionsleder VVS og Ventilation, Bygninger, Øst, COWI



Digitaliseringsbølgen er over os, og vi forsøger konstant at optimere processer ved digitalisering. Dette gælder også for byggeriet.

Alle større nybyggerier modelleres i 3D og meget af projekteringen foregår digitalt i modellerne. Men det fulde digitale potentiale er langt fra opnået, idet vi stadig modellerer den samme bygning i forskellige programmer. Nogle for at kunne koordinere terræn, konstruktioner, installationer og skabe tegninger, og

andre for at kunne energioptimere, certificere eller simulere f.eks. indeklima, bestemme et energimærke eller opnå en certificering. Driften af bygningen er som oftest ikke tænkt med i de digitale modeller af bygningen, hvilket er en skam.

I den ideelle verden modellerede man kun bygningen en gang og stod ved

ibrugtagningen af bygningen med en digital tvilling af bygningen, som kan bruges i driftsfasen til optimering af bygningens drift. Vi er



NYE STRØMNINGER 2023

stadig langt fra dette ønske scenarie... men rejsen derhen er begyndt.

Med digitaliseringen af byggeriet er der opstået nye muligheder for at effektivisere både drifts- og vedligeholdelsesarbejde i forbindelse med bygninger. Sensorer kan registrere og levere data i realtid om nærmest alle aspekter af en bygning. Digitale bygningsmodeller koblet med data i realtid er grundlaget for arbejdet med digitale tvillinger.

Den digitale tvilling er en levende digital model af en bygning, som følger bygningen gennem hele dens levetid. Ved hjælp af sensorer, der sender realtidsdata til modellen, viser den digitale tvilling et aktuelt og opdateret billede af en bygnings tilstand. En digital tvilling kan give viden om alt fra, hvornår en pære skal udskiftes, energiforbrug, indeklima, og meget mere.

COWI deltager i et samarbejdsprojekt mellem fire hovedpartnere. SDU Center for Energiinformatik ved Syddansk Universitet, COWI A/S, KMD A/S og Aarhus kommune. Projektet er støttet af EUDP (Det Energiteknologiske Udviklings- og Demonstrationsprogram), som ligger under Energistyrelsen. Formålet med projektet er at udvikle en digital tvilling-platform, der muliggør koordineret beslutningsstøtte for driftsherren i hele bygningens livscyklus. Platformen erstatter den traditionelle statiske BIM-model med en operationel model, der kombinerer fleksible og adaptive dynamiske energimodeller og realtids bygningsdata.

Intentionen er at de udviklede værktøjer og digitale tvillinger af bygninger vil give:

- Ejere mere energieffektive bygninger
- Facilitetsledere et værktøj til smartere styring gennem avanceret overvågning, automatisk kontinuerlig idriftsættelse og en rettidig vedligeholdelsesindsats
- Konsulenter og planlæggere en løsning, der hjælper med effektiv strategiplanlægning og optimal støtte til beslutningstagning gennem kørsler af "What if-scenarier"

Anvendelse af de udviklede værktøjer vil blive demonstreret gennem case-studier gennemført på tre bygninger af forskellig karakter:

- SDU Live-Lab campusbygning OU44
- Dokk 1 på Aarhus havn
- En hospitalsbygning i Bergen, Norge

Projektet startede i oktober 2021 og løber over 3 år frem til efteråret 2024. I den indledende fase af projektet har vi afdækket hvilke krav, der er til den digitale tvilling og hvilke krav der er til model og data.

COWI har i den forbindelse afholdt interviews med ejere og administratorer af større bygningsporteføljer for at få afdækket, hvad deres forventninger og ønsker er til digitale tvillinger i byggeriet. Der blev holdt interviews med i alt seks kunder:

- Aarhus Kommune
- Newsec
- FSB
- KEID (Københavns Kommune)
- Bygningsstyrelsen
- Helse Bergen (NO)

Der blev udarbejdet en interviewguide som blev fulgt ved samtlige interviews, og der blev spurgt ind til disse forskellige emner, som alle interviewede forholdt sig til:

- Hvor digitaliserede er I? Gør I allerede brug af digitale tvillinger?
- Hvad er jeres forventninger og ønsker til brugen af digitale tvillinger?
- Hvilke udfordringer ser I ved implementering af digitale tvillinger?
- Hvor giver de digitale tvillinger mest mening for jer?

Kundernes forventninger og ønsker

Generelt har kunderne et ønske om at de digitale tvillinger ikke kun skal etableres for nybyggeri, men også for eksisterende bygninger, således at hele bygningsporteføljen kan overvåges og optimeres. Optimering er her både med hensyn til energi, indeklima og om bygningen udnyttes optimalt i forhold til brugen af bygningen. Forventningen er, at de digitale tvillinger skal kunne tilgås via en platform, der er let og intuitiv for brugerne at

anvende, og som hurtigt giver et overblik over om bygningen er i optimal drift og hvor evt. udfordringer befinder sig og hvor kritiske de er. Herudover skal driftsledere hurtigt og nemt kunne lære at bruge platformen uden at have et stort kendskab til IT, CTS og automatik.

Der er ligeledes et ønske om at de digitale tvillinger skal gøre det nemt at sammenligne forskellige bygningers ydeevne med hensyn til både energi og indeklima.

Bygherrer og administratorer ønsker at have et overblik over om alle kvadratmetre i deres byggeri bliver udnyttet fuldt ud. Benyttes bygningerne som forventet eller har man ubenyttede arealer, der ikke anvendes og som kan anvendes af andre eller på en anden måde. Dette er særligt relevant for kommunerne, der stiller faciliteter til rådighed for borgerne, dette for at kunne kontrollere om faciliteterne bliver benyttet som aftalt og forventet.

Det er nærmest en selvfølge, at man har en forventning om at den digitale tvilling skal kunne spare energi ved at optimere driften. De skal kunne regulere bygningsinstallationerne til et ønsket/forventet indeklima ved mindst muligt energiforbrug. Forventningerne peger også i retning af, at man ønsker at den digitale tvilling skal kunne forudsige, hvornår det bedst kan betale sig at aftage fjernvarme og eller fjernkøling fra forsyningsnettet. Dette på baggrund af f.eks. et brugsmønster, en vejrudsigt der indikerer et behov eller når efterspørgslen/omkostningen er lav, for derved at lagre varme eller køling i bygningens termiske masse til senere afgivelse.

Der er et udpræget ønske om at de digitale tvillinger skal kunne opdage afvigelser i vand og energiforbruget, der går ud over normen og advare brugerne, måske endda tage affære ved at tilkalde en servicemedarbejder eller ekstern håndværker via SMS eller mail. Der er et ønske om at de digitale tvillinger skal kunne forudsige forskellige "What if-scenarier". F.eks. hvordan in-

NYE STRØMNINGER 2023

deklimaet og energiforbruget vil være, hvis indretningen i en bygning ændres. De ovenstående ønsker og forventninger bliver p.t. indarbejdet i projektet.

Utopia – Når rådgiveren bliver arbejdsløs

Der er også ønsker som p.t. ville kræve langt mere end hvad der er muligt i projektet. Der har f.eks. været udtalt ønsker om, at de digitale tvillinger skal kunne forklare ændringer i en ejendoms energiforbrug. Nogle ændringer ville være lette at forklare, mens andre ville kræve en dybere granskning af data, en indgående viden om bygningsdrift og regulering, brugen af bygningen og en viden om, at flere faktorer kan være i spil. Her vil de digitale tvillinger komme til kort.

Et andet utopisk ønske har været, at de digitale tvillinger skal kunne udvælge driftstrategier for bygninger, give forslag til optimeringsforbedringer og afrapportere, hvad potentialet ved disse forbedringer er. Når de digitale tvillinger kan dette, bliver de rådgivende ingeniører, inden for denne branche, arbejdsløse.

Ønsket om at gøre drift af store porteføljer lettere og mere overskueligt

Den største overraskelse for os, der deltog i projektet, var bygherres og administratorens ønske om at de digitale tvillinger ikke kun skal etableres for prestigebyggeri, men skal etableres i såvel nye som eksisterende bygninger og for hele bygningsporteføljen. Det var en stor øjenåbner for os og ønsket åbner op for et meget større marked for de digitale tvillinger. Projektet tog oprindeligt afsæt i at de digitale tvillinger skulle skabes ud fra BIM-modeller og konverteres til operationelle modeller, der kombinerer fleksible og adaptive dynamiske energimodeller og realtids bygningsdata. Den for os nye viden har ændret vores opfattelse af hvor og hvilke input og data, de digitale tvillinger skal opbygges af. Hvor BIM-modellerne oprindeligt var alfa og omega er de nu et supplement blandt flere kilder. En af grundene er at den eksisterende bygningsmasse typisk ikke har bagvedliggende 3D-modeller, hvilket dog er forholdsvis let at skabe ved bygningsscanninger. En anden grund er at modellerne ikke i høj nok grad indeholder alle de data der ønskes, for at kunne etablere de digitale tvillinger. Mere om det i det følgende.

Udfordringerne

Da projektet startede op, var der et ønske om, at man kunne trække mange af de informationer, der er behov for, for at kunne skabe de digitale tvillinger, fra BIM modellerne / IFC-formatet.

Det har dog vist sig at dette ikke er muligt. Data i modellerne er ikke struktureret og forbedrede for digitale modeller og kvaliteten varierer meget. Det er derfor nødvendigt at indhente data via andre kilder. For at man i fremtiden skal kunne drage større nytte af BIM-modellerne, skal der udarbejdes en kravspecifikation for hvad BIM-modellerne skal indeholde og hvorledes data

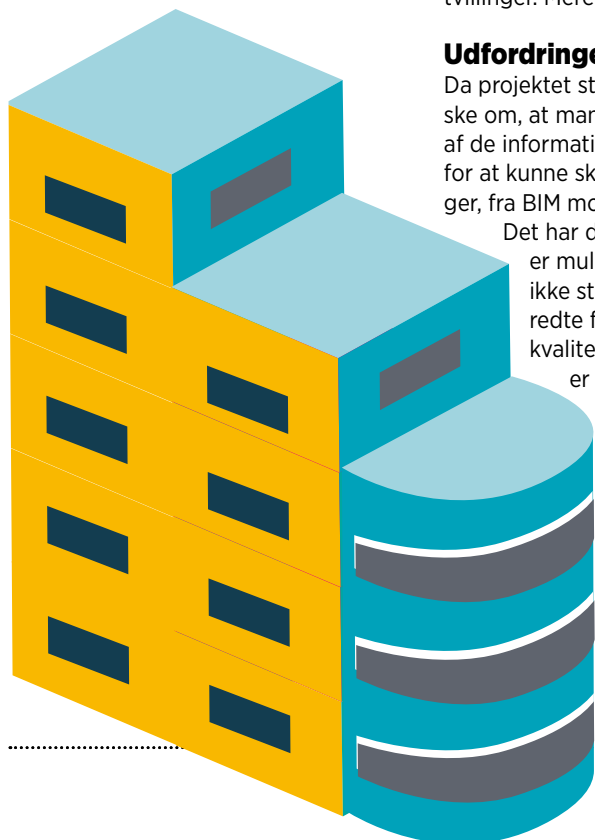
indarbejdes hvis de senere skal implementeres i digitale tvillinger. Kravsspecifikationen vil afhænge af hvad det er man vil have at den digitale tvilling skal kunne. Dette arbejde er et helt projekt i sig selv.

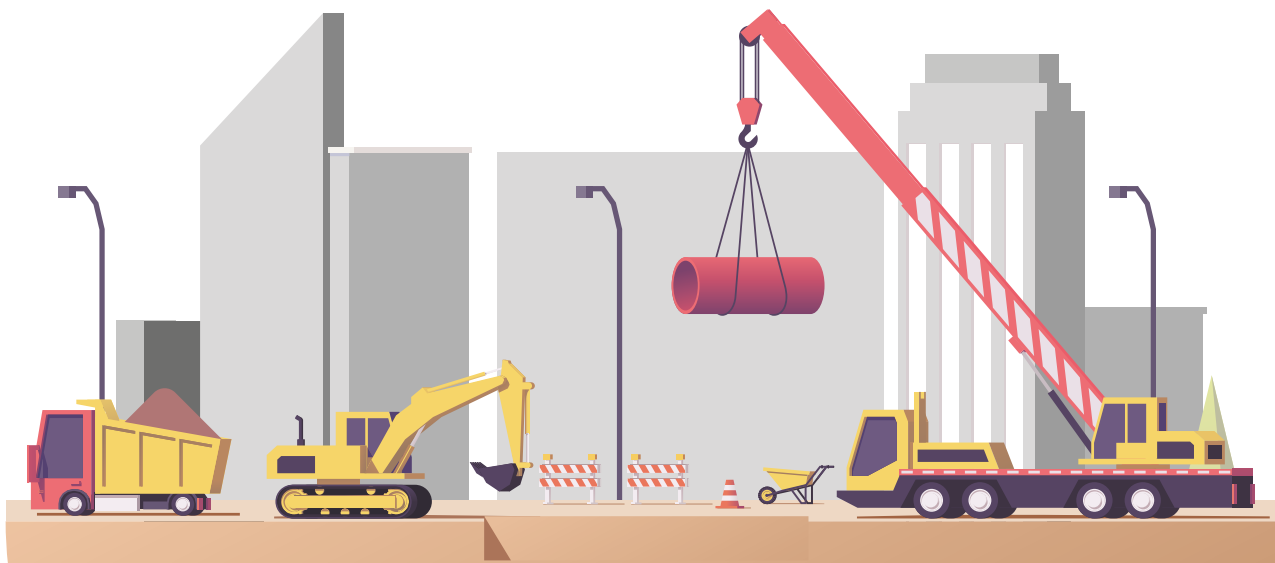
En af de større udfordringer ved de digitale tvillinger er, at de digitale tvillinger skal vedligeholdes for at give værdi gennem bygningernes totale levetid. De bagvedliggende 3D-modeller skal svare til de aktuelle indretninger og udskiftes en sensor eller udskiftes et tekniskanlæg, skal tvillingen fodres med de nye oplysninger. Er der tale om store bygningsporteføljer som f.eks. dem mange større kommuner står med, er vedligeholdelsen af de digitale tvillinger en stor og krævende opgave og spørgsmålet er om den vil stå i mål med de besparelser de digitale tvillinger vil kunne give.

De digitale tvillingers succes afhænger af om de kan skabes på baggrund af let tilgængelige data i et system der er let at opdatere. De digitale tvillinger må ikke blive så komplicerede, at ingen forstår at bruge og vedligeholde dem. Det bliver et spørgsmål om balancen mellem detalje og overordnet overblik, præcision og tilstrækkelighed.

Hvordan kommer vi videre

Nu er kundernes forventninger og ønsker klarlagt. Arbejdet med at brugertilpasse den digital tvilling-platform fortsætter i dialog med mulige kunder og andre interessenter. Der samarbejdes om udarbejdelsen af platformen for de digitale tvillinger, model og datainsamling og der er taget hul på det første casestudie, som er SDU Live-Lab campusbygning OU44. Vi glæder os og ser frem til at se den første demonstration af platformen.





Fortsat mange fejl og mangler i byggerier i dag

AF TOMMY STEEN MØLLER

Projektleder, Byggeri, Region Sjælland

Til trods for alle digitale og fysiske hjælpemidler, laver vi lige så mange fejl som for 20 år siden. Til gengæld er vi blevet bedre til at dokumentere dem

Vi oplever gang på gang, at byggeriers kvalitet er langt under, hvad vi kunne ønske os, og at byggerier ved afleveringen og for den sags skyld også i perioden efter er fyldt med alvorlige funktionelle fejl såvel som æstetiske fejl.

Jeg skal ikke her forholde mig til de allerede kendte byggeskandaler, der er dukket op gennem de senere år, men

mere generelt hvad jeg oplever i byggebranchen i dag. Gennem de mange år i byggebranchen og hos byggebranchens forskellige parter oplever jeg i dag, til trods for alle de gode intentioner og skåltaler, at der fortsat mangler samarbejde og gensidig respekt parterne imellem, samt en mangel på håndværksmæssig stolthed. Der er selvfølgelig byggerier, hvor det faktisk foregår efter planen og er mere eller mindre fejlfrit, men det er desværre, som jeg oplever det, ikke det generelle billede.

Rådgiverne bliver beskyldt for, og i flere tilfælde med rette, at lave et

ufærdigt projekt, måske ovenikøbet et ikke bygbart projekt. Det skyldes ofte afgivelsen af en for lav pris, hvor der i sidste ende ikke er tid nok til at færdiggøre projektet ordentligt. Det skyldes også, at dem, der sidder og projekterer eller tegner i modellen ikke har nok erfaring eller viden om, hvad de har med at gøre. Den manglende erfaring og viden skyldes ofte, at per-



NYE STRØMNINGER 2023

sonerne ikke har den praktiske erfaring fra byggepladsen, og at der ikke bliver kompenseret for det ved at inddrage erfarne arkitekter eller ingeniører.

Jeg har oplevet, at rådgiverne stiller spørgsmål til bygherren, hvor de glemmer, at det er dem, der er rådgivere og specialister og burde komme med forslag frem for at spørge.

Desværre er det også sådan, at det egentlige fagtilsyn, som helst skal udføres af den person, der har lavet projektet, ofte er sparet væk, hvorfor den indgående viden til projektet mangler, og det kun bliver en mere overfladisk gennemgang på pladsen.

Entreprenører og håndværkerfirmaer har i manges øjne fået slyngelolie til morgenmad. Man kan ikke stole på dem og de overholder aldrig aftaler. Det er jo noget af et prædikat, men det kommer jo ikke af ingenting. Uanset om der tale om store totalentreprenører eller den mindre håndværksmester kniber det med samarbejdet på byggepladsen, altså det rigtige samarbejde, hvor målet er at færdiggøre den stillede opgave på bedst mulig måde. Hertil kommer den manglende respekt for hinandens arbejder, hvilket kan være en total showstopper for et godt resultat.

Der sidder i dag mange personer, mange flere end tidligere, i byggepladskon-torerne, men desværre kommer mange af dem ikke længere end til bag ved computerskærmen, måske ovenikøbet fra hjemmearbejdsplads, og der løser man ikke de opståede komplikationer i byggefasen. Det skal dog nævnes, at der også bliver brugt mange ressourcer på dokumentation og rapportering, nok også mere end egentligt nødvendigt.

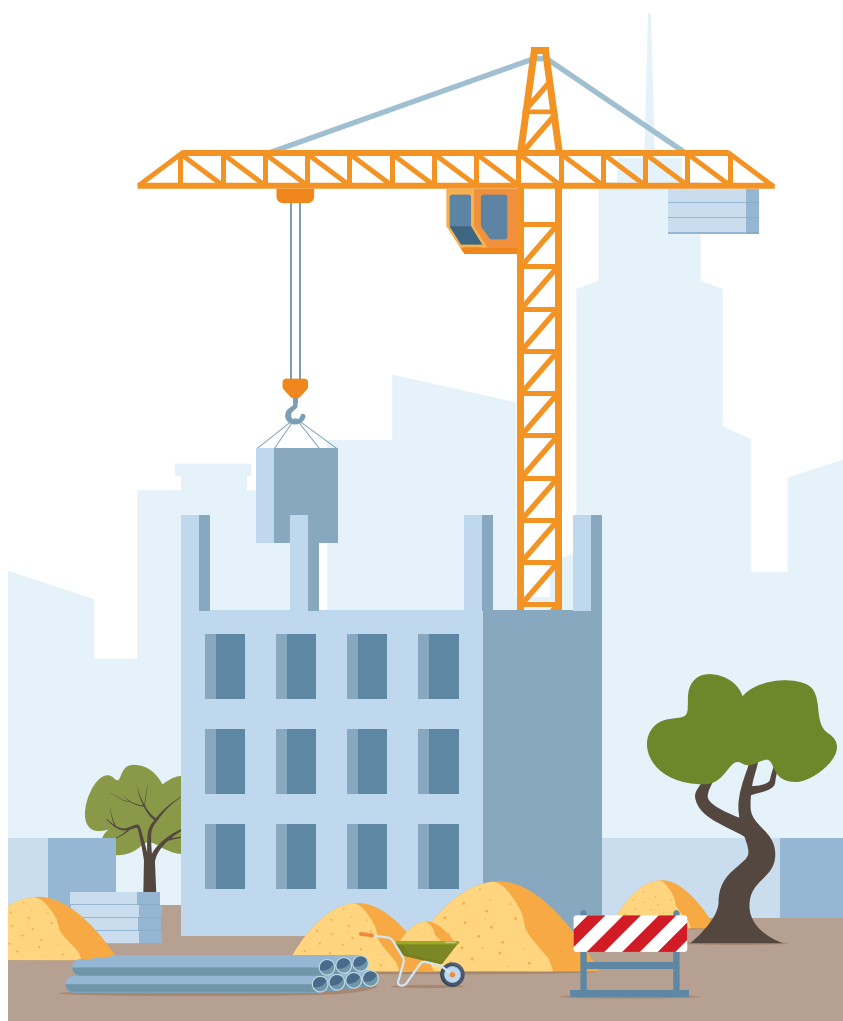
Håndværkerne på byggepladsen løber rundt mellem hinanden, samtidig med at liftene suser rundt. Sådan har det altid været, måske undtaget liftene, men det gør det ikke bedre i forhold til at udføre et ordentligt stykke arbejde. Især ikke, hvis de har en opfattelse af, at der alligevel nok kommer en æn-

dring i morgen. Jeg synes, at der er et meget stort men. Hvor er den håndværksmæssige stolthed? Der bliver i dag sjusket ved arbejdets udførelse, og det ødelægger helhedsindtrykket af det færdige arbejde, selvom det måske funktionelt kan fungere. Samtidig er der også blandt håndværkere en mangel på respekt for hinandens arbejde, især når der køres rundt med lift. Og mærkeligt nok er opståede skader på vægge, gulv eller installationer altid kommet af sig selv.

Der arbejder i dag mange forskellige nationaliteter på en byggeplads, og det gør selvfølgelig kommunikationen sværere og dermed også samarbejdet, men det er ikke det, der er hovedårsagen til den oplevede sjusk. Det sker uanset nationalitet.

Blandt byggeriets parter opleves der mindre reel samarbejdsvilje, til gengæld mere uforsonlige holdninger parterne imellem. Dette underbygges som regel af en manglende byggepladsstyring, således at parterne mere eller mindre selv kan bestemme ud fra deres forudsætninger. Så skal det gå galt.

Jeg vil gerne væk fra de store hensigtserklæringer uden reelt indhold og de flotte skåltaler. Jeg vil også gerne skelne mellem, hvad der tilføjer værdi i byggefasen, og hvad der blot skal bruges som dokumentation i tilfælde af, at vi ender i Voldgiftsretten. Jeg vil gerne have fokus på at bygge det rigtige, og helst første gang, så alle parter har noget at være stolte af.





Component-level re-commissioning of a newly retrofitted Danish healthcare building

Jakob Bjørnskov^{*}, Muhyiddine Jradi, Christian Veje

Center for Energy Informatics, Mærsk Mc-Kinney Møller Institutet, University of Southern Denmark, Campusvej 55, Odense M, 5230, Denmark

ARTICLE INFO

Keywords:

Commissioning
EnergyPlus
Heating coil
Heat recovery unit
Ventilation
Building performance simulation

ABSTRACT

The building sector is among the sectors with the highest energy use, responsible for 40% of the total energy consumption in the EU. The high energy consumption is partly due to inefficient operation and the delayed detection of faulty systems and components in the buildings, which entails both increased operational costs and CO₂-emissions. Thus, building energy performance monitoring and evaluation is key in ensuring a proper building operation and timely detection of anomalies and faults on different levels. This can be implemented through systematic building commissioning, a quality-assuring process aiming to ensure that buildings are operating according to the expected performance. While initial commissioning of newly built buildings has been the norm in the last decades, re-commissioning of deeply retrofitted buildings is generally overlooked despite the potential and positive impacts. In this work, a whole-building EnergyPlus model is developed, calibrated, and employed for the re-commissioning of building components in a Danish health care building that has recently exhibited a deep energy retrofit process on various levels including the installation of new ventilation and cooling units. Both the calibration and the commissioning of these components is performed using data collected from non-invasive clamp-on temperature and power sensors. The commissioning process is performed using a limit checking technique with a predefined threshold, allowing the detection of multiple anomalies during the commissioning period.

1. Introduction

In the EU, the building sector is responsible for 40% of the total energy consumption with electricity accounting for 48% of this share, causing buildings to be one of the major contributors to the emission of greenhouse gases, surpassing sectors such as industry and transport [1]. The high energy consumption of buildings is often caused by inefficient operation or undiscovered faults, resulting in an energy waste that could have been avoided. On average, it has been estimated that 20% of the total building energy consumption is wasted [2]. Another study showed in 2011, that the most frequently occurring faults in ventilation systems amounted to a loss of approximately 94.6 BkWh electricity equivalents or 9.6 B\$ per year in the United States [3].

Therefore, there is not only a significant need but also a significant potential for increasing the energy efficiency of buildings, which will mitigate both the emission of greenhouse gases and the cost of building operation. In the political landscape, this potential has been recognized and targeted by the EU through the 2010/31/EU *energy performance of buildings directive* [4]. In Denmark, this directive has entailed regulation of building energy performance requirements through the *Building regulations* legislation [5]. Building

^{*} Corresponding author.

E-mail address: jabj@mumi.sdu.dk (J. Bjørnskov).

owners in Denmark are thus required to test and document that their buildings comply with the various demands in this legislation before the building can be handed over and put into use.

Generally, the concept of ensuring that buildings are operating as expected is referred to as commissioning [3]. Commissioning is thus a quality-assuring process that aims to identify and correct unintended or faulty operation in buildings. Legris et al. outlined in an annex 47 report [6] four different forms of commissioning that differ based on where in the building life cycle they are implemented. Commissioning performed during the design and construction stage of a building is often referred to as initial commissioning. This process ends when the building is handed over to the owner and the occupancy stage begins. Re-commissioning is performed on buildings that have already been commissioned previously [7] and retro-commissioning is performed on buildings that have never been commissioned and are in their occupancy stage [8].

While these three types of commissioning have been demonstrated through numerous case studies to yield valuable savings both in terms of wasted energy and operational costs [9], they are also limited by their one-time execution. A Building, with its many complex systems and components, deteriorates over time and therefore, the building performance must be monitored and evaluated continuously across the whole lifetime of a building. This drawback has caused the development of continuous commissioning, which monitors the building performance continuously throughout the whole building life cycle, allowing the building operators to detect and act immediately on component faults or inefficient building operation.

1.1. State of the art

When performing commissioning on a building, an important aspect to address is how the expected performance is determined. Here, a straightforward definition is often used; the expected performance of a building must reflect an ideal scenario, where the building is operating optimally, i.e., without faulty components and inefficient control [1]. However, even under these ideal conditions, the building performance is often highly dependent on time-varying conditions such as weather conditions, occupancy, indoor comfort, etc. Therefore, evaluations of building performance are often based on forecasts that take into account the before-mentioned conditions. Forecasting building operation accurately is, thus, a complex task, which has given rise to two fundamentally different approaches; data-driven or deterministic modeling [10,11].

The deterministic or law-driven models, also commonly known as white-box models, make use of first-principles modeling to model the thermal properties and interactions in the building. This approach thus requires extensive knowledge of the building in the form of geometry, construction materials, occupancy schedules, Heat, Ventilation, and Air-Conditioning (HVAC) systems, control schemes, etc. It is deterministic in the sense that the model predictions can always be explained by the collective behavior of all the subsystems and components in the model, which are based on fundamental physical laws and principles.

To relieve the amount of work required for developing these types of models, a large number of simulation programs, have been developed over the years with some of the most popular being; EnergyPlus, DOE-2, and TRNSYS [12–14]. Conventionally, simulation tools like EnergyPlus have been used for analyzing building design and retrofit scenarios [15,16]. However, in recent years these tools have also been implemented in case studies for continuous commissioning of buildings.

Jradi et al. [17] implemented a calibrated EnergyPlus model in a continuous commissioning and performance testing framework of a campus building in Odense, Denmark. Using actual weather data collected at site, the framework continuously compares the predicted performance with actual data and reports the performance gap in a dashboard application. Using the developed framework, a malfunctioning ventilation unit was detected in the building.

Harmer et al. [18] developed and calibrated an EnergyPlus model of a campus building in Colorado, US. The model was implemented in the energy management application *SkySpark* as a dynamic baseline for fault detection using predefined thresholds. The developed framework was also able to predict energy consumption in the building up to 48 h ahead based on weather forecasts.

In data-driven modeling, the models are trained on historical data collected from the actual buildings, where the training process usually represents some sort of parameter-fitting algorithm. These models are typically divided into two categories; grey-box and black-box models. Black-box models are purely data-driven, meaning that the underlying model structure and parameters have no direct physical meaning. The term black-box model is typically used in the context of machine-learning models i.e. Neural networks, Support Vector Machines, Decision Trees, etc.

The other alternative within data-driven modeling is Grey-box models. These are a mix between the white-box models and the black-box models where the model structure is determined from first-principles while the model parameters are found through parameter-fitting techniques using actual data. For buildings, one of the most commonly used grey-box models is the Resistor-Capacitor (RC) model, which employs the electrical analogy of RC-circuits to thermal systems. This results in a system of ordinary differential equations for which optimal parameters can be estimated through parameter identification techniques [19,20]. The order of these models is typically much lower than for white-box models, which makes these models attractive for applications in advanced control such as Model Predictive Control (MPC) [21].

One of the strengths of using deterministic modeling and simulation tools for commissioning is that the model is very comprehensive and can be dissected into its smallest sub-components to a large level of detail, which can be compared with the equivalent components of the real system. However, the large amount of components and model detail also has certain downsides concerning the calibration process of the model. During the calibration process, the modeler tunes the various parameters of the energy model to minimize the discrepancy between the predictions of the model and the actual measurements from the building – the so-called *performance gap*. However, typically, many of the input parameters are unknown, which creates a vast amount of possible configurations. Reddy refers in Ref. [22] to this problem as the *identifiability problem*, i.e., there are too many candidate configurations that can explain the data used for calibration, essentially because the problem is under-determined. Therefore, the modeler is left with the decision of choosing between these candidate solutions only based on subjective judgment and experience. The manual calibration process of

large white-box building models is therefore often considered as an art rather than science.

Over the years, calibration of white box building energy models has been studied extensively in the literature. Fabrizio et al. [23] introduced the concept of different calibration levels 1–5, depending on the information available for the calibration process. Here, level 1 is the least accurate as it is based only on utility bills and as-built data. Hence, it only considers the dynamic behavior of the building in very low resolution. Level 5 is the most accurate, which considers the before-mentioned aspects as well as site visits, detailed audits, short-term monitoring and long-term monitoring of the building, typically with a time-span of several years.

Aligning with this concept Guyot et al. [24] obtained a level 5 calibrated EnergyPlus model of a 5434 m² building. The authors choose to abandon the common calibration strategy where monthly data is used. Instead they choose to evaluate the model accuracy on sub-hourly level through a dynamical analysis of the predicted cooling and heating consumption as well as predicted indoor temperatures. The study showed that most discrepancies between measured and predicted performance was due to non-modeling of transient phenomena with fast dynamics.

Jradi et al. [25] calibrated an EnergyPlus model of a university building using the methodology of Hale et al. [26]. The model was calibrated on monthly utility bill data with respect to ventilation, lighting and heating consumption by varying model parameters such as infiltration rates, equipment and electrical devices efficiency, ventilation rates and operation schedules. The calibrated model was employed in a developed tool to continuously monitor and evaluate the energy performance of the building.

Chong et al. [27] conducted a comprehensive review on the major aspects of building energy simulation models calibration. As part of the review, a detailed meta-analysis of calibration methods, model inputs, outputs, data types, resolution and key metrics was conducted. It was found that the optimization and Bayesian calibration approaches are the most common automated techniques. The authors highlighted reproducibility of simulations as a critical issue, proposing an incremental approach with clear reporting on the study contexts and processes to encourage future research's reproducibility.

Da Silva et al. [28] compared several calibration methodologies concentrating on the electrical energy consumption in buildings. They proposed a new calibration approach for performance simulation, which is separated into two evidence-based steps based on the size of the HVAC systems used. A university hospital building was used as a case study, of which an EnergyPlus model was developed. The results showed that no calibration methodology is reproducible for all buildings and that a key point to consider in every calibration process is the trade-off between increasing the model predictions accuracy and the strictness of accuracy constraints and indicators.

Royapoor et al. [29] developed a full scale EnergyPlus building model of a 5-storey office building and used a set of two calibrated environmental sensors along with weather station data to calibrate parameters in the building such as HVAC system, infiltration rates, building envelope thermal properties, and internal loads. The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) Guide 14 indices were used as a basis for the model calibration aiming to achieve a Mean Bias Error (MBE) within $\pm 5\%$ and Cumulative Variation of Root Mean Square Error (CV(RMSE)) below 10%. It was found that the developed EnergyPlus model is capable of predicting the annual hourly space air temperatures with an acceptable accuracy of $\pm 1.5^\circ\text{C}$ for 99.5% of the considered time.

In a recent study, Baba et al. [30] developed and presented a robust methodology for EnergyPlus building models calibration allowing high accuracy in predicting indoor temperatures and reducing the discrepancy in assessing overheating risks. The approach employed a variance-based sensitivity analysis to determine building parameters which have the highest influence on indoor air temperatures. Considering the identified parameters, a Multi-Objective Genetic Algorithm approach was used to calibrate different rooms in the building simultaneously. A new metric was introduced, the Maximum Absolute Difference (MAD), a new metric, to serve as a basis for the calibration process. MAD is defined as the maximum absolute difference between simulated and measured hourly indoor temperatures. School building case study was considered, and the calibration method has led to a model being able to predict the indoor temperatures in the rooms with RMSE of 0.3°C , and MAD of 0.8°C .

Lam et al. [31] presented a method for EnergyPlus office buildings models employing an occupant behavior data mining approach with the aid of empirical data. In their calibration approach, the authors generated workday occupancy schedules by mining appliance energy consumption data. Such data includes hourly and monthly power and energy consumption for lighting, equipment and HVAC systems. The result resulted calibrated EnergyPlus model was found to predict the total annual energy use intensity with an acceptable 1.27% mean bias error. Asadi et al. [32] presented an EnergyPlus building energy model calibration using automated optimization-based algorithm. To reduce the computational time in calibrating the model, a Harmony Search (HS) algorithm which is a heuristic search algorithm was developed and employed in the work in order to identify the magnitude of the model parameters to minimize the differences between the actual measured data and the energy simulation predictions. A case study of an office building in Qatar is considered and a whole building detailed EnergyPlus model was developed. The proposed optimization-based calibration method was implemented. It was found that using the HS algorithm, 500 iterations are needed for the model to be calibrated properly with a mean bias error of less than 5%.

In recent years, initial commissioning of newly built buildings at the end of the construction stage and before building handover has been the trend and the norm. In Denmark, the building legislation and the recent regulations have set strict guidelines for such initial commissioning processes to ensure a proper startup for the building. However, for re-commissioning of existing buildings, very little has been reported and implemented. The fact that existing buildings constitute the ultimate share of the building stock, along with the recent danish initiatives supporting large-scale energy retrofitting of the building stock, proves the large potential of implementing systematic retro-commissioning processes to eliminate any faults and anomalies related to modifications and upgrades implemented in the building.

In this work, a case study of a healthcare building undergoing deep energy retrofit is considered, where a subset of building components is selected for re-commissioning. The re-commissioning of building components is demonstrated through the use of a

calibrated whole-building EnergyPlus model along with non-invasive clamp-on temperature and power sensors. Using the data collected from these sensors, a transparent and reproducible calibration process is performed on a subset of components in the EnergyPlus whole-building model. The model along with the actual data is used to implement and demonstrate a recommissioning process to ensure a proper restart of the building. The model development and calibration performed in this work lay the groundwork for implementing continuous commissioning in the case building as a part of the BuildCOM research project [33].

2. Methodology

The methodology used in this work is categorized as a top-down approach. The top-down approach (opposed to bottom-up), is a method of modeling where the overall structure of the modeled system is first formulated, i.e., the system is decomposed into sub-systems and components, but without specifying any details for these components. Afterward, the model can then continuously be refined by modeling the specified components in greater detail. The four main steps of this top-down methodology, specific for this work, are shown in Fig. 1.

For the first step, the case study building is characterized. This includes the key characteristics of the building such as total area, heating and cooling equipment installed in the building, and area covered by the different ventilation systems. This is followed by a description of the method and types of sensors used for data collection. Here, the measured components are also introduced where both their physical location and their overall roles in the system are described. Next, the whole-building modeling workflow is defined. Here, the employed modeling toolchain is introduced, where the relation between the different tools is described. Important design choices, relevant for the whole-building modeling phase, are also mentioned here.

Following, in-depth analysis and calibration of the components chosen for commissioning is performed. Here, the approach is to analyze the mathematical models underlying the EnergyPlus whole-building simulation program and fit the input parameters such that the model predictions fit the measured data.

Finally, the calibrated whole-building model is used for the commissioning of the chosen components. This is done through simulation of the building for a chosen time period with the whole-building model and using actual weather data. For this simulation, the predictions of the model are representing the expected operation of the components, which is compared to the data collected from the actual components. Based on this comparison the performance gap is evaluated and any discrepancy and deviation in terms of the component operation is captured and assessed.

2.1. Case definition

In this section, the key building characteristics are first described for the case study building. Following, it is explained how data have been collected from the building, i.e., the types of sensors used and the components measured.

2.1.1. Building characteristics

The modeled building is a hospital building located in Odense, Denmark, with a total floor area of around 3000 m². It has a basement, a ground floor, and a first floor and is placed in the middle of a relatively large site with no surrounding buildings or structures that could influence the shading properties of the building. An illustration of the building, drawn in the program SketchUp, is shown in Fig. 2. It consists of various space types, e.g. offices, corridors, meeting rooms, laboratories, and medical examination and consultation rooms. The building is heated by direct district heating with radiators on all three floors and water-to-air heating coils coupled with the supply air. For cooling, the building has installed two air-cooled chillers, which supply chilled water for chilled beams and water-to-air cooling coils coupled with the supply air.

The building has 4 ventilation systems installed that supply fresh outdoor air for the occupants in the building and act as the main source of heating and cooling. In Fig. 3, the area covered by the four systems is shown in a floor plan, which has been derived from Piping and Instrumentation (PI) diagrams provided by the building operators.

The figure shows that most of the building is ventilated by the HF02 unit, which covers most of the basement and ground floor along with half of the first floor. Ventilation systems HF01 and HF03 cover the rest of the first floor and are ventilating laboratories with high airflow requirements. Finally, the ventilation system HF04 is ventilating the canteen located in the basement. All four systems are Variable Air Volume (VAV) systems. Additionally, three exhaust ventilation units and one dehumidification unit are installed in the building. However, collectively these units only provide around 5% of the total volumetric airflow in the building, so these have not been included in the figure, although they have been included in the model.

2.1.2. Data collection

Since the analysis was carried out straight after the completion of the building retrofit project, meter data from the building permanent meters were still not reconfigured to report data. Thus, an alternative approach was employed where all data for model



Fig. 1. Methodology used for commissioning of building components.

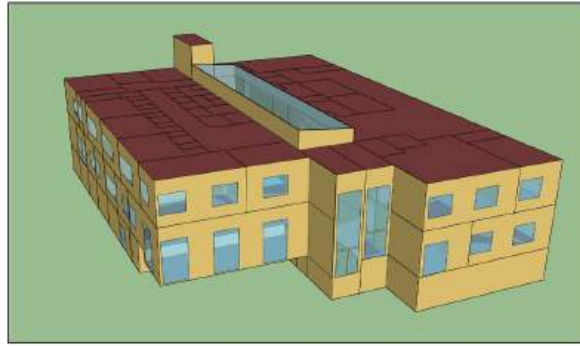


Fig. 2. Drawing of the case study building developed in the program SketchUp.

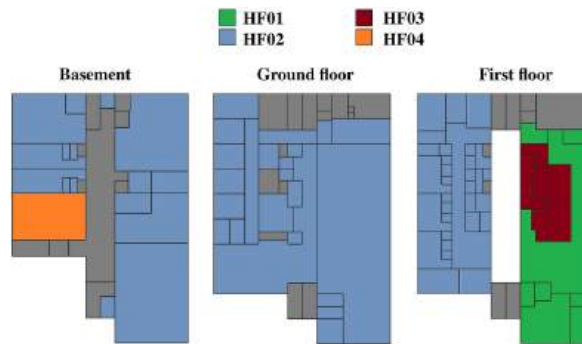


Fig. 3. Floor plan showing the area covered by the 4 outdoor air ventilation systems.

calibration and commissioning of the building was collected using two types of Remoni sensors; the HeatMoniSpot [34] and the PowerMoniSpot sensor [35]. The HeatMoniSpot sensor is a temperature sensor with 2 wired probes designed to be mounted on pipes as shown on Fig. 4a. Additionally, the sensor also has an internal probe which is located inside the sensor body, designed to measure the air temperature. The three probes have a measurement precision of $\pm 0.5^\circ\text{C}$ for temperatures between 0 and 40°C . For temperatures in the range -40 to 0°C or 40 – 80°C , it has a precision of $\pm 1^\circ\text{C}$ [36].

The PowerMoniSpot sensor can be mounted on the outside of a wire and can measure the current flow on any number of phases as long as the load is balanced. These measurements are unitless and, therefore, have to be scaled using a timestamp with a known power flow. This sensor has a measurement precision of approximately 1% and can measure down to 0.2 Ampere, which corresponds to approximately 140 Watts for a 400 line-to-line Voltage [37].

Both sensor types have 2 permanent frequency settings; 1 measurement per 5 min and 1 measurement per 1 min. All data is transmitted from the sensors to a gateway, also located in the building, which has an internet connection and stores the data in a cloud database. The time-series data can then be accessed via a web browser where it can be visualized or exported in a Comma Separated



(a) HeatMoniSpot sensor [34]



(b) PowerMoniSpot sensor [35]

Fig. 4. The two types of sensors used for data collection in the building. (a) The HeatMoniSpot sensor, which is used for temperature measurements and (b) The PowerMoniSpot sensor, which is used for power consumption measurements.

Values (CSV) format for further analysis.

Regarding the placement of sensors and the choice of components for calibration, mainly two criteria were considered. Firstly, because the building has recently been undergoing an extensive renovation, not all units installed in the building were under normal operation at the time of sensor installment. This was for example the case for the ventilation unit HF03 and also the air-cooled chillers coupled with the cold water loop. Secondly, the sensor technology that Remoni offers has an upper limit on the maximum distance between the gateway and the sensors, which Remoni declares to be larger than 750 m in plain sight [38]. However, this distance tends to be considerably reduced by obstacles such as walls, floors, and ceilings. This turned out to be a limiting factor for the data collection on the airloop heating coils, which are located on the roof of the building. It was therefore only possible to obtain a signal for the data collection on one airloop heating coil installed in the HF02 ventilation system. For this system, fan power consumption data have also been collected. A Diagram of the system and the placement of sensors is illustrated in Fig. 5.

Ventilation system HF02 (Fig. 5) supplies 55 rooms in total with 5 parallel heating and cooling water-to-air coils. The ventilation system has one cable supplying power to all fans in parallel. Practically, this means that the total fan power consumption is measured by installing one PowerMoniSpot on the cable supplying these fans. The temperature measurements made on the heating coil are for both the inlet and outlet air and water temperatures (refer Fig. 8). Based on the placement of sensors, the components chosen for commissioning are listed in Table 1.

As evident from Table 1, all fans for the ventilation system will be commissioned using power measurements. For the commissioning of the heating coil, the temperature controller, and the heat recovery unit, the heating coil temperature measurements are used. The heating coil is the last component before the supply air is supplied to the rooms, so at the heating coil air outlet, the temperature must be equal to the supply setpoint temperature. For the heat recovery unit, the measured inlet air temperature of the heating coil is used. This measured temperature is expected to be very close to the outlet supply temperature of the heat exchanger with only a small temperature rise across the fans and no temperature difference across the cooling coil, which was not in operation at the time of data collection.

2.2. Whole-building energy modeling and simulation

To develop a whole-building energy model of the building just presented, the holistic energy modeling and simulation approach implemented by Jradi et al. [39] was used. In this regard, a combination of three tools is employed, Sketchup, OpenStudio, and EnergyPlus. The workflow is illustrated in Fig. 6. As seen from the figure, the model development is comprised of two steps. First, all the geometry of the building is defined in SketchUp. SketchUp is a 3D drawing tool, which integrates with the building energy modeling program Openstudio through a plugin that can be installed. This makes it possible to draw the building geometry in SketchUp and use the defined geometry for further building modeling in OpenStudio. For the case building considered in this work, all necessary geometry data have been available through detailed floor plans and cross-sections of the building.

After the geometry of the building has been defined, OpenStudio is used to model the HVAC systems and the physical properties of the envelope along with other properties such as occupancy, lighting, and schedules. OpenStudio is an open-source whole-building modeling tool that builds on top of the EnergyPlus simulation engine. Contrary to EnergyPlus, which by default depends on text-based editing of the Input Data File (IDF), OpenStudio has a graphical interface that allows for more convenient and intuitive modeling of buildings while it also offers access to the Building Component Library (BCL), a large library of templates for HVAC components, materials, schedules, etc. Moreover, OpenStudio offers Application Programming Interface (API) capabilities through the use of the

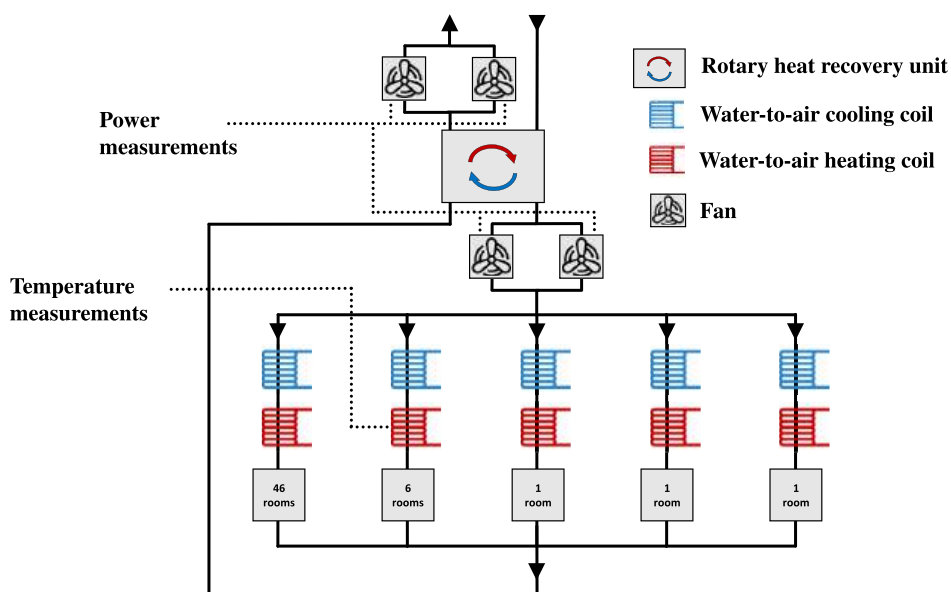


Fig. 5. Diagrams showing the placement of sensors in ventilation system HF02.

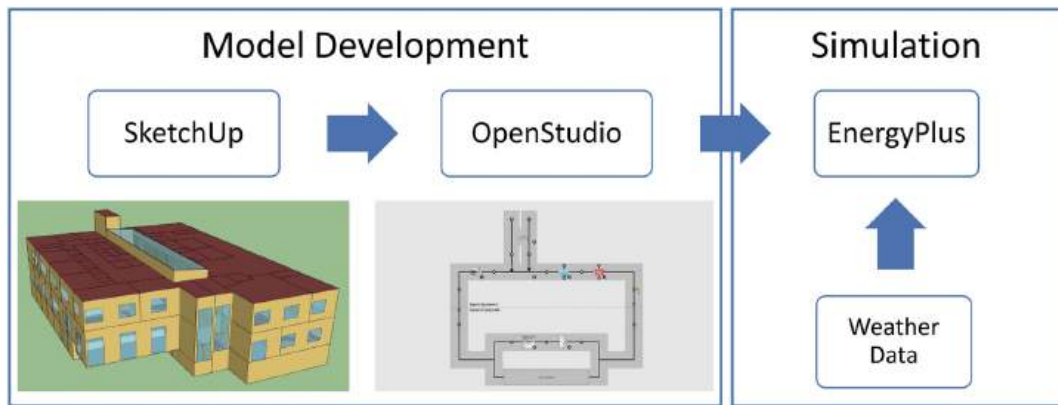


Fig. 6. Workflow of the whole-building model development.

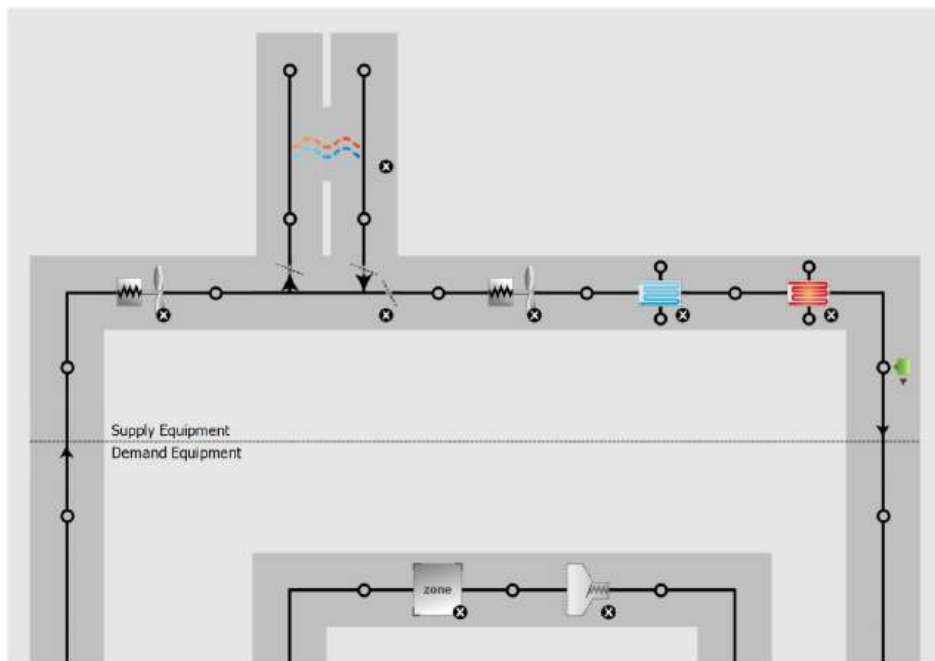


Fig. 7. OpenStudio diagram of ventilation system HF01.

Ruby scripting language [40], which can help automatize otherwise tedious tasks, such as changing the value of a specific parameter on a large number of components.

The HVAC systems were modeled based on PI diagrams with most of the important inputs available such as rated values for capacities, temperatures, and air and water massflows for components such as radiators, fans, air terminals, heating, and cooling coils, etc. Ventilation systems HF01, HF03, and HF04 were modeled with the same air-loop configurations as the actual systems. In Fig. 7 an OpenStudio diagram is shown for ventilation unit HF01. The figure shows showing the configuration of the different components. Namely, one heat recovery unit, two variable speed fans, one cooling coil, one heating coil and a VAV diffuser unit.

For components where input values have been missing, either default OpenStudio values have been used or the Autosizing option has been used. If Autosizing is enabled, the rated values for the component are sized by running several simulations for design days specified by the user (typically a summer design day and winter design day). All airflows for the ventilation systems have been modeled through values provided in the building documentation. This is the case for both the maximum system air flows along with minimum and maximum airflows for each room in the building. Furthermore, general information for the heating and cooling system models as well as ventilation system HF02 model can be found in Table 2, Table 3 and Table 4.

Many of the rooms in the building have thermostats, which are accessible by the occupants. This, therefore, makes the modeling of the temperature setpoint for each room more difficult, because thermal comfort is determined by many factors such as relative humidity, air temperature, air movement, radiation, and personal preference [41,42].

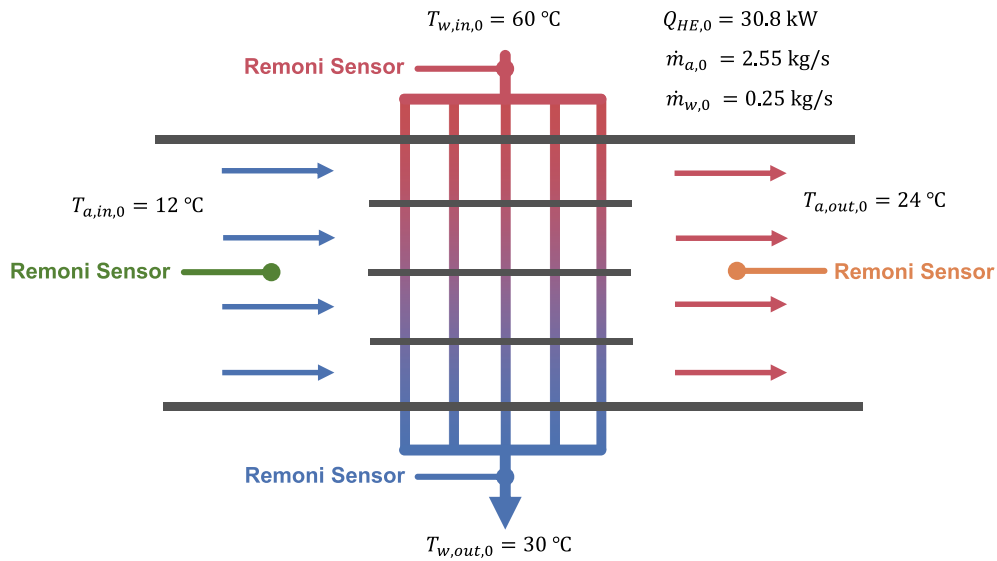


Fig. 8. Illustration of the cross-flow unmixed heat exchanger and the sensor placement. The temperatures shown are the rated operating temperatures provided by the manufacturer.

Table 1

List of components chosen for commissioning.

Component	Ventilation system	Measurement type
Fans	HF02	Power
Heating coil	HF02	Temperature (air inlet/outlet, water inlet/outlet)
Supply air temperature controller	HF02	Temperature (air outlet)
Heat recovery unit	HF02	Temperature (air inlet)

Table 2

General information for the heating system model.

Heating source	District heating
Capacity (sized by EnergyPlus)	~ 270 kW
Supply temperature setpoint	60 °C
Pump type	Variable speed
Number of radiators supplied	72
Number of air loop heating coils supplied	11

Table 3

General information for the cooling system model.

Cooling source	Air cooled chiller
Capacity	124 kW
Supply temperature setpoint	10 °C
Pump type	Variable speed
Number of air loop cooling coils supplied	8

Table 4

General information for the ventilation system HF02 model.

Capacity branch 1	13 000 m ³ /h
Capacity branch 2	7500 m ³ /h
Capacity branch 3,4,5	750 m ³ /h
Supply temperature setpoint	Temperature reset controlled (Refer subsection 2.3.3)
Fan type	Variable speed

For modeling the envelope of the building, almost all information has been available regarding the material layers and thickness for the walls, ceilings, floors, and roofs. Most of the materials used have either been found in the default OpenStudio library or have been imported from BCL with preset values for properties such as conductivity, specific heat, density, absorptance, etc.

The final step in Fig. 6 shows the transition from modeling to simulation. When the developed OpenStudio model is ready for simulation, it is translated into an IDF format by the OpenStudio software, which is then used as input for the EnergyPlus engine. Along with the IDF, an EnergyPlus Weather File (EPW) must be supplied, which contains data on an hourly resolution for variables such as outdoor temperature, atmospheric pressure, wind speed, radiation levels, etc. In this work, all data required in this weather file has been extracted from the Meteoblue website [43]. Meteoblue offers both historic weather data and real-time API services to access data from weather stations all over the world. However, the data used in this work is historic data that have been collected at a local weather station in Odense, located around 3 km from the building site [44].

2.3. Component modeling and calibration

This section is divided into four subsections covering four important components of the studied building; Airloop heating coil, Supply air temperature controller, Ventilation fans, and Heat recovery unit. These are the components that are considered for re-commissioning in this work. All sections have a description of how the component is modeled in EnergyPlus. For the ventilation fans and the heat recovery unit, it is further described how these components are calibrated.

2.3.1. Heating coil

The water-to-air heating coil is used in the case system to heat the supply air such that the supply temperature setpoint is satisfied. Usually, the recovery unit, located before the heating coil, can take care of the majority of the heating load, such that the heating coil only has a heating load of a few degrees. However, being a few degrees off the supply temperature setpoint can have significant consequences for the temperature control at room level. The heating coil is therefore important to consider in a commissioning process.

For the heating coil considered, unit specifications were obtained from the manufacturer in the form of rated water and air inlet and outlet temperatures, $T_{w,in,0}$, $T_{w,out,0}$, $T_{a,in,0}$, $T_{a,out,0}$, along with the rated capacity and water and air massflows, Q_0 , $\dot{m}_{w,0}$, $\dot{m}_{a,0}$. The heat exchanger has been illustrated in Fig. 8 with the rated values and the placement of the sensors.

In EnergyPlus, the heating coil can be modeled using the *Hot-Water-Based Air Heating Coil* model, which is a steady-state mathematical model that can be solved sequentially, given the water and air massflows and inlet temperatures. During simulation, the flow-varying UA-value is first calculated using the model developed by Wetter [45]. Following, the *effectiveness-Number of Transfer Units (NTU)* method is employed to calculate the transferred heat and the outlet temperatures of the water and airflows, assuming an unmixed cross-flow configuration of the heating coil. The full mathematical formulation is presented in the EnergyPlus Engineering Reference [46].

2.3.2. Ventilation fans

Ventilation fans are responsible for 15–50% of the total HVAC energy consumption, depending on the type of system [47] and are thus key components to consider when auditing a building. This section will first describe the mathematical models used by EnergyPlus. This includes how the electricity consumption is calculated and how the fan component affects the other modeled components concerning the temperature rise across the fan. Subsequently, it is presented how the pressure rise and the power coefficients are calibrated such that the minimum and maximum power consumption of the model match the real system at the minimum and maximum airflows. Finally, different parameters influencing the operation of airflows are investigated and tuned to mimic the power consumption peak during occupied hours.

For the fans shown in Fig. 5, the OpenStudio VAV fan is used. The most important inputs for this model are the total efficiency of the fan, the maximum flow rate, the pressure rise, and the four coefficients modeling the part-load curve of the fan. While the total efficiency and the maximum flow rate are known for all fans, the pressure rise and the coefficients are not known. Therefore, to find values for these parameters, the underlying EnergyPlus *Variable Speed Fan Model* [46] is inspected. The mathematical model describing the power consumption $\dot{W}_{fan}$ of this component is given by Equations (1)–(3).

$$f_{flow} = \frac{\dot{m}_a}{\dot{m}_{a,max}} \quad (1)$$

$$f_{pl} = c_1 + c_2 f_{flow} + c_3 f_{flow}^2 + c_4 f_{flow}^3 \quad (2)$$

$$\dot{W}_{fan} = \frac{f_{pl} \dot{m}_{a,max} \Delta P_{max}}{\eta_{tot} \rho_a} \quad (3)$$

where, f_{flow} in Equation (1) is the flow fraction, $\dot{m}_a$ is the operating air massflow and $\dot{m}_{a,max}$ is the maximum airflow. In Equation (2), the part load factor f_{pl} is modeled as a third order polynomial of the flow fraction with the coefficients c_1 – c_4 . In OpenStudio, these coefficients have a default value obtained from the EnergyPlus Input Output Reference [48]. However, later in this section, it is shown how these values can be changed to better describe the measured consumption on the ventilation systems. In Equation (3), ΔP_{max} is the pressure rise across the fan at maximum air flow, η_{tot} is the total efficiency of the fan and ρ_a is the air density. A diagram of the fan configuration for HF02 can be seen in Fig. 9.

Knowing the maximum accumulated power consumption of all the fans for each ventilation system $\dot{W}_{acc}$, the maximum air

massflows and the total efficiency η_{tot} , the pressure rise ΔP_{max} across each fan can be calculated. With $f_{pl} = 1$ at maximum flow, balanced supply and exhaust airflows and equal total efficiencies, Equation (4) can be derived from Equation (3).

$$\Delta P_{max} = \frac{1}{2} \frac{\dot{W}_{acc,max} \eta_{tot} \rho_a}{\dot{m}_{a,max}} \quad (4)$$

Using Equation (4) to find ΔP_{max} , it is thus ensured that the predicted power consumption of the model equals the maximum consumption of the actual system at maximum airflow. A minimum airflow is also defined for each room. Therefore, to ensure that the predicted power consumption also equals the consumption of the actual system at minimum air flows, the coefficients c_1 - c_4 in Equation (2) are fitted accordingly.

Through the collected consumption data, it has been estimated that the minimum electrical power of the ventilation system is $\dot{W}_{acc,min} = 7.55$ kW under normal conditions. This value is used with the already known minimum airflow value $\dot{m}_{a,min}$ to calculate the corresponding part load factor $f_{pl,min}$ and flow factor $f_{flow,min}$ given by Equation (5) and Equation (6).

$$f_{flow,min} = \frac{\dot{m}_{a,min}}{\dot{m}_{a,max}} \quad (5)$$

$$f_{pl,min} = \frac{\dot{W}_{acc,min}}{\dot{W}_{acc,max}} \quad (6)$$

The final step is then to solve a linear system of equations, which enforces certain constraints on the desired coefficient values. This is expressed in Equations (7) and (8), in terms of the two coefficients c_3 and c_4 , where $c_1 = c_2 = 0$ is assumed.

$$f_{pl,min} = c_3 f_{flow,min}^2 + c_4 f_{flow,min}^3 \quad (7)$$

$$c_3 + c_4 = 1 \quad (8)$$

Here, Equation (7) ensures that the minimum measured consumption is obtained at the minimum air flow and Equation (8) ensures that $f_{pl} = 1$ when $f_{flow} = 1$. Solving the two equations then gives the coefficient values of $c_3 = -0.3384$ and $c_4 = 1.3384$. Using the obtained values for the coefficients, the resulting power consumption as function of massflow have been plotted for ventilation system HF02 in Fig. 10. In the figure, two points have been included, which show the minimum and maximum points of operation.

Now that the missing parameter ΔP_{max} has been found for each fan and the coefficients c_1 - c_4 have been fitted to the measured data, the operation of air flows in the ventilation systems must be calibrated. Within the two bounds of minimum and maximum massflow shown in Fig. 10, the air flows are by default controlled to meet all heating and cooling room temperature setpoints. In the actual building, this temperature setpoint control is overruled if the CO₂-concentration rises above 800 parts per million (ppm), forcing increased airflows to bring down the CO₂-level. This type of control is called Demand-Controlled Ventilation (DCV), which EnergyPlus supports. To implement DCV in EnergyPlus, a CO₂ setpoint schedule must be specified. Here, a constant setpoint value of 800 ppm has been implemented in the model. Additionally, an outdoor CO₂ concentration schedule must be specified. This defines the amount of CO₂ contained in the outdoor air, which enters the rooms through both the supply air and the infiltration. It has been set to a constant 400 ppm as recommended in Ref. [49]. With this input along with the occupancy schedules and the zone air flows, EnergyPlus calculates the CO₂ concentration for all zones each timestep and decides whether any control action should be taken to meet the CO₂ setpoint.

With both temperature and CO₂ setpoint control implemented in the model, a large number of parameters influence the supply airflow demanded in each room, including lighting and equipment load schedules, infiltration, occupancy schedules, and temperature

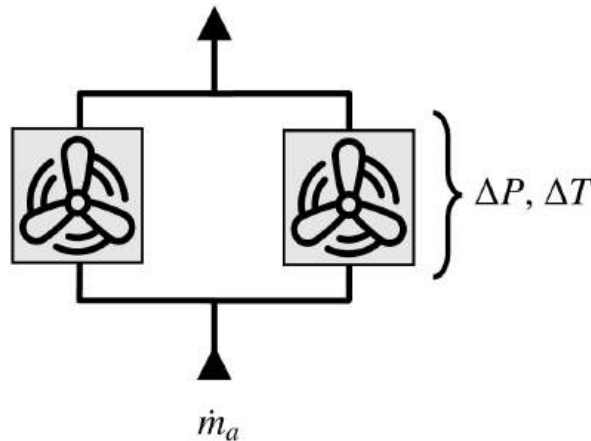


Fig. 9. Fan configuration for HF02.

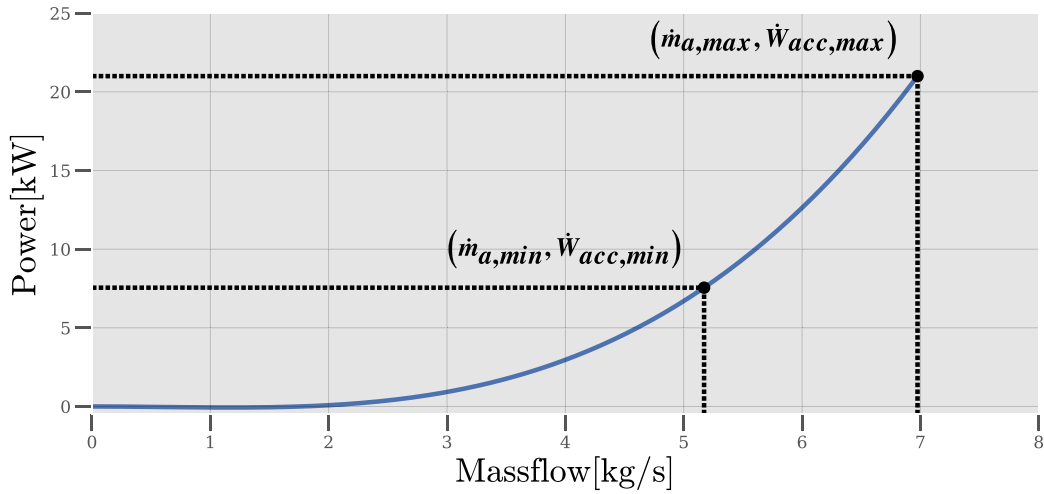


Fig. 10. Total fan power consumption as function of massflow for ventilation system HF02 after estimating the pressure rise ΔP_{max} and the two coefficients c_3 and c_4 .

setpoints. To keep the model transparent and convenient to calibrate, two types of occupancy schedules have been used for most of the building; one for frequently occupied areas such as offices, waiting rooms, meeting rooms, laboratories, etc. and one for common areas such as corridors, toilets, staircases, etc. For lighting and equipment loads, two schedules have been defined, each covering the whole building. The daily schedules for the building are shown in Fig. 11.

The implemented schedules are largely inspired by the archetype schedules recommended by ASHRAE. For lighting and equipment, capacities used are based on information collected from the building. For lighting, an average nominal load of 4.0 W/m^2 has been used for the whole building, while for electric equipment, values ranging from 1.7 to 9.0 W/m^2 has been used, depending on the space type. Finally, an infiltration rate of $1\text{e-}4 \frac{\text{m}^3/\text{s}}{\text{m}^2}$ (flow per exterior surface area) has been assumed for the whole building. To assess the impact of the calibration process, two different simulations have been run for a three-day period and the results are shown in Fig. 12.

The figure shows the actual measurements for the ventilation fan power consumption and air temperature measurements from a

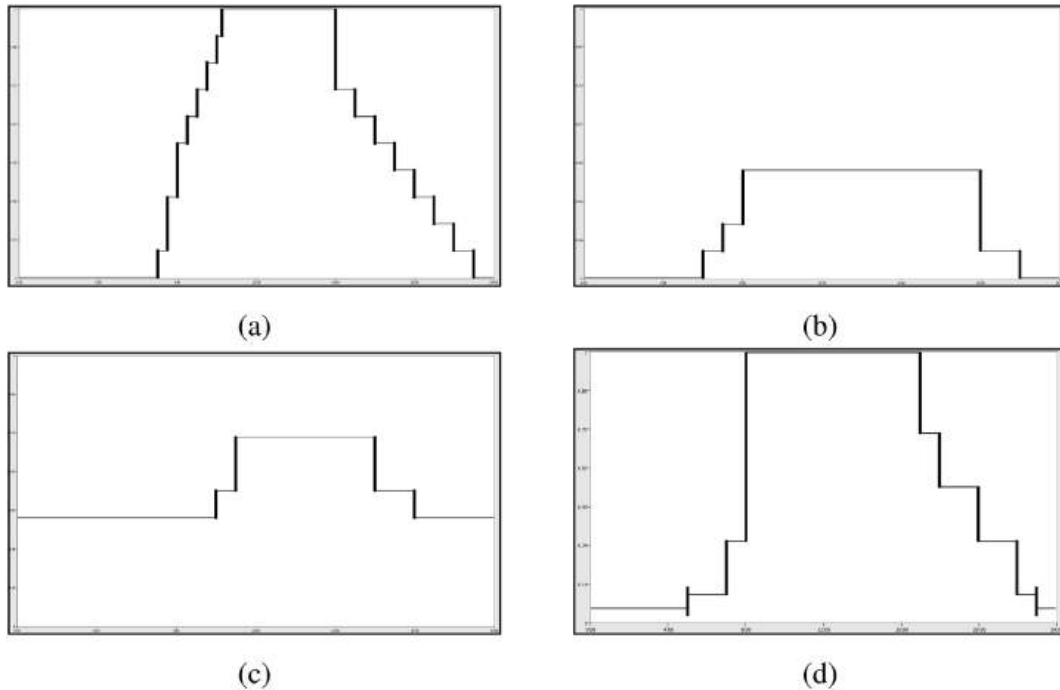


Fig. 11. Daily schedules with time on the first axis and fractional value on the second axis. (a) Occupancy schedule used for frequently occupied areas (b) Occupancy schedule used for common areas (c) Electric equipment schedule (d) Lighting schedule.

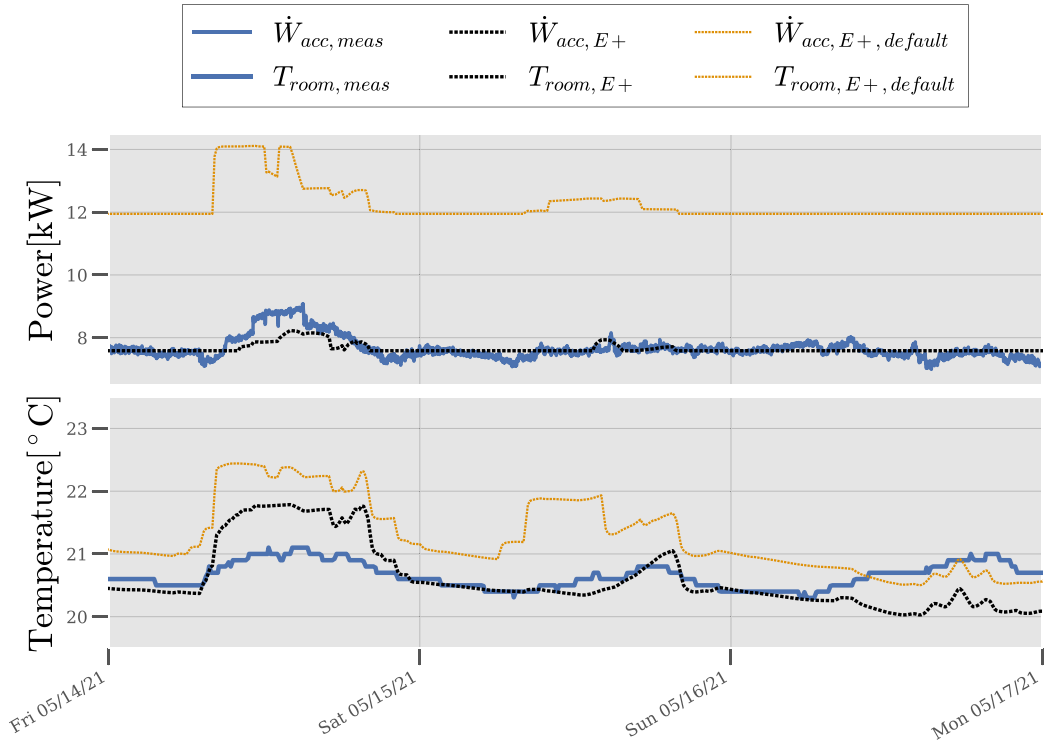


Fig. 12. Comparison of measured and predicted values for fan power consumption and room temperature for default and calibrated EnergyPlus model.

measured office room, along with the EnergyPlus predictions. The first simulation is denoted $E+$ and shows the predictions of the calibrated model. In the second simulation, denoted $E+, default$, the occupancy, lighting, and equipment schedules used in the calibrated model have been replaced by a low-resolution whole-building ASHRAE template from the 189.1–2009 standard. Here, nominal occupancy and nominal lighting and equipment loads are all defined with 1 W-per-square-meter value with one schedule for the whole building each. Additionally, the coefficients c_1 – c_4 have been set to default EnergyPlus values, found in the Input Output reference [48], instead of the values used in the calibrated model. This simulation thus acts as a baseline for an uncalibrated fan model with uncalibrated loads and schedules.

As seen from Fig. 12, there is a significant difference between the two model configurations both in terms of ventilation power

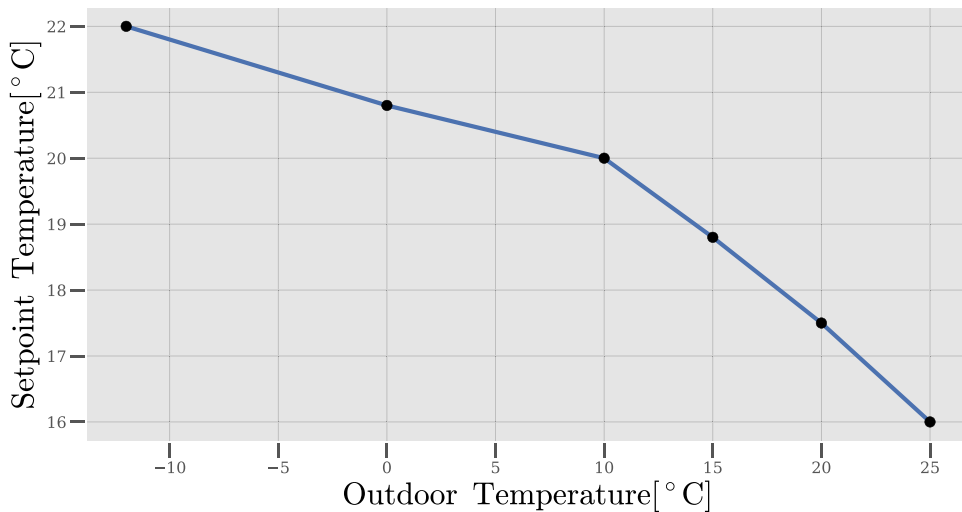


Fig. 13. Piecewise linear function defining the temperature reset strategy that is implemented in the building for ventilation system HF02.

consumption and room temperatures. While the calibrated model produces a good estimate of the measured consumption, the default model configuration generally overestimates the power consumption by around 4 kW. The downwards shift in predictions for the calibrated model can be attributed to the updated power coefficients, while the change in profile shape can be attributed to the change in schedules.

2.3.3. Supply air temperature controller

As mentioned earlier, the case study building is mainly heated by the ventilation system. Furthermore, the only source of cooling in the majority of the rooms is through the ventilation system. It is thus vital that the supply temperature of the ventilation systems is controlled properly, i.e. to heat a room using the ventilation system, the supply air temperature must be above the air temperature of the room. On the other hand, to cool a room, the supply air temperature must be below the air temperature of the room. This makes it undesirable to have a constant temperature setpoint because in many cases it would be impossible to satisfy the room temperature setpoint while also satisfying the airflow requirements of the room. Alternatively, heating and cooling must occur simultaneously to compensate, which is inefficient [50]. To solve this problem, many buildings employ a so-called *temperature reset* strategy, where the heating and cooling loads are assumed to vary as a function of the outdoor temperature. For this control method, the supply temperature is increased when the outdoor air temperature is low and decreased when the outdoor air temperature is high.

For the studied building, these functions have been provided by the building operators for ventilation system HF02 in Fig. 13. EnergyPlus has built-in support for this type of control when the temperature setpoint can be expressed with one linear function. However, it does not support multiple piecewise linear functions as presented here.

Traditionally, this type of problem has been solved using the Energy Management System (EMS) in EnergyPlus [51]. EMS allows the user to override many of the internal variables and routines performed during an EnergyPlus simulation using the EnergyPlus Runtime Language (ERL), which supports simple programming operations such as if-then statements and loops. A program would then be written in the IDF, allowing the user to implement custom control. Although this adds significant capabilities to EnergyPlus, EMS has certain limitations as ERL does not support nested loops, array structures, or user-defined functions [51,52].

To rectify these shortcomings, EnergyPlus has recently added an embedded python interpreter, which can read and write to EnergyPlus variables during simulation much like EMS. However, with the much more mature and complete programming language of Python and the extensive amount of powerful libraries such as the numerical computing library NumPy [53] or the machine-learning library PyTorch [54], adding user-defined functionality to EnergyPlus has become much more effective and reliable. Therefore, the authors have chosen to implement the temperature reset control strategy in this framework.

Using the Python API, custom routines can be injected at various calling points each timestep during the EnergyPlus simulation as shown in Fig. 14. In EnergyPlus, various load calculations and reporting procedures are performed each timestep. Therefore, the calling point must be chosen carefully, to ensure that potentially required variables are defined when the python function is called. Furthermore, the calling point must also be chosen before the variables to be overridden are used by EnergyPlus in the timestep. Otherwise, changing the values of the variables will not affect the simulation.

To implement the reset control strategy, the supply air setpoint for HF02 is therefore overridden at the very start of each simulation timestep using the calling point *BeginTimestepBeforePredictor*. At the chosen calling point, the EnergyPlus variable *Site Outdoor Air DryBulb Temperature* can be accessed and thereby used to calculate the supply air setpoint for HF02, by interpolating between the points shown in Fig. 13.

2.3.4. Heat recovery unit

The heat recovery unit is another key component in the modeled system. From documentation provided by the building operators,

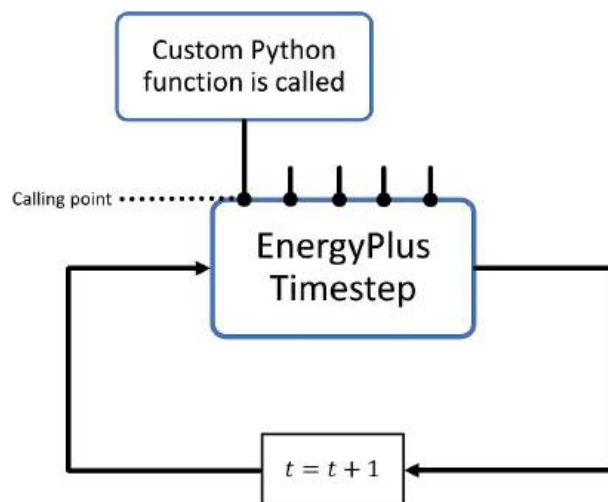


Fig. 14. Illustration of the framework used to control the supply air temperature.

the heat recovery unit for HF02 is stated to have a rated heating capacity of 388,8 kW. Poor modeling of this component will thus affect the predictions of the heating coil model drastically to either be overly passive or overly active to meet the supply air temperature setpoint. In this section, the mathematical models used by EnergyPlus to model a rotary wheel heat recovery unit are first described. Following, it is described how the outlet temperature of the heat recovery unit can be derived from the measured heating coil inlet temperature. This heat recovery outlet temperature is then used to find optimal parameter values for the EnergyPlus model.

The EnergyPlus heat recovery wheel model *Air System Air-To-Air Sensible and Latent Effectiveness Heat Exchanger* accounts for both sensible and latent heat transfer in two possible modes; heating and cooling. When the outdoor air temperature is lower than the exhaust inlet temperature and the outdoor air temperature is lower than the supply temperature setpoint, the unit operates in heating mode. Contrary, when the outdoor air temperature is higher than the exhaust inlet temperature and the outdoor air temperature is higher than the supply temperature setpoint, the unit operates in cooling mode. However, because only temperatures are measured for the actual unit, only the sensible heat transfer is calibrated for this model. The latent heat transfer part of the model is therefore left at the default configuration and is not considered further. From the Energyplus Engineering Reference [46], the governing equations for the sensible part of the model are given as Equations (9)–(14). First, the operating effectiveness of sensible heat transfer ϵ_s is calculated in Equation (9).

$$\epsilon_s = \epsilon_{s,75\%} + (\epsilon_{s,100\%} - \epsilon_{s,75\%}) \frac{f_{flow} - 0.75}{1 - 0.75} \quad (9)$$

where, $\epsilon_{s,75\%}$ is the sensible effectiveness at 75% airflow and $\epsilon_{s,100\%}$ is the sensible effectiveness at 100% airflow. The sensible effectiveness of the heat recovery unit is thus assumed to be a linear function of the flow fraction f_{flow} , by interpolating from the user-provided effectiveness at 75% and 100% flowrate. In the EnergyPlus reference, f_{flow} is given as the average volumetric flow fraction between the supply and exhaust flow as shown in Equation (10).

$$f_{flow} = \frac{\frac{1}{2}(\dot{V}_{sup} + \dot{V}_{exh})}{\dot{V}_{max}} \quad (10)$$

where, $\dot{V}_{sup}$ and $\dot{V}_{exh}$ is the volumetric flowrate of the supply and exhaust air respectively and $\dot{V}_{max}$ is the maximum volumetric flowrate. Comparing Equation (10) with the definition presented in Equation (1) for the ventilation power consumption model, it can be seen that if a balanced supply and exhaust air flow is assumed along with neglecting changes in density with temperature, then Equation (1) and Equation (10) are equivalent. If this is done, then f_{flow} found in Equation (9) can be directly applied as calibration input for the heat recovery unit. With the operating effectiveness ϵ_s , the supply outlet temperature $T_{sup,out}$ is then calculated in Equation (14). Here, $T_{exh,in}$ is the inlet temperature of the exhaust air and T_{OA} is the outdoor air temperature. The minimum heat capacity rate $\dot{C}_{min}$ is given by Equation (13), where $\dot{C}_{sup}$ is the heat capacity rate of the supply air and $\dot{C}_{exh}$ is the heat capacity rate of the exhaust air given by Equation (11) and Equation (12) respectively.

$$\dot{C}_{sup} = \dot{m}_{sup} c_{p,sup} \quad (11)$$

$$\dot{C}_{exh} = \dot{m}_{exh} c_{p,exh} \quad (12)$$

$$\dot{C}_{min} = \min(\dot{C}_{sup}, \dot{C}_{exh}) \quad (13)$$

$$T_{sup,out} = T_{OA} + \epsilon_s \left(\frac{\dot{C}_{min}}{\dot{C}_{sup}} \right) (T_{exh,in} - T_{OA}) \quad (14)$$

Again, if a balanced air massflow for the supply and exhaust side is assumed along with neglecting variations in specific heat capacities, then Equations (11)–(14) can be reduced to Equation (14) where $\frac{\dot{C}_{min}}{\dot{C}_{sup}} = 1$. This simplification will become useful during the calibration process.

The default parameter values for the model in OpenStudio is $\epsilon_{s,75\%} = 0.81$ and $\epsilon_{s,100\%} = 0.76$. However, using the collected data, it is possible to improve these parameter values to achieve more accurate model predictions. For this purpose, a least-square optimization problem has been formulated to minimize the discrepancy between the predicted outlet temperature and the measured temperature. To solve this optimization problem four inputs are needed:

- T_{OA} : Outdoor temperature. Known from the used weather data.
- $T_{exh,in}$: Inlet exhaust temperature. It is assumed that the measured room temperature for the office is representable for the inlet exhaust air temperature, although these rooms contribute a small fraction of the total exhaust air.
- $T_{sup,out}$: Outlet supply temperature. It is derived from the measured heating coil inlet temperature by accounting for the fan temperature rise, which is calculated from the work added to the air stream.
- f_{flow} : Flow fraction. It is derived from the measured electricity consumption using Equations (1)–(3).

A period of 4 consecutive days from 4th of May to May 8, 2021 have been chosen for calibration. The first 3 days are used for training while the last day is used for testing. For this period, the 4 calibration inputs $T_{OA,meas}$, $T_{exh,in,meas}$, $T_{sup,out,meas}$ and $f_{flow,meas}$ have

been shown in Fig. 15.

In Fig. 15 each graph represents approximately 5800 observations, corresponding to a measurement frequency of 1 per minute. Solving the least-squares optimization program for the data yields an optimal value for the parameters $\varepsilon_{s,75\%} = 0.827$ and $\varepsilon_{s,100\%} = 0.801$. If these effectiveness values are compared with performance testing values for other rotary recovery units certified by the Air-Conditioning, Heating and Refrigeration Institute (AHRI) [55], the found parameter values are found to be reasonable. From the database, $\varepsilon_{s,75\%}$ values span from 0.67 to 0.86 and $\varepsilon_{s,100\%}$ values span from 0.64 to 0.84. Furthermore, as also evident from this database, $\varepsilon_{s,75\%}$ is consistently higher than $\varepsilon_{s,100\%}$, which also fits with the obtained values. The resulting supply outlet temperature predictions compared with the measured temperatures have been shown in Fig. 16. In the figure, predictions using the OpenStudio default parameter values have also been included to show the impact of the calibration.

From the figure, it can be seen that the predicted temperatures $T_{sup,out,E+}$ and $T_{sup,out,E+,default}$ follow the general tendency of the measured temperatures $T_{sup,out,meas}$. Visually, it seems the calibrated model matches the measured data better. However, to quantify this the Mean Absolute Error (MAE) and the Mean Squared Error (MSE) have been calculated for both cases. The results are shown in Table 5 for both the training data and the test data.

As seen from Table 5, there is a clear improvement in both metrics for the calibrated model. It is noted that the training error is higher than the test error for the calibrated model. However, this is also the case for the default model, indicating that the explanation is found in the dataset.

3. Results and discussion

In this section, the results of the performed re-commissioning are presented and discussed. The commissioning is performed by simulating the developed EnergyPlus whole building model while monitoring and comparing the model predictions of the selected components with actual reported data measured on-site. The predictions of the model are compared with measured data using a fault detection method called *limit checking*. This method uses either constant or variable predefined thresholds to determine whether an observation should be categorized as an anomaly or normal operation [56]. Despite the simplicity of this method, it is frequently used in many different applications. For building energy performance evaluation, it is often implemented as part of a continuous commissioning framework by comparing the expected performance predicted by an EnergyPlus model with reported data [17,18,57]. Here, it is also stated that the thresholds should be chosen carefully to avoid alarm fatigue, i.e. when the thresholds are set too low and a high rate of false-alarms are triggered, which over-loads the building operators. On the other hand, a too high threshold could lead to undetected anomalies. In this work, however, this problem is not considered further, and the thresholds have therefore been chosen for demonstration purposes.

As outlined in Table 1, commissioning of the heating coil, the supply air temperature controller, and the heat recovery unit is



Fig. 15. Inputs for calibration of the heat recovery model.

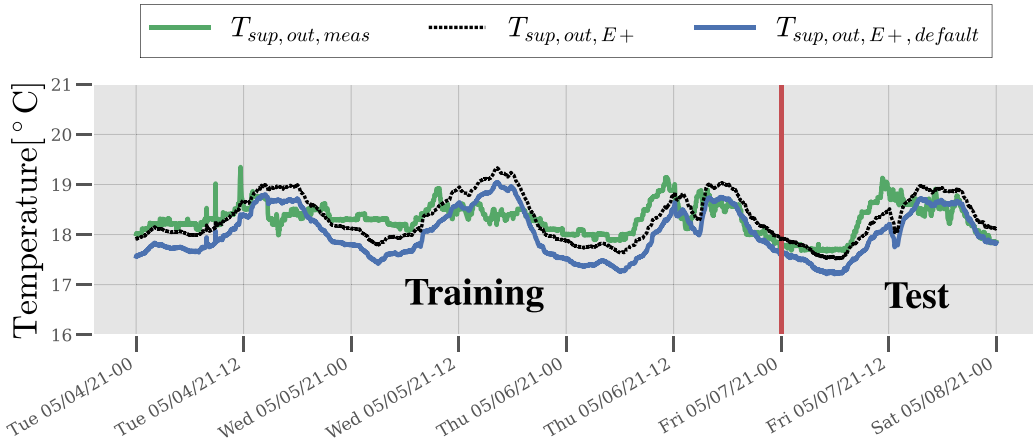


Fig. 16. Comparison of predicted and actual measured supply outlet air temperature of the heat recovery unit for both calibrated and default parameter values.

Table 5

Mean Absolute Error and Mean Squared Error for the default parameters and the calibrated parameters on the training and test data.

	Training		Test	
	MAE	MSE	MAE	MSE
Default	0.41	0.22	0.33	0.18
Calibrated	0.28	0.12	0.26	0.094

performed using temperature measurements from the heating coil. For these components, a period has been chosen of nine consecutive days from Sunday 9th of May to Monday May 18, 2021. To detect any anomaly behavior for these components, the predicted temperatures are compared to the measured temperatures using a threshold of 5%. In Fig. 17, the inlet and outlet water and air temperatures $T_{w,in}$, $T_{w,out}$, $T_{a,in}$, $T_{a,out}$ are shown for both the measured data (denoted *meas*) and the simulation (denoted *E+*). Additionally, the outdoor air temperature T_{OA} is shown.

3.1. Heating coil

From Fig. 17, it is seen that the simulated inlet water temperature is set to a constant 60 °C. In the actual building, the inlet temperature is dependent on the district heating forward temperature and is thus not controllable. Therefore, there are times where the measured temperature deviates substantially from the assumed temperature in the simulation. However, this deviation will only affect the predicted water massflow and outlet water temperature. It will not affect the outlet air temperature as long as it is physically possible to transfer the required heat to meet the supply air temperature.

From the outlet water temperature $T_{w,out}$ in Fig. 17, it can be seen that the EnergyPlus predictions only match the measured temperature for specific periods during the simulation period. This is caused by an EnergyPlus convention that sets the water outlet temperature equal to inlet temperature when the heating coil is not in operation. For the actual system, when the water flow valve is shut, the pipe temperature starts to decrease and approach the airflow temperature, explaining the multiple temperature drops of the inlet water temperature measurements (box 1). Comparing these temperature drops with the predicted water outlet temperature, they all coincide with the heating coil being off in the EnergyPlus simulation. These occurrences of sudden temperature drops, which in some cases could be misinterpreted as faults or anomalies, can thus be justified by the calibrated model.

For the first two days of the commissioning period, Sunday and Monday, the outdoor air temperature T_{OA} is considerably higher than for the rest of the period, with temperature peaks around 24 °C (box 2). As a consequence, the measured inlet and outlet air temperatures deviate substantially from the predicted values and it is clear that the 5% error threshold is violated for the inlet and outlet air temperatures. However, during this period, the heating coil in the actual building correctly turns off as seen from $T_{w,in,meas}$, causing the measured air temperature difference across the heating coil to be practically zero. The same behavior is observed for the simulated heating coil, although the air temperatures here are much lower. Altogether this leads to the conclusion that the anomaly observed in the building is not caused by the heating coil.

3.2. Supply temperature setpoint

To make sure that the low air temperatures predicted by the model are not caused by a wrong setpoint, it has been carefully verified that the EnergyPlus predictions for the outlet air temperature $T_{a,out,E+}$ exactly follows the setpoint described in subsection 2.3.3 during the whole commissioning period. As seen from the figure, the measurements $T_{a,out,meas}$ also agree well with the model predictions outside of the discussed period, verifying that the actual controller respects this setpoint. Therefore, it is most likely that the

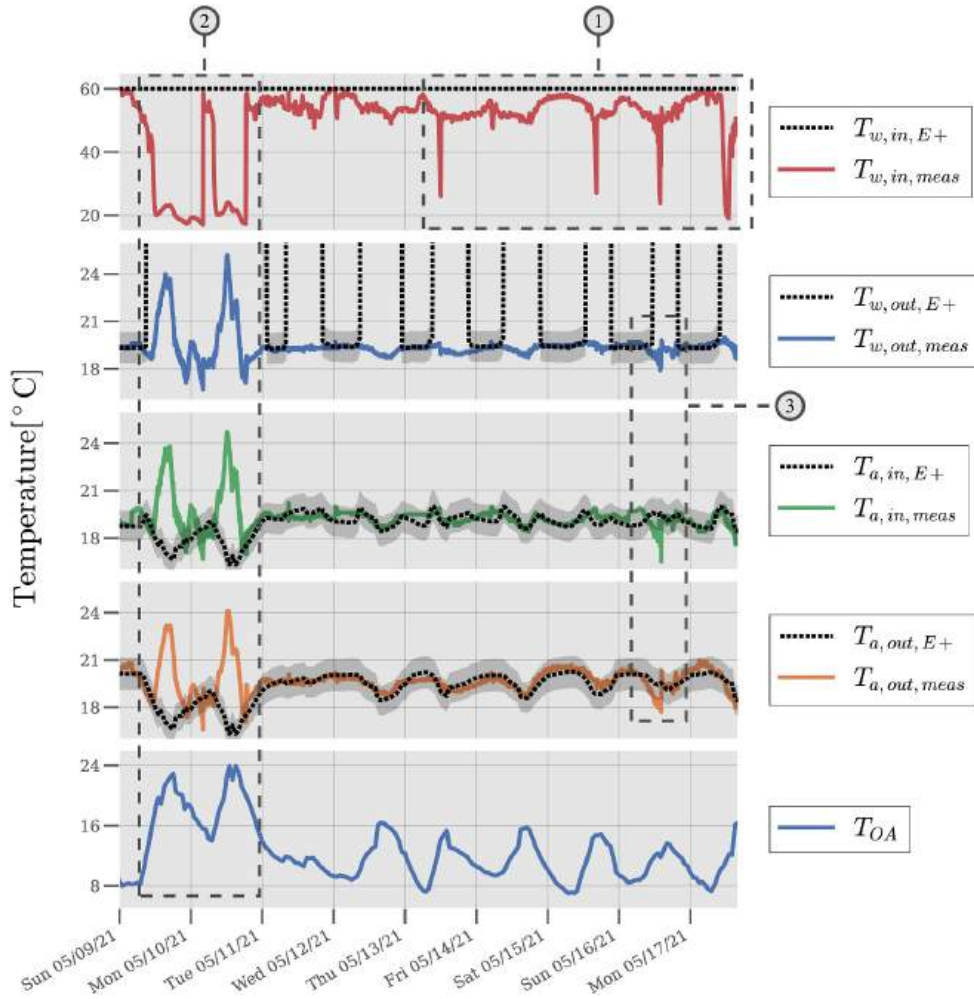


Fig. 17. Inlet and outlet water and air temperatures for the water-to-air heating coil shown with the collected data against the predicted temperatures using the calibrated EnergyPlus model. The grey area represents a 5% error margin with respect to the predicted values.

cooling coil, located just before the heating coil (refer Fig. 5) was not operating as expected. Afterward, this conclusion has been verified by the building operators who confirmed, that the chillers supplying cold water were not operational during this period.

3.3. Heat recovery unit

Although no direct measurements have been collected on the heat recovery unit, the collective behavior of the heat recovery unit and the fan can still be assessed. Comparing the predicted air inlet temperature $T_{a,in,E+}$ with the measured temperature $T_{a,in,meas}$ of the heating coil in Fig. 17, it is seen that these temperatures closely match during most of the period, i.e., the heat recovery unit performs as expected. Ignoring the first two days, that have already been discussed, one additional exception occurs during Sunday 16th of May, where there is a short period with lower inlet and outlet air temperatures than expected (box 3). Although the exact cause for this sudden temperature drop is not known, is most likely caused by the heat recovery unit. However, as also seen, the temperatures return to their expected values shortly after.

3.4. Ventilation fans

The period chosen for commissioning the ventilation fans is fourteen consecutive days from Monday 17th of May to Monday May 31, 2021. Using the calibrated EnergyPlus model, the building performance was simulated for the chosen period and the results are shown in Fig. 18, along with the measured electricity consumption. Similar to the commissioning of the other components, an error margin has been included. However, for the ventilation fans, this error margin was set as 15%, due to high uncertainties in occupancy and load schedules.

As seen from Fig. 18a, the model predictions match the measured profile reasonably well with the majority of the measured consumption being within the error margin. However, the last two days, Saturday and Sunday appear to have unusually high electricity consumption for a weekend, as also indicated by the much lower model predictions. Based on the information collected on the

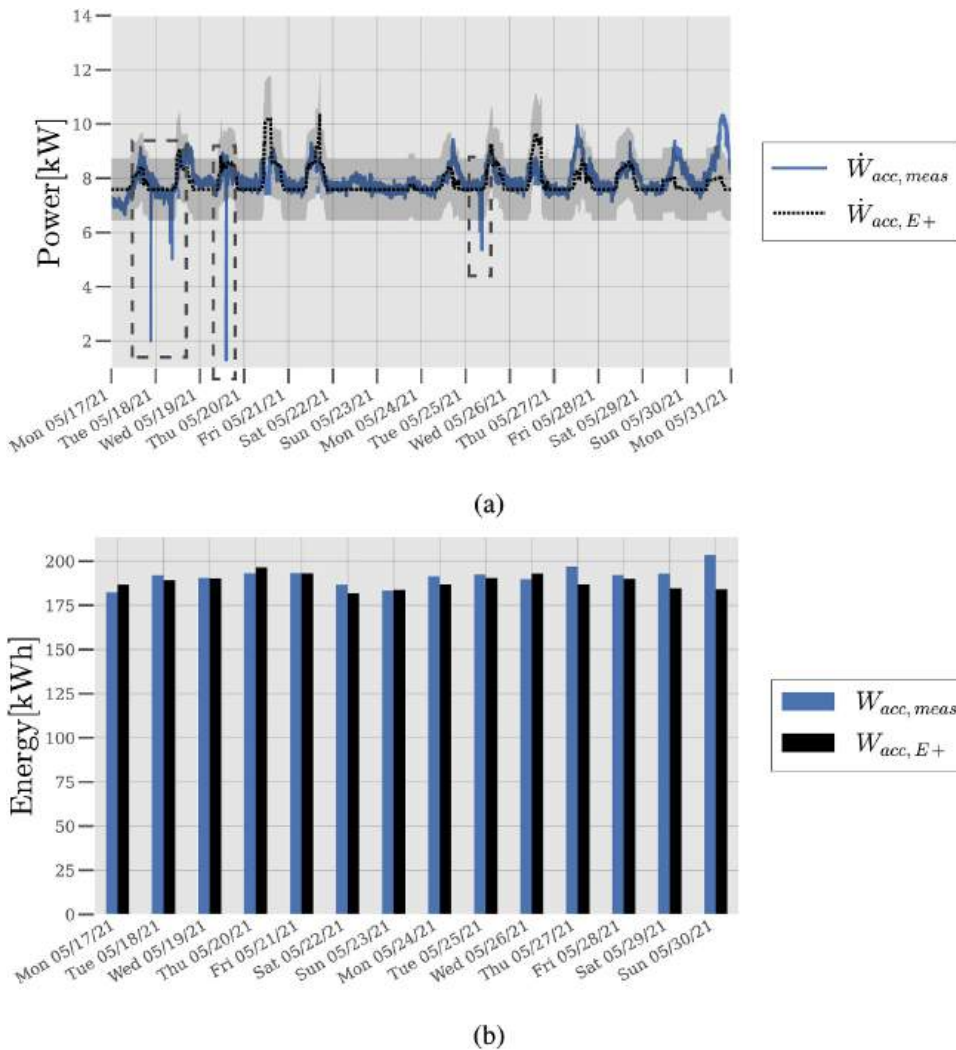


Fig. 18. (a) The total fan power consumption of ventilation system HF02 predicted by the model against the measured values with the grey area representing a 15% error margin with respect to the predicted values. (b) The resulting daily fan energy consumption predicted by the model, against the actual measurements.

building use patterns, it was highlighted that the building is heavily used in full-load from Monday to Friday. This was taken into account in the model development and is shown clearly in the simulations, for example the first weekend period. However, it is evident that some of the building rooms, including screening and laboratories, were used in the second weekend, which has led to higher energy consumption as shown in the figure. This supports our decision to increase the margin error for the power consumption of fans, as the impact of occupancy stochastic behavior and patterns is very evident.

There are four outliers in total during the commissioning period where the power consumption suddenly drops, giving the impression that the ventilation system is turned off for a brief period. Even though majority of the observed power consumption is within the predefined thresholds, it is noted that the model is not fully able to capture the stochastic and dynamic behavior in the building with frequent under and over estimations of the actual power consumption. Therefore, to obtain a more robust perspective on the performance, the daily energy consumption is shown in Fig. 18b. From Fig. 18b, it is seen that the predicted daily energy use of the ventilation system is very close to the actual measured energy use. Besides the detected outliers, the measured performance of the ventilation system is thus very close to the expected performance and the performance of the ventilation units is deemed acceptable for the considered period.

4. Conclusion

The largest-ever comprehensive Danish energy renovation strategy, “Strategy for the energy renovation of the existing building stock” was developed as part of the Danish government energy agreement established in 2012. The strategy was marked as a “significant milestone” in the country’s transition towards a highly efficient energy sector. This has been followed by a recent agreement in 2020, with a total fund of 30.2 MDKK to be invested by 2026 in the green renovation of Danish public buildings. Initial commissioning

of newly built buildings is the norm now in Denmark, and the documentation of proper performance testing at the handover stage is a requirement. However, re-commissioning of deeply retrofitted buildings is not given full consideration and is generally overlooked in the process. Considering the large potential and the technical and economic added value, this work presented a systematic retro-commissioning approach aiming to eliminate any faults and anomalies related to modifications and upgrades implemented in a building undergoing a deep energy retrofit.

A case study of a hospital building in Odense, Denmark is considered and a holistic dynamic model-driven re-commissioning of multiple components in the building is demonstrated. The commissioned components are; ventilation fans, one heating coil, the supply air temperature controller, and one heat recovery unit. The re-commissioning process is performed using a *limit checking* fault detection technique with predefined thresholds. Through the commissioning process, it is among other things found that a cooling coil, in the same air loop as the commissioned heating coil, was not operating properly during the commissioning period.

Existing building retrofit is a very hot topic nowadays in the majority of the countries in Europe but also around the world. In the contest of the current study, the lack of a systematic re-commissioning of retrofitted buildings is generally a common problem in the majority of these countries. The approach provided and implemented in this paper is generic and scalable enough to be used and implemented in a large multitude of buildings, of different type, use, age, energy systems and locations. The only modification and adaptation needed in the re-commissioning process is taking into account the specific systems in use in each case. For example, buildings in hot climates will have large and more sophisticated cooling systems compared to buildings located in countries with generally cold climate. In addition, the use of clamp on sensors, similar to the method used in this study to collect data and calibrate the energy models is a simple, flexible and generic method which can be implemented in various building types and locations without major modifications or specific considerations.

The development and calibration of the whole-building dynamic energy model is a preliminary step for implementing continuous commissioning in the building as part of the BuildCOM research project.

Author statement

Jakob Bjørnskov: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Muhyiddine Jradi:** Conceptualization, Methodology, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition. **Christian Veje:** Conceptualization, Methodology, Writing - review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is carried out under the 'BuildCOM: Automated Auditing and Continuous Commissioning of Next Generation Building Management Systems' project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64019-0081.

References

- [1] E. Markoska, A Framework for Automated Real-Time Performance Evaluation of Smart Buildings, Syddansk Universitet. Det Tekniske Fakultet., 2019.
- [2] I. Khan, A. Capozzoli, S.P. Corngati, T. Cerquitelli, Fault detection analysis of building energy consumption using data mining techniques, *Energy Proc.* 42 (2013) 557–566, <https://doi.org/10.1016/j.egypro.2013.11.057>, mediterranean Green Energy Forum 2013: Proceedings of an International Conference MGEF-13.
- [3] E. Mills, Building commissioning: a golden opportunity for reducing energy costs and greenhouse gas emissions in the United States, *Energy Efficiency* 4 (2011) 145–173.
- [4] C. o. t. E. U., European Parliament, Directive 2010/31/eu of the european parliament and of the council of 19 may 2010 on the energy performance of buildings, *Official Journal of the European Union* 153 (2010) 13–35.
- [5] B. og Planstyrelsen, Bygningsreglementet. <https://byggningsreglementet.dk/>, 2021 (last visited 30/10/21).
- [6] C. Legris, D. Choiniere, N. Milesi-Ferretti, Annex 47 report 1: commissioning overview. <https://doi.org/10.6028/NIST.TN.1725>, 2011.
- [7] W. Vink, W. Zeiler, Re-commissioning the hvac of an office building: systematic performance evaluation according I.e.a, in: *Proceedings CIBSE Technical Symposium, 2018 CIBSE Technical Symposium 2018* ; Conference date: 12-04-2018 Through 13-04-2018.
- [8] L. Liu, V. Georgieva, D. Zabusova, K. Milenkova, R. Christiansen, S. Montes, V. Vecsey, Performance evaluation and improvement of a newly-renovated Danish school building (retro-commissioning), in: *Proceedings of the 15th International Conference of the International Building Performance Simulation Association, Building Simulation Conference Proceedings, IBPSA, 2017*, Pp. 1567–1575, the 15th International Conference of IBPSA, Building Simulations, 2017, <https://doi.org/10.26868/25222708.2017.419>. Conference date: 07-08-2017 Through 09-08-2017, <http://www.buildingsimulation2017.org/>.
- [9] R.C. Coyner, S.W. Kramer, Long term benefits of building commissioning: should owners pay the price? *Procedia Eng.* 196 (2017) 429–435, <https://doi.org/10.1016/j.proeng.2017.07.220>, creative Construction Conference 2017, CCC 2017, 19-22 June 2017, Primosten, Croatia, <https://www.sciencedirect.com/science/article/pii/S1877705817330928>.
- [10] C. Deb, J.Y. Fan Zhang, S.E. Lee, K.W. Shah, A review on time series forecasting techniques for building energy consumption, *Renew. Sustain. Energy Rev.* 74 (2017) 902–924.
- [11] D. Coakley, P. Raftery, M. Keane, A review of methods to match building energy simulation models to measured data, *Renew. Sustain. Energy Rev.* 37 (2014) 123–141, <https://doi.org/10.1016/j.rser.2014.05.007>. <https://www.sciencedirect.com/science/article/pii/S1364032114003232>.
- [12] D. Crawley, L. Lawrie, F. Winkelmann, W. Buhl, Y. Huang, C. Pedersen, R. Strand, R. Liesen, D. Fisher, M. Witte, J. Glazer, Energyplus: creating a new-generation building energy simulation program, *Energy Build.* 33 (2001) 319–331.
- [13] F.C. Winkelmann, B.E. Birdsall, W.F. Buhl, K.L. Ellington, A.E. Erdem, J.J. Hirsch, S. Gates, Doe-2 supplement: version 2.1e (11 1993). <https://www.osti.gov/biblio/10147851>.

- [14] S. Klein, W. Beckman, Trnsys 16: a transient system simulation program: mathematical reference, TRNSYS 5 (2007) 389–396.
- [15] C. Miller, D. Thomas, S. Domingo-Irigoyen, C. Hersberger, Z. Nagy, D. Rossi, A. Schlüter, Bim-extracted Energyplus Model Calibration for Retrofit Analysis of a Historically Listed Building in Switzerland, 2014, <https://doi.org/10.13140/RG.2.1.1671.7285>.
- [16] A. Angelotti, M. Martire, L. Mazzarella, M. Pasini, I. Ballarini, V. Corrado, G. De Luca, P. Baggio, A. Prada, F. Bosco, C. Cornaro, Building energy simulation for nearly zero energy retrofit design: the model calibration, in: 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC/I CPS Europe), 2018, pp. 1–6, <https://doi.org/10.1109/EEEIC.2018.8493777>.
- [17] M. Jradi, N. Liu, A. Johansen, K. Arendt, C. Mattera, M. Kjergaard, C. Veje, B. Jørgensen, Dynamic energy model-based automatic building performance testing for continuous commissioning, in: V. Corrado, E. Fabrizio, A. Gasparella, F. Patuzzi (Eds.), Proceedings of Building Simulation 2019, Proceedings of the International Building Performance Simulation Association, International Building Performance Simulation Association, 2020, pp. 822–829, <https://doi.org/10.26868/25222708.2019.210200>, 16th IBPSA International Conference and Exhibition Building Simulation, Building Simulation 2019; Conference date: 02-09-2019 Through 04-09-2019, <http://buildingsimulation2019.org>.
- [18] L.C. Harmer, G.P. Henze, Using calibrated energy models for building commissioning and load prediction, Energy Build. 92 (2015) 204–215, <https://doi.org/10.1016/j.enbuild.2014.10.078>, <https://www.sciencedirect.com/science/article/pii/S0378778814009244>.
- [19] X. Yu, L. Georges, M.D. Knudsen, I. Sartori, L. Imsland, Investigation of the model structure for low-order grey-box modelling of residential buildings, Int. Build. Perform. Simulat. Assoc. 16 (9) (2019) 5076–5083.
- [20] H. Park, N. Martaj, M. Ruellan, R. Bennacer, E. Monmasson, Modeling of a building system and its parameter identification, J. Elect. Eng. Technol. 8 (2013) 975–983, <https://doi.org/10.5370/JEET.2013.8.5.975>.
- [21] S. Freund, G. Schmitz, Development of a Framework for Model Predictive Control (Mpc) in a Large-Sized Office Building Using Modelica Grey-Box Models, 2019, pp. 2864–2871, <https://doi.org/10.26868/25222708.2019.210443>.
- [22] A. Reddy, Literature review on calibration of building energy simulation programs: uses, problems, procedure, uncertainty, and tools, Build. Eng. 112 (2006) 226–240.
- [23] E. Fabrizio, V. Monetti, Methodologies and advancements in the calibration of building energy models, Energies 8 (2015) 2548–2574, <https://doi.org/10.3390/en8042548>.
- [24] D. Guyot, F. Giraud, F. Simon, D. Corgier, C. Marillet, B. Tremeac, Building energy model calibration: a detailed case study using sub-hourly measured data, Energy Build. 223 (2020) 110189, <https://doi.org/10.1016/j.enbuild.2020.110189>, <https://www.sciencedirect.com/science/article/pii/S0378778820300505>.
- [25] M. Jradi, K. Arendt, F.C. Sangogboye, C.G. Mattera, E. Markoska, M.B. Kjergaard, C.T. Veje, B.N. Jørgensen, Obepme: an online building energy performance monitoring and evaluation tool to reduce energy performance gaps, Energy Build. 166 (2018) 196–209.
- [26] E. Hale, L. Lisell, D. Goldwasser, D. Macumber, J. Dean, I. Metzger, A. Parker, N. Long, B. Ball, M. Schott, E. Weaver, L. Brackney, Cloud-based Model Calibration Using Openstudio, 2014.
- [27] A. Chong, Y. Gu, H. Jia, Calibrating building energy simulation models: a review of the basics to guide future work, Energy Build. 253 (2021) 111533, <https://doi.org/10.1016/j.enbuild.2021.111533>, <https://www.sciencedirect.com/science/article/pii/S0378778821008173>.
- [28] P. da Silva, A. Neto, I.L. Sauer, Evaluation of model calibration method for simulation performance of a public hospital in Brazil, Energies 14 (2021) 3791, <https://doi.org/10.3390/en14133791>.
- [29] M. Royapoor, T. Roskilly, Building model calibration using energy and environmental data, Energy Build. 94 (2015) 109–120, <https://doi.org/10.1016/j.enbuild.2015.02.050>, <https://www.sciencedirect.com/science/article/pii/S0378778815001553>.
- [30] F. Baba, H. Ge, R. Zmeureanu, L. Wang, Calibration of building model based on indoor temperature for overheating assessment using genetic algorithm: methodology, evaluation criteria, and case study, Build. Environ. 207 (2021) 108518, <https://doi.org/10.1016/j.buildenv.2021.108518>.
- [31] K. Lam, J. Zhao, E. Ydstie, J. Wirick, M. Qi, J. Park, An Energyplus Whole Building Energy Model Calibration Method for Office Buildings Using Occupant Behavior Data Mining and Empirical Data, 2014.
- [32] S. Asadi, E. Mostavi, D. Boussaa, M. Indaganti, Building energy model calibration using automated optimization-based algorithm, Energy Build. 198 (06 2019), <https://doi.org/10.1016/j.enbuild.2019.06.001>.
- [33] M. Jradi, N. Boel, B.E. Madsen, J. Jacobsen, J.S. Hooge, L. Kildelund, Buildcom: automated auditing and continuous commissioning of next generation building management systems, Energy Informatics 4 (2 2021), <https://doi.org/10.1186/s42162-020-00136-2>.
- [34] Remoni, HeatMoniSpot. <https://www.remoni.com/products2/product-overview/heatmonispot/>, 2021.
- [35] Remoni, PowerMoniSpot. <https://www.remoni.com/products2/product-overview/powermonispot/>, 2021.
- [36] Remoni, HeatMoni spot datasheet. <https://support.remoni.com/hc/en-us/articles/360018577138-HeatMoniSpot>, 2021.
- [37] Remoni, PowerMoni spot datasheet. <https://support.remoni.com/hc/en-us/articles/360018577078-PowerMoniSpot>, 2021.
- [38] Remoni, Data sheet ReMoni wireless communication (last visited 30/11/21), <https://support.remoni.com/hc/da/articles/360018577438-Datablad-tr%C3%A5d%C3%B8s-kommunikation>, 2021.
- [39] M. Jradi, Dynamic energy modelling as an alternative approach for reducing performance gaps in retrofitted schools in Denmark, Appl. Sci. 10 (2020) 7862, <https://doi.org/10.3390/app10217862>.
- [40] OpenStudio. <https://www.openstudio.net/>, 2021 (last visited 30/11/21).
- [41] S.A. Al-Sanea, M.F. Zedan, Optimized monthly-fixed thermostat-setting scheme for maximum energy-savings and thermal comfort in air-conditioned spaces, Appl. Energy 85 (2008) 326–346.
- [42] J. Kim, S. Schiavon, G. Brager, Personal comfort models – a new paradigm in thermal comfort for occupant-centric environmental control, Build. Environ. 132 (2018) 114–124.
- [43] Meteoblue, Weather history download Odense. https://www.meteoblue.com/en/weather/archive/export/odense_denmark_2615876, 2021 last visited 30/10/21.
- [44] Meteoblue, Weather history download Odense, last visited 30/10/21, https://www.meteoblue.com/en/weather/archive/export/odense_denmark_2615876, 2021.
- [45] M. Wetter, Simulation Model Finned Water-Air-Coil without Condensation, 1999, <https://doi.org/10.2172/7353>, <https://www.osti.gov/biblio/7353>.
- [46] U. D. of Energy, EnergyPlus Version 9.5.0 Documentation, EnergyPlus Engineering Reference, 2020. https://energyplus.net/sites/all/modules/custom/nrel_custom/pdfs/pdfs_v9.5.0/EngineeringReference.pdf (last visited 30/11/21).
- [47] C.W. Andersen, C.A. Hviid, Low-energy mechanical ventilation: a case study of two new office buildings, Proc. 10th Nordic Symp. Build. Phys. 7 (2014) 718–725.
- [48] U. D. of Energy, EnergyPlus Version 9.5.0 Documentation, EnergyPlus Input Output Reference, 2020 last visited 30/11/21, https://energyplus.net/sites/all/modules/custom/nrel_custom/pdfs/pdfs_v9.5.0/InputOutputReference.pdf.
- [49] C.I. of Building Services Engineers, Cibe Guide a: Environmental Design, 2007 (chapter 4).2, Table 4.2.
- [50] W. Guanghua, W. Turner, D. Claridge, M. Liu, Single-duct Constant Air Volume System Supply Air Temperature Reset: Using Return Air Temperature or outside Air Temperature?, 09 2003, [https://doi.org/10.1061/40699\(2003\)23](https://doi.org/10.1061/40699(2003)23).
- [51] U. D. of Energy, EnergyPlus Version 9.5.0 Documentation, Application Guide for EMS, 2020 last visited 30/10/21, https://energyplus.net/sites/all/modules/custom/nrel_custom/pdfs/pdfs_v9.5.0/EMSApplicationGuide.pdf.
- [52] O. of Energy Efficiency Renewable Energy, Python EMS: a unique EnergyPlus feature gets a serious upgrade (last visited 30/10/21), <https://www.energy.gov/eere/buildings/articles/python-ems-unique-energyplus-feature-gets-serious-upgrade>, 2020.
- [53] NumPy, NumPy. <https://numpy.org/>, 2021 (last visited 30/10/21).
- [54] PyTorch, PyTorch. <https://pytorch.org/>, 2021 (last visited 30/10/21).

- [55] R. Air-Conditioning, H. Institute, AHRI's Directory of Certified Product Performance, 2021. <https://www.ahridirectory.org/NewSearch?programId=38&searchTypeId=3> (last visited 30/10/21).
- [56] R. Isermann, *Fault-Diagnosis Systems*, Springer-Verlag Berlin Heidelberg, 2006 (chapter 7).1.
- [57] M. Jradi, N. Liu, K. Arendt, C.G. Mattera, An automated framework for buildings continuous commissioning and performance testing – a university building case study, *J. Build. Eng.* 31 (2020) 101464, <https://doi.org/10.1016/j.jobe.2020.101464>. <https://www.sciencedirect.com/science/article/pii/S2352710220302412>.

A Modular Thermal Space Coupling Approach for Indoor Temperature Forecasting Using Artificial Neural Networks

Abstract

With the increasing digitalization of buildings and the adoption of comprehensive sensing and metering networks, the concept of building digital twins is emerging as a key component in future smart and energy-efficient buildings. Such digital twins enable the use of flexible and adaptable data-driven models to provide services such as automated performance monitoring and model-based operational planning in buildings. In this context, accurate indoor temperature models are vital to ensure that the proposed operational strategies are effective, feasible, and do not compromise indoor comfort. In this work, the significance of thermal space coupling for data-driven indoor temperature forecasting is investigated by assessing and comparing the performance of an isolated and coupled Long Short-Term Memory model architecture across 70 spaces in a case study building. To construct the coupled architecture, an open-source tool is developed and presented, which allows the automated extraction of space topology from IFC-files to identify adjacent spaces. The coupled architecture is found to outperform the isolated architecture for ~84% of the investigated spaces, with significant improvements under certain operational and climatic conditions. To account for the subset of spaces where the isolated architecture performs better, it is proposed to select between the two architectures accordingly. The demonstrated modularity and embedded adaptability of the proposed model architectures provide a sound basis for implementation in a highly dynamic building Digital Twin environment.

Key Innovations

- A novel and generic tool for automatic extraction of building topology from IFC files is presented and applied
- A thermally coupled MISO architecture for modular data-driven indoor temperature modeling and forecasting is presented and assessed

Practical Implications

The presented tool and architectures provides an automated, scalable, and accurate method for modeling and forecasting indoor temperature, enabling decision makers to evaluate how different operational strategies affect the indoor comfort levels of a building.

Introduction

Energy consumption in buildings is necessary to sustain proper air quality and adequate thermal comfort for occupants. However, the building sector is currently one of the largest energy-consuming sectors with major potential for improvements in energy efficiency and flexibility. With this as the driving force, buildings are currently seeing a fast-paced digitalization in the pursuit of Internet of Things (IoT)-integrated smart buildings (Jia et al. (2019)). In this context, data-driven modeling methods have received considerable attention for numerous applications, due to the promising potential for automation and scalability. The adaptability of data-driven approaches has been demonstrated through various applications, e.g. in forecasting of indoor air quality (Wei et al. (2019)), indoor temperature (Alawadi et al. (2020)), building occupancy (Jin et al. (2021)), energy loads (Amasyali and El-Gohary (2018)) as well as for fault detection and diagnosis (FDD) (Mirnaghi and Haghighat (2020)).

Indoor temperature is considered one of the key indicators of indoor comfort and accurate prediction models are therefore crucial for exploring different operational strategies while maintaining proper indoor comfort. Recently, black-box modeling and especially Artificial Neural networks have gained increased attention for indoor temperature modeling with recent studies suggesting a potential performance gain in considering the thermal coupling between spaces in buildings using large monolithic Multiple Input Multiple Output (MIMO) architectures (Mtibaa et al. (2020); Fang et al. (2021)). However, the performance of such architectures varies greatly with the specific topology and operational patterns of the modeled building and can in some cases be worse than the simpler isolated Multiple Input Single Output (MISO) model architectures (Mtibaa et al. (2020)). Furthermore, relying on one large black-box model for modeling the whole building makes it difficult to understand and rectify the model, in case of poor model performance. Therefore, we propose a different MISO architecture, where each space is modeled individually and only locally relevant thermal interaction between spaces is considered.

The benefits of this are twofold: 1) Each space is modeled separately, adding much greater modularity

and flexibility, compared to the mentioned monolithic MIMO architecture. For instance, using the proposed MISO architectures, some spaces can be modeled as isolated while other spaces with non-negligible heat transfer to adjacent spaces can be modeled using a coupled architecture. 2) More training data is available due to the lower amount of inputs per model, which can potentially improve performance (this is further discussed in section).

For the proposed coupled architecture to be feasible for practical implementation, an automated tool, Ifc2Graph (Bjørnskov (2022)), has been developed. Based on the open Industry Foundation Classes (IFC) standard, the tool analyses the building geometry to identify adjacent spaces. The two architectures are tested on 70 spaces, highlighting the performance of the two architectures in a realistic building setting. The considered spaces are occupied and include radiators for space heating, demand-controlled ventilation, and automated shading.

The presented work is in line with the ongoing digitalization of the building sector and the need for smarter buildings. The modularity and flexibility of the proposed model architectures ensure seamless integration with future building Digital Twins, where accurate and flexible indoor temperature models play a crucial role in safely exploring and recommending optimal building operation strategies without compromising indoor comfort. The presented work is carried out as part of the international research project 'Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle', aiming to design, develop, and demonstrate a holistic Digital Twin platform allowing coordinated decision support over the whole building life cycle. The proposed Digital Twin will provide the three following services: i) Flexible and effective data collection and management employing open standard context information models; ii) Smart facility management through performance monitoring and automated commissioning; and iii) Planning support through model-based informed decision-making.

Case study

In this work, a Danish university building is considered as a case study for testing the two model architectures. The building was commissioned in 2015 and was one of the first buildings in Denmark to comply with the Danish building class 2020. It has four stories, including a basement with technical rooms, storage, and installations. The ground floor, first floor, and second floor are mainly devoted to teaching and are composed of classrooms, offices, corridors, and study zones. An image of the building is shown in Figure 1.

The indoor climate is regulated through four balanced mechanical ventilation systems, each with a capac-



Figure 1: The considered case study university building (Jradi et al. (2017)).

ity of 35,000 m³/h, while the heating source of the building is district heating, supplying air-loop heating coils as well as radiators distributed across all four stories. Most spaces are well-instrumented with temperature, CO₂, humidity, PIR, and lux sensors. Furthermore, all radiators are equipped with thermostatic valves that control the indoor temperature by actuating the valve position. To ensure a proper indoor climate, the BMS employs Demand Controlled Ventilation (DCV) to regulate the airflows supplied to the different zones in the building. Specifically, for each space, the damper positions are actuated to prevent the measured CO₂-concentration from exceeding 600 ppm. To avoid overheating in the summer months, all outside-facing windows are equipped with shades that are controlled by the BMS. With the large scheme of sensors and meters instrumentation, the building is used as a living lab to carry out investigations and testing of various innovative applications and research-based solutions.

Local weather data on hourly resolution in the form of outdoor temperature, longwave solar irradiation, and shortwave solar irradiation is available for the site in hourly resolution. In Table 1, an overview of the relevant data is given including a shorthand notation, which will be used in the following sections. For the 8 types of data, we distinguish between data collected at space-level and data collected at building level.

Table 1: Overview of relevant data collected from the building.

Notation	Type	Space level (10 min)	Building level (1 hour)
T_z	Indoor temperature	X	
C'_z	CO ₂ -concentration	X	
u_v	Radiator valve position	X	
u_d	Air diffusers damper position	X	
u_{sh}	Shades position	X	
T_o	Outdoor temperature		X
Φ_L	Longwave solar irradiation		X
Φ_S	Shortwave solar irradiation		X

Methodology

In this section, the methodology used in this work is presented. Overall, two model architectures are investigated for the prediction of indoor temperature. First, the model architectures and chosen inputs for each of the model types are presented. Following, it

is explained how the overall space topology can be automatically extracted from a Building Information Model (BIM). It is then described how the collected data has been preprocessed and how the resulting two types of data sets differ. Finally, the training and testing methodology is presented.

Model architecture

In this work, a specific sub-category of ANNs called Recurrent Neural Networks (RNN) is used. The fundamental property of RNNs, that distinguishes them from other ANN variants such as the Multi-Layer Perceptron (MLP), is the inclusion of a system state that changes over time. This additional notion of time makes RNNs especially useful for the modeling of dynamic systems. In fact, the mathematical form of vanilla RNNs is essentially a discretized form of Delay Differential Equations (DDE) (Sherstinsky (2020)). The Long Short-Term Memory (LSTM) model, developed by Hochreiter and Schmidhuber (1997), is a specific type of RNN, mainly designed to solve the *Vanishing gradient problem* that was encountered with the vanilla RNN model. The main contribution of the LSTM model is the addition of gated units that control how information is added or removed from the system state. For the LSTM model, the state is described by two state vectors, the cell state $c \in \mathbb{R}^n$ and the hidden state $h \in [-1, 1]^n$, where n is a hyperparameter that describes the size of the state vectors, i.e. how much information that can be transferred between time steps.

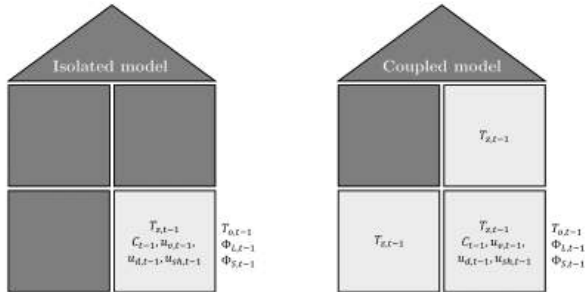


Figure 2: Inputs for the isolated and coupled architectures. The coupled architecture includes indoor temperatures from adjacent spaces as input.

As already mentioned, two types of models are investigated in this work. These models will in the remaining sections be referred to as the *isolated model* and the *coupled model*. The isolated model is based on previous work of the authors (Bjørnskov et al. (2022)), where the model inputs were determined based on well-known heat transfer mechanisms, including radiation, conduction through external surfaces, internal heat gains, ventilation, and heat added by radiators. The inputs of the coupled model consider the same heat transfer mechanisms. However, in addition to that, it also considers heat transfer through the non-external bounding surfaces, by in-

cluding the indoor temperature of the adjacent spaces as input. An overview of the two model types is shown in Figure 2. As seen, the only difference between the two model types is the inclusion of adjacent space temperatures as input in the coupled model. The network architecture of the two investigated models is seen in Figure 3. Both the isolated model and the coupled model consist of two sequential LSTM networks, A and B. Network A receives the inputs shown in Figure 2, i.e. it receives one set of inputs for the isolated model and another set of inputs for the coupled model. In addition, Network A receives as input the state vectors from the previous time step $c_{A,t-1}$ and $h_{A,t-1}$ and outputs the updated state vectors $c_{A,t}$ and $h_{A,t}$. The size n of the state vectors c_A and h_A will be referred to as n_A .

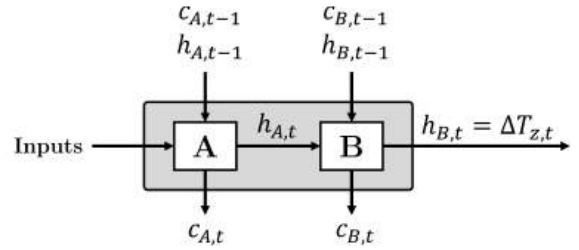


Figure 3: Model architecture for both the isolated and coupled model consisting of the sequential LSTM networks A and B. The model output is the predicted temperature change of the modeled space during a given time step.

Network B receives as input the updated hidden state $h_{A,t}$ from network A along with its own state vectors $c_{B,t-1}$ and $h_{B,t-1}$ from the previous time step. Network B outputs the updated state vectors $c_{B,t}$ and $h_{B,t}$, both with size $n = 1$. $h_{B,t}$ is treated as the model output which is fitted to the chosen prediction target during training. As seen in Figure 3, the chosen prediction target of the model is the temperature change of the modeled space during a given time step $\Delta T_{z,t}$. Through earlier investigations by the authors, it was found that this configuration (rather than directly predicting temperature) generalized better for custom control scenarios, e.g. when the model is implemented in a setpoint control loop (Bjørnskov et al. (2022)). The use of this configuration for indoor temperature predictions is later explained in section .

Automated extraction of space topology

For a coupled model to be feasible in larger buildings, an efficient and automated method of identifying adjacent spaces is required. With the significant adoption rate of Building Information Modeling (BIM), automatic extraction of geometrical building data for the use in Building Energy Modeling (BEM) tools is an active research area (BIM2BEM) and various algorithms have been developed for this purpose (Treeck and Rank (2006); Rose and Bazjanac (2013); Jones et al. (2013); Lilis et al. (2016)). Traditional

BEM tools require the paring of *space boundaries*, each associated with thermal heat transfer between two spaces. However, such detailed surface paring is not necessary for the application presented in this work, where the topology of adjacent spaces is sufficient. Furthermore, no accessible open-source tools have been identified in the literature that can fulfill this purpose. Therefore, the authors have chosen to develop such an algorithm based on the Industry Foundation Classes (IFC) standard (ISO (2018)), one of the leading BIM standards (Edirisinghe and London (2015)).

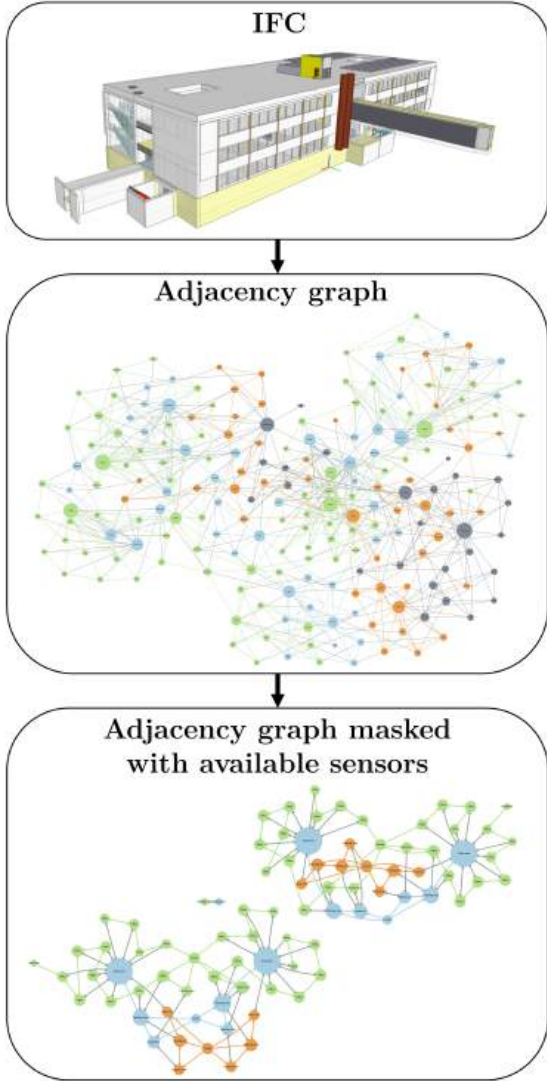


Figure 4: IFC file of the case study building (top), the derived graph that represents the adjacency of spaces (middle), and the adjacency graph masked with available sensors (bottom). The node colors represent the story levels with grey, orange, blue, and green representing the basement, ground, first floor, and second floor, respectively. The node size is scaled with respect to the number of connected edges, i.e. the number of adjacent spaces.

Generally, the IFC format offers detailed information

on building geometry and design data. Typically, all geometrical information relevant for modeling buildings is grouped under **IfcSpace** entities. The geometry of the building can thus be extracted on a space-by-space level and used to determine which spaces share boundaries. For this purpose, a python package called **Ifc2Graph** has been developed and made available as an open-source tool on GitHub (Bjørnskov (2022)). The developed tool has been tested on multiple IFC files to ensure that it generalizes beyond the case study considered in this work.

An image of the IFC file containing the architecture of the case study building is seen in Figure 4 (top). Using this IFC as input for the algorithm, yields the graph shown in Figure 4 (middle). Here, nodes represent spaces, one for each **IfcSpace** entity in the IFC file, while edges represent adjacency of the two connected spaces. The node colors show the story levels with grey, orange, blue, and green representing the basement, ground, first, and second floor, respectively. The size of each node has been scaled in correspondence to the number of connected edges, i.e. the number of adjacent spaces.

In this work, we concentrate on spaces with the energy balance described in section . Therefore, only a subset of the shown nodes with all five space-level data types described in Table 1 is selected for further modeling. This yields the masked adjacency graph in Figure 4 (bottom). The masked graph is obtained by removing all nodes from the adjacency graph that does not have all five space-level sensor types described in Table 1. Spaces with temperature measurements that are adjacent to one or more nodes fulfilling the above criteria are also kept in the graph. Hence, the nodes in this graph represent spaces that are either modeled or are supplying indoor temperature data to modeled spaces.

The case study building is almost symmetrical, meaning that comparing the north-south and west-east sections of the building yields an almost identical number of spaces and space types in each section. As seen, this is reflected in the masked adjacency graph, which consists of two large clusters that are almost identical in structure. The clusters are not connected because the central hallway of the building, dividing the building into its north and south sections has no installed temperature sensors. A small third cluster is also shown, consisting of two spaces. Again, none of these two rooms are adjacent to any space with installed temperature sensors.

Data preprocessing

To compare the performance of the isolated space models and the coupled space models, two types of data sets have been constructed, each with the inputs described in section . For each space, the inputs have been sampled on a 10 minutes interval in 24-hour sequences with 144 time steps, based on actual

sensor measurements from the building during a two-year period from January 1st 2018 to December 31st 2019.

For the coupled architecture, there can be certain drawbacks to introducing additional inputs to a given model. For instance, consider Figure 5 where time steps for three inputs and one output are shown during a certain period after preprocessing of the data. Green represents the availability of data and red represents the absence of data for each respective time step. The absence of data could be due to many reasons, e.g. missing values, removed outliers, frozen sensor readings etc. Note that for the actual data set, linear interpolation is performed for gaps of 12 time steps or less. The gaps of 1 to 3 time steps in the example would therefore not be present after preprocessing.

To find which time steps can be used for constructing training sequences, the intersection of data available across all inputs and outputs is found. Using this intersection, unique sequences with the desired step length can then be identified. For instance, if the sequence length is 3, then a data set of 8 sequences can be constructed for the given example. Therefore, introducing additional model inputs can potentially lower the number of training sequences, e.g. by adding adjacent space temperatures as model inputs.

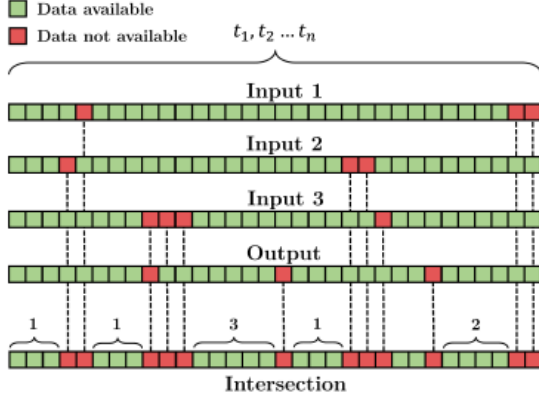


Figure 5: Example of intersection across three inputs and one output to find suitable time steps in the collected data for the construction of training, validation, and testing sequences.

The distribution of data sequences for both data sets after preprocessing across the different seasons is shown in Figure 6. Here, the sequence count across all spaces has been sorted for each season and data set, as represented by the colored area. Furthermore, these distributions have been summarized by the shown box plots. The sequence count varies notably between seasons. As seen, the inclusion of adjacent spaces in the model inputs lowers the sequence count across all seasons with the winter season having the most significant data loss. For the coupled datasets, the se-

quence count distribution is effectively shifted down for all seasons.

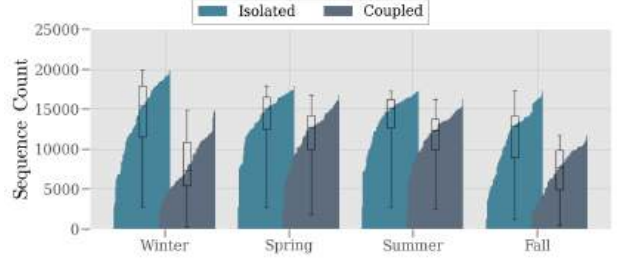


Figure 6: Sequence count per space distributed over the seasons for the two preprocessed data sets.

Training and testing method

For each space, the two data sets described in section , have been split into training, validation, and testing data sets with the splits $[1/3, 1/6, 1/6]$, respectively. To ensure that seasonality and varying building operation is reflected properly in each data set, data from each month has been distributed in proportion to the mentioned data splits in the training, validation, and testing data sets. In this process, it has been ensured that sequences between training and test data sets do not overlap to avoid data leakage.

All models have been trained using the python library Pytorch and Stochastic Gradient Descent (SDG) as optimizer with the Mean Square Error (MSE) loss function. This work considers a one-fit-all strategy where the same set of hyperparameters is used across all models. Although hyperparameter tuning performed on a model-by-model basis would likely improve performance, it would also increase computational demands and training time considerably. Through initial testing, a batch size of 32, a learning rate of 0.1, a momentum of 0.9, and a state size $n_A = 20$ was found to give satisfactory performance across all spaces. All models have been trained using an early-stopping technique to avoid overfitting. This is achieved by monitoring the loss on the validation data set while saving copies of the model instances at specific gradient step intervals during training. When the performance stops to improve or starts to degrade, the model instance with the lowest recorded validation loss is kept while the rest is discarded.

In earlier studies, the performance of indoor temperature models is often evaluated only for one-step-ahead predictions (Marvuglia et al. (2014)). However, if the step size is small, e.g. between 1-10 minutes, the obtained performance metrics are not sufficient to assess the model performance for longer forecasting horizons such as hours or days, due to the large thermal inertia and slow-moving dynamics of buildings. Therefore, to reflect model performance across longer horizons, this work considers a forecasting horizon of 24 hours. As explained in section , the models predict temperature change instead of the actual space temperature

on a 10-minute step interval. Therefore, to measure the model performance in terms of temperature prediction, the model is implemented in a closed-loop configuration as shown in Figure 7.

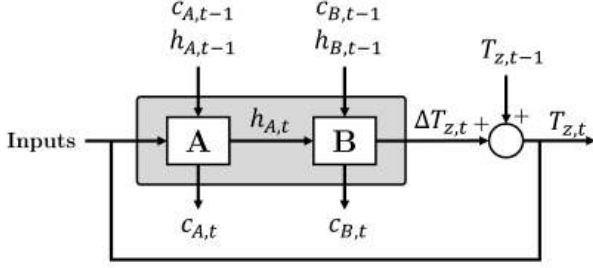


Figure 7: Closed-loop configuration for indoor temperature forecasting.

In Figure 7, the current indoor temperature change $\Delta T_{z,t}$ predicted by the model is added with the previous temperature value $T_{z,t-1}$ to obtain the current indoor temperature $T_{z,t}$. This newly obtained temperature is then fed back as input to the model, thus creating a loop. This process can be repeated for an arbitrary number of steps, as long as the defined inputs are available.

Results and Discussion

To assess and compare the performance of the isolated and coupled architectures, the closed loop configuration described in section is employed on the constructed test data sets that contain 24-hour sequences of 144 steps for all four seasons. For each of these sequences, predictions are thus obtained for the indoor temperature, which can be compared to the actual measured temperature. For this purpose, the Mean Absolute Error (MAE) is used as a metric to evaluate the performance of the trained models. Furthermore, to see how the prediction accuracy varies with different operational and ambient conditions, the MAE has been calculated on a seasonal basis for all models. The results are shown in Figure 8 where the colored areas show the sorted seasonal MAE across all spaces for both model architectures and boxplots have been added to summarize these distributions.

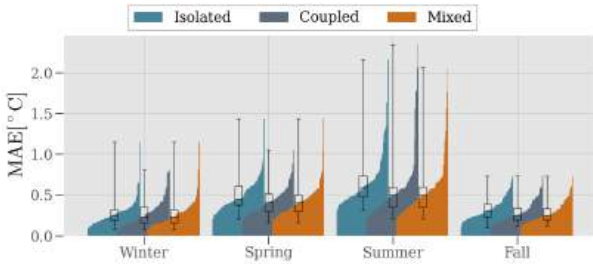


Figure 8: Mean Absolute Error per space model on the testing data set distributed over the seasons for the two model architectures.

As seen in Figure 8, the coupled model architecture yields better predictions for spring, summer, and fall with a lower median, first quartile, and third quartile. Especially for the summer season, a significant difference in performance can be observed between the two architectures. The building has no installed cooling equipment, and the building is thus mostly unconditioned during summer. For an unconditioned building, the energy balance of each space is typically dominated by heat transfer through the surrounding walls and windows, i.e. by temperature differences. Furthermore, since the indoor temperatures for this scenario are no longer forced to follow the specified setpoints, the temperature variance across the building is also greater. This explains the improved performance of the coupled model, where boundary temperatures are provided as inputs.

For the winter season, where the building is conditioned and the indoor temperature is mostly uniform, the isolated architecture performs similar to the coupled architecture with a higher first quartile, but a lower third quartile. This indicates that 25% of the best-performing isolated models perform worse than the equivalent fraction for the coupled models. However, it also follows that 25% of the worst-performing isolated models perform better than the equivalent fraction for the coupled models.

Contrary to the suggested monolithic MIMO architectures in recent studies (Mtibaa et al. (2020); Fang et al. (2021)), the presented MISO architectures allow for a modular modeling approach where the model architecture and performance can be calibrated on a space-by-space level. Spaces with strong thermal coupling to adjacent spaces can thus be modeled with the coupled architecture, while spaces having negligible thermal interaction with their adjacent spaces can be modeled using the isolated architecture.

This is demonstrated in Figure 8, where a *Mixed* category is shown. Here, either the isolated or coupled model has been selected, based on their performance on the validation data set, resulting in the selection of 11 isolated models and 59 coupled models. This slightly improves most aspects of the performance distribution across all seasons.

To validate the trained models on a qualitative basis, three of the coupled space models, representing a classroom, a study zone, and an office, have been chosen for further investigations. For this purpose, a 24-hour winter sequence from the test dataset is used as input with the closed-loop configuration described in section . The results are shown in Figure 9. Here, the individual inputs for each of the three models are shown in the six plots to the left while the common weather input is shown on the rightmost plot. As seen from the weather data, outdoor temperatures range between -5 to 5 °C, while irradiation levels are generally low. The three bottom plots on the left show

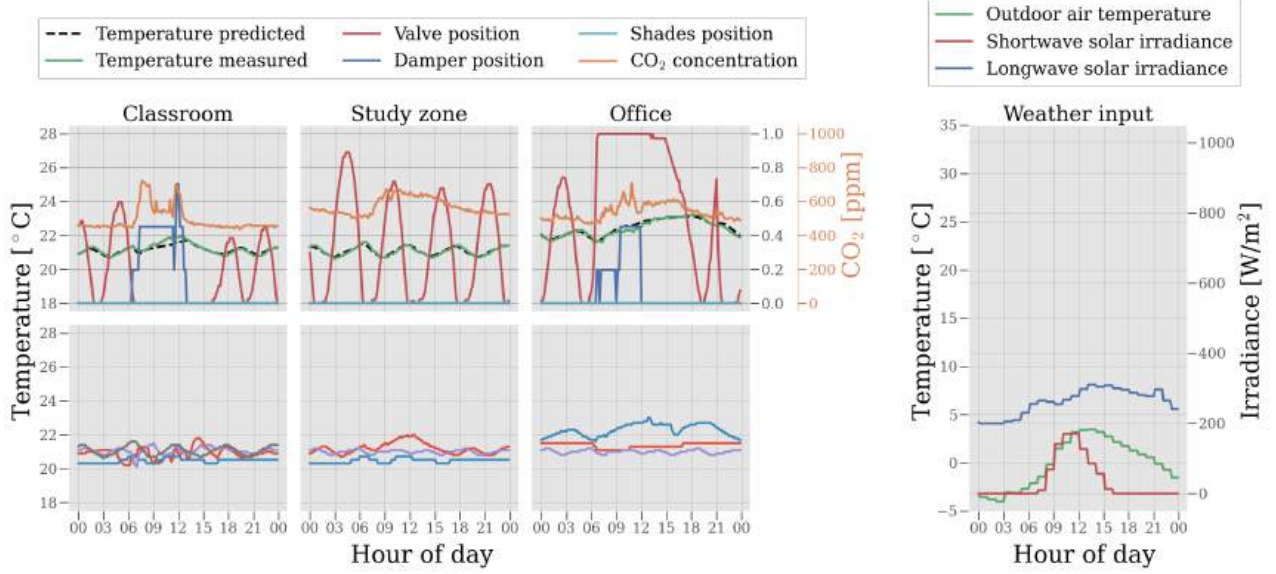


Figure 9: 24-hour forecasting for three spaces during a winter period using the coupled model architecture.

the input temperatures of the adjacent spaces, which are mostly within the range of 20 to 23 °C. The classroom includes four such inputs while the study zone and the office include three. The three top plots on the left show the remaining model inputs as well as the measured and predicted indoor temperature for each space. As seen, the model predictions agree very well with the measured temperatures throughout all 24 hours with only minor errors. During the winter period, the radiator valve position appears to be the most influential input with respect to the indoor temperature. To reach the specified setpoint temperature for each space, the valve position is opened when the measured temperature falls below the setpoint and is closed when the temperature surpasses the setpoint. For the classroom and the study zone, this results in a periodically fluctuating valve position and indoor temperature, which is well-captured by the coupled model architectures. Although the valve position of the space heater has the largest influence on indoor temperature, the impact of the dampers, regulating the airflows, can also be observed.

For the Classroom and the Office, the dampers are opened between hours 06 and 12 as a response to the increasing CO₂-concentration, which in periods exceeds 600 ppm. The isolated impact of increased air exchange on the indoor temperature is visible for the classroom, where the radiator valve is closed during this period. Here, the model predicts a slightly lower temperature increase compared with the actual measured temperature. At approximately 12, where the dampers are closed, the temperature starts to decrease until at approximately 16, when the valve position is again opened.

Conclusion

Previous studies have highlighted that the benefit of coupled MIMO models largely depends on the type of building and its operation and that the effectiveness of isolated MISO and coupled MIMO architectures must be evaluated on a building-by-building basis. In this work, it was investigated how thermal coupling between adjacent spaces affects the prediction accuracy of indoor temperature by comparing an isolated and a coupled LSTM MISO architecture across 70 spaces in a case study building. To construct the coupled models, a novel open-source tool Ifc2Graph was presented and used to automatically extract the space topology from an IFC-file of the case study building. For each space and architecture, data from various sensor and meter types across a two-year period was used to train and test the developed models. It was found that the coupled architecture improved overall performance for majority of spaces (~84 %), with significant improvements during the summer season, where the case study building was mostly unconditioned.

The modularity of the presented architectures allows the modeling of spaces with strong thermal coupling to adjacent spaces using the coupled architecture, while spaces with insignificant heat transfer to adjacent spaces can be modeled using the isolated architecture. In the broader context of future smart buildings and building digital twins, this added modularity will allow for increased adaptability to different building types, climate conditions, and data availability, while ensuring that accurate indoor temperature predictions are obtained. With indoor temperature being a key-indicator for the comfort-levels of occupants in buildings, the presented approach can be used to obtain accurate indoor temperature models,

that can help assess the feasibility of different operational strategies in buildings. For such an application to be reliable, future studies should investigate how the uncertainty of each prediction can be quantified and how well the models generalize to unseen operational conditions. Furthermore, the proposed method should be tested in environments with varying quantities of data to assess how this affects the performance, e.g. to simulate implementation in a newly built building.

Acknowledgment

This work is carried out under the ‘Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle’ project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64021-1009.

References

- Alawadi, S., D. Mera, Fernández-Delgado, F. Alkhabbas, Carl, M. Olsson, P. Davidsson, and S. Alawadi (2020, 01). A comparison of machine learning algorithms for forecasting indoor temperature in smart buildings. *Energy Systems*.
- Amasyali, K. and N. M. El-Gohary (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews* 81, 1192–1205.
- Bjørnskov, J. (2022). Ifc2graph. <https://github.com/JBjoernskov/Ifc2Graph>, Last Accessed: 06-17-2022.
- Bjørnskov, J., M. Jradi, and C. Veje (2022, 07). A fully automated and scalable approach for indoor temperature forecasting in buildings using artificial neural networks. In *Building Simulation Applications BSA 2022*.
- Edirisinghe, R. and K. London (2015, 10). Comparative analysis of international and national level bim standardization efforts and bim adoption.
- Fang, Z., N. Crimier, L. Scanu, A. Midelet, A. Alyafi, and B. Delinchant (2021). Multi-zone indoor temperature prediction with lstm-based sequence to sequence model. *Energy and Buildings* 245, 111053.
- Hochreiter, S. and J. Schmidhuber (1997). Long short-term memory. *Neural Computation* 9, 1735–1780.
- ISO (2018). International organization for standardization, "iso 16739-1:2018". <https://www.iso.org/standard/70303.html>, Last Accessed: 06-07-2022.
- Jia, M., A. Komeily, Y. Wang, and R. S. Srinivasan (2019). Adopting internet of things for the development of smart buildings: A review of enabling technologies and applications. *Automation in Construction* 101, 111–126.
- Jin, Y., D. Yan, A. Chong, B. Dong, and J. An (2021). Building occupancy forecasting: A systematical and critical review. *Energy and Buildings* 251, 111345.
- Jones, N., C. McCrone, B. Walter, K. Pratt, and D. Greenberg (2013, 08). Automated translation and thermal zoning of digital building models for energy analysis. pp. 202–209.
- Jradi, M., F. C. Sangogboye, C. G. Mattera, M. B. Kjærgaard, C. Veje, and B. N. Jørgensen (2017). A world class energy efficient university building by danish 2020 standards. *Energy Procedia* 132, 21–26. 11th Nordic Symposium on Building Physics, NSB2017, 11-14 June 2017, Trondheim, Norway.
- Lilis, G., G. Giannakis, and D. Rovas (2016, 12). Automatic generation of second-level space boundary topology from ifc geometry inputs. *Automation in Construction* 76.
- Marvuglia, A., A. Messineo, and G. Nicolosi (2014). Coupling a neural network temperature predictor and a fuzzy logic controller to perform thermal comfort regulation in an office building. *Building and Environment* 72, 287–299.
- Mirnaghi, M. S. and F. Haghighat (2020). Fault detection and diagnosis of large-scale hvac systems in buildings using data-driven methods: A comprehensive review. *Energy and Buildings* 229, 110492.
- Mtibaa, F., K.-K. Nguyen, M. Azam, A. Papachristou, J.-S. Venne, and M. Cheriet (2020, 12). Lstm-based indoor air temperature prediction framework for hvac systems in smart buildings. *Neural Computing and Applications* 32.
- Rose, C. and V. Bazjanac (2013, 04). An algorithm to generate space boundaries for building energy simulation. *Engineering with Computers* 31.
- Sherstinsky, A. (2020, 03). Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network. *Physica D: Nonlinear Phenomena* 404, 132306.
- Treeck, C. and E. Rank (2006, 01). Dimensional reduction of 3d building models using graph theory and its application in building energy simulation. *Engineering with Computers An International Journal for Simulation-Based Engineering* 23, 109–122.
- Wei, W., O. Ramalho, L. Malingre, S. Sivanantham, J. Little, and C. Mandin (2019, 06). Machine learning and statistical models for predicting indoor air quality. *Indoor Air* 29.

A Fully Automated and Scalable Approach for Indoor Temperature Forecasting in Buildings Using Artificial Neural Networks

Jakob Bjørnskov – University of Southern Denmark, Denmark – jabj@mmmi.sdu.dk

Muhyiddine Jradi – University of Southern Denmark, Denmark – mjr@mmmi.sdu.dk

Christian Veje – University of Southern Denmark, Denmark – veje@mmmi.sdu.dk

Abstract

Improving the performance of buildings is a core pillar to attaining future energy and environmental goals in different countries, considering that the building sector is a major contributor in terms of both energy consumption and carbon emissions. These ambitious goals and the call for smarter, energy-efficient, and flexible buildings have called for innovative and scalable energy and indoor thermal comfort modeling and prediction approaches. This work presents a fully automated and scalable solution using Artificial Neural Networks to forecast indoor room temperatures in buildings. A case study of an 8500 m² university building in Denmark was considered for testing and evaluating the proposed approach. An extensive dataset was constructed with sensor data from 76 rooms that contain both readings on indoor temperature, CO₂ concentrations, and actuating signals on radiator valves and dampers, as well as outdoor ambient conditions. Using this dataset, a well-performing architecture is identified, which provides accurate temperature predictions in the various rooms of the building for prediction horizons of 24 hours.

1. Introduction

Buildings are widely regarded as one of the major contributing sectors in terms of both energy consumption and CO₂ emissions. Furthermore, future energy systems with high fractions of Renewable Energy Sources (RES) depend on high demand-side flexibility. Therefore, there is a clear need for increasing not only the performance but also the flexibility of buildings. However, to achieve feasible and intelligent operation strategies for both cost minimization and flexibility services implementation without compromising the indoor

comfort levels of buildings, reliable and accurate forecasting of building indoor thermal behavior is vital. In terms of indoor temperature forecasting, Artificial Neural Network (ANN) models have shown great potential in capturing the dynamics with high prediction accuracy (Alawadi, et al., 2022). In addition, these models can also be easily adapted and scaled up to different building cases. This work presents an ANN-based approach that requires no prior specifications for the modeled building and can achieve accurate indoor temperature predictions for long prediction horizons of 24 hours or more. The models developed generalize well enough to be used for scenario planning and what-if analyses, e.g., to test the impact of custom setpoint and shading schedules on indoor temperature.

2. Case Definition

The building under consideration in this work is an 8500 m² highly energy-efficient university building from 2015. It is located in Denmark, and it primarily consists of space types such as classrooms, study zones, corridors, and offices. In each of these spaces, indoor air temperature T and CO₂ concentration C are measured through installed sensors. In addition, each space contains space heaters of specific capacities with equipped mechanical valves that control the water massflow. The position of these valves $u_v \in [0,1]$ is managed centrally by the Building Management System (BMS) with $u_v = 0$ meaning fully closed with no massflow and $u_v = 1$ meaning fully open with maximum massflow. The supply water temperature is kept constant at approximately 60 °C.

The building is also equipped with a weather station that measures outdoor air temperature T_o , longwave solar irradiance Φ_L , shortwave solar irradiance Φ_s , and wind speed. The supply air temperature set-point is constant, at either 21-22 °C, depending on each space. The supply and exhaust airflows are controlled in each space by the supply and exhaust damper positions $u_d \in [0,1]$, with $u_d = 0$ meaning fully closed with no airflow, and $u_d = 1$ meaning fully open with maximum airflow. These damper positions are also managed through the BMS with Demand Controlled Ventilation (DCV), aiming at keeping the measured CO₂ concentration below 600 ppm. Finally, each space is also equipped with shades that are controlled by the BMS through a position parameter $u_{sh} \in [0,1]$, with $u_{sh} = 0$ meaning fully exposed with no shading, and $u_{sh} = 1$ meaning fully enclosed with maximum shading. The shades are controlled based on outdoor and indoor illuminance with a safety roll-up mechanism that sets $u_{sh} = 0$ at wind speeds higher than 15 m/s.

3. Methodology

3.1 Model Architecture

This work makes use of a specific type of ANNs called Recurrent Neural Networks (RNN) to model the transient temperature dynamics of a room. Specifically, Long Short-Term Memory (LSTM) networks were chosen due to their numerous demonstrations of adaptability and robustness in time-series black-box modeling, including indoor environment modeling (Fang et al., 2021; Mtibaa et al., 2020). LSTM models are a specific kind of RNN that were originally developed to deal with the vanishing and exploding gradient problem of traditional RNN models. A detailed explanation of the LSTM model is provided in the references (Hochreiter & Schmidhuber, 1997; Van Houdt et al., 2020).

To properly account for all phenomena that can significantly influence the energy balance of the room, it is very important to choose appropriate dynamic inputs for the model. In this work, the inputs are determined by considering the following energy transfer mechanisms:

- *Heat transfer by conduction through external surfaces* is considered by including indoor air temperature T , and outdoor temperature T_o as input.
- *Heat transfer by radiation* is considered by including longwave solar irradiance Φ_L , shortwave solar irradiance Φ_s , and the position of the shades u_{sh} in the model input.
- *Internal heat gains through occupancy* are included indirectly by including measured CO₂ concentration C , and damper position u_d . These inputs can, to a certain extent, represent occupancy due to the direct correlation between CO₂ concentration, ventilation airflow, and occupancy presence (Franco & Leccese, 2020).
- *Heat added by the space heater* is considered by adding the measured valve position u_v as input, which represents the water massflow. The supply water temperature is constant and does not therefore contribute as input.
- *Heat transfer by ventilation* is considered by adding the supply and exhaust damper positions u_d as inputs, which represent the airflow rates. The supply air temperature is constant and therefore does not contribute as input.

The model architecture is seen in Fig. 1 with inputs and outputs of the model. As shown, the model consists of two sequential LSTM models, A and B. All previously mentioned weather and sensor inputs from the previous timestep are fed to LSTM A. In addition, the LSTM also receives the cell state vector $c_{A,t-1} \in \mathbb{R}^n$ and hidden state vector $h_{A,t-1} \in [-1,1]^n$, where n is a hyperparameter that determines the size of these vectors. These two state vectors are an integral part of LSTM models, which essentially dictate the state of the system modeled during a given timestep. LSTM A outputs $c_{A,t}$ and $h_{A,t}$, which represents the updated state vectors.

LSTM B has $n = 1$ and is only given three inputs, the cell state $c_{B,t-1} \in \mathbb{R}$, the hidden state $h_{B,t-1} \in [-1,1]$, and the hidden state vector $h_{A,t}$. LSTM B outputs $c_{B,t}$ and $h_{B,t}$, which represent the updated state vectors of LSTM B. The training task is then to find an optimal set of parameters in LSTM A and B to minimize the error between $h_{B,t}$ and the chosen prediction target over multiple sequences of data.

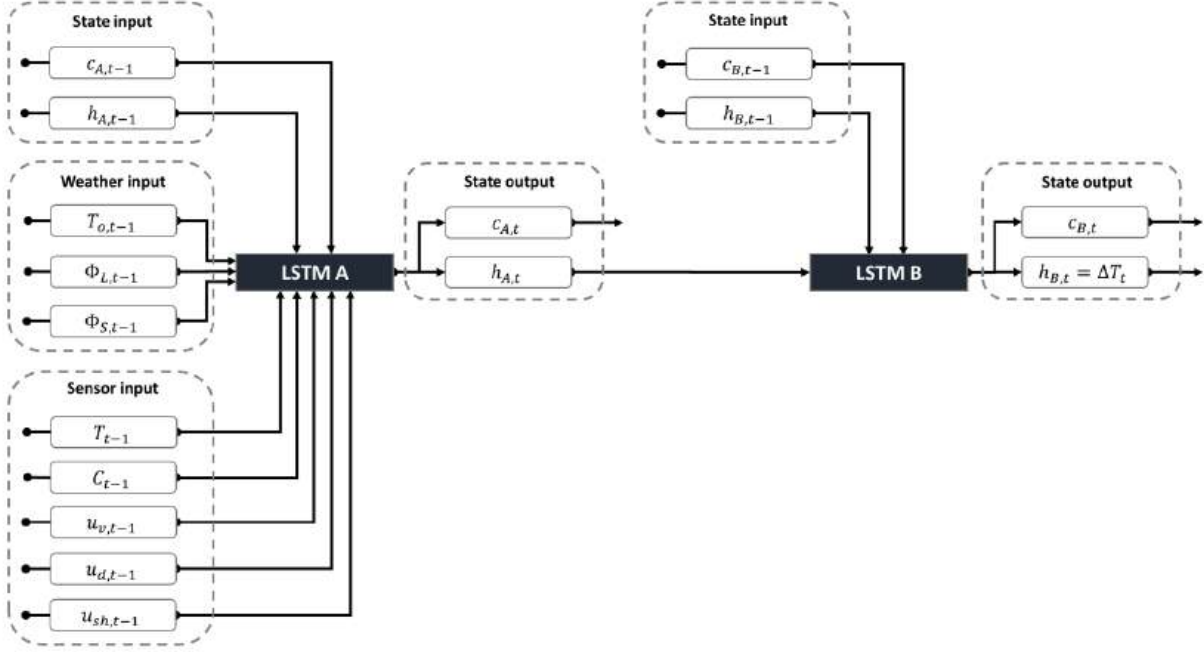


Fig. 1 – Model architecture with inputs and outputs of the sequential LSTM models, A and B

In the studies reviewed, the prediction target of the learning algorithms employed was indoor temperature in all cases. During initial testing, it was found that this configuration yielded good prediction results when testing with historical data as input. However, it was found that the models generalized poorly when fed with custom inputs instead of historical data, e.g., when using the model for setpoint control, as shown later in Section 4.2. When collecting operational data from a building, the actuation signals, e.g., valve positions, are typically directly correlated with indoor temperature through a thermostat with a simple control law, e.g., in the form of a Proportional (P) or Proportional Integral (PI) controller. Therefore, if the prediction target is temperature, the model will likely overfit the specific mode of operation that is reflected in the historical data used to train and test the model. We hypothesize that the model essentially learns to map the inverse control law of the thermostat instead of the actual thermal physics of the room. In this work, we are proposing that the model should predict the indoor temperature change ΔT instead of the actual temperature value, as this disrupts the direct correlation that is otherwise present between input and output.

3.2 Data Preprocessing

The dataset was constructed with all the weather and sensor readings introduced for 76 rooms in the case study building at a 10- minute interval for two years spanning January 1st 2018 to December 31st 2019. The raw data were pretreated and validated to ensure that proper and clean data were used. Following this, all inputs were min-max scaled between -1 and 1. After preparing the dataset, 24-hour sequences of 144 timesteps with no missing data were selected to form a collection of sequences for each space. In Fig. 2, the distribution of data sequences available among spaces is shown monthly. As shown, the number of sequences varies between rooms. The month with the most data is January, with a median of about 6000 sequences, while the month with the least data is August, with a median of about 3000 sequences.

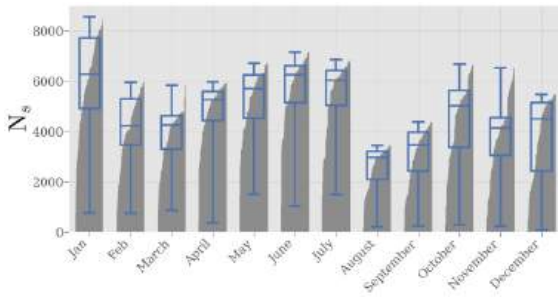


Fig. 2 – Number of sequences available per space for each month

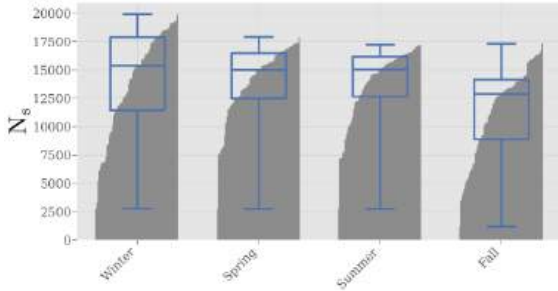


Fig. 3 – Number of sequences available per space for each season

In Fig. 3, the distribution is shown on a seasonal level. Here, the dataset generally appears to be more balanced with less variation. Hence, it is expected that the constructed datasets have enough diversity to cover most of the seasonal variance in the operation and thermal dynamics of the building.

3.3 Training and Testing Method

After preprocessing, the data is split into training, validation, and testing data sets with the splits [2/3, 1/6, 1/6], respectively. These data splits were carefully designed to ensure no overlap between sequences in the three data sets, while ensuring that seasonal and monthly variations are reflected in each dataset. For training, the machine-learning library Pytorch was used with Stochastic Gradient Descent (SDG) as optimizer, a momentum of 0.9, a batch size of 32, and a learning rate of 10^{-1} . The size n of cell state c_A and hidden state h_A was set to 20 for all space models.

During training, the model loss was evaluated on the validation dataset and saved at every 64th gradient update. At every 3000th gradient update, a copy of the model was saved, along with an average of the last 100 validation loss evaluations. After 100,000 gradient updates, the model with the lowest saved validation loss was selected. If the model had

not converged after 100,000 gradient updates, the procedure was repeated up until 400,000 gradient updates. This approach was used to train models for all 76 spaces. To properly test the trained space models, two modes of operation were presented.

The first mode was aimed at assessing the prediction accuracy of the space models given historic inputs. Here, the developed models were employed in a closed-loop configuration, as shown in Fig. 4, where future temperature predictions were based on past predictions. Perfect forecasting was assumed by feeding historical data for all weather and sensor inputs, except for the indoor temperature.

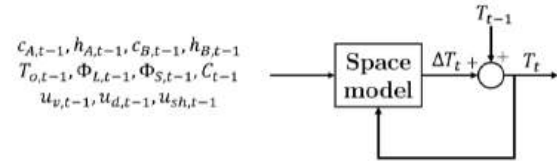


Fig. 4 – Closed-loop configuration designed to forecast indoor temperature for an arbitrary number of timesteps

The model then, for each timestep, predicted the temperature change to obtain the indoor temperature of the next time step, which was fed back to the model. This could be repeated as long as historical inputs were available. By repeating this for all timesteps in a simulation period, the produced temperature profile could then be compared with the actual measured temperature profile of the room to assess the prediction accuracy of the space model.

The second mode of operation was aimed at testing whether the developed space models generalize well enough to provide reasonable predictions under unseen operational conditions. Here, custom inputs were thus be fed to the model to observe the response. This was a very important property that made it possible to use the model for testing different operational strategies or what-if scenarios and their influence on indoor comfort in a safe environment.

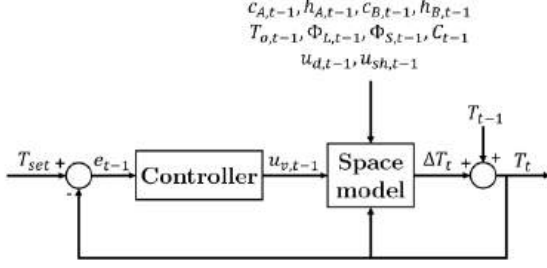


Fig. 5 – Closed-loop configuration designed for temperature set-point control through the valve position of the space heater

In this work, two model inputs were considered for this purpose; u_v and u_{sh} . In one simulation, the space heater valve position input u_v was constructed by implementing the space models in a closed-loop configuration for temperature setpoint control, as shown in Fig. 5. Here, all inputs were historical except for u_v and temperature T . The implemented controller was a simple proportional controller that each timestep scales the input signal $u_{v,t-1}$ proportionally to the error $e_{t-1} = T_{set} - T_{t-1}$, where T_{set} was the temperature setpoint in the room. In another simulation, the shades position input u_{sh} was constructed by implementing a simple predetermined schedule that operates based on the time of the day. All inputs except for u_{sh} and T were thus historical.

4. Results and Discussion

4.1 Quantitative Performance Assessment

First, the quantitative model performance was evaluated by using the first mode of operation as described in Section 3.3. Here, the Mean Absolute Error (MAE) between measured indoor temperature and predicted temperature was calculated for each space model across all 24-hour sequences in the test dataset. Fig. 6 shows the performance of each trained space model. Specifically, it shows the relationship between MAE, Standard Deviation of prediction targets $\sigma(\Delta T)$, and number of sequences N_s . The three marked space models were used for a qualitative performance assessment in Section 4.2. As seen in the figure, most of the space models (~86 %) achieve MAE values below 0.5 °C, which is lower than the required measurement accuracy of temperature sensors of ± 0.5 °C (ISO, 1998).

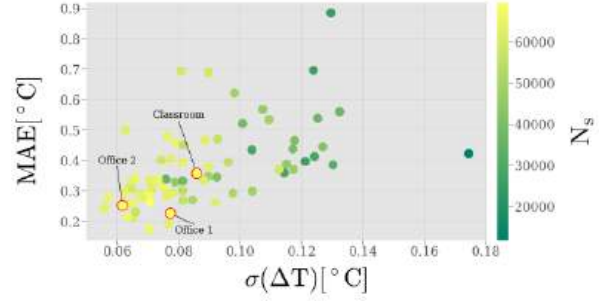


Fig. 6 – Relationship between Mean Absolute Error MAE, Standard Deviation of prediction targets $\sigma(\Delta T)$, and number of sequences N_s . The three space models chosen for qualitative assessment are marked

The best performing space model represents an office with $N_s = 66297$ and $MAE = 0.17$, while the worst performing model is of a classroom with $N_s = 24210$ and $MAE = 0.88$. As seen from the colormap in Fig. 6, there seems to be a negative correlation between the prediction error and the number of sequences available, agreeing with the general notion in machine-learning, that more data yields lower prediction error and better model generalization. Therefore, it is expected that the poor-performing space models could attain similar performance with more data. Furthermore, there seems to be a positive correlation between prediction error and the variation observed for the prediction target ΔT . This means that the prediction error will be higher for datasets that have a more fluctuating temperature profile. This is to be expected, as a fluctuating indoor temperature is generally harder to predict than a steady temperature.

4.2 Qualitative Performance Assessment And Applications

To provide a qualitative performance assessment of the models developed, three case study spaces were selected, one classroom and two offices, as also marked in Fig. 6. For each of these space models, a winter period and a summer period of 24 hours were chosen to evaluate how the space models perform under different ambient conditions. The models were first simulated for these periods using the first mode of operation, as explained in Section 3.3.

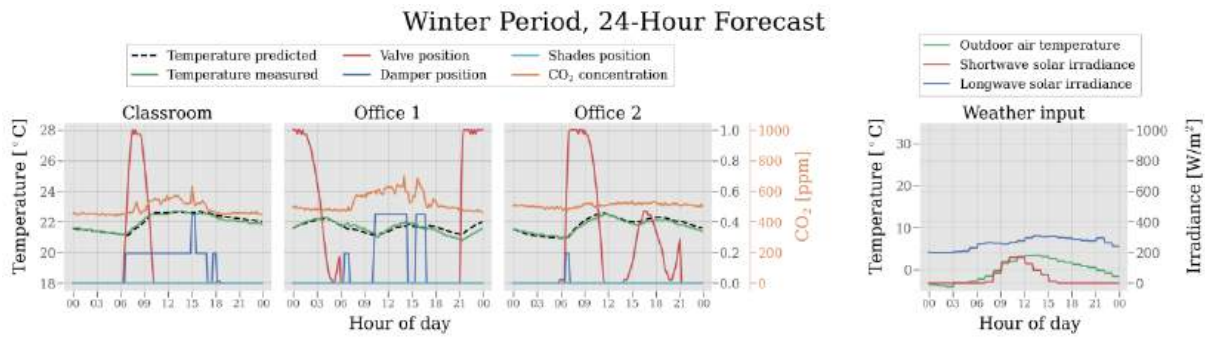


Fig. 7 – 24-hour temperature forecast in a winter month compared with actual measured temperature. The weather input is shown on the plot furthest to the right, while the individual inputs for the three selected space models are shown on the plots to the left

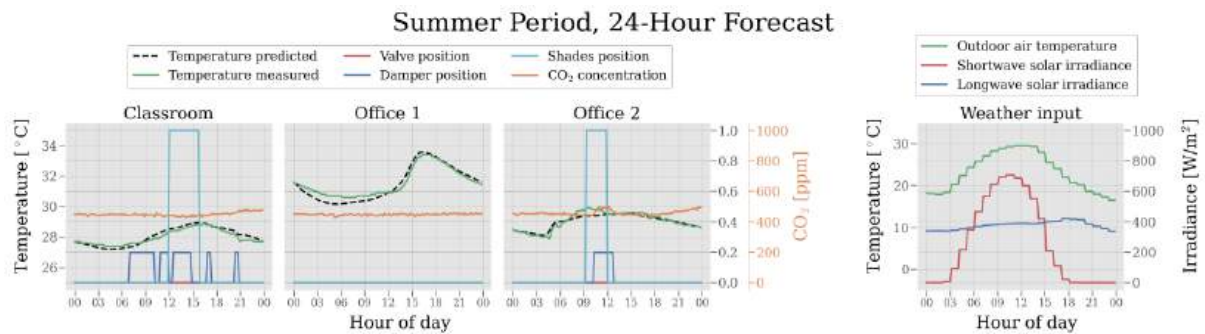


Fig. 8 – 24-hour temperature forecast in a summer month compared with actual measured temperature. The weather input is shown on the plot furthest to the right, while the individual inputs for the three selected space models are shown on the plots to the left

For the winter simulation, the results are shown in Fig. 7, with the weather inputs in the plot furthest to the right, while the results for each of the three chosen rooms can be seen on the left. As seen, all three space models accurately predict the indoor temperature, although Office 1 seems to slightly overestimate the temperature during the last 6 hours. The space heater valve position has a clear significance in all three spaces, where the general trend is that $u_v = 0$ results in decreasing room temperature, while $u_v > 0$ results in increasing room temperature. The models are also able to account for heat gains and heat losses associated with occupancy and ventilation. This is mostly seen in the Classroom and Office 1 in the period from 09:00 to 15:00, where the CO₂ concentration rises above 600 ppm and the dampers are positioned at around 50 %. Here, the model correctly predicts that the indoor temperature increases, although the space heater is not in operation. The shades are all rolled up ($u_{sh} = 0$) as they have no desirable effect during winter. Moving to the simulation results for the summer period, the results are shown in Fig. 8. The weather data inputs are again seen on the plot furthest to the

right, where the ambient temperature and irradiance levels are much higher compared with the winter period. As seen, this has a significant impact on the predicted and actual temperatures in the three spaces, especially for Office 1, where temperatures are above 30 °C during the whole 24 hour-period. The cause for large differences in both shape and peaks of the temperature profiles was different orientation, shading, and geometrical properties of the spaces. As seen, all three space models have learned to correctly account for these properties and provide accurate predictions for all 24 hours, although Office 1 slightly underestimates the temperature during the first 12 hours of the period. From the CO₂ levels, which have a very constant profile of around 450 ppm, the occupancy appears to be close to zero. However, this is expected during July and August, where the students are on summer leave. The shades are, to some extent, utilized in the Classroom and Office 2. However, it is expected that the high temperatures could be mitigated even more, by increasing the duration of the shades being rolled down ($u_{sh} = 1$). This will be investigated further in the following analysis.

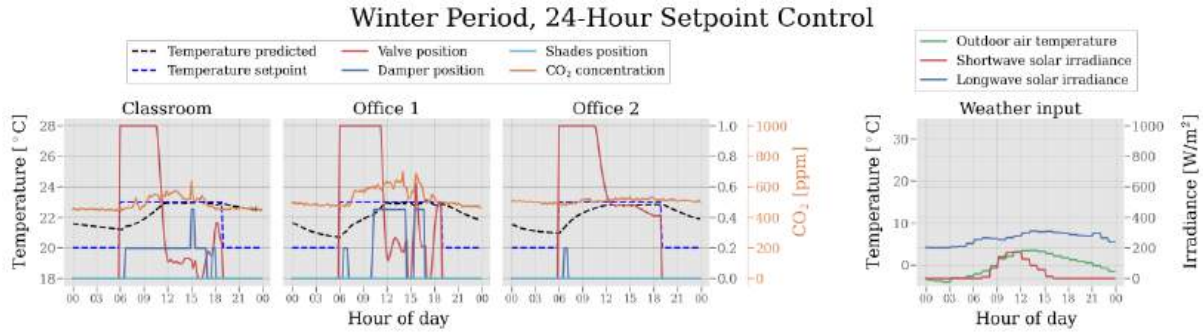


Fig. 9 – 24-hour setpoint control in a winter month. The weather input is shown on the plot furthest to the right, while the individual inputs for the three selected space models are shown on the plots to the left along with the predicted temperature and the setpoint

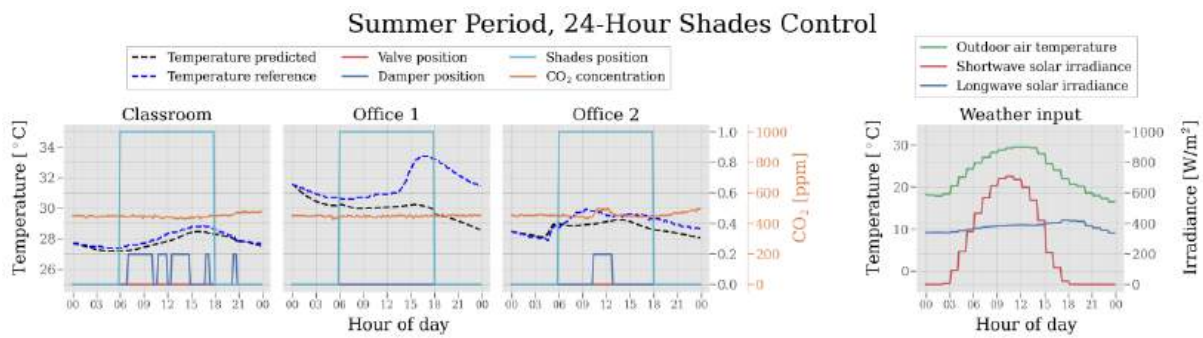


Fig. 10 – 24-hour shades control in a summer month. The weather input is shown on the plot furthest to the right, while the individual inputs for the three selected space models are shown on the plots to the left along with the predicted temperature. The original measured temperature in this period is shown as a reference

As just demonstrated, the models perform well when predicting temperatures under historical conditions. However, to demonstrate the applicability of the models, the second mode of operation, as presented in Section 3.3, was employed. The temperature setpoint control was implemented for the same winter period as shown in Fig. 7, while the shades control was implemented for the same summer period as shown in Fig. 8.

The results for the temperature setpoint control are shown in Fig. 9. As seen, the setpoint was varied, depending on the time of day. From 06:00 to 18:00, the setpoint was 23 °C, while from 18:00 to 06:00, the setpoint was 20 °C. The controller adapted to these setpoint signals by varying the valve position accordingly. For all three spaces, the valve was closed during the night, where the temperature was allowed to decrease. At 06:00, when the setpoint was raised to 23 °C, the valve was opened, and the indoor temperature increased until around 12:00, where the indoor temperature in all three rooms reached the specified setpoint. At this point, the

valve position was operated between 0 and 1 in an attempt to keep the indoor temperature at the setpoint. For Office 2, it was noticed that the temperature and valve position profile was smoother compared to the Classroom and Office 1, which had small fluctuations in temperature. However, as seen, the Classroom and Office 1 also had more disturbances in the form of varying CO₂ levels and ventilation airflows. Despite these disturbances, the simple controllers managed to keep the temperature at 23 °C in all three space models until 18:00, where the setpoint was again decreased to 20 °C. Here, the valve was again shut, and the temperature started to decrease.

Moving on to shades control during the summer period, the results are shown in Fig. 10. Here, the shades were rolled down ($u_{sh} = 1$) from 06:00 to 18:00 and rolled up ($u_{sh} = 0$) from 18:00 to 06:00. This had a significant effect on the predicted indoor temperature when compared with the original measured temperature (temperature reference), where the duration of shading was very limited. The

effect was very clear for Office 1, where the peak at 17:00 shifted from around 33 °C to 30 °C. The Classroom and Office 2 also had a reduced temperature response, although not as significant as Office 1.

5. Conclusion

In this work, a fully automated and scalable approach for temperature forecasting in buildings was presented and assessed. The presented method relied on ANNs in the form of two sequential LSTM models to predict temperature change within a room, given weather data and sensor inputs such as space heater valve positions, damper positions, shades positions, and CO₂ concentration. Hence, the developed approach needed no prior information about the building such as geometry, material properties, design data, etc.

The methodology developed was implemented considering 76 rooms of an 8500 m² university building in Denmark with 86 % of the rooms achieving a Mean Squared Error of less than 0.5 for 24-hour forecasting. The difference in prediction performance between space models was explained by differences in the amount of data available. However, more work is needed to identify more robust criteria concerning the amount, type, and quality of data that is needed to obtain accurate space models. The applicability of the models was demonstrated by implementing three selected models in a closed-loop setpoint control configuration for a 24-hour winter period. Additionally, different shading schedules were also explored to show their impact during a 24-hour summer period.

In line with the emerging initiatives toward digitalization of the building sector, building digital twins has promising technical and economic impacts. In this context, a fully scalable and automated energy modeling approach is vital, so that these twins can provide a robust, generic, and effective solution for various applications in the building sector. The modeling approach proposed in this study serves as a core for future building digital twin development and could be used as a backbone for various automated services, including performance monitoring, scenario assessment, and operational management.

Acknowledgement

This work was carried out under the ‘Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle’ project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64021-1009.

References

- Alawadi, S., D. Mera, M. Fernández-Delgado, F. Alkhabbas, C. M. Olsson, and P. Davidsson. 2022. “A comparison of machine learning algorithms for forecasting indoor temperature in smart buildings.” *Energy Systems* 13: 689–705. doi: <https://doi.org/10.1007/s12667-020-00376-x>
- Fang, Z., N. Crimier, L. Scanu, A. Midelet, A. Alyafi,, and B. Delinchant. 2021. “Multi-zone indoor temperature prediction with LSTM-based sequence to sequence model.” *Energy and Buildings* 245: 111053. doi: <https://doi.org/10.1016/j.enbuild.2021.111053>
- Franco, A., and F. Leccese. 2020. “Measurement of CO₂ concentration for occupancy estimation in educational buildings with energy efficiency purposes.” *Journal of Building Engineering* 32: 101714. doi: <https://doi.org/10.1016/j.jobee.2020.101714>
- Hochreiter, S., and J. Schmidhuber. 1997. “Long Short-term Memory.” *Neural computation* 9(8): 1735-1780.
- ISO. 1998. *ISO 7726 - Ergonomics of the Thermal Environment - Instruments for Measuring Physical Quantities*.
- Mtibaa, F., K.-K. Nguyen, M. Azam, A. Papachristou, J.-S. Venne, and M. Cheriet. 2020. “LSTM-based indoor air temperature prediction framework for HVAC systems in smart buildings.” *Neural Computing and Applications* 32: 17569–17585. doi: <https://doi.org/10.1007/s00521-020-04926-3>
- Van Houdt, G., C. Mosquera, and G. Nápoles. 2020. “A Review on the Long Short-Term Memory Model.” *Artificial Intelligence Review* 53(8): 5929-5955. doi: <https://doi.org/10.1007/s10462-020-09838-1>

A Digital Twin Platform for Energy Efficient and Smart Buildings Applications

Muhyiddine Jradi
Center for Energy Informatics
University of Southern Denmark
Odense, Denmark
mjr@mmmi.sdu.dk

Jakob Bjørnskov
Center for Energy Informatics
University of Southern Denmark
Odense, Denmark
jabj@mmmi.sdu.dk

Abstract—The employment of digital technologies is essential for creating new comfort levels, security, smartness, and efficiency in buildings. Buildings will shift from being a static, reactive element in the energy sector to being a proactive important component with the capacity to intelligently adapt to changes in the dynamic environment and interact efficiently with other structures and networks in smart communities and grids by possessing an array of sensors and meters everywhere, as well as a vast database of digital information. To capitalize on the benefits of building sector digitization and respond to rising demands for energy efficiency, comfort, and safety, we propose merging historical, real-world, and forecast data into a context information model to characterize and better comprehend the building. This is accomplished by digitally representing real-world assets as Digital Twins, enabling for integrated data-driven decision making across the building life cycle. The proposed holistic Digital Twin platform 'Twin4Build' substitutes the conventional static Building Information Model with a living deployable model that combines dynamic energy models and real-time data to provide a variety of operational services such as gathering and handling data, real-time performance monitoring and analysis, continuous commissioning, and strategic planning.

Keywords—buildings; digital twin; sensors and meters; building information modeling; data modeling

I. INTRODUCTION

Buildings in the EU consume nearly 40% of total energy consumption. In order to fulfill the future high energy and environmental targets set for 2030 and 2050, the EU has focused on the construction and buildings sector, highlighting the great potential for cost-effective efficiency gains as well as significant proportional reductions in the emission of greenhouse gases [1]. According to the International Energy Agency (IEA), the buildings and construction sector accounts for approximately 36% of total emissions: 22% for residences, 8% for non-residential buildings, and 6% for the construction projects and industries [2]. In Denmark, building standards, known as 'Danish BR' [3], have been regularly updated and tightened in past few decades, focusing mostly on enhancing the level of design and operation of newly constructed buildings and facilities. As a result, building design, operation, and management requirements have been significantly tightened, with concrete standards governing the building physical envelope, components, and specifications, in addition to indoor air quality and thermal comfort within the

built environment. Moreover, rules for the execution and accomplishment of variety of building services and energy supply systems have been established. Figure 1 depicts the progression in regard to the maximum allowable primary energy usage for a newly constructed 150 m² typical Danish residential house, taking into account the evolution and development of various building standards 'BR'. It could be highlighted that the permissible primary energy usage by a newly constructed building has fallen by 95% from roughly 350 kWh/m² per year for a building built in 1961 to just 20 kWh/m² for a building of the same type and size from 2020 [4].

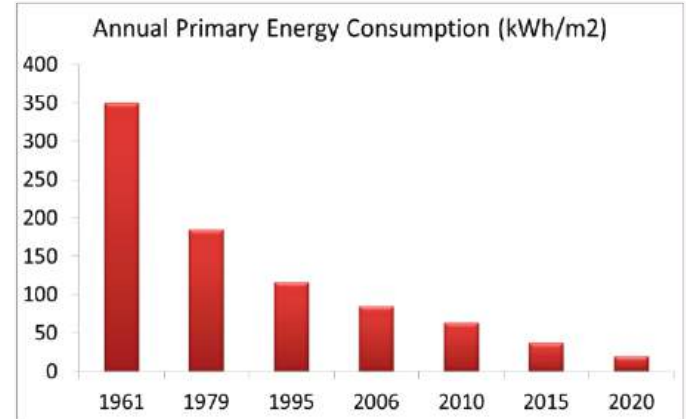


Figure 1. The evolution of the allowable primary energy consumption of a typical residential building in Denmark with the progression in the building regulations over the last 60 years [4].

Nowadays, buildings are anticipated to use and apply digital technology to achieve high comfort levels, smartness, and efficiency with the corresponding evolution in communication and information technologies, as well as the development and construction of smart and linked grids. Such facilities should provide occupants with information about present and future status, adjust operations to dynamic shifts in the energy grid, and allow strategic maintenance. Digital technologies are likely to play a significant role, resulting in smarter, more sustainable, efficient, dependable, and interconnected buildings on a wide scale around the world. The rapid advancements in data, analytics, and networking have resulted into a slew of new digital intelligent systems and devices. Digitalization and digital data are major drivers in this

respect, expanding rapidly in numerous parts of the building domain and construction business.

In this paper, an innovative digital twin (DT) platform for buildings is designed and presented. The building DT framework and design is described. The proposed DT replaces the traditional static Building Information Model with a living operational model. It combines adaptable dynamic energy models and real-time data to provide multiple operational services.

The work presented in this paper is carried out under the international research project 'Twin4Build: A holistic Digital Twin (DT) platform for decision-making support over the whole building life cycle,' which is funded by the Danish Energy Agency through the Energy Technology Development and Demonstration Program (EUDP) [5]. The project is a collaboration activity between research, industrial, and public partners, namely SDU Center for Energy Informatics at the University of Southern Denmark (SDU), COWI A/S, KMD A/S, Aarhus Municipality, NEC Laboratories Europe, NEC Laboratories India, Copenhagen Municipality and Helse Bergen Hospital in Norway.

The Twin4Build project seeks to design, develop, and demonstrate a comprehensive Digital Twin platform that enables for integrated and organized decision support across the building process life cycle. The platform substitutes the static Building Information Model with a dynamic, alive, and operational counterpart. The DT integrates building energy models that are flexible, adaptable, and scalable with real-time data information gathered by different metering and sensing devices. Overall, the proposed DT platform adds value throughout the entire value chain. As a result of implementing the provided solution:

- Property and building owners and clients will benefit from more energy-efficient facilities from the start and throughout the building's life cycle.
- Facility and property managers will have access to a complete platform that will enable smart operation and maintenance by automatically and continuously monitoring and tracking building operations, and mistakes and malfunctions on many levels will be highlighted in a timely way.
- Energy experts and planners will work together to promote successful decision making all across the entire life cycle, as the envisioned digital twin will aid in strategy and planning by guiding optimal decisions and selections and performing 'What if?' situations ahead.

II. BUILDING INFORMATION MODELING

Building Information Modeling (BIM) has to be on the very forefront when discussing the digitization of the building sector. BIM is among the most significant technology advances in the construction industry. BIM, in its broadest sense, is a virtual and digital depiction of any built environment's functional as well as physical characteristics. It is a set of procedures, technology, and rules that enable various stakeholders to plan, design, construct, and run a facility or building jointly and digitally. BIM design and concepts have

advanced dramatically in recent years, allowing for the differentiation of multiple BIM levels. This goes from BIM level 0 (no synchronization and just 2D CAD drawing) to 4D BIM (knowledge used for performance analysis), with '5D' permitting financial management and '6D' serving as the foundation for properties and facilities maintenance and management [6].

Overall, using BIM in the phases of design and construction has a significant beneficial impact, such as more reliable facility design, higher project control, decreased costs and time delivery, and improved collaboration across working groups. In present practice, BIM is used to meet the bare minimum of criteria and then discarded after the end of construction, with no value for facility management or performance optimization [7]. Yet, the owners will pay less than 20% of the facility's life cycle cost during the planning and building stages and more than 80% during the operating phase. This highlights the value of closely monitoring the building's performance, commissioning, and operation optimization.

Furthermore, as digital technologies evolve with sophisticated Internet of things devices and modern automation systems for buildings, there is an essential need to connect the enormous volume of building information and data with the present BIM framework [8]. As a result, there is the possibility for BIMs to be enhanced with information systems and apps that give interrogative objectives during operation, management, and maintenance. In this regard, it is crucial to extend and utilize BIM capabilities not just for proper and cost-effective design, but also for data-driven strategic planning and decision-making throughout the life cycle of a building. In this case, it will serve as a base for performance monitoring and analysis, continuous commissioning, and effective planning for the future.

III. DIGITAL TWINS

A. Digital Twinning Concept

A DT can be defined as a digital reproduction of real assets, systems, or processes that may be utilized for a variety of services and deliveries. So, a DT is more than simply a digital model of an element or a unit; it is also a live operational model in which real-time data is combined with the virtual replica of the system to generate a DT. This type of connection enables for continuous monitoring of the present state and informing decision making for future scenarios and applications. In recent years, the usage of digital twins has expanded, with successful implementations recorded in a variety of industrial areas, including manufacturing, production, and servicing [9]. A DT is conceptually composed of three major components: 1) Physical Component (actual component); 2) Virtual Component; and 3) Data linking them.

Data at various levels, as depicted in Fig. 2, provide the connectivity and information transmission between the real and virtual components. The successful use of DTs in the manufacturing, industrial, and production sectors has paved the way for further research into DTs for buildings and facilities. A building DT, in this way, gives a comprehensive image of the real structure in use, linking the physical and digital worlds [10]. It blends a virtual model of the building with actual data

obtained from meters, sensors, and IoT connected devices to comprehend, analyze, control, and improve building processes and guide operational decisions on various levels and considering various energy systems within the building.

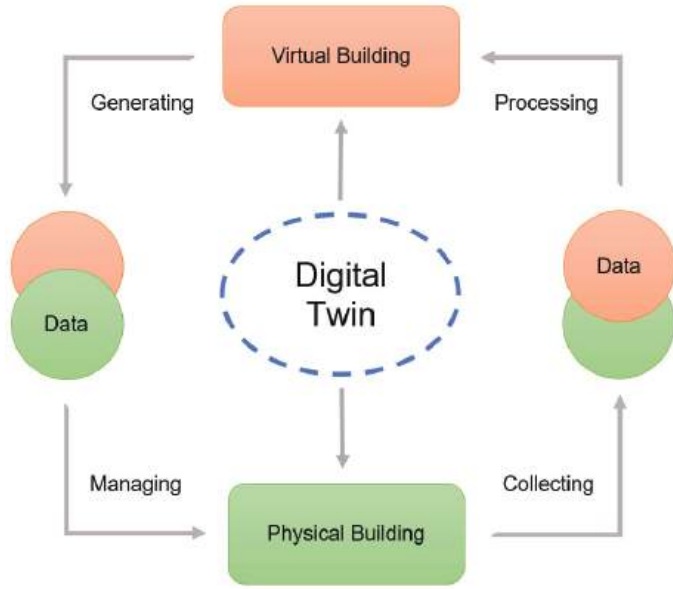


Figure 2. The main components of a standard digital twin.

A DT gives a real-time insight of how a building is operating by using structured data, allowing for fast adjustments to improve performance, leading to much more energy-efficient and cost-effective smart facilities and properties. A DT, on the full building level, offers a framework for comparing a building's actual performance to a predicted dynamic benchmark, and so serves as the foundation for commissioning operations as well as fault and anomaly identification and diagnostics. With both static facility information and dynamic data and predictions available on the same platform, building owners and facility and property managers can identify faults and problems with more precision and in less time, resulting in optimized and efficient proactive maintenance solutions. On the level of components and systems, a DT enables operational optimization by predicting and informing control, supervision, and management strategies based on different data inputs from various sources including, sensors, meters and IoT devices.

Given the concept and applications of a building digital twin, as well as the fact that existing BIMs are digital models containing all necessary building static data and information, such as geometry, structures, and systems, a DT may be viewed as the inevitable next development of the BIM model. It creates life into the static BIM by combining actual data and dynamic information, producing a living, dynamic DT that can support the building in its different phases throughout the entire life cycle to ensure energy efficient and cost-effective operation under all conditions. The building DT can employ information and operational technologies from the building information model, such as floors and spaces layouts, geometry, orientation, dimensions, and construction data. It may also collect real-time sensor operating data from the BMS,

such as data from smart HVAC systems, Internet of Things (IoT) devices, as well as information on the building services, sensor data, meters and submeters energy consumption and building use and utilization information.

B. Buildings DT Development Challenges

While it is true that building DT technologies and approaches have been around for more than a decade with promising outcomes and large demonstrated potential, it has yet to gain common grounds in global markets and industrial practices in the form of products or application lines due to significant challenges and obstacles that must be addressed in connection to the DT design and its corresponding components and systems. The main challenges hindering the promotion and the wide-scale development and implementation of such building digital twin solutions include:

- Establishing an automatic, seamless, and smooth information carryover from open platform BIM and building dynamic energy models, in addition to data and information context models, is critical to the rapid adoption of DTs in the building market.
- Dynamic models serve as the foundation for digital twin predictive, forecast, and scenarios testing services. The present modeling method is a stressful, costly, and time-consuming procedure in which modelers build models from the ground up, along with every single specification. This necessitates the requirement for modeling process automation.
- There is an essential need for an accessible and open-standard data platform that is versatile, functional, effective, and user-friendly in order to collect and manage data created within the building by diverse sources such as sensors, meters, and IoT devices.
- Significant efforts have been made in recent years to create rules and standards for effective BIM creation as well as the gathering and preservation of building data sets. Nevertheless, little progress has been made in establishing the BIM and data model prerequisites for successful DT implementations.

IV. DIGITAL TWIN PLATFORM

A. DT Platform Design

The present work tackles the four major problems mentioned above in creating and implementing efficient and adaptable building DTs. The conception, design, development, and demonstration of a novel comprehensive Digital Twin platform for building applications enhances the state of the art in regard to constructing building information modeling and open-standard data gathering and administration platforms. Fig. 3 depicts a simplified picture of the intended building digital twin structure and the essential components, as well as data and information flow.

The IFC, a BIM open-standard file format [11], is utilized as the foundation for the development since the goal is to construct a comprehensive and general DT platform. While both IFC and gbXML BIM schemes give extensive

information on the building structure, envelope, and geometry, IFC outperforms gbXML in terms of HVAC system characterization and knowledge carryover to other applications and tools. One of the primary distinctions between the two BIM formats is that the gbXML schema exclusively describes rectangle-shaped building geometry, but the IFC format may represent any geometry [12].

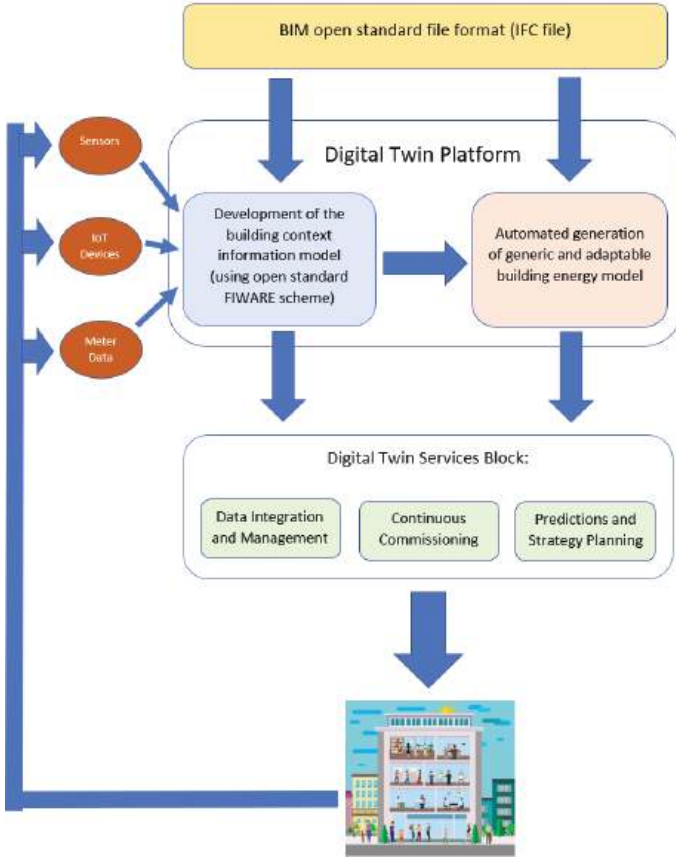


Figure 3. The proposed building digital twin framework and the main components.

As shown in Fig. 3, the proposed building DT platform is built on two cornerstones: 1) a building context information structure that integrates all sources of data acquired within the building; and 2) a series of generic and customizable grey-box building dynamic energy models. Grey-box models provide a balanced and hybrid building energy modeling technique in terms of design and implementation [13]. These models employ physics equations to depict the behavior of the building, although they are significantly simpler equations than those used in white-box models. As a result, they can be recreated and simulated in a more robust way and relatively in a quick manner. Nevertheless, such reduced physics results in a loss of accuracy, which is mitigated by model calibration with data, in a manner similar to the black-box modeling technique. A grey-box energy model, as a result, strikes a compromise between the precision of a white-box model and the speed of a black-box model. It has shown flexibility and efficacy in the development of operational management strategies as well as serving as a basis for control and maintenance applications and processes. In addition, Fig. 4 shows the major components of a

building digital twin with the flow of data between the real building and its virtual counterpart [14]. The different components, including the space heater, air diffuser, supply fan, cooling and heating coils, air heat exchangers, are to be modelled within the digital twin and acted on by the corresponding services identified.

B. Building Context Information Modeling

As part of the proposed digital twin solution in this work, FIWARE [15] is utilized as a platform for developing and creating data models inside a standardized context information framework. FIWARE is able to provide a thorough overview of the actual, physical, logical, and virtual resources and components in buildings and facilities, as well as the links and connections between them. In this regard, the Context Broker technology is at the heart of the FIWARE platform, enabling data collecting from many vertical devices and systems as well as data management, processing, share and exchange.

The FIWARE platform will serve as a backbone to provide a data-based context model of the building. For example, it will allow relating a specific ventilation system to the set of rooms supplied by this system along with the major components, including heat recovery units, fans, preheating loops, and valves, as well as the ventilation dampers actuated by the controller managing this specific system. On this basis, this context information model in FIWARE will be the backbone from which the energy models are going to read information and data on spaces and energy components and systems, in addition to being the platform where the energy model simulations and results are going to be pushed and stored.

Furthermore, establishing a seamless, direct, and secure connection between the different IoT sensors and meters in the building as well as the building management systems and the data collection and storage platform is key to the success of the automated twin solution. The deployment of a DT platform based on FIWARE overcomes the problem of comprehending the massive volume of streaming data generated by all of the building sensors and equipment. Because of this misunderstanding, the majority of sensor data is presently retained but not used effectively.

The digital twin data model is created by constructing a semantic knowledge network in NGSI-LD [16] in order to create an integrated data environment for building data, energy, and human behavior. The necessary data models for a DT are specified using the NGSI-LD model, and the corresponding runtime functions and advancements that render the Digital Twin operational are constructed. Overall, the IFC-based BIM supplies the open-source data framework with knowledge and specific information on the facility geometry, location, orientation, space allocation, and space, while still serving as a foundation for transferring building envelope and systems and components information to the corresponding dynamic hybrid energy models on various levels. The digital twin platform, as depicted in Fig. 3, will guarantee a smooth transfer of knowledge and information across the building phases and offer an integrated and coordinated decision-support service all across the building life cycle and among all the different phases from conception to operation and maintenance.

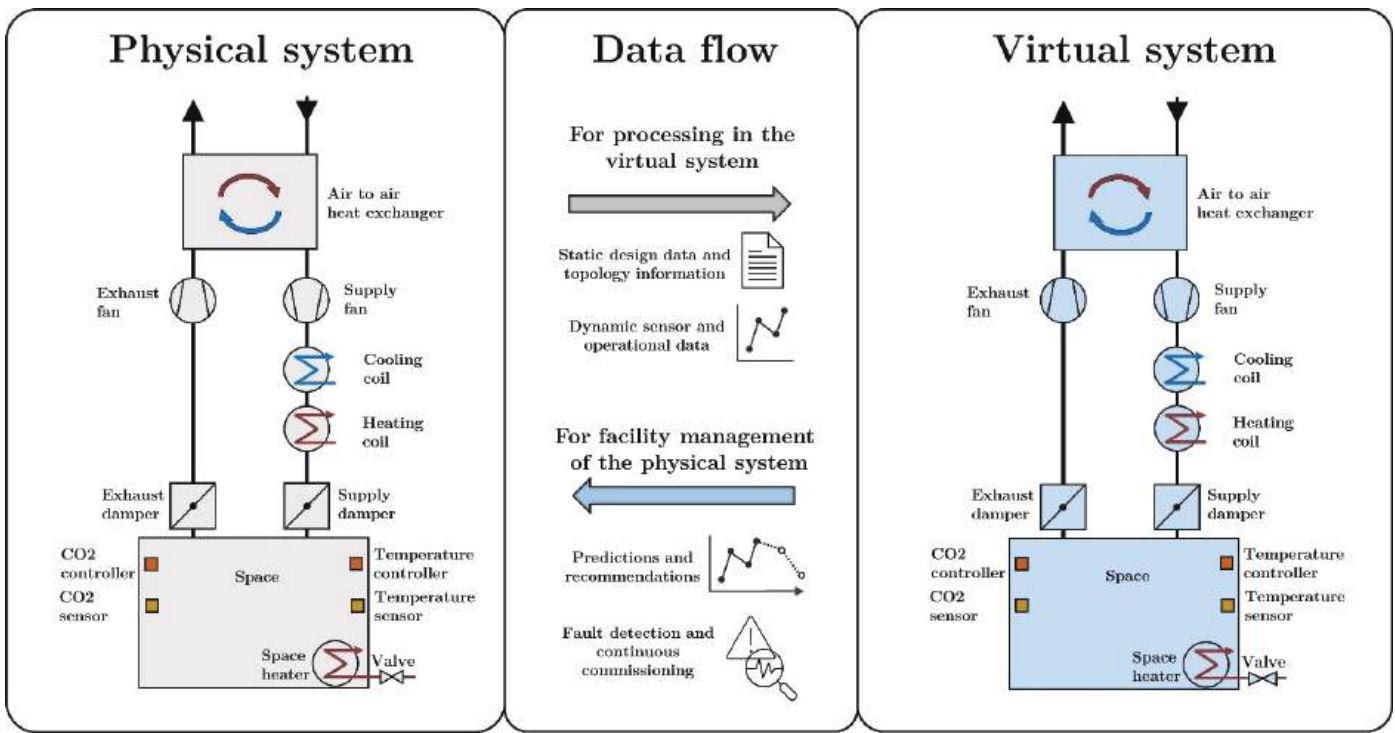


Figure 4. The major components of a building digital twin with the flow of data between the real building and its virtual counterpart.[14]

C. Building Energy Modeling

Much work has been expended in developing large-scale energy modeling tools, such as EnergyPlus, Modelica, DOE-2, and TRNSYS [17]. Models created using such methods are classified as white-box models since they are solely based on first-principles building physics, often in the form of the energy and mass balance and conservation equations, which frequently lead to enormous systems of Ordinary Differential Equations (ODE). But as is the case with the majority of big-scale white-box models, they necessitate vast construction documents in addition to being time and resources consuming, where a good building model would take a considerable amount of time to be developed and validated. Considering EnergyPlus for instance, being one of the widely used and validated white box building modeling tools, a standard modeling and simulation workflow involves 3 time-consuming steps before getting a precise energy model. Initially, detailed geometric data from blueprints, cross-sections, 3D models, and so on are used to establish the building structure. Second, envelope specifications as well as energy systems design data and building services and devices are to be introduced and specified. Moreover, the user needs to highlight the thermal qualities of the environment and occupant, lights, and equipment schedules for the areas. Lastly, in order to attain acceptable forecast accuracy, the model should be validated using actual building consumption data, a procedure that frequently depends on a manual trial-and-error method requiring specific domain knowledge from the model developer.

To address the time and resources limitations of using white box models, data-driven modelling techniques have

recently grown in popularity for developing generic and adaptable building models, owing to their low calculation demands and high prospects for integration and automation in smart buildings and facilities where sensing and metering systems collect large amount of divers data on the building performance on different levels, including the HVAC components and energy systems and services operation.

Data-driven modeling techniques are often classified as black-box or grey-box methods. Black-box models are data-driven in nature, which means that no domain-specific model structure is expected. Certain algorithms are therefore used to a collection of data to establish acceptable functional connections that relate system inputs to outputs. Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees, are examples of frequently used black-box models [118]. On the other hand, grey box models provide a balanced approach, a mix, between the physics-based white box models and the data-driven black box models, with simple physics and conservation principles included, where data is needed to fill the gaps of information and estimate the rest of the unknown parameters,

In the proposed digital twin platform in this work, a hybrid approach of energy modeling is employed, where black models are implemented and developed to model building spaces relying majorly on data collected from the building as inputs, including room temperatures, CO₂ levels, and ambient air conditions [19, 20]. In addition, grey-box models are used to develop components and energy systems models, including the components shown in Fig. 4, allowing simple modeling algorithms to be executed based on limited energy balance equations and actual data supporting this from the building. The use of grey-box models to model the different building components is key to carry on operational control and energy management services as they provide flexibility in terms of

modifying and varying specific parameters in the building and reporting the impact of the change or the effect of one strategy compared to another. For instance, such models will allow controlling the heating setpoint in the room or optimizing the supply air fan operation. To aid the models' development, the authors have developed an ontology-based innovative energy modeling framework for scalable and adaptable building digital twins. A detailed description of the modeling framework and the corresponding energy models used as well as the model-data integration and processing are presented in the authors recent publication [14]. In this regard, the modeling framework is implemented in Python as a reusable library, which will serve as a basis for the hybrid energy models' generation.

V. CONCLUSION

In this work, a comprehensive and holistic digital twin platform is designed and presented aiming to enhance the energy efficiency and smartness of buildings. The suggested proposed integrated Digital Twin platform incorporates scalable and flexible dynamic data-driven energy models with actual real-time data from various metering and sensing devices around within the building. It substitutes the current BIMs with dynamic energy models in combination with real-time data to deliver the following supporting services:

1. Gathering, processing, and managing the data created in the building by a variety of sources, such as sensing, metering, and IoT devices, inside an accessible and user-friendly open-standard context information model framework.
2. Real-time performance tracking and analysis, as well as automatic continuous commissioning and fault detection, allowing wiser and smarter building energy management, efficient systems operation, quick identification of defects, risk anticipation, and problems prevention before occurrence.
3. Planning and strategy assistance throughout the building life cycle by informing effective and optimal decisions, performing "What if?" events, and evaluating alternate control and management patterns in a virtual environment.

The proposed holistic building digital twin platform in this study adds value across the entire value chain, including owners and users, architects, designers, facility managers, consultants, planners, developers, and government agencies.

ACKNOWLEDGMENT

This work is carried out under the 'Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle' project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64021-1009.

REFERENCES

[1] European Commission, State of the Union: Commission raises climate ambition and proposes 55% cut in emissions by 2030, September 2020. https://ec.europa.eu/commission/presscorner/detail/en/ip_20_1599https://ec.europa.eu/commission/presscorner/detail/en/IP_20_1599 (accessed Feb. 23, 2023).

[2] BR18 Danish Building Regulations. <https://byggningsreglementet.dk> (accessed Feb. 25, 2023).

[3] International Energy Agency, Global Status Report for Buildings and Construction - Towards a zero-emissions, efficient and resilient buildings and construction sector, December 2019. <https://www.iea.org/reports/global-status-report-for-buildings-and-construction-2019> (accessed Feb. 19, 2023).

[4] M. Jradi, C. Veje, B.N. Jørgensen, "Deep energy renovation of the Mærsk office building in Denmark using a holistic design approach", *Energy Build.*, 151 (2017), pp. 306-319.

[5] Twin4Build - A holistic Digital Twin platform for decision-making support over the whole building life cycle. <https://www.sdu.dk/en/forskning/centreforenergyinformatics/research-projects/twin4build> (accessed Feb. 22, 2023).

[6] R. Charef, H. Alaka, S. Emmitt, "Beyond the third dimension of BIM: a systematic review of literature and assessment of professional views", *J. Build. Eng.*, 19 (2018), pp. 242-257.

[7] S. Tang, D. R. Shelden, C. M. Eastman, P. Pishdad-Bozorgi, X. Gao, "A review of building information modeling (BIM) and the internet of things (IoT) devices integration: present status and future trends", *Autom. Constr.*, 101 (2019), pp. 127-139.

[8] Y. Teng, J. Xu, W. Pan, Y. Zhang, "A systematic review of the integration of building information modeling into life cycle assessment", *Build. Environ.*, 221 (2022), Article 109260.

[9] S. S. Kamble, A. Gunasekaran, H. Parekh, V. Mani, A. Belhadi, R. Sharma, "Digital twin for sustainable manufacturing supply chains: current trends, future perspectives, and an implementation framework", *Technol. Forecast. Soc. Change*, 176 (2022), Article 121448.

[10] J. Zhao, H. Feng, Q. Chen, B. Garcia de Soto, "Developing a conceptual framework for the application of digital twin technologies to revamp building operation and maintenance processes", *J. Build. Eng.*, 49 (2022), Article 104028.

[11] BuildingSMART International. 2022. Industry Foundation Classes (IFC). <https://www.buildingsmart.org/standards/bsi-standards/industry-foundation-classes> (accessed Feb. 21, 2023).

[12] A. Andriamamonjy, D. Saelens, R. Klein, "An automated IFC-based workflow for building energy performance simulation with Modelica", *Autom. Constr.*, 91 (2018), pp. 166-181.

[13] K. Arendt, M. Jradi, H.R. Shaker, C. Veje, "Comparative analysis of white-, gray- and black-box models for thermal simulation of indoor environment: Teaching building case study", *Proceedings of the 2018 Building Performance Modeling conference and Simbuild co-organized by ASHRAE and IBPSA-USA*, ASHRAE (2018), pp. 173-180.

[14] J. Bjørnskov, M. Jradi, "An Ontology-Based Innovative Energy Modeling Framework for Scalable and Adaptable Building Digital Twins", *Energy Build.*, Article 113146.

[15] FIWARE Foundation. FIWARE: The Open Source Platform for Our Smart Digital Future, 2021. <https://www.fiware.org/> (accessed Feb.17, 2023).

[16] G. Privat, Guidelines for Modelling with NGSI-LD, ETSI White Paper No. 42, 2021.

[17] M. Jradi, K. Arendt, F.C. Sangogboye, C.G. Mattera, E. Markoska, M.B. Kjærgaard, C.T. Veje, B.N. Jørgensen, "Obepme: an online building energy performance monitoring and evaluation tool to reduce energy performance gaps", *Energy Build.*, 166 (2018), pp. 196-209.

[18] Y. Sun, F. Haghighat, B.C.M. Fung, "A review of the-state-of-the-art in data-driven approaches for building energy prediction", *Energy Build.*, 221 (2020), Article 110022.

[19] J. Bjørnskov, M. Jradi, C.T. Veje, "A Fully Automated and Scalable Approach for Indoor Temperature Forecasting in Buildings Using Artificial Neural Networks", *Proceedings of the 5th Building Simulation Applications conference (BSA 2022)*, Bolzano, Italy, 29 June – 1 July 2022.

[20] J. Bjørnskov, M. Jradi, "A Modular Thermal Space Coupling Approach for Indoor Temperature Forecasting Using Artificial Neural Networks", *Proceedings of the 6th Building Simulation and Optimisation conference (BSO 2022)*, Bath, UK, 13-14 December 2022.

Adaptable and Scalable Energy Modeling of Ventilation Systems as Part of Building Digital Twins

Andreas Hyrup Andersen¹, Jakob Bjørnskov¹, Muhyiddine Jradi¹

¹University of Southern Denmark, Odense, Denmark

Abstract

Digital twins solutions have great potential within a multitude of building applications, but flexible and scalable model backbones are essential for cost-effective large scale distribution within the building sector. In this paper, a flexible component based framework for ventilation system modeling is proposed as part of a scalable digital twin solution. The framework is applied to a case study with limited data availability and the simulated values are compared to measurements. Results show a strong similarity between predictions and data, but it is clear that inaccurate occupancy measurements can cause significant discrepancies. The developed model is used for performance comparison between different ventilation control strategies and a proportional controller is found to slightly outperform the buildings current rule-based controller.

Key Innovations

- Development and evaluation of flexible component-based modeling framework for simulation of demand-driven ventilation systems
- Consideration of case study building with multiple space types
- Evaluation of energy consumption and air quality with different control strategies

Practical Implications

This paper showcases the flexibility to combine different modeling techniques in a component-based building modelling framework within a digital twin platform. Results from control strategy comparison demonstrate significant potential for energy savings with more advanced damper controller strategies.

Introduction

Decarbonization of the building sector is critical for achieving contemporary climate goals, and to reach the European Union's goal of climate neutrality in 2050 urgent initiatives and activities are needed, focusing on enhancing the design and operation of newly built buildings as well as accelerating the rate of renovation of the existing building stock BPIE (2022). In this regard, the need for smarter, energy-efficient, and more flexible buildings calls for innovative techniques and solutions to improve the overall building energy performance and enhance the

operational efficiency. Among these innovative solutions, digital twins have demonstrated large potential with various capabilities to monitor, evaluate, predict, optimize, and manage buildings in a smarter and automated way Clausen et al. (2021).

The digital twin concept proposed by Grieves (2015) consists of three main elements: A physical system, a virtual representation and a connecting data flow. A building digital twins is a holistic representation of an existing building, which emulates the behavior of it's physical counterpart. It functions as a multi purpose tools, which facilitates a number of services such as data collection and management, automated continuous commissioning and operational strategy planning.

In recent years the use of digital twin approaches for both design and optimization in building energy simulation has accelerated as shown by Bortolini et al. (2022). Current approaches often rely heavily on existing building information such as in Zaballos et al. (2020), where an existing building information model (BIM) is used as backbone for campus building modelling. Such solutions are impressive in their capabilities, but only applicable to a small portion of the current building stock. For widespread application of digital twin technology within the building sector, a more flexible and scalable approach to building modelling is needed. Buildings are however multifarious with a wide variety of systems and devices from different contractors, which complicates the development of scalable and reusable building models. In this paper a component based framework for ventilation system modeling is proposed. Unlike traditional component based solutions such as Modelica, the framework aims to facilitate the use of a mixture of modeling techniques between component models (e.g. connecting a black-box space model to a damper model based on first principles). The goal is to deliver a flexible, efficient and robust solution for ventilation system modelling, which can be scaled and modified to a specific building with minimal manual labor. The proposed framework adopts the SAREF ontology and is developed as part larger building modelling framework which is available at the the GitHub repository <https://github.com/JBjoernskov/Twin4Build> and presented in Bjørnskov and Jradi (2023). The proposed data-driven component modeling framework will serve as a basis for building performance monitoring, continuous commissioning, operational strategy testing and strategy plan-

ning within a digital twin platform. The work in this paper is carried out as part of the Twin4Build project, aiming at the design, development and demonstration of an innovative digital twin solution to enhance efficiency and enable higher level of smartness in buildings.

The following section describes the the modeling and simulation methodology, presenting the models used to characterize each component of the ventilation system considered. Then a case study is considered to implement and assess the modeling framework proposed. Parameter estimation results are presented for the ventilation system components employing actual data from the building and the developed model is used to evaluate and assess various operational strategies and control pattern impacts.

Modelling and Simulation Methodology

The framework presented in this paper focuses purely on ventilation modelling and the thermal part connected to the ventilation systems, dealing with the preheating and precooling coils as well as the heat recovery unit are thus not considered. Figure 1 visualises the component structure in a typical heating, ventilation and cooling (HVAC) system and the reduced system considered here.

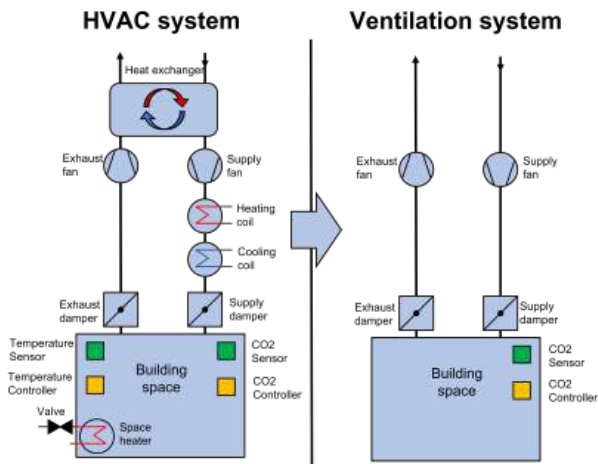


Figure 1: Component overview for typical HVAC system (left) and the considered ventilation system (right).

To fit an efficient digital twin platform it is crucial that the developed modelling framework avoids unnecessary model complexity to ease the computational demand. Therefore, the ventilation system is modelled as a quasi-steady-state system. This is a common approach in HVAC modelling as it takes advantage of the large difference between the physical time constants associated with HVAC systems and building envelopes and spaces. While HVAC systems are dominated by rapid dynamics which reach equilibrium within seconds or minutes, building envelopes and spaces are governed by much slower dynamics.

The modelling framework consist of a number of ventilation system components which are connected by the inputs and outputs. The models used for the individual components are based on validated models presented in the literature and in previous studies. The selected mod-

els have been chosen for their simplicity, performance and applicability for data-driven calibration and parameter estimation. For some component types several models are included in the framework to accommodate various degrees of available building information. For instance, a fan may be modelled using pressure difference or flow measurements depending on available building documentation and data. Only component models applied in the presented case study are described in this paper.

Fan

The fan generates the pressure difference required for the desired air flow through the ventilation system. The fan model used for both supply and exhaust fans is based on the EnergyPlus *Variable Speed Fan Model* found in Big Ladder Software (2021). It is given by equations 1-3.

$$f_{flow} = \frac{\dot{m}_a}{\dot{m}_{a,max}} \quad (1)$$

$$f_{pl} = c_1 + c_2 f_{flow} + c_3 f_{flow}^2 + c_4 f_{flow}^3 \quad (2)$$

$$\dot{W}_{fan} = f_{pl} \dot{W}_{fan,max} \quad (3)$$

The flow fraction f_{flow} is calculated as the ratio between flow rate $\dot{m}_a$ and maximal flow rate $\dot{m}_{a,max}$ in equation 1. Equation 2 determines a part load fraction f_{pl} from the flow fraction and the power coefficients c_1 - c_4 , where $c_1 + c_2 + c_3 + c_4 = 1$. The fans power consumption $\dot{W}_{fan}$ is calculated using the maximal fan power consumption $\dot{W}_{fan,max}$ and part load fraction in equation 3.

Damper

The damper controls the air flow between the ventilation system and a building space. The damper model is based on a pressure independent air damper from Modelica Buildings Library (2021) and described by equation 4. It calculates the air flow $\dot{m}_a$ as the product of the damper opening position $y \in [0, 1]$ and the maximal damper flow $\dot{m}_{a,max}$.

$$\dot{m}_a = y \dot{m}_{a,max} \quad (4)$$

Controller

Controllers are vital for building system operation and come in many different forms. Two controller types are applied in this paper; rule-based control and Proportional Integral Derivative (PID) control. Equations 5-6 from Veeranna and Yaragatti (2010) describes a PID controller in discrete time.

$$e_t = y_{set,t} - y_{meas,t} \quad (5)$$

$$u_t = u_{t-1} + K_p(e_t - e_{t-1}) + K_i \Delta t + \frac{K_d}{\Delta t}(e_t - 2e_{t-1} + e_{t-2}) \quad (6)$$

The residual e_t at time t is given by the difference between set point $y_{set,t}$ and measurement $y_{meas,t}$ in equation 5. In

equation 6 the controller signal u_t is the calculated form the previous signal u_{t-1} , the residual e_t and the parameters K_p , K_i and K_d which control the strength of the proportional, integral and derivative response respectively.

Building Space

Indoor comfort is the main motivation for HVAC system operation. In this paper, the indicator for indoor air quality is the concentration of carbon dioxide (CO_2) within building spaces. The building space model aims to describe the dynamic development of CO_2 -concentration within a space based on a mass balance equation from Macarulla et al. (2017) given by equation 7.

$$dC_{int}m_{a,int} = \dot{m}_{a,ven}(C_{ven} - C_{int})dt + \dot{m}_{a,inf}(C_{inf} - C_{int})dt + K_{occ}n_{occ} \quad (7)$$

Where $m_{a,int}$ is the mass of air in the room and the development in internal CO_2 -concentration C_{int} is described by the impact from ventilation, infiltration and occupancy. The ventilation impact depends on the CO_2 -concentration in the ventilation system C_{ven} and the ventilation mass flow $\dot{m}_{a,ven}$, whereas the infiltration impact depends on the CO_2 -concentration of the infiltrating air C_{inf} and the infiltration mass flow $\dot{m}_{a,inf}$. The occupancy impact is given by the CO_2 -generation rate pr. occupant K_{occ} and the number of occupants n_{occ} . Equation 7 is transformed with backward Euler discretization and equation 8 is obtained, where Δt is the time step.

$$C_{int,t} = C_{int,t-1} + \frac{\dot{m}_{a,ven}(C_{ven} - C_{int,t-1})\Delta t}{m_{a,int}} + \frac{\dot{m}_{a,inf}(C_{inf} - C_{int,t-1})\Delta t + K_{occ}n_{occ}\Delta t}{m_{a,int}} \quad (8)$$

It is assumed that the CO_2 -concentration of the infiltrating air and air from the ventilation system is equal to the outdoor air CO_2 -concentration level at 400 ppm.

Component overview

An overview of the included components and their inputs, outputs, parameters and constants is provided by table 1.

Table 1: Overview of inputs, outputs, parameters and constants in component models.

Components	Inputs	Outputs	Parameters	Constants
Fan	$\dot{m}_a \in \mathbb{R}^+$	$W_{fan} \in \mathbb{R}^+$	$c_1 \in \mathbb{R}$ $c_2 \in \mathbb{R}$ $c_3 \in \mathbb{R}$ $c_4 \in \mathbb{R}$ $\dot{m}_{a,max} \in \mathbb{R}^+$ $W_{fan,max} \in \mathbb{R}^+$	ρ_a
Damper	$y \in [0, 1]$	$\dot{m}_a \in \mathbb{R}^+$	$\dot{m}_{a,max} \in \mathbb{R}^+$	
Controller	$y_{set} \in \mathbb{R}$ $y_{meas} \in \mathbb{R}$	$u \in \mathbb{R}$	$K_p \in \mathbb{R}$ $K_j \in \mathbb{R}$ $K_d \in \mathbb{R}$	
Building Space	$n_{occ} \in \mathbb{N}$ $\dot{m}_{a,ven} \in \mathbb{R}^+$ $\dot{m}_{a,inf} \in \mathbb{R}^+$	$C_{room} \in \mathbb{R}^+$	$k_{occ} \in \mathbb{R}^+$ $C_{ven} \in \mathbb{R}^+$ $\dot{m}_{a,int} \in \mathbb{R}^+$ $C_{inf} \in \mathbb{R}^+$	Δt

Case Study

In order to assess the performance of the proposed modelling framework a case study is considered. The chosen building is a highly energy-efficient teaching and office building located at Odense Campus of the University of Southern Denmark Jradi et al. (2017). It was build in 2015 with an area of around 8500 m² spanning across 3 floors and a basement for storage and technical installations. The building consists largely of teaching rooms, but also features several study zones, meetings room, personal offices and recreational areas. It has 4 separate ventilation systems, which each service around one quarter of the building. Only one of these ventilation systems is considered for the case study. A schematic of the ventilation system is shown in figure 2.

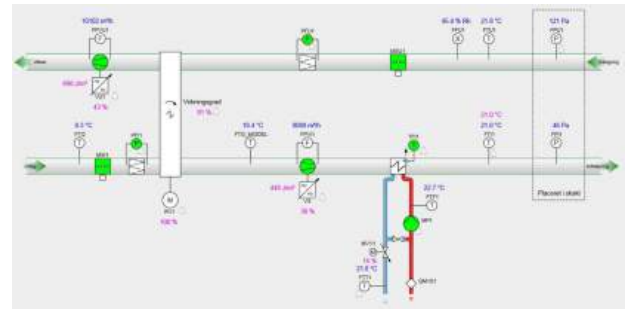


Figure 2: Schematic of the case study ventilation system.

The ventilation system operation is controlled based on CO_2 -measurements in the ventilated spaces. A four step rule-based control strategy controls the damper opening fraction based on the CO_2 -concentration range:

$$\begin{aligned} C_{int} < 600 \text{ ppm} &: 0\% \text{ opening} \\ 600 \text{ ppm} \leq C_{int} < 750 \text{ ppm} &: 45\% \text{ opening} \\ 750 \text{ ppm} \leq C_{int} < 900 \text{ ppm} &: 70\% \text{ opening} \\ 900 \text{ ppm} \leq C_{int} &: 100\% \text{ opening} \end{aligned}$$

Only a few spaces in the building are fitted with occupancy sensors, which are ideal for verifying the building space model. Of these, two spaces "Ø22-508-1" and "Ø22-511-2" are both serviced by the same ventilation system so these are chosen for the case study. Ø22-508-1 is a 125 m² study-zone intended for group work in small teams. It features a number of mobile whiteboards, TV-screens and other furniture which creates natural room separation. The estimated capacity is 40 occupants. Ø22-511-2 is a 139 m² teaching room with an estimated capacity of around 70 occupants. Classes are however often much smaller.

An overview of component connections for the case study system is presented in figure 3. Note that cycles such as the signals from the CO_2 sensors back to the controllers are hidden for the sake of simplicity.

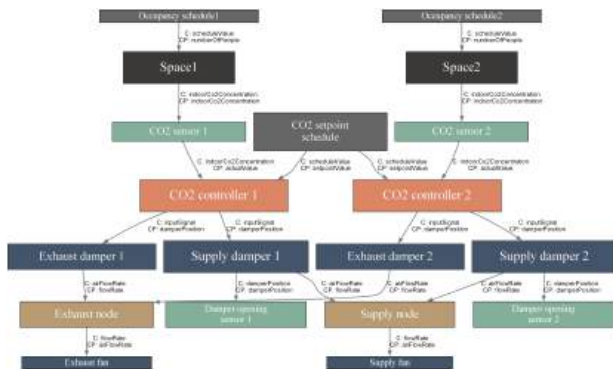


Figure 3: System graph with component connections for the case study system. Cycles are hidden.

The case study's main purpose is to verify the modelling frameworks performance by comparing measurements and simulated values for the current rule-based strategy. Additionally, a simple scheduled controller and a PID-controller are tested in the same simulation period to compare the power consumption and indoor air quality with different control strategies.

Power and Air Quality Calculations

The two spaces examined in the case study only constitute a fraction of the total ventilation demand in the system. For the fan power calculations it is assumed that the remaining spaces connected to the ventilation system cause a base load of $5 \text{ m}^3/\text{s}$ on both the supply and exhaust fan. The power demand for each control strategy is then calculated as the additional power needed to provide the required ventilation flow for Ø22-508-1 and Ø22-511-2 on top of the base load.

According to Kapalo et al. (2018) CO_2 -concentration is the most adequate parameter for measuring indoor air quality. Within the literature many different CO_2 -concentrations are suggested as thresholds when evaluating air quality. For this comparison a threshold of 700 ppm is used and the air quality KPI is defined as the time integrated residual above this value as expressed by equation 9.

$$KPI = \sum_{t=1}^N \max((C_{int,t} - C_{threshold})\Delta t, 0) \quad (9)$$

Where $C_{int,t}$ is indoor CO_2 concentrations at time t , $C_{threshold}$ is 700 ppm, Δt is the model time step and N is the number of timesteps in the simulation. The total KPI is calculated as the sum of KPI's for both spaces. Note that the KPI is a 'penalty' so the lower the value the better.

Parameter Estimation

In the case study, the values of some component parameters are missing within the available building documentation. These are instead determined through parameter estimation using historical building data. Parameters are estimated by minimizing the sum of squared residuals between measured and simulated values over a continuous training period. The estimation is conducted with the use of the "minimize" function from the SciPy (2023) Python

library. Several solvers are tested and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method is chosen for its ability to solve unconstrained nonlinear problems effectively with reasonable initial guesses.

Fans

For both the supply and exhaust fan, the power coefficients c_1, c_2, c_3 and c_4 need to be estimated. In the chosen training data there are no hours where $0 < \dot{W}_{fan} < 0.3\dot{W}_{fan,max}$, so to ensure that the lower flow rate region of the fans are modeled reasonably $c_1 = 0$ is assumed. This ensures that $\dot{W}_{fan} = 0$ when $\dot{m}_a = 0$. The training data consists of two separate days, one summer day and one winter day, with significantly differing ventilation profiles and demands.

Figure 4 shows the measured power consumption and the predicted power consumption for the exhaust fan

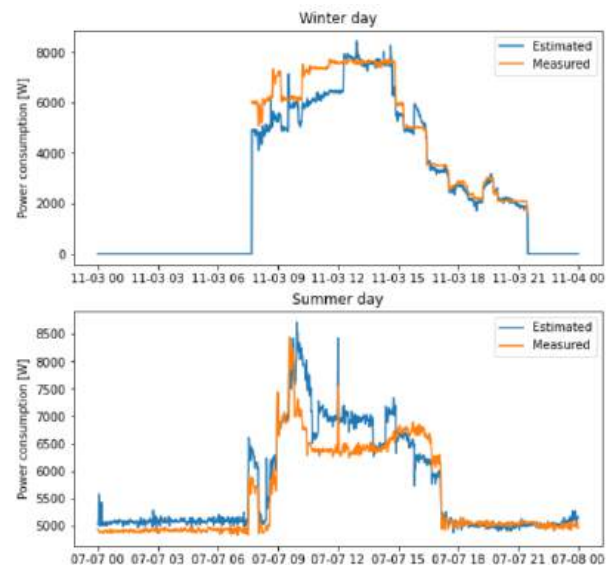


Figure 4: Measured and estimated exhaust fan power consumption for the training periods.

with estimated parameters during the two training days. As evident from the figure, the estimated parameters allow for reasonably accurate modeling of the fans electricity consumption during both training days. Notably there are visible estimation errors in some high demand hours on both days. This error could be further explored to improve the model if additional training data were available.

The power coefficients estimated from the training periods provide the fan power/flow curves displayed in figure 5.

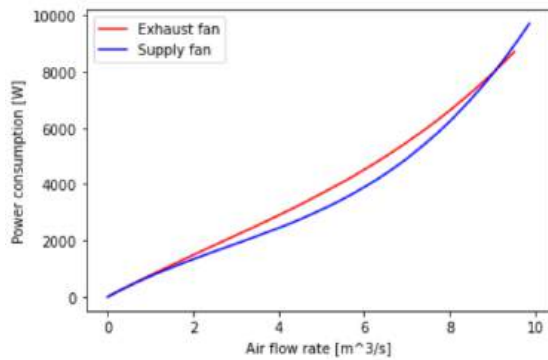


Figure 5: Relation between fan power consumption and flow rates with estimated power coefficients for the case study supply and exhaust fan.

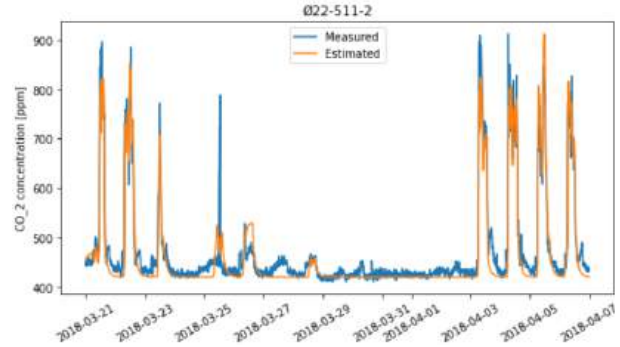


Figure 7: Estimated and measured CO₂-concentration during training period for Ø22-511-2.

Space and Damper

For the dampers in both spaces the nominal flow rates $\dot{m}_a$ are unavailable. Damper flow is not measured so the nominal flow rates are estimated based on their ventilation impact in their corresponding spaces using both the damper and building space models. Additionally, the CO₂-generation rate pr. occupant K_{occ} and the infiltration air mass flow rate $\dot{m}_{a,inf}$ are estimated for each space rather than relying on literature values. The actual infiltration air flow is expected to vary significantly depending the positions of manually operated openings (doors and windows) and indoor air pressure. Positions of these openings and the space pressure is however not measured, so the infiltration flow rate is assumed to be constant in the model. For the parameter estimation a consecutive 17 day training period is used. The measured and predicted CO₂-concentration with estimated parameters are shown in figure 6 and 7 for spaces Ø22-508-1 and Ø22-511-2 respectively.

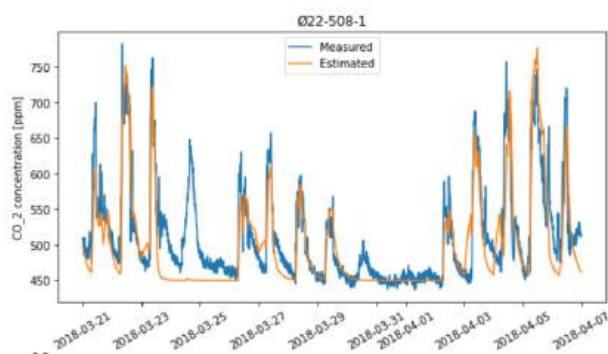


Figure 6: Estimated and measured CO₂-concentration during training period for Ø22-508-1.

Parameter overview

Table 2 shows the parameters estimated for the case study. The estimated infiltration rates correspond to air change rates of 0.20 and 0.39 room volumes pr. hour which aligns with typical values from the literature.

Table 2: Estimated parameters used in the case study.

	Ø22-508-1	Ø22-511-2
Damper	$\dot{m}_a = 0.22 \text{ kg/s}$	$\dot{m}_a = 0.46 \text{ kg/s}$
Building Space	$K_{occ} = 5.34 \text{ L/h}$ $\dot{m}_{inf} = 0.0299 \text{ kg/s}$	$K_{occ} = 9.19 \text{ L/h}$ $\dot{m}_{inf} = 0.0627 \text{ kg/s}$
Supply Fan	$c_2 = 0.8321$ $c_3 = -0.9847$ $c_4 = 1.1526$	
Exhaust Fan	$c_2 = 0.8843$ $c_3 = -0.4617$ $c_4 = 0.5775$	

According to Yang et al. (2020) the average CO₂-generation from occupants is around 15 L/h during office activities. The estimated values are significantly lower for both spaces, especially Ø22-508-1. This discrepancy may be caused by several unaccounted factors. Firstly, the occupancy sensor might undershoot the actual occupancy count, especially in Ø22-508-1 where mobile whiteboards, screens and other objects create natural room separations which can block the sensor view. A suspected instance of this can be seen in the training period (figure 6) on the 24th of March where a significant CO₂-spike was measured, but not simulated since no occupancy was recorded. It seems the sensor completely failed to register present occupants, which causes a significant model prediction error for the day.

Secondly, it is common that doors between spaces and adjacent hall areas are left open in the case study building. This increases the effective room volume and thus reduces the increase in CO₂-concentration from occupancy generation.

Simulation and Result Evaluation

The constructed model is used to simulate ventilation system operation over a 5 day period (monday-friday). The model is provided with the measured occupancy count in both spaces with a 10 minute resolution as the only in-

puts. The occupancy data and simulated airflow and CO₂-concentration for space Ø22-508-1 and Ø22-511-2 and displayed in figures 8 and 9 respectively. The developed model relies entirely on occupancy data, so a precise occupancy schedule is essential for accurate simulation. An inaccurate occupancy input will produce a roughly proportional error in the simulated ventilation requirement and resulting energy consumption. As illustrated by figures 6 and 7, the case study model is well calibrated to produce accurate CO₂-concentration predictions from the available occupancy data.

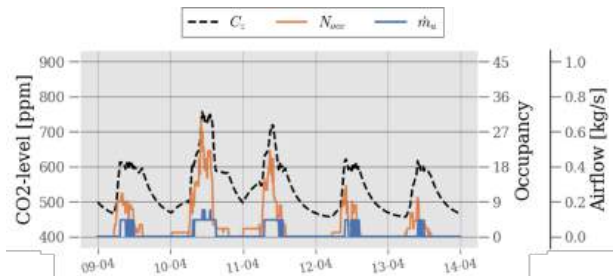


Figure 8: Occupancy data and simulated air mass flow rate and CO₂-concentration in Ø22-508-1.

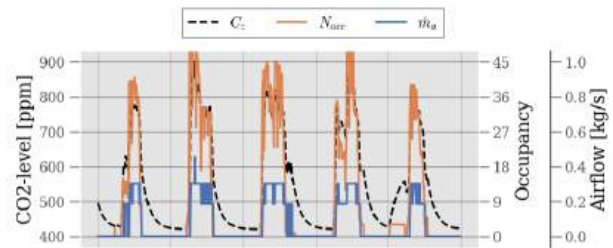


Figure 9: Occupancy data and simulated air mass flow rate and CO₂-concentration in Ø22-511-2.

The simulated CO₂-development in each of the spaces is compared to the measured values in figures 10 and 11.

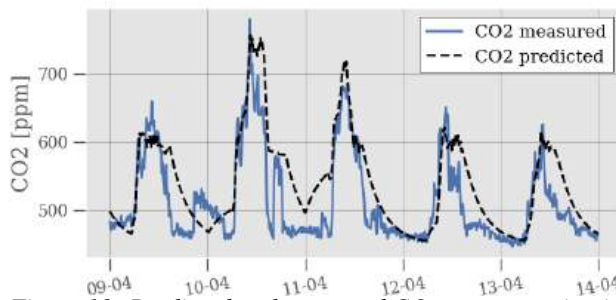


Figure 10: Predicted and measured CO₂-concentration in Ø22-508-1.

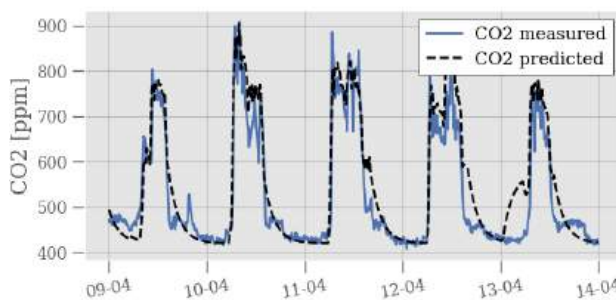


Figure 11: Predicted and measured CO₂-concentration in Ø22-511-2.

Figure 11 shows that the simulated CO₂-profile for Ø22-

511-2 is in line with the measured values. There are a few discrepancies which likely stem from various unexpected disturbances and sensor inaccuracies, but at large the model seems to emulate the actual system dynamics well. From figure 10 it is clear that the predicted CO₂-values for Ø22-508-1 also follow a similar pattern to the measured values with closely matching spike heights and times. There is however a clear discrepancy between the predicted and measured CO₂-developments after CO₂-spikes. It seems that the CO₂-concentration drops from 600 ppm (below which the damper should be closed) to around 450 ppm much faster than expected. This is seemingly, at least in part, due to some unexpected damper behavior for the Ø22-508-1 damper. The simulated and measured damper opening in Ø22-508-1 can be seen in figure 12. From the figure it is clear that the damper consistently is open for a longer period than expected. Over the 5 days period the total air mass delivered to the space by the ventilation system is twice as high as predicted. Especially on the 1st and 3rd day it is notable that the damper stays open much longer than predicted even when the CO₂-concentration is well below opening threshold of 600 ppm. This explains the unexpectedly rapid decline in CO₂-concentration observed in figure 10. It is unclear why the damper opening deviates so significantly from the expected opening strategy. It may be the case that the damper is signaled to stay open until a lower threshold is reached (e.g. 500 ppm) or for some minimum period of time after the opening, but the exact strategy is difficult to infer over the limited testing period. It is however clear that the error between simulated and measured CO₂-concentration in Ø22-508-1 is caused by some variance from the expected control strategy more-so than a faulty building space model.

In Ø22-511-2 the predicted damper position schedule is quite similar to the measurement as seen on figure 13. The total predicted air flow to the space from ventilation over the 5 day period is 86,4% of the measured total and most of the difference comes from earlier closing in the simulation - again likely due to some undocumented controller rule existing. It is concluded that the model functions as intended for ventilation system simulation, but lacks some con-

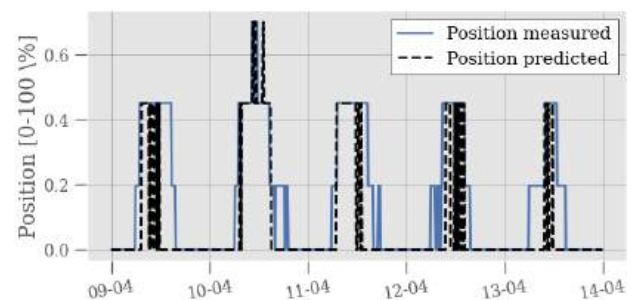


Figure 12: Predicted and measured damper position in Ø22-508-1.

troller refinement in the specific case study. The model can thus fittingly be applied for a performance comparison between different well-defined control strategies.

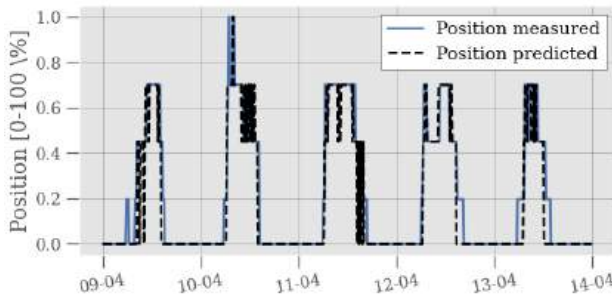


Figure 13: Predicted and measured damper position in Ø22-511-2.

Control Strategy Comparison

The modeling framework developed and implemented above will be used in this section to evaluate different operational strategies and control patterns in the building. Three different control strategies are simulated and compared with respect to power consumption and indoor air quality. The first is the rule-based controller which is currently used. The accumulative exhaust fan power consumption from ventilation of Ø22-508-1 and Ø22-511-2 with the rule-based controller is shown in figure 14 along with the total flow rate for the two spaces. The power consumption pattern for the supply fan is similar, but a little lower.

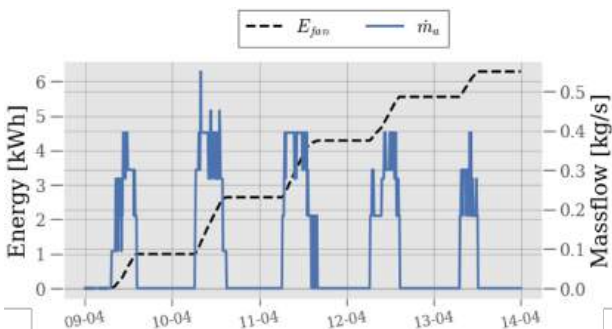


Figure 14: Simulated accumulative exhaust fan power consumption and total ventilation system flow rate in the case study with a rule based controller.

The second controlling strategy tested in the case study is a PID-controller. For the selected comparison method testing has shown that a purely proportional controller performs well since it minimizes KPI well enough and avoids overshooting the needed ventilation. Therefore the controller is given the values: $K_p = -0.005$, $K_i = 0$ and $K_p = 0$. The simulated accumulative exhaust fan power consumption with the proportional controller is shown in figure 15 and the simulated development in CO₂-concentration for Ø22-511-2 with the proportional controller is shown in figure 16. Comparison between the rule-based controller and PID-controller for managing CO₂ in Ø22-511-2 (figures 11 and 16) shows that the PID-controller is more effective at avoiding high CO₂-concentration spikes and faster at reducing the concentration to an acceptable level.

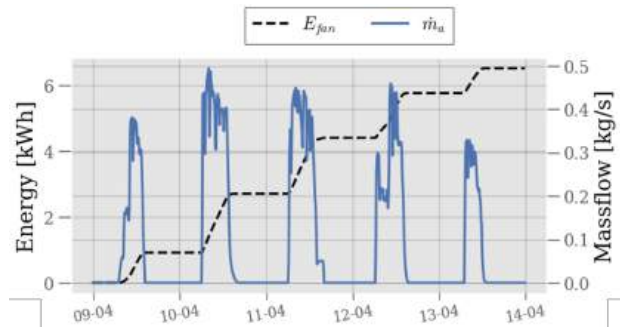


Figure 15: Simulated accumulative exhaust fan power consumption and total ventilation system flow rate in the case study with a PID-controller.

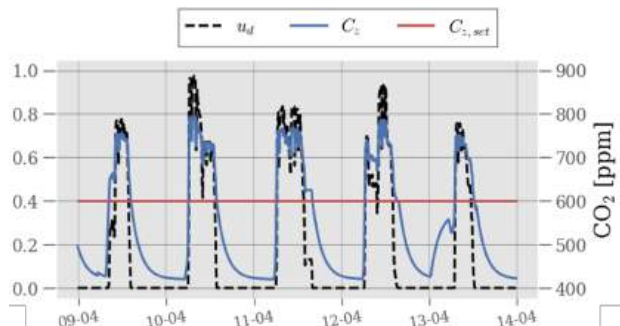


Figure 16: Simulated CO₂-concentration development and damper position in Ø22-511-2 with a PID-controller.

The final controller is a very simple rule-based controller: A scheduled controller, which follows a fixed daily schedule for damper opening regardless of CO₂-levels. It is included in the comparison to gauge the value of having a CO₂-sensor in the spaces. The scheduled controller follows the pattern:

- 0.00 – 6.00 : 0% damper opening
- 6.00 – 16.00 : 50% damper opening
- 16.00 – 20.00 : 25% damper opening
- 20.00 – 24.00 : 0% damper opening

The simulated development in CO₂ - for Ø22-511-2 with the scheduled controller is shown in figure 17.

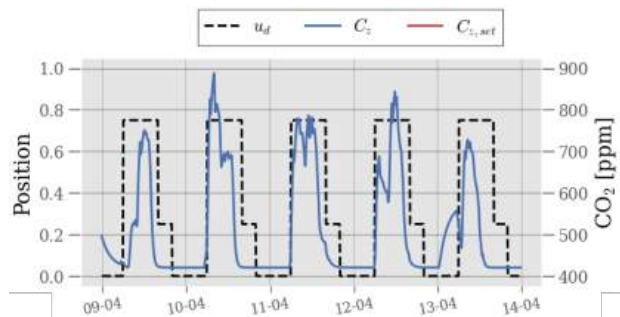


Figure 17: Simulated CO₂-concentration development and damper position in Ø22-511-2 with a scheduled controller.

The total system power consumption and KPI for each of the three controllers are presented in table 3. It is notable that the scheduled controller has a much higher

power consumption than the others while also achieving a worse KPI-value. Between the rule-based controller and the PID-controller the power consumption is almost identical, but the PID-controller obtains significantly higher air quality.

Table 3: Power consumption and KPI for each of the simulated damper control strategies in the case study.

Control strategy	Power [kWh]	KPI [ppm·h]
Fixed schedule	22.24	3374
Rule-based	10.71	2238
PID	10.77	1238

Conclusion and Future Work

This paper presented a framework for ventilation system modelling intended for building digital twins and applied it to a small case study with limited available building data. The results indicated that the developed model was fairly accurate despite the heavy reliance on parameter estimation. They did however also show some discrepancies between predictions and measurements which were difficult to explain and address without complete documentation or long testing periods, as should be expected for the very dynamic operation nature of the demand-controlled ventilation system.

From the control strategy comparison it was found that a basic ventilation scheduled resulted in a much higher energy consumption than CO₂-dependent controllers without benefiting the indoor climate, suggesting that installation of CO₂-sensors is worthwhile for most ventilation systems. The PID-controller outperformed the rule-based controller as expected, but only slightly.

As of this writing the presented modelling framework is still under development. The next step is to combine the current models with thermal components models, including the preheating and precooling coils and the heat recovery unit. Thus, a full-scale ventilation system digital twin model is developed. Another line of improvement for this work is considering FMU connections and the use of built-in parameter estimations. The considered case study was sufficient to show a functional modelling concept, but the framework should be applied to larger case studies where the viability of parameter estimations and black-box modelling can be tested more thoroughly.

References

- Big Ladder Software (2021). *Energyplus version 9.6.0 documentation*.
- Bjørnskov, J. and M. Jradi (2023). An ontology-based innovative energy modeling framework for scalable and adaptable building digital twins. *Energy and Buildings*.
- Bortolini, R., R. Rodrigues, H. Alavi, L. F. D. Vecchia, and N. Forcada (2022). Digital twins; applica-

tions for building energy efficiency: A review. *Energies* 15(19).

Buildings Performance Institute Europe (2022). *Report on the evolution of the European regulatory framework for buildings efficiency*.

Clausen, A., K. Arendt, A. Johansen, F. Sangogboye, M. Kjærgaard, C. Veje, and B. Jørgensen (2021, 09). A digital twin framework for improving energy efficiency and occupant comfort in public and commercial buildings. *Energy Informatics* 4, 40.

Grievess, M. (2015, 03). Digital twin: Manufacturing excellence through virtual factory replication.

Jradi, M., F. C. Sangogboye, C. G. Mattera, M. B. Kjærgaard, C. Veje, and B. N. Jørgensen (2017). A world class energy efficient university building by danish 2020 standards. *Energy Procedia* 132, 21–26. 11th Nordic Symposium on Building Physics, NSB2017, 11-14 June 2017, Trondheim, Norway.

Kapalo, P., F. Domnița, C. Bacoțiu, and N. Spodyniuk (2018). The impact of carbon dioxide concentration on the human health - case study. *Journal of Applied Engineering Sciences* 8(1), 61–66.

Macarulla, M., M. Casals, M. Carnevali, N. Forcada, and M. Gangoellis (2017). Modelling indoor air carbon dioxide concentration using grey-box models. *Building and Environment* 117, 146–153.

Modelica (2021). *Modelica buildings library documentation: Buildings.Fluid.Actuators.Dampers.PressureIndependent*.

SciPy (2023). *Scipy documentation: scipy.optimize.minimize*.

Veeranna, S. and U. Yarangatti (2010, 01). Analysis and implementation of discrete time pid controllers using fpga. *International Journal of Electrical and Computer Engineering* 2, 71–82.

Yang, L., X. Wang, M. Li, X. Zhou, S. Liu, H. Zhang, E. Arens, and Y. Zhai (2020). Carbon dioxide generation rates of different age and gender under various activity levels. *Building and Environment* 186, 107317.

Zaballos, A., A. Briones, A. Massa, P. Centelles, and V. Caballero (2020). A smart campus' digital twin for sustainable comfort monitoring. *Sustainability* 12(21).

Implementation and demonstration of an automated energy modeling framework for scalable and adaptable building digital twins based on the SAREF ontology

Jakob Bjørnskov¹, Muhyiddine Jradi¹

¹SDU Center for Energy Informatics, University of Southern Denmark, Odense, Denmark

Abstract

The building sector is currently seeing rapid digitalization through IoT sensor and smart meter networks as well as the development of more advanced building operation algorithms. In this work, a framework for automated energy model development is developed and assessed as a backbone for effective building digital twin applications. The framework is based on the SAREF ontology and its extensions, SAREF4BLDG and SAREF4SYST to represent the most important devices and components in the building, how they behave, and how they interact. Each component is associated with a data-driven model that is calibrated through operational data collected from the building. The framework is demonstrated and evaluated in a case study building considering a teaching room, where continuous commissioning and scenario testing are investigated as services.

Key innovations

1. The implementation of an innovative energy modeling framework based on the SAREF ontology is demonstrated in a real case study
2. A physics-informed LSTM model for indoor temperature forecasting is proposed and implemented to enable better generalization and accuracy
3. Two digital twin services are proposed and demonstrated in the context of the case study

Practical implications

The structure and relations provided by the SAREF ontology can effectively be used in combination with data-driven modeling and parameter estimation techniques to form a flexible and scalable energy model suitable for digital twins.

Introduction

The increasing availability of operational data from buildings favors the use of scalable and automated data-driven modeling methods. In this regard, recent studies introduced the concept of a Building Digital Twin Lu et al. (2019); Zhao et al. (2022), serving as a digital replica that senses and adapts according to the behavior of the actual building using the installed IoT devices. Such digital twins can be used to deliver various services such as performance monitoring, fault detection, advanced opera-

tional planning, and optimization.

In recent years, the potential of digital twins in yielding useful insights and delivering crucial services has been demonstrated in multiple sectors. However, there is still not a clear consensus on how a digital twin for a building application should be represented and which services it should deliver. Although Building Information Models (BIM) have played a major role in the digitalization of buildings, it is argued by Boje et al. (2020) that status quo BIMs are not compatible with IoT-integration due to their legacy data standards and originally intended use during the design and construction phase.

This work demonstrates the implementation of an innovative framework that builds upon the SAREF ontology, which spans across multiple domains including smart cities. The proposed framework is implemented and assessed by considering a case-study of a classroom in a university building. The building is serving as a living lab, with very detailed submetering on various levels. The classroom is equipped with various sensors and monitoring devices, measuring damper, shading, and radiator valve positions as well as temperature and CO₂-concentration. The collected data is used to automatically calibrate the components associated with the space. This includes a space model that accurately predicts indoor temperature, as well as models for valves, space heaters, and controllers.

The work presented in this paper is part of the international research project 'Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle'. The developed platform targets three high-level services; I) A user-friendly interface allowing building owners to better understand how their building behaves and performs; II) Automated Performance monitoring and continuous commissioning for improved building operation and smarter facility management; III) Model-based scenario testing and operational optimization to enable planning support.

Case study

The considered case study is a classroom located on the second floor of a Danish university building as shown in Figure 1. The building was built in 2015 and attained the Danish building class 2020 as one of the first buildings in Denmark. It has a total floor area of 8500 m² with four stories of which the ground, first, and second floor contains

offices, study zones, and classrooms while the basement contains storage rooms and technical installations.

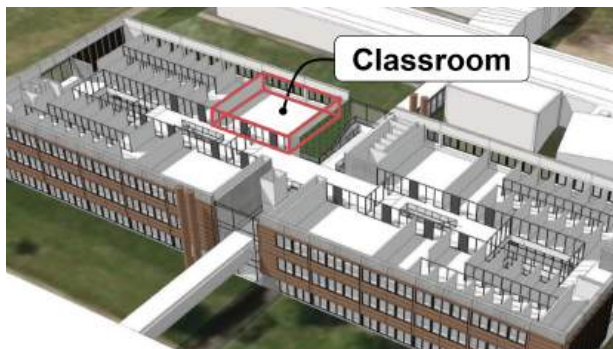


Figure 1: Section overview of the second floor where the considered classroom is located.

The building is divided into four symmetrical quadrants each supplied by a balanced ventilation system. Each system has one heat recovery unit, one heating coil, and the supply air temperature is under normal conditions controlled based on the exhaust air temperature according to a piecewise linear curve. Similarly, the supply water temperature to the radiators is controlled based on the outdoor air temperature measured by a weather station placed on the roof of the building.

The considered classroom is heated with radiators and the supply air flowrate is controlled through Demand Controlled Ventilation (DCV), where the damper opening is actuated based on the measured CO₂ concentration in the room. Shades are installed on the building facade to avoid overheating during the summer months. The actuation signal of the space heater valves, dampers, and shades is controlled centrally by the Building Management System (BMS).

The classroom is equipped with multiple sensors, namely, temperature, CO₂, valve position, damper position, and shade position. In addition, a heat consumption meter is installed on the radiators.

Modeling framework

Through the presented case-study, this paper aims to demonstrate the implementation of an energy modeling framework suitable for digital twins of buildings. The modeling methodology builds upon earlier work by the authors Bjørnskov and Jradi (2023), where the Smart Applications REference (SAREF) ontology is used as a backbone and semantic structure. SAREF was developed by Daniele et al. (2015) in close collaboration with the industry to rectify the interoperability issues encountered in the IoT-domain. The ontology has a focus on devices, e.g. a light switch or a sensor with extensions covering multiple domains, e.g. SAREF4BLDG Poveda-Villalón and Garcia-Castro (2020), SAREF4CITY Poveda-Villalón et al. (2020), SAREF4SYST Lefrançois (2019), SAREF4ENER Daniele (2020), etc. Here, SAREF4BLDG is an ontology extension dedicated to the building domain, which classes are a subset of the

Industry Foundation Classes (IFC) standard. This is considered a significant advantage as it potentially allows for direct mapping of properties and design values for the different building components such as coils, space heaters, dampers, etc.

In Figure 2, the overall Digital Twin concept is shown with its high-level components. The digital twin concept considered in this work follows the definition used by Grieves (2015) and Boje et al. (2020), i.e. that the digital twin concept is composed of three main constituents; the physical system, the virtual system, and a flow of data between these systems. The physical system is the asset to be managed, where the IoT devices play a major role, as they allow for continuous and automated data collection and feedback that can be processed in the virtual system. The role of the virtual system is to mimic the physical system as closely as possible through an accurate dynamic simulation model to enable different services such as continuous commissioning, operational optimization, and scenario testing. For the implementation presented in this work, the virtual system consists of three tightly linked components; the model library, the simulation model, and the semantic model.

The semantic model contains information about system topology, relations, and properties using concepts from the SAREF core ontology as well as the SAREF4BLDG and SAREF4SYST extensions. For instance, it might be specified that a space heater of class `s4bldg:SpaceHeater` is located inside a space of class `s4bldg:BuildingSpace` through the `s4bldg:isContainedIn` property.

The model library shown in Figure 2 consists of different component model classes that represent the dynamic behavior of the SAREF and SAREF4BLDG devices contained in the semantic model. Each model class inherits design properties from the SAREF4BLDG device classes and has tuneable parameters that can be adjusted based on data collected from the physical system.

The simulation model also (Figure 2) is obtained through the system topology and relations contained in the semantic model to connect inputs and outputs for the different component models. For the case study classroom, the information flow for the obtained simulation model is shown in Figure 3.

The Out/In labels at the edges shows the translation of output to input between models. As shown in the graph, virtual sensors and meters are placed in the model (green components), e.g. the Space temperature sensor or the Heating meter. Readings from these virtual measuring devices can be directly compared with readings from actual sensors and meters placed in the physical system.

Component Modeling and Calibration

In this section, the modeling and calibration of different model components is described in relation to the case study.

Each model is calibrated based on collected data on the

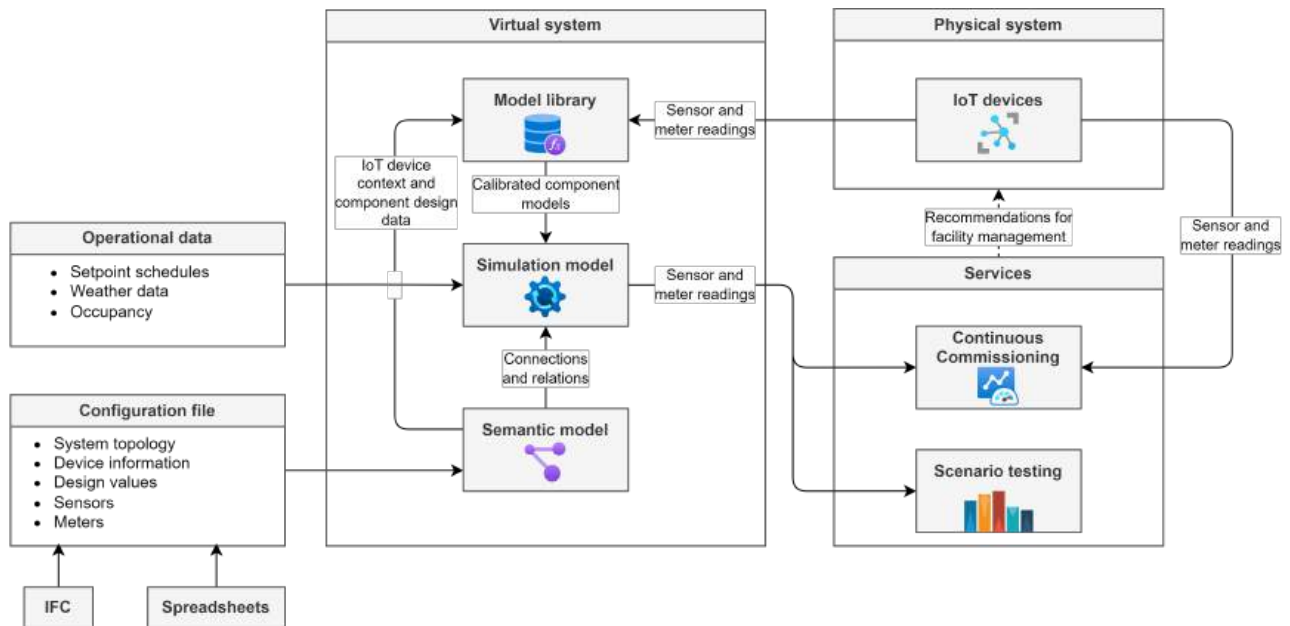


Figure 2: Block diagram showing the high-level components of the digital twin concept.

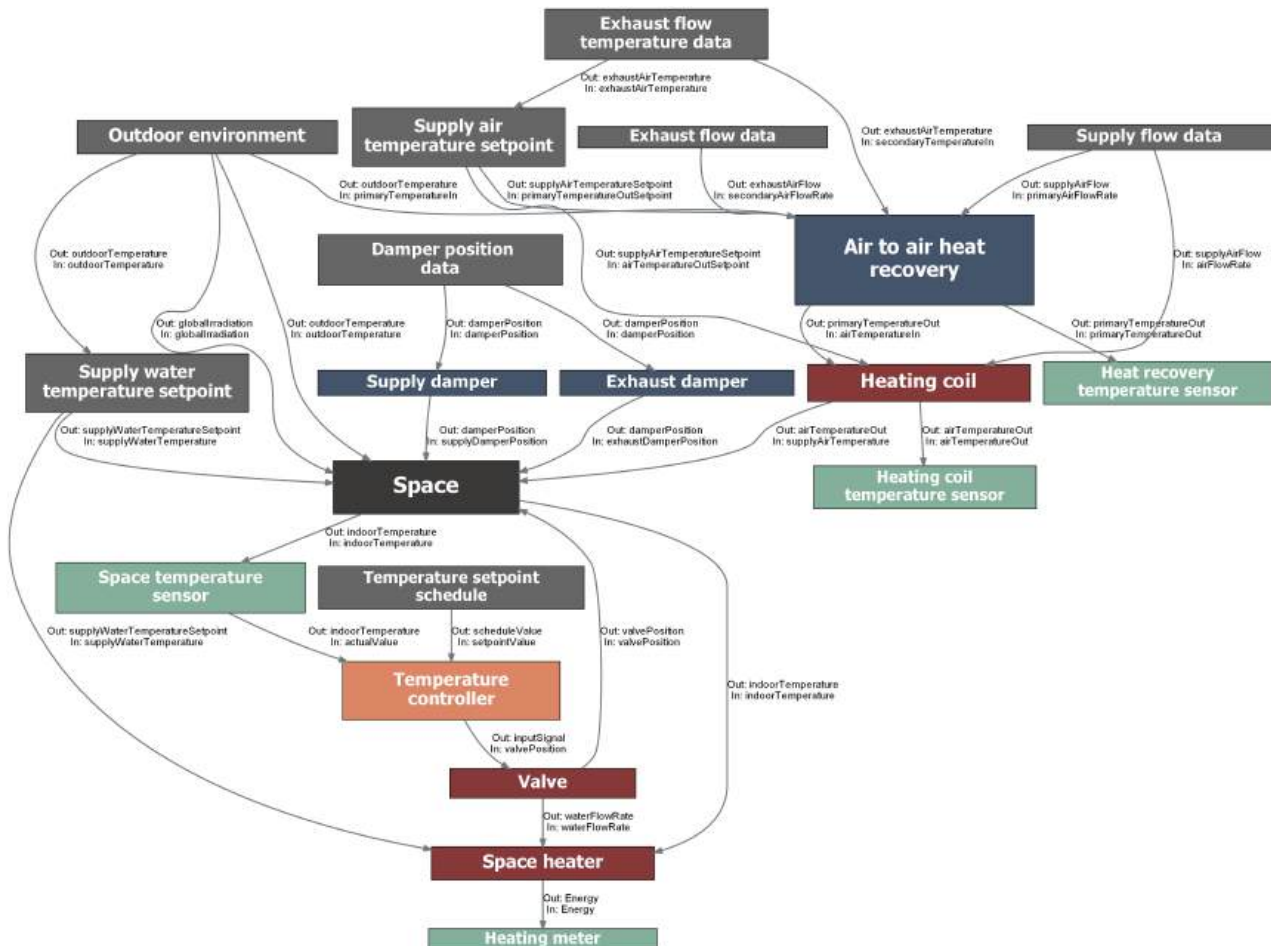


Figure 3: Graph of the generated simulation model for the classroom case study showing the information flow and logical connections between components. The graph has been automatically generated using the open-source Python package available on GitHub Bjørnskov (2023).

model inputs and outputs from the physical system. For parameter estimation of the air-to-air heat recovery, tem-

perature controller, and space heater, a least-square optimization approach is used. The models are described in

detail in Bjørnskov and Jradi (2023).

Air to air heat recovery

The employed air-to-air heat recovery model has four parameters, $\epsilon_{75\%,h}$, $\epsilon_{100\%,h}$, $\epsilon_{75\%,c}$, $\epsilon_{100\%,c}$, each denoting the efficiency for 75% and 100% airflow for heating and cooling mode. The inputs are supply and exhaust airflow, exhaust flow temperature, supply flow temperature setpoint, and outdoor temperature. The model output is the supply outlet temperature. The results are shown in Table 1.

Table 1: Overview of calibration results for the heat recovery model.

Parameter	After calibration	MAE	RMSE
$\epsilon_{75\%,h}$	0.849	0.28 °C	0.47 °C
$\epsilon_{100\%,h}$	0.851		
$\epsilon_{75\%,c}$	0.828		
$\epsilon_{100\%,c}$	0.822		

Temperature controller

The indoor temperature controller model is based on the simple Proportional Integral Derivative (PID) controller. The model has three parameters K_p , K_i , and K_d and takes indoor temperature and indoor temperature setpoint as inputs with valve position as output. The calibration results are shown in Table 2.

Table 2: Overview of calibration results for the temperature controller model.

Parameter	Obtained value	MAE	RMSE
K_p	0.438	0.0083	0.027
K_i	0.251		
K_d	0		

As shown in Table 2, the obtained value for the derivative gain K_d is 0 and the identified controller is thus a PI controller.

Space heater

The space heater model has two parameters, the overall heat transfer coefficient UA -value, and the thermal capacity C_r . The model represents an aggregated model of 5 space heaters placed in the classroom. Model inputs are inlet water temperature, indoor temperature, and water flowrate while the output is aggregated heat consumption. The calibration results are shown in Table 3.

Table 3: Overview of calibration results for the space heater model.

Parameter	After calibration	MAE	RMSE
UA	83.3 W/K	1.09 kWh	1.37 kWh
C_r	$2.73 \cdot 10^6$ J/K		

Space

The model used in this paper, for predicting indoor temperature, is based on earlier work by the authors where Long Short-Term Memory (LSTM) architectures were used Bjørnskov et al. (2022); Bjørnskov and Jradi (2022). LSTMs are a variant of Artificial Neural Network (ANN) architectures which are designed to account for transient and time-dependent behavior due to the cell state vector $c \in \mathbb{R}^n$ and the hidden state $h \in [-1, 1]^n$, which carry information between time steps. In addition, the use of Physics-Informed Neural Networks (PINN) has recently also widened the application of ANN models for thermal control applications Gokhale et al. (2022), due to the increased generalization and robustness of the obtained models. In these models, the model is guided during training by imposing physics-based loss terms in the loss function.

A diagram of the used architecture for this work is shown in Figure 4. The model consists of four units responsible for modeling certain heat transfer mechanics, namely, G(ain), S(paceheater), V(entilation), and R(adiation). Unit G models the heat gained from the surroundings of the space, which is mainly the heat transfer through external walls. Unit S models the heat gained from the space heaters placed in the room. Unit V models the heat gained from the ventilation air flows and unit R models the heat gained from solar irradiation.

Each unit has two sequential LSTM units, A and B, where LSTM_A receives the model inputs for timestep $t-1$, and outputs its hidden state $h_{A,t}$ for timestep t . This hidden state is then used as input to LSTM_B, which outputs its hidden state $h_{B,t}$, i.e. the indoor temperature change caused by the heat transfer mode modeled by the unit. Finally, all four temperature change predictions are added together to obtain $\Delta T_{z,t}$. In order to impose physical constraints on the model, a physics-based loss term is associated with each of the units, G, S, V, R.

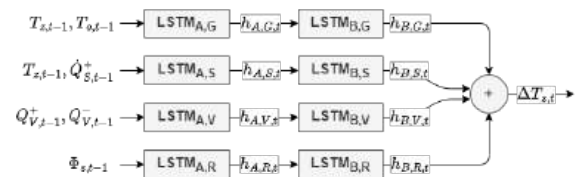


Figure 4: Physics-informed LSTM architecture for predicting indoor temperature.

For unit G, there are two inputs, the indoor temperature $T_{z,t-1}$ and the outdoor temperature $T_{o,t-1}$. The loss term is given in Equation 1.

$$\mathcal{L}_{G,t} = \text{ReLU}\left(\frac{\partial h_{B,G,t}}{\partial T_{z,t-1}}\right) + \text{ReLU}\left(-\frac{\partial h_{B,G,t}}{\partial T_{o,t-1}}\right) \quad (1)$$

The first term in Equation 1, penalizes the model if the indoor temperature change caused by the surroundings

$h_{B,G,t}$ grows with indoor temperature $T_{z,t-1}$. The second term in Equation 1, penalizes the model if $h_{B,G,t}$ decreases with outdoor temperature $T_{o,t-1}$. ReLU is the Rectified Linear Unit function, which is defined as $\text{ReLU}(x) = \max(0, x)$.

The two inputs of unit S are $T_{z,t-1}$ and a feature describing the inlet water flow energy to the spaceheater, $\dot{Q}_{S,t-1}^+ = T_{w,t-1}u_{v,t-1}$, where $T_{w,t-1}$ is the inlet water temperature and $u_{v,t-1}$ is the valve position of the spaceheater. Under the assumption that water flow is proportional to valve position, this feature is equivalent to inlet energy after min-max scaling.

$$\mathcal{L}_{S,t} = \text{ReLU}\left(\frac{\partial h_{B,S,t}}{\partial T_{z,t-1}}\right) + \text{ReLU}\left(-\frac{\partial h_{B,S,t}}{\partial \dot{Q}_{S,t-1}^+}\right) + \text{ReLU}\left(-h_{B,S,t}\right) \quad (2)$$

In Equation 2, the first term penalizes the model if the indoor temperature change caused by the space heater $h_{B,S,t}$ grows with indoor temperature $T_{z,t-1}$. The second term in Equation 2, penalizes the model if $h_{B,S,t}$ decreases with the energy added to the space heater $\dot{Q}_{S,t-1}^+$.

Unit V has two inputs, and similar to the space heater, the two inputs describe the inlet air flow energy $\dot{Q}_{V,t-1}^+ = T_{a,t-1}u_{d,t-1}$ and outlet air flow energy $\dot{Q}_{V,t-1}^- = T_{z,t-1}u_{d,t-1}$, respectively. Here, $T_{a,t-1}$ is the supply air temperature, and $u_{d,t-1}$ is the damper position. The associated loss with unit V is given in Equation 3.

$$\mathcal{L}_{V,t} = \text{ReLU}\left(\frac{\partial h_{B,V,t}}{\partial \dot{Q}_{V,t-1}^-}\right) + \text{ReLU}\left(-\frac{\partial h_{B,V,t}}{\partial \dot{Q}_{V,t-1}^+}\right) \quad (3)$$

In Equation 3, the first term penalizes the model if the indoor temperature change caused by ventilation $h_{B,V,t}$ grows with removed energy $\dot{Q}_{V,t-1}^-$. The second term in Equation 3, penalizes the model if $h_{B,V,t}$ decreases with added energy $\dot{Q}_{V,t-1}^+$.

$$\mathcal{L}_{R,t} = \text{ReLU}\left(-\frac{\partial h_{B,R,t}}{\partial \Phi_{t-1}}\right) + \text{ReLU}\left(-h_{B,R,t}\right) \quad (4)$$

From Equation 4 it follows that the model is penalized if the indoor temperature change caused by solar irradiation $h_{B,R,t}$ decreases with solar irradiation Φ_{t-1} , while the second term penalizes the model if $h_{B,R,t}$ becomes negative.

The total loss for a given time sequence with the number of time steps N is then expressed in Equation 5 as the sum of unit losses and the squared residual of the observed temperature change $\Delta \hat{T}_{z,t}$ and predicted temperature change $\Delta T_{z,t}$.

$$\mathcal{L} = \frac{1}{N} \sum_{t=1}^N \left[(\Delta \hat{T}_{z,t} - \Delta T_{z,t})^2 + \mathcal{L}_{G,t} + \mathcal{L}_{S,t} + \mathcal{L}_{V,t} + \mathcal{L}_{R,t} \right] \quad (5)$$

With the loss function defined in Equation 5, the model has been trained and evaluated on winter data using the same preprocessing and testing methods as used in earlier work Bjørnskov and Jradi (2022). In Table 4, the performance is shown on the test set on temperature prediction. Note that these performance measures are for temperature prediction on 24-hour horizons using a closed-loop configuration of the architecture shown in Figure 4.

Table 4: Overview of calibration results for the space model.

Parameter	Obtained value	MAE	RMSE
-	-	0.36 °C	0.54 °C

Digital twin service applications

In this section, it is demonstrated how the obtained model for the case study can be used to deliver different services. The services explored in this paper are continuous commissioning and scenario testing.

Continuous Commissioning and performance monitoring

To demonstrate the utility of the modeling framework for continuous commissioning, a period of 2 weeks has been selected, which includes a significant change in the operation of the case study building and classroom. This change in operation is caused by the current energy crisis in Europe and the resulting 19 °C indoor temperature policy of public buildings recently introduced by the Danish government. One major implication for the considered case study building and classroom is that the indoor temperature is reduced from 21 °C to 19 °C and the supply air temperature setpoint is reduced to constant 19 °C during the chosen period.

The results for the commissioning case are shown in Figure 5. For all four plots, the red line indicates the time stamp where the supply air temperature setpoint was changed to constant 19 °C in the physical system. In Figure 5a, Figure 5b, and Figure 5c the upper plots show the readings from the physical and virtual temperature sensors, while the lower plots show the performance gap represented through a moving average of the residual with a window size of 1 day. This moving average is used to obtain a more robust and stable metric for anomaly detection than the raw residual and avoid stochastic fluctuations. Figure 5d shows a binary anomaly signal which is 1 if the moving average is outside of the 1°C error band and otherwise 0.

Generally, as seen from Figure 5a, Figure 5b, and Figure 5c, readings from the virtual and physical sensors

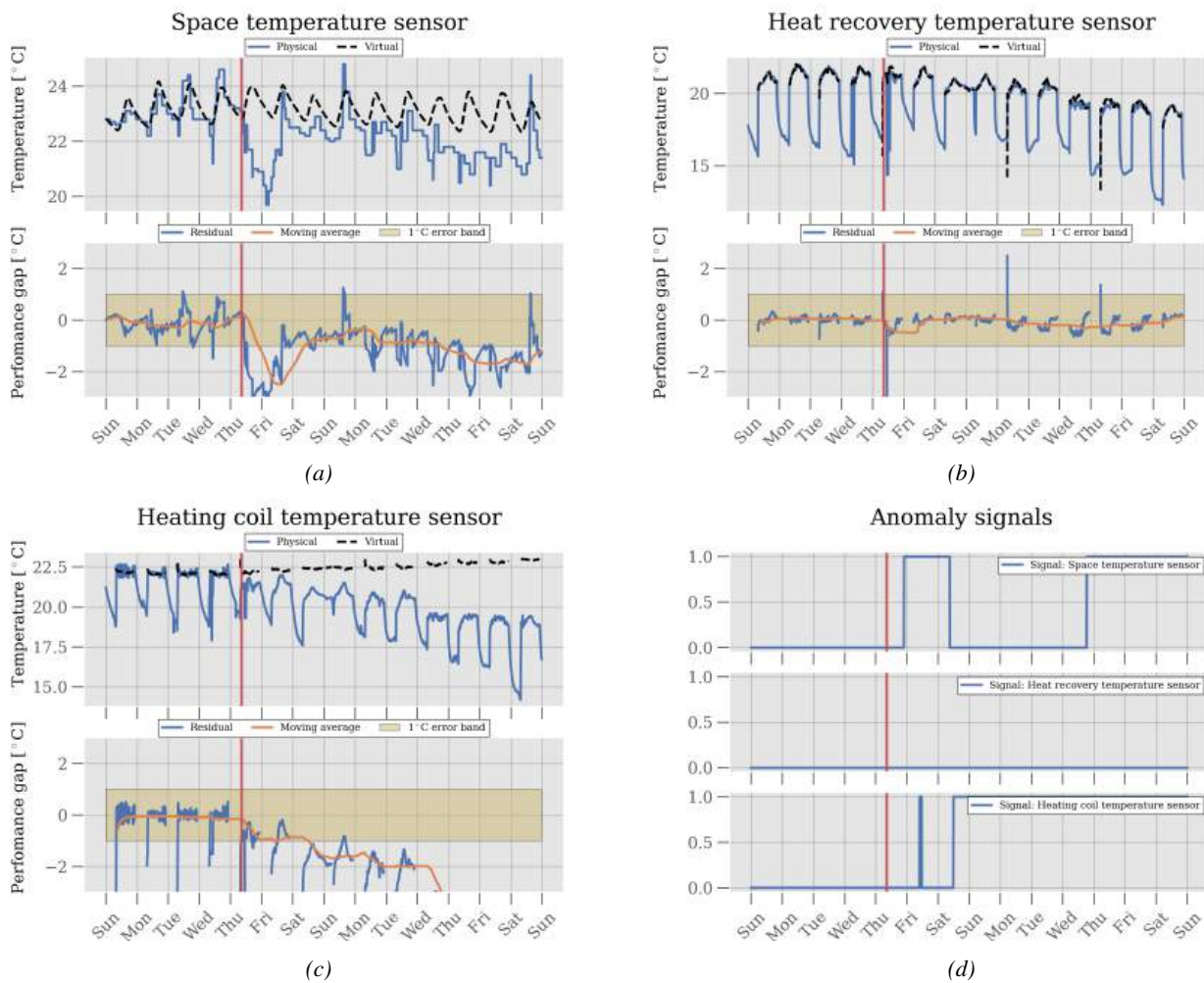


Figure 5: Continuous commissioning of a 2-week period. The red line marks the time stamp where the supply air temperature setpoint was changed to a constant 19 °C in the physical system. (a) Measurements, predictions, and residuals for the Space temperature sensor. (b) Measurements, predictions, and residuals for the Heat recovery sensor. (c) Measurements, predictions, and residuals for the Heating coil sensor. (d) Binary anomaly detection signals for the sensors, where 0 indicates expected operation and 1 indicates anomalies.

are in agreement during the first four days with a performance gap within the 1°C error band. However, after the change of operation, the performance gap between the Indoor temperature and the Heating coil sensor starts to drift outside of the error band. For the virtual space, the simulated temperature is higher due to the warmer temperature of the supply air, while the physical space receives colder air due to the change in supply air temperature. These differences in supply air temperatures are evident from Figure 5c.

Figure 5d summarizes these observations through the binary anomaly signals. As shown, the anomaly signal changes to 1 for the space sensor directly after the change in supply air due to the unexpected drop in temperature, which continue for about 1 day. Afterward, the signal returns to 0 until Thursday in the second week when the anomaly signal again changes to 1. For the Heating coil temperature sensor, the anomaly signal switches to 1 after approximately two days when the measured tempera-

ture falls below 21 °C. The signal does not switch on right away because there is a transitioning period of approximately a week before the supply air temperature reaches 19 °C.

Although the detected anomaly is not a fault, the commissioning case presented in this study demonstrates the potential of the data-driven model and its capability in capturing anomalies and performance drifting as soon as it occurs. This is translated into shorter reaction times and resource savings. In addition, the presented approach can form a basis for a building's continuous commissioning process on different levels, serving as a backbone for effective and robust fault detection and preventive maintenance.

Scenario testing

The second application of the modeling framework and the obtained case study model is scenario testing, which aims to compare different candidate operational scenarios in a zero-risk virtual environment.

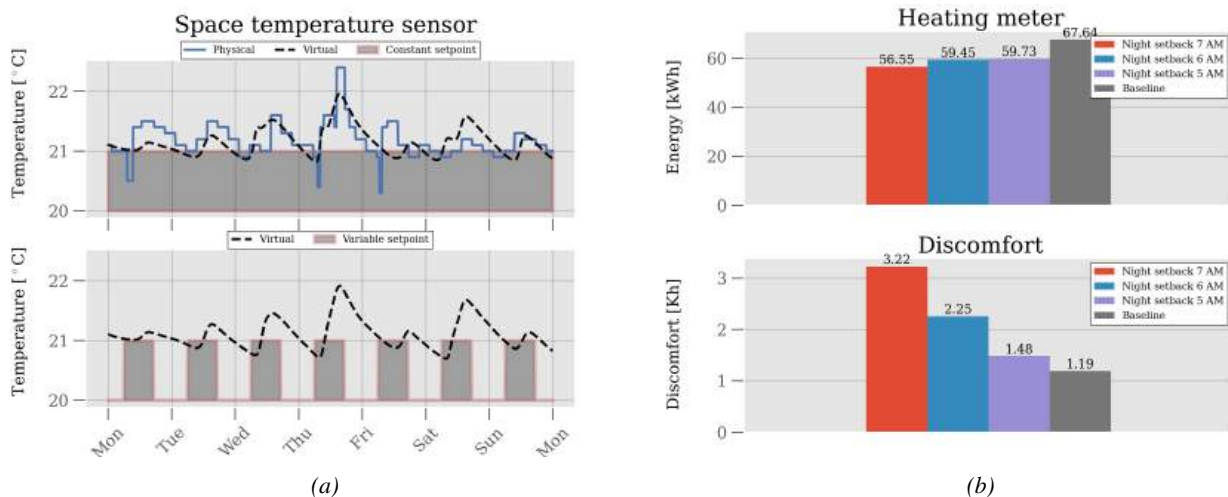


Figure 6: Comparison of four different scenarios for a 1-week period. (a) Comparison between the baseline scenario with constant setpoint (top) and the alternative scenario with varying setpoint (bottom). (b) Total predicted heating consumption on the space heaters (top) and the predicted discomfort (bottom).

During the winter periods, the building has a heating setpoint of constant 21 °C. A common strategy for reducing energy consumption in buildings is to decrease the temperature setpoint in periods with no occupancy, e.g. during nighttime. This strategy is called *night setback*, which must be implemented such that the building is pre-heated before occupation in order to avoid thermal discomfort. However, buildings behave differently under different weather and operational conditions. Exploration of different setpoint strategies and their impact on energy consumption and indoor comfort is therefore important.

In this work, we explore four scenarios, the baseline constant 21 °C setpoint of the classroom, and three variations of night setback scenario with a varying setpoint. The investigated night setback scenarios are to change the setpoint schedule to 21 °C from xx:00 to 17:00 and 20 °C otherwise, where the starting hour xx is varied between 05, 06, and 07. To measure the impact on indoor comfort, we use the same aggregated metric as proposed by Arendt et al. (2020) for measuring thermal discomfort in Kelvin-hours, which captures both the duration and intensity of setpoint violations. Thermal discomfort is only aggregated for hours with occupancy which is assumed to be between 08:00 and 17:00.

The results for the scenario testing are shown in Figure 6. Figure 6a shows the virtual temperature readings for the baseline (top) and the 07 AM night setback scenario (bottom). The baseline also includes the actual measurements from the physical temperature sensor during the period as a reference. As shown, the baseline virtual readings (as well as physical readings) fluctuate around the 21 °C setpoint. In the night setback scenario, the temperature is allowed to drop during the night until 07:00 where the temperature setpoint is again increased to 21 °C. Figure 6b shows the resulting difference in heating consumption (top) and discomfort (bottom). As shown there is a clear tradeoff between reduced energy consump-

tion and discomfort for the chosen scenarios. The sooner space heating is initiated, the lower the energy consumption. However, this comes at the cost of higher discomfort levels. Typically, maintenance of thermal comfort levels is valued higher than decreases in energy consumption. Therefore, the most viable of the three-night setback alternatives is the 5 AM scenario, which impacts the thermal discomfort the least by 0.29 Kh but saves around 12 % in heat consumption.

Conclusion

In this paper, the implementation of an innovative energy modeling framework was demonstrated in a case study. The framework builds upon the SAREF ontology and its extensions to form the backbone of a digital twin platform. The building digital twin is considered as a collection of components such as spaces, heaters, controllers, fans, etc., that can be directly related to their real counterparts, enabling detailed performance monitoring and fault detection at both component and systems-level. This further enables building owners to test different operational strategies, and how they affect the building before they are implemented. Employing the framework, two digital twin services were demonstrated. Continuous commissioning and performance monitoring was demonstrated on individual components by comparing virtual and physical sensor and meter readings, i.e. predictions and measurements for a certain period. Here, anomaly behavior was correctly detected on a period with a change in building operation. As a second service, scenario testing was demonstrated by comparing baseline operation with three alternative operation strategies in terms of indoor comfort and consumed energy.

This work provides a proof-of-concept of the modeling framework and the potential applications and services it can deliver. The use of a well-established ontology ensures the scalability of the approach and that the de-

veloped framework could potentially be expanded and adapted to clusters of buildings or cities in the future. For future work, it is intended to further scale up the efforts leading to the development of a whole-building digital twin of the case-study building. In addition, it is expected that the delivered services could be expanded to consider also operational optimization and automated actuation of components or change of setpoints at a later stage.

Acknowledgment

This work is carried out under the ‘Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle’ project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64021-1009.

References

- Arendt, K., A. Clausen, C. Mattera, M. Jradi, A. Johansen, C. Veje, M. Kjærgaard, and B. Jørgensen (2020). Multi-objective model predictive control framework for buildings. In *Proceedings of building simulation 2019*. Rome (Italy), 02-04 September 2019.
- Bjørnskov, J. (2023). Energy modeling framework. <https://github.com/JBjoernskov/Twin4Build>, Last Accessed: 03-23-2023.
- Bjørnskov, J. and M. Jradi (2022). A modular thermal space coupling approach for indoor temperature forecasting using artificial neural networks. In *Proceedings from BSO2022: Building Simulation and Optimisation Conference*. Bath (UK), 13-14 December 2022.
- Bjørnskov, J. and M. Jradi (2023). An ontology-based innovative energy modeling framework for scalable and adaptable building digital twins. *Energy and Buildings* 292, 113146.
- Bjørnskov, J., M. Jradi, and C. Veje (2022). A fully automated and scalable approach for indoor temperature forecasting in buildings using artificial neural networks. In *Proceedings from BSO2022: Building Simulation Applications Conference*. Bolzano (Italy), June 29-July 1 2022.
- Boje, C., A. Guerriero, S. Kubicki, and Y. Rezgui (2020). Towards a semantic construction digital twin: Directions for future research. *Automation in Construction* 114, 103179.
- Daniele, L. (2020). Saref4ener: an extension of saref for the energy domain. [Online]. Available: <https://saref.etsi.org/saref4ener/>, Last Accessed: 03-23-2023.
- Daniele, L., F. den Hartog, and J. Roes (2015, 08). Created in close interaction with the industry: The smart appliances reference (saref) ontology. pp. 100–112.
- Gokhale, G., B. Claessens, and C. Develder (2022). Physics informed neural networks for control oriented thermal modeling of buildings. *Applied Energy* 314, 118852.
- Grieves, M. (2015, 03). Digital twin: Manufacturing excellence through virtual factory replication.
- Lefrançois, M. (2019). Saref4syst: an extension of saref for topology of systems and their interconnections. [Online]. Available: <https://saref.etsi.org/saref4syst/>, Last Accessed: 03-23-2023.
- Lu, Q., A. K. Parlikad, P. Woodall, X. Xie, Z. Liang, E. Konstantinou, J. Heaton, and J. Schooling (2019, 10). Developing a dynamic digital twin at building and city levels: A case study of the west cambridge campus. *Journal of Management in Engineering* 36.
- Poveda-Villalon, M., R. Garcia-Castro, and P. Espinoza-Arias (2020). Saref extension for smart city. [Online]. Available: <https://saref.etsi.org/saref4city/>, Last Accessed: 03-23-2023.
- Poveda-Villalón, M. and R. Garcia-Castro (2020). Saref extension for building. [Online]. Available: <https://saref.etsi.org/saref4bldg/>, Last Accessed: 03-23-2023.
- Zhao, J., H. Feng, Q. Chen, and B. Garcia de Soto (2022). Developing a conceptual framework for the application of digital twin technologies to revamp building operation and maintenance processes. *Journal of Building Engineering* 49, 104028.

Data-driven digital twining of ventilation systems for performance optimization: A university building case study

Andres Sebastian Cespedes Cubides – University of Southern Denmark, Denmark –
acub@mmmi.sdu.dk

Jakob Bjørnskov – University of Southern Denmark, Denmark – jabj@mmmi.sdu.dk

Muhyiddine Jradi – University of Southern Denmark, Denmark – mjr@mmmi.sdu.dk

Abstract

This study introduces the creation and application of a data-driven digital twin for building ventilation systems, focusing on a university building as a case study. It employs a grey-box energy modelling framework to accurately forecast, simulate, and monitor the ventilation system's efficiency under diverse conditions. The study collects a substantial dataset to reflect various usage patterns and environmental influences, which serves to test and validate the component models of the ventilation system. These models are integrated into a digital twin platform, providing a comprehensive overview of the system's performance and critical indicators in real time. The digital twin facilitates informed decision-making for facility managers regarding energy consumption, inefficiency identification, and the recommendation of custom retrofitting actions specific to the building's characteristics and use. The findings confirm that digital twins are effective as a tool to continuously commission and detect anomalies in buildings. The study offers a ventilation modelling and monitoring method capable of recognizing rule-based control behaviours and changes in systems that occur in cycles, like system shifts from winter to summer, and can estimate total air mass flow rate with a correlation exceeding 80%.

1. Introduction

The European Union (EU) is addressing its energy and environmental objectives for 2030 and 2050 by focusing on the building sector, which accounts for nearly 40% of its energy consumption (European commission, 2018). According to the International Energy Agency (IEA), this sector is also responsible for about 36% of total emissions, broken down into

residential buildings (22%), non-residential buildings (8%), and construction projects (6%) (IEA, 2019). Digital twin technology is identified as a key innovation for improving building operations, particularly through optimizing ventilation systems which play a crucial role in maintaining indoor air quality and constitute a substantial part of building energy usage.

Recent advancements in computer science have led to the integration of digital technologies like Building Information Models (BIM) and their advanced form, Digital Twins, into building management (Lu et al., 2020). Digital Twins, particularly when combined with sensor data, enhance building energy efficiency through various methods, including model predictive control (MPC) (Smarra et al., 2018) and facilitating real-time energy-saving decisions (Agostinelli et al., 2021). These technologies enable precise and efficient modelling of a building's energy systems. A significant challenge in this field is modelling ventilation systems that must account for the variability in control actions due to different conditions such as seasonality and manual adjustments (Zhang et al., 2022). This research focuses on developing a Digital Twin for the OU44 building at the University of Southern Denmark. The building serves as a live laboratory and is equipped with CO₂ sensors and air diffuser damper position sensors. The study combines data-driven and physics-based methods to precisely estimate energy consumption and conditions in specific rooms. This Digital Twin platform offers real-time data on ventilation performance and energy utilization, employing artificial neural networks to understand cyclic behaviours in the system.

The development of the digital twin uses the *Twin4Build* framework (Bjørnskov et al., 2023), which supports grey-box energy modelling of various system components. Additionally, this research investigates continuous monitoring and anomaly detection, addressing the challenges of data collection and system integration in older systems. The findings highlight the profound impact of Digital Twin technology in building management, showcasing through the OU44 case study how such technologies can improve decision-making, enhance monitoring capacity, and adapt to other buildings equipped with basic sensors, demonstrating their broad applicability in the field.

2. Grey box, ontology-based modelling.

Grey-box modelling merges theoretical knowledge and data-driven methods to produce models that balance comprehensibility with accurate system dynamics representation, especially useful for partially understood systems. It incorporates both partial theoretical knowledge and empirical data. Ontology-based modelling enhances this approach by using structured frameworks called ontologies to organize information and define relationships specific to the domain, thereby improving interoperability and facilitating knowledge reuse across different applications. The digital twin concept organizes these models and tools to serve physical asset designers and operators, defining a digital twin as a collection of digital models that predict specific outcomes, supported by a data-acquisition system for real-time physical-digital interaction. A diagram of this model-asset interaction is depicted in Figure 1. The Twin4build framework leverages the *SAREF* core ontology and its extensions *SAREF4BLDG* and *SAREF4SYST* (SAREF, 2020).

It is composed of 5 main classes which provide a streamlined modelling flow:

- **Model:** Represents the simulation model, composed of 1 or more component models.
- **Simulator:** Simulates a model instance for a given period.
- **Monitor:** This class offers methods to

analyse estimated and measured data to assess performance and detect anomalies.

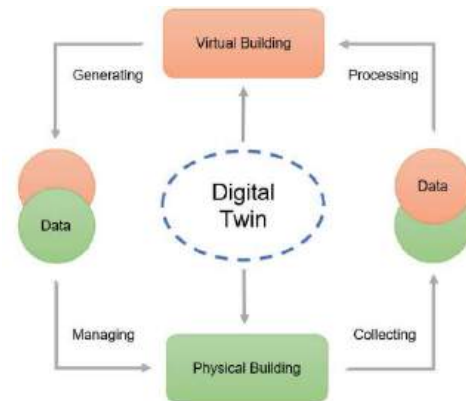


Fig. 1 – Digital twin diagram. It uses sensor data in tandem with digital models to provide a variety of services for building operators and designers (Jradi and Bjørnskov, 2023).

- **Evaluator:** Can evaluate and compare different Model instances based on user-defined quantities of interest.
- **Estimator:** Provides methods for performing parameter estimation and sensitivity analysis using a model instance and user-defined parameters.

3. Methodology

Figure 2 shows the abstract process carried out to develop the ventilation model. The methodology encompasses an initial phase of characterization and data acquisition, followed by the application of historical data to train a machine learning algorithm. This algorithm is designed to replicate the rule-based control actions for each ventilated room. Subsequently, the approach employs historical data in conjunction with model estimates to conduct parameter estimation and refine the model's accuracy through fine-tuning.

The next sections show a general diagram of the model, with inputs and outputs. Followed by a description of a room model which is replicated for each room inside the ventilation model, the room model contains a controller and a damper model. The controller model attempts to mimic the heterogenous behaviour of the rule-based controller

that controls the damper position of each room.

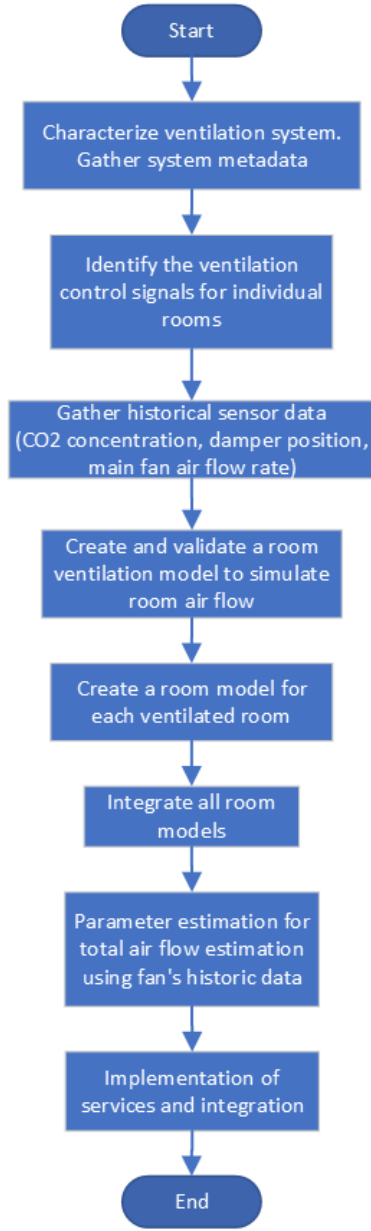


Fig. 2 – General modelling methodology.

3.1 Ventilation system model

This study considers one of the 4 nearly identical ventilation systems in the OU44 building. This subsystem provides air to a quarter of the areas on the building, distributed in 20 rooms with demand-controlled ventilation (DCV) and other 10 auxiliary, small rooms like bathrooms and copy rooms with a constant air volume (CAV). These rooms span all 4 levels of the building.

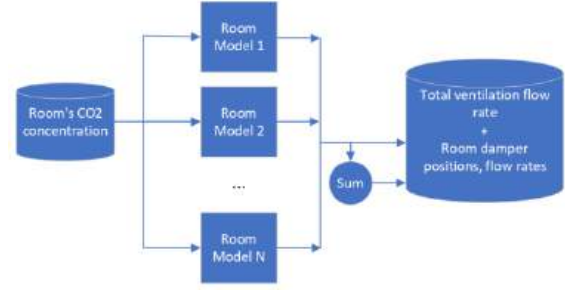


Fig. 3 – Ventilation system model

Figure 3 illustrates a model of the ventilation system. The facility's Building Management System (BMS) monitors CO₂ levels continuously, capturing real-time data every minute from rooms with DCV. This data informs the adjustment of damper positions in each room, which in turn controls the airflow according to a model specific to each room. These airflow calculations are based on the predetermined nominal flow rates designated for each room.

3.1.1 Room model



Fig. 4 – Room ventilation model

The individual room model consists of a CO₂-based controller which uses both CO₂ and time information to provide a damper opening position signal.

The damper model estimates the air flow through the damper as a function of the damper position. Using the model by Huang (Huang, 2011) in its constrained form, the model has two main parameters: a corresponding to a unitless air damping coefficient and $\dot{m}_{a,max}$ which corresponds to the maximum flow rate of the room damper. The damper's air mass flow rate is described by the equation 1, where the constraints for the coefficients c and b are given by equations 2 and 3.

$$\dot{m}_a = ae^{bu_d} + c \quad (1)$$

$$c = -a \quad (2)$$

$$b = \ln\left(\frac{\dot{m}_{a,max}-c}{a}\right) \quad (3)$$

Where:

$\dot{m}_{a,max}$ is the damper's nominal flow rate.

a is a damper coefficient.

u_d is the damper opening position [0-1]

The maximum air flow ($\dot{m}_{a,max}$) rate for each room is taken from the design values given in the buildings blueprints. A default fine-tuned value is used for the second parameter (a) of the model.

3.1.2 Controller model

One prevalent challenge encountered in modelling ventilation systems of buildings, years after they commence operations, stems from the inconsistency in the control strategy across all rooms within the building, which can fluctuate throughout the year. This variability, compounded by the lack of comprehensive historical data on adjustments and configurations applied to the control system over time, often results in the inadequacy of simple rule-based controllers' models to accurately predict the ventilation system's actual behaviour across the entire building. To address this issue, a data-driven methodology was employed, leveraging artificial neural networks, as shown in Figure 5.

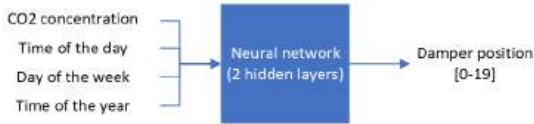


Fig. 5 – Room ventilation controller model

This approach uses both CO2 concentrations and the rooms' damper positions to train an ANN model which attempts to mimic the behaviour of the rule-based controller assigned to each room.

3.1.3 Data embeddings

To capture the configuration of the ventilation modes for each one of the rooms and to validate the controlling signals each one of the 3 main quantities were adapted for its use in the neural network model. Equation 4 presents the discretization used for the damper position value. A continuous value from 0 to 100 is converted to 20 discrete classes with an integer from 0 to 19 for each. This is the output of the neural network.

$$d = \text{round}\left(\frac{D}{100} \cdot 19\right) \quad (4)$$

Where d is the discrete value for the damper position used as output of the neural network and D is the original data point [0,100].

Equation 5 describes the normalized CO₂ concentration, with the original values in ppm, a gaussian normalization (Z-score normalization) is made:

$$z = \frac{CO_{2,i} - \mu_{CO_{2,i}}}{4\sigma_{CO_{2,i}}} \quad (5)$$

Where the mean and standard deviation are calculated with all the data points of each room (i) corresponding to the year 2023. The denominator includes four times the sigma value to ensure that 99.99% of the recorded values fall within the interval (-1,1). This range is considered more suitable for neural networks.

The time variables were divided in three categories: Day of the year, day of the week and time of the day. This division allows the model to have a notion of the effects of seasons, weekdays and day-night cycles which are typically used to define ventilation rule-based controls. To provide insight into the cyclic nature of the first two time-variables, the cyclic embedding presented in equations 6 and 7 was used.

$$x_{sin} = \sin\left(\frac{2\pi x}{P}\right) \quad (6)$$

$$x_{cos} = \cos\left(\frac{2\pi x}{P}\right) \quad (7)$$

Where x are the original time of day and day of year cyclic features, P is the period of the feature (24 for time of day and 365 for day of year). This representation produces 2 variables per cyclic feature that are fed to the neural network as inputs. The day of the week feature is encoded using one-hot-encoding.

Table 1 – Hyperparameters used for the neural networks representing the ventilation controller for each room.

Hyperparameter	Value
Input size	12
Output size	20
Hidden layer 1 size	50
Hidden layer 2 size	100
Learning rate	0.001
Number of epochs	15
Batch size	64

4. Case study

A case study using sensor data from a living lab building from the University of Southern Denmark is considered for the implementation of the developed model. This study focuses on the ventilation system of the building. The ventilation system's digital twin aims to provide continuous commissioning services, anomaly detection and insights into the power consumption of the ventilation fan through the estimation of the total air flow rate and room ventilation conditions.

4.1 OU44 Building

Presented in Figure 6, the OU44 building of the University of Southern Denmark is a multi-purpose building equipped with different sensors throughout the building. Specifically, the main rooms and offices measure indoor CO₂ concentrations, a signal that is used in the demand-controlled ventilation system of the building.

The building has 4 nearly identical ventilation subsystems which operate independently. Each one of these is comprised of an Air Handling Unit, ventilation ducts and room air dampers controlled by a centralized controller.



Fig. 6 – Façade of the UO44 building, SDU Odense

4.2 Ventilation system

This study focuses on the first ventilation sub-system, denominated VE01, where the main system components are shown in the block diagram of figure 7.

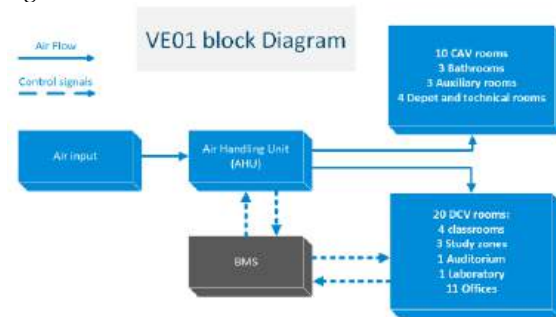


Fig. 7 – Block diagram of one of the ventilation subsystems in UO44, VE01.

5. Results

5.1 Controller model

The data-driven approach attempts to capture the rule-based controller behaviour by mimicking the measured controlling actions without having to manually craft a set of controlling rules for each room in the ventilation system. The accuracy of the Artificial Neural Network (ANN) models is evaluated by comparing them to the actual measured positions of dampers in each room. This accuracy is calculated as the percentage of instances where the ANN model's output precisely matches the discretized position of the damper at every timestamp in the test data. The results are presented in figure 8.

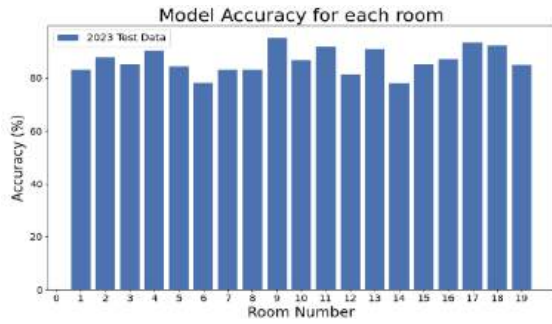


Fig. 8 – The accuracy of the Artificial Neural Network (ANN) models.

The model training process starts with the data embedding described in section 2.1.1 where the CO₂ concentration and time data are pre-processed. Then the data is split into training and test datasets in the following process: Out of the data for the whole 2023 year, the first three weeks of each month are used for training and the remaining days of each month for testing. Additionally, the months of January and February 2024 are used for validation. The compound contribution of the ANN controller trained for each one of the rooms is shown in the Figure 9.

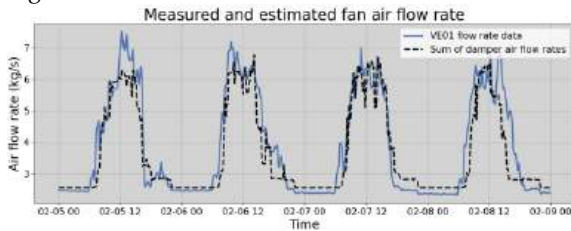


Fig. 9 – ANN-based controller estimation for a week during the month of February 2024, RMSE: 0.6581 kg/s Correlation: 0.9335

The accuracy of the ANN-based controllers is evaluated by comparing two key measurements: the Root Mean Square Error (RMSE) calculated using actual measured positions of the dampers and the RMSE calculated using the damper position signals provided by the ANN controllers. This comparison helps assess the error introduced by the ANN controllers. Table 2 presents the comparison of the RMSE obtained when estimating the total system airflow when the control signal is the measured damper position and the estimated air flow using the position signal from the ANN controller model. RMSE calculated with all validation data from 2024.

Table 2 –RMSE and correlation coefficients of mass air flow rate

with and without ANN controllers.

Control signal	RMSE [kg/s]	Correlation
Measured	0.6910	0.9684
ANN control	0.6581	0.9335
Difference	0.0329	0.0349

5.1.1 Share of air flow from rooms without DCV

The proportion of total airflow to rooms without demand-controlled ventilation was determined by analysing estimated airflow values against measurements from the air handling unit (AHU). The Root Mean Square Error (RMSE) for the initial dataset, covering January 2023, was determined to be 1.566 kg/s. This RMSE is considered indicative of the constant airflow volume in rooms equipped with continuous ventilation. The accuracy of this estimation was further assessed by comparing it against data collected in subsequent periods, as shown in Figure 10.

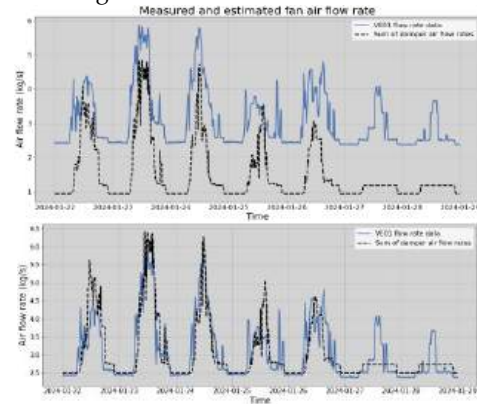


Fig. 10 – (Top) Original air flow rate estimation. RMSE: 1.6468 Correlation: 0.8241. (Bottom) Adjusted air flow rate estimation: RMSE: 0.5045 Correlation: 0.8241

5.2 Continuous monitoring

First, the digital twin is implemented for continuous commissioning service of the ventilation system case study. In this regard, the model enables the tracking and identification of irregularities in both the overall ventilation system and within each individual room. This is done by continuously

comparing an estimated signal with its measured counterpart, calculating a moving average of the error between the two signals and defining a threshold that would trigger an anomaly signal if surpassed. Figure 11 shows the continuous monitoring of the ventilation system's main air flow rate.

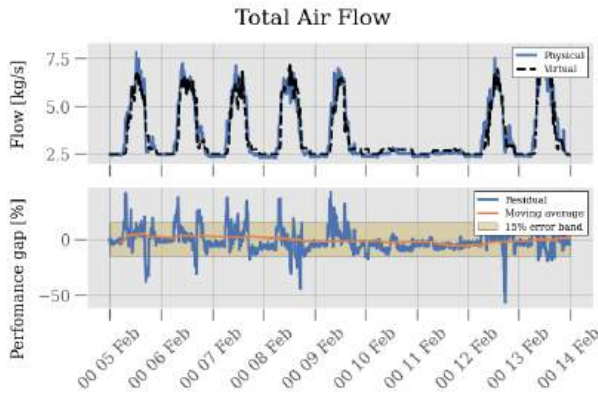


Fig. 11 – Continuous monitoring of total air handling unit inlet air flow rate from the 5th to the 14th of February 2024

The second plot shows the anomaly signal, which is calculated by identifying deviations between the expected and actual damper positions or airflow rates, an anomaly is considered when the error average exceeds 15% for total airflow and 20% for damper positions. In Figure 12, a detailed examination of the monitoring models for all these rooms revealed that the expected control signal deviated in behaviour for 4 rooms. Analysing further, for one of the offices with anomalies, it could be seen that the control signal was manually set to maintain a fully open damper position during work hours, disregarding CO₂ levels. This setup was modified in the following year, as depicted in Figure 13, to allow for adjustments in response to measured CO₂ levels, thereby optimizing ventilation performance and indoor air quality.

6. Conclusion

This study utilizes grey-box modeling and ontology-based methods to blend empirical data with theoretical insights for predicting and simulating ventilation system performance under various scenarios. This approach achieves a balance

between interpretability and accuracy, optimizing operations without fully depending on complex theoretical processes.



Fig. 12 – Anomaly signals for the rooms of the ventilation system

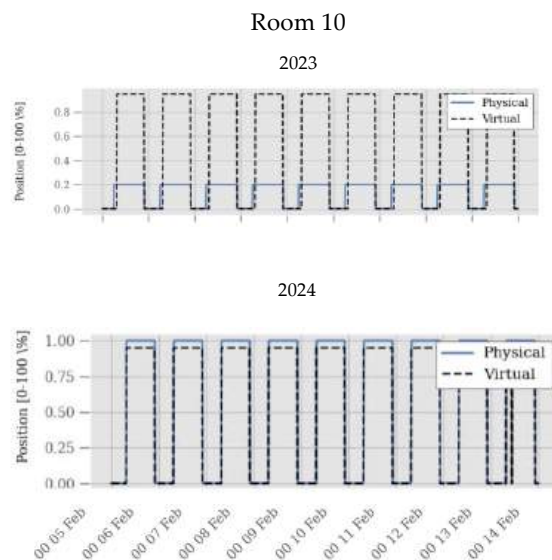


Fig. 13 – Anomalous change in the damper controller for room 10. The behaviour of the controller is different compared to the same period in the previous year.

This approach achieves a balance between

interpretability and accuracy, optimizing operations without fully depending on complex theoretical processes. It replicates rule-based controllers in larger buildings by training a neural network with control signals, adjusting to cyclic patterns and periodic control strategy changes efficiently. Although effective, it can inadvertently include isolated configuration changes, which could be mitigated with larger datasets.

This research also underscores the challenges in data collection and integration within live buildings, pointing out the difficulties of merging various subsystems into a unified digital twin platform. Future efforts involve incorporating variables such as temperature and humidity to provide a more comprehensive understanding of system performance. Additionally, a promising field of study revolves around exploring forecasting ventilation power consumption and investigating optimization strategies.

7. Acknowledgements

This research was carried out under the “DanRETwin: A Digital Twin solution for optimal energy retrofit decision-making and decarbonization of the Danish building stock” project, funded by the Danish Energy Agency (*Energistyrelsen*) under the Energy Technology Development and Demonstration Program (EUDP), ID number: 640222-496754.

References

- Bjørnskov, J. Jradi, M. 2023. “An ontology-based innovative energy modeling framework for scalable and adaptable building digital twins.” *Energy and Buildings*
<https://doi.org/10.1016/j.enbuild.2023.113146>
- Jradi, M., Bjørnskov, J. 2023. “A Digital Twin Platform for Energy Efficient and Smart Buildings Applications.” *2023 Fifth International Conference on Advances in Computational Tools for Engineering Applications (ACTEA)*
DOI: 10.1109/ACTEA58025.2023.10194071
- Lu, Q. Xie, X. Parlikad, A. Schooling, J., M. Konstantinou, E. 2020. “Moving from building information models to digital twins for operation and maintenance.” *Proceedings of the Institution of Civil Engineers: Smart Infrastructure and Construction*
DOI: 10.1680/JSMIC.19.00011
- Cespedes-Cubides, A.S., Jradi, M. A review of building digital twins to improve energy efficiency in the building operational stage. *Energy Inform* 7, 11 (2024).
<https://doi.org/10.1186/s42162-024-00313-7>
- Huang, Gongsheng 2011. “Model predictive control of VAV zone thermal systems concerning bilinearity and gain nonlinearity.” *Control Engineering Practice*
DOI: 10.1016/j.CONENGPRAC.2011.03.005
- Smarra, F. Jain, A. de Rubeis, T. Ambrosini, D. D’Innocenzo, A. Mangharam, R. 2018. “Data-driven model predictive control using random forests for building energy optimization and climate control.” *Applied Energy*
DOI: 10.1016/j.APENERGY.2018.02.126
- Agostinelli, S. Cumo, F. Guidi, G. Tomazzoli, C. 2021. “Cyber-Physical Systems Improving Building Energy Management: Digital Twin and Artificial Intelligence” *Energies* 2021, Vol. 14, Page 2338
DOI: 10.3390/EN14082338
- Zhang, Chenchen Sun, Hejiang 2022. “Analysis and Modeling of Mechanical Ventilation Operation Behaviors of Occupants in Cold Regions of North China” *Applied Sciences* 2022, Vol. 12, Page 5143
DOI: 10.3390/APP12105143
- European commission. 2018. “Energy performance of buildings directive.” European commission. Accessed February 2024.
https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en
- IEA. 2019. “Global Status Report for Buildings and Construction 2019 Analysis” IEA. Accessed February 2024. <https://www.iea.org/reports/global-status-report-for-buildings-and-construction-2019>
- SAREF. 2024. “The Smart Applications REference Ontology (SAREF)” SAREF. Accessed February 2024. <https://saref.etsi.org/>

Title: Using Digital Twins to investigate the need for active cooling systems in Danish public buildings

Muhyiddine JRADI^{1,*}

¹ Center for Energy Informatics, University of Southern Denmark, Odense, Denmark

*Email: mjr@mmmi.sdu.dk

Abstract: When it comes to energy consumption in buildings, Denmark is not an exception, where the building sector contributes to around 35% of the overall energy consumption in the country, with a comparable share in greenhouse gas emissions. Traditionally, most of the energy used in Danish public buildings is devoted to satisfying the demands for heating, hot water, ventilation, lighting, and equipment. On the other hand, such buildings are rarely equipped with cooling systems, even in newly built facilities, and rely mainly on passive techniques and mechanical ventilation systems to provide good indoor thermal comfort in the summer. However, the rise in average ambient air temperatures in recent years has opened the discussion wide on the need to equip public buildings in Denmark with active cooling systems to satisfy the cooling loads of the summer months. In this work, a case study of a public university building is considered a case study to investigate the impact of the rise in ambient air temperatures on the overall indoor thermal comfort levels in addition to assessing the need for active cooling systems in such facilities. The case study considered is relatively modern in terms of envelope design and energy supply systems. It has no active cooling systems and relies mainly on mechanical and natural ventilation to establish indoor comfort. A building energy model is developed and calibrated using actual data collected from the building. The building twin is used to predict the impacts of the increase in the ambient air temperatures and global solar irradiation on the cooling energy demand and indoor thermal comfort. The results highlighted a major increase in the cooling load of the building as well as the energy consumption to satisfy the cooling needs. In addition, major violations to the indoor thermal comfort requirements were reported which demonstrates the need for active cooling solutions in public Danish buildings.

Keywords: Buildings, energy consumption, active cooling systems, energy modelling, digital twins.

1. INTRODUCTION

Buildings within the European Union (EU) consume approximately 40% of the total energy consumed and are responsible for around 36% of CO₂ emissions (Csoknyai et al., 2016). Recognizing the need to achieve its ambitious energy and climate objectives for 2020 and 2050, the EU has placed a strong emphasis on the building sector, acknowledging its significant potential for cost-effective energy savings and the corresponding reduction in greenhouse gas emissions (Jradi et al., 2018). Over the past few decades, substantial efforts have been made in the European Union to enhance the energy efficiency of buildings, reduce overall energy consumption, and decrease associated greenhouse gas emissions. From a regulatory perspective, the Energy Performance of Building Directive (EPBD) (European Commission, 2010) which was established in 2002, laid the groundwork for improving the energy performance of buildings within the EU. As a follow-up, the EPBD underwent a recast and was adopted by the European Parliament and the Council of the European Union. This recast aimed to enhance and strengthen the energy performance requirements for buildings, replacing the original 2002 Directive. In this context, the 2010 EPBD introduced stringent requirements for implementing energy efficiency measures during major energy renovations of buildings, with an estimated potential for 30% energy savings across the EU building stock by 2020.

Throughout Europe, increasing temperatures, coupled with a growing elderly population and the trend towards urban living, are causing the populace to become more susceptible to heat-related issues. Consequently, the need for cooling within buildings is experiencing a swift escalation. Buildings, being enduring constructions, have the potential to shield individuals from heatwaves and elevated temperatures, provided they are adeptly planned, built, refurbished, and upheld. The summer of 2022, characterized by prolonged heatwaves and surging energy costs, might have intensified the sense of immediacy attributed to addressing the challenges posed by heat-induced stress. With people in the EU spending around 80%-90% of their time indoors, enhancements in the constructed buildings and facilities can play a significant role in considerably mitigating the impacts of heat stress (WHO, 2021). Given their durable nature, buildings, when appropriately designed, have the potential to safeguard their occupants from the consequences of heatwaves and increasing temperatures. Conversely, inadequate building design leading to overheating can result in indoor air temperatures reaching levels that detrimentally influence occupants' well-being, causing heatstroke, dehydration, exacerbation of chronic and respiratory ailments, and even fatalities. Research demonstrated that approximately 9%-20% of the population across various EU nations could be vulnerable to building-related overheating (Kolokotsa et al., 2015). Relatively low thermal efficiency and substandard building quality can contribute to the risks of overheating and unmanaged rises in energy consumption for cooling. In this regards, a substantial portion of the EU's building stock was constructed prior to the establishment of thermal standards, with nearly 75% of the stock exhibiting energy inefficiency. Furthermore, in the EU countries, the volume of ultimate energy consumption designated for cooling within residential buildings witnessed a threefold surge from 2010 to 2019. Across the 27 EU Member States, the energy allocation for residential cooling constituted a mere 0.4%, on average, of the complete final energy consumption within the residential sector during 2019 (Quefelec, 2022). Nonetheless, this proportion exhibited notable elevation in certain nations such as Malta, Cyprus, and Greece, reaching 11%, 10%, and 5%, respectively. Throughout the summer of 2022, Greece, Italy, Spain, and other countries encountered a pressing cooling demand due to prolonged and recurrent heatwaves (Copernicus, 2022), coupled with elevated energy costs and the Ukraine conflict. Projections concerning forthcoming cooling demand within EU buildings yield contrasting findings and encompass diverse parameters. Nevertheless, these studies unanimously foresee an escalation in energy consumption for building cooling, consequently leading to an increased proportion of energy dedicated to this purpose. Kranzl et al. (2021) highlighted that, among the energy used within buildings, the part allocated to cooling, encompassing both residential and non-residential structures, might rise to around 8% to 9% by 2050, a substantial increase compared to the 2% recorded in 2012. The surge in future cooling energy demand will be most pronounced in the southern EU countries. Jakubcionis et al. (2017) estimated that, in the impending years, Greece, Italy, Portugal, and Spain might collectively account for 71% of the average annual energy consumption designated for residential building cooling across the EU.

Regarding energy usage in buildings, Denmark is not an exception and follows the trend, where the building sector accounts for approximately 35% of the nation's total energy consumption. This is in line with the corresponding contribution of greenhouse gas emissions. Conventionally, Danish public buildings harness most of the consumed energy for heating, hot water, ventilation, lighting, and equipment. However, cooling systems are scarcely found in these buildings, even in newly constructed ones. Instead, they heavily rely on passive methods and mechanical ventilation systems to ensure comfortable indoor temperatures during summer. Nonetheless, the increasing average outdoor temperatures in recent years have triggered discussions about the necessity of incorporating active cooling systems in Danish public buildings to effectively manage cooling needs during the warmer months. To explore this matter, a public university building is taken as a case study. The aim is to evaluate how the rising outdoor temperatures influence indoor comfort levels and to assess whether there is a requirement for active cooling systems in such buildings. The chosen case study involves a relatively modern building with up-to-date envelope design and energy supply systems. Notably, it currently lacks active cooling systems and primarily relies on a combination of mechanical and natural ventilation to establish indoor comfort. To analyze this scenario, a digital replica of the building is constructed and fine-tuned using real-world data collected from the physical structure. This digital twin of the building is then employed to forecast the potential consequences of the variation in the climatic conditions in the recent years on the overall building performance, particularly in the cooling season. This encompasses analyzing the corresponding impacts on energy usage and maintaining optimal indoor thermal comfort. Based on these findings, significant recommendations are formulated, offering insights into potential future initiatives for introducing active cooling systems in Danish public buildings.

2. THE DANISH BUILDING SECTOR

Denmark has served as a remarkable example over the past thirty years, achieving economic growth and prosperity while maintaining relatively stable overall energy consumption and a significant reduction in CO₂ emissions. This accomplishment can

be attributed to the implementation of feasible energy-saving policies and regulations, substantial investments in renewable energy sources, particularly wind energy, and the promotion of social awareness regarding energy efficiency and conservation measures. To meet the EU's objective of reducing greenhouse gas emissions by 80-90% by 2050, the Danish government has taken significant measures to enhance the country's energy production, supply, and consumption patterns, while also considering the economic and social aspects and the citizens' quality of life. Ensuring a secure and prosperous energy sector is a key priority in response to potential energy shortages or fluctuations in energy prices, and to address this, the government has set an ambitious goal of achieving a fossil-free energy supply by 2050, relying solely on renewable energy sources to meet Denmark's energy demands (Lund et al., 2009). This necessitates the implementation of comprehensive, energy-efficient, and cost-effective measures, accompanied by a substantial expansion of the renewable energy sector.

In Figure 1, the evolution in terms of the maximum allowable primary energy consumption for a typical 150 m² newly constructed residential building in Denmark is depicted. This highlights the progression and evolution of different building regulations, referred to as 'BR' (Jradi et al., 2017). It is noteworthy that the permitted primary energy usage for newly constructed buildings has decreased by a remarkable 95%. For instance, buildings constructed in 1961 had an approximate energy use of 350 kWh/m² per year, whereas in 2020, the limit has been significantly reduced to just 20 kWh/m². This clearly emphasizes the substantial potential not only in constructing energy-efficient new buildings but also in improving the overall performance of existing buildings, particularly those built prior to the 1980s, through systematic retrofit processes and activities. As part of the Danish government's energy agreement established in 2012, a comprehensive energy renovation strategy, known as the "Strategy for the energy renovation of the existing building stock," was introduced, marking a significant milestone in the country's transition towards a highly efficient energy sector. The strategy encompasses 21 initiatives aimed at enhancing the energy renovation process in Denmark and improving the energy performance of existing buildings. In Denmark, energy renovation is not merely a voluntary action to reduce energy consumption and enhance the energy performance of buildings; it is an essential process due to the fact that nearly 75% of the existing building stock was constructed before 1979, using materials with a lifespan of 40-50 years. Therefore, a comprehensive and widespread energy renovation and retrofit process is set to be implemented over the coming years until 2030, and it is crucial to incorporate energy-efficient and cost-effective measures into these upcoming energy renovation projects.

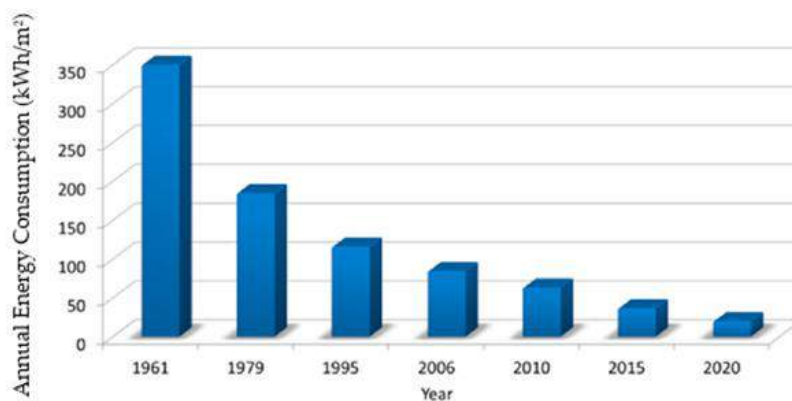


Fig. 1. The maximum allowed annual primary energy usage in connection to the progressive building regulations for a 150 m² Danish residential building.

In Denmark, the building sector, in terms of building area, has experienced an annual growth rate of 1% over the past 40 years, indicating that by 2050, new buildings constructed from now on will only constitute 20-30% of the building stock. This highlights the potential significance of focusing on existing buildings, which make up the majority share of the building stock, in order to achieve the stringent Danish energy goals by 2050. Building energy requirements have been progressively tightened since the 1970s to reduce overall energy consumption through the implementation of energy-efficient measures and standards. The latest 2020 building class in Denmark complies with the requirements for nearly-zero energy buildings, which will become mandatory by 2020. The Danish Building and Urban Research Institute has assessed the potential heat savings in existing buildings and reported that the primary sources of heat loss in buildings constructed between 1960 and 2004 are exterior walls, ventilation systems, and windows, in that order. Therefore, measures prioritizing the upgrading and insulation of exterior walls and windows, as well as improving ventilation units, should be given priority in the renovation of such existing buildings. Additionally, the Danish Building Research Institute has estimated that improving envelope insulation could result in annual energy savings of 31 PJ and 6 PJ in residential and commercial buildings, respectively, while renovation of installations could provide respective savings of 21 PJ and 3 PJ in residential and commercial buildings.

3. CASE STUDY DESCRIPTION

In this investigation, a case study of a Danish university office building is considered to evaluate how the rising outdoor temperatures influence indoor comfort levels and to assess whether there is a requirement for active cooling systems in such buildings. The building, which spans an area of 2600 m² and is depicted in Figure 2, was constructed in 1995. It encompasses two floors and a small basement that houses HVAC systems. The building primarily functions as an office space for daily use by staff and researchers. The structure consists of 110 zones, predominantly comprising personnel offices, along with meeting rooms,

research group areas, laboratories, and seminar rooms. Key details about the building's physical envelope, construction, and energy supply systems can be found in Table 1.



Fig. 2. University public building case study.

Table 1. Case study building characteristics.

Location	SDU Odense Campus
Indoor Heated Area	2600 m ²
Rooms	110 rooms
Window to wall ratio	0.26
Roof	300 mm concrete with insulation and bitumen
Exterior Walls	300 mm mainly concrete and insulation
Windows	Double-glazed windows
Heating System	In-direct district heating system with a heat exchanger on the main supply
Hot Water	Central production in an insulated storage tank
Space Heating	Radiators in every room with thermostatic valves
Ventilation System	Mainly natural ventilation with individual mechanical ventilation units in laboratories and specific rooms
Lighting	A mix of 100W halogen lights and fluorescent lights in the offices
Equipment	Mainly laptops, screens, printers, kitchen equipment.

To fulfill the requirements for space heating and domestic hot water, an indirect district heating system is employed. This system includes two heating sub-stations situated in the building's basement, as well as a heat exchanger on the main supply line. Hot water is generated within an insulated storage tank located in the basement. Radiators with thermostatic valves are utilized in each room for space heating control. The building relies largely on natural ventilation through controlled operable windows for its ventilation needs. However, a primary mechanical ventilation system is also integrated to ensure indoor air quality and thermal conditions, particularly in spaces like laboratories and kitchens. It is worth noting here that the building does not have a cooling system in place. Lighting is provided by a combination of 100 W halogen lights in corridors and LED tubes in offices. Given that the majority of the building space is designated for office purposes, the primary equipment consists of laptops, computers, printers, screens, as well as laboratory devices and kitchen equipment. The building's thermal envelope adheres to the Danish BR95 building regulation, featuring external walls constructed from dense layers of concrete with added insulation. According to the Danish energy rating classification, the building currently holds a 'D' rating. Regarding the building's measurement and sensing system, the structure is outfitted with electricity and heating meters that cover the entirety of the building, as well as specific meters for each energy-supply component, encompassing domestic hot water, space heating, ventilation, and lighting. Furthermore, individual rooms are equipped with Passive Infrared (PIR) and illuminance sensors that play a role in regulating the operation of the lighting system. For the purpose of control and oversight, the building incorporates a Schneider Electric structureware building automation system, which oversees diverse systems and services. Additionally, two occupancy counters are placed at each entrance of the building to monitor the count of occupants. Moreover, an on-site weather station is established on the roof of the building. This station enables immediate recording of climatic conditions such as ambient temperature, wind speed, and solar irradiation. The Kafka platform is employed as an interface to store data amassed from the building's measurement and sensing infrastructure at various levels. This data is accessible through a centralized platform. The utilization of the Kafka platform serves to facilitate the process of data acquisition, gathering, categorization, and preliminary processing. It also provides functionalities for further processing of data collected from different meters and sensors. Consequently, this setup eases the utilization of such data in diverse applications, including monitoring performance, controlling operations, managing systems, and making comparisons.

4. BUILDING ENERGY MODELLING

4.1. Development of Holistic Energy Model

A comprehensive and dynamic model was created to assess the energy performance of the specified case study building. This model incorporated diverse details about the building's design specifications, structural features, energy systems, utility systems, loads, and operational variables. The construction of the comprehensive energy model followed the systematic approach outlined by Jradi et al. (2017), which emphasizes a holistic perspective on whole-building energy modeling and performance simulation. This approach employs several tools, including Sketchup Pro for creating a detailed 3D architectural model, OpenStudio for defining building specifications and energy model parameters through an intuitive interface, and EnergyPlus for executing accurate energy simulations. To begin, a precise 3D architectural model of the building was crafted using Sketchup, as depicted in Figure 3. This process meticulously accounted for information such as building dimensions, orientation, floor layouts, and space allocations. This effort resulted in a comprehensive building model featuring 110 distinct thermal zones. Subsequently, the 3D model was imported into OpenStudio, where all necessary inputs for the energy model were incorporated. These inputs encompassed data related to the building envelope, energy system arrangement, utility services, loads, and activity schedules. Finally, the EnergyPlus simulation engine was employed to perform an exhaustive dynamic energy simulation, enabling the prediction of the building's overall performance under various conditions.



Fig. 3. 3D model of the case study building.

4.2. Calibration of Building Energy Model

Once the comprehensive energy model of the entire building has been constructed, it becomes necessary to fine-tune the model to ensure accurate representation of the building's performance. This involves achieving reliable outcomes and forecasts with an acceptable level of precision. This finely calibrated dynamic energy model is essential not only to form the foundation of a consistent, ongoing commissioning system but also to offer effective services for monitoring and evaluating performance. To calibrate the energy model for the specific case study building, actual on-site data collected from various meters and submeters over a 5-month operational period from January to May is employed. This time frame corresponds to a standard usage pattern for the building post-COVID pandemic, when staff and students have returned to their usual work and study routines within the building. Alongside the meter data, real weather information sourced from a weather station, encompassing factors like ambient air temperature, wind speed, and solar irradiation, is utilized as input for the energy model simulation. This assists in the calibration process and facilitates reasonable comparisons between the model's simulations and the actual energy consumption figures. Moreover, the energy model incorporates detailed schedules for occupancy and load allocations throughout the building, which are established based on a comprehensive survey of the various rooms and spaces. These specific schedules and loads supersede the default and generic ones in the model, enhancing its accuracy. Additionally, operational parameters for the systems, including setpoints, thresholds for automated control, and management patterns, are extensively integrated into the developed model.

During the process of model calibration, the utilized approach for dynamic energy modeling calibration in OpenStudio, as introduced by Hale et al. (2014), was employed. Nevertheless, this framework underwent enhancements. Instead of solely focusing on overall heating and electricity usage figures, supplementary submeters were integrated into the calibration procedure. These additional submeters encompassed energy consumption for domestic hot water and the preheating energy consumption for the ventilation system. Regarding the parameters chosen for calibration and considering the aforementioned framework established by Hale et al. (2014), the parameters that were taken into consideration within this investigation comprised space infiltration rates, pump head, lighting and equipment schedules, efficiency of devices, system setpoints, pressure rise of ventilation fans, and occupancy patterns. The Parametric Analysis Tool (PAT) within OpenStudio is employed to assist in the process of model calibration. It accomplishes this by assessing various combinations of parameters and enabling the simulation of numerous scenarios concurrently. Given that the building's construction dates back to the 1990s, the Danish building regulation BR95 was utilized to establish limits for specific parameters employed in the calibration process. These parameters encompass infiltration rates, loads, and the efficiency of systems and devices. The objective is to ensure that the calibrated parameter value remains consistent with the anticipated real-world parameters for such a building.

From the extensive range of simulated scenarios, the scenario exhibiting the smallest deviation concerning individual daily energy consumption performance indicators was selected as the reference point for the model. This chosen scenario will serve as the foundational core for continuous building commissioning and performance evaluation.

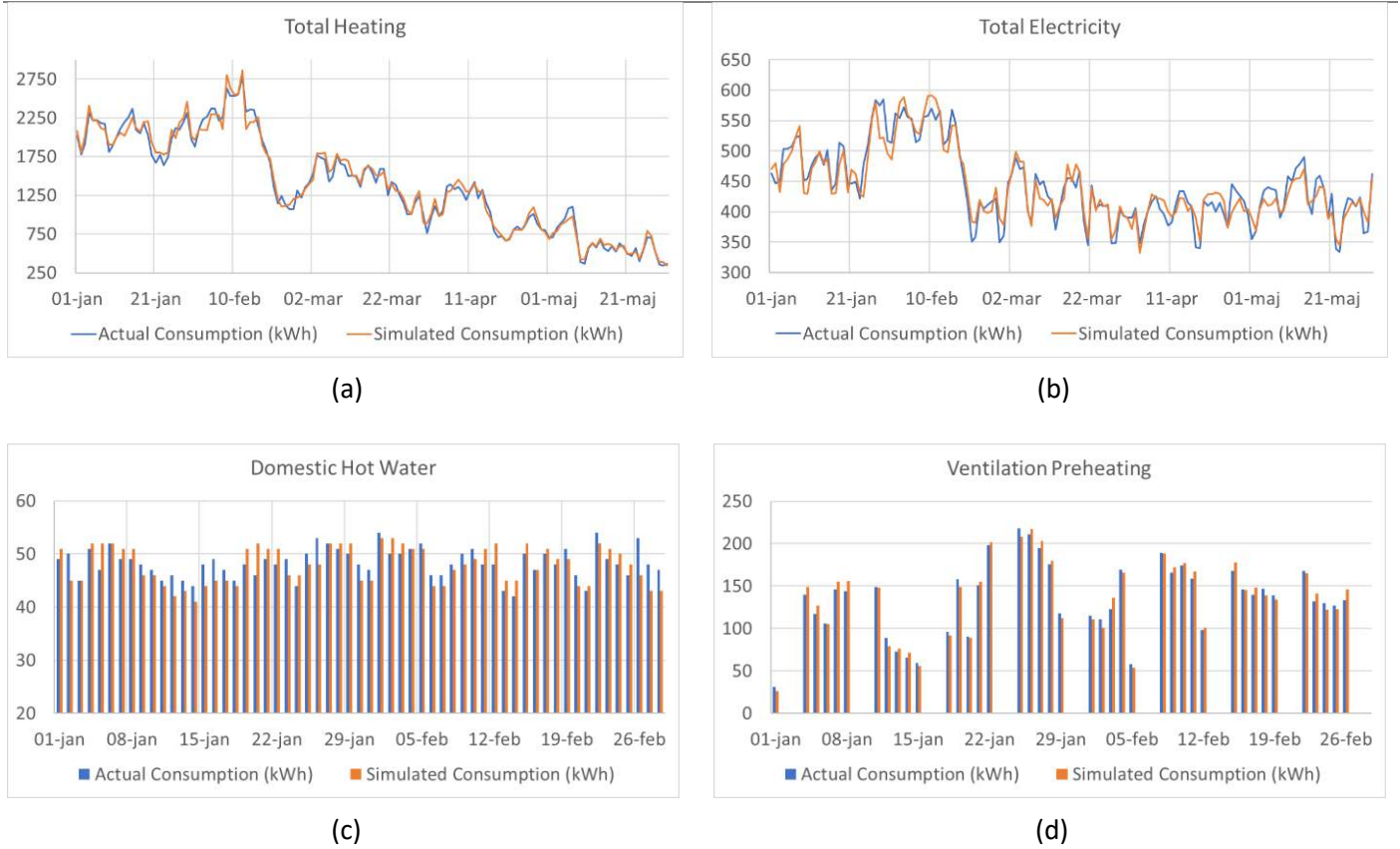


Fig. 4. Calibration of energy model results considering the (a) total heating consumption, (b) total electricity consumption, (c) domestic hot water energy, and (d) ventilation preheating energy.

Figure 4 (a to d) depicts the outcomes of the calibration process by contrasting the daily anticipated energy usage with the actual on-site energy consumption for various components: (a) total heating, (b) total electricity, (c) domestic hot water, and (d) ventilation preheating. As illustrated in the illustration, it's evident that the refined building model reasonably predicts energy consumption across different areas of the building. To assess the accuracy of the predictions, the Root Mean Square Error (RMSE) is used. Table 2 summarizes the average monthly RMSE values for the four energy performance metrics in consideration. For total heating consumption, the highest average RMSE is approximately 107 kWh in February, corresponding to a maximum daily uncertainty of 17.4% between the projected and actual values. In the case of total electricity, the most significant RMSE is observed in January, amounting to 24 kWh, with a maximum uncertainty of 14.7%. Conversely, RMSE values for domestic hot water range from 1.4 kWh in May to 3.1 kWh in March, with the highest daily uncertainty of around 13.8%. Lastly, for the ventilation preheating metric, the maximum RMSE reported is about 7 kWh in March, accompanied by a maximum daily uncertainty of 16.3% in the match between simulated and real consumption.

Table 2. Root mean square error for the model calibration.

	RMSE (kWh)				
	January	February	March	April	May
Total Heating Consumption	95.71	107.56	73.09	59.91	50.26
Total Electricity Consumption	23.97	20.03	13.88	18.91	18.41
Domestic Hot Water Energy	2.85	2.59	3.09	1.75	1.48
Ventilation Preheating Energy	5.55	6.04	7.13	5.75	4.63

5. BUILDING COOLING DEMAND

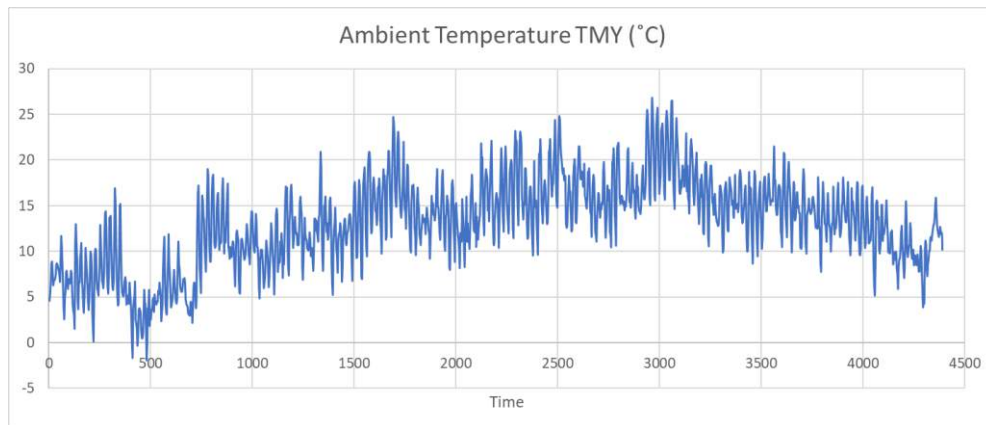
5.1. Climatic Conditions

Before investigating the case study building cooling demand and indoor thermal comfort conditions, the considered area climatic conditions are presented and analyzed as they play a major role in driving the building's need for cooling and heating. In this regard, Table 3 shows a comparison between the climatic conditions of the Typical Meteorological Year (TMY) in Odense and the average between the last three years, 2020 to 2022, denoted by (AVG). In the table, the average ambient air temperature, maximum ambient air temperature, average global solar irradiation, and maximum global solar radiation are reported for each of the cooling season

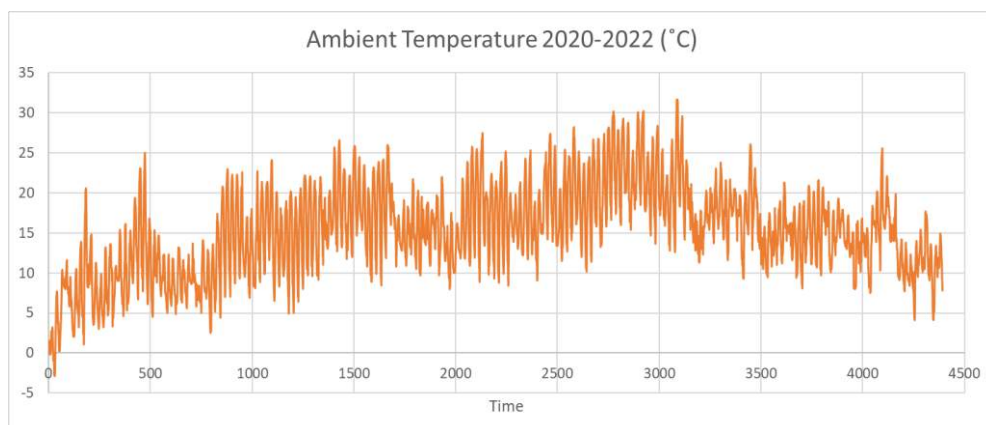
months, considered from April to September. The difference between the climatic conditions of the TMY and AVG profiles is very obvious, with notable higher average temperatures and global solar radiation in the AVG compared to those in the TMY. For example, the average ambient air temperatures in June and July for the TMY are 14.6°C and 16.6°C, respectively, compared to 18.7°C and 21.7°C, respectively, for the AVG year. In addition, the maximum recorded ambient air temperature in August in the TMY is 26.8°C compared to 31.7°C in the average year of 2020-2022. Furthermore, Figures 5 (a-d) show the timeseries data for the ambient air temperature for TMY and the average of 2020-2022 and the global solar radiation for TMY and the average of 2020-2022. Overall, it is obvious that there has been a notable increase in the ambient air temperature and the global solar radiation in the area in the last three years in comparison to the TMY case.

Table 3. Climatic conditions comparison between TMY and 2020-2022 average (AVG).

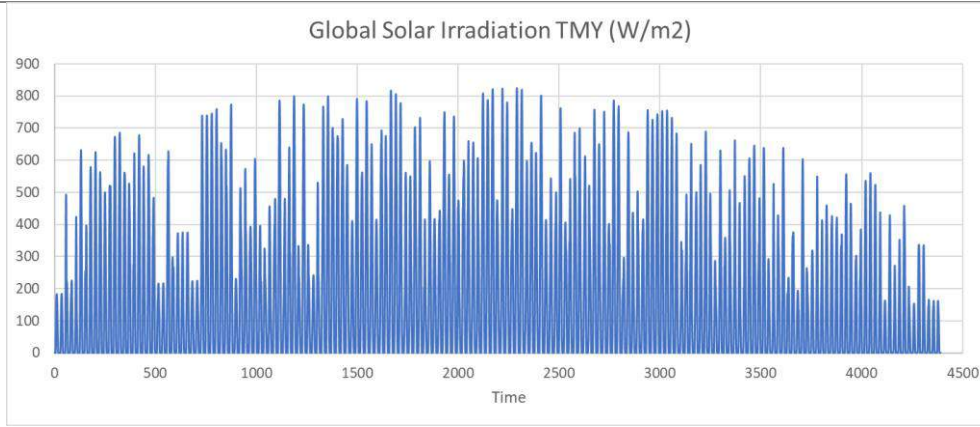
	April		May		June		July		August		September	
	TMY	AVG	TMY	AVG	TMY	AVG	TMY	AVG	TMY	AVG	TMY	AVG
Average Ambient Air Temperature (°C)	6.2	11.1	11.4	16.7	14.6	18.7	16.6	21.7	16.8	19.6	12.3	15.8
Maximum Ambient Air Temperature (°C)	16.9	24.3	20.9	26.6	24.7	27.5	24.8	30.2	26.8	31.7	19.1	25.6
Average Global Solar Irradiation (W/m ²)	144.9	180.8	205.22	243.5	231.4	267.7	217.6	247.1	174.7	212.6	105.4	144.1
Maximum Global Solar Irradiation (W/m ²)	686.2	751.6	799.66	768.3	821.6	846.3	824.76	803.6	755.28	761.3	603.8	606.3



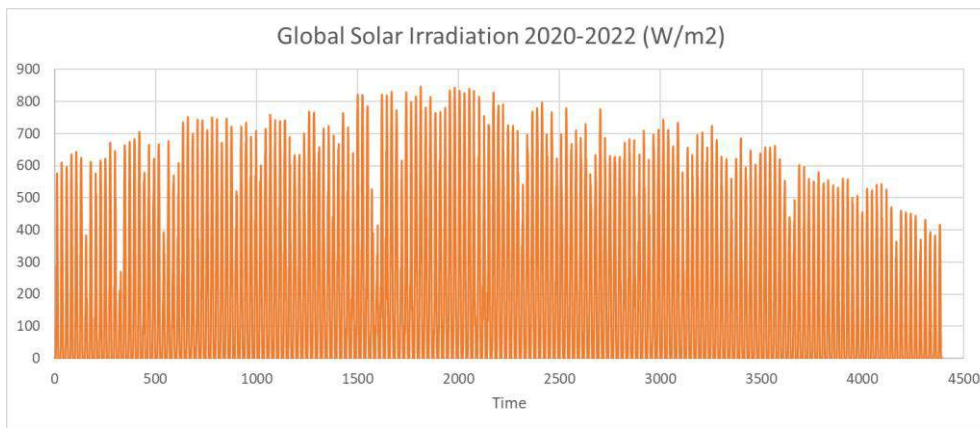
(a)



(b)



(c)



(d)

Fig. 5. Ambient air temperature for (a) TMY and (b) average of 2020-2022 and global solar radiation for (c) TMY and (d) average of 2020-2022.

5.2. Building Cooling Load

After presenting the climatic conditions in the considered area as the average of the last three years in comparison to the TMY year, the developed and calibrated building energy model in OpenStudio was employed to simulate and predict the case study building cooling load when using the TMY or the AVG climatic conditions profiles as the weather file input to the tool. A comparison of the attained maximum building cooling load in each of the six months considered in both cases is reported in Table 4. As shown in the table, using the AVG climatic conditions as weather profile input for the simulation leads to an increase in the needed cooling load in all six months considered, with the largest increase attained in July and August. In July, the maximum cooling load was increased from 55 kW to 144 kW, whereas in August, the cooling load increased from 68 kW to 131 kW.

In addition, Figure 6 presents the monthly building energy consumption to satisfy the cooling demand in each month from April to September. It is also obvious here from the figure that the cooling energy consumption increased drastically when employing the AVG year of 2020-2022 in comparison to the results obtained with the TMY. For instance, the cooling energy consumption in August increased from around 5600 kWh to around 15650 kWh, an increase of 180%.

Table 4. Building cooling load comparison between TMY and 2020-2022 average (AVG).

	April		May		June		July		August		September	
	TMY	AVG	TMY	AVG	TMY	AVG	TMY	AVG	TMY	AVG	TMY	AVG
Maximum Cooling Load (kW)	14.5	33.1	17.8	68.2	37.9	71.1	55.2	144.3	68.1	131.1	13.1	36.9

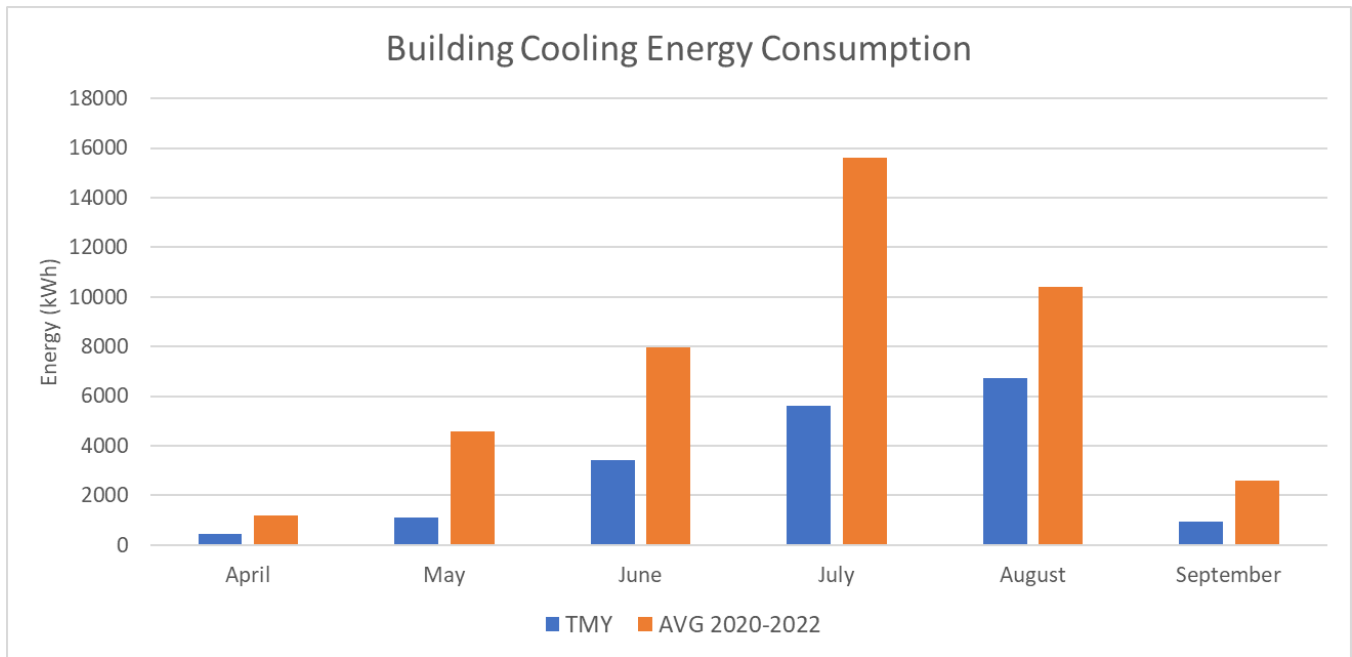


Fig. 6. Monthly building energy consumption to satisfy cooling demand.

5.3. Indoor thermal comfort

In addition to investigating the variations in the required cooling load and energy consumption in recent years, in this section, the impact on indoor thermal comfort is also presented and analysed. In this regard, the Danish building regulations BR18 highlight that for non-residential buildings, the indoor air temperature should not exceed 26 °C for more than 100 hours and not exceed 27 °C for more than 25 hours over a year. To assess this, the AVG weather profile with the average of the last three years was used to predict and simulate the indoor thermal comfort conditions in the considered case study building. Figure 7 shows the indoor air temperature profile in two office rooms in the considered building over the period from April to September. It was found that both zones violated the two constraints set above. In this regard, THERMAL ZONE 15 and THERMAL ZONE 3 violate the time constraint of 26°C for 516 and 2102 hours, respectively, as compared to the allowed 100 hours. Nevertheless, the two respective zones violate the 27°C constraint of 25 hours, with 291 and 1431 hours with an indoor air temperature above 27°C, respectively. The findings above demonstrate that the employed hybrid mechanical-natural ventilation system is not capable of establishing proper indoor thermal comfort, and an active cooling system is a must to lower the high temperatures reported.

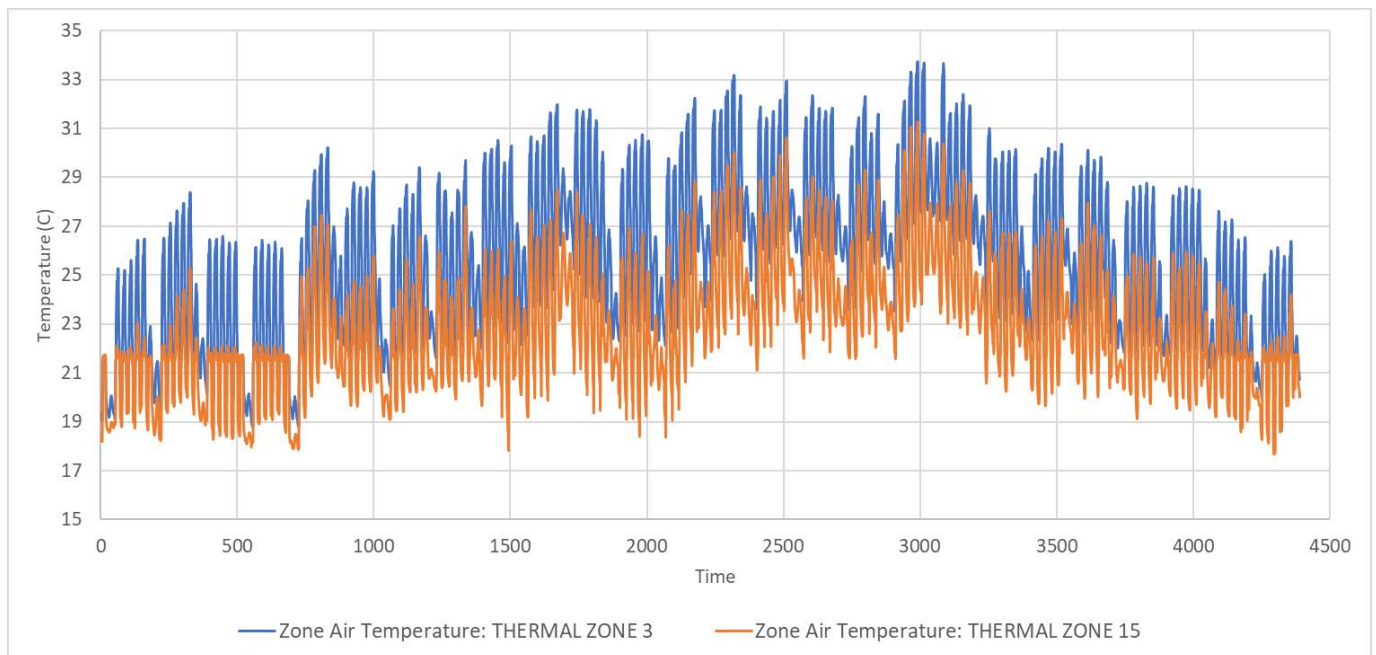


Fig. 7. Indoor air temperature profile in two zones in the considered case study building.

6. CONCLUSION

Traditionally, heating, hot water, ventilation, lighting, and equipment account for most energy consumption in Danish public buildings. Interestingly, cooling systems are seldom found in these buildings, even in newly built ones. Instead, they primarily rely on passive methods and mechanical ventilation to maintain comfortable indoor temperatures during the summer. However, the escalating average ambient temperatures in recent times, along with the heat waves experienced in Europe in recent years, have sparked discussions about incorporating active cooling systems in Danish public buildings to address summer cooling needs. This study centers on a public university building as a case study to explore how rising outdoor temperatures affect indoor thermal comfort and to evaluate the necessity of introducing active cooling systems in such facilities. The chosen case study is relatively modern in terms of its architectural design and energy supply systems. It currently lacks active cooling systems and leans on mechanical and natural ventilation to ensure indoor comfort. To investigate the situation, a model of the building's energy usage is developed and fine-tuned using real data collected from the structure. This model is employed to anticipate the consequences of elevated outdoor temperatures and increased solar radiation on the demand for cooling energy and indoor thermal comfort. The results showed a major increase in the cooling load of the building as well as the energy consumption to satisfy the cooling needs. In addition, major violations of the indoor thermal comfort requirements were highlighted, demonstrating the need for active cooling solutions in public Danish buildings. The results of this study are based on an investigation carried out considering one case study. To generalize the results and findings, additional cases are to be considered in future investigations, including modern and newly built buildings and old existing buildings. In addition, future studies could investigate the impact of different parameters, including the building construction envelope specifications, orientation, and use.

7. ACKNOWLEDGMENT

This work is carried out under the 'Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle' project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64021-1009.PV modeling part.

8. REFERENCES

- Copernicus, 2022, Europe continued to swelter in July, Copernicus Programme, European Commission. [Online]. Available at: <https://climate.copernicus.eu/europe-continued-swelter-july>
- Csoknyai, T., Hrabovszky-Horváth, S., Georgiev, Z., Jovanovic-Popovic, M., Stankovic, B., Villatoro, O., Szendro, G., Building stock characteristics and energy performance of residential buildings in Eastern-European countries, *Energy Build.*, 132 (2016) 39-52.
- Dittmann, F., et al., 2016, Space cooling technology in Europe: technology data and demand modelling, Heat Roadmap Europe. [Online], Available at: https://heatroadmap.eu/wp-content/uploads/2018/11/HRE4_D3.2.pdf
- European Commission, 2010, Directive 2010/31/EU of the European Parliament on the energy performance of buildings, 2010.
- Hale, E., Lisell, L., Goldwasser, D., Macumber, D., Dean, J., Metzger, I., Parker, A., Long, N., Ball, B., Schott, M., Weaver, E., Brackney, L., CloudBased Model Calibration Using OpenStudio, eSim 2014, Ottawa, Canada, May 7-10, 2014.
- Jakubcionis, M., Carlsson, J., Estimation of European Union residential sector space cooling potential, *Energy Policy* 101 (2017) 225-235.
- Jradi, M., Veje, C., Jørgensen, B.N., Deep energy renovation of the Mærsk office building in Denmark using a holistic design approach. *Energy Build.*, 151 (2017) 306–319.
- Jradi, M., Veje, C.T., Jørgensen, B.N., A dynamic energy performance-driven approach for assessment of buildings energy Renovation—Danish case studies, *Energy Build.*, 158 (2018) 62-76.
- Kolokotsa, D., Santamouris, M., Review of the indoor environmental quality and energy consumption studies for low income households in Europe, *Science of the Total Environment* 536 (2015) 316-330.
- Lund, H., Mathiesen, B.V., Energy system analysis of 100% renewable energy systems - The case of Denmark in years 2030 and 2050, *Energy* 34 (2009) 524-531.
- Quefelec, S., 2022, Cooling buildings sustainably in Europe: exploring the links between climate change mitigation and adaptation, and their social impacts. [Online], Available at: <https://www.eea.europa.eu/publications/cooling-buildings-sustainably-in-europe>
- WHO, 2021, Heat and health in the WHO European region: updated evidence for effective prevention, World Health Organization. [Online], Available at: <https://apps.who.int/iris/handle/10665/339462>



An ontology-based innovative energy modeling framework for scalable and adaptable building digital twins

Jakob Bjørnskov*, Muhyiddine Jradi

SDU Center for Energy Informatics, Mærsk Mc-Kinney Møller Institut, University of Southern Denmark, Campusvej 55, Odense M, 5230, Denmark

ARTICLE INFO

Keywords:

Digital twin
Data-driven
Building energy model
Building simulation
Ontology
SAREF

ABSTRACT

Digitalization of buildings and the use of IoT sensing and metering devices are steadily increasing, offering new opportunities for more autonomous, efficient, and flexible buildings. As part of this transformation and inspired by the added value demonstrated in other domains, the concept of a building digital twin that can monitor, simulate, manage, and optimize building operation has received increased interest. To aid such digital twin implementations, accurate and adaptable simulation models are required, which can effectively integrate and utilize the available data. However, traditional building modeling and digital practices, such as Building Information Modeling and white-box modeling tools, are not easily compatible with these requirements. This work presents an innovative and flexible energy modeling framework based on the SAREF ontology. With a basis in the SAREF4BLDG extension for buildings and the defined classes, different models are presented for a selection of typical systems and devices such as spaces, space heaters, dampers, coils, etc. Using the generic semantics and relations of the SAREF4SYST extension, a method for linking and simulating component models is then presented. A proof-of-concept of the modeling framework is provided, showing its application and feasibility to provide a dynamic simulation of the different systems and devices included in a demonstration case. Finally, a future line of work is identified considering the implementation of the modeling framework in an actual building case study, including integration with actual sensing equipment to demonstrate different digital twin services such as performance monitoring, strategy planning, and operational optimization.

1. Introduction

Driven by ambitious environmental goals and the promising potential for improved efficiency and automation by employing the Internet of Things (IoT) and Artificial Intelligence (AI) technologies, the building sector is currently undergoing fast-paced digitalization with increasing adoption of large sensing and metering networks. While the utilization of such data offers many opportunities, it also presents great challenges to status quo data management, building operation, and building energy modeling. As a consequence, the demand is steadily growing for more flexible and scalable data and energy models that integrate seamlessly.

As part of the digitalization of buildings, Building Information Models (BIM) have for the past two decades played a major role, unlocking significant cost savings and improving efficiency by easing communication and information exchange between contractors in the design and construction phases of buildings [1]. However, despite its demonstrated success, the BIM technology is now facing significant challenges as at-

tempts are now made to extend it beyond its intended scope to fulfill the demands of IoT-integrated smart buildings. Boje et al. [2], argued that BIM is not suitable for IoT integration due to the static legacy formats and standards used to represent data. Instead, they highlighted the emerging concept of Digital Twins (DT) in buildings as a promising solution for solving these interoperability issues.

The DT concept was first used by the National Aeronautics and Space Administration (NASA), which defined it as a “comprehensive multi-physical, multi-scale, probabilistic simulation system for vehicles or systems” [3]. Following, the DT concept has been adapted in various other scientific and engineering fields such as product manufacturing, medical sciences, and smart cities [4]. Although DTs as a concept is starting to emerge in the building sector, the concept is still in its infancy with no clear consensus on a common definition and which services it should provide.

In this work, we adopt the definition provided by Boje et al. [2] and Grieves [5]. That is, a DT refers to a concept with three main constituents; a physical system, a virtual system, and a flow of data linking

* Corresponding author.

E-mail address: jabj@mmmi.sdu.dk (J. Bjørnskov).

<https://doi.org/10.1016/j.enbuild.2023.113146>

Received 22 January 2023; Received in revised form 1 May 2023; Accepted 3 May 2023

Available online 9 May 2023

0378-7788/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

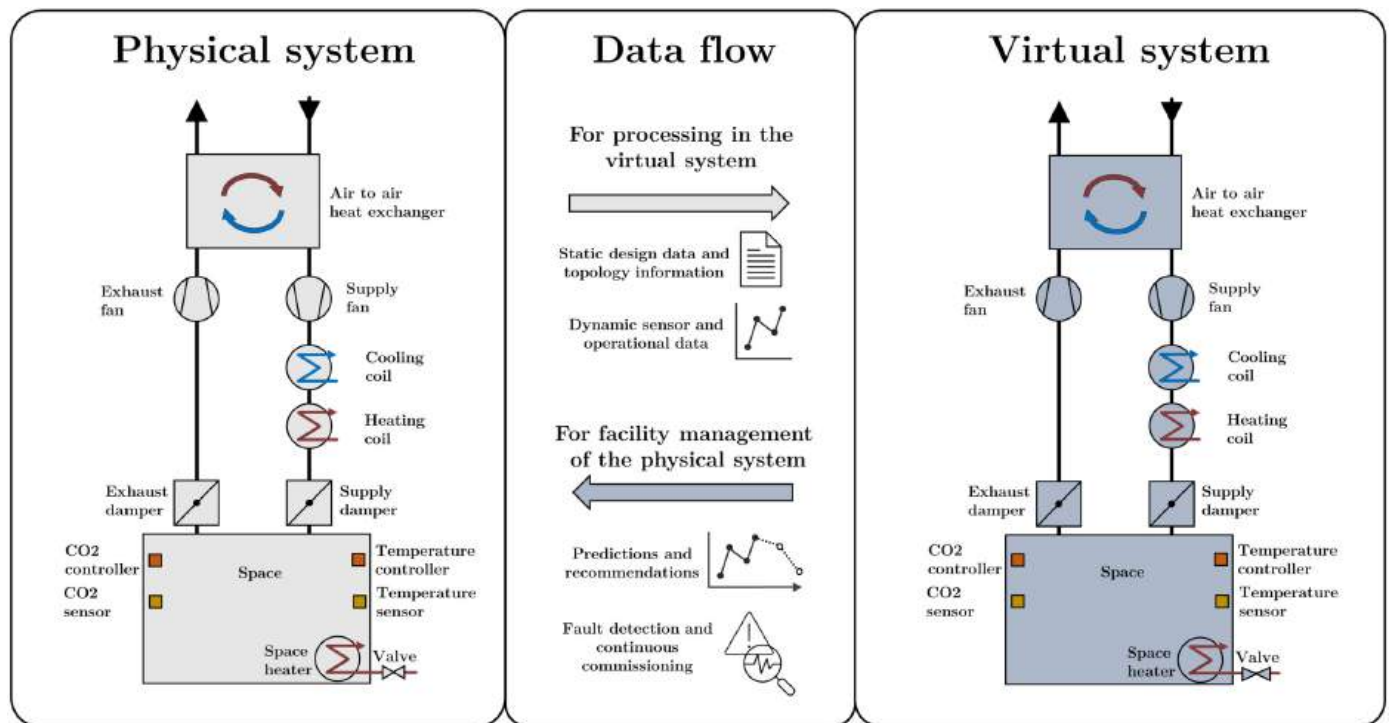


Fig. 1. The DT concept illustrated with typical systems and components present in buildings with their virtual counterpart and the bidirectional flow of data.

these two systems. This is conceptualized in the context of a generic building energy system as shown in Fig. 1. Here, the physical system is the actual asset to be managed, i.e. a building including all systems, devices, spaces, occupants, etc. The virtual system employs relevant simulation models to emulate the behavior of the physical system as closely as possible. The physical and virtual systems are linked through a bidirectional data flow, where the virtual system receives raw data in the form of static design and topology information as well as dynamic sensor and meter data. The virtual system then processes this information to adapt, monitor, simulate, and optimize based on the employed simulation models. This processed information is then communicated back to the physical system for facility management. As a basis for the work presented in this paper and the envisioned DT, three fundamental services, targeting the operational phase of the building life cycle, have been defined as follows:

- Service 1** Efficient collection and management of data generated by sensors, meters, and IoT devices through a user-friendly open-standard context information model
- Service 2** Automated continuous commissioning and performance monitoring in real-time to detect faults and malfunctions, ensuring a smarter and more cost-effective facility management
- Service 3** Operational strategy planning support to enable more informed decision-making by running different control and management scenarios in a zero-risk virtual environment

In recent years, successful implementations of fault detection and performance monitoring, as listed in **Service 2**, have been reported numerous times with different approaches. Typically, fault detection methods are divided into two categories. The first category is model-based methods, which rely on comparisons between the simulated behavior of a system with actual measurements collected from the physical unit. The second category is model-free methods, which typically utilize only operational data collected from a system to find patterns that distinguish faulty from normal operation. Although model-free methods require no models to implement, their effectiveness is also often very limited in comparison to model-based methods when ap-

plied to large-scale distributed systems such as Heating Ventilation Air-Conditioning (HVAC) systems [6]. Therefore, most efforts in fault detection and continuous commissioning of buildings have considered model-based methods, often with the use of white-box simulation tools [7–10]. In addition, an accurate dynamic energy model enables simulation and optimization of future operational scenarios, as required by **Service 3**, aiding decision-makers with operating the building on a day-to-day basis, taking into account weather forecasts, expected occupancy, and price signals. This potentially increases indoor comfort, reduces operational costs, and improves the overall energy flexibility of the building.

The added value of the outlined services is well-established and has been demonstrated numerous times in case-studies [11–13]. However, using traditional energy modeling tools and being challenged by various interoperability issues, such implementations usually result in highly specialized solutions tailored to each specific building. This building-by-building approach, relies on manual workflows, is costly, hard to scale, and acts as a major barrier to broadly implementing such systems.

Therefore, inspired by the added value provided in different domains, there is a need to promote and develop DT technology for the building sector to harness the benefits provided by advanced sensing and metering devices and provide the outlined services at scale. As a key milestone in such promotion, an automated and adaptable framework for energy modeling of buildings is needed to effectively integrate such models into a building DT.

In this work, an innovative framework is proposed and presented for automated and adaptable energy model development to provide the simulation models required by building DTs. The framework builds upon the Smart Applications REFERENCE (SAREF) ontology [14] to ensure interoperability by using existing classes, concepts, and relations. A selection of SAREF classes is extended with adaptable data-driven gray-box and black-box models, to describe their dynamic behavior. Furthermore, a generic framework based on the SAREF4SYST extension is presented to describe how these component models interact based on simple topology information. To provide a proof-of-concept, a demonstration case is considered to illustrate the framework implementation considering the simple single-zone system shown in Fig. 1.

The presented framework in this work is capable to serve as a basis for the development and implementation of scalable DTs for building applications, allowing the delivery of major services to enhance the energy efficiency and intelligence quotient of buildings. This includes performance monitoring, data management, performance optimization, and strategy planning.

As established, one of the core elements of DTs is accurate simulation models. Therefore, in the following section, an overview and discussion of different approaches for building energy modeling are provided.

2. Building energy modeling

Over the years, considerable effort has been invested in developing large-scale energy simulation tools, e.g. EnergyPlus, DOE-2, and TRN-SYS [15–17]. Models developed in such tools fall under the category of white-box models, which are based entirely on first-principles building physics, typically in the form of mass and energy balance equations, which often lead to large systems of Ordinary Differential Equations (ODE). However, being the case for the majority of large-scale white-box models, they require extensive information on the building and a substantial amount of resources and time to develop. Taking EnergyPlus as an example, a typical workflow consists of three time-consuming phases before an accurate energy model is obtained. First, the building geometry is defined with detailed geometry data from floor plans, cross-sections, 3D models, etc. Second, HVAC systems design data must be specified along with material and thermal properties of the envelope and occupancy, lighting, and equipment schedules of the spaces. Finally, to achieve acceptable prediction accuracy, the model must be calibrated using actual consumption data from the building, a process that often relies on a manual trial-and-error approach that requires domain-specific expertise from the modeler [18]. In addition to the high amount of resources and time that must be dedicated to the manufacturing of these high-resolution white-box models, high computational costs and execution time is also a well-known challenge, which makes them unsuitable for operational optimization purposes [19,20].

Therefore, to reduce the amount of manual work of white-box model development, considerable efforts have been put into the automatic translation of the design and geometrical data available in BIM to attain working Energy models for simulation tools (BIM2BEPS) such as EnergyPlus and Modelica. For instance, Ramaji et al. [21] proposed a BIM2BEPS transformation tool for conversion from the BIM-format International Foundation Classes (IFC) into an OpenStudio model.

Andriamamonjy et al. [22] proposed an automated BIM2BEPS workflow between IFC and Modelica. The workflow applies an intermediate representation called Model View Definition (MVD), on which different checks are performed to ensure a given IFC file includes all required information. If the IFC passes this test, it can be directly translated into a working Modelica model. While the workflow is successfully applied to a test case study, the authors also highlighted potential barriers of adoption. Namely, that it relies greatly on information that is typically not available in status quo IFC files.

Despite the significant efforts in BIM2BEPS research, many of the existing interoperability issues such as interdependence on extensive toolchains, inconsistent current practices by BIM practitioners, and loss of information still persist [23,24].

As a result of these shortcomings, data-driven modeling methods have recently gained increasing popularity for building modeling and simulation, due to their lower computational demands and a high potential for automation and integration in smart buildings where sensor networks and metered HVAC components collect large amounts of operational data.

Data-driven modeling methods are typically divided into black-box and gray-box methods. Black-box models are purely data-driven, meaning that no domain-specific model structure is assumed. Given a set of data, specific algorithms are thus applied to find appropriate functional

relationships that map system inputs to system outputs. Commonly used black-box models are Artificial Neural Networks (ANN) [25,26], Support Vector Machines (SVM) [27], Decision Trees [28], and state-space models [29]. Due to the absence of domain-specific assumptions, these types of models have been successfully applied across many academic disciplines as well as in industry [30,31].

Typically, black-box models are considered lighter than white-box models in terms of computational complexity and simulation speed. For existing buildings with available historical data, black-box models have been used for aggregated heating and electricity load forecasting [32–34]. These types of models are usually referred to as *monolithic models*, i.e. whole-building models that do not incorporate any knowledge of the modeled building, besides historical weather and operational data such as heating or electricity consumption. While this offers convenient and fast implementation, there are also certain drawbacks to consider. If the building operation or design is changed, e.g. if system setpoints are changed or if the building is retrofitted, new data have to be collected and the model has to be re-trained. Furthermore, if the data available is scarce or of low quality, the performance of black-box models might be affected. A promising solution to this issue is the use of *transfer learning*, a technique where knowledge is transferred from one domain to another [35]. Thereby, the knowledge contained in models that are trained on large amounts of data can be reused for training other models, given that the domains have a certain amount of overlap. This concept has been applied for both indoor temperature prediction [36] and energy prediction [37].

One of the often highlighted drawbacks of black-box models is the lack of interpretability, i.e. the model structure and the obtained parameters have no direct physical meaning and the underlying cause of model behavior is thus unknown. This is opposed to gray-box models, the second category of data-driven models. As the name suggests, these types of models are a mix of white-box and black-box models where the model structure is based on physical principles, while the unknown parameters of this model structure are estimated through parameter identification techniques. The physics-derived model structure enables interpretation and validation of the obtained parameters from parameter estimation in a physical context, e.g. as demonstrated by Macarulla et al. [38]. An example of the use of gray-box modeling in buildings is the thermal Resistor Capacity (RC) model, a thermal analog to electrical circuits used to model the thermal dynamics of buildings [39,40]. However, gray-box modeling is a general concept that can also be applied to individual building components such as heat pumps, recovery units, thermal storage tanks, cooling coils, etc., as demonstrated by Afram et al. [41].

In general, the various recent studies on data-driven modeling of building systems and components using black-box and gray-box approaches report clear benefits regarding automation potential and flexibility as well as high prediction accuracy. Considering the discussed challenges of BIM2BEPS automation, these are key properties for applications in a dynamic DT-environment where close integration between data collected from IoT networks and simulation models is fundamental for delivering the outlined DT services.

3. Ontologies and metadata schemes for improved interoperability and building energy modeling

Buildings have a high degree of heterogeneity and complexity, including a wide variety of systems and devices from different contractors and providers. This makes it difficult to develop applications and models that can be scaled and reused across different buildings and use-cases. This has caused the development of ontologies, which attempts to standardize and structure metadata and communication between devices to avoid the status-quo many-to-many translation between the multitude of protocols used.

Pritoni et al. [42] provided a review of 40 different ontology and metadata schemas for building design, energy modeling, and building

operation applications. From these 40 different ontologies, five of the most popular ontologies were selected for a deeper comparison based on their applicability in three use cases; Energy auditing, automated fault detection and diagnosis, and optimal control of HVAC systems. The selected ontologies were SAREF, Semantic Sensor Network/Sensor, Observation, Sample, and Actuator (SSN/SOSA), Building Topology Ontology (BOT), Brick, and RealEstateCore (REC).

Generally, it was highlighted that all five ontologies had varying degrees of missing concepts, which impacted their utility in the three use cases, i.e. choosing the right ontology depends on its intended application. For instance, BOT and SSN/SOSA missed central concepts for describing sensed properties, actuator idioms, and units of measurement. BOT, Brick, and REC missed equipment properties to describe e.g. nominal capacities, flowrates, efficiencies, etc. All ontologies except Brick missed schedule concepts for describing control strategies. Generally, Pritoni et al. conclude that none of the five ontologies accounts for all required concepts and that the ontologies must be customized and extended for certain applications.

In this work, SAREF is chosen as a backbone for the energy modeling framework. SAREF was developed in 2015 by Daniele et al. [43] in close collaboration with industry to define a unified reference for recurring concepts and relations in the IoT domain. The core SAREF ontology revolves around the concept of devices, e.g. a lighting switch, sensor, or meter. The ontology has extensions spanning across multiple domains, e.g. SAREF4BLDG [44], SAREF4CITY [45], SAREF4SYST [46], SAREF4ENER [47], etc. Here, SAREF4BLDG is an extension dedicated to the building domain and is based on the ISO-published Industry Foundation Classes (IFC) standard [48]. The extension has 72 classes and 179 object properties that represent typical appliances and devices in buildings. Hence, this ontology allows an almost direct mapping between IFC and SAREF for most of these objects within the building domain. This is a great advantage compared with other relevant ontologies, which could potentially allow for significant reuse of existing well-structured information, contained in the frequently available IFC models for buildings. Another ontology that also has IFC mapping capabilities is IFCowl [49]. However, as discussed by Rasmussen et al. [50], IFCowl was designed to be backwards compatible with IFC and as a result inherits two main drawbacks; complexity and size. This makes the ontology both hard to understand and hard to use. In comparison, SAREF has a higher degree of modularity and its extension SAREF4BLDG is lighter in complexity and size. In addition, the multiple domains and broader scope of SAREF could allow for further expansion of the DT concept to consider larger systems such as districts or cities at a later stage.

As is the case with the components in IFC, the component classes specified in SAREF4BLDG are data containers with no formal or mathematical description of their dynamic behavior or interaction with other components or systems. For instance, a `s4bldg:SpaceHeater` object holds the property `s4bldg:outputCapacityop`, but it has no attached model which specifies how this property should be used to simulate actual heat consumption.

Therefore, to build upon the SAREF4BLDG extension and ensure high flexibility, a modular and component-based modeling approach is proposed. Using the component definitions and class hierarchies from SAREF4BLDG, the energy model is thus split into a matching set of component models that can be automatically assembled and adapted, based on the specific use case. Each component should be able to learn from the collected data through various parameter estimation techniques and strive to improve prediction accuracy as the collected data accumulates. This approach will make it possible to continuously monitor and compare the performance of actual components in real-time with their digital counterparts in the digital twin environment, as required by **Service 2**. Furthermore, a highly modular and flexible model allows for the interchange of component models during the entire building lifecycle, which is necessary to maintain the Digital Twin as a close replica of the actual building, as the installed systems or building use might change over time. Therefore, the component models should be accurate enough

to meet the defined services to a satisfactory level, but simple enough to allow for robust and automated model calibration.

With a basis in the components depicted in Fig. 1, a description of the components and their properties from the SAREF4BLDG ontology is given in the Appendix Table B.3. The superscripts *op* and *dp* stand for object property and data property, respectively. An object property points to another object, while a data property holds data in the form of standard types such as integer, string, floating-point, etc. A complete list of available components and a description of the listed properties can be found in the SAREF4BLDG documentation [44].

4. Component modeling

When modeling buildings, it is common practice to take advantage of the major difference in physical time constants that are associated with the HVAC system and the building envelope and spaces. The HVAC system is often governed by fast-moving dynamics that reach equilibrium within seconds or minutes, while the large thermal inertia of the building envelope and spaces has very slow-moving dynamics. Therefore, it can often be justified to model the HVAC systems as quasi-steady-state systems, which decreases model complexity and lowers the computational burden considerably [51,52]. For these reasons, building simulation tools such as EnergyPlus is also heavily based on this modeling approach [53]. Therefore, with a basis in the presented components in Table B.3, potential candidate models have been found in the literature for implementation and detailed descriptions are included in Appendix A. The key selection criteria for the models have been simplicity and applicability for data-driven calibration and parameter estimation.

To provide an overview of all the identified models, the inputs, outputs, parameters, and constants have been summarized for each model in Table 1. The parameters for the models in Table 1 can either be given through design specifications or calculated through parameter estimation techniques if time-series data is available for the inputs and outputs. In the case where input-output data can be continuously provided, parameter estimation can be performed on a regular interval to attain a continuously improving performance on the different gray- and black-box component models. This method is demonstrated for indoor temperature forecasting by Ruano et al. [54]. Additionally, each physical component can be monitored closely, by comparing its operation with its digital replica to provide fault detection and continuous commissioning capabilities on a component level. If neither design specifications nor operational input-output data is available, default values must be assumed.

Based on the listed inputs and outputs in Table 1, and the model descriptions in the previous sections, certain expected relationships between components arise. For instance, the space heater model in Table 1 expects the massflow $\dot{m}_w$, and the room temperature T_z as input, which it can acquire from the valve model and the building space model, respectively. The valve model expects the valve position u_v as input, which it can acquire from the controller component etc. In the following section, these types of relations are formalized using simple but broadly applicable concepts from the SAREF4SYST extension.

5. Modeling framework

In the previous section, common building components were presented, along with generic and scalable component energy models. However, to express how the defined components interact, a domain-independent SAREF-compliant methodology for linking the inputs and outputs of component models is defined. The concept of having different self-contained component models defined by an input-output relationship and an overall framework that links these models share many similarities to the co-simulation framework defined in the Functional Mockup Interface (FMI) standard [55]. Here, models are exported from simulation tools as Functional Mockup Units (FMU) and are used as

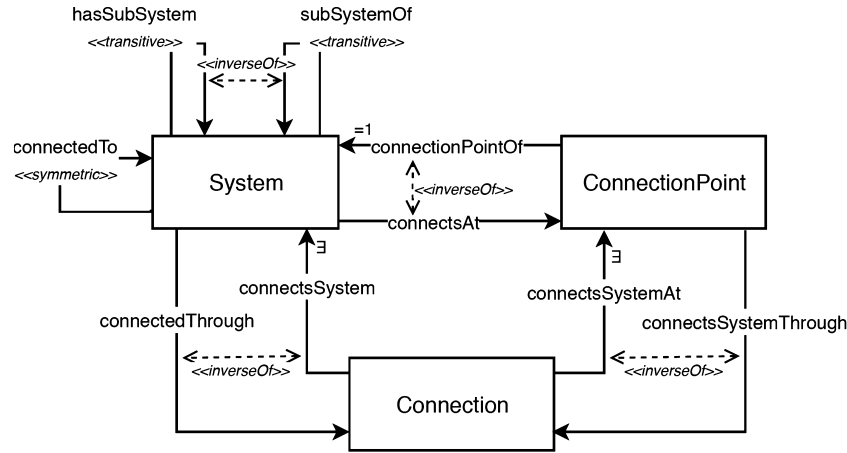


Fig. 2. Overview of the SAREF4SYST ontology extension with available classes and properties [46].

Table 1

Overview of inputs, outputs, parameters, and constants of potential mathematical models for the highlighted components.

Component	Inputs	Outputs	Parameters	Constants
Valve	$u_v \in [0, 1]$	$\dot{m}_w \in \mathbb{R}^+$	$N_v \in [0, 1]$ $\dot{m}_{w,max} \in \mathbb{R}^+$	
Space heater	$\dot{m}_w \in \mathbb{R}^+$ $T_{w,in} \in \mathbb{R}$ $T_z \in \mathbb{R}$	$\dot{Q}_h \in \mathbb{R}^+$	$UA \in \mathbb{R}^+$ $C_r \in \mathbb{R}^+$	$c_{p,w}$ Δt
Heating coil	$\dot{m}_a \in \mathbb{R}^+$ $T_{a,set} \in \mathbb{R}$ $T_{a,in} \in \mathbb{R}$	$\dot{Q}_{hc} \in \mathbb{R}^+$		$c_{p,a}$
Cooling coil	$\dot{m}_a \in \mathbb{R}^+$ $T_{a,set} \in \mathbb{R}$ $T_{a,in} \in \mathbb{R}$	$\dot{Q}_{cc} \in \mathbb{R}^+$		$c_{p,a}$
Air to air heat recovery	$\dot{m}_{a,sup} \in \mathbb{R}^+$ $\dot{m}_{a,exh} \in \mathbb{R}^+$ $T_{a,sup,in} \in \mathbb{R}$ $T_{a,exh,in} \in \mathbb{R}$ $T_{a,set} \in \mathbb{R}$	$T_{a,sup,out} \in \mathbb{R}$	$\epsilon_{75\%,h} \in [0, 1]$ $\epsilon_{75\%,c} \in [0, 1]$ $\epsilon_{75\%,h} \in [0, 1]$ $\epsilon_{100\%,c} \in [0, 1]$ $m_{a,max}$	$c_{p,a}$
Fan	$\dot{m}_a \in \mathbb{R}^+$	$\dot{W}_{fan} \in \mathbb{R}^+$	$c_1 \in \mathbb{R}$ $c_2 \in \mathbb{R}$ $c_3 \in \mathbb{R}$ $c_4 \in \mathbb{R}$ $\dot{m}_{a,max}$ ΔP_{max} $\eta_{tot} \in [0, 1]$	ρ_a
Damper	$u_d \in [0, 1]$	$\dot{m}_a \in \mathbb{R}^+$	$a \in \mathbb{R}$ $b \in \mathbb{R}$ $c \in \mathbb{R}$	
Controller	$y_{set} \in \mathbb{R}$ $y_{meas} \in \mathbb{R}$	$u \in [0, 1]$	$K_p \in \mathbb{R}$ $K_i \in \mathbb{R}$ $K_d \in \mathbb{R}$	
Building Space	$N_{occ} \in \mathbb{N}$ $\dot{m}_{a,sup} \in \mathbb{R}^+$ $\dot{m}_{a,exh} \in \mathbb{R}^+$ $T_o \in \mathbb{R}$ $\Phi_s \in \mathbb{R}^+$ $u_v \in [0, 1]$ $u_d \in [0, 1]$ $u_s \in [0, 1]$	$C_z \in \mathbb{R}$ $T_z \in \mathbb{R}$	$K_{occ} \in \mathbb{R}^+$ $m_z \in \mathbb{R}^+$ [-]	Δt

so-called slaves. The slaves are allowed to communicate at specific time intervals called communication points, managed by a master algorithm. Between communication points, the slaves are simulated independently. The modeling framework presented in this section shares many of the same principles as co-simulation and is thus well-aligned with the use of FMUs for modeling individual components.

5.1. Linking component models using SAREF4SYST

SAREF has a dedicated extension called SAREF4SYST, which aims at providing a generic framework for representing the topology of systems and how they interact [46]. Through SAREF4SYST, the system topology can be expressed through three generic classes, `s4syst:System`, `s4syst:Connection`, and `s4syst:ConnectionPoint` along with nine applicable properties as shown in Fig. 2.

As shown in Fig. 2, a `s4syst:System` object can be a subsystem of other `s4syst:System` objects specified through the property `s4syst:subSystemOf` or have subsystems itself specified through `s4syst:hasSubSystem`. Various forms of flows between systems, e.g. information or physical quantities can be expressed through the `s4syst:Connection` and `s4syst:ConnectionPoint` classes. In this work, the `s4syst:Connection` class is used to represent system outputs and the `s4syst:ConnectionPoint` class to represent system inputs. Hence, all inputs and outputs in Table 1 can be represented by `s4syst:ConnectionPoint` and `s4syst:Connection` objects respectively, while each component can be represented by `s4syst:System` objects. This concept is shown in Fig. 3, where three SAREF4BLDG components are shown. As recommended in the SAREF4BLDG documentation [56], each component is extended with its identified model as a specialized subclass from section 4. Furthermore, these components all inherit from the `s4syst:System` superclass, enabling them to interact and communicate through the SAREF4SYST ontology. It should be noted that the full SAREF4BLDG class hierarchy is not included for clarity.

In Fig. 3, simple information flows between the three components are shown. The `s4bldg:BuildingSpace` component points to a `s4syst:Connection` object through the `s4syst:connectedThrough` property. This `s4syst:Connection` object represents the indoor temperature output T_z of the `s4bldg:BuildingSpace` model extension. Through the `s4syst:connectsSystemAt` property, it further points to a `s4syst:ConnectionPoint` object, which represents the input T_z of the `s4bldg:SpaceHeater` model extension. This relation is expressed through the `s4syst:ConnectionPoint` property of the `s4syst:ConnectionPoint` object. The same concepts are then applied to relate the `s4bldg:Valve` water massflow output $\dot{m}_w$ to the `s4bldg:SpaceHeater` input $\dot{m}_w$. Altogether, this example shows how information is translated from model output to model input. Correspondingly, the inverse path from model input to model output is expressed through the `s4syst:connectsAt`, `s4syst:connectsSystemThrough`, and `s4syst:connectsSystem` properties as shown in Fig. 2. This generic framework and the translation from output to input and vice versa at the `s4syst:Connection` and `s4syst:ConnectionPoint` objects provide the flexibility to exchange information between any type of models independently of model formulation and

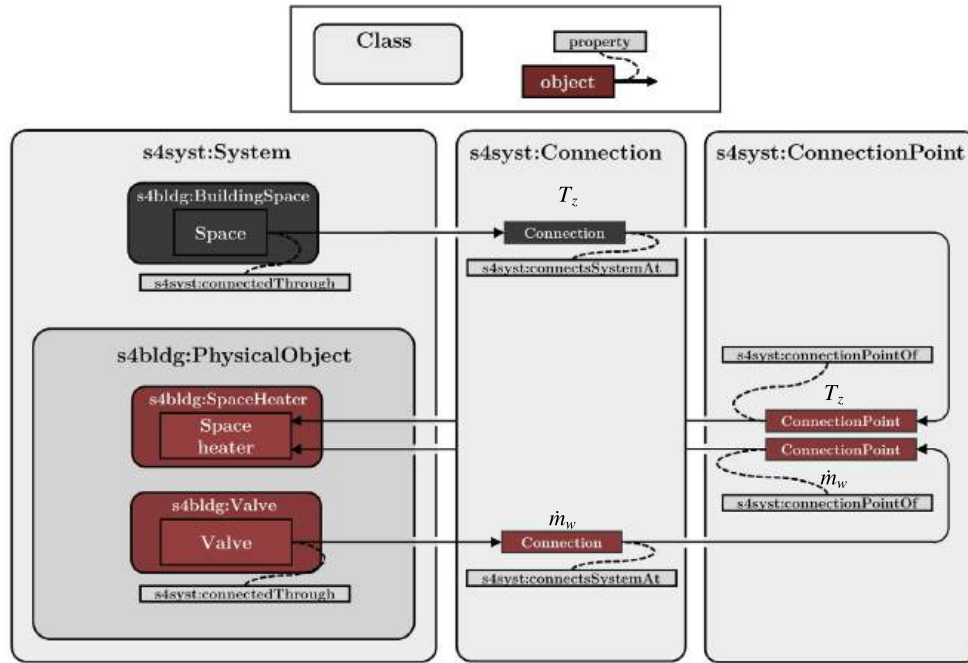


Fig. 3. Example of data structure showing how SAREF4SYST is used to link models contained in the SAREF4BLDG components through the `s4syst:System`, `s4syst:Connection`, and `s4syst:ConnectionPoint` classes.

input-output terminology. Hence, the presented framework is not specific to the models shown in Table 1, which could be modified or extended to cover other components such as `s4bldg:Chiller` or `s4bldg:Pump`, where models from external simulation tools and libraries could be considered, e.g. as FMUs.

5.2. Model execution order

The topology in Fig. 3 defines the required information exchange between components at a given timestep. This produces a scheduling problem, where it must be ensured that all required inputs are available when a model is executed during a simulation time step. For instance, in Fig. 3, both the `s4bldg:BuildingSpace` and the `s4bldg:Valve` model must be executed before the `s4bldg:SpaceHeater` model can be executed. Such task scheduling problems are well-known in computer science and are formally known as topological sorting [57]. Given a directed graph $G = (V, E)$, with V being the set of nodes and E being the set of edges, topological sorting is the ordering of nodes such that for every edge $(u, v) \in E$ the node u comes before the node v . This is only possible if the graph has no cycles, i.e. if it is a Directed Acyclic Graph (DAG), in which case it is guaranteed to have at least one solution [57,58]. Hence, if the `s4syst:System` instances are interpreted as nodes, and the `s4syst:Connection`, and `s4syst:ConnectionPoint` pairs are interpreted as directed edges, the correct execution order can be found using topological sorting, given that the graph is acyclic.

5.3. Cycle removal procedure

The no-cycles condition of topological sorting is essential to consider when applying the presented methodology to model actual buildings as feedback control loops are an integral part of all building automation systems. When such control loops are modeled, cycles are introduced. An example of a cycle produced by a controller is shown in Fig. 4. Here, a `s4bldg:BuildingSpace` model predicts CO_2 -concentration C_z , which is used as input y in a `saref:Sensor` model. The output y_{meas} is then used as input in a `s4bldg:Controller` model. An appropriate input signal u is then computed, which is sent to a

`s4bldg:Damper` model to calculate the resulting airflow m_a . Finally, this value is sent back to the `s4bldg:BuildingSpace` model for recalculation of indoor CO_2 -concentration C_z at the next time step, completing the cycle. To systematically deal with such cycles, an algorithm is proposed and is presented in Algorithm 1. Given the set of all `s4syst:System` objects H and an initially identical set H^* , the algorithm iterates through the set of all `s4bldg:Controller` components $U \in H^*$, identifies reachable components V_u from $u \in U$, and removes all edges from $v \in V_u$ to the controlled component u . `controlsProperty.isPropertyOf` if such edges exist. For instance, applying Algorithm 1 on the simple system shown in Fig. 4 the Supply damper would be identified as a reachable component from the CO_2 controller. The Supply damper is connected with the controlled component Space, and this connection is thus removed.

5.4. Topological sorting procedure

After applying Algorithm 1, topological sorting can be applied on the modified set of components H^* . This is shown in Algorithm 2 where an adaptation of Kahn's topological sorting algorithm is presented, which matches the semantics of SAREF4SYST. The input of this algorithm is the set H^* , while the output is a sequence L with all the components in H in a topologically sorted order. The algorithm recursively adds all mapped components $f(s) \in H$ with $|s.\text{connectsAt}| = 0$ (i.e. the component s has no inputs) to a sequence L while removing all edges originating from s . In addition, a sequence P is created which holds integers starting from 0 that represent the execution priority of the components in L . Components in L with the same priority can be executed in parallel. For instance, applying Algorithm 2 on the simple system shown in Fig. 3 produces the sequence $L = (\text{Valve}, \text{Space}, \text{Space heater})$ and the priority sequence $P = (0, 0, 1)$. This implies that the Valve and Space models can be executed in parallel and that they must be executed before the Space heater model.

5.5. Simulation procedure

When the execution sequence L is found using Algorithm 2, the system can be simulated by simply traversing through this sequence

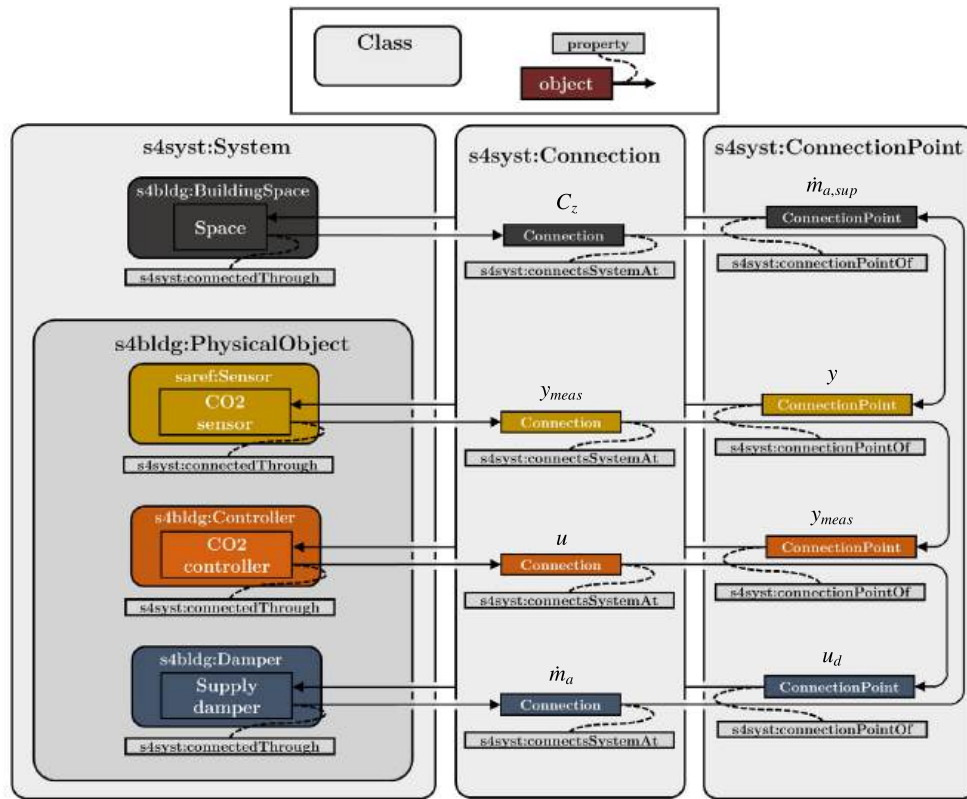


Fig. 4. Example of a cycle produced by a controller, which actuates the opening position of a supply airflow damper of a space. Some model inputs are neglected for clarity (e.g. the setpoint for the controller).

Algorithm 1 Removes cycles created by feedback control loops. $\text{DFS}(V, s)$ is a recursive depth-first search to find the set of reachable components V from the component s .

```

1: Let  $H$  be the set of all instances with superclass s4sys:System
2: Let  $H^*$  contain a copy of all the elements in  $H$  such that the elements in  $H^*$  can be modified to contain no cycles
3: Let  $U$  be the subset of all instances with the class s4bldg:Controller of  $H^*$ 

4: function  $\text{DFS}(V, s)$ 
5:    $V \leftarrow V \cup \{s\}$ 
6:   for all  $c \in s.\text{connectedThrough}$  do
7:      $cp \leftarrow c.\text{connectsSystemAt}$ 
8:      $s^* \leftarrow cp.\text{connectionPointOf}$ 
9:     if  $s^* \notin V$  then
10:       $V \leftarrow \text{DFS}(V, s^*)$ 
11:   end if
12: end for
13: return  $V$ 
14: end function

15: for all  $u \in U$  do
16:    $w_u \leftarrow u.\text{controlsProperty.isPropertyOf}$ 
17:    $V_u \leftarrow \text{DFS}(\emptyset, u)$ 
18:   for all  $s \in V_u$  do
19:     for all  $c \in s.\text{connectedThrough}$  do
20:        $cp \leftarrow c.\text{connectsSystemAt}$ 
21:        $s^* \leftarrow cp.\text{connectionPointOf}$ 
22:       if  $w_u = s^*$  then
23:          $w_u.\text{connectsAt} \leftarrow w_u.\text{connectsAt} \setminus \{cp\}$ 
24:          $s.\text{connectedThrough} \leftarrow s.\text{connectedThrough} \setminus \{c\}$ 
25:       end if
26:     end for
27:   end for
28: end for

```

of components, while gathering the needed inputs and executing the attached model. Executing this routine once completes one system timestep of length Δt and can be repeated indefinitely given the re-

Algorithm 2 A SAREF4SYST-specific adaptation of Kahn's topological sorting algorithm [59].

```

1: Let  $f : H^* \rightarrow H$  be a function that maps from the modified components  $h^* \in H^*$  to the original components  $h \in H$ 
2: Let  $S$  be the subset of  $H^*$  with no inputs, i.e.
    $S = \{s \in H^* : |s.\text{connectsAt}| = 0\}$ 
3: Let  $L \leftarrow ()$  be an empty sequence which will be populated with the components of  $H$  in a sorted order
4: Let  $P \leftarrow ()$  be an empty sequence which will be populated with the priority of the components in the sequence  $L$ 
5: Let  $p \leftarrow 0$  be the priority counter
6: The notation  $X^\frown(x)$  is used to append the element  $x$  to the sequence  $X$ 

7: while  $|S| > 0$  do
8:    $S^* \leftarrow \emptyset$ 
9:   for all  $s \in S$  do
10:     $L \leftarrow L^\frown(f(s))$ 
11:     $P \leftarrow P^\frown(p)$ 
12:    for all  $c \in s.\text{connectedThrough}$  do
13:      for all  $cp \in c.\text{connectsSystemAt}$  do
14:         $s^* \leftarrow cp.\text{connectionPointOf}$ 
15:         $s^*.\text{connectsAt} \leftarrow s^*.\text{connectsAt} \setminus \{cp\}$ 
16:        if  $|s^*.\text{connectsAt}| = 0$  then
17:           $S^* \leftarrow S^* \cup \{s^*\}$ 
18:        end if
19:      end for
20:    end for
21:  end for
22:   $p \leftarrow p + 1$ 
23:   $S \leftarrow S^*$ 
24: end while

```

quired inputs, such as weather data and schedule values, are available. This is formulated in Algorithm 3, where each component model is assumed to have an input and output dictionary which are indexed by the properties `inputName` and `outputName` of the associated `s4sys:ConnectionPoint` and `s4sys:Connection` objects, respectively. These property names are simply the mathematical notation

for the inputs and outputs shown in Table 1 translated into a representative name, e.g. “indoorTemperature” to represent T_z . In addition, each component model is assumed to have a `do_step()` method, implementing the governing equations defined in section 4 or executing external models or libraries such as exported FMU models.

Algorithm 3 An algorithm for simulating building devices and systems given the execution sequence L .

```

1: Let  $t_{start}$  and  $t_{end}$  be the start and end simulation time
2: Let  $\Delta t$  be the simulation time step
3:  $t \leftarrow t_{start}$ 

4: while  $t < t_{end}$  do
5:   for  $i \leftarrow 1 \dots |L|$  do
6:      $s \leftarrow L_i$ 
7:     for all  $cp \in s.connectsAt$  do
8:        $c \leftarrow cp.connectsSystemThrough$ 
9:        $s^* \leftarrow c.connectsSystem$ 
10:       $s.input[cp.inputName] \leftarrow s^*.output[c.outputName]$ 
11:    end for
12:     $s.output \leftarrow s.do\_step(s.input)$ 
13:  end for
14:   $t \leftarrow t + \Delta t$ 
15: end while

```

6. Demonstration case and implementation

To validate the proposed modeling framework, the single-zone system shown in Fig. 1 is considered as a demonstration case. The modeling framework is implemented in Python as a reusable library [60], following the semantic structure and guidelines provided by the SAREF ontology. However, to accurately model the system, basic topological information is required. To ensure that the model can easily be constructed and adapted without manually specifying all the logical connections between models as presented in section 5, a simple model-independent representation is instead used as a starting point. Based on this representation, the logical connections between models can then be inferred based on simple rules. For this purpose, the single-zone system has been formulated using the SAREF ontology semantics and concepts as shown in Fig. 5 only considering the actual topology of the physical system.

6.1. Available topology information

The system consists of one `s4bldg:BuildingSpace` instance, containing one `s4bldg:SpaceHeater`, one `s4bldg:ShadingDevice`, one `s4bldg:Valve`, two `s4bldg:Damper`, two `s4bldg:Controller`, and two `saref:Sensor` instances as represented by the `s4bldg:contains` property. The `s4bldg:BuildingSpace` instance is linked with one `saref:Temperature` instance representing indoor temperature and one `saref:CO2` instance representing CO₂-concentration through `saref:hasProperty`, which is inherited through the `saref:FeatureOfInterest` super class. The `saref:CO2` subclass of `saref:Property` has been added as an extension to the core SAREF ontology similar to the work carried out by Weerdt et al. [61], as it is not included by default. The measured properties of the two `saref:Sensor` instances are then determined through the `saref:measuresProperty`, pointing either to the `saref:Temperature` instance or the `saref:CO2` instance. Similarly, fashion, the controlled properties of the `s4bldg:Controller` instances are determined through the `saref:controlsProperty`. To determine which devices are actuated by the controllers, each `s4bldg:Controller` instance is associated with a `saref:LevelControlFunction` instance (subclass of `saref:Function`) through the `saref:hasFunction` property, which further points to a `saref:SetLevelCommand` instance (subclass of `saref:Command`) through the `saref:hasCommand` property. Finally, the `saref:SetLevelCommand` instances

each point to a `saref:MultiLevelState` instance (subclass of `saref:State`) through the `actsUpon` property. The specific subclasses used here indicate that the actuated `s4bldg:Damper` and `s4bldg:Valve` instances can be regulated with a level-setting, e.g. from 0-100%. Alternatively, for representing binary control of other objects such as windows or doors, the `saref:OpenCloseFunction`, `saref:OpenCommand`, `saref:CloseCommand`, `saref:OpenState`, and `saref:CloseState` could be used instead. For the shown system, the supply and exhaust dampers share the same state, resulting in balanced supply and exhaust airflows. However, for controlling supply and exhaust airflows separately, each damper instance should have its own `saref:MultiLevelState` instance each associated with two different `s4bldg:Controller` instances.

To describe the topology of the ventilation system components and their order of placement in the flow path, the `s4sys:connectedTo` property is used, which is inherited from the `s4sys:System` class. This property is symmetric, meaning that two connected components both point to each other with the `connectedTo` property (although only one path is shown in Fig. 5). Hence, the `s4bldg:AirToAirHeatRecovery` instance is connected to both the supply and exhaust `s4bldg:Fan` instances while the exhaust fan is further connected to an exhaust `Node` instance. Here, the `Node` class represents nodes in a flow network and keeps track of flow and temperature before one stream branches out or after multiple streams join together. In the demonstration case, there is only one space, resulting in the same flow properties before and after the `Node` instances. However, this addition is necessary in the case of multiple spaces and a more complex flow path topology. The exhaust `Node` instance is further connected to an exhaust `s4bldg:Damper` instance. The supply `s4bldg:Fan` is connected to the heating `s4bldg:Coil` instance, which is connected to a supply `Node` instance. Finally, this supply `Node` instance is connected to a supply `s4bldg:Damper` instance.

For Fig. 5, inverse properties also exist for most of the shown properties, although these are omitted for clarity. For example, `s4bldg:isContainedIn` and `saref:isMeasuredByDevice` are inverse properties of `s4bldg:contains` and `saref:measuresProperty`.

6.2. Obtaining a working energy model

With the SAREF representation in Fig. 5 of the considered single-zone system, the topological context of all devices and components is defined. However, this representation cannot be directly used with the presented algorithms in section 5 to simulate the system. Therefore, inputs and outputs of each model extension, as described in section 5, are first matched using the topological relationships shown in Fig. 5 and the `s4sys:Connection` and `s4sys:ConnectionPoint` classes as demonstrated in section 5.

For example, a connection between the `s4bldg:BuildingSpace` instance and the CO₂ `saref:Sensor` instance is inferred through the properties `saref:measuresProperty` and `saref:isPropertyOf` of the `saref:Sensor` and `saref:CO2` instance, respectively. A connection between the `saref:Sensor` and `s4bldg:Controller` is established through the `saref:controlsProperty` and `saref:isMeasuredByDevice` properties.

The resulting connections are shown in Fig. 4. Completing this process for all `s4sys:System` instances produces the graph shown at the top of Fig. 6. Here, each pair of `s4sys:Connection` and `s4sys:ConnectionPoint` instances linking two models are represented as one edge labeled with the shorthand “C:” and “CP:” representing the `outputName` and `inputName` properties, respectively. As mentioned previously, these property names correspond to the mathematical notation of input and outputs in Table 1. In addition to the already discussed `Node` class, two other helper classes `OutdoorEnvironment`, and `Schedule` have also been introduced. `OutdoorEnvironment` represents relevant outdoor conditions such as outdoor temperature and solar irradiation, which affects different systems, in this

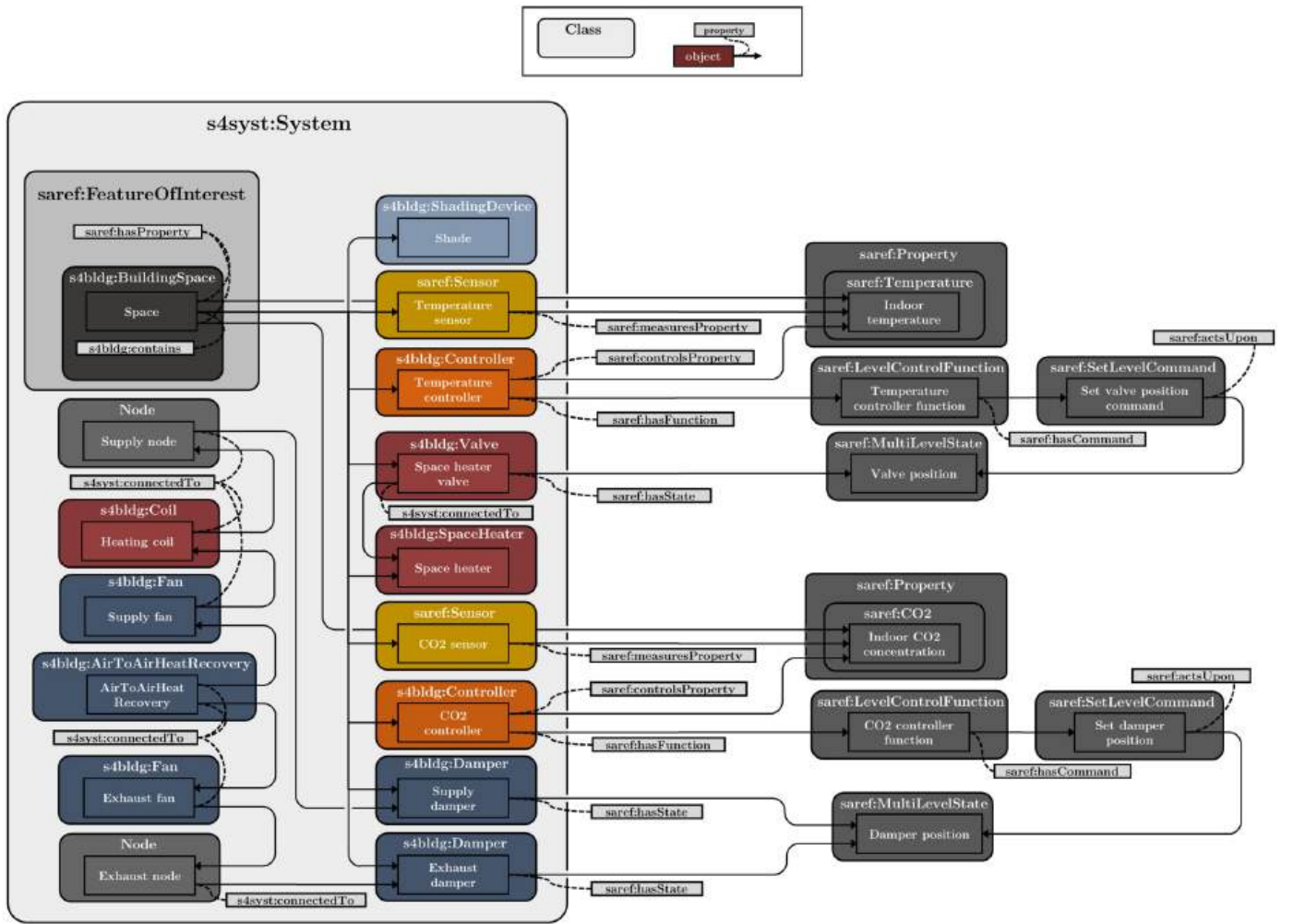


Fig. 5. SAREF representation of the demonstration case, including the required topology information for the involved components. The representation incorporates concepts from both SAREF core, SAREF4BLDG, and SAREF4SYST.

case, the `s4bldg:BuildingSpace` and `s4bldg:AirToAirHeatRecovery` instances. *Schedule* is a generic class that outputs predetermined schedules for a specific period, e.g. daily occupancy or setpoint profiles.

With the inputs and outputs defined for all components in the system, the three algorithms from section 5 can be applied to simulate the system. An overview of this workflow is shown in Fig. 6. As a first step, Algorithm 1 is applied to the set of components H , to identify and remove cycles caused by the controllers. This results in the updated set H^* , where four edges caused by the two controllers are removed. As shown in Fig. 6, all four removed edges point to the `s4bldg:BuildingSpace` instance, which holds the controlled properties. The obtained set of components H^* is then used as input for Algorithm 2 from which a sequence L is obtained, containing the components of H in a topologically sorted order. This sequence is then used in Algorithm 3 for simulating the original system H . To run this simulation, parameters have been specified for each model as listed in Table B.2. The parameters are based on a classroom in an actual university building. This is the case for the temperature prediction model for the `s4bldg:BuildingSpace` instance, which was trained on actual data collected from the classroom as described in our previous work [62,63]. For other components such as the `s4bldg:Fan` and `s4bldg:AirToAirHeatRecovery` instances, standard values, suitable for the demonstration case, are used. It should be stressed that the simulation is not intended to validate the accuracy of the presented component models, but rather to provide a proof-of-concept and

validation of the overall modeling framework, information carryover, algorithms, and workflow.

6.3. Simulation results

For simulation, a 4-day winter period using actual historical weather data from Denmark is considered. The results for some of the components are shown in Fig. 7 where the output of each model is shown with the black dashed line while the colored lines show model inputs. The weather data used for the simulation are shown in Fig. 7i as a separate plot, characterized by low solar irradiance and temperatures.

Results for the temperature and CO2 models attached to the `s4bldg:BuildingSpace` instance are shown in the two top plots, Fig. 7a, and Fig. 7b. Fig. 7c shows the inputs and outputs of the temperature `s4bldg:Controller` instance. As shown, a constant 23°C setpoint $T_{z,set}$ is provided as input to the `s4bldg:Controller` model during the whole period, resulting in a fluctuating valve position, which closes when the indoor temperature T_z exceeds $T_{z,set}$ and opens when the temperature drops below $T_{z,set}$. From Fig. 7e, the resulting heat consumption $\dot{Q}_h$ is shown for the `s4bldg:SpaceHeater` instance following the same fluctuating pattern, based on the water massflow $\dot{m}_w$.

Due to the low solar irradiation in the winter period, the shading position u_s has a limited effect on the indoor temperature and is therefore set to 0 during the entire period. However, as expected, running simulations for summer periods, it is observed that shade position has a high impact on indoor temperature, due to the higher irradiation levels.

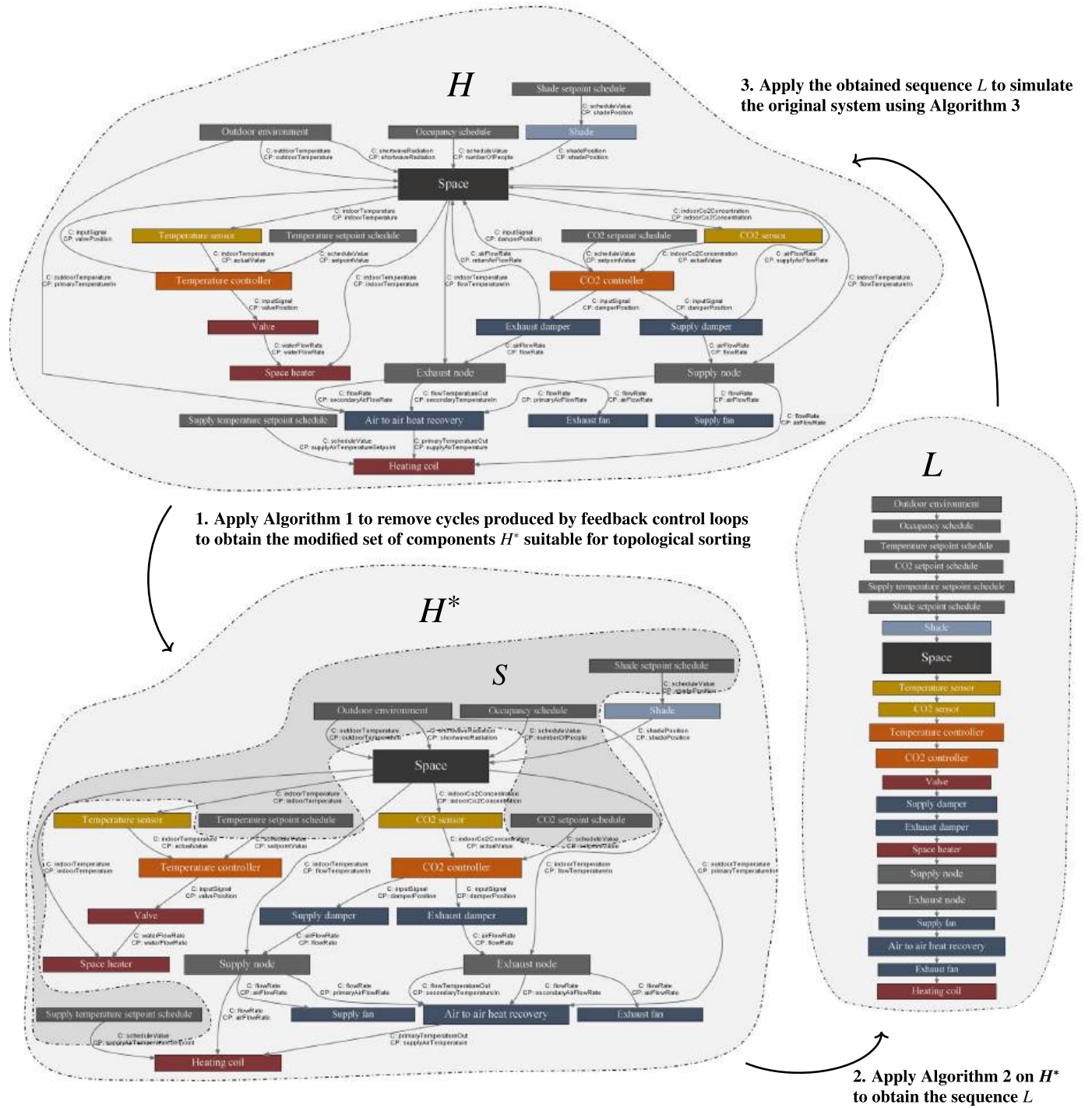


Fig. 6. Overview of Algorithm 1, Algorithm 2, and Algorithm 3 applied to the demonstration case.

Fig. 7b shows the simulation results for the CO₂ model attached to the s4bldg:BuildingSpace instance, where an artificially generated occupancy profile N_{occ} is used as input. Here, $\dot{m}_a$ represents both the supply and exhaust airflow as they are equal. Fig. 7d shows the CO₂ setpoint $C_{z,set}$ is set constant at 600 ppm, which the controller tracks by inversely actuating the supply and exhaust damper positions u_d . As shown in Fig. 7b, the predicted CO₂-concentration C_z is sensitive to rapid changes in occupancy and airflow, causing the tracking error to occasionally approach 100 ppm during occupancy periods.

The predicted power consumption $\dot{W}_{fan}$ for the supply s4bldg:Fan instance is shown in Fig. 7f. As expected, this consumption closely follows the supply airflow $\dot{m}_a$. The exhaust fan power consumption is

identical due to identical models and balanced supply and exhaust airflows and is therefore not shown.

Fig. 7g shows the predicted outlet temperature at the supply side $T_{a,sup,out}$ of the s4bldg:AirToAirHeatExchanger instance. The inputs here are inlet temperature $T_{a,sup,in}$, i.e. the outdoor temperature, the exhaust inlet temperature $T_{a,exh,in}$, i.e. the indoor air temperature predicted by the s4bldg:BuildingSpace model, and the airflows at the supply and exhaust side $\dot{m}_a$.

The simulated heat consumption for the heating s4bldg:Coil instance $\dot{Q}_{hc}$ is shown in Fig. 7h. The consumption closely follows the trend of the supply airflow $\dot{m}_a$, due to a constant setpoint $T_{a,set}$ and a relatively invariant inlet air temperature $T_{a,in}$ between 18–20 °C.

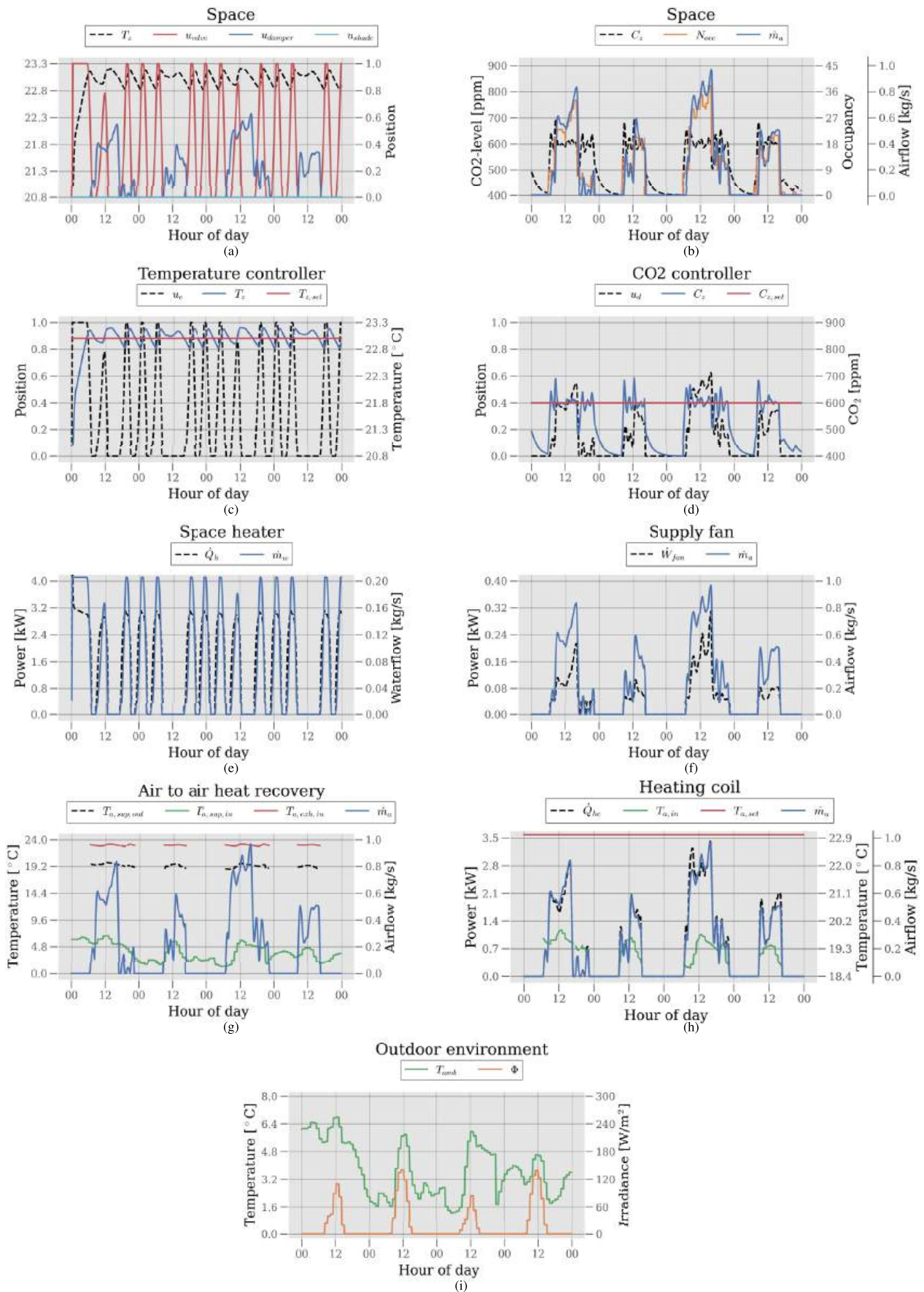


Fig. 7. Results of a 4-day winter period simulation for the demonstration case, employing the presented component models and modeling framework. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

6.4. Discussion

As demonstrated from the generated results, the modeling framework enables direct dynamic simulation of the components and systems represented through SAREF semantics. Although the demonstration takes basis in a simple single-zone system, the framework is not case-specific and can readily be applied to larger systems.

The majority of the presented component models are validated through other studies and implemented in simulation tools. However, their performance and effectiveness, as part of the presented modeling framework and a dynamic DT environment, are important to establish. Implementation of the framework in a real building case study is therefore planned as part of the next steps. Here, model parameters should be specified through a combination of available design data, and parameter estimation using measured operational data. The semantic alignment between SAREF4BLDG and IFC potentially provides an advantage in the task of obtaining the required design information and properties of the modeled devices and systems.

In addition, given the fact that sensor and meter equipment is virtually represented through the `saref:Sensor` and `saref:Meter` classes, the task of retrieving and comparing actual readings from the physical system and simulated readings from the virtual system for the same device is conceptually straightforward, which facilitates the implementation of continuous commissioning and parameter estimation. Relating to the demonstration case, the virtual temperature `saref:Sensor` instance reads the simulated indoor temperature shown in Fig. 7a while the CO₂ `saref:Sensor` instance reads the simulated CO₂-concentration shown in Fig. 7b. These readings are logged as `saref:Measurement` instances and retrieved through the `saref:hasFunction` and `saref:hasSensingRange` properties. Therefore, if actual readings from the two physical sensors are represented using the same concepts, the primary communication point from the physical system to the virtual system would thus pass through the physical and virtual `saref:Sensor` and `saref:Meter` instances. Here, robust and automated live communication with the physical IoT devices and preprocessing of the data will play a crucial role in the overall DT solution for proper use in continuous commissioning and parameter estimation.

7. Conclusion and future work

As part of the global energy transition, increasing the smartness of buildings is a key milestone, improving automation, efficiency, and flexibility. As part of this transformation, increased availability, sharing, management, and use of data from IoT devices and sensor networks both entail new opportunities as well as challenges.

Inspired by the demonstrated success and applications of the DT concept in other fields and industries undergoing a similar digital transformation as the building sector, the building DT is now emerging as a promising solution to handle and utilize these data streams for different important services such as fault detection, operational optimization, and strategy planning.

One of the fundamental constituents of a building DT which enables this is a dynamic simulation model. However, to avoid manual and time-consuming modeling workflows, and to ensure the adaptability and scalability of building DTs, close integration between the simulation models and IoT-devices at the building site is essential.

This work presents an innovative energy modeling framework that builds directly upon the semantics of the SAREF ontology with an emphasis on modular data-driven models. With a basis in the SAREF4BLDG extension, different candidate models were considered as extensions for a selection of classes. Using the existing ontology patterns from the SAREF4SYST extension, a generic method for systematically linking model inputs and outputs between components was presented. To enable direct dynamic simulation of the different components and systems, three essential algorithms were then presented. As a proof-of-concept, the framework was applied to model and simulate a generic

single-zone system with common building components and systems. The presented framework is considered the foundation for future work that will consider an actual building as a case study with the integration of actual devices, real-time sensor equipment, parameter estimation, and delivery of different services. This will demonstrate the feasibility of using the DT services to enhance the efficiency and smartness of buildings.

The use of SAREF ontology as a semantic backbone for the modeling framework improves interoperability and integration with the actual physical system and devices and avoids translation to simulation software. This potentially enables broader adoption and scalability of building DTs. In addition, using the multiple domains covered by the SAREF extensions such as industry, energy, water, smart cities, etc., the DT concept could potentially be expanded to consider streets, districts, and cities.

CRedit authorship contribution statement

Jakob Bjørnskov: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Muhyiddine Jradi:** Conceptualization, Methodology, Writing - review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Jakob Bjoernskov reports financial support was provided by Danish Energy Agency.

Data availability

Data will be made available on request.

Acknowledgements

This work is carried out under the ‘Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle’ project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64021-1009.

Appendix A

A.1. Valve

The valve model is based on the model used by Zajic et al. [64], and is given by the equations (A.1)-(A.2).

$$\tilde{u}_v = \frac{u_v}{\sqrt{u_v^2 (1 - N_v) + N_v}} \quad (\text{A.1})$$

$$\dot{m}_w = \tilde{u} m_{w,max} \quad (\text{A.2})$$

In equation (A.1), the effectively applied stem position $\tilde{u}$ is calculated based on the control input $u \in [0, 1]$, and the valve authority $N_v \in [0, 1]$. Following, the massflow can then be calculated in equation (A.2) using the maximum massflow $\dot{m}_{w,max}$.

A.2. Space heater

The space heater model is based on the energy balance equation (A.3).

$$C_r \frac{d\bar{T}_r}{dt} = \dot{m}_w c_{p,w} (T_{w,in} - T_{w,out}) - UA(\bar{T}_r - T_z) \quad (\text{A.3})$$

equation (A.3) represents the energy balance of the space heater with the stored heat on the left-hand side and the ingoing and outgoing heat flows on the right-hand side. Here, $\bar{T}_r$ is the effective temperature of the space heater, $T_{w,in}$ is the water inlet temperature, and $T_{w,out}$ is the

water outlet temperature. C_r is the heat capacity of the space heater, $\dot{m}_w$ is the water massflow, $c_{p,w}$ is the specific heat capacity of water, UA is the heat transfer coefficient, and T_z is the room temperature.

Xu et al. [65] models the effective temperature $\bar{T}_r$ of a space heater as one lumped element governed by equation (A.3) with a varying non-linear UA based on flow conditions. Tahersima et al. [66] consider the space heater as multiple finite elements that each are governed by equation (A.3) with a constant UA . For simplicity, this work considers the space heater as one lumped element with a constant UA and $\bar{T}_r = T_{w,out}$ is assumed, i.e. all energy transferred to the room is driven by the outlet water temperature. Through Backward Euler discretization [67] of equation (A.3), an explicit expression for the effective space heater temperature $\bar{T}_{r,t}$ at time t can thus be obtained as shown in equation (A.4).

$$\bar{T}_{r,t} = \frac{(T_{w,in} \dot{m}_w c_{p,w} + UA \bar{T}_{z,t-1}) \Delta t + \bar{T}_{r,t-1} C_r}{(UA + \dot{m}_w c_{p,w}) \Delta t + C_r} \quad (A.4)$$

Here, Δt is the timestep and $\bar{T}_{r,t-1}$ is the effective space heater temperature from the previous timestep. Based on the obtained space heater temperature $\bar{T}_{r,t}$, the heat transferred to the space heater from the heating system can be calculated in equation (A.5).

$$\dot{Q}_h = \dot{m}_w c_{p,w} (T_{w,in} - T_{w,out}) \quad (A.5)$$

A.3. Heating coil

The heating coil model is based on a simple heat balance equation by considering the sensible heat transfer from the coil to the air and is given by equation (A.6).

$$\dot{Q}_{hc} = \begin{cases} \dot{m}_a c_{p,a} (T_{a,set} - T_{a,in}), & \text{if } T_{a,in} < T_{a,set} \\ 0, & \text{otherwise} \end{cases} \quad (A.6)$$

In equation (A.6), $\dot{m}_a$ is the massflow of air, $c_{p,a}$ is the specific heat capacity of air, $T_{a,set}$ is the temperature setpoint of the supply airflow, and $T_{a,in}$ is the inlet air temperature of the airflow with respect to the heating coil. The model assumes that the heating coil is ideal and can always meet the specified supply air setpoint.

A.4. Cooling coil

Similarly to the heating coil, the cooling coil model also is based on a simple heat balance equation by considering the sensible heat transfer from the air to the coil and is given by equation (A.7).

$$\dot{Q}_{cc} = \begin{cases} \dot{m}_a c_{p,a} (T_{a,in} - T_{a,set}), & \text{if } T_{a,in} > T_{a,set} \\ 0, & \text{otherwise} \end{cases} \quad (A.7)$$

In equation (A.7), $\dot{m}_a$ is the massflow of air, $c_{p,a}$ is the specific heat capacity of air, $T_{a,set}$ is the temperature setpoint of the supply airflow, and $T_{a,in}$ is the inlet air temperature of the airflow with respect to the cooling coil.

A.5. Air to air heat recovery

The heat recovery model is based on the EnergyPlus Air System Air-To-Air Sensible and Latent Effectiveness Heat Exchanger model [53] and is given by equations (A.8)-(A.15).

$$f_{flow} = \frac{\frac{1}{2} (\dot{m}_{a,sup} + \dot{m}_{a,exh})}{\dot{m}_{a,max}} \quad (A.8)$$

$$\epsilon_{75\%} = \begin{cases} \epsilon_{75\%,h}, & \text{if } T_{a,sup,in} < T_{a,exh,in} \\ \epsilon_{75\%,c}, & \text{otherwise} \end{cases} \quad (A.9)$$

$$\epsilon_{100\%} = \begin{cases} \epsilon_{100\%,h}, & \text{if } T_{a,sup,in} < T_{a,exh,in} \\ \epsilon_{100\%,c}, & \text{otherwise} \end{cases} \quad (A.10)$$

$$\epsilon_s = \epsilon_{s,75\%} + (\epsilon_{s,100\%} - \epsilon_{s,75\%}) \frac{f_{flow} - 0.75}{1 - 0.75} \quad (A.11)$$

$$\dot{C}_{a,sup} = \dot{m}_{a,sup} c_{p,a} \quad (A.12)$$

$$\dot{C}_{a,exh} = \dot{m}_{a,exh} c_{p,a} \quad (A.13)$$

$$\dot{C}_{a,min} = \min(\dot{C}_{a,sup}, \dot{C}_{a,exh}) \quad (A.14)$$

$$T_{a,sup,out} = T_{OA} + \epsilon_s \left(\frac{\dot{C}_{a,min}}{\dot{C}_{a,sup}} \right) (T_{a,exh,in} - T_{OA}) \quad (A.15)$$

In equation (A.8) the average flow fraction f_{flow} is calculated based on the supply air massflow $\dot{m}_{sup}$, the exhaust air massflow $\dot{m}_{exh}$, and the maximum air massflow $\dot{m}_{max}$. The heat recovery model has two modes of operation, heating, and cooling, reflected by the conditional effectiveness parameters $\epsilon_{75\%,h}$, $\epsilon_{100\%,h}$, $\epsilon_{75\%,c}$, $\epsilon_{100\%,c}$ in equation (A.9) and equation (A.10). Here, $\epsilon_{75\%}$ and $\epsilon_{100\%}$ is the efficiency at 75% and 100% airflow, while c and h denotes cooling and heating mode, respectively. Based on these effectiveness values, the operating efficiency ϵ is calculated in equation (A.11) by linear interpolation.

$\dot{C}_{sup}$ is the heat capacity rate of the supply air and $\dot{C}_{exh}$ is the heat capacity rate of the exhaust air given by equation (A.12) and equation (A.13), respectively. Following, the minimum heat capacity rate $\dot{C}_{min}$ is calculated in equation (A.14), and finally, the outlet air temperature on the supply side $T_{a,sup,out}$ can be calculated in equation (A.15), where T_{OA} is the outdoor temperature. In cases where the heat exchanger can provide setpoint control, $T_{a,sup,out}$ is corrected with equation (A.16) to avoid excessive heating or cooling of the air beyond the setpoint temperature.

$$T_{a,sup,out} = \begin{cases} \min(T_{sup,out}, T_{a,set}), & \text{if } T_{a,sup,in} < T_{a,exh,in} \\ \max(T_{sup,out}, T_{a,set}), & \text{otherwise} \end{cases} \quad (A.16)$$

In a previous investigation, the authors have proposed and demonstrated the use of a data-driven approach to identify the parameters $\epsilon_{75\%,h}$ and $\epsilon_{100\%,h}$ through collected temperature data by formulating a least squares optimization problem [68]. The resulting model provided good prediction accuracy on the outlet air temperature compared with the measured data.

A.6. Fan

The fan model is based on the EnergyPlus Variable Speed Fan Model [53] and is given by equations (A.17)-(A.19).

$$f_{flow} = \frac{\dot{m}_a}{\dot{m}_{a,max}} \quad (A.17)$$

$$f_{pl} = c_1 + c_2 f_{flow} + c_3 f_{flow}^2 + c_4 f_{flow}^3 \quad (A.18)$$

$$\dot{W}_{fan} = f_{pl} \dot{W}_{fan,max} \quad (A.19)$$

In equation (A.17), the flow fraction f_{flow} is first calculated based on the current massflow $\dot{m}_a$ and the maximum massflow $\dot{m}_{a,max}$. Following, in equation (A.18), the part load ratio f_{pl} is calculated, based on the flow fraction and the power coefficients c_1 - c_4 , where $c_1 + c_2 + c_3 + c_4 = 1$. In equation (A.19), the part load ratio is then used to calculate the fan power consumption $\dot{W}_{fan}$, based on the power consumption at maximum airflow $\dot{W}_{fan,max}$.

A.7. Damper

The damper model describes the airflow through the damper $\dot{m}_a$ as a function of the damper position u and the parameters a , b , and c . It is based on the model presented by Huang [69] and is given by equation (A.20).

$$\dot{m}_a = a e^{b u_d} + c \quad (A.20)$$

Constraints can be imposed on the parameters a , b , c in equation (A.20) such that $m_a = 0$ when $u_d = 0$ and $m_a = m_{a,max}$ when $u_d = 1$. Hence, if a is given, b and c are given by equations (A.21)-(A.22) to satisfy these two constraints.

$$c = -a \quad (\text{A.21})$$

$$b = \ln \left(\frac{\dot{m}_{a,max} - c}{a} \right) \quad (\text{A.22})$$

A.8. Controller

Properly designed and implemented Building Management Systems (BMS) are vital for the building to adapt and react effectively to internal and external dynamic changes such as occupancy and climatic conditions [70]. The BMS is typically connected with a set of controllers that control the operation of various building systems. Controllers can come in different forms, e.g. rule-based control, conventional feedback control, predictive control, etc. However, the general aim of controllers is to modulate system inputs such that a predefined goal is achieved. For conventional controllers, the goal is typically to drive a specific system property toward the desired setpoint value. One of the most fundamental and widely used controllers is the Proportional Integral Derivative (PID) controller [71]. In discrete time, the controller behavior is described by equations (A.23)–(A.24) [72].

$$e_t = y_{set} - y_{meas,t} \quad (\text{A.23})$$

$$u_t = u_{t-1} + K_p(e_t - e_{t-1}) + K_i\Delta t + \frac{K_d}{\Delta t}(e_t - 2e_{t-1} + e_{t-2}) \quad (\text{A.24})$$

Where e_t is the residual between the setpoint y_{set} and the measured value $y_{meas,t}$ at time t . In equation (A.24), the output signal of the controller u_t is calculated based on the previous signal u_{t-1} , the previous residuals e_{t-1} , e_{t-2} , the parameters K_p , K_i , K_d , and the timestep Δt . Changing the type of controller is simply a matter of adjusting the parameters K_p , K_i , K_d , i.e., a P controller can be modeled with $K_p \neq 0$ and K_i , $K_d = 0$ while a PI controller can be modeled with $K_p \neq 0$ and $K_d = 0$. Additionally, a reverse-acting PID controller can be modeled with K_p , K_i , $K_d < 0$.

A.9. Building space

Indoor comfort is the driving force of essentially all energy use in buildings. Most BMS have automated control of indoor temperature to follow specified setpoint schedules. Additionally, Demand Controlled Ventilation (DCV) is a common strategy to ensure that mechanical ventilation is only operating during occupancy presence in the building with the aim of lowering power consumption and ensuring proper Indoor Air Quality (IAQ). For this purpose, CO₂ concentration is predominately measured as an indicator of IAQ and is controlled to follow certain setpoint schedules, by modulating the ventilation airflows [73]. Therefore, the task of accurately modeling and predicting temperature and CO₂ concentration is essential to properly represent the actual building space in the building DT. Therefore, the space component must be capable of predicting both the dynamic temperature and CO₂ concentration response, depending on the operation of the HVAC system, weather conditions, and occupancy use. The dynamic CO₂ concentration in rooms is typically modeled with the fundamental mass balance given by equation (A.25) [38,70].

$$m_z \frac{dC_z}{dt} = C_{sup}\dot{m}_{a,sup} - C_z\dot{m}_{a,exh} + K_{occ}N_{occ} \quad (\text{A.25})$$

Where, m_z is the mass of the air contained in the room, C_z is the room CO₂ concentration, and C_{sup} is the CO₂-concentration of the supply airflow, which can be assumed equal to the outdoor air CO₂-concentration level at 400 ppm [70,74]. K_{occ} is the rate of CO₂ mass generated per occupant and N_{occ} is the number of occupants in the room. Pantazaras et al. [70], estimated K_{occ} , while synthetic Energy-Plus data was used as input for air flows and occupancy. Macarulla et al. [38], extended equation (A.25) to form a stochastic differential equation. Different configurations were tested for the estimated parameters. However, for the best-performing model, the actual occupancy

count was used as input, while K_{occ} and the air flows were estimated as constants.

However, similarly to the transformation from equation (A.3) to equation (A.4), an explicit expression can be obtained for the room CO₂-concentration with Backward Euler discretization of equation (A.25). The obtained expression is given by equation (A.26), where the timestep Δt has been introduced.

$$C_{z,t} = \frac{m_z C_{z,t-1} + C_{sup}\dot{m}_{a,sup}\Delta t + K_{occ}N_{occ}\Delta t}{m_z + \dot{m}_{a,exh}\Delta t} \quad (\text{A.26})$$

Indoor temperature forecasting has been studied extensively using different modeling methods covering both white-box, gray-box, and black-box approaches. However, recently, black-box approaches in the form of ANNs have become increasingly popular. Specifically, to deal with the transient and dynamic nature of indoor temperature, a specific branch of ANNs called Recurrent Neural Networks (RNN) is typically used. Under this category, the Long Short-Term Memory (LSTM) architecture is often used, due to its gated structure, which rectifies certain undesirable traits of its predecessor, the vanilla RNN [75]. The adaptability and prediction accuracy of LSTM networks have been demonstrated across many disciplines [76], including for the task of indoor temperature forecasting [77,78].

We have previously developed and demonstrated the use of the LSTM architecture [62,63] for indoor temperature modeling of a large set of spaces in a case study building. Therefore, this black-box approach is also applied in this work, which also demonstrates the flexibility of the proposed framework and the ability to combine different modeling approaches in a unified manner. The overall model architecture is summarized by equations (A.27)–(A.31), employing two sequential LSTM networks A and B .

$$\mathcal{X}_{A,t-1} = (T_{z,t-1}, T_{o,t-1}, \Phi_{s,t-1}, u_{v,t-1}, u_{d,t-1}, u_{sh,t-1}) \quad (\text{A.27})$$

$$(c_{A,t}, h_{A,t}) = \text{LSTM}_A(\mathcal{X}_{A,t-1}, c_{A,t-1}, h_{A,t-1}) \quad (\text{A.28})$$

$$\mathcal{X}_{B,t-1} = h_{A,t} \quad (\text{A.29})$$

$$(c_{B,t}, h_{B,t}) = \text{LSTM}_B(\mathcal{X}_{B,t-1}, c_{B,t-1}, h_{B,t-1}) \quad (\text{A.30})$$

$$\Delta T_{z,t} = h_{B,t} \quad (\text{A.31})$$

equation (A.27) defines the input $\mathcal{X}_A$ for LSTM_A. In this case, the inputs include indoor temperature T_z , outdoor temperature T_o , short-wave irradiation Φ_s , space heater valve position u_v , damper opening position u_d , and shades opening position u_{sh} . These inputs were chosen considering the thermal heat balance of the modeled spaces, and commonly available data in buildings. However, the inputs should be modified based on available sensors and equipment installed for the modeled space.

In equation (A.28), the input $\mathcal{X}_{A,t-1}$ as well as the state vectors $c_{A,t-1}$ and $h_{A,t-1}$ are provided as input for LSTM_A, which computes the updated state vectors $c_{A,t}$ and $h_{A,t}$. equation (A.29) then defines the input $\mathcal{X}_{B,t-1}$ for LSTM_B as the updated hidden state $h_{A,t}$. Using this input $\mathcal{X}_{B,t-1}$ as well as the state vectors $c_{B,t-1}$ and $h_{B,t-1}$, equation (A.30) then computes the updated state vectors $c_{B,t}$ and $h_{B,t}$, where the hidden state $h_{B,t}$ is treated as prediction target $\Delta T_{z,t}$, as shown in equation (A.31). Here, $\Delta T_{z,t}$ is the indoor air temperature difference between the previous and current timestep. After training, the model is used in inference for indoor temperature forecasting through a closed-loop configuration, given by equation (A.32). Here, the indoor temperature of the next timestep is obtained by simply adding the predicted temperature change $\Delta T_{z,t}$ with the known indoor temperature of the previous timestep $T_{z,t-1}$. The newly obtained temperature $T_{z,t}$ can then be fed back as input in equation (A.27), closing the loop.

$$T_{z,t} = T_{z,t-1} + \Delta T_{z,t} \quad (\text{A.32})$$

Appendix B

Table B.2

Overview of parameter and constant values used for the demonstration case.

Component	Parameters	Constants
Valve	$N_v = 0.8$ $\dot{m}_{v,max} = 0.21 \text{ kg/s}$	
Space heater	$UA = 162 \text{ W/K}$ $C_r = 125000 \text{ J/K}$	$c_{p,w} = 4180 \text{ J/kg/K}$ $\Delta t = 600 \text{ s}$
Heating coil		$c_{p,a} = 1000 \text{ J/kg/K}$
Air to air heat recovery ^[1]	$\epsilon_{75\%,h} = 0.79$ $\epsilon_{75\%,c} = 0.79$ $\epsilon_{100\%,h} = 0.73$ $\epsilon_{100\%,c} = 0.73$ $\dot{m}_{a,max} = 1.7 \text{ kg/s}$	$c_{p,a} = 1000 \text{ J/kg/K}$
Supply Fan ^[2]	$c_1 = 0.027828$ $c_2 = 0.026583$ $c_3 = -0.087069$ $c_4 = 1.030920$ $\dot{m}_{a,max} = 1.7 \text{ kg/s}$ $\dot{W}_{fan,max} = 1500 \text{ W}$	$\rho_a = 1.225 \text{ kg/m}^3$
Supply Damper ^[3]	$a = 5$ $\dot{m}_{a,max} = 1.63 \text{ kg/s}$	
Exhaust Damper ^[3]	$a = 5$ $\dot{m}_{a,max} = 1.63 \text{ kg/s}$	
Temperature controller	$K_p = 0.05$ $K_i = 0.8$ $K_d = 0$	
CO2 controller	$K_p = -0.001$ $K_i = -0.001$ $K_d = 0$	
Space ^[4]	$K_{occ} = 8.316 \cdot 10^{-6} \text{ kg/s/person}$ $m_z = 571.5 \text{ kg}$ [-]	$\Delta t = 600 \text{ s}$

^[1] Efficiency values $\epsilon_{75\%,h}$, $\epsilon_{75\%,c}$, $\epsilon_{100\%,h}$, $\epsilon_{100\%,c}$ was selected based on the AHRI directory of certified product performance for a specific product [79].

^[2] Power coefficients c_1 - c_4 was selected based on the single-zone default values provided by the ANSI/ASHRAE/IES standard 90.1 [80].

^[3] equation (A.21) and equation (A.22) are used to calculate c and b .

^[4] CO₂-generation per occupant K_{occ} obtained for a classroom from [81]. The LSTM model contains too many parameters to show.

Table B.3

A subset of class descriptions and related properties from the SAREF4BLDG ontology extension for components commonly considered in building energy modeling [44].

Component	Description	Properties
Valve	<i>A valve is used in a building services piping distribution system to control or modulate the flow of the fluid.</i>	closeOffRating ^{op} flowCoefficient ^{op} size ^{op} testPressure ^{op} valveMechanism ^{dp} valveOperation ^{dp} valvePattern ^{dp} workingPressure ^{op}
Space Heater	<i>Space heaters utilize a combination of radiation and/or natural convection using a heating source such as electricity, steam or hot water to heat a limited space or area. Examples of space heaters include radiators, convectors, baseboard and finned-tube heaters.</i>	bodyMass ^{op} energySource ^{dp} heatTransferDimension ^{dp} heatTransferMedium ^{dp} numberOfPanels ^{dp} numberOfSections ^{dp} outputCapacity ^{op} placementType ^{dp} temperatureClassification ^{dp} thermalEfficiency ^{op} thermalMassHeatCapacity ^{op}

(continued on next page)

Table B.3 (continued)

Component	Description	Properties
Coil	<i>A coil is a device used to provide heat transfer between non-mixing media. A common example is a cooling coil, which utilizes a finned coil in which circulates chilled water, antifreeze, or refrigerant that is used to remove heat from air moving across the surface of the coil. A coil may be used either for heating or cooling purposes by placing a series of tubes (the coil) carrying a heating or cooling fluid into an airstream. The coil may be constructed from tubes bundled in a serpentine form or from finned tubes that give a extended heat transfer surface.</i>	airFlowRateMax ^{op} airFlowRateMin ^{op} nominalLatentCapacity ^{op} nominalSensibleCapacity ^{op} nominalUa ^{op} operationTemperatureMax ^{op} operationTemperatureMin ^{op} placementType ^{dp}
Air To Air Heat Recovery	<i>An air-to-air heat recovery device employs a counter-flow heat exchanger between inbound and outbound air flow. It is typically used to transfer heat from warmer air in one chamber to cooler air in the second chamber (i.e., typically used to recover heat from the conditioned air being exhausted and the outside air being supplied to a building), resulting in energy savings from reduced heating (or cooling) requirements.</i>	hasDefrost ^{dp} heatTransferTypeEnum ^{dp} operationTemperatureMax ^{op} operationTemperatureMin ^{op} primaryAirFlowRateMax ^{op} primaryAirFlowRateMin ^{op} secondaryAirFlowRateMax ^{op} secondaryAirFlowRateMin ^{op}
Fan	<i>A fan is a device which imparts mechanical work on a gas. A typical usage of a fan is to induce airflow in a building services air distribution system.</i>	capacityControlType ^{dp} motorDriveType ^{dp} nominalAirFlowRate ^{op} nominalPowerRate ^{op} nominalRotationSpeed ^{op} nominalStaticPressure ^{op} nominalTotalPressure ^{op} operationTemperatureMax ^{op} operationTemperatureMin ^{op} operationalRiterial ^{op}
Damper	<i>A damper typically participates in an HVAC duct distribution system and is used to control or modulate the flow of air.</i>	airFlowRateMax ^{op} bladeAction ^{dp} bladeEdge ^{dp} bladeShape ^{dp} bladeThickness ^{op} closeOffRating ^{op} faceArea ^{op} frameDepth ^{op} frameThickness ^{op} frameType ^{dp} leakageFullyClosed ^{op} nominalAirFlowRate ^{op} numberOfBlades ^{dp} openPressureDrop ^{op} operation ^{dp} operationTemperatureMax ^{op} operationTemperatureMin ^{op} orientation ^{dp} temperatureRating ^{op} workingPressureMax ^{op}
Controller	<i>A controller is a device that monitors inputs and controls outputs within a building automation system. A controller may be physical (having placement within a spatial structure) or logical (a software interface or aggregated within a programmable physical controller).</i>	
Building Space	<i>An entity used to define the physical spaces of the building. A building space contains devices or building objects.</i>	contains ^{op} hasSpace ^{op} isSpaceOf ^{op}

References

- [1] K. Ullah, I. Lill, E. Witt, An Overview of Bim Adoption in the Construction Industry: Benefits and Barriers, 2019.
- [2] C. Boje, A. Guerriero, S. Kubicki, Y. Rezgui, Towards a semantic construction digital twin: directions for future research, Autom. Constr. 114 (2020) 103179, <https://doi.org/10.1016/j.autcon.2020.103179>, <https://www.sciencedirect.com/science/article/pii/S0926580519314785>.
- [3] E. Glaessgen, D. Stargel, The digital twin paradigm for future nasa and u.s. air force vehicles, 2012.
- [4] J. Guo, Z. Lv, Application of digital twins in multiple fields, Multimed. Tools Appl. 81 (08.2022), <https://doi.org/10.1007/s11042-022-12536-5>.
- [5] M. Grieves, Digital twin: Manufacturing excellence through virtual factory replication, 03.2015.
- [6] F. Xiao, S. Wang, Progress and methodologies of lifecycle commissioning of hvac systems to enhance building sustainability, Renew. Sustain. Energy Rev. 13 (2009) 1144–1149.
- [7] L. Wang, S. Greenberg, J. Fiegel, A. Rubalcava, S. Earni, X.R. Yin, S. Woodworth, J. Hernandez-Maldonado, Monitoring-based hvac commissioning of an existing office building for energy efficiency, Appl. Energy 102 (2013) 1382–1390.
- [8] Z. O'Neill, X. Pang, M. Shashanka, P. Haves, T. Bailey, Model-based real-time whole building energy performance monitoring and diagnostics, J. Build. Perform. Simul. 7 (2014) 83–99.
- [9] E. Markoska, M. Jradi, B. Jørgensen, Continuous commissioning of buildings: a case study of a campus building in Denmark, in: 2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2016.
- [10] M. Jradi, K. Arendt, F. Sangogboye, C. Mattera, E. Markoska, M. Kjærgaard, C. Veje, B. Jørgensen, Obepme: an online building energy performance monitoring and evaluation tool to reduce energy performance gaps, Energy Build. 166 (2018).
- [11] C.K. Metallidou, K.E. Psannis, E.A. Egyptiadou, Energy efficiency in smart buildings: iot approaches, IEEE Access 8 (2020) 63679–63699, <https://doi.org/10.1109/ACCESS.2020.2984461>.

- [12] The cost-effectiveness of continuous commissioning® over the past ten years, 01.2008.
- [13] H.B. Gunay, M. Ouf, G. Newsham, W. O'Brien, Sensitivity analysis and optimization of building operations, *Energy Build.* 199 (2019) 164–175, <https://doi.org/10.1016/j.enbuild.2019.06.048>, <https://www.sciencedirect.com/science/article/pii/S0378778818334479>.
- [14] M. Laura Daniele, R. Garcia-Castro, M. Lefrançois, M. Poveda-Villalon, Saref: the smart applications reference ontology, Available at <https://saref.etsi.org/core/>. Online.
- [15] D. Crawley, L. Lawrie, F. Winkelman, W. Buhl, Y. Huang, C. Pedersen, R. Strand, R. Liesen, D. Fisher, M. Witte, J. Glazer, Energyplus: creating a new-generation building energy simulation program, *Energy Build.* 33 (2001) 319–331.
- [16] F. Winkelman, B. Birdsall, W. Buhl, K. Ellington, A. Erdem, J. Hirsch, S. Gates, Doe-2 supplement: Version 2.1e.
- [17] S. Klein, W. Beckman, Trnsys 16: a transient system simulation program: mathematical reference, in: TRNSYS 5, 2007, pp. 389–396.
- [18] A. Reddy, Literature review on calibration of building energy simulation programs: uses, problems, procedure, uncertainty, and tools, *ASHRAE Trans.* 112 (2006) 226–240.
- [19] J. Reynolds, M. Ahmad, Y. Rezgui, Holistic modelling techniques for the operational optimisation of multi-vector energy systems, *Energy Build.*
- [20] Y. Chen, M. Guo, Z. Chen, Z. Chen, Y. Ji, Physical energy and data-driven models in building energy prediction: a review, *Energy Rep.* 8 (2022) 2656–2671.
- [21] L.J. Ramaji, J. Messner, E. Mostavi, Ifc-based bim-to-bem model transformation, *J. Comput. Civ. Eng.* 34 (04.2020), [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000880](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000880).
- [22] A. Andriamamonjy, D. Saelens, R. Klein, An automated ifc-based workflow for building energy performance simulation with modelica, *Autom. Constr.* 91 (2018) 166–181, <https://doi.org/10.1016/j.autcon.2018.03.019>, <https://www.sciencedirect.com/science/article/pii/S0926580517308282>.
- [23] A. Andriamamonjy, Automated workflows for building design and operation using openbim and modelica, Ph.D. thesis, 10.13140/RG.2.2.31108.78729, 11.2018.
- [24] H.-Y. Chong, C.-Y. Lee, X. Wang, A mixed review of the adoption of building information modelling (bim) for sustainability, *J. Clean. Prod.* 142 (2017) 4114–4126, <https://doi.org/10.1016/j.jclepro.2016.09.222>, <https://www.sciencedirect.com/science/article/pii/S0959652616315748>.
- [25] F. Mtibaa, K.-K. Nguyen, M. Azam, A. Papachristou, J.-S. Venne, M. Cheriet, Lstm-based indoor air temperature prediction framework for hvac systems in smart buildings, *Neural Comput. Appl.*
- [26] Z. Fang, N. Crimier, L. Scanu, A. Midelet, A. Alyafi, B. Delinchant, Multi-zone indoor temperature prediction with lstm-based sequence to sequence model, *Energy Build.*
- [27] B. Dong, C. Cao, S.E. Lee, Applying support vector machines to predict building energy consumption in tropical region, *Energy Build.* 37 (2005) 545–553.
- [28] Z. Yu, F. Haghighat, B. Fung, H. Yoshino, A decision tree method for building energy demand modeling, *Energy Build.* 42 (2010) 1637–1646.
- [29] S. Royer, S. Thil, T. Talbert, M. Polit, Black-box modeling of buildings thermal behavior using system identification, in: 19th IFAC World Congress, IFAC Proc. Vol. 47 (3) (2014) 10850–10855, <https://doi.org/10.3182/20140824-6-ZA-1003.01519>, <https://www.sciencedirect.com/science/article/pii/S1474667016433392>.
- [30] P.P. Shinde, S. Shah, A review of machine learning and deep learning applications, *IEEE*, 2018, pp. 1–6.
- [31] O.I. Abiodun, A. Jantan, A.E. Omolara, K.V. Dada, N.A. Mohamed, H. Arshad, State-of-the-art in artificial neural network applications: a survey, *Heliyon* 4 (11) (2018) e00938, <https://doi.org/10.1016/j.heliyon.2018.e00938>, <https://www.sciencedirect.com/science/article/pii/S2405844018332067>.
- [32] S. Idowu, S. Saguna, C. Åhlund, O. Schelén, Applied machine learning: forecasting heat load in district heating system, *Energy Build.* 133 (2016) 478–488.
- [33] Y. Guo, J. Wang, H. Chen, G. Li, J. Liu, C. Xu, R. Huang, Y. Huang, Machine learning-based thermal response time ahead energy demand prediction for building heating systems, *Appl. Energy* 221 (2018) 16–27.
- [34] F. Pallonetto, C. Jin, E. Mangina, Forecast electricity demand in commercial building with machine learning models to enable demand response programs, *Energy AI* 7 (2022) 100–121.
- [35] G. Pinto, Z. Wang, A. Roy, T. Hong, A. Capozzoli, Transfer learning for smart buildings: a critical review of algorithms, applications, and future perspectives, *Adv. Appl. Energy* 5 (2022) 100084, <https://doi.org/10.1016/j.adapen.2022.100084>, <https://www.sciencedirect.com/science/article/pii/S2666792422000026>.
- [36] Z. Jiang, Y.M. Lee, Deep transfer learning for thermal dynamics modeling in smart buildings, in: 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 2033–2037.
- [37] X. Fang, G. Gong, G. Li, L. Chun, W. Li, P. Peng, A hybrid deep transfer learning strategy for short term cross-building energy prediction, *Energy* 215 (2021) 119208, <https://doi.org/10.1016/j.energy.2020.119208>, <https://www.sciencedirect.com/science/article/pii/S036054422023315X>.
- [38] M. Macarulla, M. Casals, M. Carnevali, N. Forcada, M. Gangoilells, Modelling indoor air carbon dioxide concentration using grey-box models, *Build. Environ.* 117 (02.2017), <https://doi.org/10.1016/j.buildenv.2017.02.022>.
- [39] X. Yu, L. Georges, M. Knudsen, I. Sartori, L. Inslund, Investigation of the Model Structure for low-Order Grey-Box Modelling of Residential Buildings, International Building Performance Simulation Association.
- [40] H. Park, N. Martaj, M. Ruellan, R. Bennacer, E. Monmasson, Modeling of a building system and its parameter identification, *J. Electr. Eng. Technol.* 8 (2013) 975–983.
- [41] A. Afram, F. Janabi-Sharifi, Gray-box modeling and validation of residential hvac system for control system design, *Appl. Energy* 137 (2015) 134–150, <https://doi.org/10.1016/j.apenergy.2014.10.026>, <https://www.sciencedirect.com/science/article/pii/S0306261914010678>.
- [42] M. Pritoni, D. Paine, G. Fierro, C. Mosiman, M. Poplawski, A. Saha, J. Bender, J. Granderson, Metadata schemas and ontologies for building energy applications: a critical review and use case analysis, *Energies* 14 (2021) 2024, <https://doi.org/10.3390/en14072024>.
- [43] L. Daniele, F. den Hartog, J. Roes, Created in close interaction with the industry: the smart appliances reference (saref) ontology, 2015, pp. 100–112.
- [44] M. Poveda-Villalón, R. Garcia-Castro, Saref extension for building, Available at <https://saref.etsi.org/saref4bldg/>. Online.
- [45] M. Poveda-Villalón, R. Garcia-Castro, P. Espinoza-Arias, Saref extension for smart city, Available at <https://saref.etsi.org/saref4city/>. Online.
- [46] M. Lefrançois, Saref4syst: an extension of saref for typology of systems and their inter-connections, Available at <https://saref.etsi.org/saref4syst/>. Online.
- [47] L. Daniele, Saref4ener: an extension of saref for the energy domain, Available at <https://saref.etsi.org/saref4ener/>. Online.
- [48] International organization for standardization, iso 16739-1:2018. Available at <https://www.iso.org/standard/70303.html>. Online.
- [49] P. Pauwels, W. Terkaj, ifcowl ontology, Available at https://standards.buildingsmart.org/IFC/DEV/IFC4/ADD2_TC1/OWL/. Online.
- [50] M.H. Rasmussen, M. Lefrançois, G. Schneider, P. Pauwels, Bot: the building topology ontology of the w3c linked building data group, *Semant. Web* (11.2020), <https://doi.org/10.3233/SW-200385>.
- [51] P. Li, Z. O'Neill, J. Braun, Development of control-oriented models for model predictive control in buildings, *ASHRAE Trans.* 119.
- [52] F. Jorissen, M. Wetter, L. Helsen, Simulation speed analysis and improvements of modelica models for building energy simulation, in: 11th International Modelica Conference, Versailles.
- [53] U. Energy, Energyplus version 9.6.0 documentation.
- [54] A. Ruano, E. Crispim, E. Conceição, M. Lúcio, Prediction of building's temperature using neural networks models, *Energy Build.* 38 (6) (2006) 682–694, <https://doi.org/10.1016/j.enbuild.2005.09.007>, <https://www.sciencedirect.com/science/article/pii/S0378778805002136>.
- [55] T. Blochwitz, M. Otter, M. Arnold, C. Clauß, H. Elmqvist, A. Junghanns, J. Mauss, M. Monteiro, T. Neidhold, D. Neumerkel, H. Olsson, J.-V. Peetz, S. Wolf, The functional mockup interface for tool independent exchange of simulation models, 2011, pp. 105–114.
- [56] E.T.S. Institute, Smartm2m, extension to saref, part 3: building domain, etsi-ts-103 410-3, Available at <https://www.iso.org/standard/70303.html>. Online.
- [57] C. Pang, J. Wang, Y. Cheng, H.L. Zhang, T. Li, Topological sorts on dags, *Inf. Process. Lett.* 115 (2015) 298–301, <https://doi.org/10.1016/j.ipl.2014.09.031>.
- [58] R. Liu, A low complexity topological sorting algorithm for directed acyclic graph, *Int. J. Mach. Learn. Comput.* 4 (2014) 194–197, <https://doi.org/10.7763/IJMLC.2014.V4.411>.
- [59] A.B. Kahn, Topological sorting of large networks, *Commun. ACM* 5 (1962) 558–562.
- [60] J. Bjørnskov, Ontology-based building modeling framework, <https://github.com/JBjornskov/Twin4Build>, 2023. (Accessed 23 March 2023).
- [61] R. van der Weerd, V. Boer, L. Daniele, B. Nouwt, Validating SAREF in a smart home environment, 2021, pp. 35–46.
- [62] J. Bjørnskov, M. Jradi, C. Veje, A fully automated and scalable approach for indoor temperature forecasting in buildings using artificial neural networks, in: Building Simulation Applications BSA, 2022, pp. 349–356.
- [63] J. Bjørnskov, M. Jradi, A modular thermal space coupling approach for indoor temperature forecasting using artificial neural networks, in: Building Simulation and Optimization BSO, 2022, https://publications.ibpsa.org/conference/paper/?id=bso2022_59.
- [64] I. Zajic, T. Larkowski, M. Sumislawska, K. Burnham, D. Hill, Modelling of an air handling unit: a Hammerstein-bilinear model identification approach, in: International Conference on Systems Engineering, Las Vegas.
- [65] B. Xu, L. Fu, H. Di, Dynamic simulation of space heating systems with radiators controlled by trvs in buildings, *Energy Build.* 40 (9) (2008) 1755–1764, <https://doi.org/10.1016/j.enbuild.2008.03.004>.
- [66] F. Tahersima, J. Stoustrup, H. Rasmussen, P. Nielsen, Thermal analysis of an hvac system with trv controlled hydronic radiator, in: 2010 IEEE International Conference on Automation Science and Engineering, 2010, pp. 756–761.
- [67] The backward Euler method, the trapezoidal method, John Wiley & Sons, Ltd, 2009, pp. 49–65. Ch. 4.
- [68] J. Bjørnskov, M. Jradi, C. Veje, Component-level re-commissioning of a newly retrofitted Danish healthcare building, *J. Build. Eng.* 51 (2022) 104277, <https://doi.org/10.1016/j.job.2022.104277>, <https://www.sciencedirect.com/science/article/pii/S235271022200290X>.
- [69] G. Huang, Model predictive control of vav zone thermal systems concerning bilinearity and gain nonlinearity, *Control Eng. Pract.* 19 (7) (2011) 700–710.
- [70] A. Pantazaras, S.E. Lee, M. Santamouris, J. Yang, Predicting the co2 levels in buildings using deterministic and identified models, *Energy Build.* 127 (2016) 774–785, <https://doi.org/10.1016/j.enbuild.2016.06.029>, <https://www.sciencedirect.com/science/article/pii/S0378778816305187>.

- [71] F. Belic, Z. Hocenski, D. Sliskovic, Hvac control methods - a review, <https://doi.org/10.1109/ICSTCC.2015.7321372>, 2015.
- [72] S. Veeranna, U. Yaragatti, Analysis and implementation of discrete time pid controllers using fpga, *Int. J. Electr. Comput. Eng.* 2 (2010) 71–82.
- [73] B. Merema, M. Delwati, M. Sourbron, H. Breesch, Demand controlled ventilation (dcv) in school and office buildings: lessons learnt from case studies, *Energy Build.* 172 (2018) 349–360, <https://doi.org/10.1016/j.enbuild.2018.04.065>, <https://www.sciencedirect.com/science/article/pii/S0378778818301348>.
- [74] Chartered Institution of Building Services Engineers, Cibse guide a: Environmental design, chapter 4.2, Table 4.2, 2007.
- [75] A. Sherstinsky, Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network, *Physica D* 404 (2020) 132306, <https://doi.org/10.1016/j.physd.2019.132306>, <https://www.sciencedirect.com/science/article/pii/S0167278919305974>.
- [76] G. Van Houdt, C. Mosquera, G. Nápoles, A review on the long short-term memory model, *Artif. Intell. Rev.* 53 (12.2020), <https://doi.org/10.1007/s10462-020-09838-1>.
- [77] F. Mtibaa, K.-K. Nguyen, M. Azam, A. Papachristou, J.-S. Venne, M. Cheriet, Lstm-based indoor air temperature prediction framework for hvac systems in smart buildings, *Neural Comput. Appl.* 32 (12.2020), <https://doi.org/10.1007/s00521-020-04926-3>.
- [78] Z. Fang, N. Crimier, L. Scanu, A. Midelet, A. Alyafi, B. Delinchant, Multi-zone indoor temperature prediction with lstm-based sequence to sequence model, *Energy Build.* 245 (2021) 111053, <https://doi.org/10.1016/j.enbuild.2021.111053>, <https://www.sciencedirect.com/science/article/pii/S0378778821003376>.
- [79] The Air-Conditioning, Heating, and refrigeration institute, directory of certified product performance, Available at <https://www.ahridirectory.org/>. Online.
- [80] S. Goel, M.I. Rosenberg, *Ansi/ashrae/ies standard 90.1-2010 performance rating method reference manual*, 2016.
- [81] A. Persily, L. De Jonge, Carbon dioxide generation rates from building occupants, *Indoor Air* 27 (03.2017), <https://doi.org/10.1111/ina.12383>.

Systematic Commissioning and Performance Testing of Next Generation Building Management Systems

Muhyiddine Jradi, PhD, MSc

Na Liu, MSc

ABSTRACT

In recent years, the energy sector as a whole has benefited from new services and effective management due to an extensive digital transformation. The major component driving this digital revolution in buildings is the building management system (BMS). One of the key issues confronting the smart building industry is the assumption at the design stage that the building management system will work as planned and described in the design documentation. Inadequate BMS system design and testing, as well as failures on multiple levels of control of various building components, are important causes of these performance disparities. This work presents the design, development, and implementation of a first-of-its-kind testing and commissioning application for building management systems. The application adopts step response testing as an approach to monitor, test, and evaluate the BMS functionalities and control capabilities. The application demonstration to test and assess the ventilation system control and operation in an actual office building is presented and discussed. The findings highlight the effectiveness of the proposed tool in ensuring a proper design of the building automation system and smooth integration and interaction with the rest of the building systems as a milestone to improve energy performance and enhance building smartness.

INTRODUCTION

With the strict energy and environmental standards and ambitious goals set by the majority of countries worldwide, enhancing buildings energy performance is a key component of every single plan due to their contribution to the overall energy and emissions schemes. In this regard, theoretical assessments, evaluations, experimental implementations, and actual applications of building performance audits and commissioning have all been prominent in recent decades [1, 2]. The majority of these investigations focused on holistic building performance auditing and commissioning, both for initial and continuous commissioning, in order to close the gap between the intended and actual building design and operation, as well as to guarantee proper building operation from the start and throughout the life cycle of the building. Additionally, numerous research investigations have concentrated on the examination and assessment of building structures and materials [3]. Likewise, there has been considerable attention given to scrutinizing specific elements of buildings, such as heating, ventilation, and air conditioning (HVAC) systems [4], as well as devices and various building services [5], with the aim of enhancing the design and optimizing the operation. Although auditing and evaluation of building performance is not a new issue, relatively few studies have focused on auditing and commissioning of the

Muhyiddine Jradi is an Associate Professor of Energy Engineering at the Center for Energy Informatics, The Maersk Mc Kinney Moller Institute, University of Southern Denmark, Odense, Denmark. **Na Liu** is a Research Engineer at the University of Southern Denmark, Odense, Denmark.

building management and control system, analyzing its design and usage patterns. One of the biggest problems confronting the smart building industry is the presumption made at the design stage that the building automation system would behave as expected and stated in the design documents when more and more building automation and control systems are added with different characteristics and functions [6]. Yet, a significant body of research has lately shown that these presumptions and assertions, which are not supported by testing and appropriate commissioning, have resulted in significant energy performance gaps and buildings that are functioning inefficiently [7, 8]. The BACS system's improper design and testing, as well as failures at various levels of control of various building components, including HVAC units and services, are some of the key reasons for these performance gaps [9, 10].

Prior to building handover, manual performance testing and first commissioning are mandated by construction regulations in the vast majority of countries [11]. Yet, the bulk of these tests are conducted at the level of the entire building or facility, paying little attention to the BMS, the building's central system. Moreover, as most of these evaluation and performance testing procedures are carried out manually [12], there is no assurance for the owner that they are carried out correctly or meet the necessary standards of quality. In addition, this manual testing and commissioning procedure is time- and resource-intensive, costly to carry out, and error-prone. Most of these problems and faults spread throughout the operating phase, resulting in a very inefficient operation and high maintenance and repair expenses. To verify the effectiveness of the management system and the reaction of the interconnected HVAC systems, a methodical automated initial and retro-commissioning action is essential. This will guarantee an advanced and properly designed and installed BMS, as well as seamless integration and communication with all of the building's energy generation and supply systems. This paper presents the design, development, and demonstration of a first-of-its-kind testing and commissioning application for building management systems. The application adopts step response testing as an approach to monitor, test, and evaluate the BMS functionalities and control capabilities. Along with the design and development of the application with all the various hardware and software components, the methodology and development of the step response methods are described. The application demonstration to test and assess the ventilation system control and operation in an actual office building is presented and discussed.

BUILDING MANAGEMENT SYSTEMS AUDITING METHODOLOGY AND DESIGN

The major goal is to ensure proper and effective management and control of various dynamic building systems, such as cooling, heating, ventilation, and lighting systems. In this regard, the proposed application is not intended to substitute the building control and automation system; it operates as a complementary component to the BMS, integrating major parameters and inputs regarding the different systems and services in the building. Thus, the application needs a BMS interface to perform standard step response tests, connect to the controllers, and establish actuation processes, in addition to gathering the required information and data to characterize the system's operation and behavior. The testing and commissioning application architecture is given in Figure 1 [14]. Overall, the application design comprises three major cores: pre-processing configuration, parallel testing and execution, and post-processing. The three major modules of the proposed application are described below.

Pre-processing Configuration

Each suggested test's functionality and assessment will be directly impacted by a number of distinct elements. In this regard, the best configuration involves the automation of as many configuration elements as possible. However, a number of the set variables can have a significant impact on the final step response outcome, potentially compromising the final response's consistency and stability in certain cases. As a result, it is the responsibility of the assessment and test engineer to specify the appropriate parameter values for the system being evaluated as well as verify the accuracy of the findings. Overall, the main parameters to be configured and introduced at the stage of pre-processing are the duration of the test, the setpoint of the variable to be controlled, and the dead-band, or acceptance zone.

Parallel Testing and Execution

The parallel execution and processing capabilities offered by the test performance engine are one of the proposed application's key characteristics. The tool enables multiple tests to run concurrently without compromising their effectiveness. Every test is going to possess its own unique set of parameters. Designing the essential parameters for various systems with various dynamics is a major difficulty. The longest test in terms of duration among the set of tests carried out will generally determine how long it will take to complete the testing and commissioning process and deliver the final evaluation results. Considering that the proposed application employs a web interface, there are two ways to carry out step-response testing and evaluation. On the one hand, the user can specify each system's input parameters as well as the system of interest for the test. In this instance, the end result will combine all reports and graphs onto a single screen and a common document. Alternatively, the user has the choice of opening a separate tab in the web browser for each test intended to run. In this case, each test will be run individually by the execution engine, and the outcomes won't be compared or evaluated on the same page. Choosing the method of execution depends largely on the commissioning process conducted, the scale of the testing planned, and the type of systems targeted.

Post-processing and Evaluation

Delivering the step response test results and comparing different outcomes are the major tasks of the post-processing unit of the application. To ensure proper results reporting and documentation, the user needs to guide the process with key inputs in the post-processing stage. The results are compiled using graphical data. The output's curves may be used to extract a variety of data. In addition, setpoint configuration involves using step modifications. Different colors of lines, both horizontal and vertical, are used to depict these modifications. They are unique in that they can pinpoint three distinct points in the step response test: the state of the system prior to the test, the transition that occurs as a result of the step response test execution, and the state of the system at the point where the variable setpoint value returns to the initial value. It is also easy to see the specified dead-band region in the resulting plots. The tests are often finished after the system attains a previous stability point since interruptions from the past might significantly reduce the test's overall effectiveness. Nevertheless, the success criterion of the test depends heavily on the time interval that follows the setpoint modification. At last, the provided results will show if the system reverts to its initial state when the test is reset, and the controlled variable goes back to its initial value. A continuous line is used to depict the operational data. The graph legend presents each of those with a distinct color to aid in recognizing their originating source.

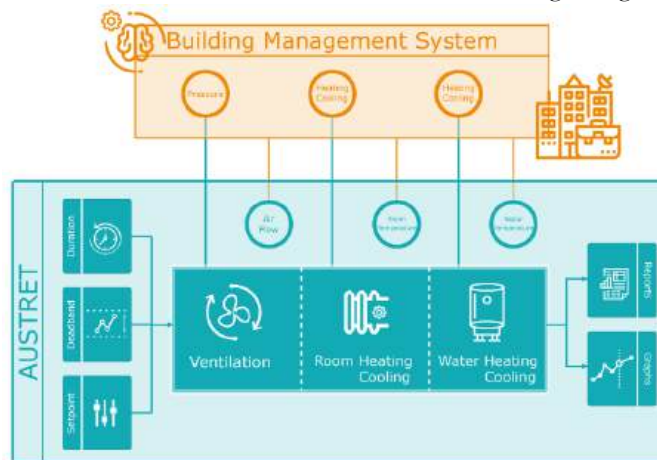


Figure 1 The testing and commissioning application architecture.

APPLICATION DEVELOPMENT AND INTERFACE

The BMS testing and commissioning application is designed and built considering the described building management systems' auditing methodology and design framework. The application seeks to connect and link to the BMS, allowing it to search, change, and record information. The BMS values used in the step response testing are logged

in an asynchronous manner. All the recorded data is archived on the file system of the evaluator and may be utilized to create a comprehensive report to serve as a basis for the building automation system auditing and functionality evaluation. Furthermore, the application is developed to manage parallel step response testing processes, considering different systems and different spaces in the targeted building. With regard to the building management system, the proposed application offers a detailed search feature based on keywords. At this stage, the auditor can identify and introduce inputs and outputs for the planned step test by searching for the specific name or phrase related to the input or output to be considered. As the searching for variables stage is done, the auditor has the option to look over the inputs and outputs that were previously included in the step response testing process before executing the planned tests. In addition, if multiple sets of tests are to be executed, maybe over different periods or days, the evaluators preferred settings and parameter definitions may be saved and reused later without spending time redefining inputs, outputs, and other parameters. Most importantly, before conducting the step response test, all the customized variables and parameters may be validated to check for potential issues, including duplicated names, missing or blank inputs, or improper execution time settings. This aids in a smooth and robust BMS performance testing process and ensures that the selected inputs and outputs will fulfill the requirements and conditions of the planned test.

In addition to the input and outputs, prior to conducting the step-response testing, the user also needs to introduce seven major parameters, as shown in Figure 2. Those parameters will characterize the type, duration, success criteria, and reported data of the specific test to be conducted. The input parameters are as follows: 1) the name of the step response test that will be conducted; 2) the test type, for instance, a temperature test or a pressure test; 3) the supplied setpoint value, which is the new imposed value of the setpoint considered; 4) the time duration of the test; 5) the time interval considered in the test, indicating the frequency of logging and recording data during the test; 6) a dead-band with the percentage, highlighting the upper and lower bounds with respect to the specified setpoint value; 7) a start forcing time, highlighting the date and time at which the test is planned to start and configuring the considered setpoint. Moreover, the BMS testing application is also capable of visualizing the variations in the inputs and outputs throughout the execution of a step response test. This is because input-related operations are handled by the primary thread in an asynchronous way. Nevertheless, tracking and storing the changes to the inputs and outputs are handled by a separate thread. The application continuously collects the data for the implemented step response test from system files in order to generate graphs that show the changes in inputs and outputs during step response tests. The data are then processed and reformed in a specific plotting dashboard.

Figure 2 Parameters definition for a step response testing process.

BMS PERFORMANCE TESTING IMPLEMENTATION

The developed BMS performance testing and commissioning tool was implemented and demonstrated considering an actual case study building, the Schneider Electric Greenhouse building, located in Ballerup, Denmark. The building

considered is a commercial complex, comprising personal offices, meeting rooms, working areas, and product showcase spaces. The case study building is equipped with an innovative management system and outfitted with modern management and operation functions and services that are primarily used by staff and employees on a regular basis to enhance the different systems operational efficiency and maintain satisfactory indoor comfort and air quality. The main ventilation system in the case study building is considered the object of investigation in this work to implement and demonstrate the BMS performance testing and commissioning tool. A screenshot of the considered ventilation system BMS platform is shown in Figure 3, highlighting the major parameters and variables as well as the main setpoints being controlled and managed in the daily operation.

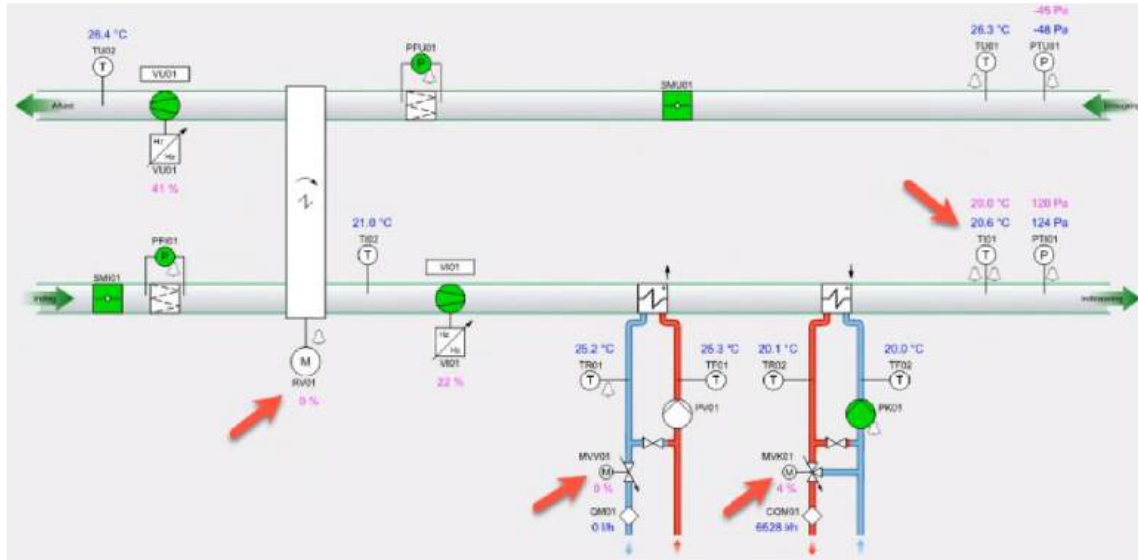


Figure 3 Ventilation system interface in the building management system.

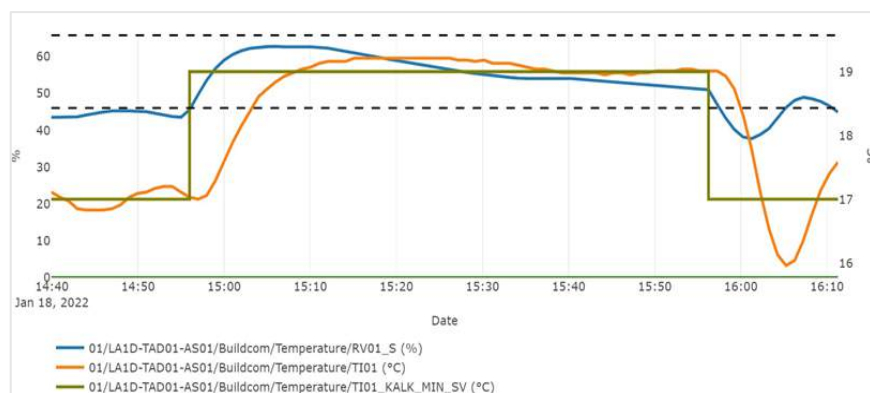


Figure 4 Test with increasing supply air temperature setpoint from 17 °C to 19 °C.

As shown in the figure, the ventilation fan comprises a supply and exhaust fan, a heat recovery unit, a pre-heating coil, and a pre-cooling coil. The variable TI01 indicates the actual temperature of the air supplied to the space. The degree of opening of the valve that permits the flow of air throughout the heat exchanger unit in the system is represented by the signal RV01-S. In addition, another valve is used to operate and control the preheating loop unit connected to the ventilation system by accessing the MVV01 variable, while the precooling loop is controlled by the valve represented by the variable MVR01, as highlighted in the figure. In this work, the step response test that will be

carried out aims to test and evaluate the performance of the ventilation system heating loop. The test involves varying the temperature setpoint of the supply air to the space and investigating the response of the ventilation preheating loop and heat recovery unit by recording the actual temperature at variable TI01. In this regard, Figure 4 shows a step response test carried out on the identified system with increasing the temperature setpoint of the supply air from 17°C to 19°C. In the figure below, the curve with an olive color denotes the heating setpoint, showing a clear increase in the setpoint from 17°C to 19°C at around 14:57 and then decreasing the setpoint again to 17°C after one hour of the test duration. The lower and upper boundaries of the dead-band are denoted by the two dashed lines with temperatures of 18.5°C and 19.5°C, respectively. The orange curve shows the actual temperature of the air supply recorded by the sensor at point TI01, introduced earlier. Also, the blue curve shows the valve opening of the heat recovery unit in percent. Overall, the TI01 signal, which reflects the real temperature reading from the corresponding temperature sensor installed in the regulated zone, shows that the system responds very well to the change in the temperature setpoint with a short delay. While the temperature setpoint shifts almost immediately, as presented in the figure, the actual room temperature (TI01) requires about ten minutes before it reaches the dead-band zone for the very first occurrence. The shift is seamless enough to prevent significant overshooting from occurring. After a few minutes, the system is found to stabilize with an actual temperature of about 19°C. Until the target setpoint is returned to its initial value (17°C), at which point a new stability dynamic has to be formed, the actual temperature is found to be maintained inside the dead-band region.

Furthermore, the signal RV01-S, which reflects the valve's opening and the system's ability to flow air across the integrated heat recovery unit, is likewise depicted in Figure 4. There is no connection between the valve RV01-S and the active heating module in the ventilation unit. On the contrary, it is an interior valve that may allow a heat transfer occurrence between the exhaust air from inside the building environment and the fresh supplied ambient air. The valve is found to be open to heat the environment from a point that corresponds to 44% and overshoots to levels close to 60% in a proper and logical reaction to the increase in heating setpoint before stabilizing at a value near 50%. Overall, it can be highlighted that the test carried out on the ventilation system air heating control performance in the case presented in Figure 4 is an actual successful test. In general, there is very little temperature overshooting, and the reaction is quite smooth. The delay is within the time range anticipated due to the dynamic nature of the heating application. In addition, the location of the duct opening is sufficient to supply the required airflow. The test ends with a fresh setpoint adjustment that returns the temperature to its initial setpoint of 17°C. After this point, stability cannot be evaluated. But the new configuration value may be assessed using the same inputs.

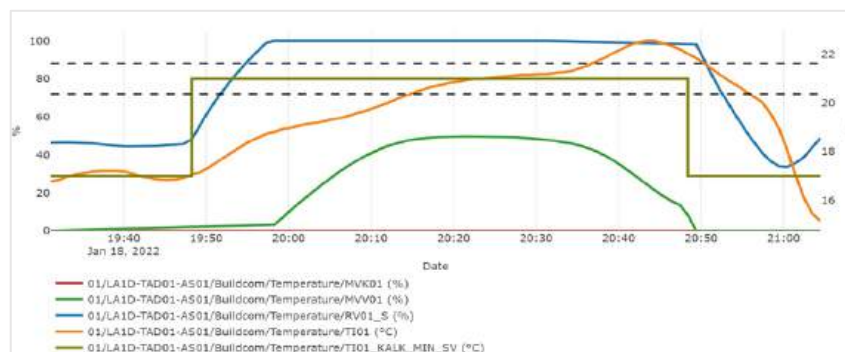


Figure 5 Test with increasing supply air temperature setpoint from 17°C to 21°C.

Another test was performed considering the same ventilation system heating unit as the investigation object. In this case, the temperature setpoint of the supply air is increased from 17°C to 21°C, as shown in Figure 5. The same variables and parameters described earlier in Figure 4 are present in the new test, except for the higher increase in the setpoint by 4°C in this test compared to 2°C in the previous one. Also, the overall duration of the test is set at one hour in this case, changing the setpoint at 19:48 and decreasing it again at 20:48. The figure highlights that the step response

test in this case did not attain stability across the examined time period. The actual temperature recorded at TI01 illustrates this. Although it reaches the dead-band zone set, the temperature does not remain stable at the changed setpoint value. The dead-band is actually broken at the top boundary, where the actual temperature exhibits a continuous increase. It is clear that the actual temperature didn't stabilize within the dead-band in the one-hour duration of the conducted test. When the temperature setpoint value changes back to the old value of 17°C, the actual temperature is shown to decrease. Looking at the figure, it is clear that both heating sources are operating in this case, in contrast to the previous test. Both valves, the air flow opening valve through the heat exchanger (represented in blue) and the preheating loop water supply valve (represented in green), are open, and the two components are providing heat during the test period. This is because the increase in the heating setpoint is more significant in this case (4°C), and thus an immediate heat supply is needed. This all led to the actual air supply temperature increasing drastically and breaking the dead-band upper limit.

While this test is considered unsuccessful as the actual temperature variable didn't stabilize within the dead-band, the main question to ask here is: would the test be successful if the time duration of the test was increased in comparison to the one-hour duration in the current scenario? To evaluate this parameter's impact, the test was carried out again, with an extended duration of 2 hours this time. As shown in Figure 6, it is clear that the test is now successful, as after overshooting at the beginning of the test, the actual temperature (in orange) tends to drop and start stabilizing within the set dead-band. The results indicate the importance of selecting the right parameters and constraints for the test to be conducted. The three tests conducted and presented above all involve an increasing temperature setpoint process. Nevertheless, multiple tests were also conducted considering the ventilation system, but with a decreasing temperature setpoint to investigate the heating components' reaction and performance response. For instance, Figure 7 shows the results of a step response test where the temperature setpoint is decreased from an initial value of 20°C to 15°C. Overall, it should be noted that while the previous tests involving the increase in the temperature setpoint called for additional heating by the ventilation heating components, the test in this case, with the decrease in the temperature setpoint, is actually pushing towards limiting the amount of heating supplied to the ventilation air supply. This is very clearly observed in the figure below. In this regard, the heat recovery unit valve opening (shown in blue) was at 100% before decreasing the setpoint and is reduced drastically to less than 40% when the temperature setpoint becomes 15°C. In addition, the preheating loop valve opening (shown in green) was also at around 35% and is decreased to 0% with a totally shut-off preheating loop when the setpoint is decreased. Moreover, the actual supply air temperature recorded and shown in orange exhibits a clear declining trend when the temperature setpoint is reduced and tends to stabilize within the dead-band after an hour of the test duration. This again demonstrates successful step-response testing of the considered heating system. It should also be noted that when the setpoint is increased again to 20°C at the end of the test, both valves of the heat recovery unit and the preheating loop show a drastic increase in value and start operating again, delivering heat to the supply air.

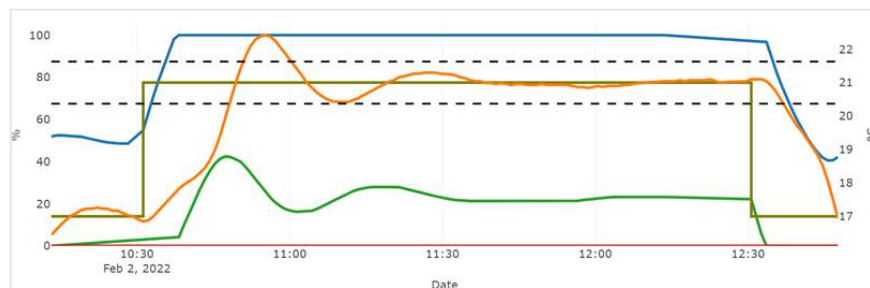


Figure 6 Test with increasing supply air temperature setpoint from 17°C to 21°C, with extended time.

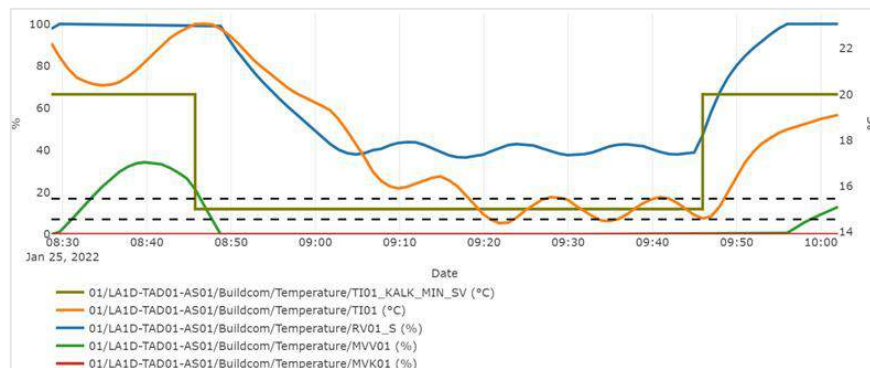


Figure 7 Test with decreasing supply air temperature setpoint from 20°C to 15°C.

CONCLUSION

This study introduced an innovative and systematic commissioning application for building management systems. The application employs step response testing as a method to observe, examine, and assess the functionalities and control capabilities of the BMS. A field demonstration was conducted to evaluate and verify the performance of the ventilation system in an actual office building. The results emphasize the effectiveness of the proposed tool in ensuring a well-designed building automation system, seamless integration, and smooth interaction with other building systems. Nevertheless, the study highlighted the importance of inputs selection for the desired tests, as well as the expertise and knowledge of the evaluator concerning the system evaluated. While a ventilation system would have a response time of minutes, a space heating application may require a testing timespan of hours. This work opens the door for future upgrades and follow-up investigations, including generalizing the approach to operate with different BMS types and architectures, evaluating other HVAC system configurations, and establishing an automated evaluation process.

ACKNOWLEDGMENTS

This work was supported by ‘Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle’ project, funded by the Danish Energy Agency under EUDP, ID number: 64021-1009.

REFERENCES

- Chauhan, R.K., and K. Chauhan. Building automation system for grid-connected home to optimize energy consumption and electricity bill”, *Journal of Building Engineering* 21:409-420.
- Dascalaki, E.G., Argiropoulou, P., Balaras, C.A., Drousa, K.G., and S. Kontoyiannidis. 2021. Analysis of the embodied energy of construction materials in the life cycle assessment of Hellenic residential buildings. *Energy and Buildings* 232: Article 110651.
- De Wilde, P. 2014. The gap between predicted and measured energy performance of buildings: a framework for investigation. *Automation in Construction* 41:40-49.
- Economidou, M., Todeschi, V., Bertoldi, P., D'Agostino, D., Zangheri, P., and L. Castellazzi. 2020. Review of 50 years of EU energy efficiency policies for buildings. *Energy and Buildings* 225: Article 11032244.
- Engvang, J.A., and M. Jradi. 2021. Auditing and design evaluation of building automation and control systems based on eu.bac system audit – Danish case study. *Energy and Built Environment* 2:34-44.
- Frei, B., Sagerschnig, C., D. Gyalistras. 2017. Performance gaps in Swiss buildings: an analysis of conflicting objectives and mitigation strategies. *Energy Procedia* 122:421-426.
- Jradi, M., Arendt, K., Sangogboye, F.C., Mattera, C.G., Markoska, E., Kjærgaard, M.B., Veje, C., and B.N. Jørgensen. 2018. ObepME: An online building energy performance monitoring and evaluation tool to reduce energy performance gaps. *Energy and Building* 166:196-209.
- Jradi, M., Boel, N., Madsen, B.E., Jacobsen, J., Hooge, J.S., and L. Kildelund. 2021. BuildCOM: Automated auditing and continuous commissioning of next generation building management systems. *Energy Informatics* 4:2.

- Jradi, M., Liu, N., Arendt, K., and C.G. Mattera. 2020. An automated framework for buildings continuous commissioning and performance testing – a university building case study. *Journal of Building Engineering* 31: Article 101464.
- Mirnaghi, M.S., and F. Haghighat. 2020. Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review. *Energy and Buildings* 229: Article 110492.
- Pan, T., Chen, J., Xie, J., Chang, Y., and Z. Zhou. 2020. Intelligent fault identification for industrial automation system via multi-scale convolutional generative adversarial network with partially labeled samples. *ISA Transactions* 101:379-389.
- Santos, A., Liu, N., and M. Jradi. 2021. AUSTRET: An Automated Step Response Testing Tool for Building Automation and Control Systems. *Energies* 14(13):3972.
- Van Dronkelaar, C., Dowson, M., Burman, E., Spataru, C., and D. Mumovic. 2016. A review of the energy performance gap and its underlying causes in non-domestic buildings. *Frontiers in Mechanical Engineering* 1:17.
- Wang, W., Brambley, M.R., Kim, W., Somasundaram, S., A.J. Stevens. 2018. Automated point mapping for building control systems: recent advances and future research needs. *Automation in Construction* 85:107-123.

Journal of Building Performance Simulation

Development and Demonstration of a Digital Twin Platform Leveraging Ontologies and Data-driven Simulation Models

Submission ID	241418114
Article Type	Research Article
Keywords	Digital Twin, Data-driven, Building Energy Model , Real-time, Ontology
Authors	Jakob Bjørnskov, Avneet Badhwar, Deep Singh , Mohit Sehgal, Rasmus Åkesson, Muhyiddine J radi

For any queries please contact:

TBPS-peerreview@journals.tandf.co.uk

Note for Reviewers:

To submit your review please visit <https://mc.manuscriptcentral.com/TBPS>

Development and Demonstration of a Digital Twin Platform Leveraging Ontologies and Data-driven Simulation Models

Jakob Bjørnskov^{1*}, Avneet Badhwar², Deep Shikhar Singh²,
Mohit Sehgal², Rasmus Åkesson³, Muhyiddine Jradi¹

¹SDU Center for Energy Informatics, University of Southern Denmark, Odense, Denmark

²NEC Corporation India Pvt Ltd., New Delhi, India

³KMD, Ballerup, Denmark

*Corresponding author, Email address: jabj@mmmi.sdu.dk

Abstract

The building sector is currently seeing rapid digitalization through IoT sensor and smart meter networks. In this regard, the concept of digital twins for building operation is a recent but promising research area concerned with the integration of IoT devices, data-driven modeling, monitoring, and optimization. In this work, a digital twin platform is presented and implemented, showcasing various applications and real-time services for improved operation of buildings. The digital twin platform is comprised of three parts: 1) a virtual representation of the physical building in the form of ontology-based and data-driven component models, 2) a data layer responsible for communication between the virtual system and the various measuring devices installed in the physical system, and 3) an interactive user interface layer, visualizing building performance. The validity and added value of the platform are demonstrated in a case study of a university building.

Keywords: *Digital twin, Data-driven, Building Energy Model, Real-time, Ontology*

1 Introduction

The current digitalization of the building sector, including the increasing availability of operational data from buildings, favors the use of scalable and automated data-driven modeling methods. In this regard, recent studies introduced the concept of a building digital twin (DT) Lu et al. (2019); Zhao et al. (2022), serving as a digital replica that senses, adapts, and optimizes operation according to the behavior of the actual building through the installed Internet of Things (IoT) devices. Such DTs can be used to deliver various services such as performance monitoring, fault detection, advanced operational planning, and optimization.

In recent years, the potential of DTs to yield useful insights and deliver crucial services has been demonstrated in multiple domains, e.g. the manufacturing, production and aerospace industry. However, one of the critical challenges in developing DTs is ensuring

semantic interoperability between different systems and data sources.

1.1 Data models

Traditionally, Building Information Models (BIM) have played a major role in the digitalization of buildings, primarily as part of the design and construction phase. The continuous adoption of BIM in the industry and research, and its capabilities as a data model, currently make it an attractive foundation for DTs. Recent studies on BIM-based DTs include Agostinelli et al. (2021), who developed a DT platform for a residential district in Rome based on the Industry Foundation Classes (IFC). Through integration with the energy simulation software Revit Suite of Mc4, the platform was used to assess the potential energy savings of a smart-grid integration. Desogus et al. (2021) and Casillo et al. (2023) both investigated real-time integration of IoT devices with a Revit-based interface.

While most of the current literature on DTs in the building domain is focused on BIM integration, other domains with IoT-centric development have seen significant adoption of ontologies and semantic web technologies to facilitate a more dynamic and interoperable environment Pauwels et al. (2017); Barnaghi et al. (2012); Boje et al. (2020). As argued by Boje et al. (2020), status quo BIMs are not easily compatible with IoT integration due to their legacy data standards and originally intended use as a static information exchange platform.

The number of studies in the building domain implementing ontology-based DTs is currently scarce. However, some studies have begun to explore both hybrid approaches combining BIM and ontologies as well as pure ontology-based DT architectures. For instance, Arsiwala et al. (2023) proposed a DT platform based on the Azure Digital Twin framework and the RealEstateCore ontology for monitoring indoor CO₂ concentration. A data-lake architecture was proposed by Merino et al. (2022), building upon both the IFC standard and the Brick ontology. Similarly, El Mokhtari et al. (2022) also employed a data-lake

architecture using the Building Topology Ontology (BOT) and Brick ontologies for their DT data model, improving interoperability and making access to data streaming endpoints more accessible

1.2 Simulation models

While an interoperable data model is crucial for managing various sources of data, in order to enable predictive capabilities and to bring the digital representation alive, simulation models must be integrated.

Within the building domain, the modeling type is typically divided into white-box, black-box, and grey-box. White-box modeling, which has been used for decades in Building Performance Simulation (BPS) tools such as EnergyPlus Crawley, Lawrie, Winkelmann, Buhl, Huang, Pedersen, Strand, Liesen, Fisher, Witte, and Glazer (Crawley et al.), relies on detailed physical equations to describe building behavior. However, their usefulness in DTs is limited compared to more data-driven approaches, due to the manual and time-consuming model development and calibration processes, which often require expert domain knowledge Bazjanac et al. (2011); Kim et al. (2016).

In opposition, black-box models are based entirely on data, typically using machine-learning and artificial intelligence modeling approaches. Studies by Agostinelli et al. (2021) and Ni et al. (2024, 2023) demonstrated the use of machine-learning techniques for building energy management and forecasting. Bárkányi et al. (2021) showcased gradient-boosting regressors for anomaly detection, while Arsiwala et al. (2023) employed a stochastic gradient descent (SGD) regressor with L2 regularization for modeling indoor CO₂ concentration. The authors noted that their model was able to capture high-level trends. However, to improve model accuracy, they concluded that more training data is required.

Recognizing the limitations of purely data-driven approaches, hybrid methods that combine domain knowledge with black-box modeling, have also been explored. Hosamo et al. (2022, 2023) combined hand-made expert rules with machine-learning techniques such as Bayesian networks, support vector machines, and artificial neural networks. They argue that this approach reduces the high requirements for data quality and quantity, typically associated with black-box models. Similarly, Bjørnskov and Jradi (2023a) demonstrated the use of a physics-informed neural network for predicting indoor temperatures as part of a larger energy modeling framework.

As a mix of white-box and black-box models, Grey-box models incorporate both physical principles and data-driven techniques. Clausen et al. (2021) employed such models for Model Predictive Control in a DT framework focused on energy efficiency and occupant comfort. Andersen et al. (2023) implemented a ventilation model using grey-box component mod-

els for CO₂ and fan power consumption predictions. In a case study, Bjørnskov and Jradi (2023a) combined a physics-informed neural network with grey-box models of various building components such as a space heater, heating coil, heat recovery unit, and controller, to deliver performance monitoring and operational strategy planning services.

In summary, the development of DTs is a multidisciplinary effort that leverages advancements in data modeling, semantic interoperability, and data-driven simulation models. By providing comprehensive and real-time insight into the operational performance of buildings, DTs have the potential to improve the efficiency and management of the built environment. The main challenge and focus of this paper is the integration of ontology-based data models with advanced simulation techniques, ranging from black-box machine learning approaches to grey-box models. This will enable a new level of understanding and optimization of building operations through services such as performance monitoring and advanced operational planning.

2 Aim and contribution

This work demonstrates the development and implementation of a real-time DT platform that uses data-driven component models and the Smart Applications REFERENCE (SAREF) ontology as a common data model to improve interoperability. The proposed platform is implemented and assessed by considering a case study of a classroom in a university building, serving as a living lab. The building has detailed sub-metering on various levels, including the classroom equipped with sensors and monitoring devices measuring damper and radiator valve positions, as well as temperature and CO₂-concentration. The collected data is used to calibrate the component models associated with the space. This includes a space model that accurately predicts indoor temperature, as well as models for valves, space heaters, and controllers.

The developed platform targets three high-level services; I) A user-friendly interface allowing building owners to better understand how their building behaves and performs; II) Real-time automated performance monitoring and continuous commissioning for improved building operation and smarter facility management; III) Model-based scenario testing and operational optimization to enable planning support. The platform includes real-time services such as performance monitoring and forecasting which has been running continuously for an extended period.

The paper is structured as follows: First, we present the case study in section 3 as this forms the basis for the proceeding sections, section 4 presents an overview of the developed platform along with the architecture, its constituents, and the flow of information. Following, in section 5, the virtual system

and simulation models used to replicate the behavior of the real system are described. The data processing layer, responsible for connecting the physical and virtual system, is described in section 6. Finally, the services delivered by the DT platform are presented and evaluated in section 7, followed by a conclusion in section 8.

3 Case study

To validate and showcase the potential of the developed DT platform, it is implemented on a case study. The considered case study is a classroom located on the second floor of a Danish university building as shown in Figure 1. The building was built in 2015 and attained the Danish building class 2020 as one of the first buildings in Denmark. It has a total floor area of 8500 m² with four stories of which the ground, first, and second floors contain offices, study zones, and classrooms while the basement contains storage rooms and technical installations.

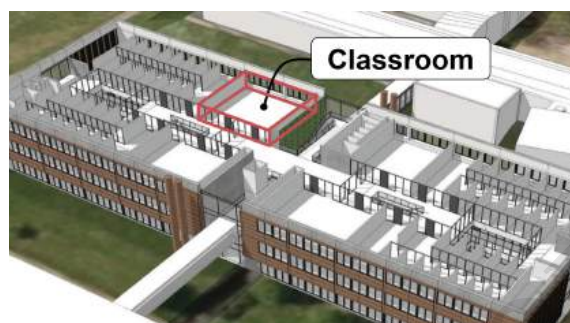


Figure 1: Section overview of the second floor where the considered classroom is located.

The building is divided into four symmetrical quadrants each supplied by a balanced ventilation system. Each system has one heat recovery unit, one heating coil, and the supply air temperature is under normal conditions controlled based on the exhaust air temperature according to a piecewise linear curve. Similarly, the supply water temperature to the radiators is controlled based on the outdoor air temperature measured by a weather station placed on the roof of the building.

The considered classroom is heated with radiators and the supply air flowrate is controlled through Demand Controlled Ventilation (DCV), where the damper opening is actuated based on the measured CO₂ concentration in the room. Shades are installed on the building facade to avoid overheating during the summer months. The actuation signal of the space heater valves, dampers, and shades is controlled centrally by the Building Management System (BMS).

The classroom is equipped with multiple sensors, namely, temperature, CO₂, valve position, damper position, and shade position. In addition, a heat consumption meter is installed on the radiators.

4 Digital twin Architecture overview

In this section, the overall architecture of the DT platform is presented along with the high-level considerations behind the design. The DT concept considered in this work follows the definition used by Grieves (2015) and Boje et al. (2020), i.e. that the DT concept is composed of three main constituents; the physical system, the virtual system, and a flow of data between these systems.

The overall architecture of the platform is shown in Figure 2 where four high-level groups are highlighted. The physical system is the asset to be managed, where the IoT devices such as indoor climate and flow temperature sensors, energy consumption and flow meters, valve and damper position sensors and actuators, etc. play a major role, as they allow for continuous and automated data collection and feedback that can be processed in the virtual system. The role of the virtual system is to mimic the physical system as closely as possible through an accurate dynamic simulation model to enable different services such as continuous commissioning, operational optimization, and scenario testing. These services are exposed to the user of the platform through the user interface layer. The data processing layer acts as the communication link between the physical and virtual systems by collecting, cleaning, and storing data.

Based on the discussed importance of interoperability and data fluidity in a potential DT solution, one of the early design decisions of the platform was to decide on a common semantic model. Therefore, both the data processing layer and the virtual system make use of the SAREF ontology Daniele et al. (2015). SAREF was developed in close collaboration with the industry to rectify the interoperability issues encountered in the IoT-domain. The ontology has a focus on devices, e.g. a light switch or a sensor with extensions covering multiple domains, e.g. SAREF4BLDG Poveda-Villalón and Garcia-Castro (2020), SAREF4CITY Poveda-Villalón et al. (2020), SAREF4SYST Lefrançois (2019), SAREF4ENER Daniele (2020), etc. Here, SAREF4BLDG is an ontology extension dedicated to the building domain, whose classes are a subset of the Industry Foundation Classes (IFC) standard. This is considered a significant advantage as it potentially allows for direct mapping of properties and design values for the different building components such as coils, space heaters, dampers, etc.

5 Virtual system

For the implementation presented in this work, the virtual system consists of four closely connected components; a model library, a semantic model, a translator, and a simulator. This is encapsulated by the general-purpose and ontology-based modeling framework and Python library developed as part of earlier

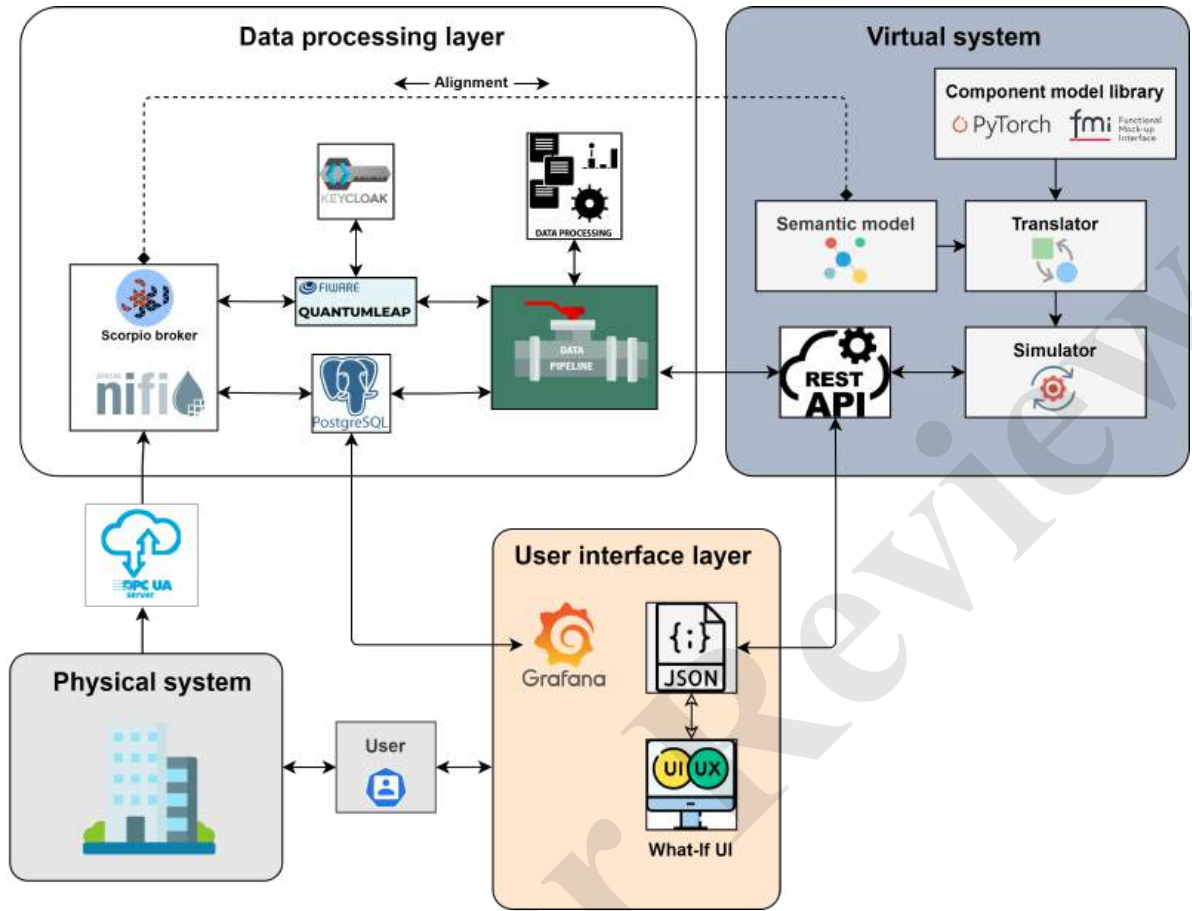


Figure 2: Architecture of the DT platform and the information flow between high-level components. Adapted from Bjørnskov et al. (2024).

work Bjørnskov and Jradi (2023b); Bjørnskov (2023).

5.1 Semantic model

As a method to represent the physical system in a comprehensive, accurate, and machine-readable format, a semantic model is used to represent information about system topology, relations, and properties using concepts from the SAREF core ontology, the SAREF4BLDG and SAREF4SYST extensions, and the Flow Systems Ontology (FSO). The semantic model for the case study is shown in Figure 3. Using the qualified relations and classes of the applied ontology to represent the classroom, a graph-like structure emerges. For instance, it is specified that the `<valve>` $\rightarrow$ `hasProperty` $\rightarrow$ `<openingposition 2>` and that the `<openingposition 2>` $\rightarrow$ `isControlledBy` $\rightarrow$ `<temperature controller>`. In addition, the `<temperature controller>` $\rightarrow$ `observes` $\rightarrow$ `<temperature 0>` and the `<temperature 0>` $\rightarrow$ `isPropertyOf` $\rightarrow$ `<space>`. In short, this specifies that a motorized valve is controlled to regulate the indoor temperature in a space. As demonstrated through this example, the use of an ontology enables an accurate representation of the real system in a

machine-readable and structured format.

5.2 Component model library

The model library consists of different grey-box and black-box component model classes that represent the dynamic behavior of the SAREF and SAREF4BLDG devices contained in the semantic model. Each model class inherits design properties from the SAREF4BLDG device classes and has tuneable parameters that can be adjusted based on data collected from the physical system. For the considered case study, the used component models were presented and validated in earlier work Bjørnskov and Jradi (2023a).

5.3 Translator

The translator is responsible for assembling and connecting models from the library according to the system topology and relations contained in the semantic model. The translation method is described in detail in Bjørnskov et al. (2024). For the case study classroom, the information flow for the obtained simulation model is shown in Figure 4. The Out/In labels at the edges show the translation of output to input between models. As shown in the graph, virtual

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

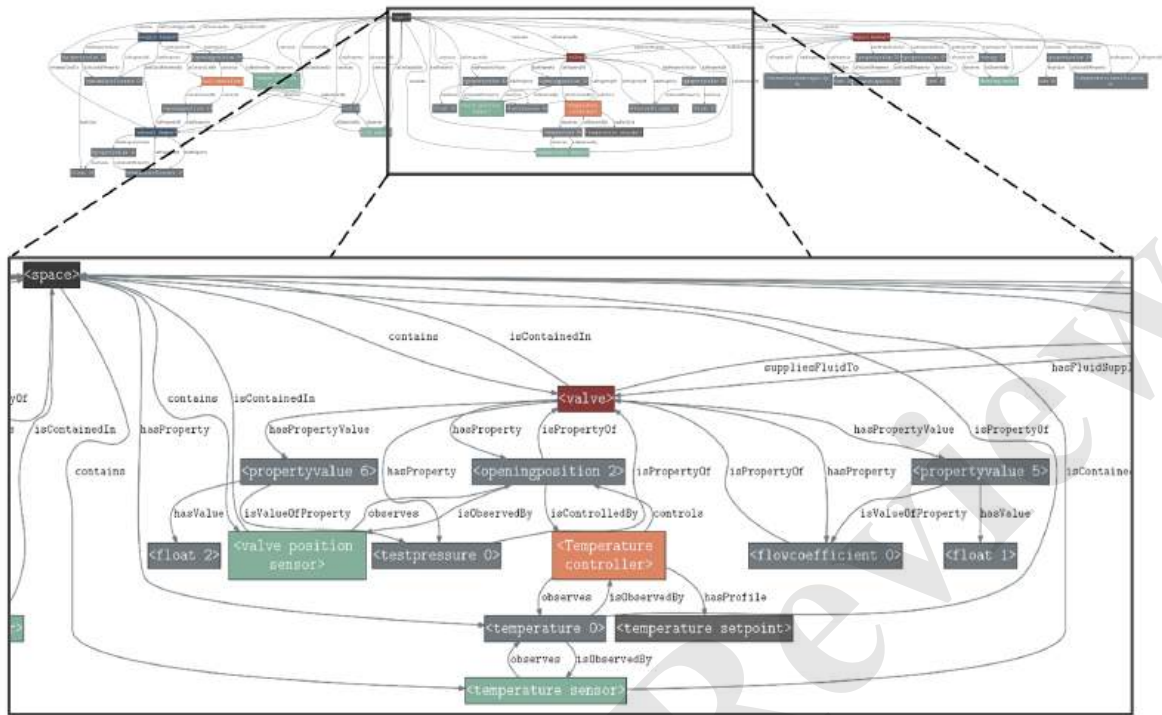


Figure 3: Semantic model of the classroom case study representing the system topology and equipment design information of the physical system. The model is based on concepts from the SAREF ontology extended with concepts from SAREF4BLDG, SAREF4SYST, and FSO.

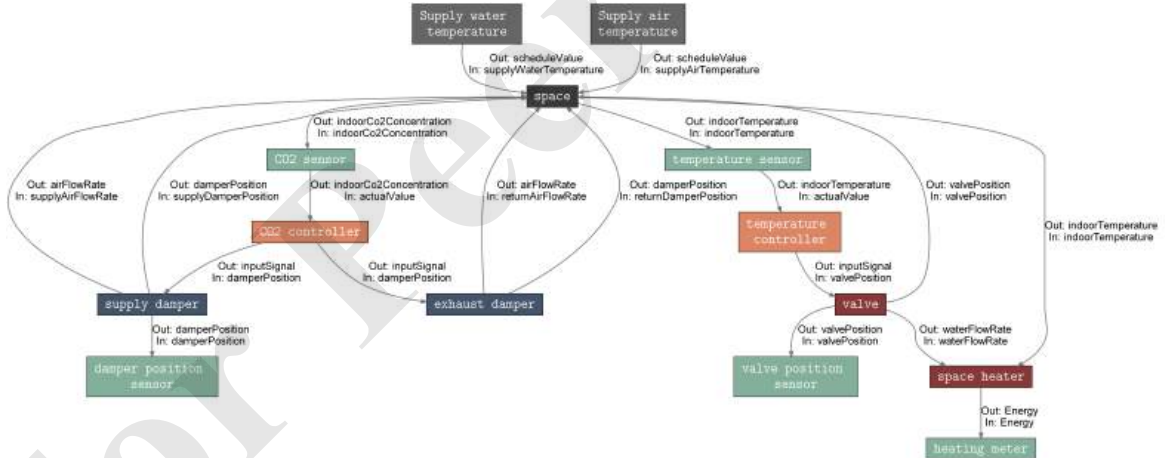


Figure 4: Graph of the generated simulation model for the classroom case study showing the information flow and logical connections between component models. The graph has been automatically generated using the open-source Python package available on GitHub Bjørnskov (2023).

sensors and meters are placed in the model (green components), e.g. the **temperature sensor** or the **heating meter**. Readings from these virtual measuring devices can be directly compared with readings from actual sensors and meters placed in the physical system.

5.4 Simulator

Given a model of the physical system, the simulator is responsible for simulating the dynamic model response for a given period and set of inputs. This is done by first performing a topological sort of the model graph shown in Figure 4 and then traversing through the obtained list of components models for each timestep in the simulation. The details of this

process are explained in Bjørnskov and Jradi (2023b).

6 Data processing layer

Data integration is a fundamental aspect of DT technology. In this section, it is explained how data is ingested into simulation models from a database, the cleaning and formatting of data for use in API, and the creation of a FastAPI application for deploying simulation models as containerized services. Furthermore, the deployment of this application using Docker and Kubernetes is explored.

6.1 Data Ingestion from the Physical System

The data ingestion process for the DT entails interfacing with an OPCUA server, serving as the primary data source. Initiated through predefined mappings stored in taglist or room-name files, this process enables the retrieval of attributes and their corresponding values. Subsequent to data retrieval, pre-processing procedures are conducted via NiFi processors and Python scripts, aimed at distilling pertinent attributes while structuring the data into a JSON format. Following transformation, the data is transposed into NGSI-LD format to align with the specifications of the Scorpio Broker, responsible for managing entity creation and subscription. Real-time data dissemination to Quantumleap for historical data management is facilitated through Scorpio Broker subscription. Integration with Kafka affords the direct extraction of Energy Key data from pre-configured topics, obviating the necessity for preliminary processing. The framework's scalability is bolstered by provisions for dynamic updates to the request file, accommodating novel room details and sensor information. This adaptability is seamlessly integrated with Python scripts, ensuring compatibility with evolving data structures. Leveraging Apache NiFi, Scorpio Broker, Quantumleap with Timescale DB, Python, and Kafka technologies, this comprehensive approach ensures the efficient and streamlined acquisition of data from physical systems for DT environments.

6.2 Data Retrieval from Database

The process of ingesting data is where it all begins. In the context of DTs, two distinct methods are employed for data retrieval from a PostgreSQL database. SQLAlchemy, a versatile Python library, acts as a reliable means to fetch data from a PostgreSQL database. Keycloak and QuantumLeap APIs provide an additional layer of security and data harmonization during the data retrieval process. Keycloak ensures that only authorized individuals gain access to the data, maintaining the integrity of the DT's resources. QuantumLeap, working in tandem with Keycloak, standardizes data retrieval from various sources, acting as a translator for data. This ensures a common language is spoken, much like bridging communication gaps between different systems.

6.3 Data Cleaning and Formatting

After acquiring the data, the next critical step is to prepare it for use within the DT. It involves identifying and rectifying errors and inconsistencies, ensuring data accuracy and consistency. This makes the data more accessible and understandable within the DT.

6.4 Creating a REST-based API for simulation models

The final stage of data integration involves the development of an API application tailored to deploy models as containerized services. This application serves as a bridge, enabling the ingestion of formatted data into simulation models. FastAPI is a rapid web framework that simplifies the creation of application interfaces, much like a fast delivery service. With FastAPI, an application is made that allows simulation models to quickly access clean and organized data.

6.5 Deploying with Docker and Kubernetes

Once the FastAPI application is ready, it can be deployed using Docker and Kubernetes. Docker packages the application, while Kubernetes manages its deployment, scaling, and maintenance. Docker constitutes a container for the application, making it easy to transport, and Kubernetes acts as an organizing agent, making sure it runs smoothly.

6.6 Data Integration with Visualization layer

The process of transforming raw data into actionable insights through visualization is of great importance. In this section, the choice of Grafana as the key visualization tool and its integration with the data source PostgreSQL is highlighted. The combination of PostgreSQL and Grafana makes it possible to create dynamic custom dashboards that provide clear and insightful representations of the data stored in the database, facilitating informed decision-making and real-time monitoring.

Grafana is a leading open-source data visualization and monitoring platform. It offers a versatile and user-friendly environment for designing, building, and sharing custom dashboards. Its capabilities extend to a wide range of data sources, making it a versatile choice for visualization.

Custom dashboards tailored to specific needs play a crucial role in the data integration process. These dashboards are designed to visually represent the data stored in the PostgreSQL database, allowing for quick and intuitive data exploration. It enables displaying key performance indicators, monitoring system metrics, and visualizing critical data points. This visualization facilitates both data comprehension and decision-making.

Grafana's core functionalities are instrumental in fetching data from the PostgreSQL database and enabling dynamic visualization. By configuring data sources and queries, the database can seamlessly be



(a)



(b)

Figure 5: Dashboard in the Grafana interface showcasing real-time performance monitoring of the classroom. (a): Two days of normal operation. (b) An anomaly detected, most likely caused by an opened window.

accessed to retrieve real-time data and generate interactive visualizations. This integration is pivotal in transforming raw data into actionable insights, making Grafana an indispensable component of the data integration process.

In summary, data integration and Visualization are the backbone of DT technology. It involves data retrieval, cleaning, formatting, and making it accessible for simulation models and an easy-to-understand visual representation of data. The deployment of the API application using Docker and Kubernetes ensures that everything operates seamlessly. This integration is fundamental in enhancing decision-making,

monitoring performance, and facilitating a deeper understanding of the information stored in the database.

7 Digital twin service applications

The presented platform architecture shown in Figure 2 has been implemented based on the case study classroom and has been running continuously throughout the test period from December 2023 to April 2024. In this section, the different services are demonstrated.

7.1 Real-time performance monitoring

The developed dashboard for real-time performance monitoring of the classroom case is shown in Fig-

ure 5. Based on the results of the test period, the space heater is used as an example for the provided monitoring service. Two periods are shown: Figure 5a shows a four-day period with normal operation and Figure 5b shows a one-day period with a detected anomaly in the classroom. In both examples, readings from the **Temperature sensor** and **Valve position sensor** are shown. The upper plots show a direct comparison between the physical and virtual systems, while the lower plots show the performance gap calculated with a moving average of the residual with a window size of 1 day. This moving average is preferred to the raw residual as it produces a more robust and stable metric for anomaly detection. For Figure 5a, the readings from the virtual and physical sensors are generally in agreement with a performance gap within the 2°C allowed gap for the <Temperature sensor> and 50% for the <Valve position sensor>.

As mentioned, Figure 5b has been included to show the detection of an anomaly. As shown, the physical and virtual readings agree in the first part of the period. However, at around 08.00, the temperature in the physical system drops around 1 °C and the performance gap starts to increase. It has not been possible to confirm the exact cause of this, however, it is most likely caused by the opening of a window when the occupants arrive. The virtual system assumes normal operation, causing the gap between the virtual and physical systems to increase. As shown, the physical space heater in the classroom turns on to reach the 21 °C setpoint. At 18.00, the sun sets causing a rapid temperature drop in the classroom, further supporting the assumption of an opened window. As a result, the space heater operates at full capacity during the whole night, while the consumed energy is lost to the surroundings. From 08.00 when the window is opened, it takes 42 hours for the classroom temperature to return to normal, amounting to around 115 kWh of wasted energy or 30 % of the monthly heating consumption for the classroom. This waste of energy could have been avoided if action had been taken as soon as the anomaly was detected at 16.00 for the valve and 18.00 for the indoor temperature.

Although this anomaly stems from occupant behavior rather than a system fault, the monitoring case presented in this study demonstrates the potential of the developed platform and its capability to capture anomalies and performance drifting as soon as they occur. This is translated into shorter reaction times and resource savings. In addition, the presented approach can form a basis for a building's continuous commissioning process on different levels, serving as a backbone for effective and robust fault detection and preventive maintenance.

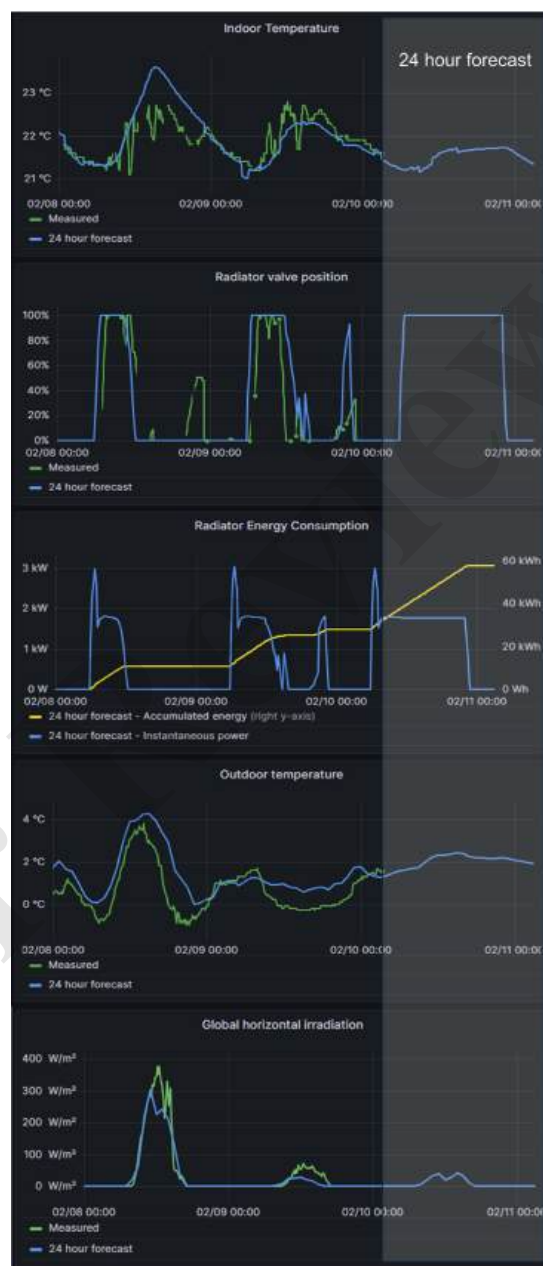


Figure 6: Dashboard showing the real-time forecasting service.

7.2 Real-time forecasting

One of the central capabilities of the DT platform is real-time forecasting. This is crucial for providing services such as operational optimization and making the building reactive to price signals and demand response measures in response to prices and environmental changes. Therefore, a real-time forecasting service has been implemented and has been running continuously on the considered case as shown in Figure 6. Here, five predicted variables are shown: Indoor temperature, valve position, outdoor tempera-

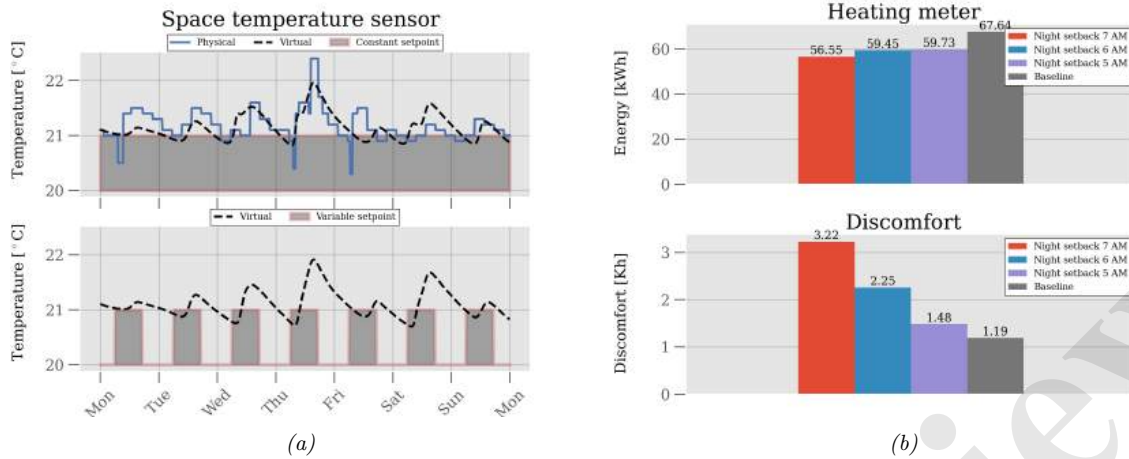


Figure 7: Comparison of four different scenarios for a 1-week period. (a) Comparison between the baseline scenario with a constant setpoint (top) and an alternative scenario with a varying setpoint (bottom). (b) Total predicted heating consumption on the space heaters (top) and the predicted discomfort (bottom).

ture, global horizontal irradiation, and radiator energy consumption. The radiator energy consumption plot shows both the predicted heating power and the aggregated energy. The forecasted outdoor temperature and global horizontal irradiation are obtained through a weather model API developed by the Danish Meteorological Institute (2024) (DMI) and provided as input or boundary to the virtual system. For the shown dashboard, a 24-hour forecast is considered and compared to the measured values as validation. As shown, the forecast generated by the virtual system generally agrees with observations.

However, larger prediction uncertainties should be expected compared to historical periods where actual observed data can be used as input. As shown in Figure 6, some additional uncertainty is introduced by inaccurate weather forecasts. However, as for the performance monitoring service, the largest source of uncertainty is occupancy and related factors such as open windows and doors, internal loads, etc.

7.3 Scenario testing

The third explored service of the DT platform is scenario testing, which aims to compare different candidate operational scenarios in a zero-risk virtual environment.

During some winter periods the building operates with a heating setpoint of constant 21 °C. A common strategy for reducing energy consumption in buildings is to decrease the temperature setpoint in periods with no occupancy, e.g. during nighttime. This strategy is called *night setback*, which must be implemented such that the building is preheated before occupation to avoid thermal discomfort. However, buildings behave differently under different weather and operational conditions. Exploration of different setpoint strategies and their impact on energy con-

sumption and indoor comfort is therefore important. In this work, four scenarios are explored, the baseline constant 21 °C setpoint of the classroom, and three variations of night setback scenario with a varying setpoint. The investigated night setback scenarios are to change the setpoint schedule to 21 °C from xx:00 to 17:00 and 20 °C otherwise, where the starting hour xx is varied between 05, 06, and 07. To measure the impact on indoor comfort, the aggregated metric proposed by Arendt et al. (2020) is used for measuring thermal discomfort in Kelvin-hours, which captures both the duration and intensity of setpoint violations. Thermal discomfort is only aggregated for hours with occupancy which is assumed to be between 08:00 and 17:00.

The results for the scenario testing are shown in Figure 7. Figure 7a shows the virtual temperature readings for the baseline (top) and the 07 AM night setback scenario (bottom). The baseline also includes the actual measurements from the physical temperature sensor during the period as a reference. As shown, the baseline virtual readings (as well as physical readings) fluctuate around the 21 °C setpoint. In the night setback scenario, the temperature is allowed to drop during the night until 07:00 where the temperature setpoint is again increased to 21 °C. Figure 7b shows the resulting difference in heating consumption (top) and discomfort (bottom). As shown there is a clear tradeoff between reduced energy consumption and discomfort for the chosen scenarios. The sooner space heating is initiated, the lower the energy consumption. However, this comes at the cost of higher discomfort levels. Typically, maintenance of thermal comfort levels is valued higher than decreases in energy consumption. Therefore, the most viable of the three-night setback alternatives is the 5 AM scenario, which impacts the thermal discomfort the least by

0.29 Kh but saves around 12 % in heat consumption.

8 Conclusion

In this paper, the implementation of a DT platform delivering various services was demonstrated in a case study. The platform is divided into three main parts, 1) a virtual system that emulates the physical system, 2) a data layer tying together the virtual and physical system, and 3) a user interface through which a user can interact with the virtual system and data layer.

The virtual system builds upon the SAREF ontology as a collection of components such as spaces, heaters, controllers, fans, etc., that can be directly related to their real counterparts, enabling detailed performance monitoring and fault detection at both component and system levels. This also enables building owners to test different operational strategies and how they affect the building before they are implemented.

Three DT services were demonstrated. Real-time performance monitoring was demonstrated by comparing virtual and physical sensor and meter readings for the indoor temperature of the classroom and valve position of the space heater. Here, anomaly behavior was detected on a period with an open window, which could potentially yield energy savings of 115 kWh or 30% of the monthly heating consumption, had the anomaly been detected earlier. As a second service, real-time forecasting of 24 hours was demonstrated, forming the basis for operational optimization. Finally, scenario testing was demonstrated by comparing baseline operation with three alternative operation strategies in terms of indoor comfort and consumed energy.

This work provides a proof-of-concept of the DT platform and the potential applications and services it can deliver. It therefore addresses the current needs for more practical DT implementations.

The use of a well-established ontology ensures the scalability of the approach and that the developed framework could potentially be expanded and adapted to clusters of buildings or cities in the future. For future work, it is intended to further scale up efforts, leading to the development of a whole-building DT. In addition, it is expected that the services provided could be expanded to consider also operational optimization and automated actuation of components or change of setpoints at a later stage.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability statement

The data that support the findings of this study are available from the corresponding author, JB, upon reasonable request.

Acknowledgment

This work is carried out under the ‘Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle’ project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64021-1009.

References

- Agostinelli, S., F. Cumo, G. Guidi, and C. Tomazzoli (2021). Cyber-physical systems improving building energy management: Digital twin and artificial intelligence. *Energies* 14(8).
- Andersen, A. H., J. Bjørnskov, and M. Jradi (2023). Adaptable and scalable energy modeling of ventilation systems as part of building digital twins. Volume 18 of *Building Simulation Conference Proceedings*, pp. 3046–3053.
- Arendt, K., A. Clausen, C. Mattera, M. Jradi, A. Johansen, C. Veje, M. Kjærgaard, and B. Jørgensen (2020). Multi-objective model predictive control framework for buildings. In *Proceedings of building simulation 2019*. Rome (Italy), 02-04 September 2019.
- Arsiwala, A., F. Elghaish, and M. Zoher (2023). Digital twin with machine learning for predictive monitoring of co2 equivalent from existing buildings. *Energy and Buildings* 284.
- Bárkányi, Á., T. Chován, S. Németh, and J. Abonyi (2021). Modelling for digital twins—potential role of surrogate models. *Processes* 9(3).
- Barnaghi, P., W. Wang, C. Henson, and K. Taylor (2012). Semantics for the internet of things: Early progress and back to the future. *International Journal on Semantic Web and Information Systems* 8(1), 1 – 21. Cited by: 436.
- Bazjanac, V., T. Maile, J. O Donnell, C. Rose, and N. Mrazović (2011, 01). Data environments and processing in semi-automated simulation with energypus.
- Bjørnskov, J. (2023). Energy modeling framework. <https://github.com/JBjoernskov/Twin4Build>, Last Accessed: 03-23-2023.
- Bjørnskov, J. and M. Jradi (2023a, September). Implementation and demonstration of an automated energy modeling framework for scalable and adaptable building digital twins based on the saref ontology. In *Proceedings of Building Simulation 2023*:

- 18th Conference of IBPSA, Volume 18 of *Building Simulation*, Shanghai, China, pp. 1192–1199. IBPSA.
- Bjørnskov, J. and M. Jradi (2023b). An ontology-based innovative energy modeling framework for scalable and adaptable building digital twins. *Energy and Buildings* 292, 113146.
- Bjørnskov, J., M. Jradi, and M. Wetter (2024). Automated model generation and parameter estimation of building energy models using an ontology-based framework. *Energy and Buildings*. submitted.
- Boje, C., A. Guerriero, S. Kubicki, and Y. Rezgui (2020). Towards a semantic construction digital twin: Directions for future research. *Automation in Construction* 114, 103179.
- Casillo, M., L. Cecere, F. Colace, A. Lorusso, F. Marongiu, and D. Santaniello (2023). An internet of things approach to support a high-tech house. *Colloquium in Information Science and Technology, CIST*, pp. 323–328.
- Clausen, A., K. Arendt, A. Johansen, F. C. Sangogboye, M. B. Kjærgaard, C. T. Veje, and B. N. Jørgensen (2021). A digital twin framework for improving energy efficiency and occupant comfort in public and commercial buildings. *Energy Informatics* 4.
- Crawley, D., L. Lawrie, F. Winkelmann, W. Buhl, Y. Huang, C. Pedersen, R. Strand, R. Liesen, D. Fisher, M. Witte, and J. Glazer. Energyplus: creating a new-generation building energy simulation program. *Energy and Buildings* 33, 319–331.
- Daniele, L. (2020). Saref4ener: an extension of saref for the energy domain. [Online]. Available: <https://saref.etsi.org/saref4ener/>, Last Accessed: 03-23-2023.
- Daniele, L., F. den Hartog, and J. Roes (2015, 08). Created in close interaction with the industry: The smart appliances reference (saref) ontology. pp. 100–112.
- Danish Meteorological Institute (2024). Dmi open data. [Online]. Available: <https://opendatadocs.dmi.govcloud.dk/>, Last Accessed: 06-10-2024.
- Desogus, G., E. Quaquero, G. Rubiu, G. Gatto, and C. Perra (2021). Bim and iot sensors integration: A framework for consumption and indoor conditions data monitoring of existing buildings. *Sustainability (Switzerland)* 13(8).
- El Mokhtari, K., I. Panushev, and J. J. McArthur (2022). Development of a cognitive digital twin for building management and operations. *Frontiers in Built Environment* 8.
- Grieves, M. (2015, 03). Digital twin: Manufacturing excellence through virtual factory replication.
- Hosamo, H. H., H. K. Nielsen, D. Kraniotis, P. R. Svennevig, and K. Svidt (2023). Digital twin framework for automated fault source detection and prediction for comfort performance evaluation of existing non-residential norwegian buildings. *Energy and Buildings* 281.
- Hosamo, H. H., P. R. Svennevig, K. Svidt, D. Han, and H. K. Nielsen (2022). A digital twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics. *Energy and Buildings* 261.
- Kim, H., Z. Shen, I. Kim, K. Kim, A. Stumpf, and J. Yu (2016). Bim ifc information mapping to building energy analysis (bea) model with manually extended material information. *Automation in Construction* 68, 183–193.
- Lefrançois, M. (2019). Saref4syst: an extension of saref for topology of systems and their interconnections. [Online]. Available: <https://saref.etsi.org/saref4syst/>, Last Accessed: 03-23-2023.
- Lu, Q., A. K. Parlikad, P. Woodall, X. Xie, Z. Liang, E. Konstantinou, J. Heaton, and J. Schooling (2019, 10). Developing a dynamic digital twin at building and city levels: A case study of the west cambridge campus. *Journal of Management in Engineering* 36.
- Merino, J., X. Xie, N. Moretti, J. Y. Chang, and A. Parlikad (2022). Data integration for digital twins in the built environment. *Proceedings of the European Conference on Computing in Construction*, pp. 199–206.
- Ni, Z., C. Zhang, M. Karlsson, and S. Gong (2023). Leveraging deep learning and digital twins to improve energy performance of buildings. 2023 IEEE 3rd International Conference on Industrial Electronics for Sustainable Energy Systems, IESES 2023.
- Ni, Z., C. Zhang, M. Karlsson, and S. Gong (2024). A study of deep learning-based multi-horizon building energy forecasting. *Energy and Buildings* 303.
- Pauwels, P., S. Zhang, and Y.-C. Lee (2017). Semantic web technologies in aec industry: A literature overview. *Automation in construction* 73, 145–165.
- Poveda-Villalon, M., R. Garcia-Castro, and P. Espinoza-Arias (2020). Saref extension for smart city. [Online]. Available: <https://saref.etsi.org/saref4city/>, Last Accessed: 03-23-2023.

837 Poveda-Villalón, M. and R. Garcia-Castro (2020).
838 Saref extension for building. [Online]. Available:
839 <https://saref.etsi.org/saref4bldg/>, Last Ac-
840 cessed: 03-23-2023.

841 Zhao, J., H. Feng, Q. Chen, and B. Garcia de Soto
842 (2022). Developing a conceptual framework for
843 the application of digital twin technologies to re-
844 vamp building operation and maintenance pro-
845 cesses. *Journal of Building Engineering* 49, 104028.

Automated Model Generation and Parameter Estimation of Building Energy Models Using an Ontology-based Framework

Abstract

This study presents a methodology for automated model generation and parameter estimation of building energy models using semantic modeling and Bayesian estimation. Semantic modeling techniques are used to represent the system components and their interactions, facilitating the automatic generation of a simulation model from dynamic component models. The proposed approach is applied to a case study of a ventilation system where a simulation model is generated, calibrated, and assessed through different performance metrics. These metrics demonstrate the accuracy and reliability of both model point estimates and probabilistic prediction intervals across all model outputs. Overall, the proposed methodology offers a systematic and automated approach to model development and calibration in building energy systems, with potential applications in building performance analysis, monitoring, and optimization.

Keywords: Digital Twin, Ontology, Data-driven, Building Energy Model, Building Performance Simulation

1. Introduction

The global transition towards future low-carbon energy systems calls for more energy-efficient and flexible buildings [1]. However, increasing the efficiency and flexibility brings additional complexity to both the design and operation of buildings. According to Hong et al. [2], the field of building performance simulation (BPS) will play a crucial role in this transition through services such as automated fault-detection, model-aided decision-making, and improved control of buildings. BPS represents the use of computational methods for modeling and simulating energy and indoor climate and has played an important role in the design of energy-efficient and sustainable buildings with decades of simulation program development, e.g. EnergyPlus, DOE-2, and TRNSYS [3, 4, 5].

Although the potential benefits of BPS-based services are well-established, broader adoption by the industry and practitioners is currently limited by several challenges, including, insufficient data availability and quality, time-consuming development, validation, and calibration of energy models, and missing interoperability between tools [2, 6]. This paper focuses on the operational stage and the automation of two tasks fundamental to the BPS modeling workflow: model development and model calibration.

1.1. Model development

Model development is the process of gathering and specifying the set of inputs required by BPS tools, including the specification of geometry, equipment design properties, and system topology. This information is often document-based and spread across diverse sources and formats that lack machine-readability, e.g. piping and instrumentation diagrams (P&ID), representing the system topology of buildings, or design specifications from equipment vendors. The process of gathering and

mapping this information is to a large degree carried out manually by a modeling expert, making it time-consuming, costly, and error-prone [7, 8].

Therefore, automation of model development for BPS has received significant attention in recent research projects such as Annex 60 [9] under the International Energy Agency (IEA), and Project 1 [10] under the International Building Performance Simulation Association (IBPSA). The work conducted in these projects mainly considers Building Information Models (BIM) as the primary information source for the automated generation of simulation models and two general approaches can be highlighted.

The first approach is partly automated and relies on mapping relevant information from the source to an intermediate data model. At this step, the user then manually inserts any additional information required by the simulation tool. As a one-time task, algorithmic mapping rules are then formulated to map from the data model to the simulation tool. The work of Cao et al. [11, 12] and Wimmer et al. [13, 14] uses IFC as the information source, SimModel [15] as the intermediate data model, and Modelica as the simulation environment.

In the second approach, information is mapped directly from the source to the simulation tool. For instance, Andriamamonjy et al. [16] developed a direct translation workflow from IFC to Modelica through a set of mapping rules. The IFC must follow a set of requirements on the format and structure. Compared with the semi-automated approach, the direct translation approach enables full automation of model development, but it comes with the cost of stricter requirements on the information source. This limits its usefulness in practical use cases where incomplete and inconsistent IFC files frequently occur [17]. In addition, it is unclear whether introducing strict requirements on the information source simply shifts manual labor from the energy modeler to the BIM practitioners with no net gain in

time spent.

More recent development in automation of model development and information mapping revolves around the use of ontologies, which represent a type of data model that describes generalized and structured relationships between concepts for specific domains. The core motivation behind the use of these structures is to improve interoperability [18]. This is crucial for highly complex and heterogeneous systems such as buildings, as also emphasized with the recent launch of the IEA Annex 91, dealing with openBIM and ontologies for building services [19]. Recent work by Roa et al. [20] demonstrated how the Brick ontology combined with the Control Description Language (CDL) can be used to streamline the implementation of advanced building control strategies. As part of a field demonstration, a hot water supply temperature setpoint reset strategy was implemented in a case study building.

Zheng et al. [21], presented a data framework based on the Brick ontology for BPS and BMS interoperability. The study considered Brick as a common data model between the BPS and BMS software. Both the BPS model development and mapping to Brick was performed in a semi-automatic fashion, relying on some level of manual work.

Wu et al. [22] presented a direct translation framework from the Brick and Bot ontologies to EnergyPlus. Considering a single floor of a university campus building, the study showed that the energy model development can be fully automated once the semantic model is obtained.

1.2. Model calibration

Even in the ideal case where simulation models can be reliably and effortlessly generated from available design data, additional work is still required to unlock operational services such as performance monitoring or predictive control. Most energy models developed from design conditions suffer from a so-called performance gap, i.e. a gap between predictions and actual measurements [2, 23]. Chaudhary et al. [24] report performance gaps varying from 23% to 97% in several studies. To close this gap, energy models are therefore calibrated, traditionally through an iterative trial-and-error process. Here, the model parameters are manually adjusted until the discrepancy between model predictions and actual measurements is acceptable. However, as already established, relying on manual labor is costly and requires domain-specific expertise [25, 26]. The recurrence of studies investigating automated approaches is therefore growing [27], with the majority using either optimization-based algorithms or Bayesian parameter estimation techniques [27, 28].

Among these studies, Chong et al. [29] proposed a method for automated and continuous calibration of BPS models using the Bayesian formulation proposed by Kennedy and O'Hagan (KOH) [30]. The method also partially automates model development by importing geometry and material information from BIM models following the Green Building eXtensible Markup Language (gbXML). However, specification of system topology and properties as well as the mapping between model outputs and measuring devices is performed manually.

1.3. Aim and contribution

Recent studies have focused on the automation of model development and the implementation of automated model calibration approaches mostly as two isolated processes. However, to calibrate a model, information about which model parameters and measured data to calibrate against is required. This information is not considered in current automation schemes and needs to be established in an automated, yet systematic and comprehensive way. In this work, we propose an ontology-based method that unifies automated model development and calibration. It is aimed at the operational stage of buildings combining design data from semantic models and operational data from measuring devices. For automated model generation, the concept of a signature pattern is introduced, which describes in the language of an ontology how and where a simulation model applies. By solving a graph search problem, these patterns are matched against a semantic model to generate a simulation model. With a standardized format of this simulation model, a general Bayesian parameter estimation problem is formulated, which enables probabilistic parameter estimation and simulation. Both the model generation and parameter estimation methodology are implemented and validated on an end-to-end case study of a ventilation system. Here, the SAREF [31] ontology is used in combination with the SAREF4BLDG [32] and SAREF4SYST [33] extensions to represent the topology and physical properties of the system.

The presented work extends previous efforts as part of a larger digital twin (DT) framework [anonymous] and is implemented as a reusable open-source Python library [anonymous]. Following the definition by Boje et al. [34] and Grieves [35], a DT refers to a concept with three main constituents; a physical system, a virtual system, and a flow of data linking these two systems. Therefore, the automated model generation and calibration method presented in this paper supports this concept by allowing the virtual system to automatically adapt and re-calibrate based on both structural and operational changes in the building.

The paper is structured as follows: in [section 2](#), an overview of the workflow and methodology is provided. In [section 3](#), a method is proposed to generate the simulation model by coupling FMU component models through a comparison of predefined signature patterns with information in the semantic model. Following, the problem formulation for Bayesian parameter estimation of the generated model is presented in [section 4](#). To validate the methodology, it is implemented on a case study of a ventilation system in [section 5](#). The results are presented and discussed in [section 6](#). Finally, a conclusion with the main findings and reflections is presented in [section 7](#).

2. General methodology and workflow

An overview of the overall methodology and information flow adopted in this study is presented in [Figure 1](#). Starting with automated model generation, the core component is a translator that reads information, applies mapping rules, and generates a simulation model. As shown, two information sources are used

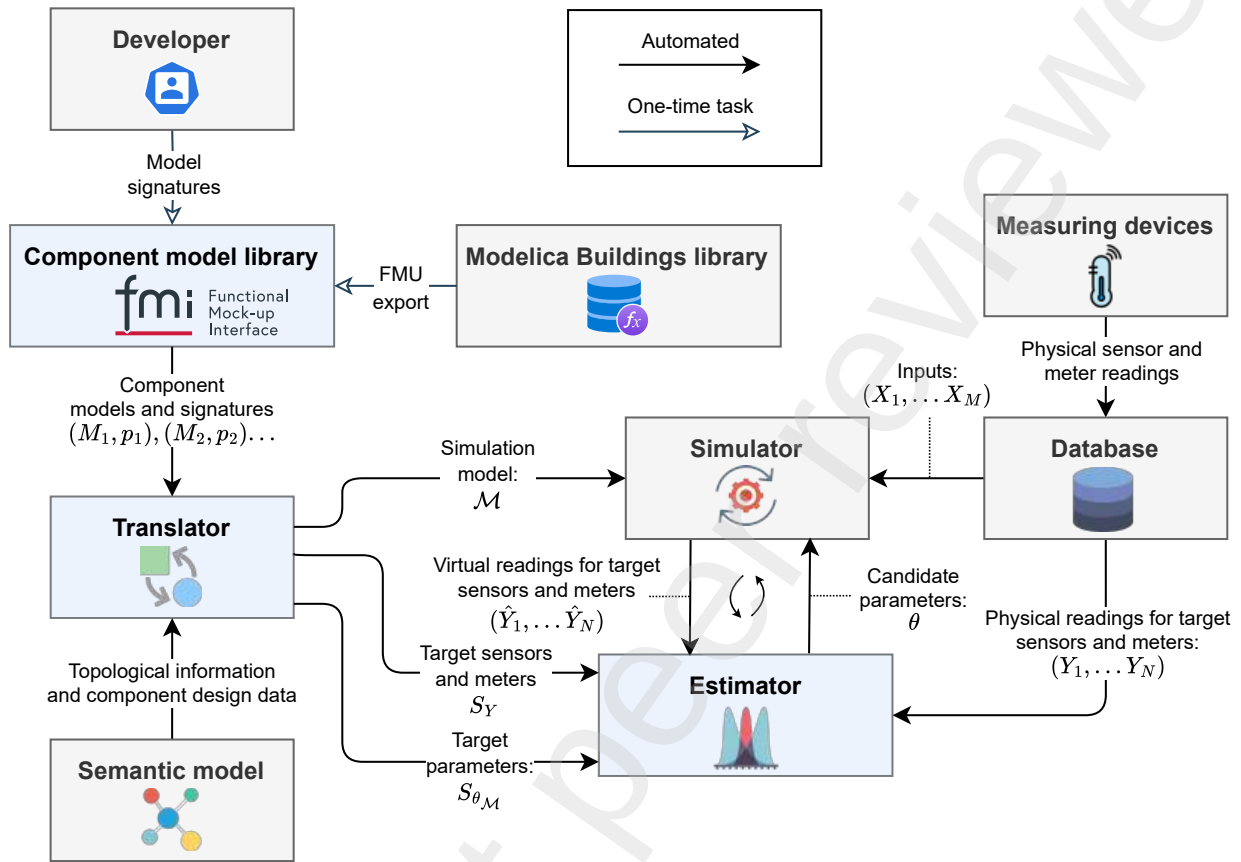


Figure 1: Flow diagram showing the information flow between high-level components. The main contributions of this paper are highlighted in light blue boxes.

for this process: 1) a semantic model with information about the real system such as sensor placement, system topology, and device properties, and 2) a component model library storing the simulation component models and signatures that describe how they can be connected to other component models and how they relate to the real system.

Post-translation, three main outputs emerge that are used in the calibration process: a simulation model, the target measuring devices, and the target parameters in the simulation model. Here, the target measuring devices represent the sensors and meters the simulation model is calibrated against, where the target parameters represent the set of model parameters considered in the parameter estimation problem. The specific notation used in the diagram is explained further in [section 3](#) and [section 4](#).

3. Ontology-driven model generation

In this section, the translation process is further described including, the semantic model, the component model library,

and the translator.

3.1. Semantic model

One of the information sources required by the translation method presented in this section is a semantic model with detailed information about the physical building, such as system topology and equipment specifications. Grounded in an underlying ontology, the format and structure of such a data model are organized using a construct called a *triple*, comprising subject, predicate, and object. In addition, types or classes are assigned to specify its meaning, e.g. $\text{BuildingSpace} \rightarrow \text{isSpaceOf} \rightarrow \text{Building}$ (using the SAREF4BLDG ontology). Triples can be combined in a graph-like structure such that the object of one triple might be the subject of another to form arbitrarily large semantic models of real systems in a rich and machine-readable format. Here, the detail and scope by which the semantic model can represent the real system is defined by the ontology. In this paper, the semantics are based on the SAREF ontology [31] with the SAREF4BLDG [32] and SAREF4SYST [33] extensions as introduced in [anonymous].

To properly represent the flow system topology and the placement of equipment, SAREF4SYST is extended with concepts from the Flow Systems Ontology (FSO) according to the alignments described by Kukkonen et al. [36].

An example of a semantic model is shown in Figure 2. Here, the nodes represent instances of classes, e.g. <flow temperature sensor 1> of class Sensor or <supply fan> of class Fan, where the brackets <> are used to distinguish instances from classes. In this example, the instances are interrelated through the six types of predicates as shown in the legend, e.g. observes which is used to describe the property observed by a sensor. The suppliesFluidTo predicate from the FSO ontology describes the physical placement of equipment in flow streams. The design properties of equipment, in this case of the nominal airflow of the <supply fan>, are described with the hasPropertyValue, isValueOfProperty, and hasValue predicates.

With a basis in this example, it is shown in the following sections how any semantic model can be searched for specific patterns for the automatic generation of a simulation model. Although the description of the developed methodology takes basis in specific ontologies, it should be noted that it generalizes to any ontology.

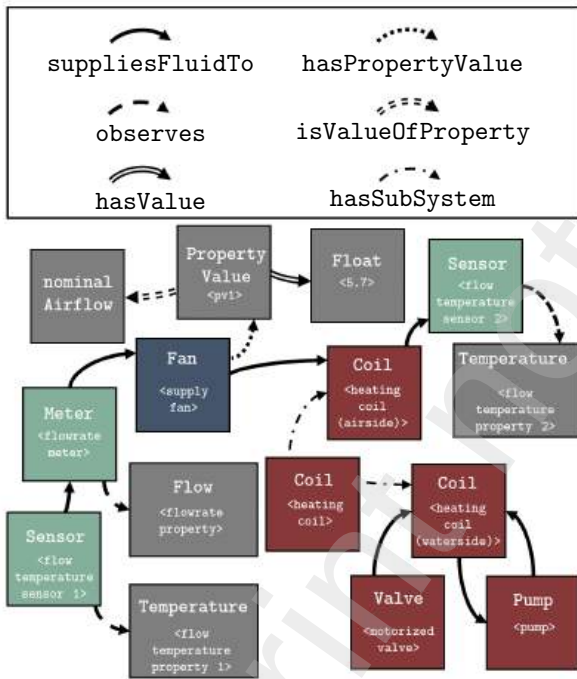


Figure 2: Example semantic model based on the SAREF, SAREF4BLDG, SAREF4SYST, and FSO ontologies. Inverse predicates have been omitted for clarity.

3.2. Component model library

For this purpose, a modular approach is used where the simulation model is assembled from a set of component models in different configurations to mimic the behavior of the real system. In this work, we consider the Functional Mockup Interface (FMI) standard [37] as it facilitates the reuse and sharing

of dynamic simulation models across a variety of simulation tools and platforms through Functional Mockup Units (FMU). An FMU is a self-contained dynamic simulation model with a set of inputs, outputs, and parameters. In this section, we use five the model classes M_1 - M_5 as placeholders for such dynamic models to explain the methodology. For the actual application in section 5, we use component and system models from the Modelica Buildings library [38].

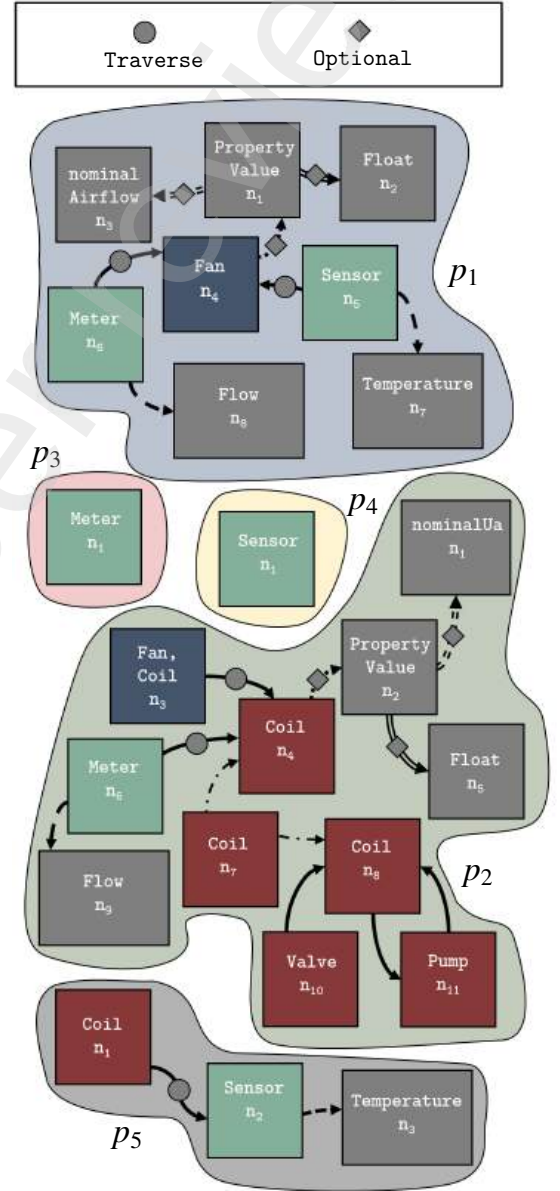


Figure 3: Semantic triples of the signature patterns p_1 - p_5 defined for the five example model classes.

To ensure the adaptability of the developed models and the automated reuse of the different component models in different applications, a *signature pattern* is defined for each model as a generalized pattern of subject, predicate, and object relations, akin to a blueprint outlining the semantic patterns where the models apply. This concept is illustrated in Figure 3 where examples of signature patterns p_1 - p_5 are shown for M_1 - M_5 .

Each pattern is specified using the triples available from the applied ontology with the nodes representing the ontology classes and the edges representing the predicate types. The nodes are marked with a local id, e.g. n_1 in p_1 , to distinguish nodes with identical classes. Each node can represent multiple classes to support different contexts where the models might apply, e.g. n_3 in p_2 , where instances of both class Fan and Coil are allowed. Taking p_1 as an example, we see that model M_1 applies if the instances of seven triples are present in the semantic model:

1. Sensor $\rightarrow$ observes $\rightarrow$ Temperature
2. Sensor $\rightarrow$ suppliesFluidTo $\rightarrow$ Fan
3. Meter $\rightarrow$ observes $\rightarrow$ FlowRate
4. Meter $\rightarrow$ suppliesFluidTo $\rightarrow$ Fan
5. Fan $\rightarrow$ hasPropertyValue $\rightarrow$ PropertyValue
6. PropertyValue $\rightarrow$ isValueOfProperty $\rightarrow$ nominalAirflow
7. PropertyValue $\rightarrow$ hasValue $\rightarrow$ Float

To increase the generality of the signature pattern, rules can be assigned to triples. For instance, as shown in Figure 3, the second and fourth triple in the list is augmented with the rule *Traverse*, which allows the semantic model to be traversed using the given predicate until an instance of the subject class is found or until the model cannot be traversed any further with the given predicate. This is useful for recognizing certain patterns in flow systems, e.g. to find the nearest temperature sensor or flow meter as shown in the example. In addition, triples can be augmented with the *Optional* rule, which makes the given triple optional, i.e. the pattern can still be matched even though the marked triple is not present in the semantic model. This is used when the mapped information is not strictly required but useful when available, e.g. equipment design parameters that can be estimated through measured data.

In addition to this specification of triples, input, output, and parameter mapping information is also required for each model as shown in Table 1. The first column specifies the nodes whose behavior is modeled.

Table 1: Example input, output, and parameter mapping of the signature patterns p_1 - p_5 .

	Nodes	Input		Output	Parameter	
	Modeled	Model $\leftrightarrow$ Source			Model $\leftrightarrow$ Source	
p_1	n_4	$T_{a,in}$	$n_5: y_{meas}$ $\dot{m}_a$ $n_6: y_{meas}$	$T_{a,out}$ P	$\bar{m}_a$	n_2
p_2	n_7 n_{10} n_{11}	$T_{a,in}$ $\dot{m}_a$	$n_3: (T_{a,out}, T_{a,out})$ $n_6: y_{meas}$	$T_{a,out}$ Q_{coil}	$\bar{U}A$	n_5
p_3	n_1			y_{meas}		
p_4	n_1			y_{meas}		
p_5	n_2	y_{meas}	$n_1: T_{a,out}$			

Again, taking p_1 as an example, we see that M_1 models the behavior of n_4 (a Fan) and requires two inputs. The first input, the inlet air temperature $T_{a,in}$, it receives from the output y_{meas} of the model representing the instance of n_5 while the

second input, the air massflow $\dot{m}_a$, is received from the output y_{meas} of the model representing the instance of n_6 . The model requires one parameter, the nominal air flow $\bar{m}_a$, which is mapped from n_2 (if available). Note that the unit assertions or translations of these parameter values could also be performed by extending the pattern signature with appropriate triples, e.g. PropertyValue $\rightarrow$ isMeasuredIn $\rightarrow$ UnitOfMeasure. Finally, the model has two outputs $T_{a,out}$ and P , representing the outlet air temperature and the power consumption of the fan, respectively.

3.3. Translator

By defining signature patterns for each model in the component model library, the semantic model can be searched systematically for matching patterns and the corresponding models can be employed in the correct context. For this task, it is useful to represent the semantic model and the signature patterns as a directed graph (V, E, L) with the set of nodes V , the set of edges E , and a labeling function L , mapping nodes or edges to a corresponding label or set of labels. Given this representation, the task of searching for signature patterns in the semantic model is directly related to the *subgraph isomorphism* problem. In the context of this paper, the subgraph isomorphism problem can be stated as follows [39]. Given the pattern signature represented by the graph $p = (V_p, E_p, L_p)$ and the semantic model represented by the graph $G = (V_g, E_g, L_g)$, find the map $f: V_p \rightarrow V_g$ such that $\forall u \in V_p, L_g(f(u)) \subseteq L_p(u)$ and $\forall (u, v) \in E_p$, and $L_p(u, v) = L_g(f(u), f(v))$ and $(f(u), f(v)) \subseteq E_g$. The condition $L_g(f(u)) \subseteq L_p(u)$ requires that the node label, i.e. the ontology class, of the semantic model is a subset of the pattern node label. Similarly, $L_p(u, v) = L_g(f(u), f(v))$ ensures that the edge label, i.e. the ontology predicate, of the semantic model matches the pattern edge label. Finally, $(f(u), f(v)) \subseteq E_g$ ensures that the mapped pattern edge $(f(u), f(v))$ also exists in the semantic model.

Various algorithms that solve this problem already exist. In this work, we have adapted the VF2 algorithm which applies a depth-first search that simultaneously traverses through the signature pattern and semantic model to add or eliminate nodes as matching candidates based on local information. The original algorithm is described in detail in [40].

Solving the subgraph isomorphism problem for the semantic model in Figure 2 and each of the patterns p_1 - p_5 in Figure 3 generates the maps f_1 - f_5 shown in Figure 4. As shown, all five signature patterns can be correctly matched once in the semantic model. For each match, the corresponding models are instantiated (referred to as m_1 - m_5) and connected using the generated maps and the information shown in Table 1.

For each model instance, the input connections are established by intersecting the identified instances between the 'Input, Source' column and the 'Nodes Modeled' column. For example, to identify the $T_{a,in}$ input of m_1 , we use the f_1 map to find that the source of this input is $f_1(n_5) = \langle \text{flow temperature sensor} \rangle$. Since $f_4(n_1) = \langle \text{flow temperature sensor} \rangle$ with n_1 being a modeled node, the model representing this instance is m_4 . Altogether, this means that the output y_{meas} of m_4 maps to the input $T_{a,in}$ of m_1 . Completing this process for all

Match of signature p_4

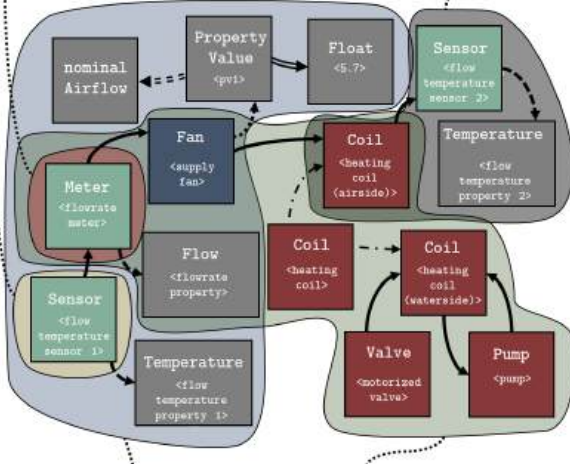
$$f_4 = \{(n_1, \langle \text{flow temperature sensor} \rangle)\}$$

Match of signature p_3

$$f_3 = \{(n_1, \langle \text{flowrate meter} \rangle)\}$$

Match of signature p_5

$$f_5 = \{(n_1, \langle \text{heating coil (airside)} \rangle), \\ (n_2, \langle \text{flow temperature sensor 2} \rangle), \\ (n_3, \langle \text{flow temperature property 2} \rangle)\}$$



Match of signature p_2

$$f_2 = \{(n_3, \langle \text{supply fan} \rangle), \\ (n_4, \langle \text{heating coil (airside)} \rangle), \\ (n_6, \langle \text{flowrate meter} \rangle), \\ (n_7, \langle \text{heating coil} \rangle), \\ (n_8, \langle \text{heating coil (waterside)} \rangle), \\ (n_9, \langle \text{flowrate property} \rangle), \\ (n_{10}, \langle \text{motorized valve} \rangle), \\ (n_{11}, \langle \text{pump} \rangle)\}$$

Match of signature p_1

$$f_1 = \{(n_1, \langle \text{pv1} \rangle), \\ (n_2, \langle 5.7 \rangle), \\ (n_3, \langle \text{nominalAirflow} \rangle), \\ (n_4, \langle \text{supply fan} \rangle), \\ (n_{13}, \langle \text{flow temperature sensor} \rangle), \\ (n_{17}, \langle \text{flowrate meter} \rangle), \\ (n_{16}, \langle \text{flow temperature property} \rangle), \\ (n_{18}, \langle \text{flowrate property} \rangle)\}$$

Figure 4: Matching of patterns in the semantic model to generate maps between signature nodes and semantic model instances.

inputs and model instances generates the model $\mathcal{M}$ shown in Figure 5

3.4. Identifying target measuring devices and parameters

After generating a simulation model $\mathcal{M}$, the required information for the parameter estimation stage is gathered. A vital step for both parameter estimation, as well as downstream BPS applications such as performance monitoring, is the comparison of simulated versus measured values. As introduced in earlier work [41], a key concept of the used modeling framework is to incorporate measuring devices into the simulation model on equal terms as other models, and use these as an interface for straightforward comparison between simulated and measured

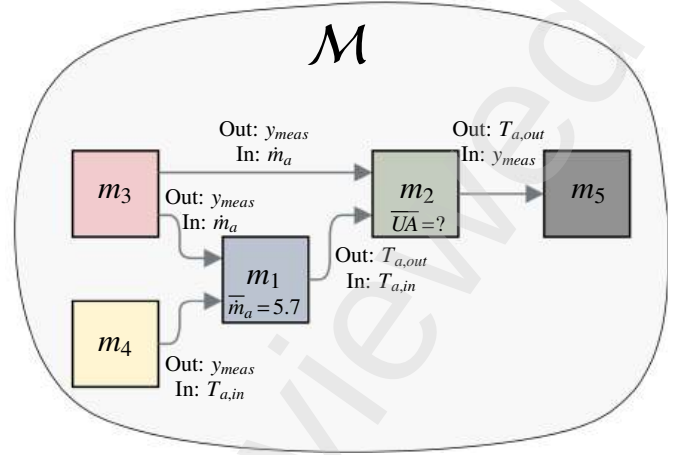


Figure 5: Resulting simulation model after matching and connecting the component models based on their signature. The boxes represent the component models while the labeled edges represent the inputs and outputs of models.

values. Using the notation from [41], the set of measuring devices used for parameter estimation S_Y is given by Equation 1

$$S_Y = \{s \in H_y : |s.\text{connectsAt}| = 1\} \quad (1)$$

In Equation 1, H_y is the set of component models in $\mathcal{M}$ that models nodes with a measuring device class. In the SAREF ontology for instance, these classes would be saref:Sensor or saref:Meter while in the Brick ontology they would be brick:Sensor. The $|s.\text{connectsAt}| = 1$ condition filters for models with 1 input. From the generated maps f_1 - f_5 , only three contain measuring devices that are also modeled:

- $f_3(n_1) = \langle \text{flowrate meter} \rangle$
- $f_4(n_1) = \langle \text{flow temperature sensor 1} \rangle$
- $f_5(n_2) = \langle \text{flow temperature sensor 2} \rangle$

This gives $H_y = \{m_3, m_4, m_5\}$ and $S_Y = \{m_5\}$ as m_3 and m_4 acts as sources with 0 inputs. The model m_5 thus acts as a virtual sensor and is used as a calibration target for the example model, as its readings can be directly compared to the actual sensor represented by $\langle \text{flow temperature sensor 2} \rangle$.

In addition to identifying the target measuring devices, the target set of parameters S_{θ_M} must also be identified. This set is defined in Equation 2

$$S_{\theta_M} = S_{\theta} \setminus (S_{\theta} \cup S_{\theta, \text{mapped}}) \quad (2)$$

Here, S_{θ} is the total set of parameters in $\mathcal{M}$, i.e. the entries in the 'Parameter, Model' column in Table 1. $S_{\theta, \text{mapped}}$ is the set of parameters in the simulation model that are successfully mapped from the semantic model with an entry in the generated map. Hence, for the example model $\mathcal{M}$ in Figure 5, the target parameter set becomes $S_{\theta_M} = \{UA\}$ as $S_{\theta} = \{\dot{m}_a, UA\}$ and $S_{\theta, \text{mapped}} = \{\dot{m}_a\}$ with $\dot{m}_a = 5.7$.

4. Bayesian parameter estimation

Given the simulation model $\mathcal{M}$, the calibration targets S_Y , and the target parameters of the model S_{θ_M} , parameter estimation can now be performed. A wide variety of methods exist for this purpose, e.g. gradient-based algorithms or heuristic-based methods such as genetic algorithms. These optimization-based algorithms typically provide a point estimate of the best-fitting model parameters given an objective function. In contrast, in a Bayesian framework, model parameters are treated as random variables and the outcome of an estimation process is probability distributions for each parameter in the model. This fundamental difference provides several advantages in favor of Bayesian estimation. 1) It explicitly handles problems where fundamentally different parameter sets reproduce the observed data of the modeled system equally well, addressing the identifiability problem discussed by Reddy [23]. 2) Using the obtained parameter distributions, uncertainty can be propagated through the model to give uncertainty estimates on predictions.

This quantification of uncertainty is especially valuable in the context of automation, where the calibration process is not supervised by an expert modeler and if the model is used as part of a decision-making process [27]. In the following section, the Bayesian estimation problem is introduced along with the sampling technique used to solve this problem.

4.1. Problem formulation

Consider a dynamic simulation model $\mathcal{M}$ parameterized by the vector of parameters $\theta_M \in \mathbb{R}^P$ with the time series inputs $X = (X_1, \dots, X_M)$ with $X_i \in \mathbb{R}^T \forall i \in 1..M$, outputs $\hat{Y} = (\hat{Y}_1, \dots, \hat{Y}_N)$ with $\hat{Y}_i \in \mathbb{R}^T \forall i \in 1..N$, and timestamps $t \in \mathbb{R}^T$. Here, $P = |S_{\theta}|$, $M, N = |S_Y|$, and T represents number of parameters, inputs, outputs, and timesteps, respectively. The relationship between inputs and outputs is then given by Equation 3.

$$\hat{Y} = \mathcal{M}(X, t, \theta_M) \quad (3)$$

Using KOH's Bayesian model formulation [30], the relationship between observations Y_y , model response $\hat{Y}_y$, and discrepancy $r_y(\cdot)$ is given by Equation 4.

$$Y_y = \hat{Y}_y + r_y(\cdot) \quad \forall y \in 1..N \quad (4)$$

Here, the purpose of the discrepancy term $r_y(\cdot)$ is to account for structural inadequacies in the model and is modeled as a Gaussian process with a mean function of 0 and covariance function $k_y(x_i, x_j)$ as given in Equation 5.

$$r_y(x_i, x_j) \sim \mathcal{GP}(0, k_y(x_i, x_j)) \quad \forall y \in 1..N \quad (5)$$

In this work, we model the covariance function using the Matern 3/2 kernel as shown in Equation 6.

$$k_y(x_i, x_j) = a_y \left(1 + \sqrt{3}d_y(x_i, x_j) \right) \exp \left(-\sqrt{3}d_y(x_i, x_j) \right) \quad \forall y \in 1..N \quad (6)$$

Where a_y is a parameter controlling the variance of the kernel function. Using this kernel, the covariance decays with the

distance $d_y(x_i, x_j)$, which is computed between two points x_i and x_j using Equation 7.

$$d_y(x_i, x_j) = (x_i - x_j)C_y^{-1}(x_i - x_j) \quad \forall i, j \in 1..T \quad (7)$$

The metric C_y is populated with the hyper parameters $I_y = [I_{y,1}, \dots, I_{y,|x_i|}]^T$ on the diagonal, determining the scale of the inputs as shown in Equation 8.

$$C_y = \text{diag}(I_y) \quad (8)$$

All the gaussian process hyperparameters are then given by the vector $\theta_{\mathcal{GP}}$, as shown in Equation 9.

$$\theta_{\mathcal{GP}} = [a_1, I_1^T, \dots, a_N, I_N^T]^T \quad (9)$$

Following, the full vector of parameters θ is then given by Equation 10.

$$\theta = [\theta_M^T, \theta_{\mathcal{GP}}^T]^T \quad (10)$$

Having defined the independent gaussian processes for each output y , the likelihood of making all observations $(Y_1, \dots, Y_N)$, given the parameters θ , is given by Equation 11.

$$\ln \mathcal{L}(Y_1, \dots, Y_N | \theta) = \sum_{y=1}^N -\frac{1}{2} \left(r_y^T K_y^{-1} r_y + \ln \det K_y + T \ln 2\pi \right) \quad (11)$$

Where r_y is the residual between observations and model response, as given in Equation 12.

$$r_y = Y_y - \hat{Y}_y \quad \forall y \in 1..N \quad (12)$$

Each element of row i and column j of the covariance matrix K_y is computed using the kernel function and is given by Equation 13.

$$K_{y,i,j} = k_y(x_{y,i}, x_{y,j}) + \sigma_y^2 \delta_{ij} \quad \forall i, j \in 1..T \quad \forall y \in 1..N \quad (13)$$

Where the points $x_{y,i}$ and $x_{y,j}$ are defined as the i^{th} and j^{th} sample from a subset B_y of the model inputs, i.e. $x_{y,i} = [X_{k,i}, \dots]^T_{k \in B_y}$ with $B_y \subseteq S_X$. The set B_y is used as a filter to make sure that only relevant model inputs are used in the Gaussian process for model output y . Measurement error is modeled as i.i.d. white noise by adding variance σ_y^2 at the diagonal through the Kronecker delta function δ_{ij} which is 1 for $i = j$ and otherwise 0.

As earlier stated, the goal of Bayesian estimation is to estimate the distributions of a set of model parameters, given some observed data of the modeled system. Following Bayes' rule, this distribution is given by Equation 14.

$$\pi(\theta | Y) = \frac{\mathcal{L}(Y | \theta) \pi_0(\theta)}{\int_{\mathbb{R}^P} \mathcal{L}(Y | \theta) \pi_0(\theta) d\theta} \quad (14)$$

Where $\pi(\theta | Y_1, \dots, Y_N)$ is the posterior density of observing the parameters θ given the observations $Y_1, \dots, Y_N$. The integral in the denominator, which evaluates to a constant, ensures that the

posterior integrates to 1. $\pi_0(\theta)$ is the prior density of observing θ . This distribution encodes information about the parameters before observing $Y_1, \dots, Y_N$, typically in the form of domain-specific knowledge. For the application shown in this work, we assume uniform distributions defined through upper and lower bounds on the estimated parameters. Uniform priors are also known as uninformative priors because they do not encode any information about the parameters other than the defined bounds. Another commonly used prior is the normal distribution. However, defining a meaningful prior normal distribution for each parameter would require case-specific expert knowledge of the modeled system, which defeats the purpose of the applied automation methodology.

4.2. Markov Chain Monte Carlo sampling

Although both the likelihood $\mathcal{L}(Y|\theta)$ and the prior $\pi_0(\theta)$ in Equation 14 are known, in most cases, obtaining a closed-form solution or performing any form of expectation calculations using traditional numerical integration methods is infeasible [42]. Instead, the numerical sampling method Markov Chain Monte Carlo (MCMC) is frequently used for drawing samples from the posterior using different variants, e.g. Metropolis-Hastings [43, 44] and the Affine Invariant MCMC Ensemble sampler [45].

4.3. Parallel Tempering Ensemble MCMC

In an automated model generation and calibration framework, where important characteristics of the generated model and estimation problem vary (e.g. number of target measuring devices and target parameters), a robust MCMC sampler that can deal with a wide variety of likelihood functions is important. Here, traditional MCMC schemes are often ineffective as they can get stuck at local high-probability regions of the parameter space, which prevents the algorithm from converging to the posterior in a reasonable amount of time [46]. In this work, a specific version of MCMC called Parallel tempering MCMC (PTMCMC) is therefore used. PTMCMC makes use of multiple chains that are run in parallel to sample from tempered versions of the posterior $\pi_T(\theta|Y)$, which is given in Equation 15

$$\pi_T(\theta|Y_1, \dots, Y_N) \propto \mathcal{L}(Y_1, \dots, Y_N|\theta)^{1/T} \pi_0(\theta) \quad (15)$$

The exponent in Equation 15 flattens the likelihood function to make the parameter space easier to explore. When $T \rightarrow \infty$, the posterior becomes the prior, which in this work is a uniform distribution that can easily be explored. During runtime, samples are swapped periodically between chains of different temperatures to allow solutions to propagate from hot chains to colder chains. Further details on the PTMCMC algorithm are provided by [45, 46].

4.4. Posterior predictive distribution and inference

As a key property of all MCMC variants, the distribution of the drawn samples converges to the desired posterior distribution in the limit as the chain progresses and the number of samples increases. The first part of the chain where convergence has not been reached is known as the burn-in phase and is discarded.

The sampler convergence, i.e. the number of burn-in steps, can be determined by computing the integrated autocorrelation time (IAT) and comparing it with the total number of steps [42]. The remaining k samples drawn from the posterior distribution, i.e. $\Theta_k = \{\theta_1, \theta_2, \dots, \theta_k\}$ where $\theta_p \sim \pi(\theta|Y_1, \dots, Y_N) \forall_{p \in 1..k}$ can then be used for various applications. In the context of BPS, the straightforward application is to produce the *posterior predictive distribution* which represent probabilistic predictions for future or unseen conditions of the modeled system. Given the parameter sample $\theta \in \Theta_k$, the training data (X, Y) and test data (X^*, Y^*) , the predictive distribution of model discrepancy r_y^* is first obtained by conditioning on the observations as given in Equation 16 [47].

$$r_y^*|\theta, X, r_y, X^* \sim \mathcal{N}(\bar{r}_y^*, \text{cov}(r_y^*)) \quad (16)$$

Where $\bar{r}_y^*$ is the mean vector given by Equation 17 and $\text{cov}(r_y^*)$ is the covariance matrix given by Equation 18.

$$\bar{r}_y^* = K_y^*(K_y)^{-1} r_y \quad (17)$$

$$\text{cov}(r_y^*) = K_y^{**} - K_y^*(K_y)^{-1} (K_y^*)^T \quad (18)$$

The elements of matrix $K_{y,ij}^*$ is calculated as the covariance between the test and training inputs as shown in Equation 19 while the elements of matrix $K_{y,ij}^{**}$ is calculated as the covariance between the test inputs as shown in Equation 20.

$$K_{y,ij}^* = k_y(x_{y,i}^*, x_{y,j}) \quad \forall_{i \in 1..T^*} \quad \forall_{j \in 1..T} \quad \forall_{y \in 1..N} \quad (19)$$

$$K_{y,ij}^{**} = k_y(x_{y,i}^*, x_{y,j}^*) \quad \forall_{i,j \in 1..T^*} \quad \forall_{y \in 1..N} \quad (20)$$

Finally, the posterior predictive distribution Y_y^p for one parameter sample is given by the sum of the model response $\hat{Y}_y^*$ as shown in Equation 21 and the predictive distribution for the discrepancy $r_y^*|\theta, X, r_y, X^*$ as shown in Equation 22

$$\hat{Y}_y^* = \mathcal{M}(X^*, t^*, \theta_M) \quad (21)$$

$$Y_y^p = \hat{Y}_y^* + r_y^*|\theta, X, r_y, X^* \quad \forall_{y \in 1..N} \quad (22)$$

5. Case study and implementation

To validate the proposed methodology, we apply it to a case study. The presented methodology has been implemented as part of a reusable Python library [anonymous]. The considered case study is a section of a ventilation system in a university building that supplies around 20 zones with air and has a capacity of 34.000 m³/h. A diagram is shown in Figure 6 with the installed measuring devices and equipment.

Four components from the supply side are modeled: a fan, a heating coil, a controller, and a valve. The system employs demand-controlled ventilation (DCV) and the fan is operated at variable speed to maintain a constant supply air pressure. The heating coil is responsible for heating the supply air to a desired

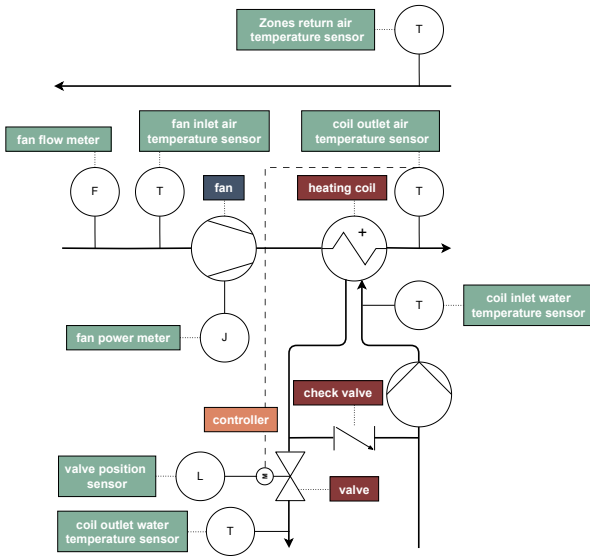


Figure 6: Diagram of the ventilation case study system.

addition, the hot water loop configuration for the case study system is fairly common which further justifies the development of a specialized model.

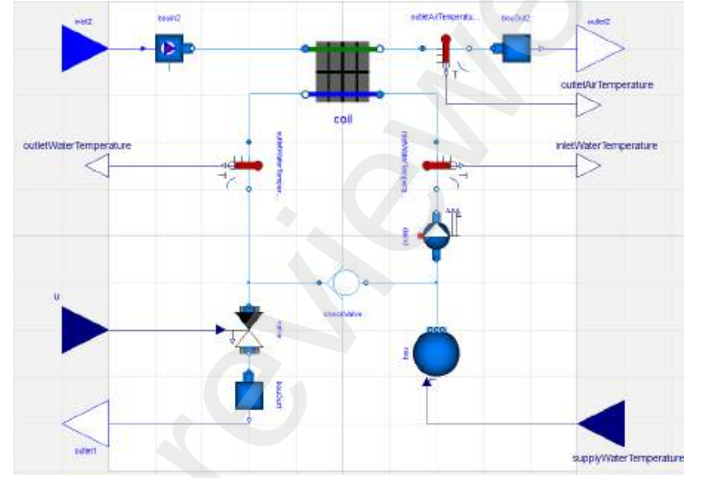


Figure 8: Modelica diagram of the exported hot water loop model.

setpoint temperature, which is calculated using a piece-wise linear function of the measured return air temperature. This is achieved by actuating the position of a motorized valve to regulate the waterflow. The required control signal for the valve position is determined by a PI controller.

The semantic model used to represent the case study system is shown in Figure 7. Similar to the example presented in section 3, the SAREF, SAREF4BLDG, SAREF4SYST, and FSO ontologies are used. As recommended by the FSO ontology [36], the <heating coil> is split into an airside and water-side to avoid ambiguous configurations where it is impossible to determine whether components placed before and after the coil are on the airside or waterside.

As shown in Figure 7, the `suppliesFluidTo` predicate is used to encode that the <fan> is placed before the <heating coil (airside)> in the flow path. Many of the shown predicates have an inverse, e.g. `hasFluidSuppliedBy` pointing from the <heating coil (airside)> to the <fan>.

The use of these predicates also enables the representation of measuring devices as part of the flow stream, e.g. the <fan flow meter>, which is placed between the <fan inlet air temperature sensor> and the <fan>. Measuring devices not physically part of the flow stream can also be related to a specific device through the `saref:observes` and `saref:isPropertyOf` as is the case for the <fan power meter>.

5.1. Hot water loop model

The hot water loop containing the heating coil, motorized valve, check valve, and pump is modeled jointly in one FMU as shown in Figure 8. Although it is generally preferable to export each component model separately to maximize reusability, a more specialized component is preferred in this case to reduce complexity and increase efficiency of the generated model. In

For the heating coil, the `Buildings.Fluid.-HeatExchangers.DryCoilDiscretized` model is used, which models the heating coil as a dynamic heat exchanger. The considered parameters for this model are the nominal heat transfer coefficient UA , the time constants of the waterside, airside, and metal τ_w , τ_a , τ_m , as well as the water and air nominal mass flows $\dot{m}_{c,w}$, $\dot{m}_{c,a}$. Assuming a linear valve characteristic, the motorized valve is modeled using `Buildings.Fluid.-Actuators.Valves.TwoWayLinear`, which is parameterized by the flow coefficient K_v , the nominal water flowrate $\dot{m}_v$, and a fixed resistance ΔP_{res} , that represents all system pressure losses after the motorized valve. The check valve is modeled with `Buildings.Fluid.FixedResistances.CheckValve` and is also parameterized by a flow coefficient K_{cv} and the nominal water flowrate $\dot{m}_{cv}$. The pump is assumed to provide a constant pressure rise ΔP_p and the pressure drop between the source and sink is assumed to be constant and is given by the parameter ΔP_s . The signature pattern of the model is given by Table 2 and Figure 9.

5.2. Controller

The controller is modeled using `Buildings.Controls.-Continuous.LimPID`, configured as a reverse-acting PI controller parameterized by the proportional gain K_P and the integral time constant T_I . The Modelica diagram is shown in Figure 10. Compared to the hot water loop model, the controller model applies for PI controllers in general, as also indicated by the signature pattern as given in Table 3 and Figure 11.

5.3. Fan

The fan component is modeled using the polynomial power curve used in EnergyPlus [48]. This model is parameterized by the nominal power consumption P_f and the nominal air mass-flow $\dot{m}_{f,a}$ along with the power coefficients c_1 - c_4 .

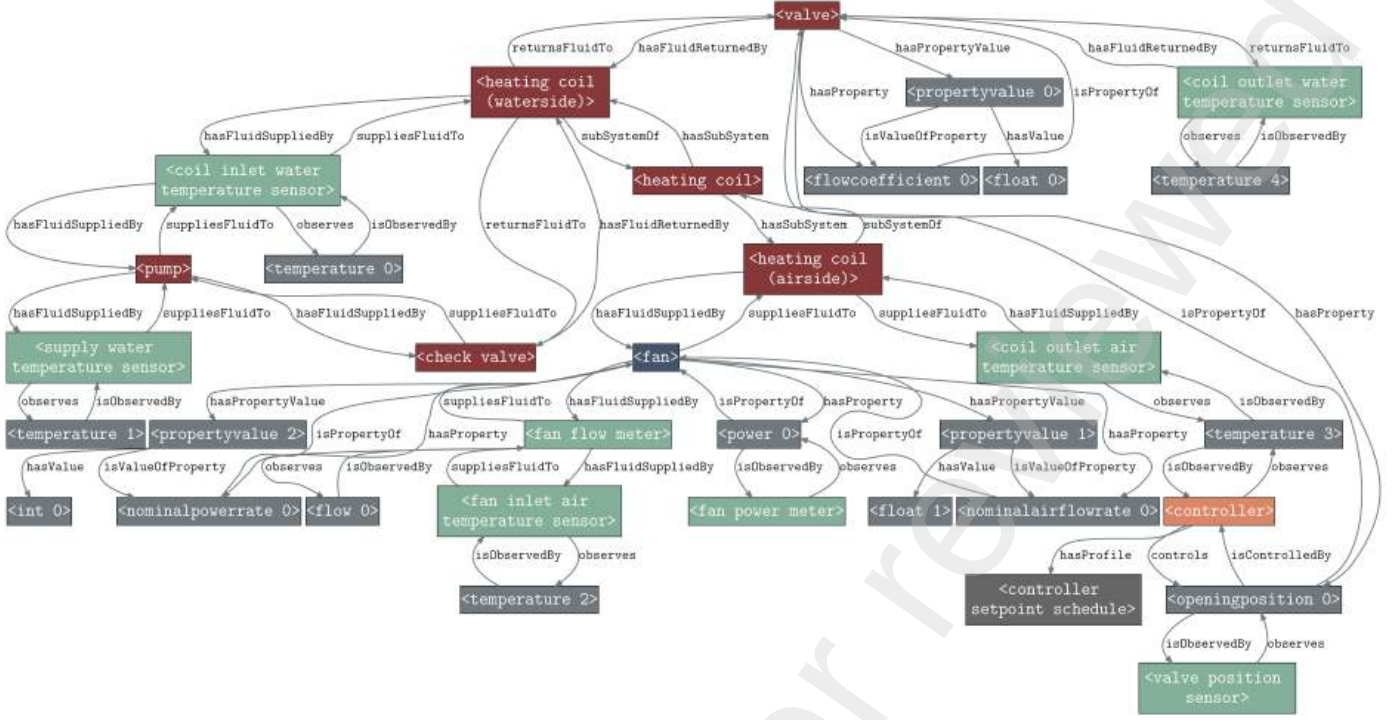


Figure 7: Semantic model of the case study system.

Table 2: Mapping of inputs, outputs, and parameters of the hot water loop signature pattern.

Nodes	Input	Output	Parameter
Modeled	Model $\longleftrightarrow$ Source		Model $\longleftrightarrow$ Source
n ₂	$T_{c,a,in}$	n ₁₆ : ($T_{f,out}, T_{c,a,out}$)	$\overline{UA}$ n ₁₈
n ₆	$T_{w,sup}$	n ₁₅ : y_{meas}	τ_w —
n ₁₁	$\dot{m}_{c,a}$	n ₁ : $\dot{m}_a$	τ_m —
n ₁₂	u_v	n ₁₄ : u	τ_a —
			$\overline{m}_{c,w}$ —
			$\overline{m}_{c,a}$ —
			ΔP_c —
			K_v n ₂₄
			$\overline{m}_v$ —
			K_{cv} n ₂₁
			$\overline{m}_{cv}$ —
			ΔP_p —
			ΔP_s —
			ΔP_{res} —

Table 3: Mapping of inputs, outputs, and parameters of the controller signature pattern.

Nodes	Input	Output	Parameter
Modeled	Model $\longleftrightarrow$ Source		Model $\longleftrightarrow$ Source
n ₁	y_{set}	n ₄ : y_{sched}	K_P —
	y_{meas}	n ₂ : y_{meas}	K_I —

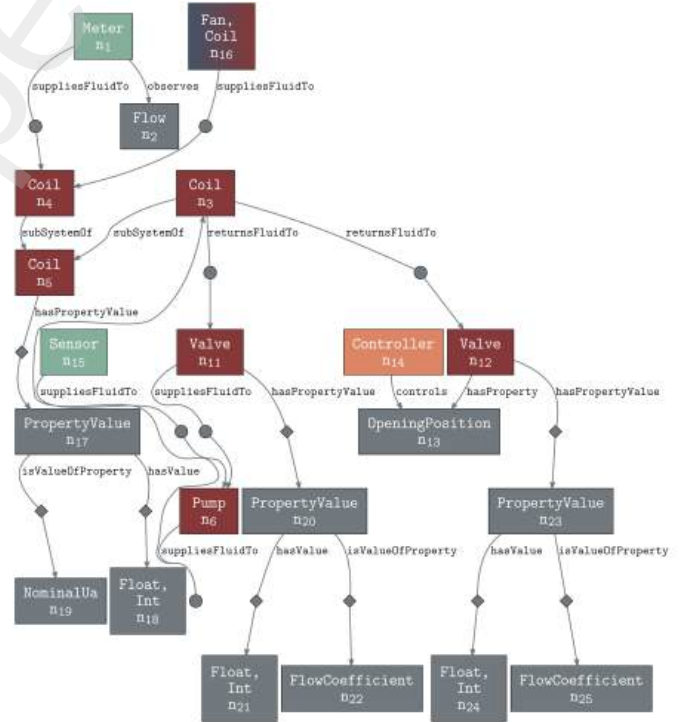


Figure 9: Semantic triples of the hot water loop signature pattern.

shown in Equation 23.

$$\dot{Q}_{toAir} = f_{tot} P_f \quad (23)$$

The signature pattern of the model is given by Table 4 and Figure 13. As shown in Table 4, 2 of the 7 model parameters

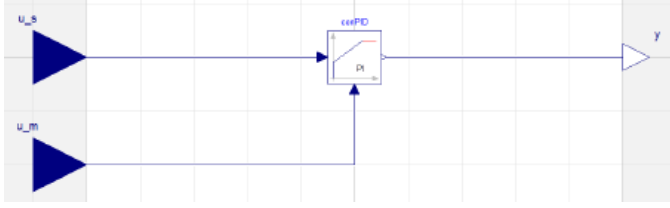


Figure 10: Modelica diagram of the exported controller model.

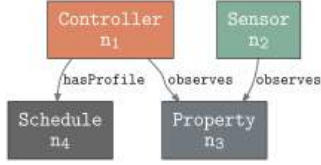


Figure 11: Semantic triples of the controller signature pattern.

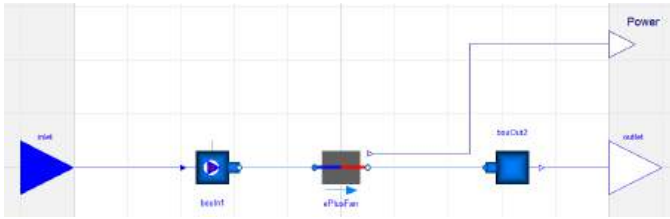


Figure 12: Modelica diagram of the exported fan model.

Table 4: Mapping of inputs, outputs, and parameters of the fan signature pattern.

Nodes	Input	Output	Parameter
Modeled	Model $\longleftrightarrow$ Source	Model $\longleftrightarrow$ Source	Model $\longleftrightarrow$ Source
n_3	$\dot{m}_{f,a}$ $T_{f,in}$	$n_1: \dot{m}_a$ $n_2: y_{meas}$ P_f $T_{f,out}$	$\bar{P}_f$ $\bar{m}_{f,a}$ c_1 c_2 c_3 c_4 f_{tot}
			n_5 n_8 — — — — —

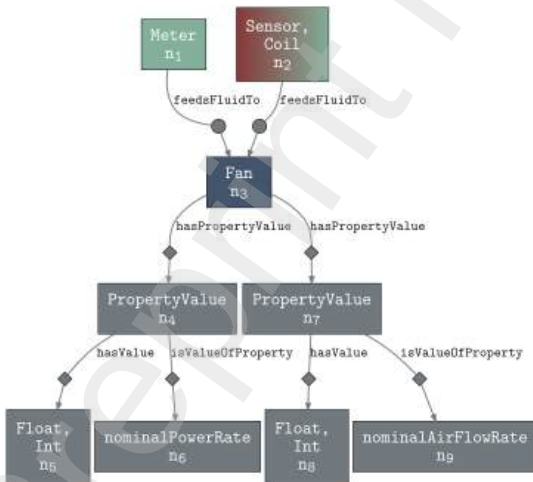


Figure 13: Semantic triples of the fan signature pattern.

can be mapped from the semantic model.

6. Results and Discussion

In this section, we present and discuss the case study implementation results, including the output of the model generation stage, and the parameter estimation process.

6.1. Model generation

Using the proposed model generation method with the semantic model of the case study system and the signature patterns defined throughout section 5, a simulation model is generated as shown in Figure 14. The three signature patterns defined for the fan, controller, and heating coil water loop models have all been identified in the semantic model with the models and connections instantiated and established accordingly. The instantiated heating coil water loop model represents four instances from the semantic model: <heating coil>, <check valve>, <valve>, and <pump> as dictated by Table 2. The remaining models represent the instances from the semantic model in a one-to-one fashion. Note that Figure 14 contains a feedback loop caused by controller model, meaning that both estimation and inference is performed in a closed-loop configuration.

The <fan flow meter>, <controller setpoint schedule>, <fan inlet air temperature sensor>, and <supply water temperature sensor> constitute the system boundary and provide actual measured data as input, while the remaining 5 measuring devices are identified as the target set S_Y considered for estimation. As part of the translation process, 3 parameter values are automatically mapped from the semantic model to the simulation model, i.e. $S_{\theta, \text{mapped}} = \{\bar{m}_{f,a}, \bar{P}_f, K_v\}$. Hence, with a total set S_{θ} of 24 parameters, the target set of parameters S_{θ_M} contains 21 of the listed parameters from tables 2-4.

6.2. MCMC parameter sampling

For parameter estimation, sensor and meter data for a one-day winter period is used to sample from the posterior distribution using the algorithm described in subsection 4.3. The resulting traceplots are shown in Figure 15, where one sample represents the vector of parameters θ . Only the model parameters θ_M are shown for readability, although similar trace plots can be shown for the Gaussian process hyper-parameters θ_{GP} . The first axis is the sampling step, while the second axis is the sampled values of the parameters. To determine sampler convergence, the IAT-based criteria proposed by Foreman-Mackey et al. [45] is used. In this work, this criteria is considered met when the number of steps is 20 times IAT. In Figure 15, this is indicated by the intersection of the IAT in red and the black dashed line, occurring around step 2000. The samples after this intersection represent the posterior distribution of interest while samples before are discarded as burn-in. In qualitative terms, convergence is identified by the transition to a scattered distribution with no systematic trends. After convergence, the chain is run for an additional 4000 steps to produce a population of samples from the posterior distribution.

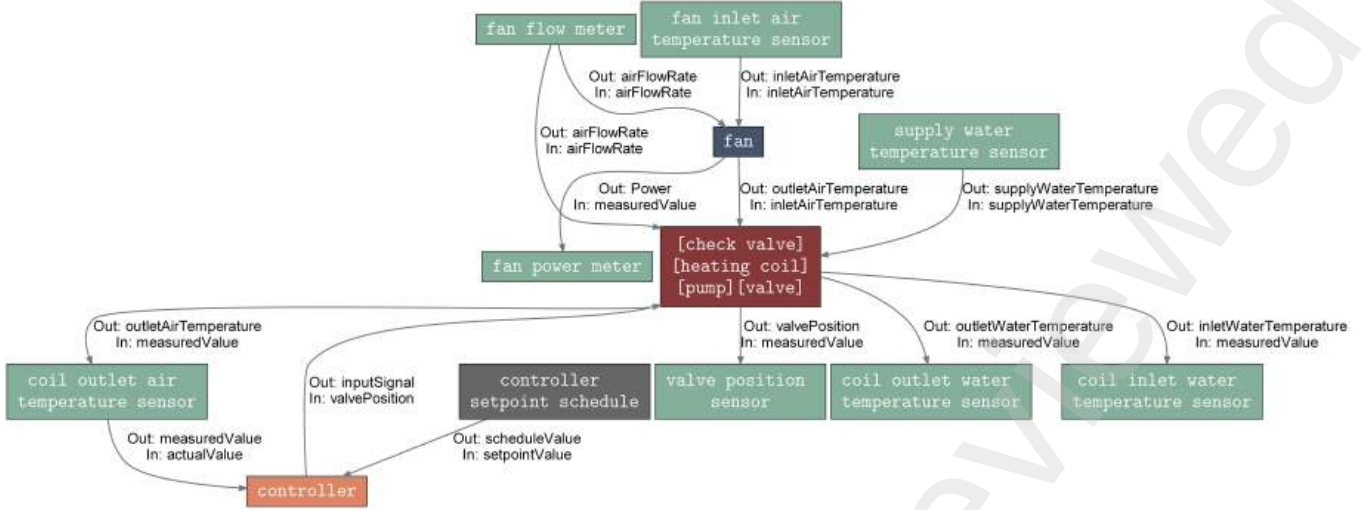


Figure 14: Simulation model of the case study system generated from a semantic model and signatures using the methodology described in [section 3](#). The boxes represent the component models while the labeled edges represent the inputs and outputs of models.

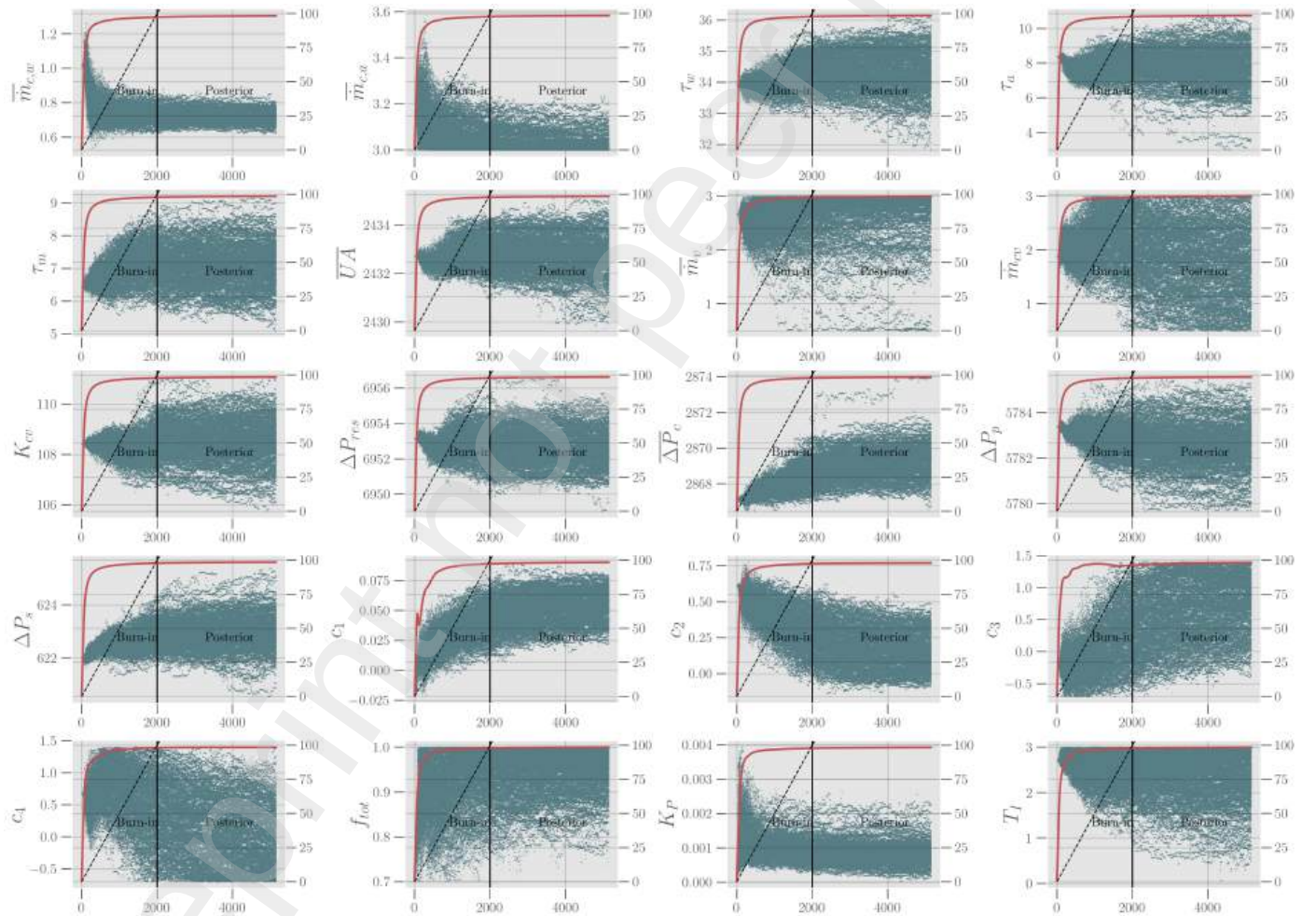


Figure 15: Trace plots of the sampled parameters θ_M . The red line represents the IAT while the black dashed line represent the convergence criteria.

6.3. Posterior distribution

As a general tool to visualize and summarize the sampled posterior, a corner plot is often used. [Figure 16](#) shows the cor-

ner plot of the resulting samples from the posterior distribution. The pairwise correlations are shown between the parameters as a 21x21 matrix with the marginalized distributions of

the parameters along the diagonal represented by a histogram. Again, only the model parameters θ_M are shown for readability, although a full 49x49 matrix plot can be generated, which includes the Gaussian process hyperparameters θ_{gp} . The quantile statistics are shown at the top of each histogram in the format $\theta = Q_{50\%}^{+(Q_{84\%}-Q_{50\%})} - (Q_{50\%}-Q_{16\%})$, corresponding to $\pm 1\sigma$ for a normal distribution. For instance, the nominal water massflow of the heating coil $\bar{m}_{c,w}$ has a median value of 0.72 kg/s across all samples with an estimated upper bound of 0.76 kg/s and lower bound 0.69 kg/s. The corner plot can also be used as a diagnostic tool to identify unidentifiable or correlated parameters. For instance, the power coefficient pairs c_1, c_2 and c_2, c_3 and c_3, c_4 , and c_1, c_4 are inversely correlated. This implies that lowering e.g. c_1 can be compensated by increasing c_2 . The opposite is true for the pairs c_1, c_3 and c_2, c_4 which are positively correlated.

6.4. Inference and posterior predictive distribution

To measure the performance of the calibrated model and the validity of the sampled posterior distribution, the posterior predictive distribution is calculated across a week of test data. As discussed earlier, one of the advantages of Bayesian parameter estimation is the inherent capability to provide uncertainty estimates or prediction intervals (PI) on predictions. However, to ensure that these estimates are reliable, the average coverage error (ACE) is used. The ACE metric measures the reliability of probabilistic models by calculating the difference between the expected and actual percentage of observations within a given PI and should be as close to 0 as possible [49]. In addition, the root mean square error (RMSE) and mean absolute error (MAE) of the median prediction $Q_{50\%}(Y_y^p)$ is also calculated, which enables comparison with deterministic models. Here, the median is chosen as the point estimator for robustness compared to other options such as the mode or mean, which can be sensitive to outliers.

Table 5: Performance metrics of the calibrated model including assessment of the PIs 50%-99% for a week of test data.

Metric	u_v	$T_{c,w,in}$	$T_{c,w,out}$	$T_{c,a,out}$	$\dot{P}_f$
MAE($Q_{50\%}$)	0.06	0.54 °C	0.37 °C	0.19 °C	46 W
RMSE($Q_{50\%}$)	0.07	0.72 °C	0.49 °C	0.29 °C	62 W
ACE _{50%}	-4%	6%	-3%	-5%	6%
ACE _{55%}	-3%	7%	-3%	-6%	5%
ACE _{60%}	-3%	7%	-3%	-8%	4%
ACE _{65%}	-2%	7%	-3%	-9%	3%
ACE _{70%}	-2%	8%	-2%	-10%	2%
ACE _{75%}	-2%	8%	-2%	-11%	1%
ACE _{80%}	-1%	7%	-1%	-11%	0%
ACE _{85%}	0%	6%	1%	-11%	-1%
ACE _{90%}	2%	4%	3%	-10%	-2%
ACE _{95%}	2%	2%	2%	-8%	-3%
ACE _{99%}	0%	0%	0%	-2%	-3%

The results of applying these metrics on the posterior predictive distribution are shown in Table 5. As shown, the automated model generation and calibration method provides fairly accurate point predictions for all five measuring devices with an RMSE of 0.72 °C, 0.49 °C, 0.29 °C for the <inlet water temperature sensor>, <outlet water

temperature sensor>, and <outlet air temperature sensor>, respectively. The <valve position sensor> and <fan power meter> has an RMSE of 0.07 and 62 W, respectively.

The ACE is calculated for a broad range of PIs from 50% to 99%. As shown, the 99% PI is the most reliable with absolute ACE values less than or equal to 3% across all 5 measuring devices. For the <valve position sensor>, <inlet water temperature sensor>, and <water outlet temperature sensor>, ACE_{99%} is 0%, meaning that 99% of all observations are contained in the PI as expected. The PIs between 75 % and 85% are not as reliable with -11% ACE for the <outlet air temperature sensor>. This means that the PI is too narrow and contains 11% fewer observations than expected. If the model predictions are used as part of a decision-making process, commonly chosen PIs are 90%, 95% or 99%, but they should be chosen based on the specific application and the tradeoff between risk and utility of the generated interval [50].

The model predictions are shown for a 1-day period of the test data in Figure 17. As also indicated by the before-mentioned RMSE metrics, the median of the predictive distributions (black dashed) agrees well with the observed data (red) and is able to capture the dynamics of the system. For instance, the flow temperatures of the system is highly sensitive to the valve position (u_v), which is set by the controller. At around 10:00 the setpoint of the controller is changed from 21 °C to 22 °C, causing the valve position to open slightly, resulting in the desired outlet air temperature change. The median point estimates are not able to capture all fast dynamics, especially towards the end of the period where the flow temperatures and valve position start to fluctuate in the real system. However, this uncertainty is reflected in the 99% PI, which covers these fluctuations.

6.5. Limitations and future work

In general, the flexibility of the proposed method is highly dependent on the number and diversity of components present in the component model library. Therefore, future work should apply the proposed methodology to a diverse set of case studies and expand the component model library accordingly. It is expected that the required effort should decrease and the level of automation increase for each new use case as new components are added to the library. In this regard, as the component library grows, a robust method for prioritizing between multiple fitting models in the translation stage must be developed. A simple option is to allow the user to determine the desired detail of the model by setting a high-level complexity option similar to the translation method proposed by Reynders et al. [51].

The proposed translation and parameter estimation methodology is independent of the ontology used for the semantic model. However, for any given ontology and semantic model used as input, a corresponding library with ontology-specific signature patterns is required. The signature patterns developed for the case study in this work are thus only viable for semantic models following the SAREF ontology and its extensions. Therefore, to accommodate other ontologies and improve interoperability, earlier work on domain ontology mapping between

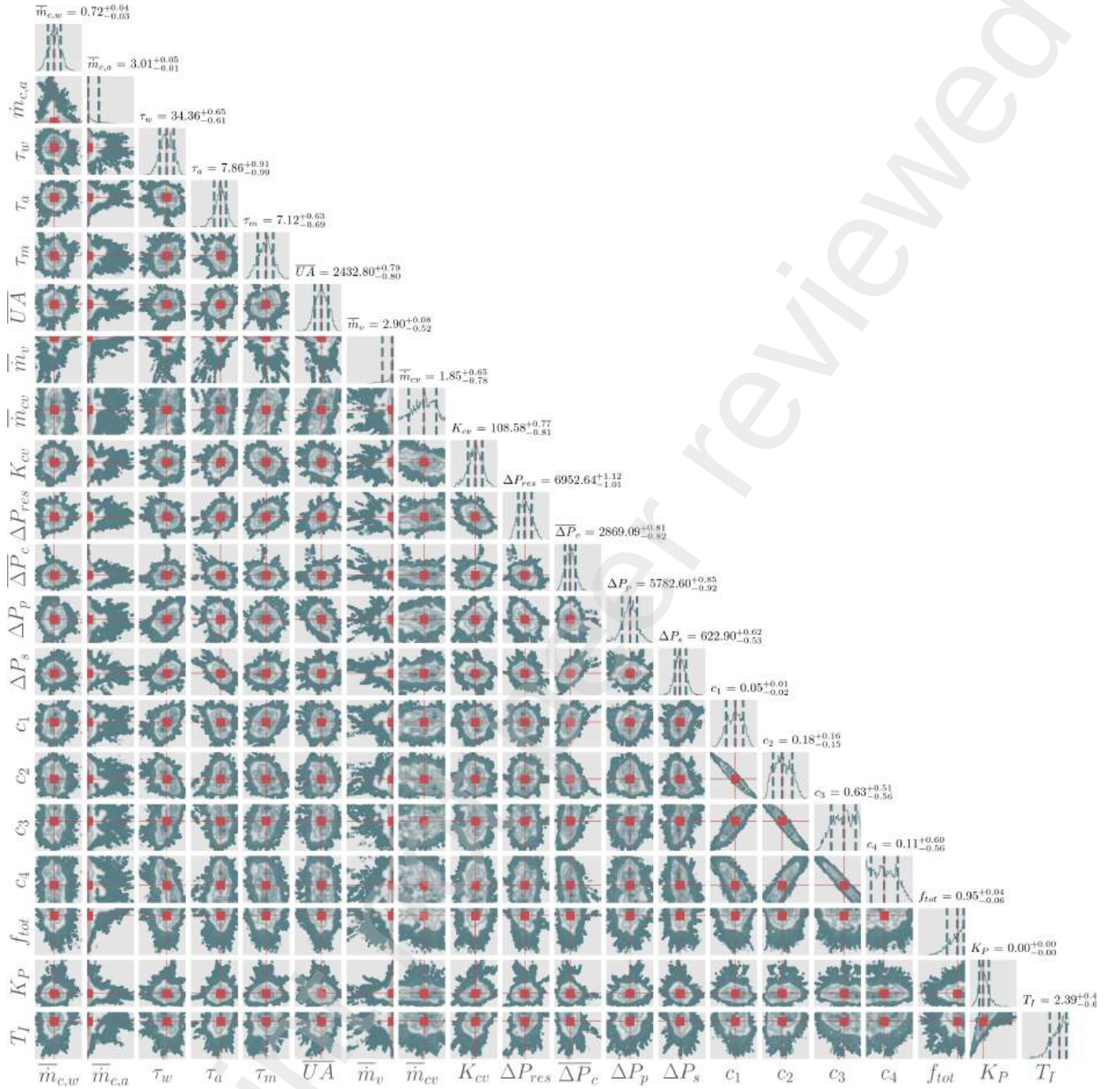


Figure 16: Corner plot summarizing the posterior distribution of the sampled parameters. The pairwise correlation between the parameters are shown as a matrix plot while the marginalized distribution for each parameter is shown at the diagonal as a histogram. The statistics at the top of the histograms show the quantiles in the format $\theta = Q_{50\%}^{+(Q_{84\%}-Q_{50\%})}_{-(Q_{50\%}-Q_{16\%})}$.

e.g. SAREF and Brick [52, 53] could be leveraged for automated translation of either the signature patterns or the semantic model used as information source.

7. Conclusion

In conclusion, this paper addresses the pressing need for automation and interoperability in the development and calibration

of BPS models. By examining the challenges faced in current practices, particularly in model development and calibration stages, the paper proposes a comprehensive approach that integrates automated model generation and calibration processes using an ontology-based method.

The proposed approach leverages semantic models to automate model generation, introducing the concept of signature patterns to describe component model applicability. Through

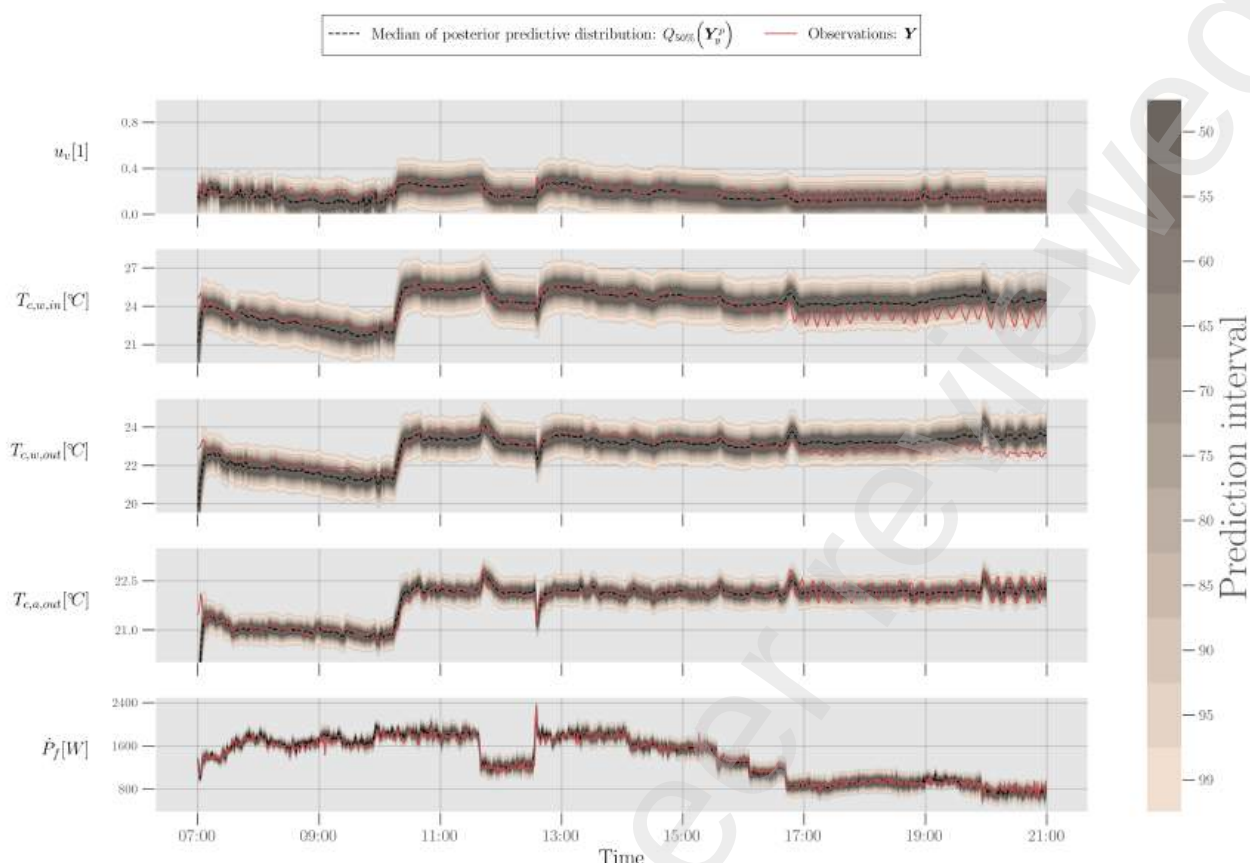


Figure 17: The median and PIs of the posterior predictive distributions compared with observations for a 1-day period of the test data.

a graph search process, these patterns are matched against semantic models to generate simulation models, thus streamlining the model development process. Based on the generated simulation model, the paper presents a general Bayesian parameter estimation problem to enable probabilistic parameter estimation and simulation based on operational data from measuring devices.

The proposed methodology is implemented and validated through a case study on a ventilation system, comprised of 8 measuring devices, a controller, a fan, and a heating coil water loop. This system is represented through a semantic model using the SAREF ontology with the SAREF4BLDG and SAREF4SYST extensions. Signature patterns are developed and dynamic component and system models from the Modelica Buildings library are exported as functional mockup units (FMU). Using the proposed methodology, a simulation model is generated, calibrated, and assessed through different performance metrics. These metrics demonstrate the accuracy and reliability of both model point estimates and probabilistic prediction intervals across all model outputs.

In summary, this paper contributes an automated approach to the development and calibration of BPS models, addressing important challenges in the field. By integrating semantic models with data-driven simulation models, this work could drive forward

the development of building digital twins and facilitate the implementation of various building services that can enhance building operation. In this regard, the probabilistic properties of the generated and calibrated model could provide a strong basis for operational decision-making.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

anonymous

Author statement

anonymous

References

- [1] B. Bleys, S. Jensen, A. Marszal, Annex 67 – energy flexible buildings, REHVA Journal (6) (2018) 32–35.

- [2] T. Hong, J. Langevin, K. Sun, Building simulation: Ten challenges, *Building Simulation* 11 (03 2018). doi:10.1007/s12273-018-0444-x
- [3] D. Crawley, L. Lawrie, F. Winkelmann, W. Buhl, Y. Huang, C. Pedersen, R. Strand, R. Liesen, D. Fisher, M. Witte, J. Glazer, Energyplus: creating a new-generation building energy simulation program, *Energy and Buildings* 33 319–331.
- [4] F. Winkelmann, B. Birdsall, W. Buhl, K. Ellington, A. Erdem, J. Hirsch, S. Gates, Doe-2 supplement: Version 2.1e”.
- [5] S. Klein, W. Beckman, Trnsys 16: A transient system simulation program: mathematical reference, *TRNSYS* 5 389–396.
- [6] M. Wetter, M. Sulzer, A call to action for building energy system modelling in the age of decarbonization, *Journal of Building Performance Simulation* 17 (3) (2024) 383–393. doi:10.1080/19401493.2023.2285824. URL <https://doi.org/10.1080/19401493.2023.2285824>
- [7] V. Bazjanac, T. Maile, J. O Donnell, C. Rose, N. Mrazović, Data environments and processing in semi-automated simulation with energyplus, 2011.
- [8] H. Kim, Z. Shen, I. Kim, K. Kim, A. Stumpf, J. Yu, Bim ifc information mapping to building energy analysis (bea) model with manually extended material information, *Automation in Construction* 68 (2016) 183–193. doi:https://doi.org/10.1016/j.autcon.2016.04.002. URL <https://www.sciencedirect.com/science/article/pii/S0926580516300656>
- [9] M. Wetter, C. van Treeck, IEA EBC Annex 60: New Generation Computing Tools for Building and Community Energy Systems, 2017. URL <https://www.iea-annex60.org/pubs.html>
- [10] M. Wetter, C. van Treeck, L. Helsen, A. Maccarini, D. Saelens, D. Robinson, G. Schweiger, IBPSA project I: BIM/GIS and modelica framework for building and community energy system design and operation – ongoing developments, lessons learned and challenges, *IOP Conference Series: Earth and Environmental Science* 323 (2019) 012114. doi:10.1088/1755-1315/323/1/012114. URL <https://doi.org/10.1088/1755-1315/323/1/012114>
- [11] J. Cao, T. Maile, J. O Donnell, R. Wimmer, C. Treeck, Model transformation from simmodel to modelica for building energy performance simulation, 2014.
- [12] J. Cao, R. Wimmer, M. Thorade, T. Maile, J. O Donnell, J. Rädler, J. Frisch, C. Treeck, A flexible model transformation to link bim with different modelica libraries for building energy performance simulation, 2015. doi:10.26868/25222708.2015.2446
- [13] R. Wimmer, T. Maile, J. O Donnell, J. Cao, C. Treeck, Data-requirements specification to support bim-based hvac definitions in modelica, 2014.
- [14] R. Wimmer, J. Cao, P. Remmen, T. Maile, J. O Donnell, J. Frisch, R. Streblow, D. Mueller, C. Treeck, Implementation of advanced bim-based mapping rules for automated conversions to modelica, 2015. doi:10.26868/25222708.2015.2424
- [15] J. O Donnell, R. See, T. Maile, V. Bazjanac, P. Haves, Simmodel: A domain data model for whole building energy simulation, 2011.
- [16] A. Andriamamonjy, D. Saelens, R. Klein, An automated ifc-based workflow for building energy performance simulation with modelica, *Automation in Construction* 91 (2018) 166–181. doi:https://doi.org/10.1016/j.autcon.2018.03.019. URL <https://www.sciencedirect.com/science/article/pii/S0926580517308282>
- [17] K. Katsigarakis, G. Lilis, G. Giannakis, D. Rovas, An ifc data preparation workflow for building energy performance simulation, 2019. doi:10.35490/EC3.2019.188
- [18] Ieee standard glossary of software engineering terminology, *IEEE Std 610.12-1990* (1990) 1–84. doi:10.1109/IEEESTD.1990.101064
- [19] IEA Energy in Buildings and Communities, EBC Annex 91: Open BIM for Energy Efficient Buildings, Tech. rep., IEA Energy in Buildings and Communities (2023). URL <https://annex91.iea-ebc.org/>
- [20] C. D. Roa, p. Raftery, A. K. Prakash, T. Pfeffer, Field demonstration of the brick ontology to scale up the deployment of ashrae guideline 36 control sequences (6 2023). doi:10.20357/B7J88Q. URL <https://www.osti.gov/biblio/1996791>
- [21] Z. Zheng, E. Yahia, E. Farzadaghi, R. EL Meouche, F. Ababsa, P. Béguery, A data framework enabling bem and bms interoperability based on semantic web technologies and brick ontology, 2023. doi:10.26868/25222708.2023.1366
- [22] Z. Wu, J. C. Cheng, Z. Wang, H. H. Kwok, An ontology-based framework for automatic building energy modeling with thermal zoning, *Energy and Buildings* 296 (2023) 113267. doi:https://doi.org/10.1016/j.enbuild.2023.113267. URL <https://www.sciencedirect.com/science/article/pii/S0378778823004978>
- [23] A. Reddy, Literature review on calibration of building energy simulation programs: Uses, problems, procedure, uncertainty, and tools, *ASHRAE Transactions* 112 226–240.
- [24] G. Chaudhary, J. New, J. Sanyal, P. Im, Z. O’Neill, V. Garg, Evaluation of “autotune” calibration against manual calibration of building energy models, *Applied Energy* 182 (2016) 115–134. doi:https://doi.org/10.1016/j.apenergy.2016.08.073. URL <https://www.sciencedirect.com/science/article/pii/S030626191631159X>
- [25] J. Sanyal, J. New, Simulation and big data challenges in tuning building energy models, in: 2013 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES), IEEE, 2013, pp. 1–6.
- [26] A. Garrett, J. New, T. Chandler, Evolutionary tuning of building models to monthly electrical consumption, *ASHRAE Transactions* 119 (2013) 89–100.
- [27] A. Chong, Y. Gu, H. Jia, Calibrating building energy simulation models: A review of the basics to guide future work, *Energy and Buildings* 253 (2021) 111533. doi:https://doi.org/10.1016/j.enbuild.2021.111533. URL <https://www.sciencedirect.com/science/article/pii/S0378778821008173>
- [28] O. Vera-Piazzini, M. Scarpa, Building energy model calibration: A review of the state of the art in approaches, methods, and tools, *Journal of Building Engineering* (2023) 108287. doi:https://doi.org/10.1016/j.jobe.2023.108287. URL <https://www.sciencedirect.com/science/article/pii/S2352710223024701>
- [29] A. Chong, W. Xu, S. Chao, N.-T. Ngo, Continuous-time bayesian calibration of energy models using bim and energy data, *Energy and Buildings* 194 (2019) 177–190. doi:https://doi.org/10.1016/j.enbuild.2019.04.017. URL <https://www.sciencedirect.com/science/article/pii/S0378778818339094>
- [30] M. C. Kennedy, A. O’Hagan, Bayesian calibration of computer models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63 (2001). URL <https://api.semanticscholar.org/CorpusID:119562136>
- [31] M. Laura Daniele, R. Garcia-Castro, M. Lefrançois, M. Poveda-Villalon, Saref: the smart applications reference ontology, [Online]. Available:.. URL <https://saref.etsi.org/core/>
- [32] M. Poveda-Villalón, R. Garcia-Castro, Saref extension for building, [Online]. Available:.. URL <https://saref.etsi.org/saref4bldg/>
- [33] M. Lefrançois, Saref4syst: an extension of saref for typology of systems and their inter-connections, [Online]. Available:.. URL <https://saref.etsi.org/saref4syst/>
- [34] C. Boje, A. Guerriero, S. Kubicki, Y. Rezgui, Towards a semantic construction digital twin: Directions for future research, *Automation in Construction* 114 (2020) 103179. doi:https://doi.org/10.1016/j.autcon.2020.103179. URL <https://www.sciencedirect.com/science/article/pii/S0926580519314785>
- [35] M. Grieves, Digital twin: Manufacturing excellence through virtual factory replication (03 2015).
- [36] V. Kukkonen, A. Küçükavci, M. Seidenschnur, M. H. Rasmussen, K. M. Smith, C. A. Hviid, An ontology to support flow system descriptions from design to operation of buildings, *Automation in Construction* 134 (2022) 104067. doi:https://doi.org/10.1016/j.autcon.2021.104067. URL <https://www.sciencedirect.com/science/article/pii/S0926580521005185>
- [37] Modelica Association, Functional Mock-up Interface for model exchange and co-simulation version 2.0, Tech. rep., Modelica Association,

- Linköping (Jul. 2014).
URL <https://fmi-standard.org>
- [38] M. Wetter, K. S. Benne, A. Gautier, T. S. Noudui, A. Ramle, A. Roth, H. Tummescheit, S. G. Mentzer, C. Winther, [Lifting the garage door on spawn, an open-source bem- controls engine](#), 2020.
URL <https://api.semanticscholar.org/CorpusID:221857040>
- [39] J. Lee, W.-S. Han, R. Kasperovics, J.-H. Lee, [An in-depth comparison of subgraph isomorphism algorithms in graph databases](#), *Proc. VLDB Endow.* 6 (2012) 133–144.
URL <https://api.semanticscholar.org/CorpusID:9694765>
- [40] L. Cordella, P. Foggia, C. Sansone, M. Vento, A (sub)graph isomorphism algorithm for matching large graphs, *Pattern Analysis and Machine Intelligence*, *IEEE Transactions on* 26 (2004) 1367 – 1372. doi:10.1109/TPAMI.2004.75
- [41] J. Bjørnskov, M. Jradi, [An ontology-based innovative energy modeling framework for scalable and adaptable building digital twins](#), *Energy and Buildings* 292 (Aug. 2023). doi:10.2139/ssrn.4341936
- [42] J. R. Goodman, J. Weare, [Ensemble samplers with affine invariance](#), 2010.
URL <https://api.semanticscholar.org/CorpusID:54842399>
- [43] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, [Equation of state calculations by fast computing machines](#) (3 1953). doi:10.2172/4390578
URL <https://www.osti.gov/biblio/4390578>
- [44] W. K. Hastings, [Monte Carlo sampling methods using Markov chains and their applications](#), *Biometrika* 57 (1) (1970) 97–109. arXiv:<https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf>, doi:10.1093/biomet/57.1.97
URL <https://doi.org/10.1093/biomet/57.1.97>
- [45] D. Foreman-Mackey, D. W. Hogg, D. Lang, J. Goodman, [temcee/tt: The MCMC hammer](#), *Publications of the Astronomical Society of the Pacific* 125 (925) (2013) 306–312. doi:10.1086/670067
URL <https://doi.org/10.1086/670067>
- [46] W. D. Voudsen, W. M. Farr, I. Mandel, [Dynamic temperature selection for parallel tempering in markov chain monte carlo simulations](#), *Monthly Notices of the Royal Astronomical Society* 455 (2) (2015) 1919–1937. doi:10.1093/mnras/stv2422
URL <https://doi.org/10.1093/mnras/stv2422>
- [47] C. E. Rasmussen, C. K. I. Williams, *Gaussian processes for machine learning*, Adaptive computation and machine learning, MIT Press, 2006.
- [48] U.S. Department of Energy, [Energyplus version 9.5.0 documentation, energyplus engineering reference](#), Tech. rep., U.S. Department of Energy (2020).
URL https://energyplus.net/sites/all/modules/custom/nrel_custom/pdfs/pdfs_v9.5.0/EngineeringReference.pdf
- [49] P. Pinson, G. Kariniotakis, H. Nielsen, T. Nielsen, H. Madsen, *Properties of Quantile and Interval Forecasts of Wind Generation and Their Evaluation*, Vol. 2, 2006.
- [50] J. Landon, N. D. Singpurwalla, [Choosing a coverage probability for prediction intervals](#), *The American Statistician* 62 (2) (2008) 120–124. doi:10.1198/000313008X304062
- [51] G. Reynders, A. Andriamamonjy, R. Klein, D. Saelens, [Towards an ifc-modelica tool facilitating model complexity selection for building energy simulation](#), 2017. doi:10.26868/25222708.2017.621
- [52] G. F. Schneider, [Towards aligning domain ontologies with the building topology ontology conference](#), 2017.
URL <https://api.semanticscholar.org/CorpusID:29323176>
- [53] G. Schneider, [Automated ontology matching in the architecture, engineering and construction domain - a case study](#), 2019.

Large-scale field demonstration of an interoperable and ontology-based energy modeling framework for building digital twins

Jakob Bjørnskov^{a,*}, August Thomsen^a, Muhyiddine Jradi^a

^aSDU Center for Energy Informatics, Maersk Mc-Kinney Moeller Institute, The Faculty of Engineering, University of Southern Denmark, Campusvej 55, Odense, 5230, Denmark

Abstract

Digital twins have emerged as a promising concept for improving building energy efficiency, but their implementation faces challenges in interoperability and adaptability. This paper presents a large-scale field demonstration of an interoperable energy modeling framework for building digital twins, using ontology-based semantic models as data sources for automated model generation and calibration of data-driven component models. The study focuses on a single floor of a hospital building, comprising 12 conditioned zones and data from 45 measuring devices. Across the 45 sensors, the model achieved on average mean absolute errors of 0.40°C for temperature, 32 ppm for CO₂ concentration, 0.06 for valve position, and 0.04 for damper position predictions. These results demonstrate the framework's ability to generate and calibrate accurate and flexible building energy models with reduced effort. The paper also showcases the framework's practical application in exploring system modifications to improve indoor comfort, highlighting its potential for scenario analysis and decision support. The proposed approach significantly streamlines the process of creating and maintaining accurate, up-to-date energy models, offering a robust foundation for digital twin applications in the built environment.

Keywords: Digital Twin, Ontology, Data-driven, Building Energy Model, Automation, Interoperability

1. Introduction

In recent years, digital solutions and the use of Internet of Things (IoT) devices have shown significant potential to improve efficiency and reduce the costs of building operation through advanced services such as performance monitoring and operational optimization. As part of this digitalization, the digital twin (DT) concept has gained traction as a promising approach for managing and utilizing the large amounts of data generated in modern buildings. The DT concept comprises three elements: 1) a physical system, 2) a virtual system, and 3) a bidirectional data flow between the physical and virtual systems [1, 2]. Here, the virtual system should act as a digital replica or twin of the physical system, incorporating relevant data and simulation models. These models should be closely integrated, react to changes in the physical twin, and their granularity and accuracy should support the functional outputs or services delivered by the DT [3]. Despite decades of development in building data modeling and simulation, along with significant efforts to improve their interoperability and integration [4, 5, 6, 7, 8], they largely remain segregated silos [9, 10].

1.1. BIM-based Digital Twins

Within the built domain, one of the major enabling technologies of the past decades is Building Information Modeling

(BIM). The concept of BIM was initially designed as a file-based data-sharing platform, easing collaboration between various stakeholders during the design and construction phase of buildings [11]. The success and increased adoption of BIM for this purpose have sparked optimism that its capabilities can be extended to the operational phase of buildings, enabling DTs as a central data model. Much of the current research on DTs in the building sector focuses on developing BIM-based solutions and improving the BIM environment to support DTs.

For example, Opoku et al. [12], presented a BIM-based DT platform focusing on real-time indoor climate monitoring of a university living lab case study. The considered variables for monitoring included temperature, humidity, CO₂, and Total Volatile Organic Compounds (TVOC). Based on the historical data collected, static thresholds were considered to detect whether indoor conditions were acceptable.

Peng et al. [13], developed a whole-lifecycle BIM-based DT platform for a large hospital case study. For the operational stage, the study implemented abnormal electricity usage detection using k-means clustering, fault prediction of the air-handling unit using a Long Short-Term Model (LSTM) for classification, and a pattern recognition algorithm to analyze the repair messages reported by the technical staff. The study reported 1% energy savings and a 10% reduction in facility faults over one year of operation. Moreover, automated actuation and control of the physical system were not implemented due to interoperability issues with the Building Management System (BMS).

Although BIM might prove to play an important role as a

*Corresponding author

Email address: jabj@mmmi.sdu.dk (Jakob Bjørnskov)

Nomenclature

$\dot{m}_a$	Massflowrate of supply and exhaust air to a space	f_a	Fraction of solar irradiance absorbed as heat in the indoor air
$\dot{m}_w$	Massflowrate of water supplied to a space heater	f_w	Fraction of solar irradiance absorbed as heat in exterior walls
$\dot{Q}_{occ,gain}$	Heat generated by people and internal loads correlated with number of people	N_{occ}	Number of people in a space
$\bar{Q}_{rad}$	Nominal heat output from a spaceheater	R_{bou}	Thermal resistance of boundary surface between C_{bou} and T_z , and C_{bou} and T_{bou}
$\bar{T}_{w,ret}$	Nominal return water temperature of a space heater	R_{int}	Thermal resistance of interior wall between C_{int} and T_z , and C_{int} and $T_{z,adj}$
$\bar{T}_{w,sup}$	Nominal supply water temperature of a space heater	R_{in}	Thermal resistance of exterior wall between C_w and T_z
$\bar{T}_z$	Nominal air temperature at which $\bar{Q}_{rad}$ is calculated	R_{out}	Thermal resistance of exterior wall between C_w and T_{out}
Φ_{glob}	Global horizontal solar irradiation	$T_{a,sup}$	Supply air temperature
C_a	Thermal heat capacity of air, furniture, and objects contained in a space	T_{out}	Outdoor temperature
C_{bou}	Thermal heat capacity of a generic boundary surface	$T_{w,sup}$	Supply water temperature
C_{int}	Thermal heat capacity of interior walls	$T_{z,adj}$	Indoor temperature of an adjacent space
C_{out}	Outdoor CO ₂ concentration		
C_w	Thermal heat capacity of exterior walls		

valuable data source for DTs, it is currently uncertain whether it can fulfill the role as a central data model. This is because BIM was not originally designed to manage the dynamic and real-time data environments that are characteristic of DTs [2, 11].

1.2. Ontology-based Digital Twins

Recent developments in adjacent IoT-centric domains have highlighted the use of ontologies as a promising concept that can help align and integrate diverse data sources [2, 11, 14, 15]. Ontologies provide a web-based and standardized schema to share, represent, and integrate diverse data sources for improved interoperability [16, 17].

Using BIM as a data source, Merino et al. [18], proposed a data-lake architecture that leverages both the Industry Foundation Classes (IFC) standard and the Brick ontology to integrate data from various sources.

Similarly, Mokhtari et al. [19], developed a cognitive DT for building management and operations, utilizing BIM along with the Building topology Ontology (BOT) and Brick ontologies. Their work demonstrated the integration of semantic models with various analytics for improved building management, exploring the synergy between BIM and semantic web technologies in DT applications.

Ni et al. [20], demonstrated the use of the RealEstateCore ontology and the Microsoft Azure Digital Twins IoT platform [21] for real-time monitoring of a historic city theater. In a subsequent study, the Brick ontology was further applied to represent systems and equipment [22]. As part of a data analytics layer, different deep learning models were trained and tested for 1-hour-ahead forecasting of heating and electricity consumption.

1.3. Data-driven energy modeling and interoperability

Simulation and prediction capabilities are an integral part of DT solutions and are highly prioritized in other domains with more mature DT development, such as the manufacturing and aerospace industry [2]. However, the building domain is still facing challenges in improving interoperability and reducing implementation costs of building simulation technology [9].

Within the building domain, model types are often categorized as white-box, black-box, or grey-box. White-box models are physics-based, built from domain-specific expert knowledge and first principles. Traditional whole-building simulation programs, such as EnergyPlus [23], fall under the white-box category. On the other hand, black-box models are purely data-driven, relying on machine-learning techniques to extract statistical patterns in data. Grey-box models are a mix of white and black-box models, where physics-based principles are used to define the model structure and data is used to estimate the unknown model parameters.

Current research in the built environment that incorporates predictive features as part of DT solutions primarily relies on monolithic black-box models [13, 18, 19, 22, 24], i.e. singular end-to-end black-box models. These models are particularly appealing in the context of DTs because they are straightforward to deploy, require little domain knowledge, and can handle large data volumes effectively. However, their generalization capabilities are often insufficient for tasks such as what-if analysis, retrofitting, and model-predictive control [25, 26]. Moreover, the inherent nature of monolithic black-box models offers limited insight and interpretability for decision-makers [27], essential for ensuring the model's adoption and trust [28].

In comparison, grey-box models have been shown to generalize better outside the training data region due to the incorporated physics [25, 26]. However, grey-box models are currently not as popular for DT applications, as modeling expertise and system topology information is required to define the model structure, making them time-consuming to build and maintain in a DT context. Improving interoperability between data models and simulation models is therefore crucial to overcome these challenges and unlock proven model-based services that can greatly improve building operation efficiency [9, 29, 30].

1.4. Aim and contribution

This paper addresses the aforementioned challenges by presenting a large-scale field demonstration of an interoperable energy modeling framework for building DTs. The framework, built on previous work [31, 32] and it combines ontology-based semantic models with data-driven component models to improve the interoperability and adaptability of energy simulation models. It is available as an open-source Python library [33].

Through a comprehensive case study of a hospital building, we demonstrate how this approach can be scaled to complex, real-world scenarios with reduced effort. The contributions of this paper are summarized as follows:

1. **Demonstrating the potential of semantic models:** We showcase how semantic models can serve as effective data sources for the automated generation and calibration of building energy models. This approach significantly reduces both the effort required for modeling and the ongoing maintenance needs, streamlining the process of creating accurate and up-to-date energy models.
2. **Establishing seamless information carryover:** We validate and demonstrate the automated and seamless transfer of information from semantic models to energy models, showcasing their ability to serve as a foundational framework for DT applications. By leveraging semantic models, we enable a streamlined integration that not only improves the fidelity of energy simulations but also supports the dynamic, real-time requirements of DT environments. This approach facilitates the continuous updating and refinement of energy models, aligning them with real-world conditions as they evolve. This capability is essential for the effective deployment of DTs, as it ensures that the virtual representations remain in sync with their physical counterparts.
3. **Proving scalability through a case study:** We validate the scalability of this approach by analyzing a comprehensive case study involving a building with 12 conditioned and thermally coupled zones. The study includes data from a total of 45 measuring devices, against which the model is calibrated. This case study highlights the model's ability to handle complex and diverse real-world conditions.
4. **Applying the generated model in a practical use case:** We apply the generated and calibrated model to a specific use case, demonstrating its practical applications and the benefits of improved interoperability between simulation

models. This application highlights how improved model integration can lead to more accurate simulations, better decision-making, and optimized building energy performance.

The paper is structured as follows: In section 2, an overview of the methodology, scope, and broader context of the paper is provided. In section 3, the case study is introduced, including relevant information on spaces, equipment, type of control, etc. section 4 then describes how this information is represented in the developed semantic model of the case study, and how this information is used to generate a simulation model. In section 5, an automated parameter estimation process is applied to calibrate the generated model and relevant performance metrics are reported. The potential of the method is further demonstrated in section 6, where the fully adaptable simulation model is used to explore different system configurations to improve indoor climate. Limitations and future work is then discussed in section 7. Finally, a conclusion with the main findings and reflections is presented in section 8.

2. Methodology and context

As established, a detailed and accurate energy model is a vital element of DTs to reflect the physical system as closely as possible and unlock services based on predictive capabilities. This paper demonstrates how combining semantic information with data-driven models can produce a fully interoperable and adaptable energy model. This concept is summarized in Figure 1 within the context of DTs and its three high-level parts: The physical system, the data processing layer, and the virtual system, replicating the behavior of the physical system. For the application shown in this article, two types of information sources are required from the physical system: 1) a semantic model representing the topology of systems, sensors, flow paths, and equipment design information, and 2) operational time-series data from the various measuring devices installed in the real building.

The scheme and structure used to represent the information in the semantic model are specified using an ontology. In this work, the Smart Appliances REference (SAREF) ontology [34, 35] is used with the extensions SAREF4BLDG [36], SAREF4SYST [37], and the Flow Systems Ontology (FSO) [38]. However, it should be mentioned that the method generalizes to support any ontology.

The semantic model should be synchronized with the physical system, reflecting any changes such as equipment replacement, removal, or addition. This allows the simulation model to update accordingly, as is demonstrated later. Based on the information contained in the semantic model, the translator then performs a graph-based pattern match search to identify and assemble the appropriate component models from the component model library to generate a simulation model. Examples of the semantic model and the component model library, related to the case study considered, are provided in subsection 4.2 and subsection 4.3. After model generation, the model $\mathcal{M}$ is used for

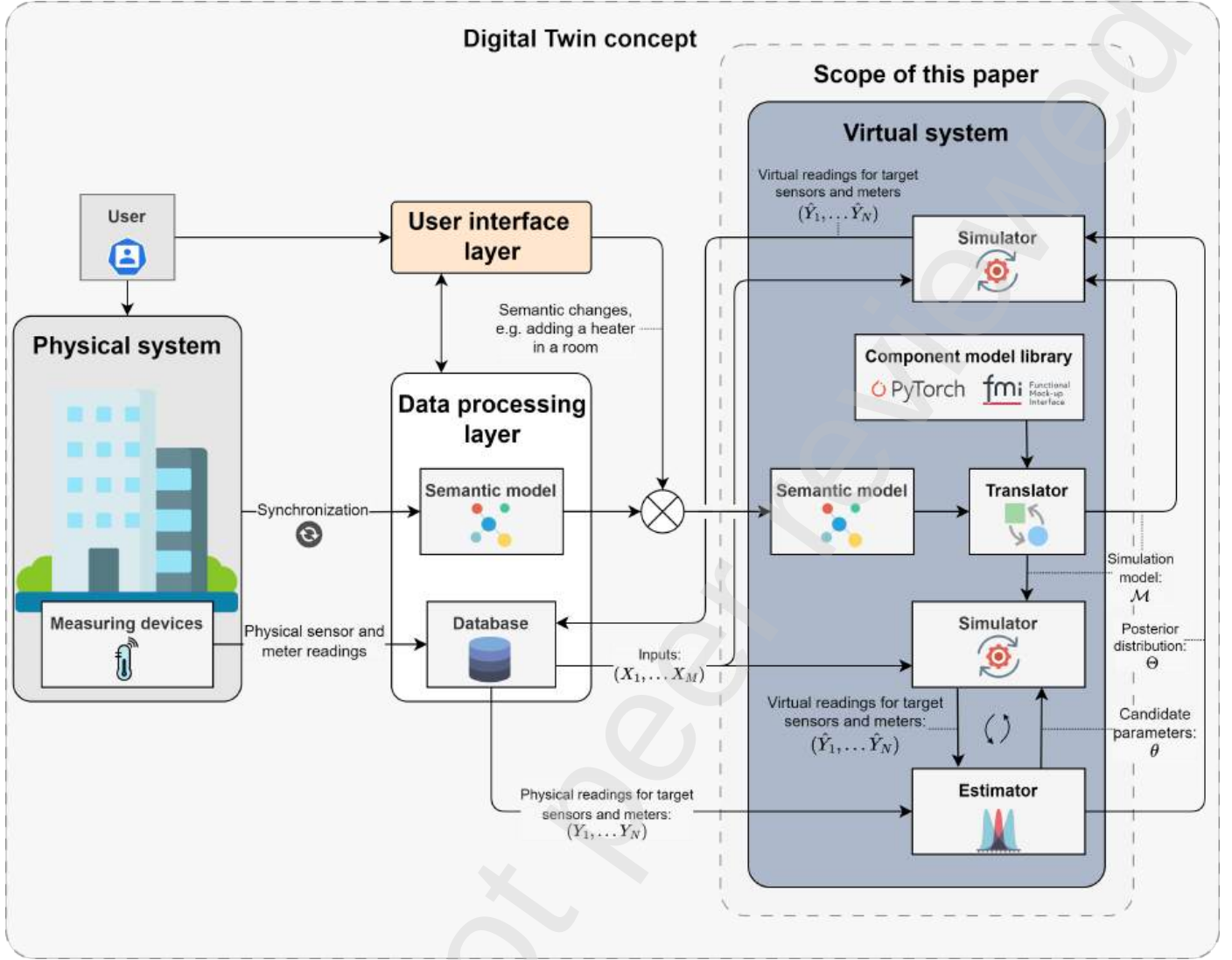


Figure 1: Methodology overview showing the role of the semantic model and how a simulation model can be continuously adapted to reflect the current state of the physical system or for testing alternative designs or configurations. Adapted from [31].

simulation, first as part of parameter estimation (bottom), and afterward for inference (top).

A generic user interface layer is also added in the figure to show how a potential user can take advantage of the automated model generation by adding or removing objects to test alternative designs or system configurations. This is explored further in [section 6](#).

Note that both the user interface layer and data processing layer have been largely simplified in [Figure 1](#), as they are outside the scope of this work. In addition, integration with other important virtual representations, such as 3D graphical models, is not included. In previous studies, different high-level DT architectures have been proposed to deal with such integrations [18, 19, 39].

3. Case study

To demonstrate and evaluate the proposed energy modeling and simulation approach, an actual case study building is considered in this paper. The case study is a two-story building within a Danish hospital with a kitchen at the ground floor and office spaces, meeting rooms, and changing rooms at the first floor. The focus of this demonstration is exclusively on the first floor, as illustrated in [Figure 2](#). The hospital itself is situated in an open area, while the case study building is next to the main hospital and is exposed to outdoor environmental conditions on most of its outer surfaces.

3.1. Spaces

As shown in [Figure 2](#), the first floor of the building comprises six office spaces, with floor areas ranging from 18 m² to 31 m². These offices are equipped with desks, computers, monitors, and multiple lighting sources. The number of occupants



Figure 2: Graphical overview of the considered floor section of the hospital case study.

Table 1: General information on the spaces included in the demonstration case.

Space	Area [m ²]	Type	Occupancy	Measuring devices				Climate control	
				Temperature	CO ₂	Valve position	Damper position	Temperature	DCV
1	82.2	Accommodation		X	X	X	X	X	X
2	15.6	Storage/Copy					X		
3	23.1	Office	3	X	X	X	X	X	X
4	20.7	Showers		X	X	X	X	X	
5	23.1	Office	1	X	X	X	X	X	X
6	41.2	Showers		X	X	X	X	X	
7	46.6	Meeting Room	9	X	X	X	X	X	X
8	46.6	Meeting Room	9	X	X	X	X	X	X
9	23.1	Office	1	X	X	X	X	X	X
10	23.1	Office	1	X	X	X	X	X	X
11	17.5	Office	1	X	X	X	X	X	X
12	30.7	Office	6	X	X	X	X	X	X

varies from single-person offices to shared offices that can accommodate up to six people. Two large meeting rooms with an area of 46 m², are separated by a folding wall, allowing them to be combined into a single larger space. The first floor also contains two separate changing rooms with areas of 21 m² and 42 m², these changing rooms have 2 and 6 walled toilet stalls respectively. Furthermore, the floor contains a shared staff accommodation with an area of 82 m² and a kitchenette. Other facilities on this floor, such as a stair wall, elevator, and multiple single-stall toilets, are not considered during the demonstration. Table 1 presents a summary of the spaces, including their respective areas, their designated use, and occupancy. Each of the spaces is equipped with multiple sensors and systems, including sensors that measure temperature, CO₂, radiator valve positions, and damper positions. As shown in Table 1, Space 2 does not contain temperature, CO₂, and valve position sensors. These sensor types are installed in all remaining spaces.

3.2. Indoor climate control

As shown in Table 1, demand-controlled ventilation (DCV) is employed in most rooms, regulating air supply flow by actuating the dampers using proportional-integral controllers (PI) to maintain the measured CO₂ concentration within comfortable ranges. All spaces except space 2 are equipped with radiators and motorized valves, also controlled using PI controllers. Most dampers are operated in an on-off fashion to also provide heating in a 0%-30% opening position range.

3.3. HVAC systems

The first floor has a dedicated ventilation system with a capacity of 11,000 m³/h, equipped with a rotating heat exchanger and cooling- and heating water coils. The building is supplied by district heating with an outdoor-compensated supply water temperature, operating between 50 to 60 °C when the outdoor temperature is between -12 and 5 °C.

3.4. Weather data

To provide boundary conditions for the generated simulation model of the case study building, temperature and solar radiation data collected from the Danish Meteorological Institute's weather API [40] are used. The weather data was sampled from one of the nearest weather stations to the hospital.

4. Ontology-driven model generation

This section introduces the semantic model representing the hospital case study, providing concrete examples of its structure and semantics. Following, an example model, part of the component model library, is presented along with its semantic representation, denoted as a signature pattern. Finally, using the component models and patterns from the library and the semantic model of the case study, a simulation model is generated and presented.

4.1. Semantic model

A semantic model is employed to represent the physical system in a comprehensive and machine-readable format. The information contained in the semantic model includes the system topology, relationships, and equipment properties of various elements. The semantic model is organized using the so-called triple construct, denoted by $\text{subject} \rightarrow \text{predicate} \rightarrow \text{object}$. Examples of such relations are shown for the case study as a graph in Figure 3, where subjects and objects are represented by nodes, and predicates are represented by edges. The brackets $\langle \rangle$ are used to represent an instance of a class, e.g. $\langle \text{Space_4_space_heater_valve} \rangle$ as an instance of the class SpaceHeater . The figure shows both the full semantic model as well as a highlighted section related to room number 4.

As shown in Figure 3, the semantic model contains the relevant design data and topological information for the case study described in section 3. For instance, it contains the following four triples (only triple 2, 3, and 4 are visible in Figure 3).

1. $\langle \text{Space_1} \rangle \rightarrow \text{connectedTo} \rightarrow \langle \text{Space_4} \rangle$
2. $\langle \text{Space_4_space_heater_valve} \rangle \rightarrow \text{suppliesFluidTo} \rightarrow \langle \text{Space_4_space_heater} \rangle$
3. $\langle \text{Space_4_space_heater_valve} \rangle \rightarrow \text{isContainedIn} \rightarrow \langle \text{Space_4} \rangle$
4. $\langle \text{Space_4_space_heater} \rangle \rightarrow \text{isContainedIn} \rightarrow \langle \text{Space_4} \rangle$

The first triple specifies that $\langle \text{Space_1} \rangle$ and $\langle \text{Space_4} \rangle$ are adjacent while the second triple specifies that $\langle \text{Space_4_space_heater_valve} \rangle$ is placed before $\langle \text{Space_4_space_heater} \rangle$ in the flow path of the heat carrying fluid. The third and fourth triple specifies that $\langle \text{Space_4_space_heater_valve} \rangle$ and $\langle \text{Space_4_space_heater} \rangle$ are both placed within $\langle \text{Space_4} \rangle$. Using the concept of triples, the whole building can accurately be represented as a large graph or network of objects and relationships. The full semantic model of the case study contains 459 instances of dampers, spaces, controllers, heaters, properties, etc. and 1254 triples connecting these instances.

4.2. Component model library

The component library contains the building blocks from which the simulation model is automatically assembled. It consists of modular self-contained grey-box and black-box models for which inputs and outputs can be connected in different configurations and adapted based on specific use cases. This is achieved using a so-called signature pattern, attached to each component model. The signature pattern acts akin to a blueprint of the component model, describing the semantic context in which the model applies and how it connects to other components.

An example of a component model related to the case study and its accompanying signature pattern is shown in Figure 4 and Figure 5. Here, Figure 4 shows a Modelica diagram representing the heat and mass dynamics of a space. The model is largely based on the grey-box model developed by Blum et al. [30]. It includes an outdoor boundary, one adjacent space, a space heater, and balanced supply and exhaust ventilation. The accompanying signature pattern shown in Figure 5 represents the system using the semantics of the applied ontologies SAREF, SAREF4BLDG, SAREF4SYST, and FSO. In this pattern, each node represents one or more ontology classes, e.g. BuildingSpace , and are given a local id, e.g. n_3 , to distinguish nodes. Below, a subset of four triples, included in the signature pattern, is listed.

1. $\text{BuildingSpace}(n_{10}) \rightarrow \text{connectedTo} \rightarrow \text{BuildingSpace}(n_3)$
2. $\text{Valve}(n_4) \rightarrow \text{suppliesFluidTo} \rightarrow \text{SpaceHeater}(n_5)$
3. $\text{Valve}(n_4) \rightarrow \text{isContainedIn} \rightarrow \text{BuildingSpace}(n_3)$
4. $\text{SpaceHeater}(n_5) \rightarrow \text{isContainedIn} \rightarrow \text{BuildingSpace}(n_3)$

The first triple specifies that the modeled space (n_3), must have an adjacent space (n_{10}). The second, third, and fourth triples specifies that a valve (n_4) must be placed before the space heater (n_5) in the flow path of the heat carrying fluid and that they must both be placed in the modeled space (n_3). Comparing these four triples with the four triples of the semantic model, highlighted in subsection 4.1, reveal a match. If the remaining triples in the signature pattern can also be matched in the semantic model, the component model fits the topology of the physical system. Completing this pattern matching procedure on the various models contained in the component model library, supporting different configurations of spaces, valves, controllers, sensors, etc. then produces a set of candidate models.

As the final step, input, output and parameter mapping information is used to assemble the matching component models into one simulation model. For the example space model, this information is provided in Table 2.

As shown in the first column, the model represents the behavior and dynamics of the n_3 and n_5 nodes in Figure 5. The example model has 9 inputs, including massflowrates for air and water $\dot{m}_a$ and $\dot{m}_w$, with temperatures $T_{a,sup}$ and $T_{w,sup}$. N_{occ} represents the number of occupants in the zone while the outdoor conditions are provided through the outdoor temperature T_{out} ,

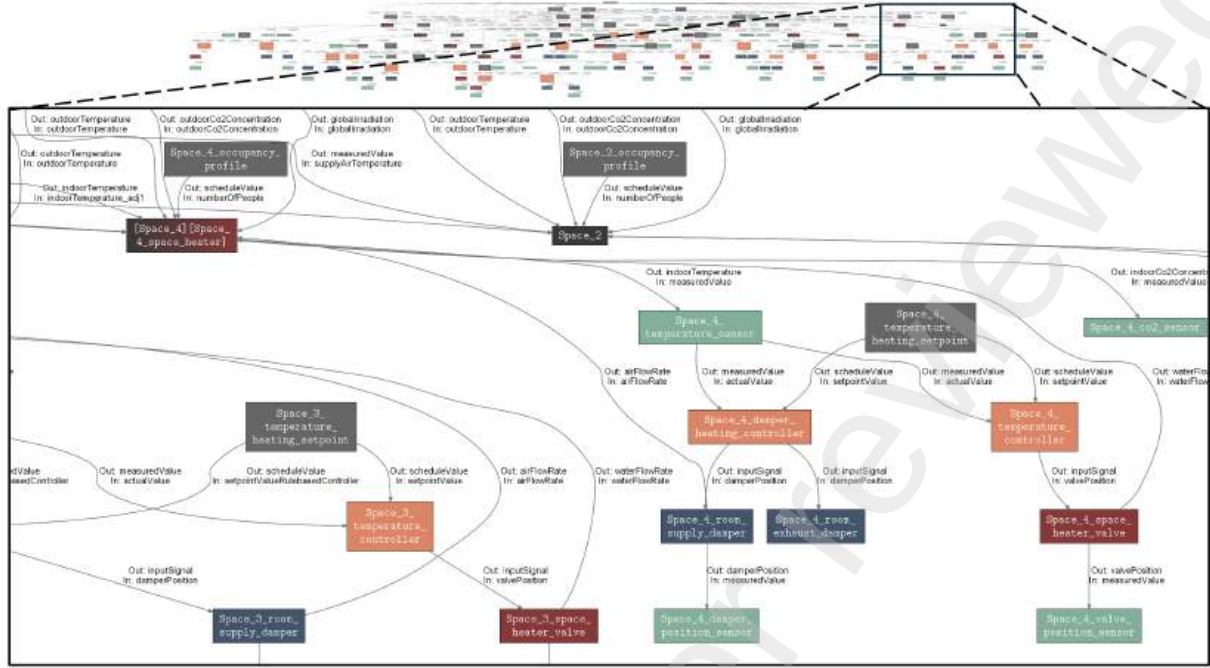


Figure 6: Simulation model of the hospital case generated automatically from the semantic model using the open source Python package available on GitHub [33]. The model contains 152 interconnected component models of dampers, spaces, controllers, etc. For readability, a section related to space 4 is enlarged.

Table 2: Example mapping of inputs, outputs, and parameters of the space model signature pattern.

Nodes	Input		Output	Parameter	
Modeled	Model	Source	Model	Model	Source
n_3	$\dot{m}_a$	$n_1: \dot{m}_a$	T_z	C_w	—
n_5	$\dot{m}_w$	$n_4: \dot{m}_w$	C_z	C_{int}	—
	N_{occ}	$n_6: y$		C_{bou}	—
	T_{out}	$n_7: T_{out}$		C_a	—
	C_{out}	$n_7: C_{out}$		R_{out}	—
	Φ_{glob}	$n_7: \Phi_{glob}$		R_{in}	—
	$T_{a,sup}$	$n_8: y_{meas}$		R_{int}	—
	$T_{w,sup}$	$n_{11}: y_{meas}$		R_{bou}	—
	$T_{z,adj}$	$n_{10}: T_z$		f_w	—
				f_a	—
				$\dot{Q}_{occ,gain}$	—
				$\dot{Q}_{rad}$	n_{14}
				$\bar{T}_{w,sup}$	n_{17}
				$\bar{T}_{w,ret}$	n_{17}
				$\bar{T}_z$	n_{17}

temperature, and n is the heat transfer exponent. As shown in Table 2, $\dot{Q}_{rad}$, $\bar{T}_{w,sup}$, $\bar{T}_{w,ret}$, and $\bar{T}_z$ can be mapped from the n_{14} and n_{17} equivalents in the semantic model.

4.3. Model generation

The semantic model serves as the basis for automatically generating the simulation model by matching and connecting component models. Performing this procedure on the semantic model of the hospital case study shown in Figure 3 generates the simulation model shown in Figure 6. Here, the

graph visualizes the information flow and the logical connections between each of the component models in the simulation model. For example, [Space_4] [Space_4_space_heater] represents an instance of the example model presented in subsection 4.2, showing some of the listed inputs and outputs. For instance, the [Space_4] [Space_4_space_heater] model receives the input airFlowRate (named $\dot{m}_a$ in Table 2) from the Space_4_room_supply_damper model.

As shown in Figure 6, the generated simulation model also contains component models of controllers, dampers, valves, sensors, setpoint schedules, etc. Component models with no inbound edges act as sources or inputs for the simulation model, e.g. the Space_4_temperature_heating_setpoint which produces a setpoint signal for the Space_4_damper_heating_controller and the Space_4_temperature_controller. In addition, measuring devices are directly represented in the simulation model, e.g. Space_4_temperature_sensor. These components act as interfaces or outputs of the simulation, enabling a direct comparison with the readings of real sensors and meters.

With the inclusion of controllers in the simulation model, the predictions and estimations in the following sections are all based on a closed-loop configuration. This means that the essential inputs for the model is weather data, setpoint schedules, and occupancy schedules.

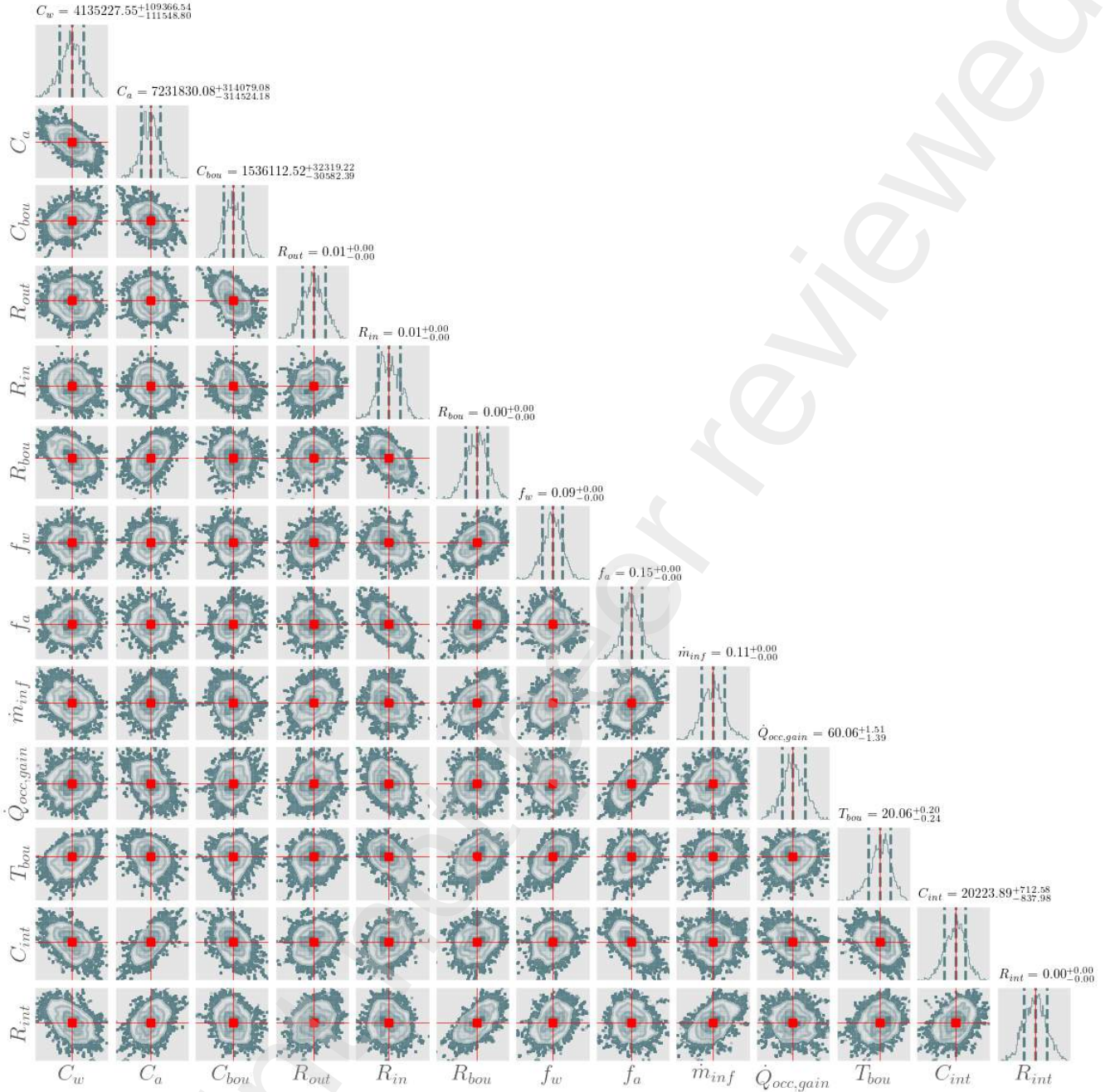


Figure 7: Corner plot summarizing the posterior distribution of the 13 sampled parameters related to the component model of space 4. The pairwise correlation between the parameters are shown as a matrix plot while the marginalized distribution for each parameter is shown at the diagonal as a histogram. The statistics at the top of the histograms show the quantiles in the format $\theta = Q_{50\%}^{+(Q_{84\%}-Q_{50\%})}_{-(Q_{50\%}-Q_{16\%})}$.

5. Bayesian parameter estimation

After model generation, parameters not mapped from the semantic model, are estimated using Monte Carlo Markov Chain (MCMC) sampling as described in an earlier work [31]. The goal of MCMC is to sample from the so-called posterior distribution, a distribution of the model parameters, given some training data. This posterior distribution can then be used

to produce posterior predictive distributions, i.e. probabilistic predictions. For the generated model of the case study, the resulting estimation problem targets 148 model parameters. Due to the significant size and dimensionality of the estimation problem, proper initialization of the MCMC algorithm is crucial to reach convergence within a reasonable time. Therefore, for initialization of parameters such as thermal capacities and resistances, high-level geometrical information about the

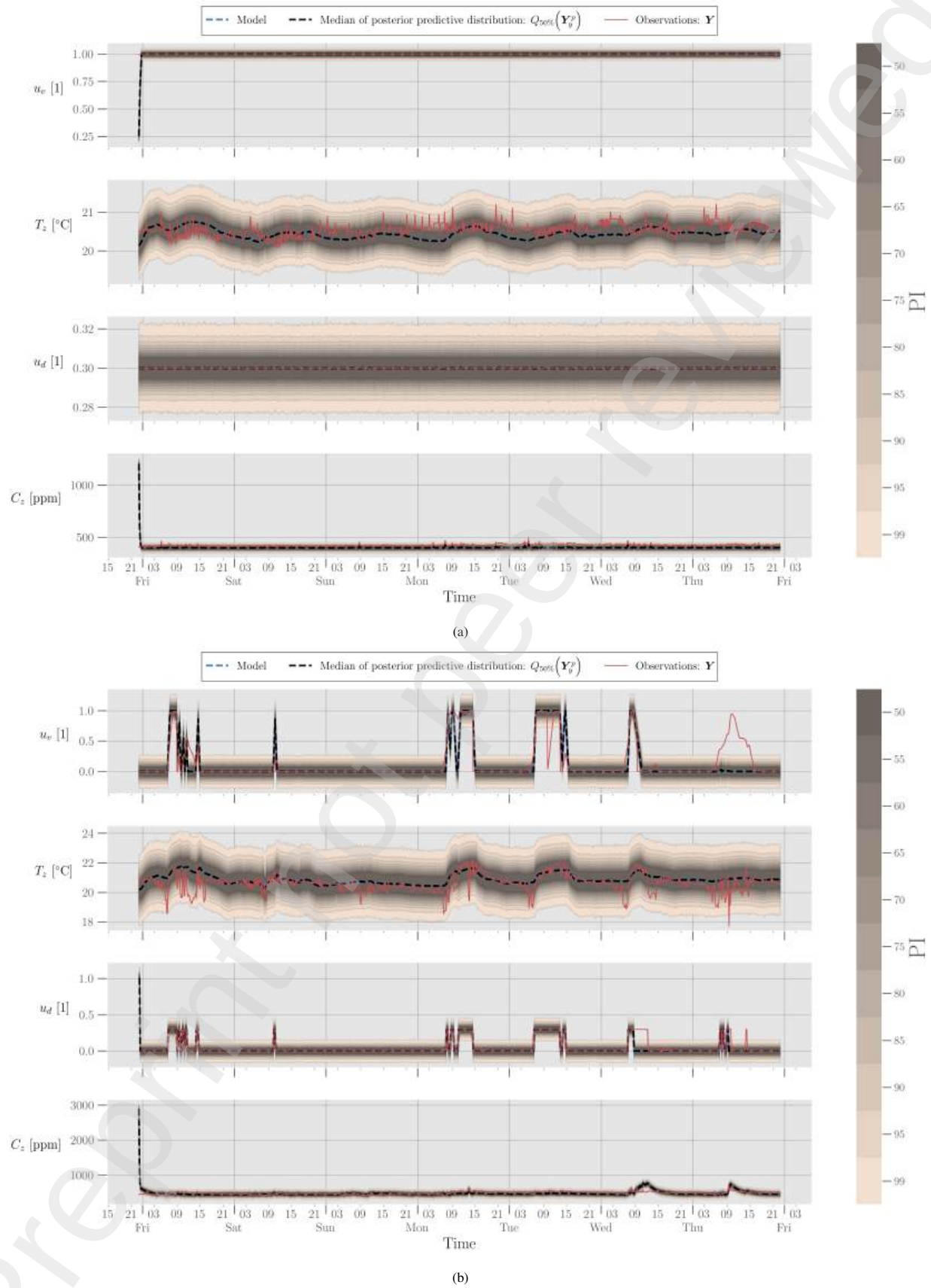


Figure 8: The median and PIs of the posterior predictive distributions compared with observations for a 1-week period of the test data. (a) space 4, a changing room. (b) space 5, an office.

building such as areas and volumes is used to calculate rough estimates.

5.1. Parameter sampling

A five-day winter period from December 2023 is used to sample the posterior distribution for the estimated model parameters. The data includes measurements from all sensing devices listed in Table 1 along with setpoint logs and weather data. To visualize and summarize the sampled posterior distribution, a corner plot is generated as shown in Figure 7. For improved readability, Figure 7 displays a 13 x 13 subset of the full 169 x 169 correlation matrix. This plot includes pairwise correlations between each sampled parameter, marginalized distributions along the matrix diagonal as histograms, and relevant quantile statistics. Pairwise correlation plots can help identify simple correlations between parameters. As shown, for most of the parameters, e.g. $\dot{m}_{inf}$ and R_{in} , there is no clear correlation. However, for other parameters such as C_w and C_a , there is a weak inverse correlation, meaning that an increase of C_w can be compensated by a decrease of C_a . The physical interpretation is that both parameters have roughly the same damping effect on the indoor temperature.

5.2. Posterior predictive distribution and inference

To assess the validity of the sampled posterior distribution, the posterior predictive distribution is calculated for the two-week period of the test data. Here, the Average Coverage Error (ACE) is used as a metric to evaluate the reliability of the resulting prediction intervals (PI). The ACE measures reliability by determining the difference between the actual and expected percentages of observations that fall within the specified PI [41]. Hence, if the AEC is negative, it indicates that fewer measurements than expected fall within the given PI range, and vice versa. Additionally, the root mean squared error (RMSE) and mean absolute error (MAE) of the median prediction are also calculated as metrics that can be compared with predictions from deterministic models. Using these metrics, the results are summarized in Table A.4. As shown, the MAE of the temperature point estimates ranges from 0.15°C for space 4 and 0.66°C for space 1. Space 1 is the largest zone with thermal coupling to all other zones on the floor, connected by a long corridor on each side. The more complex thermal dynamics and geometry combined with a higher occupancy variation makes the indoor temperature harder to predict, causing a lower accuracy. The indoor CO₂ predictions have MAE values ranging between 10 and 38 ppm for all spaces except space 12 with a MAE of ~133 ppm. For some of the rooms, the MAE for the MAE of valve and damper position predictions are 0 or close to 0, e.g. for space 4. For other rooms, e.g. space 5, the MAE is larger. To show the difference in dynamics causing this, the model predictions are presented for space 4 and 5, over a one-week winter period in Figure 8a and Figure 8b. The week shown is part of the test data utilized to evaluate the performance metrics in Table A.4. As shown, the median of the predictive distribution (represented by a black dashed line) aligns with the sensor data (depicted in red) and effectively captures the system dynamics for both rooms and all measuring devices. As shown in

Figure 8b, the temperature of space 5 fluctuates with the actuation of the dampers u_d and the space heater valve u_v . Most of discrepancies are caused by stochastic occupant behavior, e.g. opening of windows and doors. This causes sudden temperature drops and subsequent activation of the space heater, e.g. as shown Thursday at the end of the period.

On the other hand, there is much less dynamics for space 4. Here, the installed space heater is found to be greatly undersized as the heat delivered is insufficient for maintaining the indoor heating setpoint, which never drops below 21°C during the period shown. This causes the opening position of the space heater valve and dampers to be constantly open. As shown, this is well-captured by the model. For both spaces, the indoor CO₂ concentration is constant at 400 ppm with no significant occupancy.

The PIs of the posterior predictive distribution are also visualized in Figure 8a and Figure 8b as a colored area varying from dark grey at 50% to light brown at 99%. As shown, the majority of the measurements fall within the 99% PI. As validation of the PI accuracy, the ACE metric is presented for all measuring devices in Table A.4. For most measuring devices of space 4 and 5, the ACE_{95%} ranges between -3% to 5%, showing the reliability of the estimated PIs. For the valve position PIs of space 5, the ACE_{95%} is higher at -16%. This means that 16% fewer than expected observations fall within the PI.

6. Adaptive model generation and simulation

This section presents a concrete example demonstrating how changes to the semantic model directly and automatically affect simulation model predictions, illustrating the potential of closely integrating semantic and simulation models. This concept is used to explore alternative system configurations that can improve the indoor environment of the building. As noted earlier, the heater in space 4 is found to be greatly undersized and is constantly operated with a fully open valve. In addition, the room temperature never reaches the desired setpoint of 21 °C (see Figure 8a).

The investigated scenarios are listed in Table 3 where four options are compared, #1: the baseline, i.e. the current system configuration and operation, #2: the current system configuration with an increase of supply water temperature to 65 °C, #3: adding an additional heater of the same model to space 4, #4: adding an additional heater to space 4, but of a larger model, and finally, #5: A combination of #2 and #3 where both a space heater is added and the supply water temperature is increased. For each scenario, the model and specifications of the replaced or added heater is varied according to actual models available on the market [42] and are based on the EN442 standard [43]. The primary specifications considered are the output heat capacity and temperature classification of a space heater.

The structural differences between the scenarios are shown in Figure 9, where Figure 9a represents the baseline with one space heater in space 4 and Figure 9b represents scenarios where an additional space heater has been added. Note that the structures are shown in isolation and are still part of the full

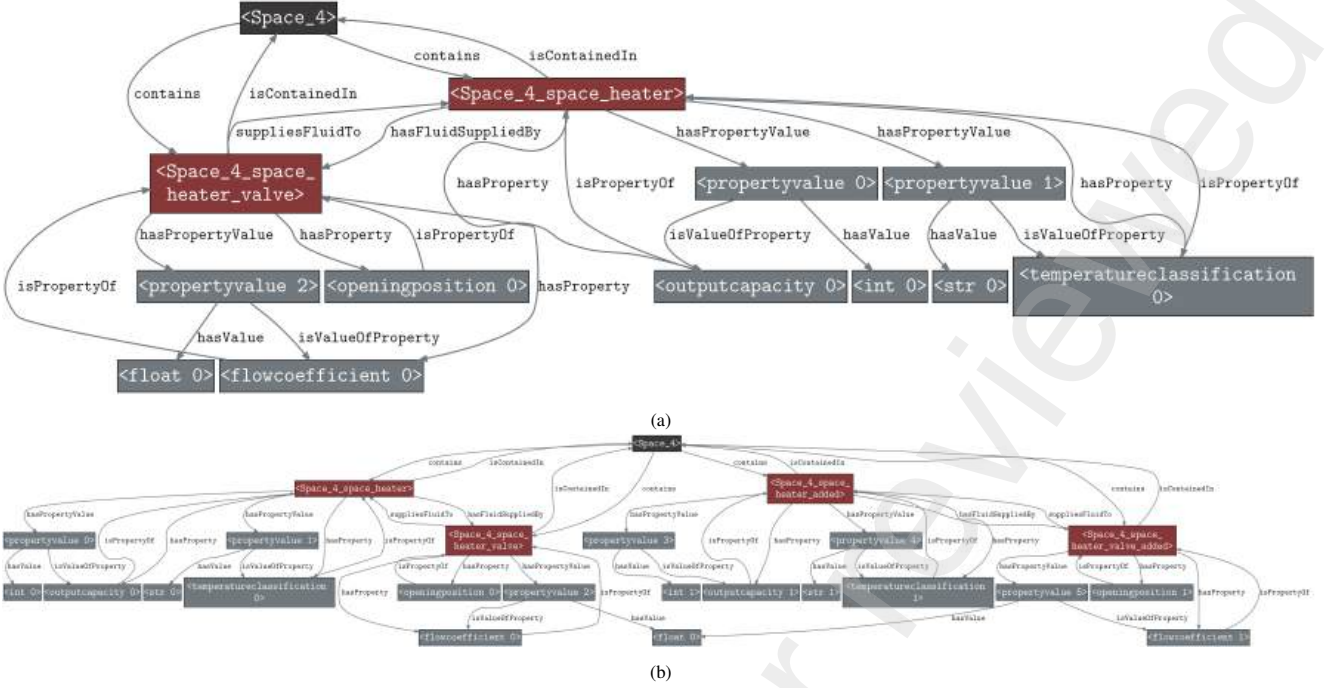


Figure 9: Structural differences in the semantic model between space 4 with one and two space heaters. Note that the structures are shown in isolation and are still part of the full semantic model of the case study shown earlier in Figure 3. (a) The baseline configuration of space 4 with one space heater (scenario #1 and #2). (b) An alternative configuration of space 4 with two space heaters (scenario #3, #4, #5).

Table 3: Scenarios to reduce heating setpoint violations in space 4. The specifications on the radiators reflect actual models available on the market [42] and are based on the EN442 standard [43].

Scenario	Radiator model	Instance	Value
#1: Baseline (Figure 9a)	1000x400 C21	<int 0> <str 0>	432 55/40/20
#2: Baseline, $T_{w, sup} = 65^\circ\text{C}$ (Figure 9a)	1000x400 C21	<int 0> <str 0>	432 55/40/20
#3: Add same (Figure 9b)	1000x400 C21 (left)	<int 0> <str 0>	432 55/40/20
	1000x400 C21 (right)	<int 1> <str 1>	432 55/40/20
#4: Add large (Figure 9b)	1000x400 C21 (left)	<int 0> <str 0>	432 55/40/20
	1000x400 C33 (right)	<int 1> <str 1>	747 55/40/20
#5: Add same, $T_{w, sup} = 65^\circ\text{C}$ (Figure 9b)	1000x400 C21 (left)	<int 0> <str 0>	432 55/40/20
	1000x400 C21 (right)	<int 1> <str 1>	432 55/40/20

semantic model of the case study shown earlier in Figure 3. Besides these structural differences, the information held by the different semantic instances is also varied, as indicated in Table 3. For each scenario, the simulation model is automatically generated based on the data in the semantic model. To evaluate the scenarios, the accumulated deviation from the heating setpoint, given by Equation 1, is used as a comparison metric

[30, 44, 45]. The metric captures both the intensity and duration of setpoint violations and has the unit Kelvin-hours.

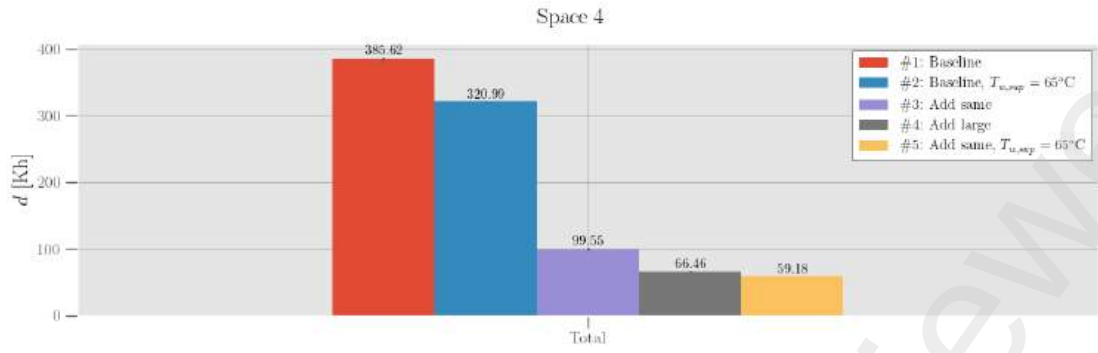
$$d = \int_0^T \max(T_{z, set, t} - T_{z, t}, 0) dt \quad (1)$$

In Equation 1, $T_{z, set, t}$ represents the indoor temperature heating setpoint while $T_{z, t}$ denotes the indoor temperature at time t .

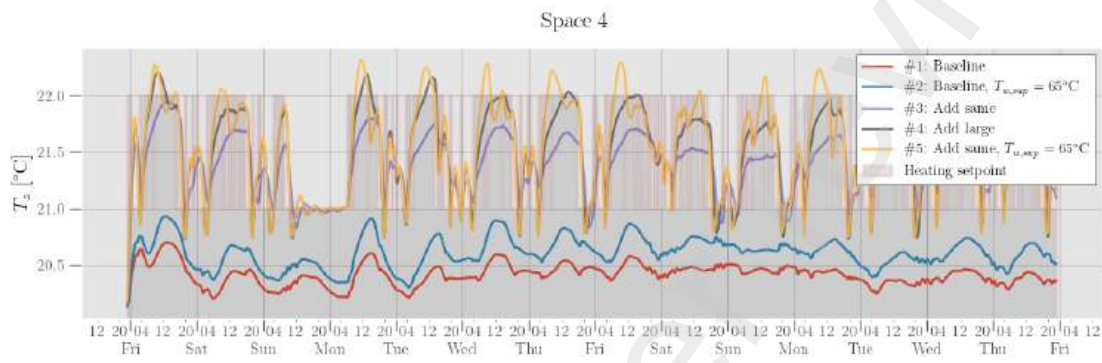
The same two-week testing period discussed in subsection 5.2 is also applied here to measure the impact of the suggested scenarios, using the historic heating setpoint signals for all spaces. The results are shown for space 4 in Figure 10, where Figure 10a shows the calculated discomfort and Figure 10b shows the temperature predictions for each scenario.

As shown in Figure 10a, all the alternative scenarios (#2, #3, #4, #5) reduce the indoor setpoint violations and improve the indoor comfort in comparison to scenario #1, the baseline, which has a setpoint violation of 385.62Kh. Scenario #2, with the increased supply water temperature, reduces the setpoint violation to 320.99Kh, a reduction of 17%. However, as shown in Figure 10b, the heating output of the currently installed space heater is still insufficient to meet the desired setpoint. Scenario #3 with an additional heater of 432W, shows a reduction of 74%, while scenario #4 with a larger space heater of 744W has an even further reduction of 83%. Finally, scenario #5 shows that an additional space heater of 432W combined with an increase of supply water temperature provides the largest reduction of 85%.

While the primary aim is to improve the indoor comfort of space 4, the different scenarios impact all considered spaces. For instance, consider the results for Space 6 (Figure 11), lo-



(a)

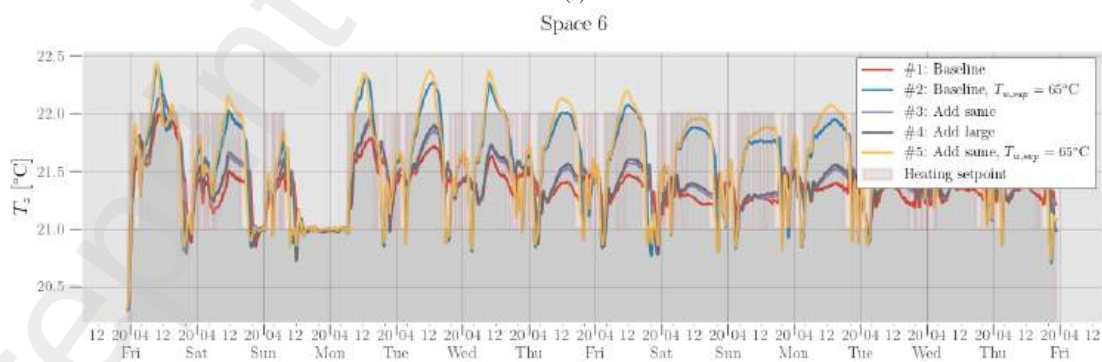


(b)

Figure 10: Impact of four scenarios on indoor comfort in Space 4 over a two-week period. **(a)** Accumulated setpoint violations in Kelvin-hours per scenario. **(b)** Predicted indoor temperatures compared to historical setpoint profile.



(a)



(b)

Figure 11: Impact of four scenarios on indoor comfort in Space 6 over a two-week period. **(a)** Accumulated setpoint violations in Kelvin-hours per scenario. **(b)** Predicted indoor temperatures compared to historical setpoint profile.

cated in the corridor opposite Space 4. As shown in [Figure 11a](#), the baseline setpoint violation for Space 6 is 124Kh, which means that the heating capacity currently installed is better sized compared to Space 4. Scenarios #2 and #5, involving increased supply water temperature, have the largest impact, reducing setpoint violations by approximately 50%.

The generated simulation model also accounts for thermal heat transfer between rooms, revealing secondary effects of the interventions. Scenarios #3 and #4, which focus on adding heaters to Space 4, also slightly improve the indoor comfort in Space 6, reducing setpoint violations by approximately 10%. This improvement is attributed to the additional heat added in Space 4, creating warmer interior boundary conditions for Space 6. The same effect is also observed for the remaining 10 rooms, although rooms closest to space 4 are affected the most. These observations align with fundamental heat transfer principles, further validating both the generated model and the overall framework.

In summary, significant improvements of the indoor comfort in space 4 can be achieved through simple measures such as installing additional space heater capacity and increasing the supply water temperature. The example presented in this section shows how different system configurations can be explored through changes in the semantic model only. As the semantic model represents the actual physical equipment and system topology of a building, a potential user only needs knowledge of the real system to test strategies. Hence, all abstractions specific to the simulation model (connections, inputs, outputs, parameters, etc.) are hidden away from the user. In comparison, performing such calculations in traditional energy simulation software such as EnergyPlus requires tool-specific expert knowledge, e.g. which parameters or inputs to change and how the different model abstractions maps to physical changes in the real building.

7. Limitations and future work

This paper demonstrated the potential of combining semantic modeling with data-driven component models to create a highly adaptable and interoperable energy modeling framework for building digital twins. The approach streamlines the process of generating and calibrating building energy models, and can be scaled to larger cases, as demonstrated through the considered hospital case study. Although the generated model was shown to provide accurate predictions across 45 measuring devices in the case study building, the approach must be further validated on other building types and HVAC system configurations to demonstrate its versatility.

One of the primary challenges in this study was the estimation of parameters due to the scale of the case study. Some manual initialization of the MCMC sampler was necessary, particularly for envelope parameters such as thermal capacitances and resistances. Here, high-level geometric data such as areas and volumes was used to calculate initial estimates. Therefore, to fully automate this process and allow further scaling, integration with additional ontologies such as the Building Topology Ontology (BOT) [\[46\]](#) should be explored.

Also, while this paper primarily addresses automated model generation and calibration, it lays the groundwork for broader advancements in building energy management. The automated generation of models presented in this work addresses the time-consuming and costly process of model development, a significant barrier to widespread implementation of model-based services such as MPC [\[9, 29, 30\]](#).

Based on these observations, promising directions for future research can be summarized as follows:

1. Consider additional case studies with different building types and HVAC system configurations to assess the adaptability of the framework and to further test its limitations.
2. Explore integration with additional ontologies, such as BOT, to access and utilize high-level geometric data for parameter initialization.
3. Integrating the framework with more sophisticated control strategies and optimization techniques to further improve building energy efficiency and performance.

By addressing these limitations and pursuing these future research directions, the proposed framework can be further improved, potentially leading to more widespread adoption of digital twins and advanced energy management strategies in the built environment.

8. Conclusion

This paper demonstrated the potential of combining semantic modeling with data-driven component models to create a highly adaptable and interoperable energy modeling framework for building digital twins. Through a large-scale field demonstration involving a hospital building, it was shown that this approach can significantly streamline and automate the process of generating and calibrating building energy models while maintaining high accuracy and flexibility. The key contributions of this work include:

1. The successful implementation of a semantic model as a comprehensive data source for automated model generation and calibration, reducing the effort required for modeling and ongoing maintenance.
2. Validation of seamless information transfer from semantic models to energy models, establishing a robust foundation for digital twin applications in the built environment.
3. Demonstration of scalability through a real-world case study involving 12 conditioned zones and 45 measuring devices, proving the approach's viability for complex building systems.
4. Practical application of the generated model to explore system modifications, showcasing the framework's potential for building performance optimization.

The predictions of the calibrated model achieved on average mean absolute errors of 0.4 °C for temperature, 32 ppm for CO₂ concentration, 0.06 for valve position, and 0.04 for

damper positions across various zones, demonstrating its accuracy and reliability. Moreover, the ability to easily test alternative system configurations by modifying only the semantic model highlights the framework's potential for rapid scenario analysis and decision support. Future work should explore further integration with other ontologies that support high-level geometric data representations as well as the development of advanced services built on top of the generated and calibrated models. Additionally, exploring the application of this framework to a wider range of building types and systems will help to further validate its generalizability and impact. In conclusion, this research represents a significant step forward in the development of interoperable and adaptable energy models for more widespread adoption of advanced building energy management strategies. This has the potential to improve both building energy efficiency and occupant comfort of buildings.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is carried out under the 'Twin4Build: A holistic Digital Twin platform for decision-making support over the whole building life cycle' project, funded by the Danish Energy Agency under the Energy Technology Development and Demonstration Program (EUDP), ID number: 64021-1009.

Author statement

Jakob Bjørnskov: Conceptualization, Methodology, Software, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **August Thomsen:** Software, Investigation, Writing - original draft, Visualization. **Muhyiddine Jradi:** Writing - review & editing, Supervision, Funding acquisition.

Appendix A.

Table A.4: Performance metrics of the calibrated model including assessment of the PIs 95%-75% for two weeks of test data.

ID	MAE	RMSE	ACE _{95%}	ACE _{90%}	ACE _{85%}	ACE _{80%}	ACE _{75%}
Temperature sensors	[°C]	[°C]	[%]	[%]	[%]	[%]	[%]
1	0.66	0.84	-21	-24	-26	-25	-25
3	0.57	0.7	-5	-8	-11	-11	-12
4	0.15	0.19	4	8	12	14	17
5	0.44	0.63	2	4	8	11	15
6	0.29	0.36	-10	-16	-16	-16	-15
7	0.29	0.4	-7	-6	-5	-4	-3
8	0.44	0.5	0	1	-1	-4	-9
9	0.42	0.49	-1	-2	-1	-2	-4
10	0.3	0.37	4	7	10	12	15
11	0.35	0.44	-1	-3	-4	-5	-4
12	0.45	0.6	-5	-3	-3	-3	-3
Absolute Mean	0.40	0.50	5	7	9	10	11
CO2 sensors	[ppm]	[ppm]	[%]	[%]	[%]	[%]	[%]
1	10.81	20.34	-2	1	5	8	12
3	22.57	54.8	4	9	13	18	23
4	18.38	27.79	-3	-8	-11	-13	-14
5	26.45	66.5	2	6	11	14	19
6	20.67	23.52	-50	-54	-53	-51	-50
7	38.0	67.5	-6	-4	-2	1	5
8	22.66	42.32	-10	-7	-4	0	4
9	10.02	27.77	5	10	15	20	25
10	33.12	72.36	1	5	8	11	15
11	19.58	41.37	4	8	12	15	18
12	132.56	178.23	-23	-25	-27	-27	-25
Absolute Mean	32	57	10	12	15	16	19
Valve position sensors	[1]	[1]	[%]	[%]	[%]	[%]	[%]
1	0.03	0.07	-50	-45	-40	-35	-30
3	0.12	0.25	-15	-11	-7	-3	1
4	0.0	0.03	5	10	15	20	25
5	0.12	0.26	-16	-11	-7	-2	2
6	0.13	0.21	-10	-12	-11	-9	-6
7	0.01	0.04	0	5	10	15	19
8	0.11	0.15	-1	1	4	7	8
9	0.03	0.1	3	8	12	17	22
10	0.0	0.0	-40	-35	-30	-25	-20
11	0.04	0.11	-1	3	8	12	17
12	0.08	0.18	1	5	9	13	17
Absolute Mean	0.06	0.13	13	13	14	14	15
Damper position sensors	[1]	[1]	[%]	[%]	[%]	[%]	[%]
1	0.0	0.01	5	10	15	20	25
2	0.0	0.02	5	10	15	20	25
3	0.01	0.05	3	8	13	18	23
4	0.0	0.01	5	10	15	20	25
5	0.02	0.07	0	5	10	15	20
6	0.0	0.02	5	10	15	20	25
7	0.09	0.16	0	-10	-10	-6	-1
8	0.1	0.17	-28	-23	-18	-13	-8
9	0.08	0.15	-21	-16	-11	-6	-1
10	0.08	0.15	-6	-10	-6	-2	2
11	0.01	0.04	4	8	13	18	23
12	0.07	0.16	-15	-10	-5	0	5
Absolute Mean	0.04	0.08	8	11	13	14	15

References

- [1] M. Grieves, Digital twin: Manufacturing excellence through virtual factory replication (03 2015).
- [2] C. Boje, A. Guerriero, S. Kubicki, Y. Rezgui, [Towards a semantic construction digital twin: Directions for future research](#), *Automation in Construction* 114 (2020) 103179. doi:<https://doi.org/10.1016/j.autcon.2020.103179>
URL <https://www.sciencedirect.com/science/article/pii/S0926580519314785>
- [3] H. Boyes, T. Watson, [Digital twins: An analysis framework and open issues](#), *Computers in Industry* 143 (2022) 103763. doi:<https://doi.org/10.1016/j.compind.2022.103763>
URL <https://www.sciencedirect.com/science/article/pii/S0166361522001609>
- [4] J. Cao, T. Maile, J. O'Donnell, R. Wimmer, C. Treeck, Model transformation from simmodel to modelica for building energy performance simulation, 2014.
- [5] J. Cao, R. Wimmer, M. Thorade, T. Maile, J. O'Donnell, J. Rädler, J. Frisch, C. Treeck, A flexible model transformation to link bim with different modelica libraries for building energy performance simulation, 2015. doi:10.26868/25222708.2015.2446
- [6] R. Wimmer, T. Maile, J. O'Donnell, J. Cao, C. Treeck, Data-requirements specification to support bim-based hvac definitions in modelica, 2014.
- [7] R. Wimmer, J. Cao, P. Remmen, T. Maile, J. O'Donnell, J. Frisch, R. Streblow, D. Mueller, C. Treeck, Implementation of advanced bim-based mapping rules for automated conversions to modelica, 2015. doi:10.26868/25222708.2015.2424
- [8] A. Andriamamonjy, D. Saelens, R. Klein, [An automated ifc-based workflow for building energy performance simulation with modelica](#), *Automation in Construction* 91 (2018) 166–181. doi:<https://doi.org/10.1016/j.autcon.2018.03.019>
URL <https://www.sciencedirect.com/science/article/pii/S0926580517308282>
- [9] T. Hong, J. Langevin, K. Sun, Building simulation: Ten challenges, *Building Simulation* 11 (03 2018). doi:10.1007/s12273-018-0444-x
- [10] M. Wetter, M. Sulzer, [A call to action for building energy system modelling in the age of decarbonization](#), *Journal of Building Performance Simulation* 17 (3) (2024) 383–393. doi:10.1080/19401493.2023.2285824
URL <https://doi.org/10.1080/19401493.2023.2285824>
- [11] S. Tang, D. R. Shelden, C. M. Eastman, P. Pishdad-Bozorgi, X. Gao, [A review of building information modeling \(bim\) and the internet of things \(iot\) devices integration: Present status and future trends](#), *Automation in Construction* 101 (2019) 127–139. doi:<https://doi.org/10.1016/j.autcon.2019.01.020>
URL <https://www.sciencedirect.com/science/article/pii/S0926580518305764>
- [12] D. G. J. Opoku, S. Perera, R. Osei-Kyei, M. Rashidi, K. Bamdad, T. Famakinwa, [Digital twin for indoor condition monitoring in living labs: University library case study](#), *Automation in Construction* 157 (2024). doi:10.1016/j.autcon.2023.105188
URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85176097493&doi=10.1016%2fj.autcon.2023.105188&partnerID=40&md5=c652c487a60ef0a3148d6c19fccae38>
<https://www.sciencedirect.com/science/article/pii/S092658052300448X?via%3Dihub>
- [13] Y. Peng, M. Zhang, F. Yu, J. Xu, S. Gao, [Digital twin hospital buildings: An exemplary case study through continuous lifecycle integration](#), *Advances in Civil Engineering* 2020 (2020). doi:10.1155/2020/8846667
URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85097320262&doi=10.1155/2f2020/2f8846667&partnerID=40&md5=ab753b9d468b135bfe571873deb26874>
<https://onlinelibrary.wiley.com/doi/pdfdirect/10.1155/2020/8846667?download=true>
- [14] P. Pauwels, S. Zhang, Y.-C. Lee, Semantic web technologies in aec industry: A literature overview, *Automation in construction* 73 (2017) 145–165.
- [15] P. Barnaghi, W. Wang, C. Henson, K. Taylor, [Semantics for the internet of things: Early progress and back to the future](#), *International Journal on Semantic Web and Information Systems* 8 (1) (2012) 1–21, cited by: 436. doi:10.4018/jswis.2012010101
URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-84870318243&doi=10.4018%2fjswis.2012010101&partnerID=40&md5=c77e3fc55be04eed4fc0847aef1c69c>
- [16] S. Hu, E. Corry, E. Curry, W. J. Turner, J. O'Donnell, [Building performance optimisation: A hybrid architecture for the integration of contextual information and time-series data](#), *Automation in Construction* 70 (2016) 51–61. doi:<https://doi.org/10.1016/j.autcon.2016.05.018>
URL <https://www.sciencedirect.com/science/article/pii/S092658051630098X>
- [17] E. Corry, P. Pauwels, S. Hu, M. Keane, J. O'Donnell, [A performance assessment ontology for the environmental and energy management of buildings](#), *Automation in Construction* 57 (2015) 249–259. doi:<https://doi.org/10.1016/j.autcon.2015.05.002>
URL <https://www.sciencedirect.com/science/article/pii/S0926580515001041>
- [18] J. Merino, X. Xie, N. Moretti, J. Y. Chang, A. Parlikad, [Data integration for digital twins in the built environment](#), *Proceedings of the European Conference on Computing in Construction*, 2022, pp. 199–206.
URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85143710436&doi=10.35490%2fEC3.2022.172&partnerID=40&md5=cd72e96b4f1e66b96397230e395b5eb0>
https://ec-3.org/publications/conferences/EC32022/papers/EC32022_172.pdf
- [19] K. El Mokhtari, I. Panushev, J. J. McArthur, [Development of a cognitive digital twin for building management and operations](#), *Frontiers in Built Environment* 8 (2022). doi:10.3389/fbuil.2022.856873
URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85139394134&doi=10.3389/2ffbuil.2022.856873&partnerID=40&md5=e14857ed95c47e14fb60c45ee76d070c>
<https://www.frontiersin.org/journals/built-environment/articles/10.3389/fbuil.2022.856873/pdf>
- [20] Z. Ni, Y. Liu, M. Karlsson, S. Gong, Enabling preventive conservation of historic buildings through cloud-based digital twins: A case study in the city theatre, norrköping, *IEEE Access* 10 (2022) 90924–90939. doi:10.1109/ACCESS.2022.3202181
- [21] Microsoft Azure, [Azure digital twins](#), [Online]. Available:
URL <https://azure.microsoft.com/en-us/products/digital-twins/>
- [22] Z. Ni, C. Zhang, M. Karlsson, S. Gong, [Leveraging deep learning and digital twins to improve energy performance of buildings](#), 2023 IEEE 3rd International Conference on Industrial Electronics for Sustainable Energy Systems, IESSES 2023, 2023. doi:10.1109/IESSES53571.2023.10253721
URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85174266992&doi=10.1109/2fIESSES53571.2023.10253721&partnerID=40&md5=84d52d5bea3a16c8245ca5749e3eedaf7>
<https://ieeexplore.ieee.org/document/10253721/>
- [23] D. Crawley, L. Lawrie, F. Winkelmann, W. Buhl, Y. Huang, C. Pedersen, R. Strand, R. Liesen, D. Fisher, M. Witte, J. Glazer, Energyplus: creating a new-generation building energy simulation program, *Energy and Buildings* 33 319–331..
- [24] W. Kritzing, M. Karner, G. Traar, J. Henjes, W. Sihn, [Digital twin in manufacturing: A categorical literature review and classification](#), *IFAC-PapersOnLine* 51 (11) (2018) 1016–1022, 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018. doi:<https://doi.org/10.1016/j.ifacol.2018.08.474>
URL <https://www.sciencedirect.com/science/article/pii/S2405896318316021>
- [25] A. Afram, F. Janabi-Sharifi, [Supervisory model predictive controller \(mpc\) for residential hvac systems: Implementation and experimentation on archetype sustainable house in toronto](#), *Energy and Buildings* 154 (2017) 268–282. doi:<https://doi.org/10.1016/j.enbuild.2017.08.060>
URL <https://www.sciencedirect.com/science/article/pii/S0378778817301743>

- [26] A. Afram, F. Janabi-Sharifi, [Black-box modeling of residential hvac system and comparison of gray-box and black-box modeling methods](#), *Energy and Buildings* 94 (2015) 121–149. doi:<https://doi.org/10.1016/j.enbuild.2015.02.045>
URL <https://www.sciencedirect.com/science/article/pii/S0378778815001504>
- [27] Y. Li, Z. O'Neill, L. Zhang, J. Chen, P. Im, J. DeGraw, [Grey-box modeling and application for building energy simulations - a critical review](#), *Renewable and Sustainable Energy Reviews* 146 (2021) 111174. doi:<https://doi.org/10.1016/j.rser.2021.111174>
URL <https://www.sciencedirect.com/science/article/pii/S1364032121004639>
- [28] C. Rudin, [Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead](#), *Nature Machine Intelligence* 1 (2019) 206–215. doi:<https://doi.org/10.1038/s42256-019-0048-x>
- [29] J. Cígler, D. Gyalistras, J. Široky, V. Tiet, L. Ferkl, [Beyond theory: the challenge of implementing model predictive control in buildings](#), in: *Proceedings of 11th Rehva world congress, Clima*, Vol. 250, 2013.
- [30] D. Blum, Z. Wang, C. Weyandt, D. Kim, M. Wetter, T. Hong, M. A. Piette, [Field demonstration and implementation analysis of model predictive control in an office hvac system](#), *Applied Energy* 318 (2022) 119104. doi:<https://doi.org/10.1016/j.apenergy.2022.119104>
URL <https://www.sciencedirect.com/science/article/pii/S0306261922004895>
- [31] J. Bjørnskov, M. Jradi, M. Wetter, [Automated model generation and parameter estimation of building energy models using an ontology-based framework](#), *Energy and Buildings* Submitted (2024). doi:<https://dx.doi.org/10.2139/ssrn.4884154>
- [32] J. Bjørnskov, M. Jradi, [An ontology-based innovative energy modeling framework for scalable and adaptable building digital twins](#), *Energy and Buildings* 292 (Aug. 2023). doi:<https://doi.org/10.2139/ssrn.4341936>
- [33] J. Bjørnskov, [Ontology-based building modeling framework, 2023](#), <https://github.com/JBjoernskov/Twin4Build>. Last Accessed: 03-23-2023.
- [34] L. Daniele, F. den Hartog, J. Roes, [Created in close interaction with the industry: The smart appliances reference \(saref\) ontology](#), 2015, pp. 100–112. doi:https://doi.org/10.1007/978-3-319-21545-7_9
- [35] M. Laura Daniele, R. Garcia-Castro, M. Lefrançois, M. Poveda-Villalon, [Saref: the smart applications reference ontology](#), [Online]. Available: <https://saref.etsi.org/core/>
- [36] M. Poveda-Villalón, R. Garcia-Castro, [Saref extension for building](#), [Online]. Available: <https://saref.etsi.org/saref4bldg/>
- [37] M. Lefrançois, [Saref4syst: an extension of saref for typology of systems and their inter-connections](#), [Online]. Available: <https://saref.etsi.org/saref4syst/>
- [38] V. Kukkonen, A. Küçükavcı, M. Seidenschur, M. H. Rasmussen, K. M. Smith, C. A. Hviid, [An ontology to support flow system descriptions from design to operation of buildings](#), *Automation in Construction* 134 (2022) 104067. doi:<https://doi.org/10.1016/j.autcon.2021.104067>
URL <https://www.sciencedirect.com/science/article/pii/S0926580521005185>
- [39] Q. LuVivi, P. Kumar, WoodallPhilip, R. Don, HeatonJames, [Developing a dynamic digital twin at a building level: using cambridge campus as case study](#), in: *International Conference on Smart Infrastructure and Construction 2019 (ICSIC)*, pp. 67–75. arXiv:<https://www.icevirtuallibrary.com/doi/pdf/10.1680/icsic.64669.067>, doi:<https://doi.org/10.1680/icsic.64669.067>
URL <https://www.icevirtuallibrary.com/doi/abs/10.1680/icsic.64669.067>
- [40] DMI Open Data: Meteorological Observation Data.
- [41] P. Pinson, G. Kariniotakis, H. Nielsen, T. Nielsen, H. Madsen, [Properties of Quantile and Interval Forecasts of Wind Generation and Their Evaluation](#), Vol. 2, 2006.
- [42] Purmo, [Panels, purmo compact](#), [Online]. Available: <https://www.purmo.com/da-dk/produkter/opvarmning-og-koeling/radiatorer/panelradiator/panelradiatorer/compact>
- [43] European Standard, [En 442. specification for radiators and convectors](#).
- [44] J. Bjørnskov, M. Jradi, [Implementation and demonstration of an automated energy modeling framework for scalable and adaptable building digital twins based on the saref ontology](#), Vol. 18, *Tsinghua University Press*, 2023, pp. 3083–3090, 18th IBPSA Conference on Building Simulation, BS 2023 ; Conference date: 04-09-2023 Through 06-09-2023. doi:<https://doi.org/10.26868/25222708.2023.1595>
- [45] K. Arendt, A. Clausen, C. Mattera, M. Jradi, A. Johansen, C. Veje, M. Kjærgaard, B. Jørgensen, [Multi-objective model predictive control framework for buildings](#), in: *Proceedings of building simulation 2019, 2020, rome (Italy)*, 02-04 September 2019.
- [46] M. H. Rasmussen, M. Lefrançois, G. Schneider, P. Pauwels, [Bot: the building topology ontology of the w3c linked building data group](#), *Semantic Web* (11 2020). doi:<https://doi.org/10.3233/SW-200385>