Ministry of Environment
of Denmark
Environmental
Protection Agency

THE STANFORD GROUNDWATER ARCHITECTURE PROJECT

Utilizing Advanced Geophysical and Computational Methods for the Development of Hydrogeologic Conceptual Models

MUDP Report

September 2021

Udgiver: Miljøstyrelsen

Redaktion:
Rosemary Knight, Stanford University
Anders Vest Christiansen, Aarhus University
Max Halkjær, Rambøll
Niels-Peter Jensen, I•GIS
Ted Asch, Aqua Geo Frameworks

ISBN: 978-87-7038-346-2

Indhold

1. Abstract

An ambitious two-year project was conducted to develop, for application in California, U.S.A., the optimal workflow for utilizing advanced geophysical and computational methods in the construction of a hydrogeologic conceptual model. We partnered with three water agencies: The Butte County Department of Water and Resource Conservation, in the Sacramento Valley, the Indian Wells Valley Water District, located just east of the Sierra Nevada in the southern part of the state, and San Luis Obispo County, on the central California Coast. All three water agencies are faced with the challenge of developing groundwater sustainability plans in the next three to five years and had all recognized the tremendous potential value of the airborne electromagnetic (AEM) method in developing the conceptual model, the foundation for much of the groundwater sustainability plan.

One of the first steps in all of the study areas was to set up the data management system (DMS). All project data were stored and shared through the use of this DMS. Approximately 800 kilometers of AEM data were acquired in each of the study areas and processed for use in this project. Novel forms of data analysis were developed and applied for the inversion of the data so as to recover models of the electrical resistivity of the subsurface. In general, in the three areas, we were able to determine resistivity to depths of approximately 300 m. Integrating well data, digitized and available through the DMS, novel methods were developed to transform the derived electrical resistivity models to models displaying sediment type. These models were then used, along with well data and all other available data/information, to build the hydrogeologic conceptual model using, what we refer to as, the more standard approach and a geostatistical approach based on the multi-point statistics (MPS) method. Through the project we were able to demonstrate the significant value in the adoption of an MPS method in that it can be used to obtain information about spatial heterogeneity at a finer scale than can be resolved with the AEM data. An approach was developed that can be used to quantify uncertainty in the final models of sediment type and to communicate to end-users the level of confidence that can be placed in various components of the developed conceptual model.

Throughout this project, we benefited from close collaboration with those who have the local knowledge and an understanding of the information needed to support decision-making related to groundwater management. In order to ensure that the acquisition of AEM data results in a product that is valued by the local water agency, this collaboration needs to drive, and be present throughout, the workflow for building hydrogeologic conceptual models from AEM data.

2. Preamble and acknowledgments

This is the final report for the Groundwater Architecture Project (GAP). An appropriate way to begin this final report, is with the closing lines from our proposal:

The passage of groundwater legislation in California has provided the regulatory framework required for sustainable groundwater management. The key challenge is obtaining the subsurface data required for SGMA implementation. In this project we focus specifically on the need to develop a hydrogeologic conceptual model, and on the use of the airborne electromagnetic (AEM) method to meet the data needs. By conducting three studies, we will develop the optimal workflow that will encourage and support the adoption of the AEM method, and related data management/modeling methods, as central to model development. We will address the critical need to provide an organizational structure and quality control for compiling well data, noting that this step in the workflow is not only essential for working with AEM data, but also for other SGMA-related data needs. Building on lessons learned in Denmark, but working in three different locations in California, we can adopt best practices, and adapt as needed. The result will be an AEM-centered workflow that is optimized for widespread use in California; providing water agencies with an unprecedented ability to acquire the data needed to develop hydrogeologic conceptual models. Such models are the foundation for groundwater modeling, which in turn is the foundation for sustainable groundwater management.

Over the two years of the GAP, we have defined an optimal workflow for the use of AEM data, along with advanced computational methods, to develop hydrogeologic conceptual models for the groundwater basins of California. We can already see that our efforts are encouraging, and supporting, the adoption of the AEM method, and related data management/modeling methods. We are grateful to have had the opportunity to contribute in this way to sustainable groundwater management in California.

The project leads of our research team were Rosemary Knight (GAP lead), Professor of Geophysics at Stanford University; Esben Auken, Professor in Hydro Geophysics at Aarhus University, replaced in September 2020 by Anders Vest Christiansen; Max Halkjær, Ramboll; Niels-Peter Jensen, I•GIS; and Jim Cannia, Aqua Geo Frameworks (AGF), replaced in August 2020 by Ted Asch following Jim's tragic death. Other members of the research team were Seogi Kang and Noah Dewar from Stanford University; Jesper Bjergsted Pedersen from Aarhus University; Paul Thorn, Tillie M Madsen, and Peter Thomsen from Ramboll; and Thomas Bager Rasmussen, Mats Lundh Gulbrandsen, and Le Thanh Vu from I•GIS.

Our project partners in the three study areas were: Paul Gosselin and Christina Buck from the Butte County Department of Water and Resource Conservation, Don Zbeda from the Indian Wells Valley Water District and Catherine Martin, Angela Ruberto, and Carolyn Berg from San Luis Obispo County Public Works. This project could not have happened without their enthusiastic participation, and the time they committed to ensuring its successful completion.

At the state level we wish to thank the following individuals for their support and input: Tim Godwin, Katherine Dlubac, Ben Brezing, and Stephen Springhorn, from the Department of Water Resources; and Natalie Stork and Sam Boland-Brien from the State Water Resources Control Board.

We benefited tremendously from interaction with many other individuals. Our collaborators included Jared Abraham and Ted Asch from AGF; Tim Parker from Parker Groundwater; Graham Fogg, Professor of Hydrogeology at U.C. Davis; Todd Greene, Professor of Geological Sciences at California State University, Chico; Sharla Stockton, Water Resource Specialist with Glenn County; Jef Caers, Professor at Stanford University; David Shimabukuro, Professor of Geology at California State University, Sacramento. We also wish to thank Paul Sorensen and Jeff Barry from GSI from their assistance with the work in San Luis Obispo. We are very grateful to Tobias Kvorning and Jesper Hannibalsen from the Danish EPA (MUDP) for their support throughout this project, to Lauren Nicholson at Stanford for the many hours of administrative assistance, and to Lasse Thomassen with the Water Technology Alliance for his interest in the project. And finally, the GAP came together because of the dedication and high level of commitment of Jacob Vind, Danish Ministry of Foreign Affairs and the Danish Water Technology Alliance, who worked between with many individuals, in California and in Denmark, to engage the team members and obtain the funding to launch this ambitious project.

In addition to funding from MUDP, matching funding for this project was received from the California Department of Water Resources and the State Water Resources Control Board. Funding for the larger research project, which provided the data and supported other activities that contributed to the research reported here, was received from the three partner water agencies. Stanford research activities received additional funding from the Gordon and Betty Moore Foundation (grant GBMF6189).

3. Project Description

3.1 Purpose

California has embarked on a historic journey to achieve groundwater sustainability with the passage of the Sustainable Groundwater Management Act (SGMA) by the California Legislature in 2014. Local agencies are vested with the responsibility for achieving sustainability, with the first step for each agency being the preparation of a groundwater sustainability plan (GSP). For the 127 medium and high priority basins in California, these plans are due in 2020 or 2022. A specific legislative requirement is the development of a hydrogeologic conceptual model (GSP Regulation §354.14); such a model provides the information about the subsurface architecture needed for generating a groundwater model and for quantifying the water balance. Historically the "conceptual model" was often a very general description of subsurface layering. The "conceptual model" required by SGMA is a 3D hydrostratigraphic model capturing the spatial heterogeneity of the subsurface needed as the input for flow modeling. Many agencies, however, lack sufficient information about the subsurface to develop such a detailed 3D hydrogeologic conceptual model with an acceptable level of uncertainty; uncertainty in the conceptual model propagates through to uncertainty in the decision-making required to achieve sustainability. There is widespread recognition of the need to acquire more subsurface data so as to reduce the uncertainty in the conceptual models. But the currently deployed, traditional methods of characterizing aquifers through the drilling of wells with testing and logging are slow, expensive and insufficient in terms of data coverage.

The same challenges facing California today were faced by Denmark in the early 1990s when groundwater legislation was passed that required all municipalities to characterize and manage the groundwater systems. What emerged – a geophysical airborne electromagnetic (AEM) method that was found to be the most cost-effective way to get the subsurface coverage needed for compliance with the regulatory requirement; a significant benefit being that there is no need for land access for data acquisition. AEM data are acquired using a helicopter system that moves a geophysical instrument 40 m above the ground, covering approximately 300 kilometers a day, imaging subsurface properties to a depth of ~400 m. The acquisition of AEM data throughout California could play a central role in providing critical information to inform the development of hydrogeologic conceptual models, that can then be used for the development of groundwater models, to guide recharge efforts, to assess geologic controls on observed subsidence, and to aid in the siting of monitoring wells. The critical issue, and the motivation for this project: how to take this on in a state where there has been no history of using the AEM method for groundwater management?

In our ambitious two-year project, we partnered with three water agencies in California to develop a template for the optimal workflow that would use AEM data as the foundation for the development of a hydrogeologic conceptual model; a key step in the implementation of SGMA. This included not only the deployment of the AEM technology to acquire AEM data, but also designing the supporting computational infrastructure for data analysis, interpretation, and archiving. Significant advancements can be made by studying ways in which we can implement, in California, a workflow that builds on the Danish experience. Our approach focused on lessons learned in Denmark, identifying transferable knowledge and technologies. Our project also involved basic research to discover new methods of data analysis, inversion, and interpretation appropriate for the specific geologic environment and management needs of California.

Three local water agencies in California worked closely with us in this project. The Butte County Department of Water and Resource Conservation is in the Sacramento Valley. Butte County has areas dependent upon groundwater and an interest in sustainable groundwater management projects. This agency brought to the project in-house expertise in hydrology, and an existing groundwater model. The Indian Wells Valley Water District (IWV) is located in southern part of the state, just east of the Sierra Nevada and relies exclusively on groundwater for all water supply needs. Groundwater levels have been steadily declining since at least the 1950s, with the area now categorized as one of 21 "Critically Overdrafted Basins" within the state. San Luis Obispo (SLO) County is on the central California Coast, and is also grappling with a dependence on groundwater and concerns about saltwater intrusion. Associated with each water agency is a large group of interested stakeholders, including members of the public and government officials at the local and county level. All three water agencies are currently facing the challenge of developing, revising and/or implementing groundwater sustainability plans, and all recognized the tremendous potential value of AEM data in developing the conceptual model, the foundation for much of the groundwater sustainability plan.

The MUDP-funded project was part of a larger project, made up of ten work packages that allowed us to develop the optimal workflow for the use, in California, of AEM data to develop a hydrogeologic conceptual model. In the MUDP project, we included four work packages, that represent the research and development components of the ten in the larger project. The main body of this report contains descriptions of these four work packages. The deliverable from the larger project, the set of recommendations describing the optimal workflow, is presented in the Appendix.

3.2 Assessment of market potential

The GAP project has created the foundation for a large-scale procurement from the State of California Department of Water Resources (DWR) for mapping of the Californian geology in all high and medium priority groundwater basins. The first large scale project was tendered out in autumn 2020 requesting use of advanced AEM technology similar to SkyTEM.

The project has not led to any patents or patent applications, but Stanford University and Aarhus University have progressed in the innovation within the AEM technology making the procedures and standards applicable to application in the US. I•GIS, through the project, has pursued the development of new modeling techniques. The Multiple Points Statistics (MPS) algorithms that have been implemented in this project are open source and for the moment I•GIS does not consider software patents to be a viable option, as it is often quite easy to surpass these in some way. But for a typical user to be able to use these algorithms in a good stable workflow has been considered of value to potential customers, so has been the approach taken by I•GIS in this project.

During the GAP Danish as well as US companies have started delivering AEM services to the Californian water sector both county-wide and state-wide. The Danish SMEs involved in the GAP project, SkyTEM, have provided AEM surveys on several projects (total of ca 3 million DKK) resulting in increased employment in Denmark (DK share of value 80-100%) and the Americas (US share of value 0-20%).

Based on the sales pipeline, especially SkyTEM expects the strong growth in California to continue for at least the next 3-5 years. The Danish consultancy Ramboll has provided consultancy services within geophysics for more than 4 million DKK over the last 3 years with annual 150% growth rates, implying significantly increasing employment in Denmark (share of value: 50%) as well as in California, US (share of value: ca 50%). Based on the sales pipeline, Ramboll expects the exponential growth in California within geophysics to continue for at least the next 3-5 years. Finally, Stanford and Aarhus University have increased their research within

geophysics and further developed the existing AEM tool box likewise creating geophysical research related jobs in the US as well as Denmark.

During the GAP project period I•GIS has only had very limited success in the Californian market. The primary reason for this is that the adoption of the geophysical mapping approach is still to come. The current geophysical mapping is being carried out by the front-runners Rambøll and AGF who are already customers of I•GIS and the market is still awaiting the general roll out of the mapping approach before investing in technology and upgrading personnel to handle this. Thus, it is the expectation that, when airborne geophysics starts to be incorporated generally in CA, the customer base should be broadened for the services and software from I•GIS as seen in other parts of the world. If I•GIS can hold a reasonable market share in California, the marked is estimated to a sale of 10 mio dkr and an ongoing subscription sale of 3 mio/year. This would result in a further increase in employees of 3-5 persons (FTE), predominantly situated in Denmark.

4. Project Deliverables

The project was defined with four deliverables, listed below and each described in the following sections.

WP	Deliverables
1. Project management and dissemination of results	A system for project management, allowing for the flow of information among all project members, tracking of progress, revision of timelines as needed, and dissemination of results. Communication between project team partners, external partners, and the broader California community is included in this step.
2. Geophysics to hydrogeology transform	A methodology that can extract from the acquired data, the desired subsurface information, e.g. mapping of aquifer and aquitard units, and any structural features. The work package integrates a DMS solution with local well data; explores advanced methods for analysis/inversion of the geophysical data; develops novel methods to transform derived electrical resistivity models to lithology.
3. Model development through data integration	A computational framework that can integrate the AEM data with all other available data to generate the hydrogeologic conceptual model. Structured modeling methods and computational Multi-Point Statistics (MPS) methodologies are developed on the basis of results from WP2.
4. Uncertainty analysis	An approach to quantify uncertainty and rigorously account for its propagation through the workflow that leads to the development of the conceptual model. The conceptual model is used as the basis for the development of the groundwater model, so uncertainty must be quantified in such a way that it informs decision-making. The development of this analytical approach will be done in close collaboration with and on the basis of local specialists with the appropriate local and contextual knowledge.

5. WP1: Project management and dissemination of results

Major contributors: All members of the project team.

A website was developed and maintained by the Dean's Office in the School of Earth, Energy and Environmental Sciences at Stanford University: https://mapwater.stanford.edu/ This website presents the project objectives, key findings, and recommendations.

Presentations at key conferences ensured that there was state-level, national and international awareness of this project. The following conference presentations were made:

September 2018: Groundwater Resources Association of California, oral presentation: R. Knight

December 2018: Dewar, Noah, and R. J. Knight, Estimating the Depth to the Water Table from Airborne Electromagnetic Data, *AGU Fall Meeting Abstracts*. 2018.

March 2018: SAGEEP – Oral Presentation: Where do we lack information? MPS realizations can tell you where to drill!- M. L. Gulbrandsen, N.-P. Jensen, and T. M. Hansen

November 2018: Association of California Water Agencies (ACWA) in Fall Conference in San Diego, California: The Indian Wells Valley Brackish Groundwater Study, A multi beneficial groundwater study. A 1½ hour long dedicated session. Oral presentations by Ryan Alward, Don Zdeba, Tim Parker, Max Halkjaer.

January 2019: Danish Water Forum: The Indian Wells Valley basin in California – Danish solutions as part of reaching sustainable water resources management, M. Halkjaer, E. Auken, P. Thorn, and Ahmad A. Behroozman

February 2019: American Groundwater Trust (AGWT) Annual Groundwater Conference in Ontario, California: Salt and fresh water interaction and airborne electromagnetic surveying-Opportunities and limitations. Max Halkjaer (The IWV was used as one of two cases.)

June 2019: AGU-SEG Airborne Geophysics Workshop: Oral Presentation: Multiple Point Statistics – A Case Study from Indian Wells Valley, California, M. L. Gulbrandsen, T. B. Rasmussen, N.-P. Jensen, V. T. Le, T. M. Pallesen, P. Thorn, and M. Halkjaer

June 2019: Dewar, Noah, Kang, Seogi, and R. J. Knight, Estimating the Depth to the Saturated Zone from Airborne Electromagnetic Data, *AGU-SEG Airborne Geophysics Workshop*. 2019.

September 2019: 46th IAH congress held in Malaga, Spain: Poster Presentation: Application of Airborne Electromagnetics to Characterize Structure and Stratigraphy in the Indian Wells Valley Groundwater Basin, California, Timothy K. Parker, Max Halkjaer, Ahmad Ali Behroozmand, Paul Thorn, Don Zdeba

September 2019: Groundwater Resources Assoc. of California, Sacramento, CA: Oral Presentation: Update from the GAP (Groundwater Architecture Project): Advancing the Use of Airborne Electromagnetic Data for Groundwater Management, Rosemary Knight, Jared Abraham, Ryan Alward, Ted Asch, Esben Auken, Ahmad Behroozmand, Bill Brown, Christina Buck, Jim Cannia, Noah Dewar, Ray Dienzo, Paul Gosselin, Todd Greene, Max Halkjær, Courtney Howard, Niels-Peter Jensen, Seogi Kang, Catherine Martin, Casper Mejer, Tim Parker, David Shimabukuro, Sharla Stockton, Jacob Vind, Don Zdeba

September 2019: Groundwater Resources Assoc. of California, Sacramento, CA: Oral Presentation: Noah Dewar, Seogi Kang and Rosemary Knight: Estimating the Depth to the Saturated Zone from Airborne Electromagnetic Data

December 2019: American Geophysical Union, San Francisco, CA: Invited Oral Presentation: Rosemary Knight, Ryan, Smith Matt Lees, Tom Lauknes, Meredith Goebel, Jared Abraham, Theodore Asch, Jim Cannia: Advancing the Use of InSAR Data through Integration with Geophysical Data

December 2019: American Geophysical Union, San Francisco, CA: Oral Presentation: Seogi Kang, Todd Greene, and Rosemary Knight: Interrogating the model space of AEM inversion to answer hydrogeologic questions

December 2019: Invited Presentation: Near-Surface Geophysics – A Century of Knowledge into Action, American Geophysical Union, San Francisco, CA., December 2019, oral presentation by Rosemary Knight

August 2020: Invited Presentation: Rosemary Knight, The Use of Geophysical Methods to Support Groundwater Management in California, Bay Area Geophysical Society

September 2020: Groundwater Resources Assoc. of California, Sacramento, CA: Oral Presentation: The Stanford Groundwater Architecture Project: Utilizing Advanced Geophysical and Computational Methods for the Development of Hydrogeologic Conceptual Models, Groundwater Resources Association of California, Rosemary Knight, Esben Auken, Christina Buck, Jim Cannia, Noah Dewar, Paul Gosselin, Max Halkjær, Niels-Peter Jensen, Seogi Kang, Catherine Martin, Jesper Bjergsted Pedersen, Don Zdeba

September 2020: Invited Presentation: Rosemary Knight, Use of Satellite and Airborne Remote Sensing to Support Sustainable Groundwater Management in California, American Water Works Association

October 2020: Society of Exploration Geophysicists Annual Convention: MC-TexT: An MCMC approach to developing the resistivity-to-sediment-texture transform, Noah Dewar and Rosemary Knight, SEG Technical Program Expanded Abstracts 2020, 2499-2504.

October 2020: Society of Exploration Geophysicists Annual Convention: Interrogating the model space of airborne electromagnetic inversion to answer a hydrogeologic question, Seogi Kang, Rosemary Knight, and Todd J. Greene

October 2020 Invited Presentation: Rosemary Knight, Advancing the Use of the Airborne Electromagnetic Method for Groundwater Management in California, Society of Exploration Geophysicists Postconvention Workshop on Applied Geophysics - Addressing the Top Challenges Facing Humanity

The following papers have been published:

July 2019: Fasttimes, Vol 24, 2-2019, AEM for Investigation of Natural and Managed Aquifer Recharge in the Indian Wells Valley Basin, California, p. 44- 47: M. Halkjaer, B. Brown, P. Thorn, D. Zdeba, T. K. Parker.

Dewar, N. and Knight, R., Estimation of the top of the saturated zone from airborne electromagnetic data, *GEOPHYSICS* 85: EN63-EN76, https://doi.org/10.1190/geo2019-0539.1, 2020.

Kang, S., Dewar, N., and Knight, R., The Effects of Water Wells on Time-domain Airborne Electromagnetic Data, Geophysics, in press, 2021.

The following papers have been submitted for publication:

Dewar, N. and Knight, R., Constructing the resistivity-to-sediment-type transform for the interpretation of airborne EM data, submitted to *Geophysics*, November 2020.

Kang, S., Knight, R., Greene, T., Buck, C., and Fogg, G., Exploring the Model Space of Airborne Electromagnetic Data to Delineate Large-Scale Structure and Heterogeneity within an Aquifer System, submitted to Water Resources Research, January 2021.

6. WP2: Geophysics to hydrogeology transform

Major contributors: Noah Dewar, Seogi Kang, Rosemary Knight (Stanford University); Jesper Pedersen and Esben Auken (Aarhus University); Niels-Peter Jensen (I-GIS)

AEM data can be used to recover a 3D model of the electrical resistivity of the subsurface. That model is then transformed to obtain properties relevant for representing the hydrogeology of the subsurface, and modeling groundwater flow. Our objective in this WP was to advance the methods used in the processing and inversion of the AEM data to obtain the 3D resistivity model, and to develop an improved methodology for constructing the required transform between resistivity and sediment type.

6.1 Local well data

The starting point in this WP was gathering, sorting, and preparing local well data, as working with these data is a key element in developing the resistivity-to-sediment-type transform. The well data included information about sediment type, extracted from the drillers' logs in the well completion reports, and information about water quality from measurements of total dissolved solids (TDS) made in the wells. Contributors from Stanford developed approaches that allowed for improved accuracy in the reported locations of the wells and improved the efficiency in classifying the information about sediment type. In the three study areas we classified the geological materials as two or three sediment types: sand and gravel, also referred to as the coarse-grained type; silt, also referred to as a mixed type; clay, also referred to as fine-grained type. Contributors from IGIS developed the data management system (DMS), into which all of the well data were uploaded. This provided one central DMS that all project members and partners could access.

6.2 Assessment of powerline effect in AEM data

Focusing on the setup of the AEM survey flown over Butte County, we designed a simulation study to understand and quantify the powerline effect in AEM data using a 3D electromagnetic simulation approach. We carried out systematic analyses investigating how the powerline effect varies with 1) near surface resistivity (top 4 m), 2) transmitter height, 3) number of powerline poles. In general, increasing near surface resistivity, decreasing transmitter height, and increasing the number of powerline loops resulted in an increased powerline effect. In addition, to understand the powerline effect in the field AEM data acquired in the study area, we calculated the powerline effect using a vertical resistivity profile obtained from the inversion of the AEM data; this resistivity profile was obtained by averaging all resistivity profiles from the inversion of the AEM data. In Figure 6.1 we show the calculated powerline effect along an AEM flight line crossing the powerline at x = 0 m. We determined that the powerline effect was negligible after 174 m away from the powerline. Note that this result change when data are acquired in other areas having a resistivity structure that differs from our study area.

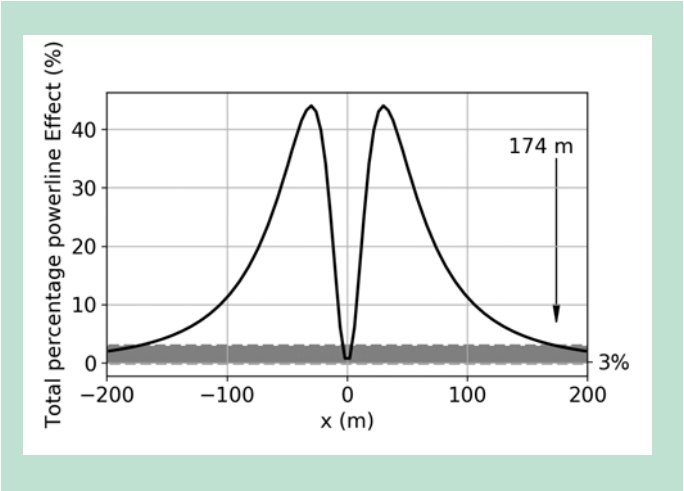


FIGURE 6.1. Profile of the total percentage powerline effect (TPPE). Approximately 174 m away from the powerline location (x = 0 m), the TPPE falls below 3 %, shown as the grayed area at the base of the plot, indicating negligible powerline effect.

6.2.1 Results from acquisition of data in Butte County

Our examination of 41 water wells within 100 m separation distance from AEM flight lines is summarized in Table 6.1. In all cases when low-quality AEM data were found at the sounding closest to the well (14 cases), we identified the presence of a powerline as the source of the noise. After removal of all impacted soundings, we found that the separation distance between a well and the closest high-quality sounding was, on average, 108 m. The separation distance between the high-quality sounding and the powerline was, on average 104 m, which is less than the 174 m predicted by our simulation. In 24 cases where we had high-quality data at the sounding closest to the well, we found no powerlines within 200 m. For the last three cases of high-quality data, we found distances to a powerline of 11 m, 15 m, and 68 m, all significantly less than our predicted value of 174 m. We attributed the smaller observed distance between the high-quality sounding and the powerline to the deterioration of the ground electrodes or their poor contact with the surrounding material.

TABLE 6.1 Summary statistics from examination of the effects related to the presence of water wells in the AEM data. IRR, DOM, and MON in the column "type of well" correspondingly indicate irrigation well, domestic well, and monitoring well. In the column listing data quality, high-quality means there was no noise observed at the closest sounding location, and low-quality means AEM data at the closest sounding were affected by noise.

Number of Wells	Type of Well	Separation Distance (Average) Between Closest Sounding and Well (m)	Data Quality	Separation Distance (Average) Between Closest Sounding and Powerline (m)	Separation Distance (Average) Between High-Quality Sounding and Well (m)	Separation Distance (Average) Between High-Quality Sounding and Powerline (m)
6	IRR	13-95 (45)	Low	7-35 (23)	80-160 (115)	73-100 (95)
5	DOM	56-92 (75)	Low	4-130 (63)	100-120 (104)	70-150 (109)
3	MON	6-92 (49)	Low	2-50 (22)	100-120 (106)	100-115 (108)
18	IRR	0-94 (48)	High	> 200 m	0 – 94 (48)	> 200 m
3	DOM	50-93 (75)	High	> 200 m	50 – 93 (75)	> 200 m
3	MON	19-67 (46)	High	> 200 m	19 – 97 (46)	> 200 m
2	IRR	2-74 (38)	High	11-15 (13)	2-74 (38)	11-15 (13)
1	DOM	84	High	68	84	68

6.2.2 Conclusions/recommendations

In our specific study area, we found that the powerlines are the only noise source related to the wells that has an effect in the AEM data. From our numerical simulations, we found that the effect of the powerline will be seen in the AEM data until there is a separation distance of 174 m from the powerline. In the field data, we found no effects from a powerline at soundings much closer to a powerline than this predicted distance of 174 m. We attributed this to high ground contact resistance caused by degradation of the ground electrodes at the bottom of the powerline poles. For our 41 wells, after removing any soundings impacted by powerline noise, we were left with high-quality soundings located 0 m to 160 m from the water wells, with an average distance of 77 m. In areas such as this, where co-located AEM and lithology data are desired and powerlines are the only noise source associated with the water wells, the AEM survey should be planned so as to minimize the distance between the wells and the AEM flight lines while avoiding the regions where a significant powerline effect in the AEM data would be expected.

6.3 Processing and inversion of AEM data

With funding from other sources, AEM data were acquired in the three study areas of Butte/Glenn Counties, Indian Wells Valley, and San Luis Obispo County. The way in which AEM data are processed and inverted is a critical step in the use of the geophysical data to derive information about the hydrogeology of the subsurface. We here describe the research conducted in this step, working with the data from SLO.

The processing is as important as acquiring precise data. Poor data processing can reduce the resolution significantly and poorly described information affecting the signal can create biased and misleading model results. The processing entails using appropriate data averaging schemes, removing distorted data, and having accurate sensor altitude, GPS and pitch/roll data. Major noise sources for the AEM data are man-made installations including power lines, roads, cables, railways, windmills, houses, antennas and vineyards etc. Typically, all data within a distance of 100–200 m from the known noise sources should be culled manually. Even for an experienced geophysicist identifying all noisy portions of the data and completely removing them is a challenging task. Therefore, rather than considering the processing and inversion as independent steps, they should be considered an iterative process, where one first carries out the processing of the data, then inverts the data to determine if there are any suspicious resistivity structures imaged and to check the misfit between the observed and predicted data. If there are suspicious structures imaged or portions of data showing high misfits, then the processing and inversion procedure should be repeated until a satisfactory quality of the resistivity model and data misfit is obtained.

Little research has been previously performed with respect to the interference from vineyards, which potentially could contaminate the results of mappings. Yet, throughout California, there will undoubtedly be interest in acquiring AEM data in areas with vineyards. Vineyards would be seen as added signal to the measurements, and therefore it would be difficult to discriminate the added signal from a clay layer or a vineyard since the wires supporting the vines cause noise in the data. Consequently, a number of extra flight lines were flown in the SLO area to investigate the effect. A detailed GIS layer with the exact location of all vineyards was created, and selected vineyards were visited to investigate their construction. The data acquired close and on top of the vineyards were scrutinized, and it was found that in most cases the vineyards introduced an added signal, which in the modelling occurred as a fake conductor. Consequently, all data were culled in the close vicinity of the vineyards.

Geophysical inversion is the process by which we estimate the resistivity model of the subsurface fitting the data. An inversion algorithm uses an iterative optimization technique and allows

inclusion of prior information that can be obtained from well data (e.g. lithology, water level, salinity, etc.) To produce unbiased results the inversion algorithm needs full descriptions of the AEM system (e.g. geometry, altitude and attitude of the transmitter/ receiver, transmitter waveforms, receiver bandwidth characteristics etc.). There are a number of tuning parameters in any inversion algorithm that require geophysical knowledge of AEM data as well as inversion methodology. Therefore, it is critical to have an experienced geophysicist involved in this process. Similar to the philosophy of viewing the processing step as an iterative process, the inversion and subsequent hydrogeological interpretation of the recovered resistivity model should not be viewed as a turn-key operation.

In the SLO we had close interaction among geophysicists, hydrogeologists, water managers, and local stakeholders such that we could run multiple AEM inversions based upon these interactions. We applied the commonly used inversion approach of a 1D spatially- (or laterally) constrained inversion technique, developed by Aarhus University and commercially available through the Aarhus Workbench. This method assumes a 1D layered earth structure when simulating the AEM data, which is a robust way to invert AEM data proven in various groundwater mapping projects as well as in our GAP project. For the SLO area both a smooth and a sharp model inversion have been carried out. Both inversion types use the SCI-setup, but the regularization scheme is different. The smooth regularization scheme penalizes the resistivity changes, resulting in smooth resistivity transitions both vertically and horizontally. The sharp regularization scheme penalizes the number of resistivity changes of a certain size, resulting in model sections with few, but relatively large, resistivity transitions. Normally the airborne EM dataset is fit almost equally well with the two inversion types. Assuming a layered geological environment, picking geological layer boundaries will be less subjective in a sharp model result compared to a smooth model. Contrary, the smooth inversion result can reveal vague resistivity signatures that may be suppressed in the sharp results. By carrying out both inversions we found that we were able to describe the geological setting in greater detail.

Additionally, system response modelling of the airborne electromagnetic system was carried out. With the system response modelling scheme, the waveform, low-pass filters etc. are not modeled separately, but as a system response measured for the specific SkyTEM setup used in the SLO survey. This approach enables accurate modeling of gates in the ramp down time.

6.4 Identifying the top of the saturated zone

The amount of water in the pore space of a sediment has a significant impact on the resistivity values associated with a sediment type. It is thus essential, in transforming resistivity to sediment type, to treat separately the regions above and below the top of the saturated zone (TSZ), the depth at which there is a change from the existence of both air and water in the pore space of the sediments, to the pore space being fully saturated with water. This requires knowing, prior to construction of the transform, the location of the TSZ at the time of AEM data acquisition.

6.4.1 Overview of developed methodology

Stanford developed a methodology based on the assumption that there is some information in how the distribution of resistivity changes with depth that can be used to estimate the TSZ. The application of the methodology in a given study area requires measurement of the water table elevation (WTE). In Butte and Glenn Counties, water level measurements were available from 29 wells, 22 of them are continuously monitored wells built and maintained by the California Department of Water Resources (DWR); water table measurements are made automatically in these wells at least every hour. The data for these wells were downloaded from the state-run Water Data Library (California Department of Water Resources, 2020a). The remaining seven WTE measurements came from wells that are part of the DWR's California

Statewide Groundwater Elevation Monitoring (CASGEM) program; these measurements were downloaded from the Groundwater Information Center Interactive Map Application (GICIMA) website (California Department of Water Resources, 2020b).

After collecting the WTE measurements, the next step in our methodology is to gather the resistivity data, recovered from the AEM measurements, that is close to the wells where we have measurements of WTE. The amount of resistivity data to be gathered is determined by a variable search radius. Distributions of resistivity data are then formed over regularly spaced intervals in the subsurface, and various statistical properties calculated on each of these distributions. The depth at which the change in a given statistical property is a maximum is then taken as the depth of the TSZ estimated by that combination of statistical property and search radius. The optimal combination of search radius and statistical parameter is then defined as that which produced estimates of TSZ with the lowest root mean squared (RMS) error when compared to the well based measurements of WTE. Figure 6.2 captures the finalized version of the developed methodology.

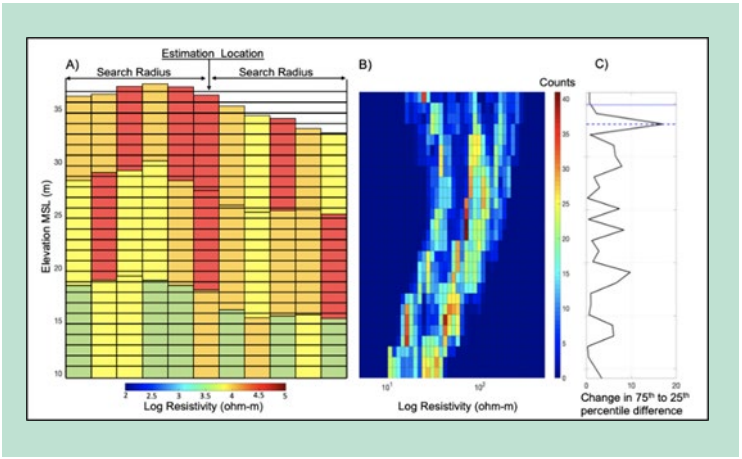


FIGURE 6.2 Schematic illustrating the process of producing resistivity distributions upon which various statistical properties are calculated as part of the developed methodology. A) Each colored block represents a layer, while a stack of blocks represents a single 1D resistivity model. The horizontal black lines are the 1 m–thick depth intervals that the resistivity distributions are built upon. The elevation at which each 1D resistivity model starts is different due to the changing ground surface elevation. B) The depth registered distribution of resistivity for a selected estimation point. C) The gradient with depth of a selected statistical property with the depth interval that contains the maximum annotated as the estimated depth to the TSZ for the same estimation point as in B).

6.4.2 Results from Butte County

When this methodology was applied to the AEM data acquired in Butte County, the RMS error between the optimal TSZ estimates and the well-based measurements of WTE was found to be 3.8 m. This level of error is only ~15% greater than the average layer thickness in the resistivity model upon which these estimates are based. Figure 6.3 shows an example of the methodology applied across an entire line of AEM data at regularly spaced points. The resulting TSZ estimates have been smoothed with a moving average filter before being plotted.

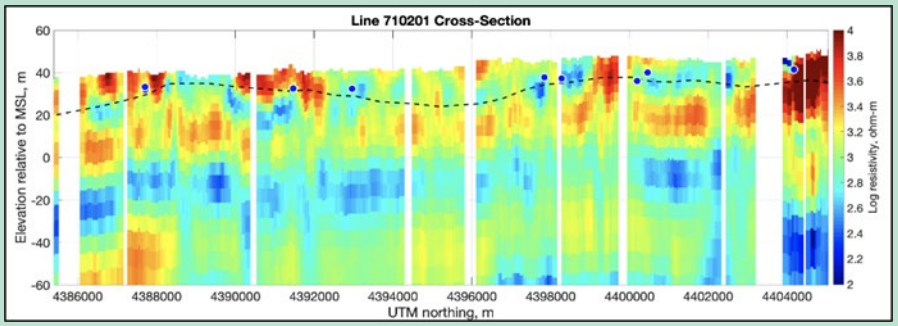


FIGURE 6.3 Resistivity cross-section of a line from the AEM dataset acquired in the Butte survey area. The dashed black line indicates the estimated depth to the TSZ while the blue circles with white outlines are the locations of the WTE measured in wells within 2 km of the line.

6.4.3 Conclusions/recommendations

In this study, we developed and tested a new methodology to map the TSZ from AEM resistivity data with sufficient accuracy to inform the development and application of resistivity-sediment type transforms that account for the saturation state. The benefits of using the developed methodology can be realized in any area where AEM data are acquired from saturated and unsaturated zones; however, these benefits will be greatest in areas with significant spatial or temporal fluctuations in the TSZ, or with few well-based measurements of WTE.

There are some conclusions that can be drawn from this study about the search radius that can inform the transfer of this methodology to other locations: (1) The maximum search radius tested should be informed by the regional groundwater gradient, and (2) the step size at which the search radius is increased should be large enough to ensure that, given the spatial distribution of AEM soundings, the number of soundings used increases with each increase in the search radius. The developed methodology, when optimized for local conditions, can be used to produce the best possible estimates of the TSZ from the AEM data acquired in a survey area. Due to the TSZ's impact on the relationship between resistivity and sediment type, the ability to estimate the TSZ from the AEM data can dramatically improve the accuracy of a resistivity-sediment type transform and it thus represents a significant contribution to advancing the adoption of airborne geophysics for mapping and managing groundwater systems.

6.4.4 References

Dewar, N. and Knight, R., 2020, Estimation of the top of the saturated zone from airborne electromagnetic data, GEOPHYSICS 85: EN63-EN76, https://doi.org/10.1190/geo2019-0539.1

California Department of Water Resources, 2020a, Water data library, http:// wdl.water.ca.gov/waterdatalibrary/docs/Hydstra.

California Department of Water Resources, 2020b, Groundwater information center interactive map application, https://gis.water.ca.gov/app/ gicima.

6.5 Construction of the rock physics transform

Through the process of inversion, models of subsurface resistivity are derived from acquired AEM data. However, what decision makers require for the development of the HCM are models of subsurface sediment type. Therefore, a resistivity-to-sediment type, or rock physics, transform must be developed. We do this using information on sediment type acquired when wells are drilled and the resistivity models obtained through inversion from the AEM data.

6.5.1 Overview of developed methodology

Current rock physics methodologies have several limitations due to issues of resolution and sensitivity as the sediment type data from the wells and the AEM resistivity have different levels of sensitivity to subsurface features and are also acquired at different vertical resolutions. Sediment type is recorded by a driller while examining the cuttings that are produced during the well drilling process; as a result the sensitivity of the acquired sediment type information from wells is higher for sands and gravels as they are easier to see and feel in the cuttings while fine sediments such as silt or clay may be lost. The vertical resolution of the sediment type information from wells is independent of depth. In contrast, due to the diffusive nature of the AEM measurement, it is more sensitive to conductive layers, such as clays and silts, than coarse layers like sands or gravels. The vertical resolution of the AEM method also decreases with depth. The resistivity data is also often spatially interpolated in order to obtain co-located resistivity and sediment type information. The interpolation of the resistivity data introduces uncertainty because in order to perform the interpolation, assumptions must be made about the correlation structure of the geology in the area. Because this correlation structure is generally unknown, the magnitude of the uncertainty introduced by this interpolation is difficult to quantify.

To circumvent the limitations of current rock physics transforms, we have developed a methodology based on Markov Chain Monte Carlo (MCMC) with Metropolis-Hastings sampling and the physics of the AEM method to build distributions of resistivity for each sediment type category. The spatial interpolation issue raised above is avoided entirely as during the survey planning phase for each project, flight lines were designed to pass as close as possible to the locations of wells where there exists high quality sediment type information. As a result of this, truly co-located sediment type and AEM data were obtained. Due to the design of the MCMC-based methodology and the use of the physics of the AEM method, the issue of resolution is also completely avoided. The difference in sensitivity between the sediment type logs and the AEM resistivity data is taken into account by using different rock physics transforms for different sensitivity thresholds calculated on the AEM data. We defined six sensitivity thresholds, based on the work of Behroozmand et al. (2013).

6.5.2 Results from Butte County

The distributions of resistivity for each sediment type category and the first four sensitivity thresholds (there was little change beyond the fourth threshold) produced using our developed methodology and the log logistic distributions fit to them are shown below in Figure 6.4.

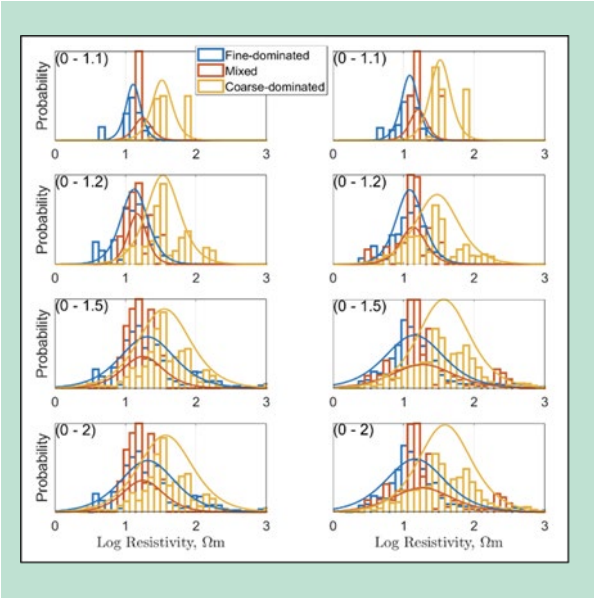


FIGURE 6.4 Resistivity distributions for each sediment type category for the first four sensitivity thresholds. The left column shows the distributions retrieved above the TSZ while the right column shows those retrieved below the TSZ. The sensitivity threshold is relaxed from top to bottom along the rows of the plot with the strictest in the top row and the most relaxed in the bottom row. The X axis of each plot shows log resistivity in ohm-m, while the Y axis is normalized counts and probability respectively for the histograms and the log logistic distributions.

The fit log logistic distributions can be used either deterministically or probabilistically to transform the AEM-derived resistivity data into sediment type. The deterministic option is to simply take the most likely sediment type category for a given resistivity and sensitivity and assign that sediment type category to the given resistivity layer. The probabilistic option is to sample many times directly from the fit log logistic distributions or to use geostatistical simulation algorithms in order build up a suite of possible models of subsurface sediment type. The most likely model resulting from the application of our developed transform to the AEM data acquired in the Butte County study area along a selected flight line is shown in Figure 6.5. Also shown in the figure is the sediment type information from all wells within 1000m of the selected flight line.

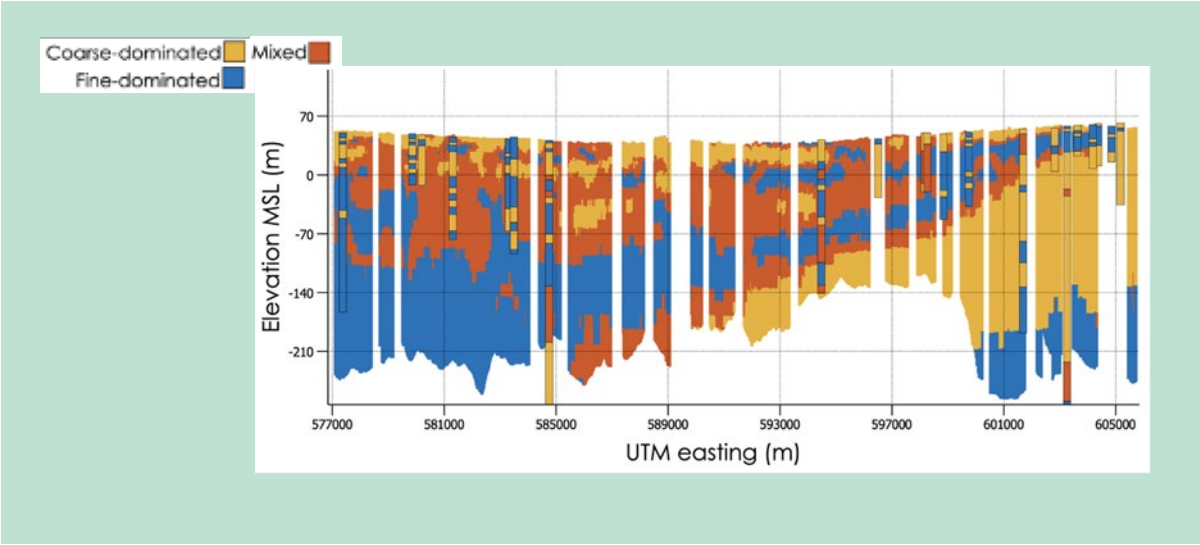


FIGURE 6.5 The most likely lithology model for the selected portion of the flight line is plotted along with the sediment type information from every well within 1000 m. The vertical exaggeration is 30.

The resistivity ranges both above and below the TSZ and for each sensitivity bin for which each sediment type category is the most likely are shown in Table 6.2 below. The minimum and maximum values of 0.5 and 395.6 ohm-m shown in Table 6.2 are, correspondingly, the maximum and minimum resistivity values in the AEM resistivity model. The resistivity values at which one sediment type category becomes more likely than another are visible in Figure 6.4 as the points where one of the fit log-logistic PDFs crosses over, and thus has a higher probability than the previously highest fit log-logistic PDF.

TABLE 6.2 Resistivity ranges for which each of the three sediment type categories are the most likely both above and below the TSZ and for each of the sensitivity bins.

Sensitivity bin	Fine-dominated range (ohm-m)	Mixed fine and coarse range (ohm-m)	Coarse-dominated range (ohm-m)
Above the TSZ			
1	0.5 – 17.9	18 – 19.9	20 – 395.6
2	0.5 – 19.4	–	19.5 – 395.6
3	0.5 – 15.6	–	15.7 – 395.6
4	0.5 – 16.5	–	16.6 – 395.6
Below the TSZ			
1	0.5 – 17.4	17.5 – 19.4	19.5 – 395.6
2	0.5 – 17.7	–	17.8 – 395.6
3	0.5 – 15.8	–	15.9 – 395.6
4	0.5 – 16.4	–	16.5 – 395.6

6.5.3 Conclusions/recommendations

The methodology we have developed here is data driven, agnostic of the geophysical method, and independent of the inversion workflow and thus represents a general and robust method for constructing resistivity-to-sediment-type transforms. The transforms provide a most likely sediment type model that can provide valuable information about the large-scale structure of the aquifer system, and also provide information about uncertainty in determining sediment

type that should be taken in account in the interpretation of the AEM data. The methodology, with minimal adjustments, can be applied to any other survey area given sufficient co-located AEM data and sediment type logs. The resulting probabilistic representations of subsurface sediment type can be used as the starting point for numerous geostatistical interpolation workflows or can be directly sampled from and used to inform structure in numeric groundwater flow models. Due to the qualities of the developed methodology, especially the use of sensitivity thresholds to create separate transforms that respect the varying sensitivity of the AEM method, we believe that the use of our methodology will advance our ability to obtain reliable information about sediment type from AEM data.

6.5.4 References

Behroozmand, A. A., E. Dalgaard, A. V. Christiansen, and E. Auken, 2013, A comprehensive study of parameter determination in a joint MRS and TEM data analysis scheme: Near Surface Geophysics, **11,** 557–67, doi: https://doi.org/10.3997/1873-0604.2013040.

7. Model development through data integration

Major contributors: Paul Thorn, Tillie M. Madsen and Max Halkjær (Ramboll); Niels-Peter Jensen, Thomas Bager Rasmussen, Mats Lundh Gulbrandsen, and Le Thanh Vu (I•GIS).

The development of hydrogeologic conceptual model for the study areas was done both in the more conventional way and by the use of a statistical approach. The conventional approach is time consuming but creates very valuable input for the training images and serves as a reference during the validation of the statistical model.

Early on in the project, the team discovered that only a very limited amount of data was available in a digital tabular format, which is a necessity for the development of the HCM. Therefore, outside the MUDP project much time and effort was spent on gathering and digitization of existing data. A significant lesson from the digitization and the use of the digitized data was that the quality of the lithological descriptions is highly variable and often it was a challenge to obtain reliable coordinates, the result being several wells having the same coordinates in the center of section or center of tract.

In the two pilot study areas Indian Wells Valley and in the Paso Robles area the hydrogeology is highly influenced by recent and old tectonic activities which had to be handled during the development of the HCMs. In IWV the HCM development was further complicated by brackish and very saline water in certain parts of the basin, which affected the AEM resistivity signal.

7.1 Development of conventional hydrogeologic conceptual models

7.1.1 Overview

Ramboll has led the development of conventional hydrogeologic conceptual models for the two pilot study areas Indian Wells Valley basin and the Paso Robles basin in San Luis Obispo County. The task was a demonstration of the importance of integration of a number of different data sets including existing geological and hydrogeological interpretations to support the AEM data. The workflow was based on the many years of experience gained from the Danish National Groundwater project though we found that adaptation to the local conditions had to take place. We found that data availability, data quality and meta-data was a significant challenge.

A Hydrogeologic Conceptual Model (HCM) is developed with the purpose of providing an understanding of the geometry and physical characteristics of the groundwater systems, which forms the basis for a numerical groundwater model. The HCM of the basins were constructed following the phases:

Phase 1. Data acquisition
Phase 2. Defining the geologic settings
Phase 3. Development of the 3D geological model
Phase 4. Development of the HCM

A diagram illustrating the workflow is presented in Figure 7.1a below. Each phase is described in more detail in the following sections.

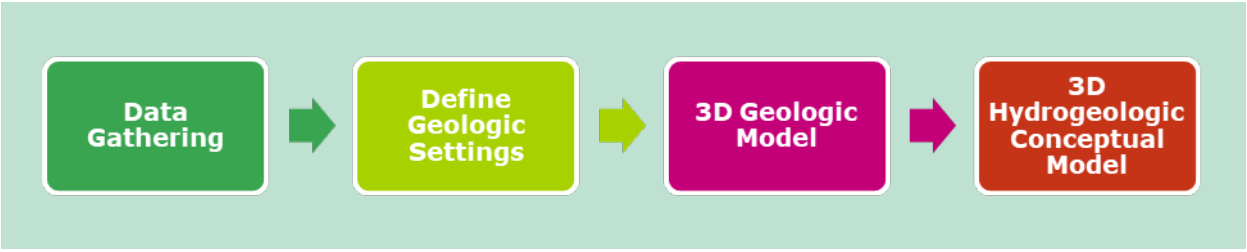


FIGURE 7.1a Workflow HCM development.

DWR defined the general boundaries of the basins. Often the boundaries are defined by the presence of bedrock. We found that the AEM data can be used to confirm and, within certain geographical areas, challenge the current boundaries.

It has not been possible to obtain a satisfactory dataset to be able to define aquifer specific water levels as we often did not have information on the screened interval for the well and thus could not say precisely which aquifer the water level measurement represented. For that reason, we were also not able to determine vertical pressure gradients.

The traditional HCM was challenged and quality controlled by IWVWD and their consultants as well as SLO County and their consultant.

7.1.2 Data gathering

The first step in the development of the HCM is data gathering. Data are collected and processed, so they can be uploaded to the modelling software GeoScene3D. Data often used include digital elevation models, geological maps, geophysical surveys (e.g. seismic, AEM, gravity and aeromagnetic data), and borehole information (e.g. well logs, well screen, water level, water quality and electrical logs). This task was done outside the MUDP project. Even though a significant amount of time was spent on this task we found that the quality of the data, especially the lithological descriptions and the geographical location of the boreholes, was still questionable.

7.1.3 Defining the geologic setting

The next step is to perform a review of the collected data, as well as previous studies on the subbasin's geology and hydrogeology. The object of the review is to obtain an initial understanding of the area's geology, knowledge that will support construction of the 3D geological model and HCM. The review is performed by providing a description of the landscape, the geological units and important structural elements (e.g. faults). The text is often accompanied with illustrations (typically cross sections) showing the major architecture of the geological units and geological structures. The geological formations and the known tectonic structures within the basins were described. In subareas within the study area it was found that the aquifers were brackish or even very saline which is crucial information for the development of the 3D geological model as the AEM data are strongly influenced by variations in the water quality.

7.1.4 Development of the 3D geological model

The third step is the construction of a 3D geological model using the modelling software GeoScene3D by I-GIS. In a 3D geological model, lithostratigraphic units (the basic geologic units described by physical properties and sequence) are correlated to known geological formations, which are then modelled. The result is a 3D rendering of the thickness and distribution of the individual geological units (i.e. geological formations) within the study area. In a 3D geological model, the accuracy and quality of the interpretations are dependent upon the amount of chronostratigraphic information (the relative age of rock strata in relation to time) available. As the geological settings even for the younger generations of sediments were very influenced by tectonics it was essential to include folds and faults in the interpretations.

7.1.5 Development of the HCM

The final step is the development of the HCM. The aim of the 3D geological model is to subdivide the geological strata into chronostratigraphic units (i.e. formations), and as a result, the geological units may contain a wide range of lithologies characterized by different hydraulic properties. In an HCM, the aim is to subdivide the geological strata into hydrostratigraphic units (i.e. aquifers and aquitards) using the interpretations from the 3D geological model. Thickness and the elevation of the layers were presented as thematic maps. The hydrostratigraphic units define the groundwater system, and the HCM provides the foundation of the groundwater flow model. In addition, the HCM provides useful information on the location of groundwater recharge areas.

In both the Indian Wells Valley pilot study area as well as in the Paso Robles pilot study area the two projects have demonstrated the value of having a large amount of digitized data available and how the HCM development benefit from existing interpretations about the depositional environment and the ability to combine a large range of datasets (borehole information, existing cross sections, seismic, water quality, water levels, etc.) in a 3D interpretation software package (Geoscene3D). In IWV there were no existing models available. In the Paso Robles Basin, six existing cross sections across the area were present. The ease with which borehole data and AEM data could be combined in a 3D interpretation environment really demonstrated the importance and the value of bringing the WCRs into a tabular database format.

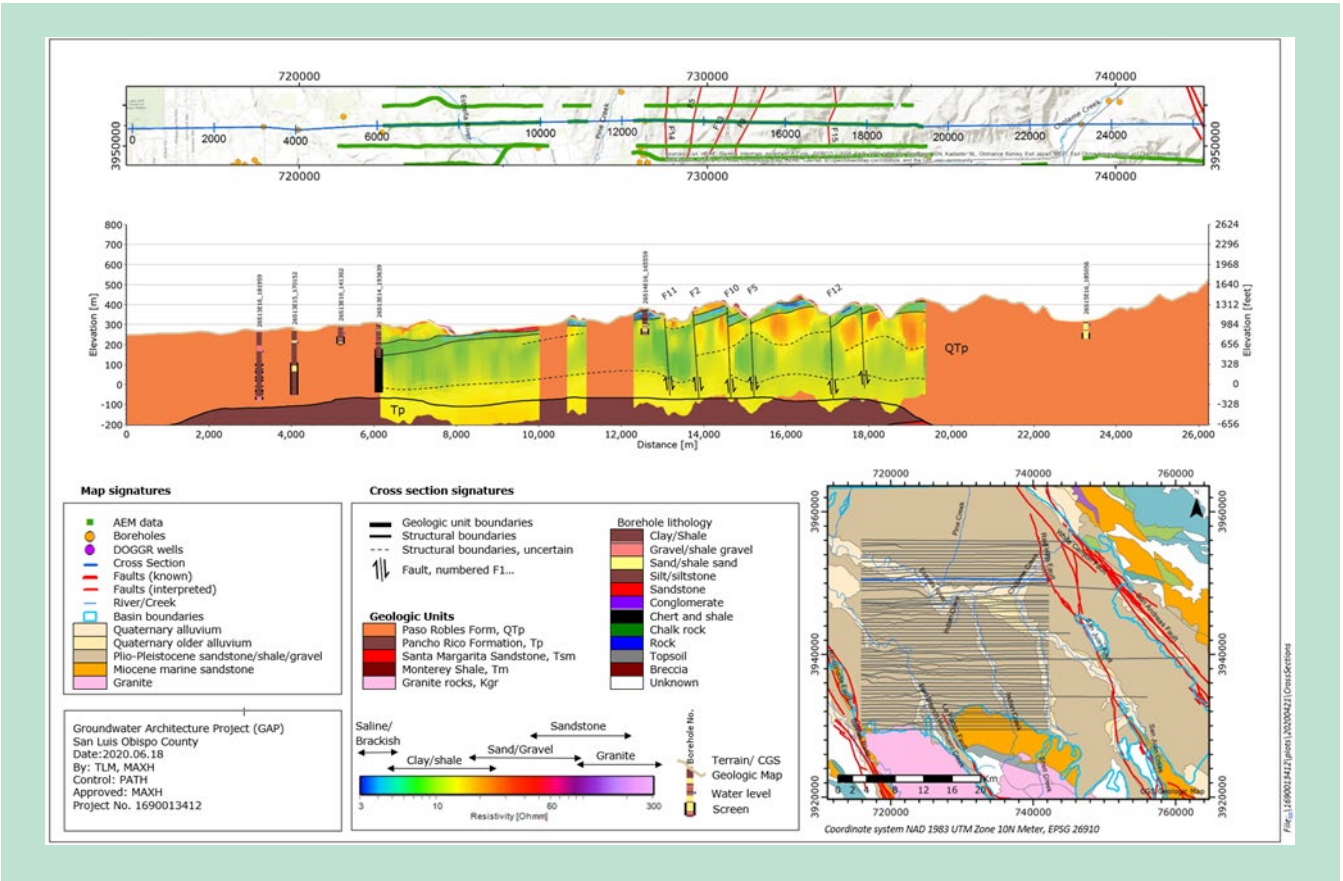


FIGURE 7.1b Example from the Paso Robles pilot study area. The vertical cross section includes information from existing geological interpretations, boreholes digitized by I-GIS, AEM data and existing geological maps. New faults are identified. The interpretation shows how the Quaternary Paso Robles formation has been subdivided into hydrogeological units.

For more details about the traditional HCMs for the IWV basin and the Paso basin see the extensive reports developed specifically for the two agencies.

7.2 Developments of the Multiple Point Statistics method

I•GIS led the developments of the Multiple Point Statistics (MPS) method. MPS is a numerical methodology that can integrate AEM data, well data, other datatypes and prior knowledge of geologic patterns (e.g. fluvial environment) in a way that can generate multiple realizations of lithology/sediment type models (i.e. categorial models) of the subsurface capturing the uncertainty.

The MPS method results in an ensemble of models all of which are compatible with the data input and the geological understanding of the depositional environment in the model area. This is done by applying a training image (TI) which is a 3D model incorporating the conceptual geological features of the model area. Each point in the model, is thereby assigned a probability of a certain lithology in the model area, usually in two fractions, sand and clay, or fine and coarse, which in terms of hydrology is the permeable and non-permeable fraction.

Using the ensemble of equally possible models leads to the ability to calculate statistics and thereby quantify uncertainty of the resulting models considering the different data used in the model.

7.2.1 Development and enhancement of MPS

I•GIS had a task in the project of developing tools for running Multiple Point Statistics (MPS) on datasets provided by the project. The basic MPS code has been in development for several years. Within the GAP further enhancements were made to that code. But foremost, efforts have been done to make MPS more accessible and applicable for geology professionals performing consultancy projects and not only for highly trained academics. This was done by further enhancing the GeoScene3D software (GS3D) as the platform on which MPS can be run. The benefits of using a platform like GeoScene3D, is that all the data preparation tools are present. GeoScene3D is a high-end and widely used modeling platform for building HCM models especially when AEM datasets are involved. Therefore, tools for loading data, model building, and visualization of models and results are already present.

Only in rare cases can an MPS model be run on a full area of interest, i.e. a basin. Often one stratigraphic or hydro stratigraphic unit (often the aquifer of interest) must be modelled individually. Furthermore, the conceptual understanding of the unit could have different depositional patterns that may change throughout the basin, i.e. alignments in N-S in one part changing to NW-SE in another part. This change is not always possible to describe in one TI, but will need two or more TI to cover.

Description of software development

Approximately half of the development process was used to improve the workflow of using MPS simulation inside GS3D. As described below, the User Interface (UI) has also been improved, simplified, and been made more user-friendly. GS3D has also been upgraded from a 32bit application to a 64bit application to be able to accommodate a large dataset.

The other half has been used outside to improve the performance of the Simulation library MPSLib. Secondly efforts were undertaken to ensure that simulations can be run in parallel and to ensure that performing simulations within different areas having different Training Images (TI) is possible.

GeoScene3D software improvements

Software improvements were done throughout the project, but most improvements were done in the period Oct 2019 to April 2020. The software development was done in an agile process, thus testing, modeling exercises and software improvements were done along with coding.

Within GeoScene3D tools were enhanced to facilitate building TI. Geological shapes and structures can be made allowing the user to create TI models that resembles “real” geology. Furthermore, tools allowing users to prepare softdata, (in this project primarily AEM soundings), and harddata (boreholes) for the simulation process. Other tools to mention are functionality to build the rock transform relation between AEM and well logs.

Below we describe major developments that have been done in GS3D:

a. **GS3D 64 bit**

Up until now, GS3D has been a 32-bit application which has a memory limitation of 3Gb. It was impossible for GS3D to handle multiple big 3D grids at the same time. Therefore, work was done to upgrade GS3D to a 64 bit application; GS3D code together with its several main libraries have been reviewed and upgraded to 64 bit. This was a long, iterative process between the development team and included both internal and external testers.

A beta public test version (version 11.0.13.632) was out on the 22nd of January 2020. Follow up was an official release (version 11.0.13.645) on the 1st of May 2020. Now GS3D can use all the available RAM and can open multiple big 3D grids at the same time.

b. **Introduction of boundary constraints**

During actual modeling done in the GAP project, it was realized that the simulation is often run inside a volume within a bigger model between an upper and lower surface and surrounding limits. Therefore, boundary constraints (area + surfaces) have been introduced as a feature in the workflow to limit the simulation area. This extra step was included in the simulation process and in the process of creating soft + hard data. These constraints also helped to speed up the simulation process, thanks to a smaller number of nodes to be simulated.

c. **Improvement of user interface**

At the beginning of GAP project a basic MPS test project (with 2D and 3D data) was prepared. Many explanations have been added in simulation's wizard so that users in a step-by-step wizard's UI has also been simplified and organized so GS3D's user can follow the process easily, together with an introduction of using MPS workshop. Besides this, enhancements have been done to facilitate the resistivity to lithology transform, a central element in using MPS with AEM data. The aim has been to have GeoScene3D to fully cover all the tasks related to setup and run MPS simulations.

d. **Enhancements of the MPSLib**

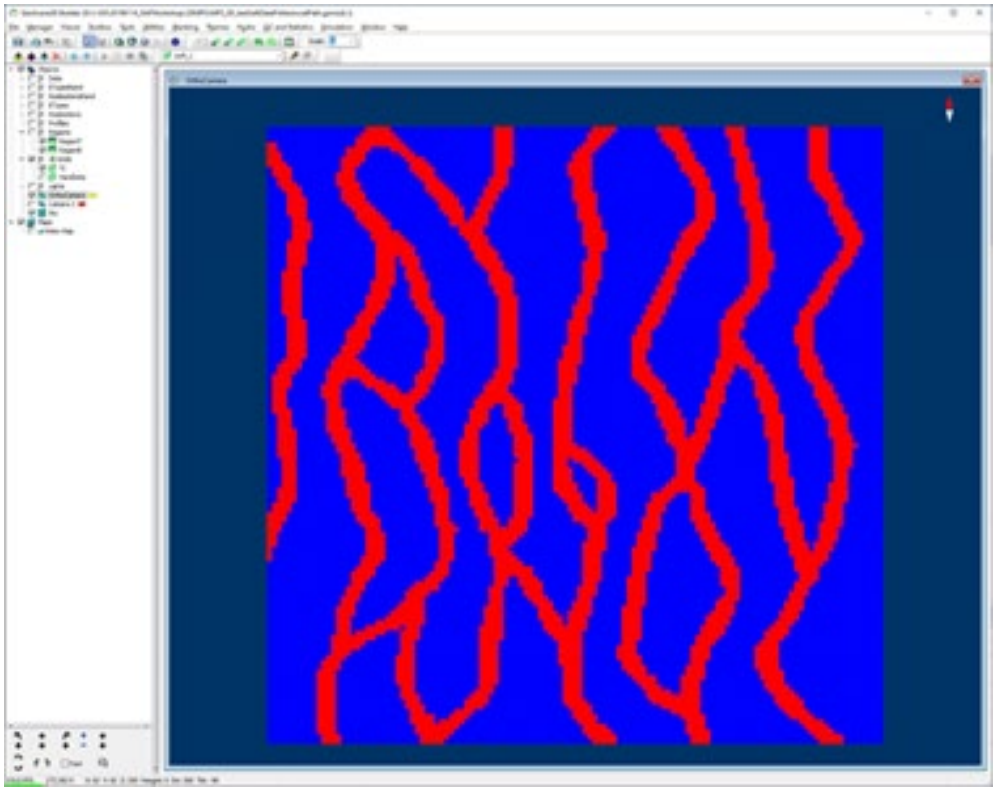
Enhancements of MPS were done throughout the project, but most enhancements were done in the period Oct 2019 to April 2020. The code revisions are done in an agile process, thus test, modeling exercises and software improvements were done along with coding.

MPSLib is the main library using from GS3D to perform MPS simulation. This multiplatform opensource library is developed and maintained together by I-GIS, Niels Bohr Institute and Aarhus University (https://github.com/ergosimulation/mpslib).

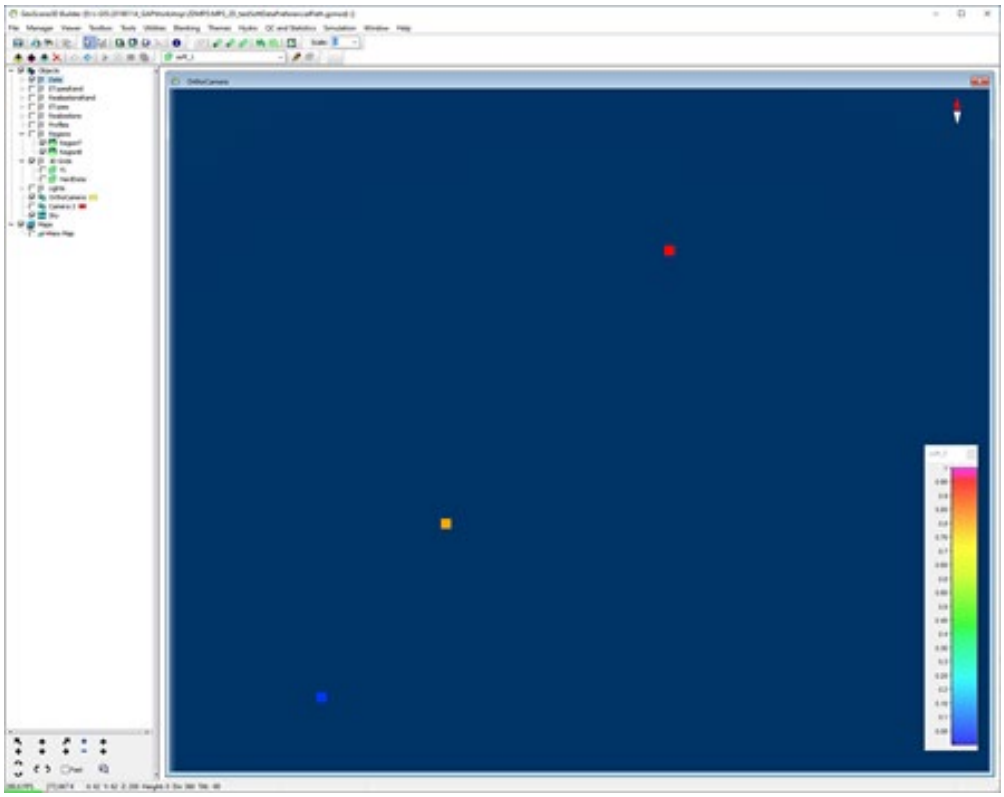
e. **Preferential path implementation**

Earlier all simulation algorithms in MPSLib used a random path to perform the simulation. An implementation of a preferential path algorithm that is using the location of softdata for a better simulation quality where soft data are non-colocated. This is the case with AEM data, that are collected in lines, and is also often the case with well logs. A preferential path article can be found here: https://www.sciencedirect.com/science/article/pii/S0098300417305654?via%3Dihub

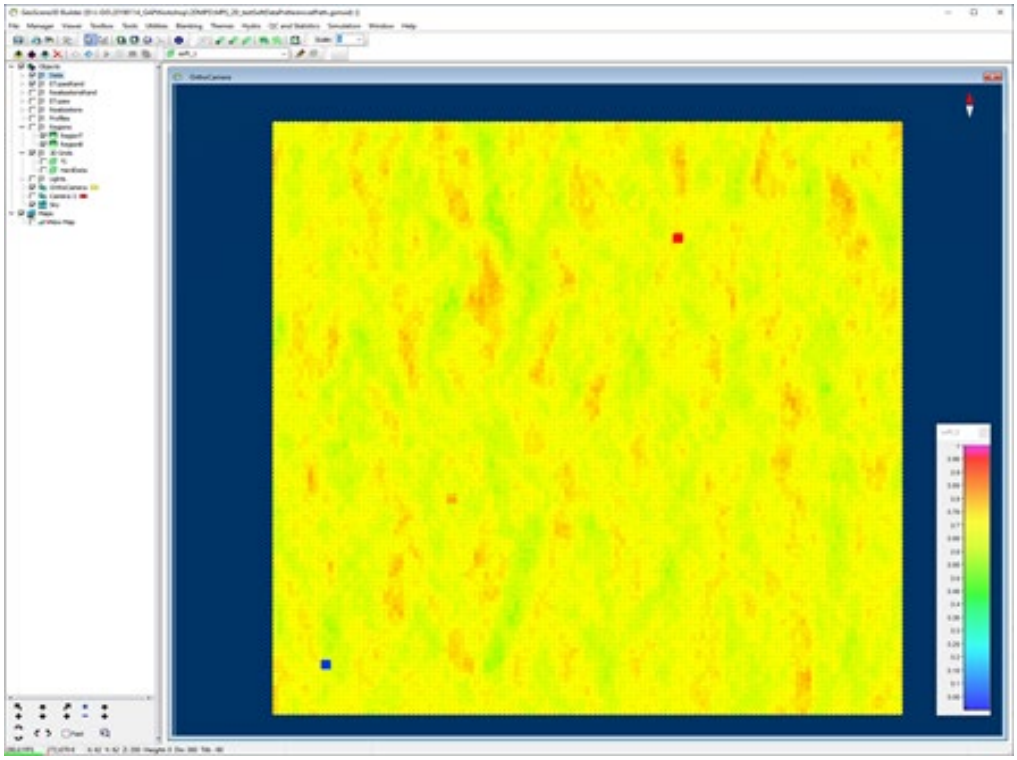
Following is a simple example of using preferential path in 2D, we can see the difference between using a random path and preferential path. There is a higher certainty where softdata are located



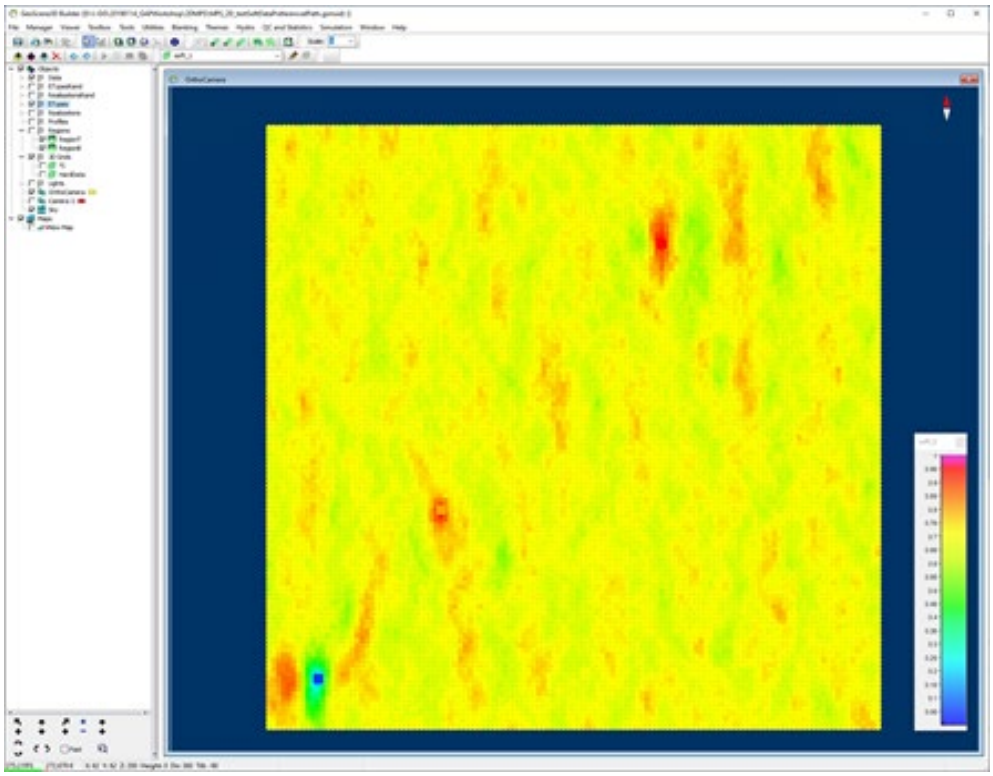
Training Image



Softdata with 3 non-colocated points.



Probability of sand from 200 realizations using Random Path (3 points of softdata are overlaid)



Probability of clay from 200 realizations using Preferential Path (3 points of softdata are overlaid)

f. **Mask grid**

Together with the introduction boundary constraints in GS3D, a mask grid has been introduced inside MPSLib (MPSLib version 1.3 released the 25th March 2019). This improvement allows MPSLib to perform the simulation in a selected subarea of a larger modelarea.

g. **Performance**

Several bug fixing and CPU and memory improvements have been implemented. Performance of the SNESIM tree algorithm in MPSLib has resulted in a speed improvement by a factor of 50. A new hybrid algorithm has been examined and tested that uses less memory and performs 24 times faster than the current SNESIM tree algorithm.

h. **Scripting and scaling performance**

MPS simulation is a very computationally intensive process and requires high computing power. Each MPS simulation must be done sequentially, which means the user can only perform one simulation at a time. Scripting and parallelizing the code is a way to overcome this and has been implemented so that simulation can be done faster and to a larger extent.

i. **Complex simulation**

During the modelling process of Indian Wells Valley (IWV), there was a need to perform multiple TIs in several areas. The former MPS simulation solution allowed the use of only one TI at a time. A new workflow that can use multiple TIs during the simulation in different areas has been developed. Also this functionality can be run in parallel thus allowing for reduced processing time.

7.2.2 Indian Wells Valley - I•GIS development and application of MPS model

7.2.2.1 Introduction

In the following our work, applying a multi-point statistics (MPS) method, to quantify uncertainties within the constructed Hydrogeological Conceptual Model (HCM) in Indian Wells Valley (IWV), California, U.S.A, will be summarized.

As shown in Figure 7.2, The area and basin of Indian Wells Valley is subdivided into two minor basins. In the south-western area we have the El Paso Basin whereas the main basin, the China Lake basin, is found in the larger region in the north-eastern part of IWV. The subdivision of the basin is due to a mapped fault, thought to be a hydrogeological barrier, dividing the basin into two minor basins. Furthermore, the two basins also vary in hydrogeological setting. This subdivision resulted in a focus on the main basin, China Lake, leaving the El Paso for future studies. AEM data were acquired in both the El Paso and the China Lake area.

FIGURE 7.2 Overview of Indian Wells Valley, showing the basin outlined in blue. El Paso Area is seen to SW whereas the main basin, China Lake, is seen to the NE.

China Lake is from the HCM divided into four major hydrogeological zones, summarized as Hydrogeological Zone 1 (HGZ1) consisting of mainly sand, HGZ2 consisting of mainly clay, but with sand lenses, HGZ3 consisting of mainly sand, and HGZ4 consisting of mainly consolidated sand.

The focus of the MPS study of IWV is on the HGZ2, as the presence of sand bodies, within what is defined from the HCM to be a clay layer, was known but widely undefinable. Local knowledge implies the existence of the sand within HGZ2. Furthermore, the bodies of sand were also visible within the AEM data, but the sparseness of the data made the mapping of the sand lenses impossible. With the MPS, the approach was to define a categorical uncertainty map, translated to a probability map, of the HGZ2. This was done with the scope of visualizing, from the categorical uncertainties, possible sand lens locations that were otherwise undiscovered/unmappable and to be able to estimate the plausible total storage potential of HGZ2. Uncertainties can, in this approach, be used to some extent, as a tool to designate areas where the possibility of having a sand lens is relatively high, directly translated from a low uncertainty of sand from the MPS result, thereby directly helping in the mapping of the sand bodies within the HGZ2.

In the study area of IWV, the MPS code used is MPSLib, which is completely integrated with the GeoScene3D software and runs with the implementation of a user-friendly wizard. The three key inputs to the MPS method are (1) the softdata being probability grids, holding the estimated probabilities directly from data, (2) the harddata which in the case of this study area are previous realizations from neighboring regions, and (3) the trainings images which are artificially created geological models that encapsulate to the maximum extent possible all prior geological knowledge and expected patterns. With these three inputs the MPS method simulation outputs an ensemble of categorical models (realizations) composed of the categories held within the training image, being clay and sand from HGZ2 of IWV. The MPS results are thereby developed from all three inputs to the MPS method and captures the variation and knowledge to be gained from each.

For the softdata both AEM data and well data were acquired and used. Also, as IWV is heavily dominated by salt/brackish water, affecting the AEM data, total dissolved solids (TDS) measurements were taken into account. Due to, in general, poor well data, extensive work was conducted sorting out the wells and we were forced to use the wells as only softdata.

7.2.2.2 Application of workflow

Working with the MPS method in IWV inspired the development of a workflow to be used in GeoScene3D and the MPS developments within. The developed workflow used for the MPS in IWV can be outlined in six steps, (1) Understand the conceptual geology, (2) Understand data, (3) Define an approach, (4) Create TIs, (5) Process data and (6) Run MPS.

(1) - It is a key to all later work to get the right conceptual understanding of the geology. The HCM developed in this project as well as the relevant literature was used as a base for the geological understanding. Also, discussion with partners with geological understanding of the local setting was invaluable for the work to be done.

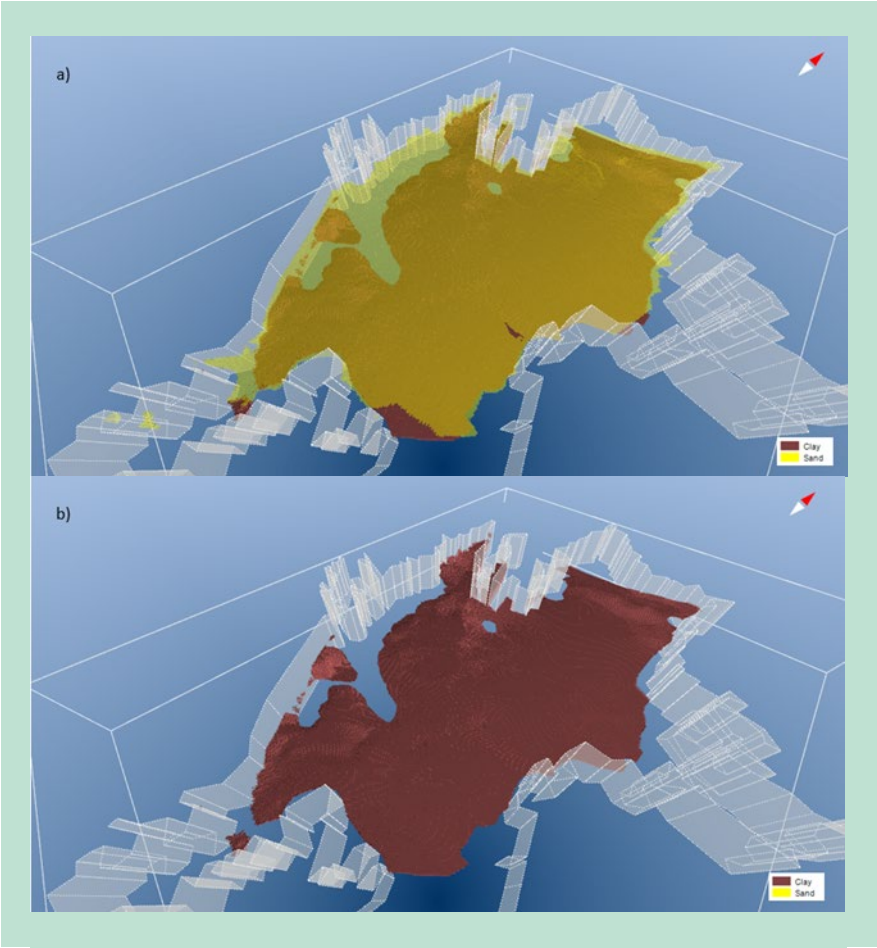


FIGURE 7.3 Hydrogeological conceptual model. a) visualizes the full HCM of the China Lake Basin in IWV, with sand as yellow and clay as brown. The basin boundaries are visualized by the grey, slightly transparent, region. The clay of HGZ2 can be seen between the sand of HGZ1 and sand of HGZ3 which have been turned slightly transparent giving a brown-green color. On b) can HGZ2 be seen, where HGZ1 and HGZ3 have been removed. The top of the model visualized on a) equals the surface boundary.

(2) - When moving forward in the workflow the key is to understand the extend of appropriate use as well as the limits of the available data. Quality, resolution, data distribution and data interpolations acceptable within (1) are all factors that will have to be considered.

(3) - The understanding of the geology drawn from the HCM work (1) and the knowledge of what were and what were not possible within the reach and resolution of the data (2), led to the focus of the China Lake area as well as the approach to define the uncertainties within the HGZ2.

(4) - Using the knowledge gathered from (1), the need for five training images became apparent and could be created encapsulating prior geological knowledge of the focus area settled from (3).

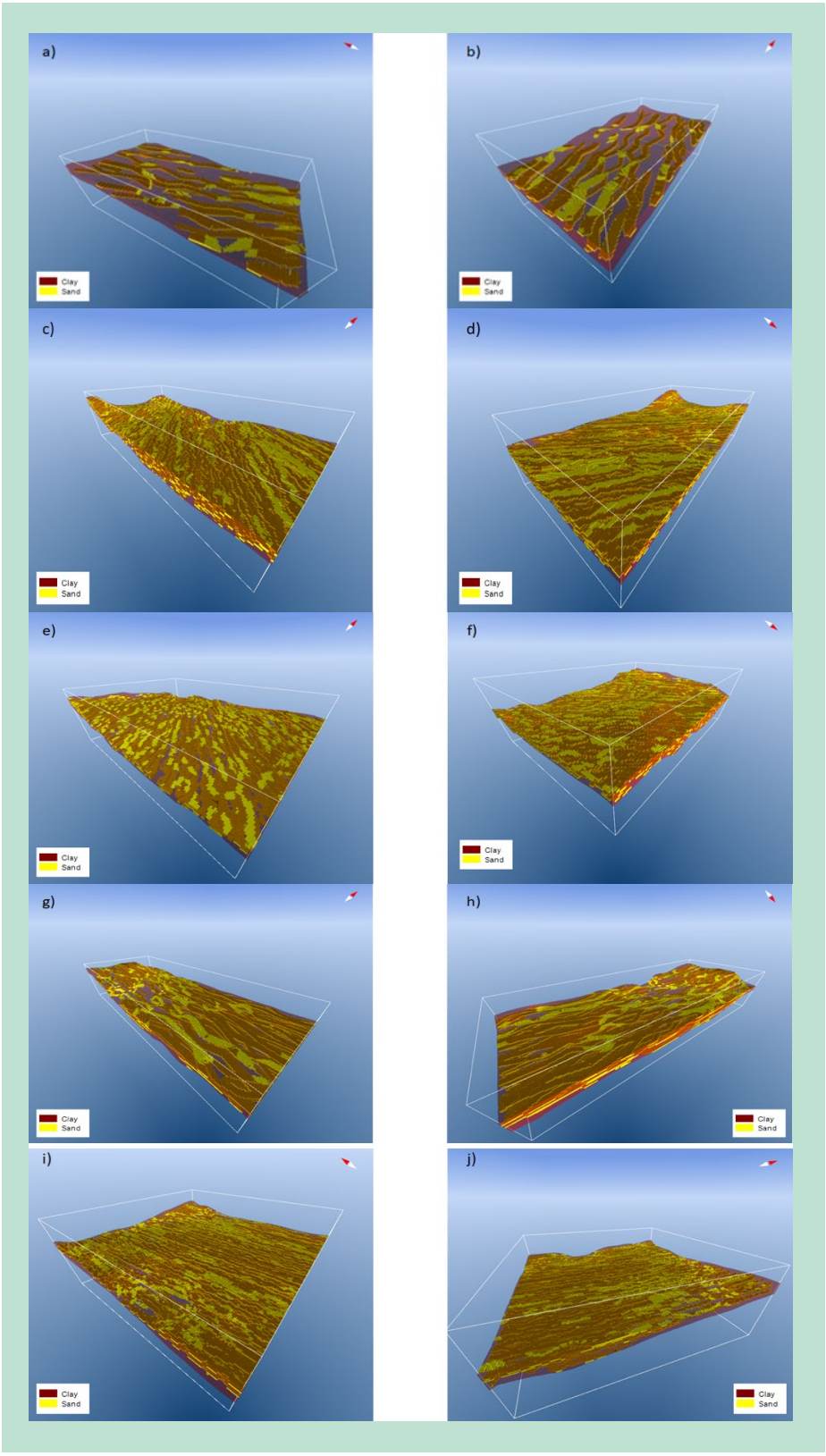


FIGURE 7.4 TIs used for the five regions. a) and b) shows the TI for Region 1, c) and d) the TI for Region 2N, e) and f) the TI for Region 2S, g) and h) the TI for Region 3 and i) and j) the TI for Region 4. Clay is visualized as brown, which have been made slightly transparent to be able to inspect the sand bodies within the clay layer. The sand bodies are yellow. The TI for Region one has been created with a fraction of 10% sand, whereas the rest of the TI's have been created with a fraction of 25% sand. The fractions have been calculated based on all

borehole data. The quality of all the wells have in this case been estimated to be good enough to be included for this calculation. More about the creation of the TI's can be found in the separate document "MPS Training ImagesIWV.pptx".

(5) - Looking at the learnings in regards to data from (2) and also the knowledge from (1) in terms of the regional geology, it was possible to process the data in the appropriate way, leaving data processing results applicable to the MPS method. For example, the TDS values were taking into account when processing AEM data and the well data were divided into groups dependent on their quality both in concern to lithology description and knowledge of their specific location.

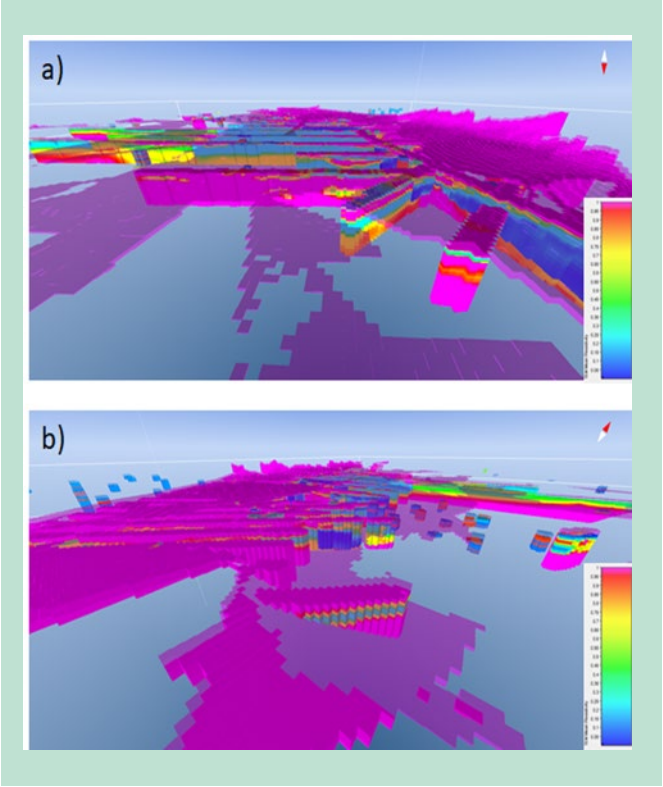


FIGURE 7.5 Softdata. Probability of sand. On a) the data is seen from the north looking south, whereas the data on b) is seen from the SE looking towards NW. The softdata added from the water table (interpreted as the having a probability of sand of 100%) have been made slightly transparent making it possible to see through and inspect the softdata seen beneath created from the AEM and well data. The warm colors indicate high probabilities of sand whereas the cold colors are representing low probability of sand, reflecting a high probability of clay.

(6) – The processed data resulting in softdata (5) and the created TIs (4) were used as the inputs to the MPS. The MPS were run in a succession, changing the TIs respectively to the region, resulting in a one final complete realization from one succession. This were repeated multiple times to output an ensemble of realizations as the result of the MPS method.

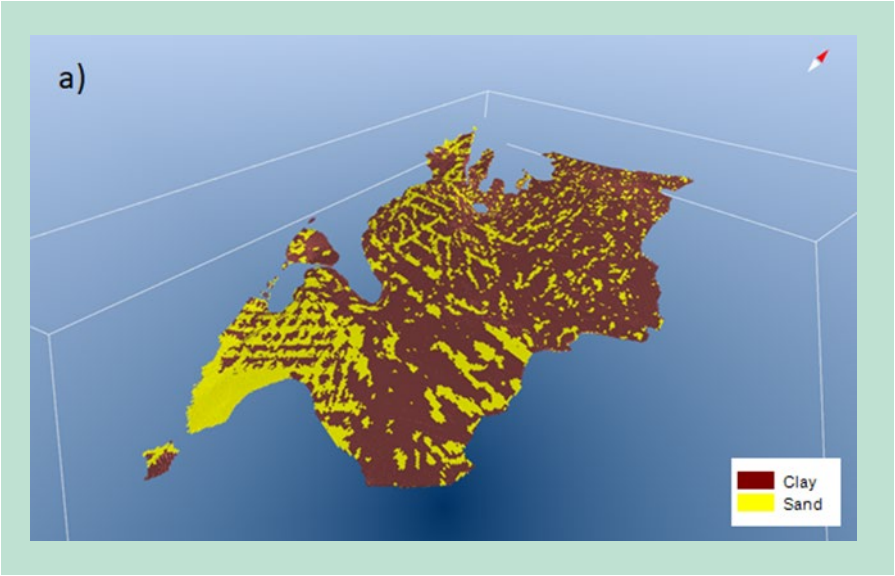


FIGURE 7.6 MPS results - A single realization of the full China Lake basin, following the outlined workflow. Clay is visualized in brown, whereas sand is visualized in yellow.

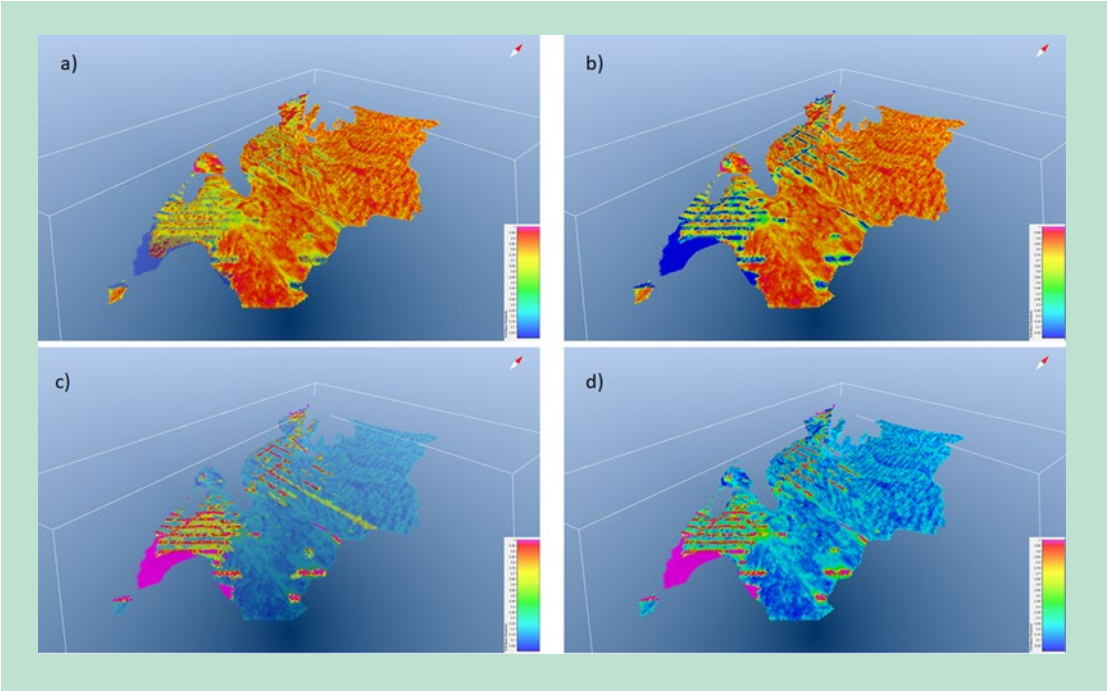


FIGURE 7.7 E-Type – Probability. On a) and b) is seen the probability of clay, whereas on c) and d) is seen the probability of sand. On a) and c) all probabilities lower than 67% have been made slightly transparent, leaving only the areas within the E-Type with a higher probability of the reflected lithology. It is seen that the probability of sand within the HGZ2 clay layer in some areas is high.

The workflow was applied in the outlined order and used as such. IWV is a complex project area and the understanding of the geology led to an understanding of a need to subdivide the China Lake Basin into five independent regions. Each region was defined based on differences in sand lens orientation, dip and size (horizontally and vertically). As the MPS algorithm is stationary, in the sense, that the TI cannot vary within different subareas of the area, i.e. it cannot catch the change in the overall direction or size of features going from one part of the basin to another. To adapt to the stationarity of the MPS algorithm the need to create a TI for

each individual region became important. This allowed us to capture the change in overall direction or size of features, that is believed to be present in different subparts of the larger basin, in the TIs that relates to one region each. The multiple TIs, trying to capture the overall variation in geological patterns as a set, resulted in a need to create a special workflow that made it possible to use all TIs in a single MPS simulation capturing the regional variations. Developments were made to make it possible to run the full China Lake basin, outputting resulting realizations without boundary effects, thus securing continuity between the five regional zones. The development is based on a user defined succession between the regions, where each region will use a specially designed TI coherent to that specific region. Running the first region will create a realization, which is then used as harddata input for the simulation of the next area, using a different TI, ensuring a seamless transition between the two regions. This continues in a user defined order until the full area of the China Lake basin has been simulated. The result is one complete realization. Softdata from the full area of the China Lake basin will be used in each part of the simulation outputting the complete realization of the China Lake basin. This process is then repeated a defined number of times, giving the desired number of realizations as the MPS output and result.

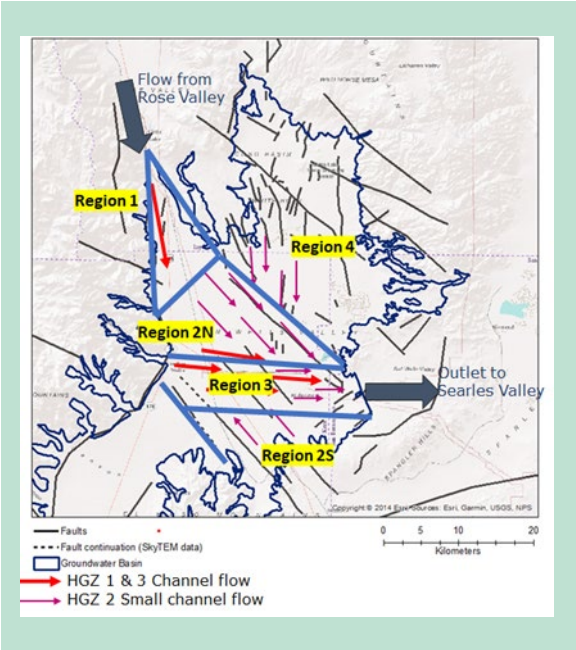


FIGURE 7.8 The five regions and their interpreted differences in orientation, dip, and size. It is seen that “Region 1” contains larger channel structures, orientated and dipping from NNW-SSE. “Region 2N“contains smaller channel structures. The orientation and dip are from NW-SE. “Region 2S is similar to Region 2N but differs in dip. The orientation is from SE-NW. “Region 3” is primarily larger channel structures in an orientation and dip from W-E. The last region, “Region 4”, is defined of smaller channels in an N-S orientation and dip.

Furthermore, a three zone TDS structure was created, varying both horizontally and vertically. The three TDS zones were defined as freshwater, brackish water and highly saline water, which all varied in extent dependent on the HGZ. Taking the TDS into account is critically moving forward when looking to process AEM data. The AEM data have been divided with respect to the TDS zones and their extent in the horizontal and vertical orientation. This resulted in three individual AEM sub-datasets, one for each TDS zone. For each of the zones a transfer function, transferring resistivity to lithology, specific to the TDS value of that zone, was created and applied to the coherent AEM sub-dataset.

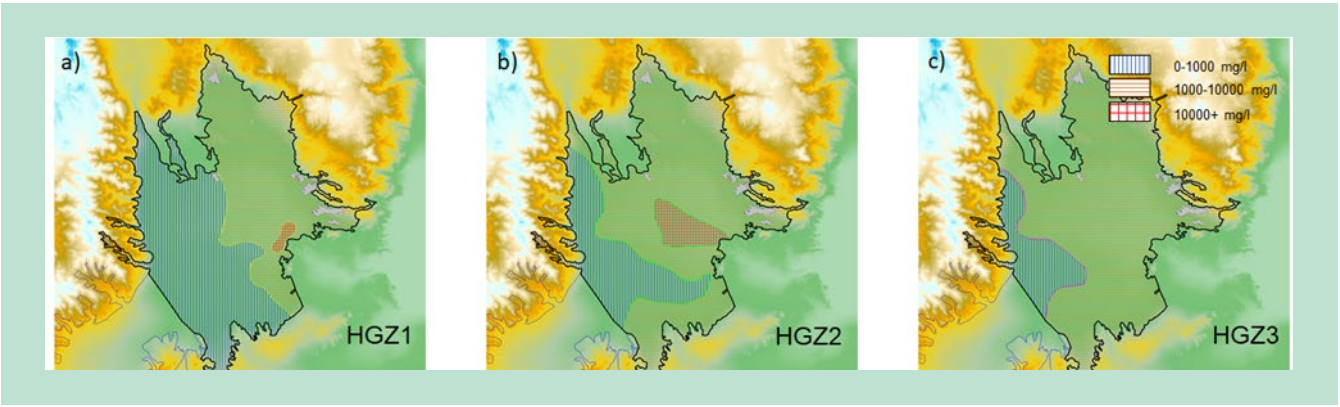


FIGURE 7.9 TDS Zone variation. On a) is seen the three TDS zones of HGZ1, b) shows the three TDS zones of HGZ2 and c) shows the three TDS zones of HGZ3. The freshwater zone is shown in blue, brackish water in orange and the highly saline zone in red. Freshwater is defined from 0-1000mg/l, brackish water from 1000-10000mg/l and highly saline water is 10000+mg/l.

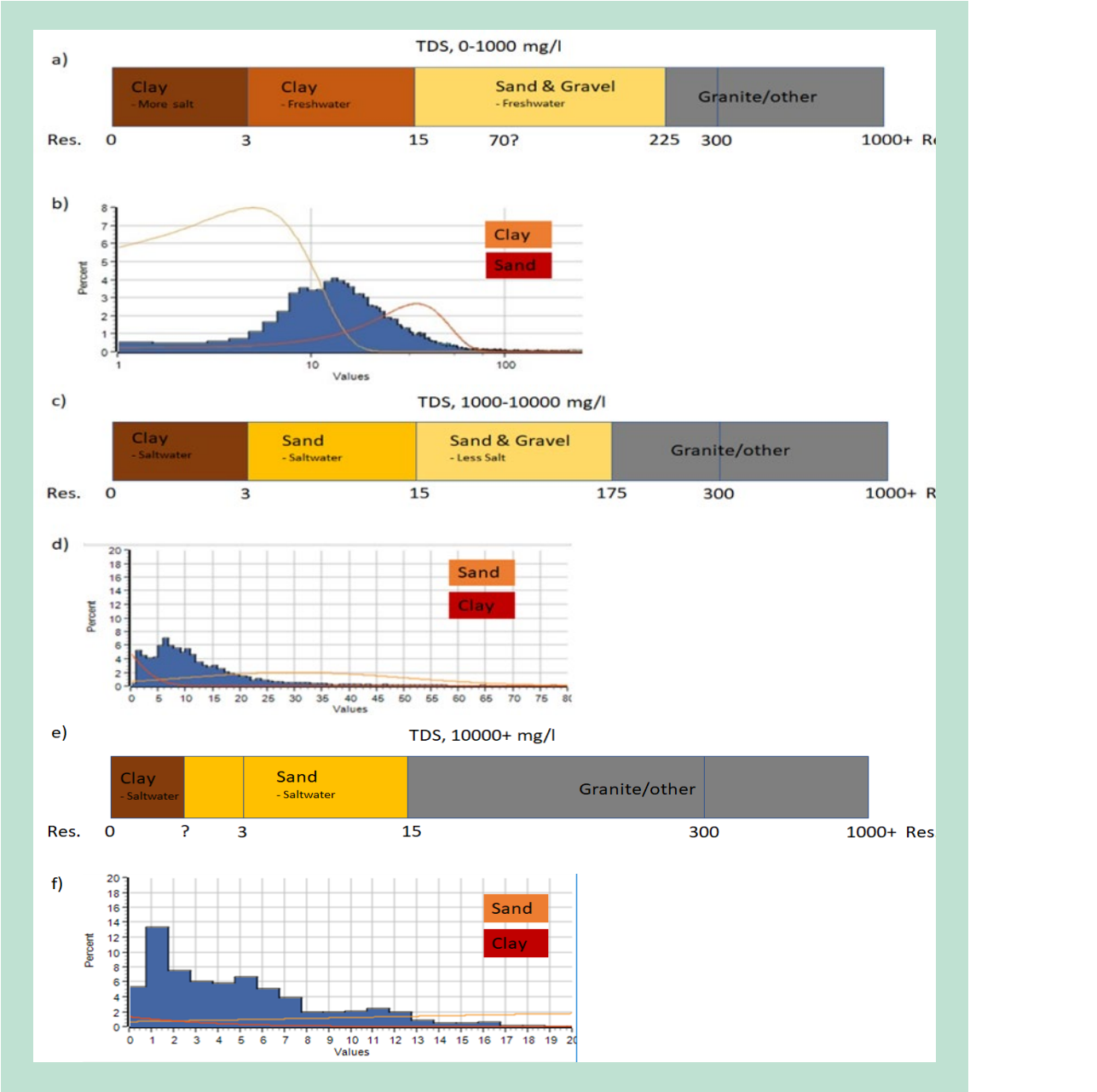


FIGURE 7.10 Visualizing the resistivity to lithology interpretations used to create the transfer functions within GeoScene3D. Three resistivity to lithology transforms were created to take the TDS data into account. a) represents the interpretations made for the freshwater zone, c) for the brackish water zone and e) for the highly saline zone. The transfer functions do not hold sharp shifts between the two categories (sand and clay) as can be seen on the interpretations above (a, c and e), but are gradually changing the resulting probabilities. The transfer functions used can be seen on b), d) and f). E.g. on a), the probability of sand will gradually increase moving above the resistivity of 15 Ohmm, whereas the probability of clay will gradually decrease which can be seen on b). The curves in b), d) and f) are representing the two lithologies, whereas the histogram seen behind the curves visualize the count of a certain resistivity within the AEM data. Resistivities in Ohmm.

Data preparation has been a large part of the work done in IWV. The data preparation has been done to ensure that data were loaded and usable to the MPS method in the best possible way. Regarding the well data, a quality system was developed to handle the differences in quality of the individual wells. All wells have been categorized into quality groups based on their lithological description and the knowledge to the well location. The wells were given a quality of 0-4, where 4 is the best, based on their lithological description. Furthermore, the wells were given a quality of “Bad”, “Medium” and “Good” based on the knowledge to the wells actual position in the X and Y coordinate. For the MPS it was defined that all wells with the quality “Bad” were of such a poor quality, that they would only add more uncertainty to the MPS simulation than they would benefit it. Therefore, they were left out.

Within the qualities of “Medium” and “Good” the wells would only have qualities of 2-4 in lithology description, leaving six groups of wells. These six groups of wells were given individual confidence ratings which resulted in individual probability settings from the data processing in the creation of softdata. The well probabilities, dependent on their quality, varied from 55% to 90% probability of either of the two categories clay or sand. Overlap in X and Y coordinate could, to some extent, still be seen between the six groups. This was handled by a priority merge, where softdata created from wells in the group of “Good” with a lithology stamp of “4”, were prioritized before wells with e.g. “Good” and “3” respectively to their location knowledge and their lithology descriptions. The priority merge was handled in an intelligent way, where a deep “Medium”, “2” quality well would still leave information in the deeper intervals if an overlapping “Good”, “4” quality well was of shallower depth. The resulting softdata grids from both AEM data and well data was merged into two complete softdata grids, one for each of the two lithologies defining the probabilities based on data and used for the MPS method.

The result was 75 realizations which were gathered into several E-Types. The E-Types define the overall uncertainty within HGZ2 of IWV, for each lithology, based on all 75 realizations. The E-Types can be used, for example, as input to conclusions on total storage capacity of HGZ2 and possible mapping of sand bodies.

To estimate the total storage capacity of the clay layer HGZ2 in IWV, we used the generated E-Type showing the probability of sand in the unit (Figure 7.7). Assuming a 20% porosity, a cumulated water volume was calculated using different thresholds. Thresholds are based on the probability of sand in the E-Type. For example, setting the threshold at 80% probability, which means that all areas with a probability of sand equaling 80% or higher results in a sand volume, gives an estimated water volume of approx. 560.000.000 m^3 (Figure 7.12).

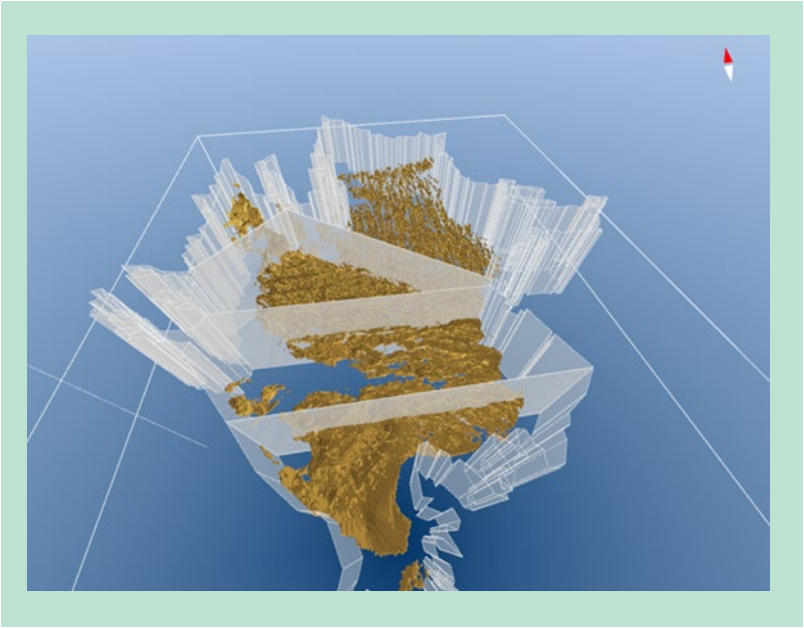


FIGURE 7.11 Shows a realization of Indian Wells Valley. Only sand volumes within HGZ2 are visualized. The visualization is by an isosurface. The Indian Wells Valley China Lake Basin is shown by a slightly transparent region boundary. Within is seen the boundaries defined and used by the MPS in terms of used TIs and outputs.

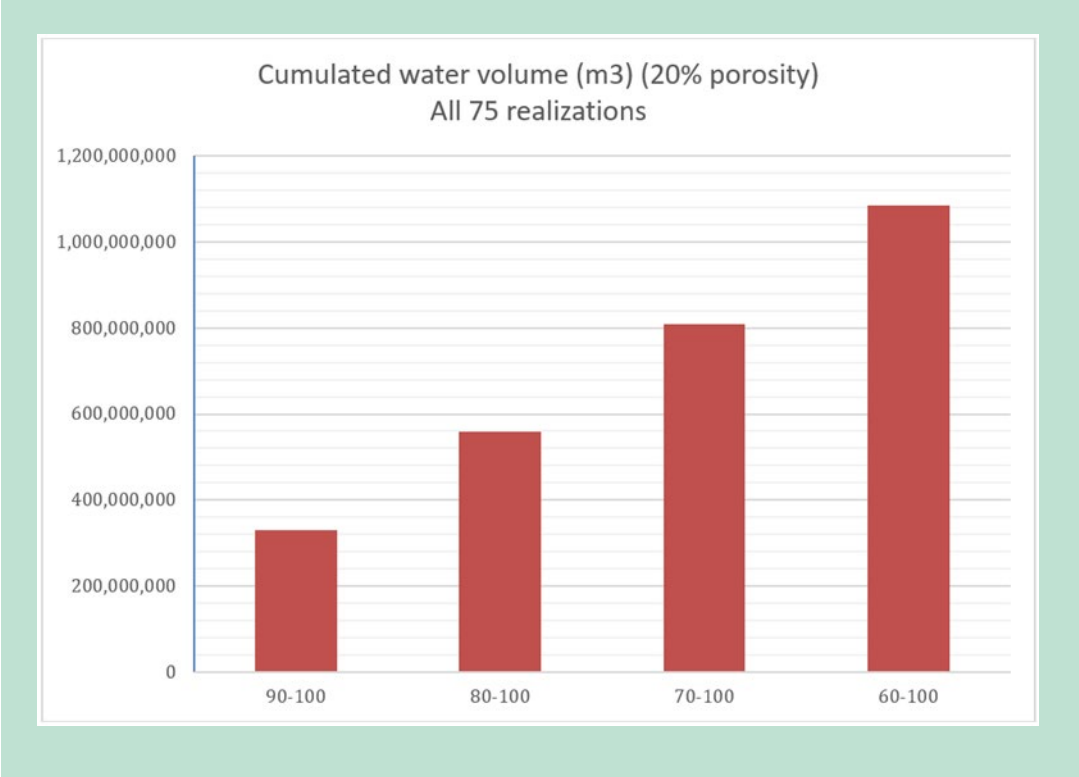


FIGURE 12. Shows the estimated storage capacity of HGZ2 in Indian Wells Valley, based on different threshold values with an assumed porosity of sand of 20%. On the Y-axis the estimated water volume in m^3 can be found. On the X-axis the threshold values are found, giving an interval from the threshold to 100% probability. Threshold values are in percentage. E.g. a threshold of 90% probability leaves an interval of 90% to 100%. The estimated water volume within this interval is approx. 330.000.000 m^3 of water.

7.2.2.3 Preferential path in Indian Wells Valley (IWV)

As part of the GAP a preferential path algorithm was implemented in the MPS software. The preferential path has been tested on the IWV dataset.

Using preferential path makes MPS realizations truer to softdata. The softdata used for this area were, among others, from the AEM data, which was collected in lines and considered non-collocated. By using the preferential path algorithm in the IWV study area it was seen that wells, for instance, though given a lower trust worth in terms of their probability in the softdata, could create vertical elongated structures non-connected to other structures of equal lithology. Such structures are interpreted as unrealistic in the area of interest. Therefore, the random path algorithm was chosen to be used in the IWV case.

The reason for this seems to be inconsistencies between the different datasets and the conceptual model but looking deeper into these problems is outside the scope of the GAP project.

Below are results from running preferential path realizations and examples comparing the random and the preferential path algorithm.

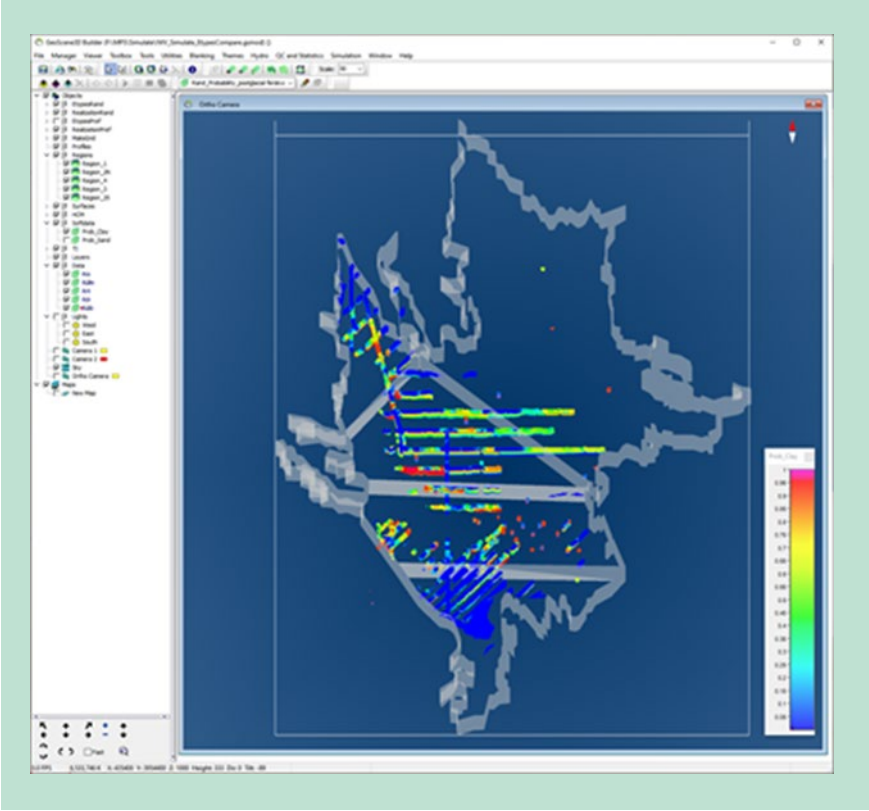


FIGURE 7.13 IWV softdata are lines from AEM data.

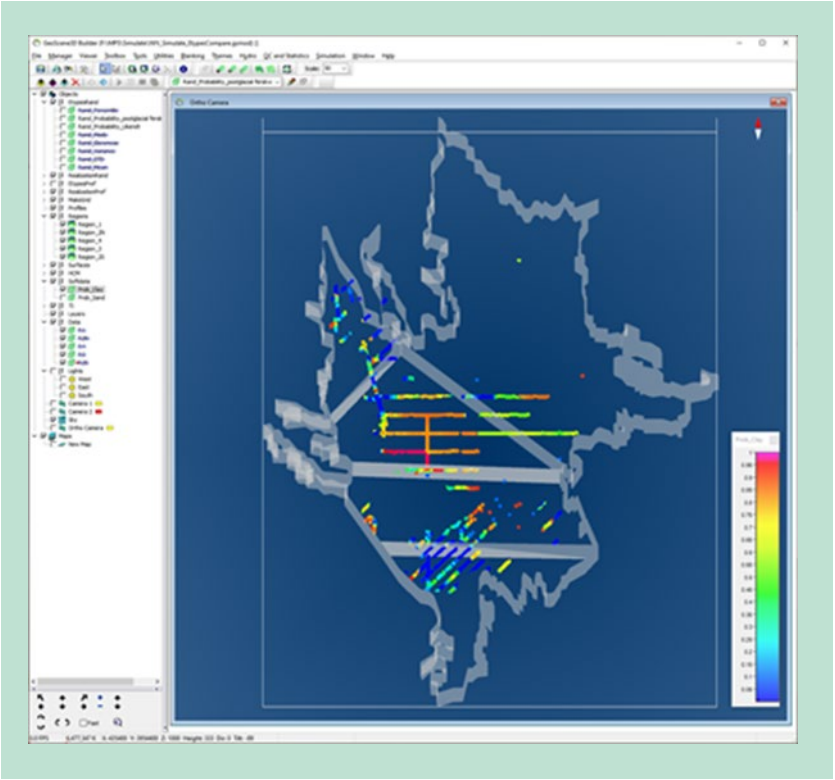


FIGURE 7.14. Soft data slice at elevation 580.

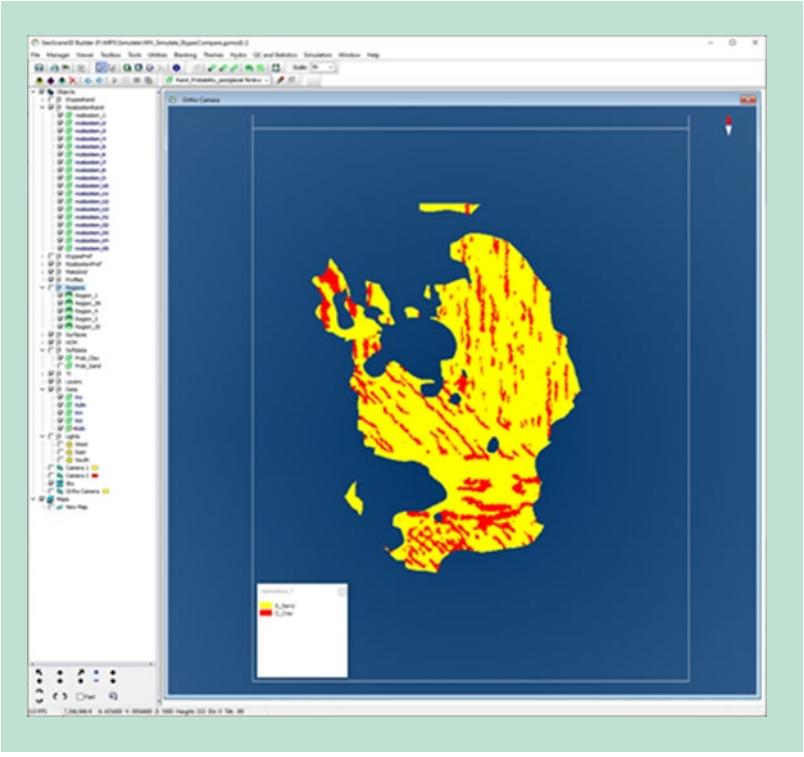


FIGURE 7.15 One realization using random path. Slice at elevation 580

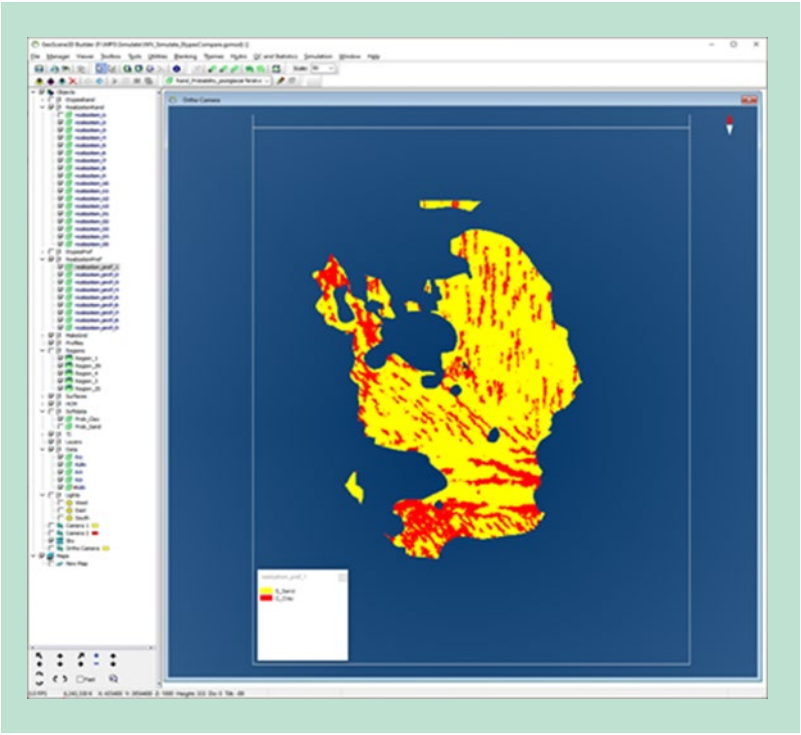


FIGURE 7.16 One realization using preferential path. A slice at elevation 580.

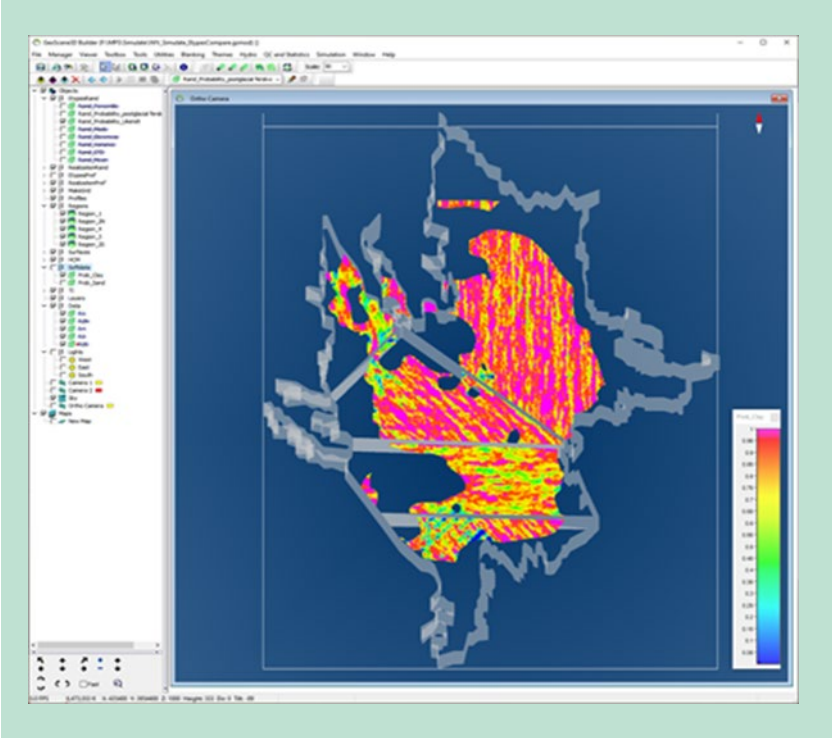


FIGURE 7.17 Probability of clay from 9 realizations using random path. This is a slice at elevation 580.

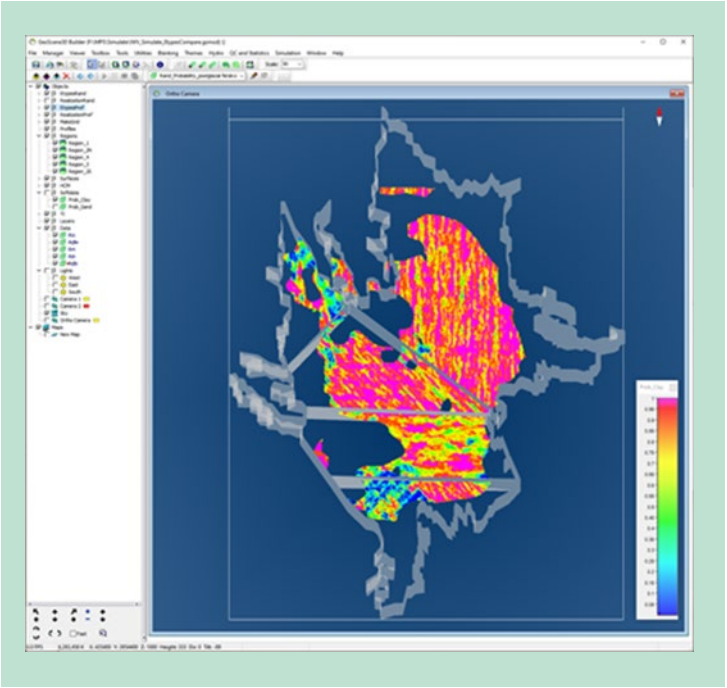


FIGURE 7.18 Probability of clay from 9 realizations using preferential path. This is a slice at depth 580.

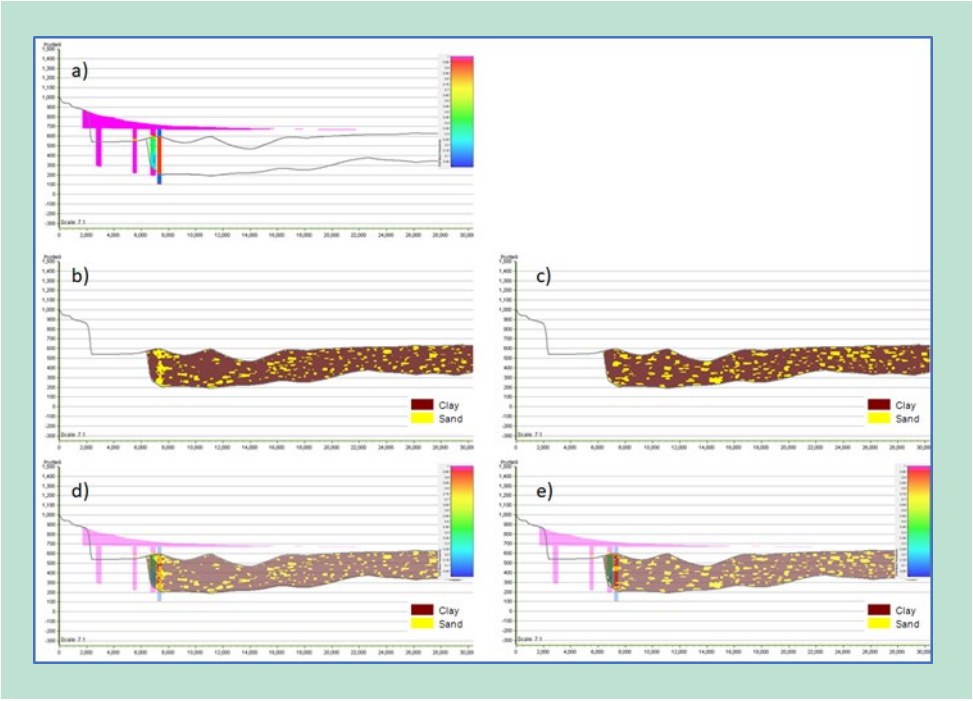


FIGURE 7.19 Visualization of the issues in the IWV data using preferential path. On a) is shown a cross section of the softdata, with the HCM grids represented as grey lines. b) and c) visualize two independent realizations. b) is from using preferential path, whereas c) is from using random path. On d) and e) it is possible to see that by using the preferential path (d)), the realization creates an unnatural vertical elongated structure, not representing the geological understanding of the China Lake basin. e) shows that the random path supports the geological understanding and creates more natural geological structures.

7.2.3 Conclusions/recommendations

The MPS workflow and software developments in this WP was applied to the real-life case of Indian Wells Valley abstraction problem. The actual case proved to be very complex as several geological and geophysical challenges were identified. The MPS approach proved applicable and was able to answer questions valuable for the study area, although the need for further data to support the model was also evident in order to solve some of the uncertain aspects of the model. The new tool developments were also able to streamline workflow and actually make it possible for an average skilled user to construct MPS models.

8. WP4: Uncertainty analysis

Major contributors: Seogi Kang and Rosemary Knight (Stanford University); Todd Greene (California State University, Chico); Christina Buck (Butte County); Graham Fogg (University of California, Davis)

An approach is required that makes it possible to quantify uncertainty and rigorously account for its propagation through the workflow that leads to the development of the conceptual model. The conceptual model is used as the basis for the development of the groundwater model, so uncertainty must be quantified in such a way that it informs decision-making. A novel approach to quantifying and communicating uncertainty was developed working in close collaboration with local specialists with the needed local and contextual knowledge.

Our study area covered two regions within Butte and Glenn Counties in the Central Valley of California. The primary interest of the water agency in this area, in acquiring AEM data, was to obtain an improved delineation of the large-scale structure and heterogeneity of the aquifer system so as to better understand the extent of vertical connectivity. The existing groundwater model in the area, developed using well data, lacked sufficient detail to understand the connection between various depths within the aquifer system.

8.1 AEM interpretation workflow

Our workflow, developed to obtain and interpret an ensemble of sediment-type models from AEM data and well data, included two main steps: the inversion of the AEM data and the resistivity-to-sediment-type transform, which is often referred to as the rock physics transform. In the AEM inversion step, we first sought to recover a resistivity model fitting the observed AEM data for a given prior model, then subsequently applied posterior sampling to obtain multiple resistivity models. This procedure was repeated with a variable prior model. Our prior model includes two terms. The first one is the spatial constraint, favoring a smooth transition of resistivity values between adjacent sounding locations. The other one is a reference model that can integrate prior information constraining the inversion based on ancillary data (e.g., well data), or an understanding of the expected variation of electrical resistivity in a survey area (e.g., the average resistivity). In the second step, we constructed a rock physics relationship between resistivity and sediment type to transform all of the resistivity models into models containing information about sediment type, and the percentage of sediment type. Interpretation of the models integrated prior knowledge of the aquifer system, which included head measurements from multi-completion wells. We first describe how the workflow was implemented then present the results of the workflow.

In the AEM inversion step, we ran six different AEM inversions and obtained six different recovered resistivity models; the six sets of inversion parameters are summarized in Table 8.1. Inversion 1 only uses the spatial constraint in the prior model, whereas Inversion 2, 3, 4 use both the spatial constraint and the reference model by setting an inversion parameter, α_s, to be 1; each inversion uses different resistivity values to set a homogenous reference model. Inversion 5 and 6 used the resistivity logs to create an inhomogeneous reference model, ρ_int^log, and a cell-based weighting, w_IDW; the cell-based weighting was to compensate AEM soundings far away from the resistivity logs. For each of six inversions, we applied the posterior sampling 1000 times, so a total of 6006 resistivity models were obtained: 6 recovered models and 6000 from posterior sampling.

The second step of the workflow is the rock physics transform. We developed and applied two forms of the transforms to obtain information about sediment type from the resistivity models. One form was used to map each resistivity value to sediment type, defined as either sand/gravel or clay/silt. The other form was used to map resistivity to the percentage of sand/gravel. A total of 55 pairs of co-located 1D layered-resistivity and lithology logs were used in this process. Given that water saturation will have a significant impact on the resistivity of lithologic units, we developed separate distributions for the sediment types above and below the top of saturated zone (TSZ), using the estimates of TSZ throughout the study area obtained by Dewar and Knight (2020). Using both forms of the rock physics transform, we obtained 6006 sediment-type models and 6006 coarse-fraction models.

TABLE 8.1 Six different sets of inversion parameters. For Inversion 5 and 6, we generated an interpolated resistivity model using 152 resistivity logs, ρ_int^log, and this was used as a reference model. The inverse distance weight, w_IDW, used for the interpolation, was used as a cell-based weighting, w_s, for the reference model.

Inversion number	$\boldsymbol{m_0}$	$\boldsymbol{m_{ref}}$	$\boldsymbol{\alpha_s}$	$\boldsymbol{\alpha_z}$	$\boldsymbol{\alpha_r}$	$\boldsymbol{w_s}$
1	10 Ωm	N/A	N/A	1	5	1
2	10 Ωm	10 Ωm	1	1	5	1
3	10 Ωm	20 Ωm	1	1	5	1
4	10 Ωm	30 Ωm	1	1	5	1
5	10 Ωm	ρ_{int}^{log}	0.1	1	5	w_{IDW}
6	10 Ωm	ρ_{int}^{log}	1	1	5	w_{IDW}

8.2 Results from Butte County

Our interpretation workflow generated a model space that included 6006 resistivity models, and a corresponding 6006 models displaying sediment type and 6006 models displaying coarse fraction. In contrast to the approach typically taken in the interpretation of geophysical data (a single resistivity model with a single interpretation), this is an enormous model space that can be explored to address the issue that defined this study: an improved delineation of the large-scale structure and heterogeneity of the aquifer system so as to better understand the extent of vertical connectivity.

The six resistivity models, recovered using the parameters for the six inversions described in Table 2, all captured very similar large-scale structure, but varied significantly in areas where the AEM data were less sensitive. Given the high quality of the resistivity logs from the survey area, we felt that incorporating information from the logs in Inversion 6 significantly improved the accuracy of the recovered model. Given our level of confidence in this model, it was used to construct the rock physics transform. The sediment-type model transformed from the primary resistivity model was selected as the primary sediment-type model.

8.2.1 Interpretation of sediment-type models to obtain large-scale structure

In Figure 8.1a we show a 3D view of the primary sediment-type model displaying only those regions classified, through use of the transform, as being sand/gravel. The depth-of investigation (DOI) is presented as a blue transparent interface. We knew that there was significant spatial heterogeneity within the aquifer system in this area, making it highly unlikely that thick

homogeneous packages of sand/gravel or clay/silt would be present. We therefore, in interpreting sediment-type models, referred to the two units as “sand/gravel-dominated” and “clay/silt-dominated”. In Figure 8.1b we show the same regions, displaying the probability of sand/gravel-dominated calculated using the other 6005 sediment type models, where the probability in a cell is equivalent to the percentage of models having the same sediment-type mapped in that cell; the probability ranges from 30% to 100%.

The results derived from our AEM data were used to obtain information about the large-scale structure and heterogeneity within the aquifer systems. In Figure 8.1c, we show a vertical section through the model in Figure 8.1b with our interpretation. Seen in this section is the large-scale structure observed in other sediment-type models. Most of the probability values in the displayed regions are above 50% indicating that this structure is present in the majority of models. In the eastern part of the study area, there appears to be a continuous package of sand/gravel-dominated starting at the East-to-West (E-W) boundary and extending to the eastern limit of the study area; this E-W boundary separates the eastern and western parts of the region as shown in Figure 8.1c. We refer to this package of sand/gravel-dominated as the eastern zone; the top and base of the eastern zone are mapped with red and yellow surfaces, respectively, in Figure 8.1c. In the western part of the study area, which corresponds to the western side of the E-W boundary, we identified three zones: an upper zone which is primarily sand/gravel-dominated, a middle zone which is primarily clay/silt-dominated, and lower zone which is primarily sand/gravel-dominated. The extent of these zones is shown in Figure 8.1c using three surfaces: the base of the upper, middle, and lower zones shown as blue, red, yellow surfaces.

It is the information about probability, extracted from our model space, that we used to quantify the uncertainty, in each cell, in our ability to identify sediment type. When a cell has a probability of 100%, in terms of corresponding to one of the sediment types (sand/gravel or clay/silt), the uncertainty is defined as 0. As the probability for either sediment type,moves towards 50%, the uncertainty increases, reaching a maximum of 1 when the probability of both sediment types equals 50%. This uncertainty is displayed in Figure 8.2. The dominant factor determining uncertainty is an inability to resolve the variation in resistivity; i.e. we cannot accurately determine the resistivity value in each cell. The ability to resolve the resistivity decreases with depth, with the depth of investigation (DOI) defined to indicate the depth at which we lose resolving ability. It is important to note that above the DOI there can be significant levels of uncertainty in areas where there is high spatial complexity so that resistivity is changing rapidly vertically and/or laterally. At the boundaries between the two sediment types, the smoothness constraint used in the inversion will result in recovered resistivity values close to the threshold resistivity value separating the two sediment types. Thus, small variations in recovered resistivity can easily change the resulting sediment type determined from the transformation. This results in a high level of uncertainty in sediment type at any interface between the two sediment types.

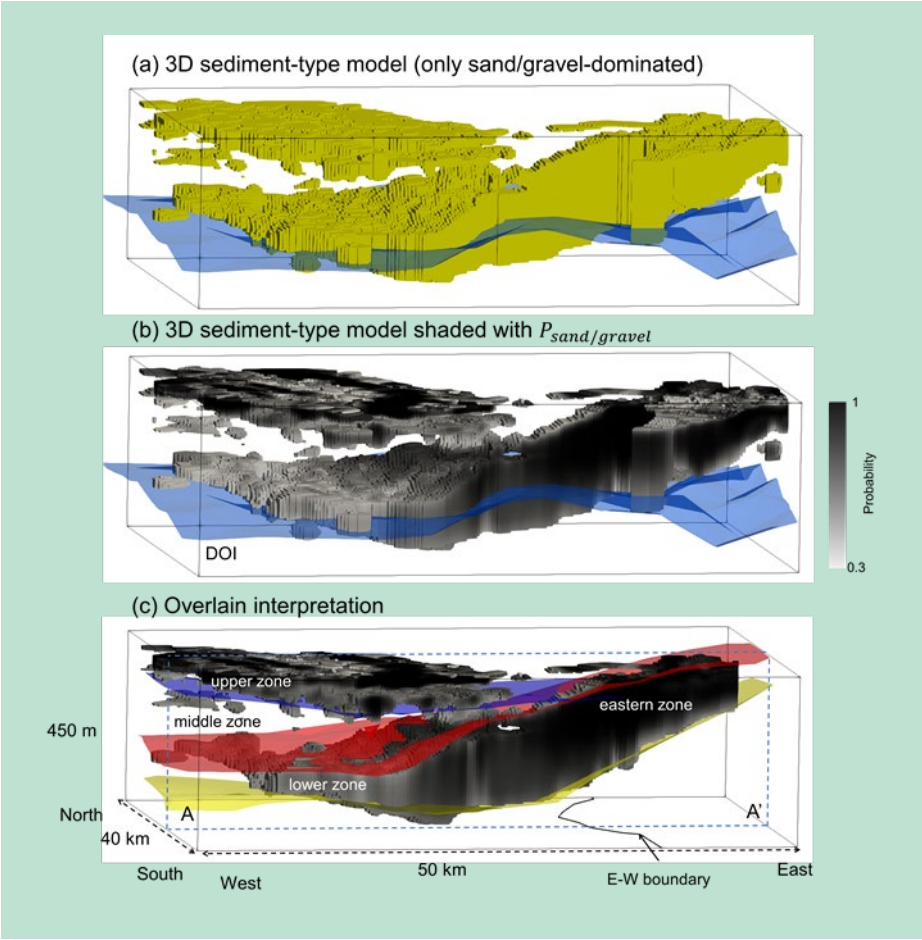


FIGURE 8.1 The results of the AEM interpretation workflow in a 3D view. (a) A display of sand/gravel-dominated (yellow color). (b) The regions shown in (a) but shaded with the probability of sand/gravel-dominated, P_(sand/gravel). (c) Overlain interpretation. The E-W boundary shown as a black line separates the eastern and western part of the regions. In the western part of the region, the upper, middle and lower zones are delineated by the three surfaces - the base of the upper zone (blue), the base of the middle zone (red), the base of the lower zone (yellow). In the eastern part of the region, a single sand/gravel-dominated zone is identified, which is denoted as the eastern zone; the red and blue surfaces in this region correspond to the top and base of the eastern zone.

The image of uncertainty in Figure 8.2 displays higher uncertainty with increasing depth but also maps surfaces displaying high uncertainty at shallower depths. At the large scale, we found high uncertainty coincident with the surfaces defined as separating zones in the aquifer system. These surfaces separate zones dominated by different sediment types so, as described above, are expected to be surfaces of high uncertainty. At the smaller scale, within the zones, there are differences in both the average level and in the variability of the uncertainty which we also interpret as related to the presence of interfaces across which there are changes in sediment type. The observed complex patterns in uncertainty within the upper zone indicate a complex interlayering of sand/gravel and clay/silt. The eastern zone shows a low level of uncertainty throughout the zone, suggesting less fine-scale spatial variability in sediment type compared to the other zones. The image of uncertainty in Figure 8.2 reveals numerous areas where we cannot interpret, with certainty, sediment type but we can interpret – with certainty – the presence of changes in sediment type, over short vertical distances.

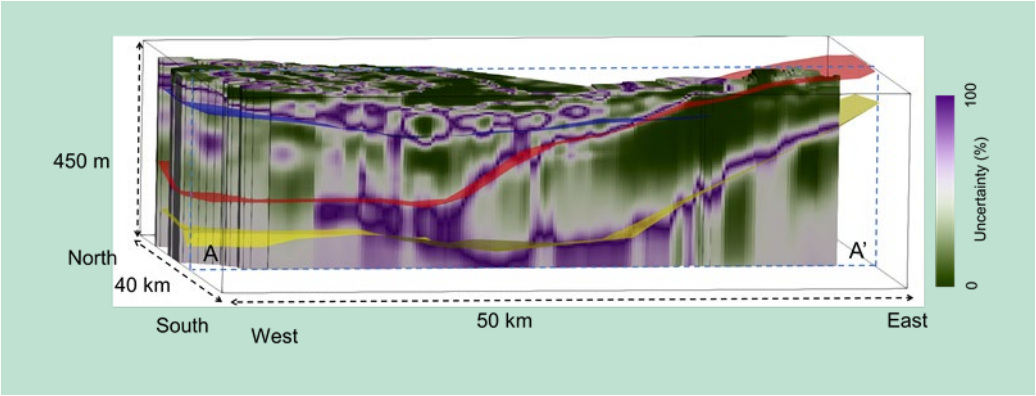


FIGURE 8.2 The estimated uncertainty in sediment type from AEM data in a 3D view. The 3D uncertainty extends from the vertical section A-A' the location of which is shown in Figure 5, and extends to the north. Blue dashed lines indicate the location the A-A' vertical section in the 3D view.

8.2.2 Interpretation of coarse-fraction map to obtain information about vertical connectivity

The AEM method is limited in terms of the ability to resolve spatial variability in sediment type, but the recovered resistivity values will be highly sensitive to the volume fractions of the sediment types that are present. This can provide information about vertical connectivity in an aquifer system. For this, we calculated the vertical average of the primary coarse-fraction model, between the TSZ and the base of the eastern zone in the eastern part of the study area, and between the TSZ and the base of the lower zone in the western part of the study area. The resulting coarse-fraction map is shown in Figure 8.3a. By repeating this process for all coarse-fraction models, we obtained many coarse-fraction maps and calculated the standard deviation displaying the uncertainty of the coarse-fraction map as shown in Figure 8.3b. The open circles shown in Figure 8.3a are the locations of the twelve multi-completion wells in the study area; these wells are labeled A to L, going from east-to-west. Each well is composed of two wells with head measurements made in two screened intervals. As an indication of vertical connectivity within the aquifer system, we calculated, for each multi-completion well, the minimum and maximum magnitude of the vertical head gradient, dh/dz between 2013 and 2018. Good vertical connectivity produces lower values of dh/dz, whereas poor vertical connectivity produces much higher values of dh/dz. The solid circles showing the well locations in Figure 8.3a are color-coded to display the maximum |dh/dz|. We observe a correlation between the values of dh/dz and coarse fraction from the AEM data, with the general trend of decreasing vertical hydraulic gradient with increasing coarse fraction.

There is not enough information available in the AEM data alone to make it possible to accurately quantify vertical connectivity within an aquifer system. We found, however, that deriving the coarse fraction from the AEM data provided an indicator of connectivity that could be used, in conjunction with well data, to assess the variability in connectivity throughout our study area.

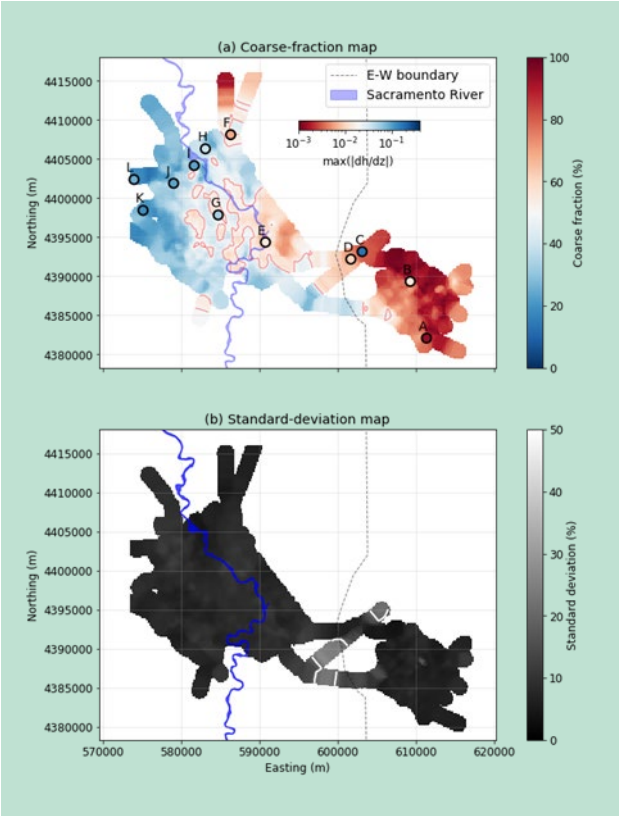


FIGURE 8.3 Interpolated maps providing information about the vertical connectivity of the aquifer system and the uncertainty. (a) The coarse-fraction map with the locations of 12 multi-completion wells indicated by circles, color coded with the maximum vertical hydraulic gradient. Letters A-L indicate the names of the twelve multi-completion wells. Red contours indicate 50% coarse-fraction values. The black dashed line indicates the E-W boundary delineating the horizontal boundary between the lower and eastern zones. The blue polygon indicates the Sacramento River. (b) The standard-deviation map represents the uncertainty; darker colors indicate low uncertainty (i.e., low standard deviation); white contours indicate 10% standard-deviation values.

8.2.3 WP4 Conclusions/recommendations

The AEM method can provide valuable information about aquifer systems that can support groundwater management efforts. The focus of our research is to develop ways of maximizing the information about aquifer systems that can be obtained from AEM data while capturing and communicating the relevant uncertainty. We advocate an approach that works with an extensive model space, so that mapping of uncertainty can become the standard practice when interpreting AEM data. As we found in this study, quantifying uncertainty can be used not only to communicate the level of confidence in the identification of sediment type, but can assist in identifying interfaces between sediment type, at both the large and fine scale. Continued research, focused on exploring the optimal way to interpret AEM data, and integrate these data with other forms of hydrologic data, will undoubtedly lead to the increased adoption of the AEM method as an integral component of subsurface characterization for groundwater science and management.

8.2.4 References

Dewar, N., and R. Knight, 2020, Estimation of the top of the saturation zone from airborne electromagnetic data: Geophysics, **85**, EN63–EN76.

9. Conclusions

The Groundwater Architecture Project has been a tremendous success. Over the past few years, we have moved from geophysics having very low visibility at the level of local and state agencies to one in which agencies are coming to appreciate its potential value. During the two years of the GAP, we acquired and interpreted AEM data in three areas – Butte and Glenn Counties, Indian Wells Valley and San Luis Obispo County. Through the GAP, working with local and state partners, there has been demonstration of the role that the AEM method can play in developing an improved understanding of the large-scale structure and the spatial heterogeneity of groundwater systems

The GAP was designed as a pilot project, with the expectation that we would develop an optimal workflow for applying AEM data in the development of hydrogeologic conceptual models. The GAP produced key findings that will inform the implementation of other AEM projects in California. As detailed in the preceding sections, it is the partnership between the AEM-geophysicists and those in the local study area that is the most important component in ensuring a successful project; in this respect we were extremely fortunate. With a focus on the three study areas, we completed the research required to identify the optimal workflow and addressed challenges in numerous components of the workflow. The full set of recommendations, for all components of the workflow, is given in the Appendices. In each of the components many of the recommendations are transferrable to other geographic areas.

With funding from MUDP, significant advancements were made in the way in which we process and invert data, construct the transform between resistivity and sediment type, integrate all available data to develop a model of the subsurface, and quantify and communicate uncertainty. Funding from other sources allowed us to focus on additional issues that can impact the success of an AEM-focused project. In particular, considerable time was spent on dealing with the multiple challenges associated with the well data. Descriptions of sediment type, which are essential for both planning and interpreting an AEM survey, were not available in machine-readable format; and well locations were not sufficiently accurate. There was also the challenge of establishing a data management system given the state of the well data and the lack of a central repository for the geophysical data. Through our work in the three study areas, we developed solutions that are transferrable elsewhere in California; these are detailed in the Appendices.

The Groundwater Architecture Project has laid the foundation for the widespread adoption of the AEM method as an integral part of groundwater management throughout California. We came into this project with the example of the Danish adoption of AEM for groundwater mapping. We ended the project with an improved understanding of numerous aspects of the workflow – from AEM data to the hydrogeologic conceptual model. While the GAP was designed to be of specific benefit to the application of the AEM method for groundwater management in California, what we have learned through the GAP will advance the adoption of the AEM method worldwide.

Appendix 1. Final recommendations from the Stanford Groundwater Architecture Project for the optimal workflow

RECOMMENDATIONS: ENGAGEMENT WITH LOCAL AGENCIES

There can be various starting points in an AEM project. Two examples:

1) The agency has limited understanding of the area and is considering the use of the AEM method to conduct a reconnaissance survey to understand the large-scale structure of an area. This typically involves AEM flight lines spaced 500 m to 5 km apart, with selected areas for denser lines to be flown at a later date. The line separation is determined using existing knowledge of the subsurface hydrogeology. As the line spacing increases, so does the possibility that the interpretation will be more challenging and more uncertain.
2) The agency has an existing conceptual model and/or groundwater model but is considering the use of the AEM method to obtain additional information to build and refine improved models. This typically involves dense AEM flight lines spaced 100 m to 500 m apart. The line separation is determined using existing knowledge of the subsurface hydrogeology.

The first step in the workflow is to determine how/if AEM data can assist with groundwater management. As examples of the types of questions that can be addressed using AEM data, we have provided links on the results page of the GAP website https://mapwater.stanford.edu/results the questions defined for Butte and Glenn Counties, for Indian Wells Valley, and for San Luis Obispo. While an initial set of questions were posed by the agencies, what you see here are the questions that were identified after considerable back and forth between the agency and the geophysicists, sometimes involving modeling steps later in the workflow. This iterative process ensures that there is a clear understanding of the information that can be derived from the AEM data.

At the start of the iterative process, a feasibility study is conducted. This study is designed to determine whether the predicted resolution and depth of investigation of the AEM method, in the study area, is sufficient to provide useful information. The data typically used for this feasibility study include geologic cross-sections, lithology logs and electrical logs. It is important to note that a lithologic model, displaying sediment type/texture and rock type, is the product that can typically be directly derived from the AEM data alone. Integration with other data sources is needed to obtain a model of hydrostratigraphy.

Engagement with the local agency continues throughout the project, as incorporating local knowledge at all steps improves the quality of what is done and ensures that the final product supports the agency's groundwater management needs.

Agencies often want to explain the purpose of the AEM survey to residents of the area and to their technical advisory groups. The power point presentation, linked to the GAP website, can be used for that purpose.

RECOMMENDATIONS:
ACQUISITION/ORGANIZING/DIGITIZING EXISTING GEO-DATA

The defined objective is the integration of AEM data with other available data to build a hydrogeologic conceptual model (HCM). The existing data serve two purposes: 1) to support the acquisition and interpretation of the AEM data and 2) to provide additional information used in building the HCM.

Determining the amount of existing data needed relies on local knowledge of the hydrogeologic system but, in general, the more high-quality data the better. Accurate locations of the wells, the well data (lithology, screened intervals, water levels, specific capacity tests, water chemistry and borehole geophysical logs) and geological and hydrogeological interpretations are essential for the design and interpretation of an AEM survey. Given that well data will eventually be needed for the interpretation and building of the HCM, the procedure should be to get the well data organized and digitized so they can be used in the design of the survey.

When gathering existing data, it is prudent for the local water agency to quality screen and record all available data in the area, documenting why certain data are omitted on the basis of poor or unusable quality. This is likely to be a very wise long-term investment as it avoids having to revisit old well completion reports and other datasets and reports at some later time. The data will all be stored and maintained in a robust open database structure.

For budgetary and time constraints, selecting certain wells to develop a representative dataset can be a cost- and time-effective alternative. When doing so, it is important to set up a selection and quality screening process to ensure that the selected wells represent as large a portion of the area of interest (for AEM data acquisition and/or HCM development) as possible; in areas with a high degree of spatial heterogeneity there is likely a need for a higher density of wells. When working with the selected dataset, we recommend capturing and quality screening all available information from the wells.

In the following sections, we describe the processes that Stanford researchers have adopted and developed for reviewing and compiling the well data needed to acquire and interpret AEM data.

Lithology Logs (also referred to as Geologic Logs and Well Logs)

We first address the need for lithology logs, which provide essential information about sediment texture/type and rock type in the subsurface. The experience of Stanford researchers, in the acquisition of data in the Central Valley of California, in Butte and Glenn Counties and in the Kaweah subbasin, suggests that for AEM flight lines spaced up to 2 km apart, we should aim for two high-quality lithology logs per section. (A section is 1 mile x 1 mile or 1.6 km x 1.6 km.) We have found that this provides adequate lithologic data for 1) the initial modeling (to estimate the depth of imaging and resolution of the AEM data), 2) flight line planning, and 3) the rock physics transform (to transform the resistivity models to lithology). In developing the rock physics transform, co-located lithology and AEM data are required, so the flight lines need to be located within ~50 m of the high-quality lithology logs. If the line spacing is greater than 2 km, we should aim for two high-quality logs per section, in the sections containing AEM flight lines. The location of the lithology logs should be such that the AEM lines can be moved to go over a number of them, so as to have the desired co-located data.

Various procedures have been adopted or developed by researchers at Stanford University for identifying, locating, digitizing, and categorizing lithology logs. We first identify high quality lithology logs in the study area, defining high quality as a lithology log that is relatively deep,

given the depth of wells in the area, with relatively fine depth discretization in the lithology descriptions. The next step is to accurately locate the wells, ensuring that they can be located to within 20 m. We then digitize the logs and categorize the lithology descriptions. Table 4.1 summarizes the work required to go through the complete process to obtain two high quality lithology logs per section for 50 sections, i.e. 100 lithology logs; total time 12 to 38 hours, so .25 to .36 hours for each high-quality lithology log that is required. The developed procedures are described in the documents linked below.

Step	**Time per high-quality lithology log or section**	**Time for 100 high quality lithology logs from 50 sections**
	–	
Data dump and initial review to identify 100 high-quality lithology logs.	1 – 5 minutes per section (This might involve opening 10 to 50 WCRs per section to find high quality lithology logs)	50 to 250 minutes = ~ 1 to 6 hours
Locating high-quality lithology logs	5-10 minutes per lithology log (2 per section)	500 to 1000 minutes =~8 to 17 hours
Digitization of high-quality lithology logs	2-3 minutes per log for shallow wells 5-7 minutes per log for deep wells	200 to 700 minutes =~3 to 12 hours
Categorizing lithology	Instant (using machine learning) to 1 minute per lithology log	0 to 200 minutes =~0 to 3 hours
Total	–	**12 to 38 hours**

Table 4.1. Time required to identity, locate and digitize high quality lithology logs.

Procedure for Identifying High Quality Lithology Logs

Procedure for Digitizing Lithology Information

Procedure for Categorizing Lithology Information

Geophysical Logs

While there are many types of geophysical logs we focus on normal-resistivity logs. These are the most common type of log acquired in water wells, and provide information about the geophysical property of the subsurface measured with AEM - the electrical resistivity. The resistivity log measurements of resistivity can greatly improve the accuracy of the initial modeling performed prior to data acquisition and are also used in the interpretation of the AEM data.

Resistivity logs are rarely acquired in private water wells used for domestic purposes, are sometimes acquired in private agricultural irrigation wells, and are often acquired in private wells providing a municipal supply. They are now commonly acquired when monitoring wells are installed by a state or local government agency. Oil and gas wells generally have resistivity logs but the logs may not start until much deeper than the portion of the subsurface of interest in an AEM project. As such it can be challenging to obtain shallow resistivity logs for a given survey area. Thus, our recommendation for the number of resistivity logs to obtain at least ten resistivity logs that contain information over the expected depth interval of the AEM measurements and that are spatially distributed over the survey area. This number of resistivity logs will provide sufficient information to inform both the initial modelling and the approach taken to

transform resistivity measurements to lithology. Procedures have been developed by researchers at Stanford University for the purposes of obtaining high-quality resistivity logs.

Procedure for Obtaining High Quality Normal-Resistivity Logs

Digitizing Normal-Resistivity Logs

Water Table Elevation Measurements

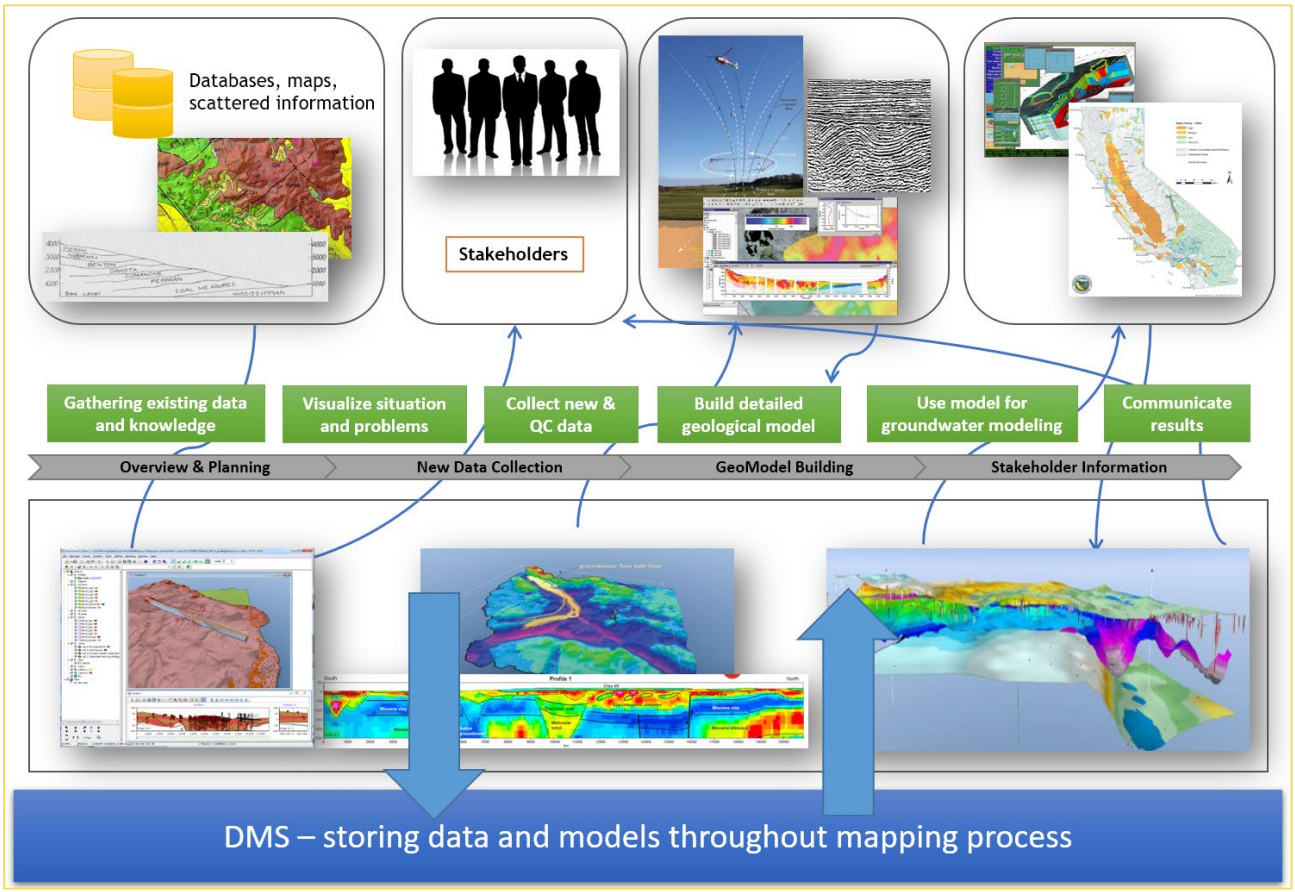
Saturation state can have a significant impact on electrical resistivity, so the water table elevation – so as to divide the subsurface into the unsaturated and saturated zones - should be accounted for in interpreting the AEM data to obtain information about sediment type or lithology. It is important to determine the locations where water table measurements will be available during AEM data acquisition and consider this in planning the timing of the survey and flight line locations. In some cases the measurements will be made as part of an ongoing monitoring program; in other cases local partners might be able to make measurements. We note that, to be accurate, it is the depth to the top of the saturated zone (TSZ) that is required to separate unsaturated from saturated; but given the resolution of the AEM measurement and the lack of measurements of the TSZ, we recommend using the water table elevation.

RECOMMENDATIONS: COMPILING OTHER DATA/INFORMATION

In addition to compiling the geo-data, there are other data and information important for the design of the AEM survey, and data acquisition and interpretation. These are described below:

1. Any TDS data that are available as variation in salinity, and its impact on electrical resistivity, needs to be accounted for in the interpretation of the AEM data.
2. Infrastructure information, both type and location, is important for flight line planning and post-data-acquisition processing and inversion. This information includes power lines both above ground and underground, metallic pipelines, railways, highways, confined animal feeding operations, dwellings, towns, canals, and any other known infrastructure that affects the AEM data collection.
3. Any areas with flight restrictions, e.g. military or environmental protection areas or other areas as designated by the Federal Aviation Authority.
4. Airports and logistical support locations including internet access, wireless services, fuel supply, hospitals and law enforcement availability.
5. Location of fire and rescue facilities and services.

RECOMMENDATIONS: DATA MANAGEMENT SYSTEM



To determine the characteristics of a given groundwater interest area and develop a 3D hydro-geological model, many different data and datatypes are collected and interpreted together. As collecting data and compiling these into models involves several disciplines and competences hardly present in a single person or even a single organization, it is mandatory that the data are collected and stored in such a way that they are readily available and continually updated. Given the cost of acquiring the data, and the people-months spent working with the data, it is a relatively small investment to develop the DMS.

The GAP project has utilized a data management system and databases capable of storing the datatypes involved in the pilot project areas, enabling the different stakeholders to access these data across various organizations and agencies. The figure above shows the flow of data and modeling. All of the data are readily accessible through a DMS.

An ideal process in building an HCM starts with gathering existing data and adding these to a preliminary 3D model. This allows for delineating discrepancies and uncertainties, thus identifying data gaps. Once these are identified and pointed out, new data should be collected and added to the model. In this process, all data should to be stored in a DMS. When new data are collected these data must flow into the DMS and from there be used to refine the model. This is in principle an iterative process where data gaps and uncertainties are defined and reduced with new data. The DMS will throughout this process store data revised models enabling a close cooperation between the different parties and persons involved. This has been the case in the GAP project.

The data and results of the modeling should also be available for public use, especially for authorities, GSA's and in general for people with interest in the topic. This requires an open formats. Most data have a spatial relation, which means that maps are the common way of displaying. Open maps formats as WMS or WFS which can be used by a variety of client software must be supported. Dedicated web map applications tailored to this purpose are very useful for public outreach. 3D applications able to visualize the often complex hydrogeologic settings are also useful in communicating results.

DMS on a Regional/State-Wide Scale

The DMS should be able to compile, store and distribute data collected in the mapping campaign and beyond for future work. Often data and models are compiled into models aimed for a specific smaller project area. Although this makes sense in a local ad hoc perspective, learnings are that these models are not maintained over the years and therefore of limited use on a longer term and often redone again, when new knowledge arises. The DMS should therefore aim to be able to compile onto one, or a least a few unified state-wide model(s). In order to do so, careful work must be done to set up a basic model and model parameters.

DMS on a Local/GSA level

On a local level, the DMS should be focused on delivering results, rather than the data itself. This means establishing specialized webpages with relevant thematic maps should be in focus, thus excluding the more technical background data, although these should also be accessible if needed. Relevant thematic maps could be isopach maps (aquifer- and aquitard thicknesses), water levels, infiltration areas etc. besides predefined (conceptual) cross sections. A live 3D model would also give the user a more comprehensive of the hydrogeological situation.

Data Uploaded for Use in Developing the HCM

Many kinds of data will contributing to the development of the HCM: well data with hydraulic properties and water chemistry, topographic maps, geologic maps etc, previous established cross sections and hand drawings, geophysics – airborne, and ground-based – ERT, DC, EM, TEM, seismics, geophysical logs, and existing 3D models that can be used as guide for new models. As these data will be used in the model, they'll have to be in a format that is possible to work with in a 3D modeling software. Suggested guidelines:

1) All data must be digital and have qualified 3D coordinates.
2) A common geographic reference system must be defined, and used, in order minimize problems related to mixed coordinate systems.
3) Data formats should be of well described standard – like comma separated files, common file formats (e.g. GeoSoft exchange format), or databases of different kinds.
4) Well data should be stored in an open database format that can be used for updating content querying and quality assessment including a user interface. This should preferably be web based and distributed via web services.
   Some geophysical data formats can be complex, and a standard should be defined. In the GAP project a common database format called GERDA is used.

We provide a table listing data that should be collected and included in the DMS in mapping the groundwater systems of an area. Not all these datatypes will be collected in a specific mapping campaign. The acquired data typically varies depending on availability, geological setting, time and budget etc.

Sharing Data

Sharing data between stakeholders is important in order to succeed in developing HCM's that can establish the foundation for resilient planning. The DMS needs to be able to handle different permissions for different stakeholders/users. Some users will have administrative privileges and be able to upload data, assign access to data and also be able to give permissions to other users and remove them again. The administration scheme should have both permission system on user level, but also on role level for easy administration.

	Data compilation
Well data	• Well data including position, casing, pumps, lithology, stratigraphy, grouting, screens, water table, photos, water chemistry etc. • Hydrological information; discharge, pumping rates, pump tests, drawdown, flow logs etc. Note: Considering defining guidelines and standards for an even level of soil sample description is considered very important for future comparability of these data.
Geophysical well logs	• Geophysical properties like resistivity, gamma, SP, Seismic velocity, neutron etc.
Geophysical data	• AEM data, predominantly TEM and Magnetics (raw data, processed data and inversion results) • Ground based TEM, tTEM, WalkTEM or other • Earth Resistivity Tomography (ERT) profiles • DC Electrical soundings (VES) and profiles • GeoRadar profiles. • EM data like DualEM and EM31 • Seismic data • NMT soundings • Earth Magnetic data • Gravity data
GIS data, maps	• Topographic maps. • Geologic maps, including soil type, stratigraphic boundaries, isopach maps, morphology, geologic structures, faults etc. • Hydrological maps, water sheds, groundwater level, water chemistry etc. • Digital elevation models. • Geomorphological maps • Soil maps • Models and related information. 2D cross sections, hydrostratigraphic models, existing reports, 3D geologic models. •
3D models and model contents	• 3D Geological models and other model results

RECOMMENDATIONS: IDENTIFYING DATA GAPS AND AEM SURVEY DESIGN

The design of the AEM survey includes selecting the AEM system and the data acquisition parameters, and determining the location, orientation, and density of the flight lines. The survey should be designed so as to best address the questions posed by the agency and to add spatial coverage of critical regions of the subsurface. One of the first steps in survey design is to select the general flight areas so as to avoid populated areas, powerlines, and other infrastructure likely to interfere with the data acquisition. Knowledge of the local geology, obtained by a review of existing data and reports, assists with the general survey design. More detailed information about the variation in lithology and electrical properties of the subsurface, also obtained from existing borehole, geophysical, and geological data, is needed in order to simulate, through forward modeling and inversion, the acquisition and interpretation of AEM data in the area. This makes it possible to predict the quality, spatial resolution, and depth of reliable imaging that can be obtained. Such an assessment is important both for survey design and for

setting realistic expectations about what questions can be addressed, in the study area, through the use of the AEM method.

The recommended approach is to use existing lithology logs and geophysical data to generate a lithology model of the subsurface, with corresponding resistivity values, and then simulate the AEM systems to recover the subsurface model. The various AEM systems available are considered for the simulations so that the appropriate system is selected given the survey objectives. Current systems (of which we are aware) that are likely to be available in California are the CGG HeliTEM, MultiPulse, Resolve and Tempest; Geotech VTEM ET and VTEM Max; SkyTEM 304 and 312. This analysis of the various systems is so crucial for a successful survey that any potential vendors need to provide reliable and comprehensive information about their systems. It is key to these simulations that realistic data noise levels, flight speeds, and flight altitudes are taken into consideration. Likewise, the full system transfer function needs to be modelled, and if bias in the data is expected, this must be considered as an additional source of data uncertainty. The analysis is carried out by simulation of how well the given model is recovered given all these parameters and how well the resistivities and thicknesses are determined by the model sensitivity analysis. Through repeating the numerical exercise, it is possible to estimate the spatial resolution of layers, the vertical resolution of shallow layers, and the maximum depth of reliable imaging (depth of investigation) that would be obtained with the different systems in one or more parts of the survey area, representative of different lithologic variation. Through this process the various elements of the survey design are assessed and determined.

A high level of data accuracy is crucial for the quality of the resulting images derived from the AEM data. Any system needs to be able to deliver data that are not systematically biased and it needs to be able to output not only a mean value for a given data point, but also the statistical uncertainty for that point. A pre-qualification of the potential systems should be designed so they all fly a test line of 1 to 10 km length for assessment of the ability of the system to deliver accurate data, the system noise level, the ability to resolve the near-surface layer, and the depth of investigation. Determination of the bias level is done by comparison to calibrated ground-based measurements made along the test line.

In the pilot study, we explored the acquisition of AEM data directly over the wells from which the high-quality lithology data were obtained. The co-location of lithology and resistivity significantly improves the accuracy of the transform linking resistivity to lithology. As a result, we recommend acquiring data as close as possible to selected wells, while staying the required separation distance from the powerlines supplying the wells with power. The required separation distance is determined by the electrical properties of the subsurface, both the background resistivity and that in the top ~6 m, and the grounding of the powerline. While it is typically assumed that one should stay at least 150 to 200 m from a powerline, the highly resistive nature of the top ~6 m in parts of California has been found to result in no powerline effect in the AEM data even when flight lines are within 100 m of a powerline.

Upon completion of the survey design, contracting for a geophysical vendor is then pursued based on cost, proposed system characteristics, and demonstrated ability to successfully complete the survey. Often this is done by requesting an RFP or quotes for cost and timing to complete the work and deliver data. In addition to the elements of the survey design discussed above, the final contract needs to include specified flight height limits, specified flight speed limits, and specified tilt angles (X, Y) limits.

RECOMMENDATIONS: ACQUISITION OF AEM DATA

Procedures for Notification Prior to Data Acquisition

An agency is advised to send advanced notice of the survey to residents through their usual channels of communication (e.g. email distribution, notices in the local paper, newsletter, radio, etc.) We suggest that notification that goes out prior to the earliest possible flight time but the wording should accommodate changes in schedule that can occur. It is very difficult to accurately predict the start date. We have produced an example of an announcement, along with a figure that compares the magnetic field associated with the data acquisition to that of household appliances. We also advise notifying the emergency responders in the area (e.g. police, fire department) as they might receive calls from concerned individuals.

Contractual Agreements and Pricing Structure

The contract for acquisition of geophysical data needs to include clear explanations of what will be done and the pricing structure. There needs to be an agreement as to which party will absorb the risk and costs associated with downtime due to weather, equipment malfunction, or for any other reason. There must be agreement early in the project planning on the payment schedule. It is typical for the AEM service company to request an advance payment to cover, for example, the booking of the helicopter, while most public agencies can only pay for services delivered.

The agreement with SkyTEM that was put in place for this project:
SkyTEM agrees to forego the 40% advance payment and instead issue an invoice for 90% of the total project on the last day of flying. Payment is due 14 days from receipt of invoice. The final invoice of 10% will be issued upon delivery of the final report from SkyTEM. Given this payment schedule, it is crucial that all data be handed over to the group responsible for QA/QC oversight on a daily basis while the field operation takes place. SkyTEM will perform quality control checks as well. Those responsible for oversight will do a final approval of the data no later than the day before the first invoice is due (13 days after the last day of flying). This was an acceptable agreement as a majority of the services provided by SkyTEM were completed before the first invoice was due.

Water Table Elevation Measurements

Water table elevation measurements should be acquired as close as possible to the time of AEM data acquisition so as to be able to differentiate between unsaturated and saturated materials in the interpretation of the AEM data. Saturation state has a significant impact on electrical resistivity that, if not accounted for, could confuse the interpretation of the data in terms of sediment type or lithology.

Guidelines and Standards for AEM Measurements, Processing, and Initial Inversion

When the geophysical mapping of Denmark with the AEM method began, a collaboration was established between the HydroGeophysics Group at Aarhus University, Denmark and the Danish EPA. This group was responsible for providing oversight and training, and for conducting the applied research required to continually improve the application of the AEM method. A report was produced that documents the guidelines and standards for AEM measurements, processing and inversion. As described in this report and in this publication, a national test site was developed in Denmark that was used for calibration and validation purposes. The Danish

EPA used the guidelines from the report as part of the tender process, requiring that consultants and AEM contractors follow the guidelines in order to be successful in the bidding process. DWR should establish a test-site; this would require a neutral party to maintain the test-site and develop the guidelines, as was done by the HydroGeophysics Group and the EPA in Denmark. Ideally other test sites would be established throughout California that would provide easier access. With only one DWR test-site, a ground-based system could be used and validated at the test-site and then used at a local site in order to calibrate and validate the AEM system.

Delivery/Archiving of AEM Data

An AEM survey should be considered to be a long-term investment, that can be utilized and revisited in the future. Use of a well-designed data management system (DMS) will ensure that all relevant data are provided by the survey company that acquired the AEM data and those responsible for the quality control, processing and inversion of the data. A powerful DMS will furthermore ensure that data can be readily shared and retrieved.

Processing and inversion schemes are continually being improved. It is therefore crucial that raw data, processed data and meta-data are documented and stored in a database.

We have chosen to divide the AEM data into three categories:

1. Raw data as recorded in the field
2. Processed data
3. Inverted data

Raw data

It is recommended that the raw unprocessed data, as recorded by the acquisition instruments, be requested from the survey company. If data have been filtered and averaged, information about this process is essential to be archived. If the raw data are delivered in a proprietary binary format we recommend that the data also be delivered in a plain text format (e.g. comma-separated values (CSV), Geosoft xyz) with the meta-data explaining the data types in the file. All information necessary to describe the entire setup of the AEM system shall be provided. This includes filter definitions from the sensor to the digital voltage data, hardware configuration of the AEM system and all settings used to define the system setup (e.g. waveform, time gates, etc.).

Processed data

Processed data where auxiliary data have been merged with the primary electromagnetic data are to be provided in a simple plain text format (e.g. comma-separated values (CSV), Geosoft xyz). The AEM sounding spacing can vary, but typically a 1 Hz or up to 10 Hz spacing is used. All information about the AEM system, data averaging and filtering must be documented.

Inverted data

Inverted data describing the resistivity model and the uncertainties related to the parameters are to be provided in a plain text format (e.g. comma-separated values (CSV), Geosoft xyz). Information about algorithms, software and parameters used for the inversion must be documented. We further recommend that the inversion results be provided in formats easily readable by third party presentation software packages.

RECOMMENDATIONS: DATA PROCESSING AND INVERSION

The process of filtering, editing and merging auxiliary data with the primary electromagnetic data is typically called data processing. When the data are processed, the next step is inversion, which is the process where a subsurface resistivity model is estimated which fits the measured AEM data (i.e. voltage) for a given prior model (e.g. smooth subsurface structure).

Data Processing

The data processing recommendations are based upon hydrogeological applications of the AEM method. The processing is as important as acquiring precise data. Poor data processing can reduce the resolution significantly and poorly described information affecting the signal can create biased and misleading model results. The processing entails using appropriate data averaging schemes, removing distorted data, and having accurate sensor altitude, GPS and pitch/roll data.

Major noise sources for the AEM data are installations including powerlines, roads, cables, railways, windmills, houses, antennas etc. The noise sources could also be atmospheric noise caused by thunderstorms or sun-storms. Although the survey should be designed to avoid the impact of these noise sources, in practice, it is impossible to avoid all of them, and hence a portion of the measured data will be contaminated by noise. Typically, all data within a distance of 100–200 m from the known noise sources should be culled. Even for an experienced geophysicist identifying all noisy portions of the data and completely removing them is a challenging task. Therefore, rather than considering the processing and inversion as independent steps, they should be considered an iterative process, where one first carries out the processing of the data, then inverts the data to determine if there are any suspicious resistivity structures imaged and to check the misfit between the observed and predicted data. If there are suspicious structures imaged or portions of data showing high misfits, then the processing and inversion procedure should be repeated until a satisfactory quality of the resistivity model and data misfit is obtained. Below is an example processing workflow from Aarhus University (Auken et al., 2009).

Recommended data processing workflow, in overview:

- Review of the raw data delivered from the airborne survey company including calibration and reference measurements.
- Processing GPS, tilt/roll and altitude data.
- Processing of voltage data
    - Removal of noisy and coupled data and averaging of raw data.
    - Visual assessment and editing along flight lines.
- Preliminary inversion to support processing
    - Possibly redo parts of processing based on preliminary inversion

Inversion

Geophysical inversion is the process by which we estimate the resistivity model of the subsurface fitting the data. An inversion algorithm uses an iterative optimization technique and allows inclusion of prior information that can be obtained from well data (e.g. lithology, water level, salinity, etc.) To produce unbiased results the inversion algorithm would need full descriptions of the AEM system (e.g. geometry, altitude and attitude of the transmitter/ receiver, transmitter waveforms, receiver bandwidth characteristics etc.). There are a number of tuning parameters in any inversion algorithm that require geophysical knowledge of AEM data as well as inversion methodology. Therefore, it is critical to have an experienced geophysicist involved in this process.

Similar to the philosophy of viewing the processing step as an iterative process, the inversion and subsequent hydrogeological interpretation of the recovered resistivity model should not be viewed as a turn-key operation. We recommend interaction among geophysicists, hydrogeologists, water managers, and local stakeholders such that the geophysicist can run multiple AEM

inversions based upon these interactions. In addition, the computer software handling the inversion needs to deliver effective displays giving the geophysicist a tool to readily adjust various parameters in the inversion and evaluate the result of the inversion.

For groundwater mapping, a commonly used inversion approach is the 1D spatially- (or laterally) constrained inversion technique, developed by Aarhus University and commercially available through the Aarhus Workbench. This method assumes a 1D layered earth structure when simulating the AEM data, which is a robust way to invert AEM data proven in various groundwater mapping projects as well as in our GAP project. We recommend using the 1D spatially-constrained inversion technique as the main tool for inverting AEM data. In addition to the Aarhus Workbench, there are other open-source tools providing a similar capability of the spatially constrained inversion. For instance, in our GAP project, the simpegEM1D code was used to invert the AEM data sets acquired at all three pilot study areas in conjunction with the Aarhus Workbench. Other freely available 1D AEM inversion codes are: geobipy (from USGS) and GA-AEM (from Geoscience Australia). In general, a higher level of expertise is needed to run these codes compared to the commercial software.

If surveying in areas with strong 3D effects (e.g. vertical bedrock-sedimentary boundaries, rough terrain, or vertical freshwater-saltwater transitions) the recovered resistivity image may show 3D artifacts in the recovered resistivity model. These features can be identified and flagged by a skilled geophysicist to avoid misinterpretation. In these scenarios, it may be attractive to consider higher dimensional (2D and 3D) inversion codes. These simulate the physics in higher dimensions, which can capture more complex geologies, but they require much greater computational power and more expert knowledge of inversion to obtain meaningful results. Hence, we do not recommend that 2D and 3D inversions be the primary tool used to invert AEM data, but rather considered as an option when 3D effects are present in particularly important areas. SimPEG's EM module provides a capability to simulate and invert AEM data in 3D.

RECOMMENDATIONS: GEOPHYSICS TO LITHOLOGY TRANSFORM

The resistivity model, derived through processing and inversion of the AEM data, must be transformed in order to obtain information about subsurface properties to which a measurement of electrical resistivity can be sensitive (e.g. sediment type, lithology, hydrostratigraphy, principal zones in the aquifer system, water quality). In the GAP, our focus was on obtaining the information required to inform the development of the hydrogeologic conceptual model, and to map TDS (total dissolved solids) to understand variation in water quality.

The first step in conducting the transform should be the use of geophysical logs, lithology logs, existing cross-sections, and other relevant measurements to determine what can/cannot be related to a variation in resistivity. In general, there is more likely to be a well-defined relationship between resistivity and lithology or sediment type than between resistivity and individual hydrostratigraphic units. Each hydrostatigraphic unit can be composed of a mix of lithology or sediment types, so can correspond to a wide range of resistivity values; this results in significant overlap in the resistivity signature of the various units. In areas with variable water quality, significant variation in TDS can be the dominant parameter influencing the observed variation in resistivity.

We first describe our approach for the case where the impact of TDS on resistivity can be considered negligible. Starting with the best resistivity model recovered from the AEM data, the best available well data, and trusted cross-sections, there are, in general, two approaches. One, the more qualitative approach, is to view all the data, typically using visualization software, to develop a qualitative understanding of which materials or packages of materials correspond to which ranges in resistivity values. The objective with this approach is to identify zones tens of meters in thickness that represent the large-scale architecture of the aquifer system. This zonation can be done to capture the variation in sediment type, lithology, hydrostratigraphy, or other divisions that are useful in describing the aquifer system.

The second approach is to develop a more quantitative relationship between resistivity and lithology or sediment type, described, for example, in terms of gravel, sand, silt and clay, sometime classified as coarse-grained (gravel and sand) and fine-grained (silt and clay). Examples of this approach are the clay-fraction method developed by Christiansen et al. (2014) and the bootstrapping method developed by Knight et al. (2018). (Christiansen, A.V., Foged, N. and Auken, E., 2014. A concept for calculating accumulated clay thickness from borehole lithological logs and resistivity models for nitrate vulnerability assessment. Journal of Applied Geophysics, 108, pp.69-77. Knight, R., Smith, R., Asch, T., Abraham, J., Cannia, J., Viezzoli, A., Fogg, G., Mapping Aquifer Systems with Airborne Electromagnetics in the Central Valley of California, Groundwater, 56:6, 893-908, https://doi.org/10.1111/gwat.12656, 2018.)

Using the derived relationship, the resistivity model can be transformed into a subsurface model mapping the variation in sediment type or lithology. It is important to note that this is not differentiating the hydrostratigraphic units unless there exists a resistivity contrast between the units. The large-scale architecture can be defined using the observed variation in the resulting subsurface models, the structure seen in the AEM data and/or interpolated well data, and prior knowledge of the aquifer system.

If TDS varies throughout the survey area, our recommendation is to use all measurements of TDS made on water samples, along with the large-scale structure seen in the AEM data, to define regions in which TDS varies over a defined range. Each of the regions can then be treated independently, using one or both of the approaches described above to extract information about the variation in sediment type, lithology, hydrostratigraphy, or other divisions that are useful in describing the aquifer system.

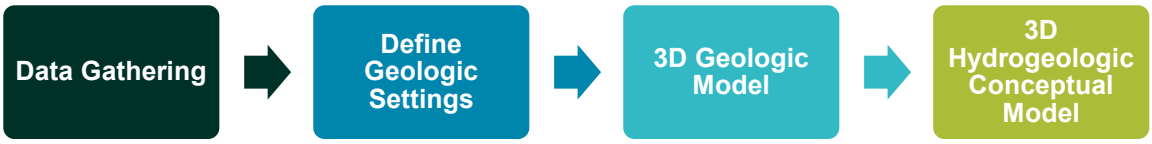
The first approach is very flexible, so one can potentially take into account various factors but it is not reproducible in a quantitative sense. The second approach, currently designed to derive the lithology or sediment type, is less flexible, but reproducible and quantitative. We recommend using both approaches. As an example, one could identify the large-scale hydrostratigraphic structure using the first approach, and then use the second approach to refine the definition of the large-scale structure and map the variation in lithology or sediment type within the large-scale structure.

RECOMMENDATIONS:
MODEL DEVELOPMENT THROUGH DATA INTEGRATION

Development of the HCM

A Hydrogeologic Conceptual Model (HCM) is developed with the purpose of providing an understanding of the geometry and physical characteristics of groundwater systems. It can also provide the starting point for the development of a numerical groundwater model.

Geophysical surveys, including electromagnetic and geoelectric, when combined with other downhole lithological and geophysical data, can provide very useful information on interpreted changes in the lithology. These data can be used in the mapping of hydrostratigraphic units. Surveys such as airborne electromagnetic (AEM) surveys, can provide 3D coverage which can significantly improve the quality of the model. We recommend, whenever possible and particularly in areas where data coverage is thin, that AEM or similar surveys be used in the development of the HCM. The following summarizes the recommended workflow for the development of the HCM.



Recommended Workflow in the development of a Hydrogeologic Conceptual Model (HCM)

Data Gathering

We recommend collecting all available data that can be used in the development of the HCM. This includes

1. Existing hydrogeologic interpretations from previous studies, including descriptions of the interpreted depositional environment and other geological events;
2. Geologic maps;
3. High resolution digital elevation models;
4. Information from boreholes including lithology, screen intervals, water levels, water quality and geophysical logs;
5. Geophysical investigations including seismic, gravity, magnetic, electrical, and electromagnetic surveys.

Data should be collected, processed, and stored in a format so they can be uploaded to the modelling software.

Defining the geologic settings

The next step in the development of the HCM is to perform a review of previous studies on the geology and hydrogeology. The object of the review is to obtain an initial understanding of the area's geology, development and the considerations taken in previous studies. The review will provide the conceptual geologic framework on which the development of the 3D models will be based. Thereafter it is recommended that this *a priori* knowledge is compared with the collected AEM data. This is important to see if the conceptual geologic framework should be modified, or in the cases where previous studies were not in agreement, to determine which interpretation fits the data best. The result of the review will be a detailed description of the landscape, the geological units and important structural elements (e.g. faults), accompanied with illustrations (often as cross-sections) showing the major architecture of the geological units and geological structures.

Development of the 3D geological model

The third step is the construction of a 3D geological model. In a 3D geological model, lithostratigraphic units are correlated to known geological formations, which are then modelled. The result is a 3D rendering of the thickness and distribution of the individual geological units (i.e. geological formations) and structures within the focus area. In a 3D geological model, the accuracy and quality of the interpretations are dependent upon the amount of chronostratigraphic information available. Due to the areal coverage typically obtained by an AEM survey as well as the penetration depth of up to 1000 feet the AEM data plays an important role in the development of the 3D geological model.

Development of the HCM

The final step is the development of the HCM. The 3D geological model is based on the geological history, with geological formation as the basic unit. As a result, the geological units may contain a wide range of lithologies characterized by different hydraulic properties. In an HCM, the aim is to subdivide the geological strata into hydrostratigraphic units (i.e. aquifers and aquitards) using the interpretations from the 3D geological model. It is, thus, the hydrogeologic units that define the groundwater system, and the HCM, therefore, provides the foundation of the groundwater flow model. In addition, the HCM provides useful information on the location of groundwater recharge areas, the geographical distribution of water storage and areas where groundwater resources may be most vulnerable to pollution.

The Use of the Multi-Point Statistics (MPS) Method

Multiple Point Statistics (MPS) is a numerical methodology that can integrate AEM data, well data, other datatypes and prior knowledge of geologic patterns (e.g. fluvial environment) in a way that can generate multiple realizations of lithology/sediment type models (i.e. categorial models) of the subsurface capturing the uncertainty.

The MPS method results in an ensemble of models all of which are compatible with the data input and the geological understanding of the depositional environment in the model area. This is done by applying a training image (TI) which is a 3D model incorporating the conceptual geological features of the model area. Each point in the model, is thereby assigned a probability of a certain lithology in the model area, usually in two fractions, e.g. sand and clay which in terms of the hydrology are the high permeability permeable and low permeability fractions.

Using the ensemble of equally possible models leads to the ability to calculate statistics and thereby quantify uncertainty of the resulting models considering the different data used in the model.

By performing the MPS in the modeling process extra steps are required. The gains of this is getting the probabilistic models that can resolve more details in the data, combined with a geostatistical based assessment of the probability and uncertainty of a given model. The possibility to provide objective quantified model uncertainties is important when models and their findings are challenged by other opinions.

The MPS model can provide insights in lateral and horizontal hydraulic contacts within the model. This is important if a low fraction of sand lenses in layers dominated by clay can provide enough hydraulic continuity so as to act as an aquifer.

Looking forward, MPS, if used as a tool within actual mapping campaigns, can pinpoint areas with large uncertainty and thereby focus data collection to these areas. Research and development related to generating flow models are increasingly concerned with estimating uncertainty in model predictions. Here MPS will be able to provide equally probable models as input for a series of flow models, thus expanding statistical assessment into these models

Although the extra work involved in creating MPS models also involves extra costs, this cost is only a fraction of the cost of collecting the data.

Workflow recommendations:

- A good understanding of the geology is vital to provide the framework for a MPS model. Normally a fully modelled HCM will be the base for the MPS model and provide the outline for the MPS model, i.e. distinct changes in rock/sediment characteristics.
- The process for development of the training image/s is critically important and should involve input from hydrogeologists familiar with the study area. Only in rare situations can one training image cover be used for modeling the entire area. More typical is the division of the area, and thus the model, into subareas which must be handled separately with individual TI's (and possibly also resistivity transforms). This should be done while maintaining continuity, where appropriate, across subarea borders.
- Careful quality check of the input data is important. Low quality in the well data resulting in e.g. uncertainty in the exact location of the well and other attributes of low quality can introduce uncertainties into the model that can be hard to resolve and can lead to misinterpretations. Often low quality well information will have problematic direct influence on soft- and/or hard data as well as the rock physics transform as contradicting information might arise.
- MPS is a computationally intensive method. A larger number of realizations must be calculated to support proper statistics and uncertainty calculations. Therefore, careful planning of the model size can optimize calculation time. Starting with a low-resolution model or a smaller subarea for testing hypotheses before going into high resolution modeling and the full size of the area can optimize time spent.
- When tuning simulation parameters for best results, it is recommended to try optimizing them carefully. One strategy is to run initial realizations without data, using only using the TI. In this way, the number of variables can be kept limited, realizations can be examined, and variables tuned accordingly. After this process data can be introduced in the calculations, thus resulting in more complex results.

RECOMMENDATIONS: QUANTIFYING UNCERTAINTY

We focus on the uncertainty in the hydrogeologic conceptual model (HCM) derived from AEM data and well data. Five major sources of the uncertainty are:

1) Noise in the AEM data (e.g. powerline noise)
2) The AEM inversion
3) Well data (e.g. location, lithologic description)
4) Rock physics transform (e.g. between resistivity and lithology or sediment type)
5) Data density of the AEM and well data
6) Geologic and hydrogeologic knowledge (e.g. depositional history)

The first two items are related to the AEM method. In the inversion step, we fit the observed AEM data for a given (or estimated) noise level. There can be multiple resistivity models that fit the observed AEM data resulting in uncertainty in the recovered resistivity model. The higher the noise level, which can be present post-processing, the greater the uncertainty. Noise in AEM data can be described as external noise (sferics/thunder, metal objects like pipelines, powerlines etc.) and internal noise in the instrumentation. We recommend that the uncertainty related to a specific AEM system will be quantified and documented by high altitude flight tests and by flying a test line where the data are compared to ground truth information.

Transforming the resistivity model to obtain information to inform the development of the HCM requires the use of well data. The distance between the wells and the AEM data introduce uncertainty, as does the quality of the lithology log. There is also inherent uncertainty in transforming resistivity to obtain other information about the subsurface. For example, one sediment type can correspond to a range of resistivity values, with overlap in the resistivity ranges for different sediment types.

The HCM model is derived from the AEM and well data, and therefore, the density of both data is a crucial source of the uncertainty. We often suffer from the lack AEM data in the region where the population density is high. In contrast, often there are more available well data in that region providing additional data to compensate.

The sixth item is geologic (or hydrogeologic) knowledge of the study area as this provides key information about the expected geometry and correlation structure of units.

Uncertainty quantification is an active research topic. The approach taken will depend on the scale of the survey and the specific defined questions being addressed through the acquisition of the AEM data and development of the HCM. Our recommendation is that the sources of uncertainty be quantified when possible and acknowledged in the presentation of the developed HCM. One approach that has been explored in the GAP, is the presentation of multiple models all of which are compatible with the data. The multi-point statistics (MPS) method is one computational framework that is well-suited for communicating uncertainty.

THE STANFORD GROUNDWATER ARCHITECTURE PROJECT - Utilizing Advanced Geophysical and Computational Methods for the Development of Hydrogeologic Conceptual Models

California has embarked on a historic journey to achieve groundwater sustainability with the passage of the Sustainable Groundwater Management Act (SGMA) by the California Legislature in 2014. A specific legislative requirement is the development of a hydrogeologic conceptual model; such a model provides the information about the subsurface architecture needed for generating a groundwater model and for quantifying the water balance. There is widespread recognition of the need to acquire more subsurface data so as to reduce the uncertainty in the conceptual models. But the currently deployed, traditional methods of characterizing aquifers through the drilling of wells with testing and logging are slow, expensive and insufficient in terms of data coverage.

The MUDP-funded project (GAP) was part of a larger project, made up of ten work packages that allowed us to develop the optimal workflow for the use, in California, of AEM data to develop a hydrogeologic conceptual model. In the MUDP project, we included four work packages, that represent the research and development components of the ten in the larger project. The main body of this report contains descriptions of these four work packages. The deliverable from the larger project, the set of recommendations describing the optimal workflow, is presented in the Appendix.

The GAP project has created the foundation for a large-scale procurement from the State of California Department of Water Resources (DWR) for mapping of the Californian geology in all high and medium priority groundwater basins. The first large scale project was tendered out in autumn 2020 requesting use of advanced AEM technology similar to SkyTEM.

Miljøstyrelsen
Tolderlundsvej 5
5000 Odense C

www.mst.dk