Stochastic multiple data integration for the characterization of quaternary aquifers

Neven Alexis

https://doi.org/10.35662/unine-thesis-3030

University of Neuchâtel, Switzerland

Faculty of Sciences

Centre for Hydrogeology and Geothermics (CHYN)

Stochastic multiple data integration for the characterization of quaternary aquifers

Thesis presented for the degree of

Docteur ès sciences

by

Alexis Neven

Thesis advisor:	Prof. Philippe Renard, University of Neuchâtel, CH
Examiners:	Prof. Philip Brunner, University of Neuchâtel, CH Prof. Florian Wellmann, RWTH Aachen, DE Prof. Anders V. Christiansen, Aarhus University, DK Dr Vincent Badoux, Geotest Bern, CH

Defended on the 16th of May 2023

unine
Université de Neuchâtel
Faculté des sciences

Rue Emile-Argand 11
CH-2000 Neuchâtel
doctorat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Alexis Néven

Titre :

"Stochastic multiple data integration for the characterization of quaternary aquifers"

sur le rapport des membres du jury composé comme suit:

- Prof. Philippe Renard, directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Philip Brunner, Université de Neuchâtel, Suisse
- Prof. Florian Wellmann, RWTH, Aachen, Allemagne
- Prof. Anders Vest Christiansen, Aarhus University, Danemark
- Dr Vincent Badoux, Geotest, Berne, Suisse

Neuchâtel, le 6 juin 2023

Le Doyen, Prof. R. Bshary

Research is formalized curiosity.
It is poking and prying with a purpose.

Zora Neale Hurston, 1942

Acknowledgements

I would like to warmly thank all the people who participated in the realization of this thesis. For me, this thesis started on a glacier, the Tsanfleuron glacier, far away from the Bernese plain of the Aare. There I met Philippe Renard, who later became my thesis director. In front of his enthusiasm and his motivation, I made the choice to undertake these four years of thesis, abandoning the glaciers for the quaternary sediments. Today, I would like to warmly thank him for these 4 years. For his commitment, for his generosity, and for his unfailing support. We must recognize his ability to reformulate in an instant a confused idea into a crystal clear explanation. Thanks to him, I had the opportunity to work on many different subjects, some of which are included in this thesis. At the same time, I want to warmly thank the members of my thesis jury: Prof. Philip Brunner from the University of Neuchâtel, Prof. Florian Wellmann from the University of Aachen, Prof. Anders Vest Christiansen from the University of Aarhus, and Dr. Vincent Badoux from GeoTest A.G. in Zollikofen.

But a thesis is a journey that one does not take alone. So I would like to thank first of all my thesis companions who accompanied me during these 4 years: Valentin Dall'Alba, Arnaud Rüegg, Francois Miville and later Adeline Cojean. Through COVID-19, my days in the deserted university would not have been the same without them. Special thanks to Valentin for all these exchanges about my thesis, his, Python, life, and for all these projects and conferences done together. In these last 2 years, I had the opportunity to be joined by Ludovic Schorpp on the PheniX project and to share his office. I want to thank him for all the exchanges we had on a wide variety of problems that really helped me to get out of dead ends and to apologize for my musical tastes. My thesis would not be what it is without him and without his work. I would also like to thank Julien Straubhaar, in whose office I could come at any time of the

day to understand a GeoNe bug, to have a mathematical explanation, or to discuss PyVista philosophy. Finally, I thank Lucile, who helped me to start this thesis, for the permitting and the acquisition of the first year. Thanks also to Ilias and Cyprien for their help during the tTEM acquisition!

I also want to name the rest of the CHYN team, who made coming to the University a pleasure. In stochastic order: Przemek, Morgan, Hugo, Giona, Reza, Eleonore, Léa. Whether it was a coffee break, a conversation on the fly in a hallway or a collaboration on a project, you all helped make these moments great.

I also want to thank the great team at Aarhus University. Firstly, Anders Vest for welcoming me to his research group and for his enthusiasm for my research. Secondly, I would like to thank especially Pradip Kumar Maurya, Jesper Bjergsted Pedersen, Nikolaj Foged, for their help during the data processing and the development of the ACT algorithm. Finally, I would like to thank Rune Kraghede for the great month we spent together on the field and on the slopes !

Enfin j'aimerais remercier ma famille et mes amis proches, qui ont été là durant ces 4 dernières années. Ils ont toujours attentivement écouté ma réponse à "mais en fait tu fais quoi dans ton doc ?", avec plus ou moins d'intêret, mais toujours de la bienveillance. Merci pour le support infaillible jour après jour !

This research was supported by the Swiss National Science Foundation through the project Phenix (grant no. 182600).

Abstract

Groundwater resource management often requires the development of geological and hydrogeological models. However, constructing accurate models can be a challenging and time-consuming task, especially in large areas with complex Quaternary deposits. However, these areas are often the most frequently subject to resource exploitation and pollution. To address this issue, several studies have proposed innovative methodologies to integrate various types of data, including wells, geophysical, and hydrogeological data. The objective is to facilitate the construction of these models within coherent and reproducible frameworks with robust error estimation. In these, we present four studies that present novel methodologies to address this challenge. The first study presents a large and dense Time Domain ElectroMagnetic (TDEM) dataset acquired in the upper Aare Valley, Switzerland, to improve knowledge of the spatial variations of Quaternary deposits. The inverted resistivity models derived from this acquisition were published and could be used for various future studies. It also highlights the data set's potential for data integration algorithm development because of the abundance of various freely available data on the same zone.

The second study proposes a new methodology to combine boreholes and geophysical data with a propagation of the uncertainty to predict the probability of clay at the scale of a valley. A spatially varying translator function was used to estimate the clay fraction from resistivity. The parameters of this function are inverted using the description of the boreholes as control points. They combine this clay fraction estimation with a nondeterministic 3D stochastic interpolation framework based on a Multiple Points Statistics algorithm and Gaussian Random Function to obtain a 3D realistic high spatial resolution model of clay fraction for the upper Aare valley. The study demonstrates the quality of the predicted values and their corresponding uncertainties using cross-validation.

The third study addresses the possibility of integrating boreholes, geophysical, and hydrogeological data, while keeping the geological concept of the models coherent. We used a stochastic geological model generator to construct a set of prior models based on the boreholes. We then propose a multiscale inversion approach that combines low-fidelity and less accurate models with high-fidelity and more accurate models to reduce the time needed for the inversion to converge. Both geophysical and hydrogeological data are integrated, using an Ensemble Smoother with Multiple Data Assimilation Algorithm (ES-MDA) algorithm. The workflow ensures that the models are geologically consistent and robustly estimate the associated uncertainty with the final model. The study demonstrates the effectiveness of this approach for a controlled synthetic case. It shows that ArchPY and ES-MDA are capable of generating plausible subsurface realizations for Quaternary Sedimentological Models.

Finally, the fourth study presents an innovative methodology that combines the ES-MDA algorithm with an open-source hierarchical geological modeling code to integrate multiple data sources and construct geologically consistent models with robust error estimation. The methodology is applied to a field case in the upper Aare Valley, Switzerland. In order to benchmark the methodology, a cross-validation framework is implemented. The approach results in final models that effectively balance accuracy and uncertainty and can take into account various data sources, including geophysical data, regional conceptual knowledge, boreholes, and hydrogeological measurements at a valley scale. In summary, this thesis presents several innovative methods that could be applied on small to large scale hydrogeological model realization.

Keywords

geostatistics; uncertainty; Groundwater resource management; hydrogeological models; Quaternary deposits; Data integration; Time Domain EM; Stochastic inversion; Ensemble Smoother with Multiple Data Assimilation Algorithm (ES-MDA); Geologically consistent inversion.

Résumé

La gestion des ressources en eaux souterraines nécessite souvent le développement de modèles géologiques et hydrogéologiques. Cependant, la construction de modèles précis peut s'avérer une tâche difficile et longue, en particulier dans les vastes zones présentant des dépôts quaternaires complexes. Or, ces zones sont souvent celles qui sont le plus fréquemment soumises à l'exploitation des ressources et à la pollution. Pour résoudre ce problème, plusieurs études ont proposé des méthodologies innovantes pour intégrer différents types de données, notamment des données sur les puits, des données géophysiques et des données hydrogéologiques. L'objectif est de faciliter la construction de ces modèles dans des cadres cohérents et reproductibles avec une estimation robuste des erreurs. Nous présentons ici quatre études qui proposent de nouvelles méthodologies pour relever ce défi. La première étude présente un vaste et dense ensemble de données électromagnétiques dans le domaine temporel (TDEM) acquises dans la haute vallée de l'Aar, en Suisse, afin d'améliorer la connaissance des variations spatiales des dépôts quaternaires. Les modèles de résistivité inversée dérivés de cette acquisition ont été publiés et pourraient être utilisés pour diverses études futures. Cette étude met également en évidence le potentiel de l'ensemble de données pour le développement d'algorithmes d'intégration de données en raison de l'abondance de diverses données librement disponibles sur la même zone.

La deuxième étude propose une nouvelle méthodologie pour combiner les forages et les données géophysiques avec une propagation de l'incertitude pour prédire la probabilité d'argile à l'échelle d'une vallée. Une fonction de translation variant dans l'espace a été utilisée pour estimer la fraction d'argile à partir de la résistivité. Les paramètres de cette fonction sont inversés en utilisant la description des forages comme points de contrôle. Ils combinent cette estimation de la fraction d'argile avec un cadre

d'interpolation stochastique 3D non déterministe basé sur un algorithme de statistiques à points multiples et une fonction aléatoire gaussienne afin d'obtenir un modèle 3D réaliste à haute résolution spatiale de la fraction d'argile pour la haute vallée de l'Aar. L'étude démontre la qualité des valeurs prédites et leurs incertitudes correspondantes en utilisant la validation croisée.

La troisième étude porte sur la possibilité d'intégrer des données de forage, géophysiques et hydrogéologiques, tout en conservant la cohérence du concept géologique des modèles. Nous avons utilisé un générateur stochastique de modèles géologiques pour construire un ensemble de modèles préalables basés sur les forages. Nous proposons ensuite une approche d'inversion multi-échelle qui combine des modèles peu fidèles et moins précis avec des modèles plus fidèles et plus précis afin de réduire le temps nécessaire à la convergence de l'inversion. Les données géophysiques et hydrogéologiques sont intégrées à l'aide d'un algorithme ES-MDA (Ensemble Smoother with Multiple Data Assimilation Algorithm). Le flux de travail garantit que les modèles sont géologiquement cohérents et estime de manière robuste l'incertitude associée au modèle final. L'étude démontre l'efficacité de cette approche pour un cas synthétique contrôlé. Elle montre que ArchPY et ES-MDA sont capables de générer des réalisations plausibles de la subsurface pour les modèles sédimentologiques du Quaternaire.

Enfin, la quatrième étude présente une méthodologie innovante qui combine l'algorithme ES-MDA avec un code de modélisation géologique hiérarchique open-source pour intégrer des sources de données multiples et construire des modèles géologiquement cohérents avec une estimation d'erreur robuste. La méthodologie est appliquée à un cas de terrain dans la haute vallée de l'Aar, en Suisse. Un cadre de validation croisée est mis en œuvre afin d'évaluer la méthodologie. L'approche aboutit à des modèles finaux qui équilibrent efficacement la précision et l'incertitude et qui peuvent prendre en compte diverses sources de données, y compris des données géophysiques, des connaissances conceptuelles régionales, des forages et des mesures hydrogéologiques à l'échelle d'une vallée. En résumé, cette thèse présente plusieurs méthodes innovantes qui pourraient être appliquées à la réalisation de modèles hydrogéologiques à petite ou grande échelle.

Mots-clés

géostatistique; incertitude; gestion des ressources en eaux souterraines; modèles hydrogéologiques; Dépôts du Quaternaire; Intégration des données; Domaine temporel EM; Inversion stochastique; Ensemble Smoother with Multiple Data Assimilation Algorithm (ES-MDA); Inversion géologiquement cohérente.

Contents

Chapter 1

Introduction

1.1 Context

The shallow underground in most of Switzerland's urban and agricultural areas is composed of Quaternary geological deposits. These formations typically consist of loose materials such as gravel, sand, or clay deposited by glaciers, rivers, gravitational forces (such as landslides), or in lake environments. The interplay of these processes leads to complex three-dimensional patterns of various sediment materials. Quaternary deposits can be found just below a few centimeters of organic soil and can range from a few tens to a thousand meters in thickness in deep Alpine valleys. They are also a key research topic for Quaternary geologists as they provide valuable information for reconstructing recent climate changes. Quaternary formations are of great significance, as they serve as a source for water supply for cities or irrigation, construction materials extraction, civil engineering work, waste storage, and geothermal exploitation. The groundwater resource represents 99% of the total liquid fresh water on Earth and provides approximately 50% of the total volume of water withdrawn for domestic use (United Nations Educational Scientific and Cultural Organization 2022). In Switzerland, 80% of drinking water is sourced from groundwater according to the Swiss Office of Environment (NAQUA 2022). Of these 80%, half of it is even fed directly into the drinking water supply system without being treated. In this context, due to the appearance of persistent pollutants that are difficult to absorb in the subsurface, redifining protection areas considering not only residence time, but also recharge volume, will be necessary in the future. Figure 1.1 shows the results of the National Groundwater Monitoring NAQUA, QUANT module, which records the groundwater levels and spring discharge rates at around 100 monitoring stations across Switzerland (NAQUA 2022). The indicator calculates the annual number of monitoring stations with low, normal, and high groundwater levels and spring discharge rates, based on daily averages and percentiles over the standard period 2001-2020. Low discharge corresponds to the number of days with groundwater levels and spring discharge rates below the 10th percentile from 2001-2020. On the other hand, High levels is the number of days with groundwater levels and spring discharge rates above the 90th percentile from 2001-2020. We can denote some variability in time, but conclude that so far, climate change has not significantly impacted the large trend in terms of groundwater availability in Switzerland in the last 30 years. However, it is likely to cause changes in terms of water recharge and temporal repartition in the very near future. Moreover, the use of groundwater-based heat pumps in shallow aquifer environments

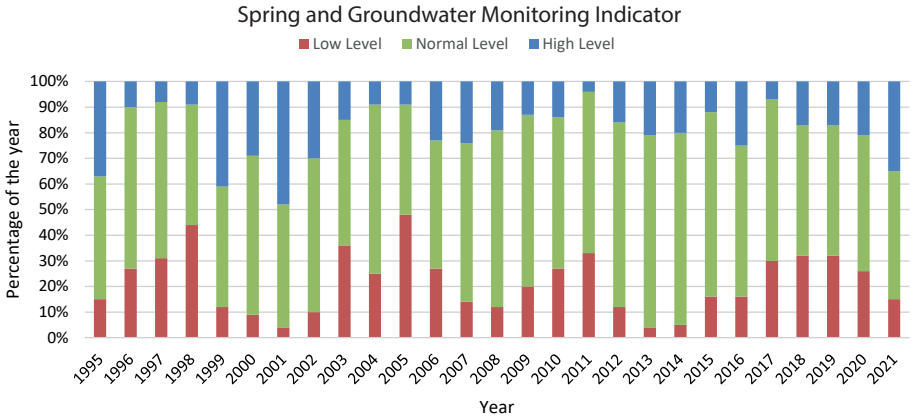

Figure 1.1: Groundwater and spring discharges in Switzerland, from the National Groundwater Monitoring NAQUA (NAQUA 2022). Low and High discharge corresponds respectively to groundwater levels and spring discharge rates below the 10th percentile and above the 90th percentile compared to the 2001-2020 distribution

is growing. However, a maximum absolute change of 3 °C in the aquifer must be respected compared to the original temperature, calling for large models to identify where the installation of a new system is possible.

For all the examples presented above, the realization of a medium- to large-scale hydrogeological model is necessary. The ultimate purpose of a hydrogeological model is to properly reproduce a complex underground system in a mathematical discretization in order to compute the theoretical response of an aquifer and represent the problem of interest. The response compute can take various forms, from advective flow to the transport of a solute. The complexity can vary, from fine large-scale 3D models to simplified 2D transmissivity models. In general, the workflow for the construction of a model is separated into a few steps (Pyrcz et al. 2014; Ringrose et al. 2015; Wellmann et al. 2018, e.g.). Typically, it consists of building a structural model, ideally with properties, and then simulating the flow within. In a typical classical workflow, boreholes and geophysical data are acquired first in the field. Geophysical data are inverted and result in a parametric model that depends on the method used. For all the electric methods, the main parameter of inversion will be the resistivity, when for seismic the density and impedence contrast will be outlined, for example. This parametric model is often

smoothed due to the regularization scheme used in the inversion and may present noise-indused artifacts. In addition, for most deterministic inversion techniques, this parametric model is only one amongst others that fits the acquired geophysical data, and its final shape will depend on the inversion parameters. This parametric model will be interpreted using colocated boreholes and expert knowledge and will produce a so-called structural model. This model usually represents large-scale structures, such as lithological units. This large-scale unit model is then filled with facies, either uniform or using geostatistical methods. Multiple facies can be present in the same unit: for example, river systems will consist of gravel in the channels, and clay in the floodplain. The petrophysical parameters are then assigned according to the unit and facies. In the case of a hydrogeological model, it can be porosity, hydraulic permeability, or storativity, for example. In the end, the exact distribution of these values, as well as other non-geological parameters such as recharge or boundary conditions, will be achieved during the model calibration. At this stage, hydrological measurements from the site are incorporated and the model is slightly modified to satisfactorily predict the measured value. In the end, the calibrated hydrological model can be used to simulate the problem of interest and take decisions based on its predictions. For most of the day-to-day applications in the private sector, a single calibrated model is used.

Considering a single model and not considering uncertainty can be problematic, especially in environements that can display sharp spatial discontinuities. Quaternary deposits, for example, consist of unconsolidated deposits with intertwined processes both spatially and temporally. A large part of Switzerland is covered with such sediments (see Fig. 1.2). These flat areas are generally more favorable for construction and agriculture and are often heavily anthropized. 90% of underground uses take place in unconsolidated rock layers (Volken et al. 2016). It consists mainly of the use of drinking water and the excavation of construction materials. Due to this, large amounts of data exist and are available; however, they are owned by various instances. This led the Swiss Geological Survey (Swisstopo) to initiate the *GeoQuat* project (Volken et al. 2016) to centralize existing geological data related to Quaternary Formations in Switzerland and build 3D underground models in four pilot regions: Birrfeld, Lake Lucerne, Upper Aare Valley, and Visp. In each of these valleys filled with Quaternary deposits, all available data were standardized and digitalized, and geological models were constructed using interpolation

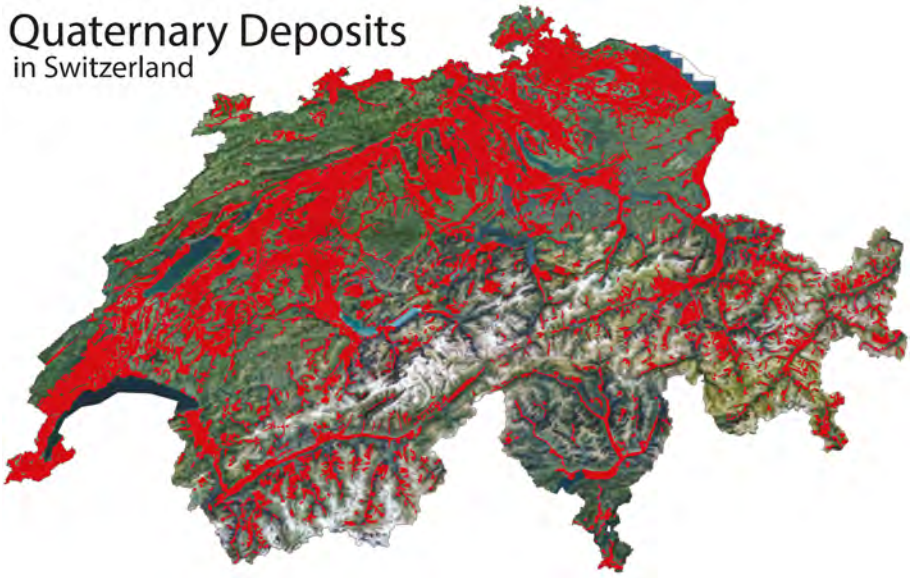

Figure 1.2: Location in red of the Quaternary deposits in Switzerland, modified from Volken et al. (2016)

methods. Similar projects have also been carried out in The Netherlands and Denmark (Barfod et al. 2018; Christensen et al. 2017; Stafleu et al. 2011). However, the construction of these models revealed certain limitations in existing techniques, such as the inability to automatically incorporate all available stratigraphic knowledge and ensure geological consistency. When the datasets are becoming large, manually combining them is time-consuming. Secondly, the interpretation and construction are extremely user-dependent. If the work is split into teams, important spatial inconsistencies arise. Finally, even if stochastic processes are sometimes used at some step of the modeling (for example for the facies simulation), the uncertainties are often not propagated through the entire workflow. The final models are often just one that fits the data, and no robust uncertainty estimation is provided. In summary, despite numerous studies dedicated to the construction of aquifer models in Quaternary formations, there is currently no methodology that integrates geological concepts, geophysical and hydrogeological constraints, and stochastic inverse modeling at a regional scale.

Multiple challenges exists for the development of such methods. First, almost all geological modeling software is proprietary and often expensive.

Most of them are designed for deterministic modeling and do not have an Application Programming Interface (API). In addition, there are often not suitable models for complex quaternary deposits, where drastic spatial variations can occur within the same unit. However, some changes are happening. For example, *GemPy* (Varga et al. 2019) is a Python open source geomodeling library that implements a generic implicit modeling approach. It supports the generation of faults and folds and can be coupled with a gravity anomaly forward, for example. Its implicit approach uses cokriging as the interpolation method. This package is very well suited for large-scale unit modeling, with, for example, a large density difference between the units, making the use of a homogeneous value relevant. In the Quaternary deposits, important facies, and consequently petrophysical, variations can occur in space. The use of stochastic methods and more advanced geostatistical methods is needed, but to our knowledge, no framework for efficient stochastic modeling in Quaternary environments existed. Due to the lack of suitable modeling tools, in the framework of this project, an open source hierarchical modeling software was developed (Schorpp et al. 2022). ArchPy can automatically generate realistic geological and parameter models. It operates hierarchically using input data consisting of a set of borehole data and a conceptual stratigraphic pile. The stratigraphic pile is a description of how the geological model should be constructed and contains information such as the order of the different stratigraphic units, their conformability, surface interpolation methods, filling methods for lithologies, and petrophysical properties. The modeling process is fully automated, which means that once the stratigraphic pile and its related parameters have been defined, the software will generate the model automatically. The process involves simulating the stratigraphic unit boundaries (units), filling them with lithologies (facies), and simulating petrophysical properties inside the lithologies. The approach is very flexible and provides a framework for generating end-to-end stochastic models. Additionally, the method allows for uncertainty quantification at any level and may be used for full inversion.

Second, even with the appropriate modeling tool, integrating data within these models and keeping geological realism is challenging. As noted in Linde et al. (2015), it is common for geological models to include millions of elements. For an inversion, computational resources are often demanding due to the high number of forward and simulation calls needed. If the number of parameters is important, classic stochastic approaches such

as Markov Chain Monte Carlo would will need several thousand to hundreds of thousands iterations (Irving et al. 2010; Linde et al. 2015). To overcome this important computational cost, serveral approaches have been considered. Model reduction is a technique that aims to find a lower-dimensional model that is still representative of the higher dimension. This can be achieved by working on a higher scale, inverting a parameter that has no physical meaning, but is controlling others. The inversion can then be performed on these parameters and on larger-scale parameters. There are several recent examples. For example, Wang et al. (2022) introduced a hierarchical inversion approach, where the method first estimates the posterior distribution of global variables, such as hyperparameters of geostatistical models. Also, Lam et al. (2020) successfully applied stochastic inversions on variables linked with Gaussian pyramids to the model. Finally, Hansen et al. (2022) used different approaches to directly estimate the distribution of the posterior models, greatly reducing the dimension of the problem. However, in these cases, keeping the models geologically consistent can be difficult. Another approach is to implement a sampling-based method. At each iteration, some fraction of the model is sampled and will be considered fixed, and the rest of the model is updated. This allows us to generate a model that is a combination of an ensemble of models, compatible with the prior. This approach can then be coupled with Markov chain Monte Carlo (MCMC). A good example of this methodology is the work by Mariethoz et al. (2010) and their idea of iterative spatial resampling. More recently, Amaya et al. (2022) proposed a sampling-based method coupled with an innovative adaptive sequential Monte Carlo, where information is passed through the parallel Markov chains to improve the convergence of the inversion. Finally, Khambhammettu et al. (2020) used pilot points to condition geostatistical realizations. But instead of working on the infinite space of parameter value, it only considered the spatial position of the conditionning point.

All of these methods succeeded in integrating significant geological knowledge. However, even with some dimension reduction strategies, the few thousands of MCMC iterations needed are just too computationally expensive when large-scale models are considered. To address this issue, Ensemble Smoothers with Multiple Data Assimilation (ES-MDA) have been developed and have been shown to efficiently solve complex nonlinear inverse problems, surpassing both traditional Ensemble Smoothers (ES) (Emerick et al. 2013) and Markov Chain Monte-Carlo

(MCMC) methods (Juda et al. 2022) in terms of finding an optimum quickly. Originally, this method was introduced to be applied for meteorological forecasting models. It uses a Monte Carlo approximation of the Kalman Filter (Kalman 1960) and estimates the relationships between state variables and parameters using an ensemble of models. The covariance matrix is therefore approximated using the experimental covariance of a finite set of models. ES-MDA and PESTPP-IES (White 2018) are variants of the Ensemble Smoothers algorithm proposed by Leeuwen et al. (1996). ES-MDA performs iterative ES corrections of the parameters by assimilating the data from the previous iteration, while PESTPP-IES optimizes an objective function using a modified form of the Levenberg-Marquardt algorithm. A comparison study by Lam et al. (2020) indicated that PESTPP-IES outperforms ES-MDA with small ensembles (200 in the study), but that ES-MDA improves with larger ensembles while PESTPP-IES does not. ES-MDA may be more appropriate for models with high-dimensional state spaces and large amounts of data, while PESTPP-IES may be more appropriate for models with complex parameter spaces and limited data. ES-MDA has been successfully implemented in several groundwater studies (Kang et al. 2019; Lam et al. 2020a; Xu et al. 2022), and will be used in this thesis. It is important to denote that ES-MDA requires the assumption of a normal or multi-Gaussian distribution of state variables and parameters. When this assumption is not met, a normal score transform can be applied to ensure that marginal distributions are Gaussian (Zhou et al. 2011).

To address the current limitations in hydrogeological modeling and inversion in Quaternary environments, the PheniX project, funded by the Swiss National Fund (SNF) was started in 2019, and this thesis was realized in its framework. The project has several partners, both in the private sector, public authorities, and universities. In the framework of this particular thesis, we can especially mention the Swiss Geological Survey (Swisstopo) and the University of Aarhus in Danemark. This large SNF-funded project aims to create a comprehensive set of tools to assess uncertainty and integrate data in groundwater modeling, with a focus on quaternary aquifers, as they are the most significant and frequently utilized groundwater resources. It focuses on three key methodological components (see Fig. 1.3): novel random function models, integration of geological knowledge, and integration of variables by inverse conditioning. In addition to the methodological tasks, it also includes large data acquisitions to benchmark the newly developed algorithm in a test

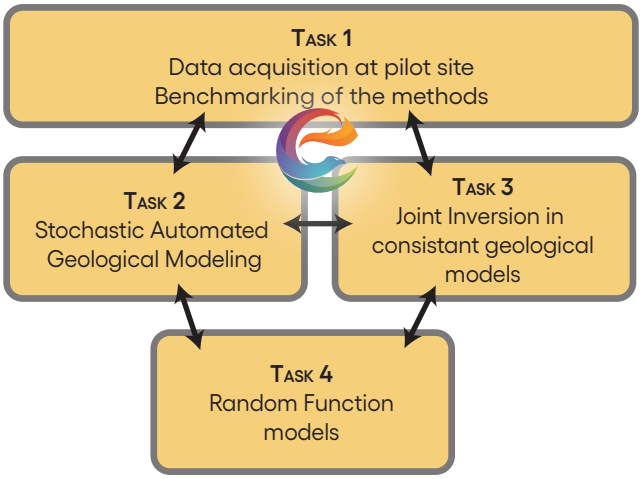

Figure 1.3: PheniX project organisation

site. The project involves two PhD students, Alexis Neven and Ludovic Schorpp.

This thesis aims to propose new methodologies to combine serveral datatypes to improve the knowledge of the underground. This thesis is divided into six chapters, including the present introduction. First, this thesis will introduce the pilot test site used in this study, *The Upper Aare Valley* in Switzerland. A rapid geological and hydrological context will be established in Chapter 2. Then, because geophysical data was lacking in the area, a large geophysical survey was carried out in close collaboration with the Hydrogeophysics group of the University of Aarhus and their new transient electromagnetic ground based instrument tTEM (Auken et al. 2019). The results of the acquisition, as well as the processing and validation of the data, will be presented in Chapter 3. Chapter 4 presents an approach to predict the probability of clay from deterministically inverted resistivity models and boreholes at the scale of the entire valley. This was done by expanding the existing work of Foged et al. (2014), and implementing it in a stochastic interpolation framework using Multiple Point Statistics combined with non-deterministic Training Images. In doing so, we were able to predict the probability of clay with a robust propagation of uncertainty from the data to the final model. To improve this approach, where no geological concepts are involved, we decided to focus the study on the interaction of data in consistant geological models. Using a hierarchical stochastic

geological modeling code developed within the PheniX project (Schorpp et al. 2022), Chapter 5 presents synthetic examples, where geophysical and hydrological data are intergrated in complex geological models in controlled synthetic cases. The methodology proved to be reliable on the controlled case and efficiently reduced the uncertainty on the structures themselves, but also on the hydrological predictions derived from the models. Finally, Chapters 6 and 7 present a similar approach applied in a part of the upper Aare Valley using real field data. Real head measurements and geophysical data acquired in the valley are combined and produce realistic and complex models that honor the conceptual geology of the area, the boreholes, and the data, within a stochastic inversion scheme. This thesis introduces several approaches that could be widely applied in the future to model complex quaternary deposits, where the data are abundant.

Chapter 2

Pilot Site : The Upper Aare Valley

2.1 Hydrostratigraphy and context

The Pilot Site, located in central Switzerland, is situated in the Upper Aare Valley, extending over a section of approximately 20 km from the city of Thun to Bern (see Fig. 2.1. This region showcases the typical Quaternary geology of Alpine valleys, with a basement layer typically a few hundred meters deep and covered by a complex intertwinding of glacial, fluvial, and fluvioglacial deposits. The Swiss basin has undergone multiple cycles of glacial advance and retreat over time, leading to significant changes in the depositional processes.

The Aare valley belongs to the northern Alpine Molasse Basin, Swiss Molasse Basin (SMB) in Switzerland, a typical peripheral foreland basin (Ibele 2011). The landscape of the Swiss Alpine foreland is the result of multiple glaciations and deglaciations that occurred during the Quaternary period over the past 2.6 million years and the study area represents one of the main valleys through which glaciers reached the lowlands of Switzerland (Keller 2021; Preusser et al. 2011). The bedrock on which the quaternary depostits lie in the Aare Valley is the Molasse, which consists mainly of sandstone with subordinate conglomerates. According to Ibele (2011), fifteen major glacial advances and retreats occurred in the Swiss Alps during the Quaternary period. At least eight glacial cycles can be identified in deep boreholes, but a complete description of the lithostratigraphy based solely on borehole descriptions is nearly impossible due to the possibility of similar deposits of different ages being superimposed and intertwined. A comprehensive hydrogeological analysis of the valley was carried out during the late 1970s and early 1980s, as documented by Kellerhals et al. (1981). Subsequently, additional data has been collected through various projects in different areas of the valley (Preusser et al. 2004; Schlüchter 1989). Despite these efforts, no updated and comprehensive hydrogeological synthesis has been produced and made available to the public.

More generally, the regional hydrostratigraphy has been described by Messan (1992), and can be summarized as follows, from top to bottom :

- A surface aquifer composed of the most recent alluvial deposits is present throughout the valley, with the water table found between a few dozen centimeters and 2 meters below the surface. The exact extent and thickness of the upper aquifer are not entirely known,

but it is estimated to have a thickness ranging from 2 to 20 meters depending on the area. This aquifer is composed of sandy-gravel with very good permeability.

- A thick layer of highly variable thickness lacustrine clay constitutes the bottom of the shallow aquifer. Its average thickness is estimated to be 50 m, but some seimic lines have suggested that it can reach 125 m. It can be considered as an impermeable layer.

- A second deeper aquifer has been identified in some deep wells, separated from the shallow one by the lacustrine clay layer. However, the details of this aquifer remain largely unknown and its depth, extent, and connectivity with the upper aquifer are yet to be determined. It is constituted of sandy to silty gravels, and will be referred to in this thesis as the lower aquifer.

The Aare River flows through the Upper Aare Valley. This 290 km long stream links the Aare Glaciers to the Rhine River through the city of Bern, feeding on the way the Brienz lake, the Thun lake and the Biel Lake. Toward the north of the pilot site, another stream is crossing the valley. It comes from the parallel Gürbe valley. However, no contribution to the aquifer was established from this side valley. Both rivers are monitored and historical data are available. In the study area, several hundred wells are in use either for shallow geothermal exploitation or drinking water, as well as 5 quarries. With such an extensive use, a 3D model would be of immense benefit to local authorities. It would not only help to assess the vulnerability of the upper aquifer, but would also provide valuable insights into the characteristics of the second deeper aquifer. This information would be invaluable to ensure sustainable management of water resources in the Upper Aare Valley and their long-term viability. Furthermore, the model could be used to identify areas with the highest potential for geothermal energy generation, providing a critical resource for the region's energy needs. In general, a 3D model of the study area would provide an essential tool for managing and preserving the valuable water resources in the Upper Aare Valley. This is why this valley was retained as a test site for the *GeoQuat* project.

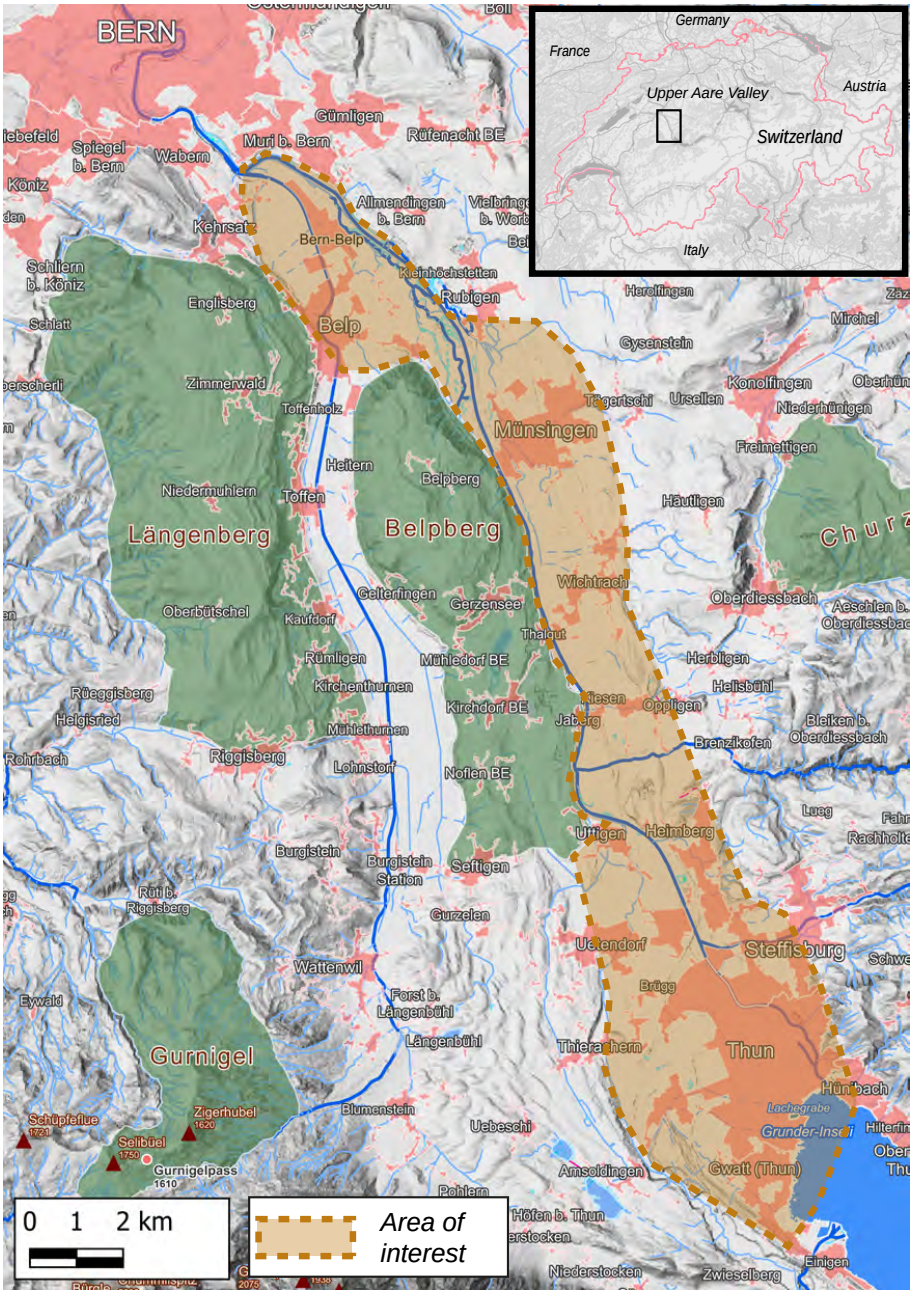

Figure 2.1: Situation of the Pilot site

2.2 Available Data

To create 3D models of the shallow subsurface, uniform data must be accessible. Due to the widespread use of the shallow underground, a large amount of information is available on its structure. Drilling operation records are the most significant data source and show the arrangement of geological layers at a particular location. Other information sources include the outcomes of geophysical surveys (seismic, gravimetric, geoelectric, electromagnetic, etc.) and geological cross sections. The experimental project *GeoQuat* was carried out from 2014 to 2018 by the Swiss Geological Survey (Swisstopo). It aimed to establish bases and standards for the description and management of wells in Quaternary environements, and propose modeling methodologies. Two main delivrables were obtained from this project : 1) a database of standardized and described boreholes that will be used for all the models presented in this thesis, and 2) a deterministic nearest neighbor unit model, that will be used for comparison with our methodology.

Figure 2.2 A) shows the quality rating statistics given to the wells present in the GeoQuat database. As mentioned above, the Quaternary data available come from different sources and have varying degrees of quality. For this reason, a grade was given to each borehole during the digitalization of the data. The grade comprises of three criteria: the quality of the geological description of a particular layer, the positionning in space, and the altitude. The first criterion takes into account the amount of description given in the lithological logs. Quality 1 or 2 only consists of a succinct description of the main component, such as "gravel", "sand", or "clay". Quality 4 or 5 in contrast includes the grain size, quantity of the different components, and measured parameters (permeability, density, porosity, ...). There is also significant uncertainty in the location of wells in the XYZ coordinates. Therefore, the second and third criteria evaluate the quality of the available information on the XY and Z positionning. "Very bad data" have a position accuracy in the hectometer scale (>100m), while "excellent data" have a position accuracy on the decacentimeter scale (<1m). The Z location is strongly connected to the precision of the XY coordinates. Using the available XY coordinates, the Z value of a borehole (surface level) is compared to the SwissALTI3D digital elevation model (DEM). The quality of the information on the elevation of Z is then assessed based on the match or mismatch with the data from the SwissALTI3D DEM. The detailed methodology can be found in Volken et al. (2016). More than 75% of the wells were graded

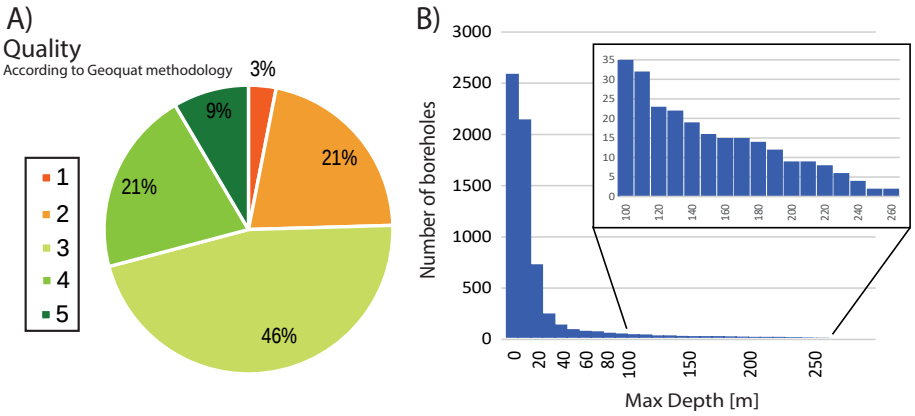

Figure 2.2: Analysis of the borehole depth and quality in the GeoQuat Aare Valley database

above or equal to 3. This QC grade will be used to estimate the uncertainty on the input data.

Figure 2.2 B) shows the distribution of the vertical depth of the wells (TVD) in the valley. Due to the shallow depth of the aquifer and since most of them were conducted for shallow geothermal, water pumping or geotechnical investigation, only about 10% exceed the depth of 20 m. This sampling bias as seen in numerous sites and causes an oversampling of the shallow and porous layers. The lower aquifer is only identified in a few dozen wells.

In addition to the borehole measurements, hundreds of head measurements were conducted in the valley by Kellerhals et al. (1981). This campaign was financed by the Swiss authorities for the realization of the regional hydrogeological map. The data were acquired in 1976. A study was carried out within the PheniX project to show that the measures were still representative of the current conditions (Vallat 2022). Using automated measurement stations from the Bern state (see Fig. 2.3), they first analyzed that the time series of hydraulic heads do not appear to show large variations. By interpolating the values with the positions of the 1976 measurements, they concluded that there has not been a significant change or global trend in terms of groundwater levels in the upper Aar valley compared to the piezometry conducted by Kellerhals et al. (1981). The hydraulic heads are almost identical. This indicates

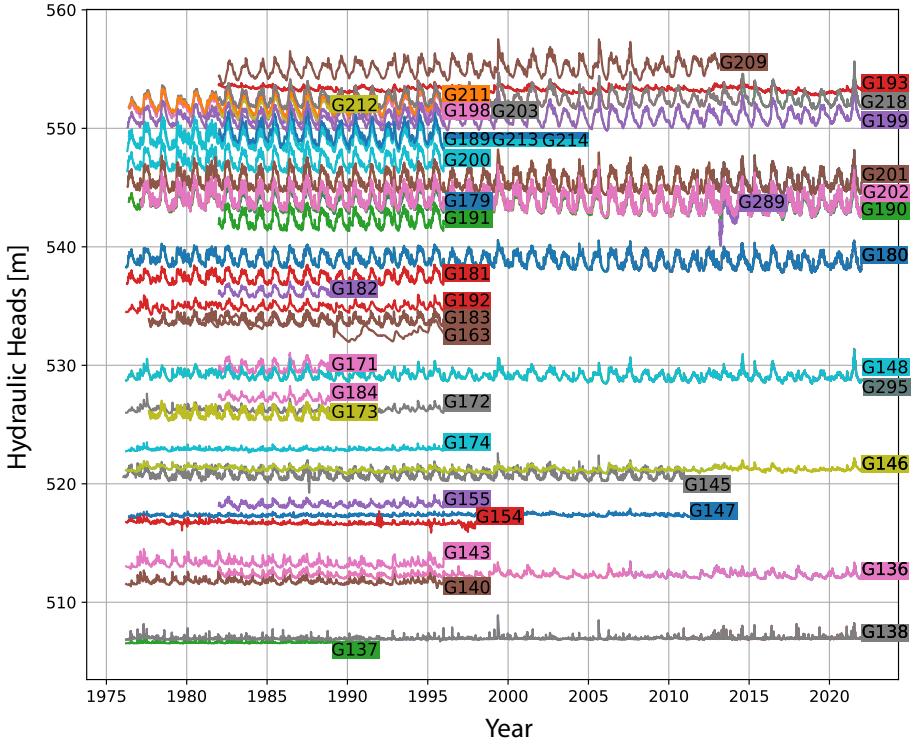

Figure 2.3: Time series of hydraulic heads of the Aare valley stations from 1976 to 2022. Modified from Vallat (2022).

that the variation in hydraulic load has remained relatively stable over the past 40 years. However, they indicate that the upstream section of the aquifer experiences more significant variations in hydraulic load compared to the downstream area. One possible explanation for this is that the installation of numerous drains in the downstream area can mitigate fluctuations in the water level at that location, or can be caused by changes in terms of river connectivity with the groundwater system. The complete statistical analysis is available in Vallat (2022).

Finally, table 2.1 summarizes the available data and their availability.

2.3 Conclusion

As showed in this chapter, the upper Aare Valley is an extensively sampled quaternary valley. Its upper aquifer is an important ressource for

Table 2.1:

Type	Data	Availability
Boreholes	Boreholes described lithologically in USCS, with indication of Quality. Presence of stratigraphy in part of them.	Upon request to Swisstopo Source: Volken et al. (2016)
Boreholes	Boreholes description available as PDF format. Some information such as Quaternary thickness or aquifer depth is encoded.	Available on Canton Bern Geoportal Source: Canton Bern
tTEM	Raw Soundings and inverted resistivity models.	Available Source: Neven et al. (2021)
Heads	Regional piezometric measurements and River measurements.	Available Source: Kellerhals et al. (1981)

the region. Due to the abundance of data present, it make this area an ideal place to test and benchmark data integration workflow. However, so far no wide geophysical coverage has been available. For this reason, in the framework of this thesis, a large Towed Time-Domain EM was conducted in the valley. The details of the acquisition and the results are presented in the next chapter.

Chapter 3

tTEM20AAR: a benchmark geophysical dataset for unconsolidated fluvio-glacial sediments[1]

[1]This chapter was published as : Alexis Neven, Pradip Kumar Maurya, Anders Vest Christiansen, and Philippe Renard (June 2021b). "tTEM20AAR: a benchmark geo-physical data set for unconsolidated fluvioglacial sediments". In: *Earth System Science Data* 13.6, pp. 2743–2752. DOI: 10.5194/essd-13-2743-2021

Abstract

Quaternary deposits are some of the most complex geology to model, despite the fact that they are some of the most abundant and extensively used ones. In order to improve the knowledge about spatial variations of such deposits, we acquired a large (1500 hectares) and dense (20m spacing) Time Domain ElectroMagnetic (TDEM) dataset in the upper Aare Valley, Switzerland. TDEM is a fast and reliable method to measure the magnetic field directly related to the resistivity of the underground. In this paper, we present the inverted resistivity models derived from this acquisition. The depth of investigation ranges between 40 to 120m depth, with an average data residual contained in the standard deviation of the data. Possible reuse of these data is generation of quaternary geological and hydrological models. In addition, It is an ideal dataset for data integration algorithm development due to the abundance of various data freely available on the same zone, such as described boreholes, other geophysical methods and water table monitoring stations.

3.1 Background & Summary

In most urbanized and agricultural areas of Switzerland, the shallow underground is constituted of Quaternary deposits. The thickness can vary from few meters to hundreds of meters. These recent sediments are deposited by various agents such as river, lake, glaciers or even landslides. Each time, the associated sediment will have a different composition, permeability, and spatial extents. Furthermore, they might be all intertwined (a lake deposit can be partially eroded by a glacier, and refilled with river sediments). This leads to a spatial variability that is often higher than expected in such deposits.

In Switzerland, these formations cover almost 30% of the country, and includes all the most populated areas. In addition, these formations are some of the most solicited : water supply for cities, extraction of construction materials, geotechnical constructions and shallow geothermal exploitation. Often, the construction of geological models using only boreholes can miss most of the spatial heterogeneity, and conduct to inadequate models and wrong conclusions. Increasing the number of boreholes to reduce the uncertainty is often difficult and expensive. A good example of these highly exploited Quaternary zones is the upper Aare Valley in Switzerland. In the 20 km by 3 km rectangle defined by the

valley side limits, the city of Thun and the city of Bern, the Aare Valley includes 4 quarries, several hundreds of pumping (Shallow geothermic or drinkable water) and injection wells (re-inject water after geothermal heat pump). A previous valley size model was designed using only boreholes and surface data (Volken et al. 2016), but the model does not represent all the internal heterogeneities of the Quaternary formations and can show unrealistic sharp variations due to the nearest neighbors interpolation method used during the workflow. Therefore, there is a need for a better understanding of Quaternary sedimentary heterogeneity, in order to better constrain geological models, knowledge that could be applied in the Aare valley or for any fluvio-glacial filling area.

Near-surface geophysics such as DC resistivity, electromagnetic or seismic can bring important information in terms of the spatial distribution of facies. However, they are usually carried out in restricted areas to answer specific local questions, and do not help to understand the variations of geology at the valley scale. Consequently with such datasets, it is difficult to develop general modeling methods that can be applied to fluvio-glacial deposits, or even to come up with simple conceptual models for large scale filling. In order to fill this gap of information, and provide a valley scale fluvio-glacial resistivity map, we conducted in January 2020 a large geophysical survey using tTEM (towed Transient Electromagnetic) system (Auken et al. 2019) in the upper Aare Valley, Switzerland. The tTEM-system provide a both vertically and horizontally, very detailed resistivity model. The dataset covers a section of the valley of approximately 1-2 km width and 16 km long. The fields were mapped with a line spacing of 20 meters, resulting in about 1500 hectares of covered land (see fig. 3.3). The raw tTEM data were processed to suppress and removed noisy data parts, and then inverted to a resistivity model using specially constrained inversion algorithm (Viezzoli et al. 2008). The resulting resistivity model consists of 57000 1D models of 30 layers. The depth of investigation varies, from 40 to 120 meters depth, primary driven by lithological/resistivity variations. The obtain resistivity model explain (fits) the recorded data well within the estimated data uncertainty. The resistivity model reveals new and very interesting geophysical/geological structures of the subsurface in a fine detain.

In addition, the Upper Aare Valley has been extensively studied and a consequent amount of data are freely distributed by the Swiss authorities

: A few thousand of logged boreholes, 46 station of present or historic groundwater level measurements, a complete bedrock elevation model and Quaternary deposits thickness, as well as a complete hydrological and geological map. All these data cover the same site, and in combination with the geophysical data presented in this paper, it forms an ideal test site for multiple data integration. Especially, we consider this site ideal for benchmarking new methods that integrates hydrological and geophysical data, or more generally that may link porosity, permeability or lithology with resistivity. Improvement in such methods may strongly improve hydrogeological modeling in such environments, subject to high local facies variations.

3.2 Methods

The tTEM-system

The tTEM-system used for the data acquisition is developed by the HGG-group at Aarhus University, Denmark (Auken et al. 2019). The tTEM-system is a towed, ground-based, transient electromagnetic system, designed for high efficient data collection and detailed 3D-mapping of the shallow subsurface (the upper 80 m). TEM-methods build on the principle of induction (Faraday's law of induction) for mapping the electrical conductivity (conductivity=1/resistivity) of the subsurface. A detailed description of the TEM principle can be found in Christiansen et al. (2009). The layout of the tTEM system is shown in Figure 3.1.

The tTEM-system consists of an ATV carrying the instrumentation and towing the transmitter frame (Tx coil) and the receiver coil (Rx coil) in an offset configuration. The Tx and Rx coils are mounted on sleds for all terrain capability. All frame parts and sleds are built of non-conductive composite materials. Driving path and various data quality control parameters are monitored in real time by the driver on a mounted screen. Operation speed is up to 20 km/h. We used an off-set configuration, where the receiver coil is 7 meters behind the transmitter coil. Both of them are horizontal, allowing us to measure the z component of the secondary magnetic field. A GPS in mounted on the frame to ensure correct positioning of the data. The transmitter loop consists of one loop of 4x2 m, creating an area of 8 m^2. We used a standard dual moment TEM configuration: a high moment (HM) with a high inductive current of 30 A and a low moment with a lower inductive current of 5 A.

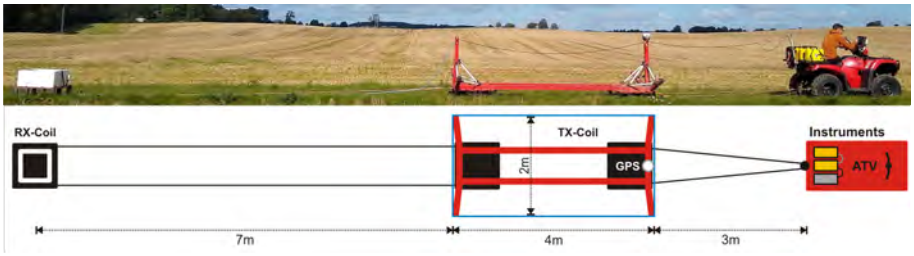

Figure 3.1: The tTEM system.

Parameters	**LM**	**HM**
No. of turns	1	
Tx Coil area	8 m^2	
Transmitter Current	5 A	30 A
Peak moment	30 Am^2	240 Am^2
Repetition Frequency	1055 Hz	315 Hz
Stacks	422	252
Total cyclus time	0.22 s	0.40 s
Tx time	0.2 ms	0.45 ms
Turn off Time	2.8 μ s	4.5 μ s
Number of gates	4	23
Gate Size	4 μ s - 10 μ s	10 μ s - 900 μ s
First gate start	4.38 μ s	10.30 μ s

Table 3.1: Specifications of the High and Low moment used in the acquisition. The gate size increases with time in order to counterbalance less good signal to noise ratio due to the wave attenuation.

Such configuration has the advantage of being able to resolve shallow targets with the low moment and its associated fast turn off time, and to reach higher penetration depth with the high moment. Both moments are stacked few hundreds times. Detailed parameters are summarized in the table 3.1. The gate is the time interval in which we will average the received amplitude. Due to the signal attenuation, further we get in the listening time, lower is the signal to noise ratio. In order to partially counterbalance this effect, we used a logarithmic increasing gate size related to listening time.

To ensure the quality of the data, prior to the survey the tTEM-instrumentation were calibrated at the Danish national TEM test site

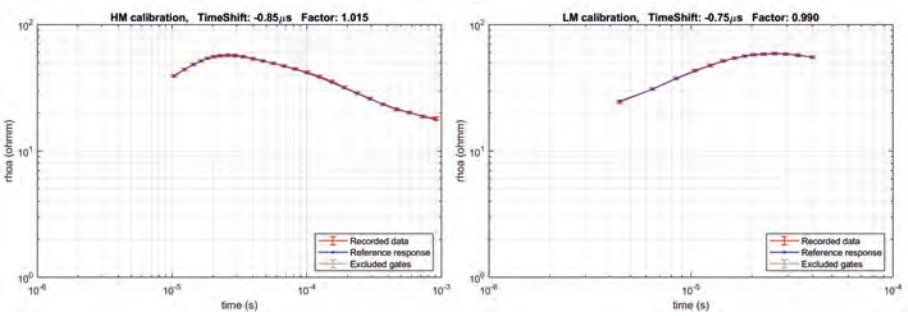

Figure 3.2: Calibration of the High and Low moment. The resulting time shift and scale factor are respectively -0.75 μs and 0.99 for the LM, and - 0.85 μs and 1.015 for the HM.

following the calibration procedure described by Foged et al. (2013). The two parameters calibrated are a time shift and an amplitude factor. The calibration was done with the ATV connected to the equipment, in order to account for any shift caused by it. Figure 3.2 shows the match between the test site reference response and the measured tTEM-response after calibration, which results in a fully acceptable match.

Field Site

The field site is the Upper Aare Valley, in central Switzerland (see figure 3.3). The survey took place in January 2020. During approximately 15 working days, we covered all the accessible farming fields in the valley along a 26 km long section. The driving speed was between 10 to 20 km/h, depending of the terrain. Since the acquisition rate is time dependant, and not distance triggered, we also lowered the speed in noiser or less responsive areas in order to acquire a denser dataset. The spacing between the lines was approximately 20 meters. The average covered surface par day was 112 hectares, for a total of 1425 hectares.

Data Processing

The voltage data from the receiver is measured continuously, and need to be cleaned of man-made noise and coupling. Data processing and inversion were carried out with the tTEM processing module in the Aarhus Workbench software. The objective of the processing of the tTEM-data

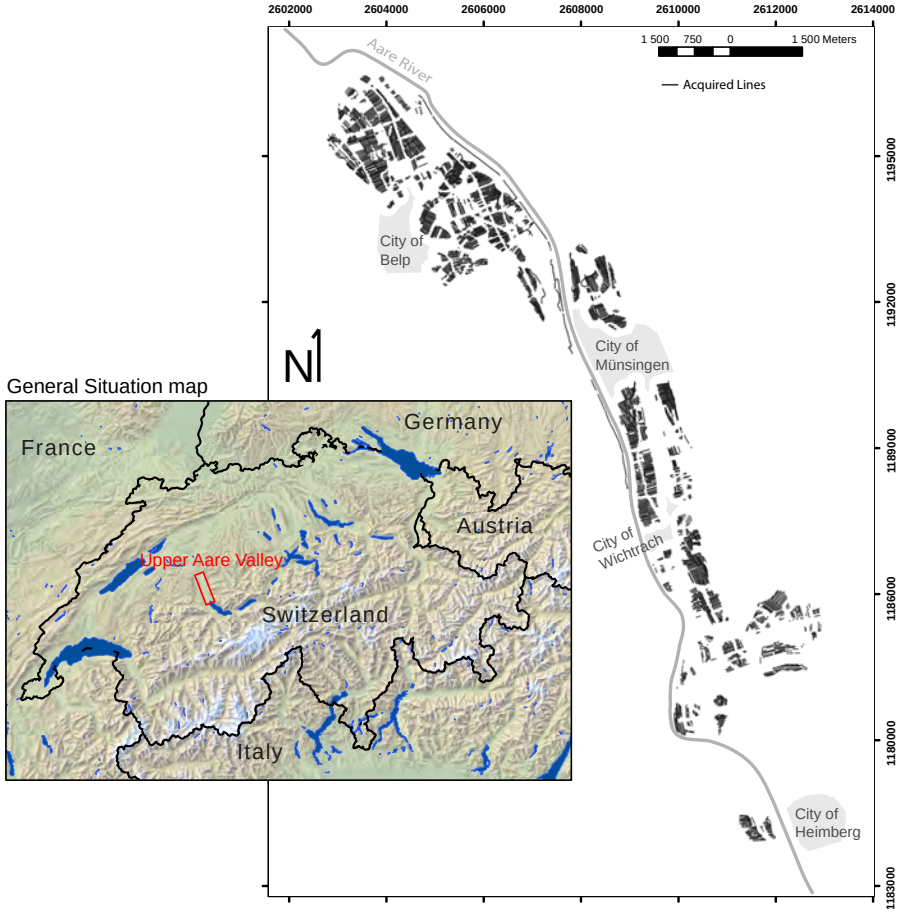

Figure 3.3: Acquired lines during the January 2020 field session. Coordinates are in UTM 32N.

is to remove any interference in the data from man-made installation (coupled data), suppress random noise by staking, and finally discard the noisy late time data entering the background noise. Thus, we ensure that the resulting resistivity model represents geological structures of the subsurface without artifact from man-made installation. Processing of the dB/dt data comprises of the following steps:

- Automatic filtering of raw data for removal of coupled data.

- Averaging of raw data to suppress random noise. Raw data are averaged using a moving average filter with narrow time windows in early times and wider in the late times.
- Following the averaging of raw data, sounding is created every 2.5 s which creates sounding spacing every 10 m distance along lines approximately, however, it may vary depending on driving speed.
- Automatic filtering of the averaged data for removal of late-time data points entering the background noise.
- Visual assessment of all dB/dt data and manual removal of coupled data not detected by the automatic filtering.
- Evaluation and adjustment of the data processing based on preliminary inversion results.

Furthermore, GPS data are lag-corrected to geographically position data models at center between transmitter and receiver coils. The uncertainty of the data consists of 3% of uniform data STD plus the STD calculated from the data stacking. Averaged data resulting with STD over 30% are discarded from inversion.

Inversion

The electrical resistivities of the underground are then estimated using a series of 3D constrained 1D-inversions. The 1D inversion is based on the AarhusInv code (Auken et al. 2015; Kirkegaard et al. 2015). This code is a implementation of a 1D non-linear damped least-squares solution, with a modeled transfer function for the TEM instrumentation. This function takes into account the transmitter waveform, the instrument low pass filters, the receiver bandwidth, the system geometry, the gate widths and the instrument front gate. However, in such an standalone 1D inversion, each model is totally independent of the neighboring ones unlike geology. To improve that the spatially constrained inversion (SCI) (Viezzoli et al. 2008) method was used in our inversions. It applies 3D constraints to 1D inversion models both along and across the mapping lines, with a weight that is decreasing with distance. All the inversions were carried out with the Aarhus Workbench software.

Table 3.2: Settings used for the model setup, the smooth and the sharp regularization.

Item	**Parameter**	**Value**
Model Setup	Number of layers	30
	Model resistivity start value (uniform - no prior)	40 ohmm
	Thickness of first layer (m)	1 m
	Depth to last layer (m)	120 m depth
	Thickness of layers	Log increasing with depth
Smooth Constraints	Factor of horizontal contrains on resistivites	1.5
	Factor of vertical contrains on resistivites	2.0
	Reference distance	10 m
	SCI Constraints with distance	$1/distance^{0.75}$
	Prior, thickness	Fixed
	Prior, resistivities	None
	Minimum number of gates per inversion point	2
Sharp Constrains	Factor of horizontal contrains on resistivites	1.12
	Factor of vertical contrains on resistivites	1.08
	Reference distance	10 m
	SCI Constraints with distance	$1/distance^{0.75}$
	Prior, thickness	Fixed
	Prior, resistivities	None
	Minimum number of gates per inversion point	2
	Sharp vertical constrains	500
	Sharp horizontal constrains	300

The SCI inversion can be used with two different regularizations : smooth or sharp. The smooth scheme tends to minimize abrupt changes in resistivity, in the vertical and horizontal directions. On the other hand, the sharp regularization scheme tends to minimize the number of resis-

tivity changes, but will consequently result in more abrupt resistivity transitions and a potential more blocky model appearance. Both regularizations were used, and are included in the output data.

For each resistivity model, we estimate the depth of investigation (DOI) using a method based on the Jacobian Sensitivity matrix (Christiansen et al. 2009a). This method has the advantage of taking into account the full transfer function, including system geometry, data uncertainty and the resistivity model. Two DOI thresholds values in the sensitivity matrix were used to provide the reported DOI-standard, and the DOI-conservative values. As a guideline, the resistivity structures above the DOI conservative value are strongly data driven, while resistivity structures below the DOI standard value are weekly represented in the data. Normally one would blank the resistivity models below DOI standard value.

Inversion setup for the smooth and sharp inversions are summarized in the table 3.2.

3.3 Data Records

Processed data, the resistivity models and the associated forward responses from the smooth and sharp inversions are provided in column based ASCII files. Each file structure is outlined in the following sections.

Processed data file

The *Processed_ Data.dat* file contains the processed tTEM data and data uncertainties. Each line in file corresponds to a LM or HM data stack for a given location. The RECORD number links the LM and HM data to a given resistivity model in the *.inv files. Number 9999 marks discarded data points or data points not present for the given moment. If all the data points of LM or HM are discarded then the data line is not present in the file. Gate center time and other info is stated in the header lines. The data uncertainty is given as relative in log space. The upper and lower bounds of the data are then defined as:

$$unc_{down} = \frac{DATA}{1 + DATASTD}$$

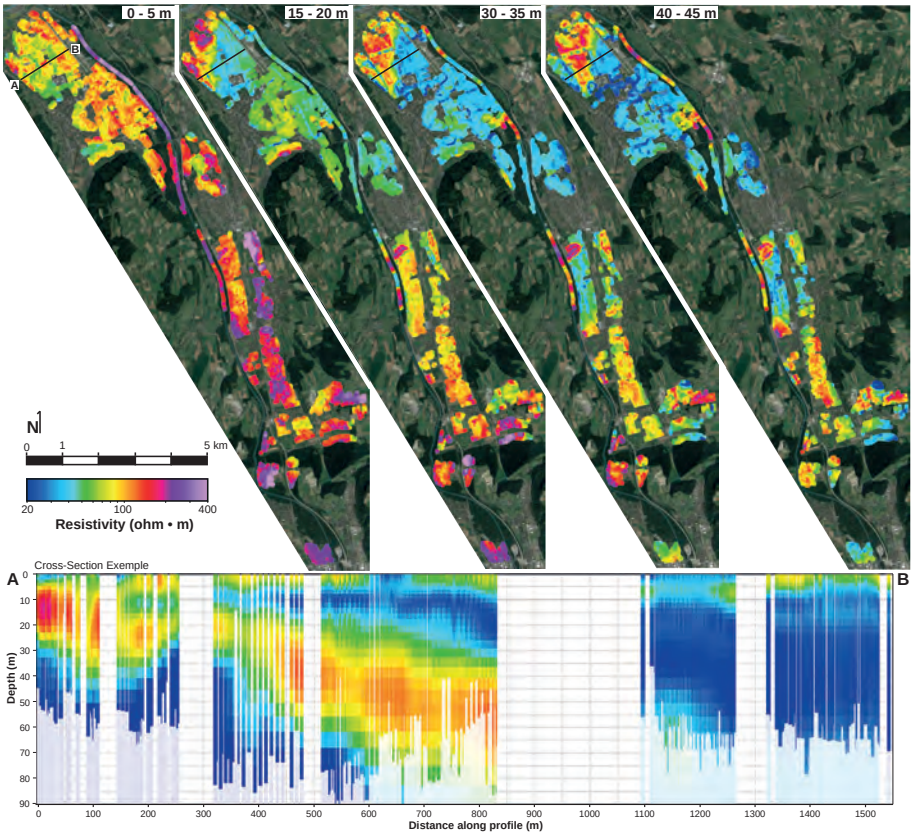

Figure 3.4: Mean Resistivity maps at different depth intervals from the smooth regularization model. In addition, a cross section is included. Base map from Swiss Federal Topographic Office

$$unc_{up} = DATA * (1 + DATASTD)$$

with $unc_{down\#}$ and $unc_{up\#}$ being the absolute lower and upper uncertainties, $DATA_{\#}$ the processed z-component dB/dt data value and $DATASTD_{\#}$ the relative uncertainty. The structure is outlined in the table A.1.

Inversion Model File

The Sharp_Model.inv and Smooth_Model.inv files file contains the resistivity models (layer resistivity and layer thicknesses). Each line hold a 30-layer resistivity model. The RECORD links the model to the data in the process data and forward data files. The file also contains the DOI, and the data fit. Note that last layer (layer 30) does not have a thickness since in continues to infinity depth in the modeling. Normally the DOI standard values are used to blank the models in depths. Detail file structure is provided in table A.2.

Synthetic response file

The Forward_Data_Sharp.dat and Forward_Data_Smooth.dat files contains the forward responses of the sharp and smooth resistivity models. The structure of the forward data files is the same as the Processed_Data.dat file except that the forward responses does not have associated data uncertainties. Detailed file structure is provided in table A.3.

3.4 Technical Validation

The quality of inversion is assessed by a quality control parameter called data misfit. We compare the forward geophysical response of our final resistivity model, with the field data, normalized by the square of the standard deviation of our data. The indicator is defined by the following equation 3.1.

$$DataMisfit = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{(d_{obs,i} - d_{frw,i})^2}{\sigma_{d,i}^2}} \tag{3.1}$$

where d_obs is the observed data, d_{fwr} is the forward data, σ_d is the uncertainty of the observed data and N is the total number of data point. A data residual below 1 indicates that our final model response is within one standard deviation of the data, when a value above 1 indicates a response out of one standard deviation. Figure 3.5b show, a single data curve (error bars) and the forward response (line) from the resistivity model in figure 3.5a. The associated data-misfit for this model is 0.27. The data misfit for the all smooth inversion model is plotted in Figure 3.6. As seen in figure 3.6, the data fit is in general well below one and

fully acceptable. 95% of the data is withing 1 standard deviation, with a global misfit average of 0.65 and 0.52 for respectively the sharp and smooth model.

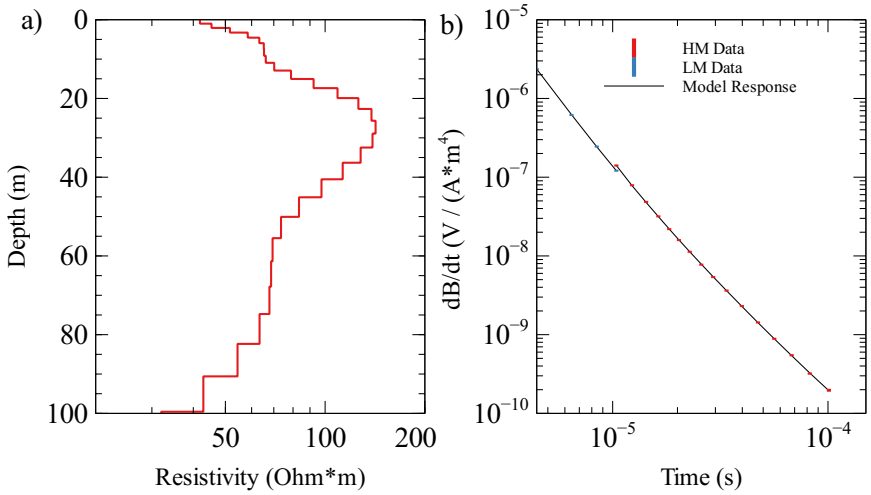

Figure 3.5: Exemple of one 1D smooth model at a location. Number 20350 at position 384744.1875/5196856 UTM 32 N. a) Smooth resistivity model b) its associated forward response in black line, with the LM & HM data point with red & blue errorbars. The normalized datafit (see text) for this model/data curve is 0.27.

A manual inspection of the high data misfit models revealed that they are all associated to highly resistive models, and/or are close to man made electromagnetic noise such as roads, fences, or train tracks. A good example is the extreme south of the acquisition, that is one of the most resistive areas. This situation logically leads to a lower signal to noise ratio, and due to the spatial constraints of the inversion, it will consequently leads to an higher data misfit. However, they are usually restricted to only a few local data-points, and the models are similar to neighbouring ones that has acceptable misfit. We therefore decided to keep them in the dataset.

Finally, users of the data should be aware that the footprint of the equipment is at least 9m at the surface (size of the equipment) and is increasing with depth and wave diffusion. Consequently, a sharp vertical transition in the geology for example, will tend to appear oblique in the resistivity

data due to this effect. The resistivity models proposed here are only one of the realisation, yet the one that fits the best our data.

Additional information

Data availability Since the file data format is a standard ASCII file, it can be open with any program supporting xzy format. The data can be downloaded at the following url: `http://dx.doi.org/10.5281/ZENODO.4269887`. **Accession codes** All the data importation, processing and SCI inversions were done using Aarhus Workbench commercial software developed by Aarhusgeosoftware. The 1D inversion code used is AarhusInv developped by the Aarhus University Hydrogeophysics group (Auken et al. 2015; Kirkegaard et al. 2015). The AarhusInv code is free to use for research purpose.

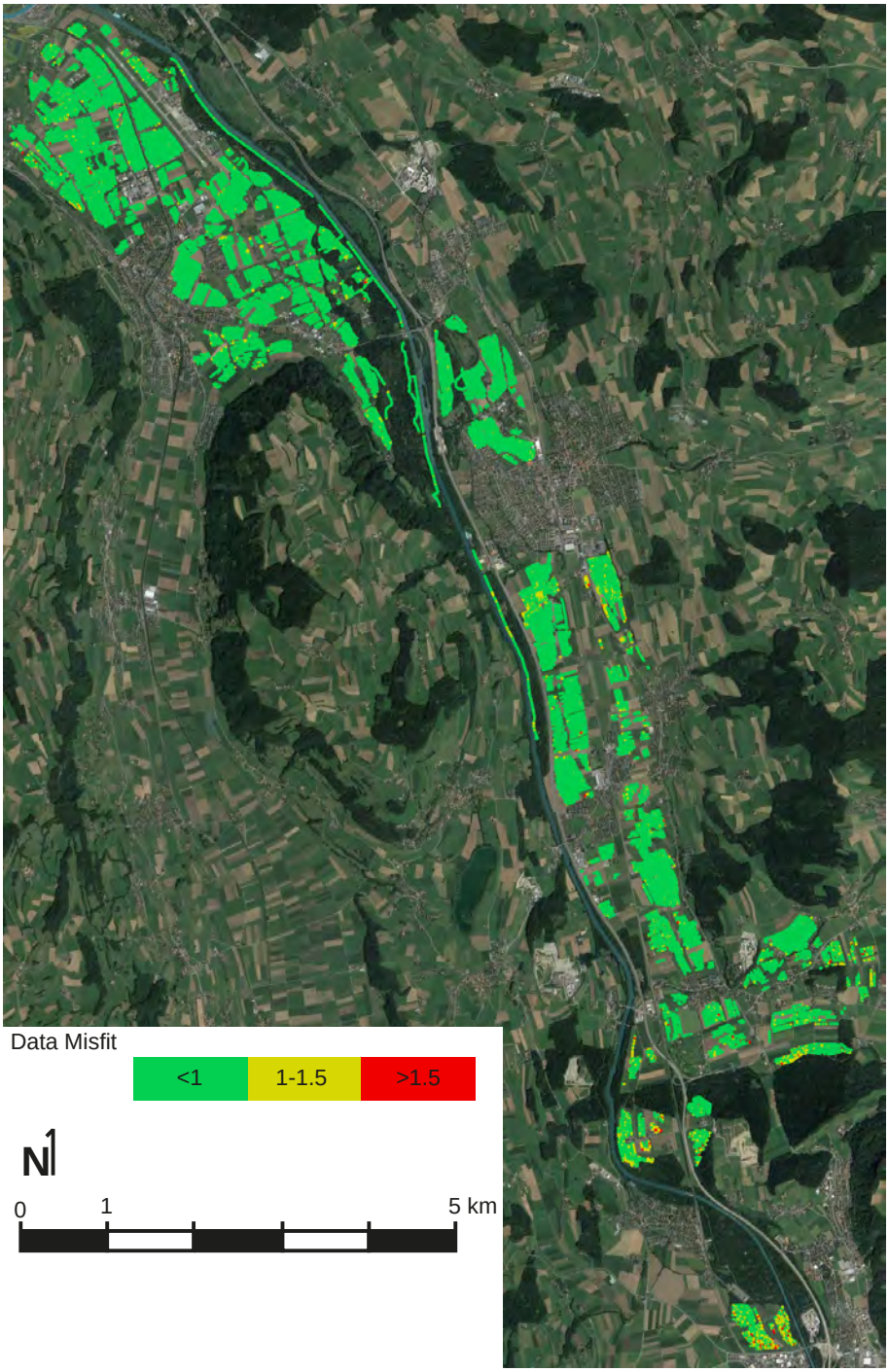

Figure 3.6: Data Misfit over the acquisition area. Base map from Swiss Federal Topographic Office.

Chapter 4

Automatic stochastic 3D clay fraction model from tTEM survey and borehole data[2]

[2]This chapter was published as : Alexis Neven, Anders Vest Christiansen, and Philippe Renard (Oct. 2022a). "Automatic stochastic 3D clay fraction model from tTEM sur-vey and borehole data". In: *Scientific Reports* 12.1.

Abstract

In most urbanized and agricultural areas of central Europe, the shallow underground is constituted of Quaternary deposits which are often the most extensively used layers (water pumping, shallow geothermic, material excavation). All these deposits are often complexly intertwined, leading to high spatial variability and high complexity. Geophysical data can be a fast and reliable source of information about the underground. Still, the integration of these data can be time-consuming, it lacks realistic interpolation in a full 3D space, and the final uncertainty is often not represented. In this study, we propose a new methodology to combine boreholes and geophysical data with uncertainty in an automatic framework. A spatially varying translator function that predicts the clay fraction from resistivity is inverted using boreholes description as control points. It is combined with a 3D stochastic interpolation framework based on a Multiple Points Statistics algorithm and Gaussian Random Function. This novel workflow allows incorporating robustly the data and their uncertainty and requires less user intervention than the already existing workflows. The methodology is illustrated for ground-based towed transient electromagnetic data (tTEM) and borehole data from the upper Aare valley, Switzerland. In this location, a 3D realistic high spatial resolution model of clay fraction was obtained over the whole valley. The very dense data set allowed to demonstrate the quality of the predicted values and their corresponding uncertainties using cross-validation.

4.1 Introduction

Shallow alluvial Quaternary aquifers are frequently used for water supply or shallow geothermal energy exploitation. In this context, a wide range of associated questions often needs to be addressed, such as evaluating groundwater resources, studying contaminant migration, evaluating interactions with surface water, or avoiding an overlap of influence zones between neighboring geothermal wells. All of these questions can only be properly answered after modeling the structure and internal heterogeneities of those aquifers.

These models are often constructed in several steps (Pyrcz et al. 2014; Ringrose et al. 2015). For small-scale models, the use of boreholes as the only source of data is common. However, such an approach is often blind to most of the spatial heterogeneity and thereby may lead to in-

adequate models and wrong conclusions. Boreholes are only one source of information to infer the local and vertical distribution of the facies. They are often of limited help to estimate complex 3D structures. When the area of interest is wide, increasing the number of boreholes to reduce the uncertainty to an acceptable level is often difficult, time-consuming, and expensive. A solution is to combine less expensive geophysical data and borehole data. Geophysical data are often interpreted and combined manually into a structural model, which is then filled with lithological or facies simulations and finally with physical properties. Such a workflow has proven to be efficient in generating geological models from aerial electromagnetic or seismic data for example (Jørgensen et al. 2015, 2013). But these modeling steps are complex and often need different software. Furthermore, even if some stochastic methods are often used (Jørgensen et al. 2015; Wellmann et al. 2017), the uncertainties are not always propagated through the complete workflow. Often some steps are considered deterministic and the resulting geological model is the one that fits most of the comprehensive knowledge available (Henriksen et al. 2003; Kollet et al. 2006; Lemieux et al. 2008; Sophocleous et al. 1999). Finally, when working with a large dataset, the manual construction of the structural model using both boreholes and geophysical data can be time-consuming. Therefore, there is a need for an automatic approach that would be able to integrate multiple data types and produce structural or parametric models. For example, the fast generation of 3D clay models with an automatic algorithm could be of great benefit to local authorities who often do not have the capacity to run manual integration of the data.

However, in most cases, the main data available along boreholes are lithological descriptions and not resistivity or density values. But these physical parameters are the ones infusing the geophysical data. Linking directly resistivities and lithologies is difficult due to the wide variety of factors affecting resistivity (Linde et al. 2006), and the often incomplete description of the lithological facies. The most fundamental relationship is Archie's Law (Archie 1942), linking resistivity to saturation, water conductivity, tortuosity, and porosity. This empirical relationship is based on the assumption that the matrix is non-conductive, an assumption not valid as soon as we have the presence of clay minerals conducting current at their surface. This pore-surface conductivity will depend on the surface area, grain size, clay type, and clay content. Estimation of all these parameters that can have a high spatial variability is tricky. A recent review (Knight et al. 2021) points out scale issues as well. Most

of the laboratory empiric laws are measured at a core scale, where the sample is in the range of a dozen of cubic centimeters. And, the upscaling of such laws to the field scale is not straightforward. Finally, most of the lithological descriptions associated with boreholes are qualitative and not quantitative (Knight et al. 2021). Also, it is way too simplistic to apply a function that would link the lithological description to a single resistivity. Each lithological description can be associated with a wide range of resistivity with some overlap between different lithologies (Knight et al. 2018).

An answer to that issue is to define the probability of having a given lithological facies conditioned on the resistivity value. These probability distributions can be estimated from a sampling of boreholes and coinciding resistivity models (Knight et al. 2018). However, such an approach does not take into account the spatial distribution of the boreholes and ignores the possible spatial dependence between the type of sediment and its resistivity. The Probability Distribution Function (PDF) is calculated from boreholes and resistivity models all over the domain. If a given lithological facies is always more resistive in a subarea of the domain, the PDF calculated over the whole domain will not reflect this.

To solve that issue, instead of trying to estimate a single PDF, Foged et al. (2014) proposed a method based on the inversion of a spatially varying translator function between resistivity and CF. The function gives the best fit between the observed CF in boreholes and the CF computed from the resistivity models. This method has the advantage of not relying on any prior parameter estimation and is only inferred from observed data. It also has the advantage of transforming the lithological description to a continuous variable making the upscaling possible, while taking the co-location of the function into account. However, even if this method can estimate the CF at the position where geophysical resistivity models exist, it still needs to be interpolated to obtain a full 3D continuous model.

To this extent, after predicting the CF, Vilhelmsen et al. (2019) used a geostatistical approach. Various geostatistical methods exist (Chilès et al. 2009; De Marsily et al. 2005; Matheron 1963) and are capable of interpolating data and producing realistic models and simulations. Their use is widespread in disciplines such as risk assessment, resources management, mining, petroleum engineering, or geological modeling. Multiple

Point Statistics (Mariethoz et al. 2014; Strebelle 2002) (MPS) is one of these geostatistical techniques. It is a non-parametric method that relies on the use of a Training Image (TI) to infer the spatial variability of one or multiple variables. MPS has proven to be capable of generating realistic complex spatial variability in a broad of situations (Carvalho et al. 2017; Dall'Alba et al. 2020; Neven et al. 2021a; Pirot et al. 2014; Strebelle et al. 2002). Vilhelmsen et al. (2019) clustered the CF in units and used it as TI in an MPS procedure. To increase variability in the simulation, they only considered the Hard Data as being the zones where the cluster is the most certain, mainly at really low and high resistivity. The other areas are only constrained by soft data. This allows one to reflect one type of uncertainty on the CF data and to show variations in uncertain model zones. But this method requires choosing the discrete threshold value at which the belonging to a cluster becomes certain. In addition, it does not allow the propagation of uncertainty from the data. The uncertainty on resistivity or on CF is lost after the clustering. Therefore, a value that would be close to the cluster limits but associated with a large uncertainty will be considered as certain as a value that would enter a cluster beyond any doubt. Moreover, due to the shape of the translator function, small changes in resistivity value in the transition zone between clay and non-clay can have huge impacts on the estimated CF. Finally, since the uncertainty of the data is used as normalization in the objective function of the CF inversion, we argue that it should also be considered when applying the function and interpolating the results. This shows the limitation of using a deterministic TI when the uncertainty should be taken into account.

In this paper, we propose an extended workflow to automatically generate a 3D clay fraction model, with a robust uncertainty propagation from the data to the final model. The paper is structured as follows. We first present the three main steps of the methodology: the CF deterministic inversion, the stochastic interpolation framework, and finally the cross-validation implementation. Then, we present the application of the methodology for a ground-based towed Transient Electromagnetic (tTEM) geophysical dataset (Neven et al. 2021b) acquired in the upper Aare valley, Switzerland.

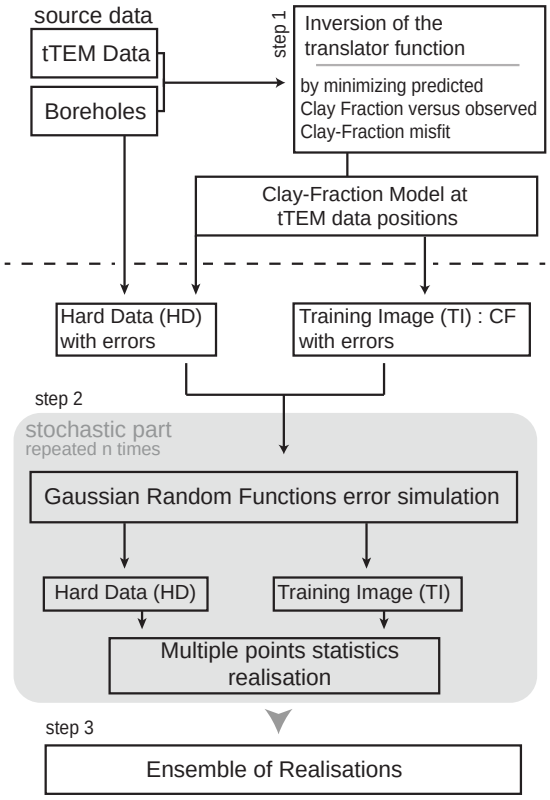

Figure 4.1: Overview of the mains steps of the proposed methodology. The shaded area highlights the stochastic part of the workflow.

4.2 The automated workflow

The proposed methodology allows generating automatically a 3D CF model of a complete Quaternary valley. A key aspect of this approach is to rely on the data itself to infer most of the parameters automatically. The following sub-sections of this paper will describe in detail all the steps. But first, a quick overview of the complete workflow is presented in Fig. 4.1.

The inputs are, on the one hand, a dense data set of resistivity logs obtained by inverting the geophysical measurements (here tTEM data) and, on the other hand, a much sparser data set of geological logs which can be used to estimate the reference CF along the boreholes. To produce a 3D CF model over the whole domain, we first invert a CF Translator

Function (Step 1 in Fig. 4.1) using the method of Foged et al. (2014) and both data sets. The translator function is then used for estimating the CF from the resistivity for all the available geophysical soundings (including locations where no borehole data is available). The main novelty of the proposed workflow is the methodology employed to interpolate the resulting CF values in 3D (Step 2 in Fig. 4.1). As compared to previous works, there are two major improvements. One is to avoid having to classify the data; the CF values are used directly. The second, and maybe most important, is that we consider the various sources of uncertainties and propagate them in the complete workflow. For this, the interpolation is done with a Multiple Point Statistics algorithm, using the CF data themselves in 3D as a Training Image (TI). The dense geophysical coverage permits using directly the data without having to add external information. To account for uncertainties in both the boreholes and the tTEM data, we use Gaussian Random Functions to simulate error maps, that are included in the conditioning data and TI images prior to the MPS simulations. Finally (Step 3 in Fig. 4.1), a cross-validation is performed to check the quality of the results. A part of the conditioning data are not used during the interpolation and compared with the results. We then assess how the method performs both in terms of predicted value and predicted uncertainty.

Clay fraction estimation

This part of our methodology (Step 1 in Fig. 4.1) follows closely the work of Christiansen et al. (2014) extended by Foged et al. (2014). We apply their idea of inverting a function that fills the gap between resistivity and CF. Since the method has been described in detail in these previous publications, only a brief summary is presented here. This approach was first developed to assess nitrate contamination risk (Christiansen et al. 2014) and called Accumulated Clay Thickness (ACT). Later, an extension was proposed to estimate the CF with the same methodology (Foged et al. 2014). The main assumption is that the major changes in resistivity in saturated and unconsolidated deposits are caused by variations in the amount of clay. With this in mind, a translator function linking the resistivity and the CF can be constructed.

Although, a few challenges need to be discussed. First, the resistivity models derived from the geophysical data are always a smoothed version of the reality, because of the footprint of the instrument, the resolution

capability, and the inversion procedures. Second, the composition of the clay sediment itself can change and cause different resistivity values in different places with similar lithological descriptions. This implies that a translator function based using only resistivity as an input cannot represent properly the link between resistivity and hydraulic conductivity.

Foged et al. (2014) therefore proposed a procedure based on the inversion of two parameters of a translator function that predicts a CF from a resistivity value and allows these parameters to vary in space. In other words, two identical resistivity values that are far from each other will not necessarily correspond to the same CF value. The basic input data, in this procedure, are boreholes geological logs as illustrated in Fig 4.2b and geophysical resistivity profiles at the same locations such as the one shown as a red curve in Fig. 4.2d. The two data sets have been acquired using independent methods and are co-located.

Based on the geological description, the clay fraction CF_{log} along the borehole is estimated by dividing the geological log into regular depth intervals and by computing the proportion of lithological clay for each interval. Because the geological description is often qualitative, the resulting CF is uncertain and the result is a mean value and range of CFs for each interval. This result is illustrated as gray boxes in Fig. 4.2c.

The next step is to define a parametrized translator function:

$$W(\rho) = 0.5 \cdot \text{erfc}\left[\text{erfc}^{-1}(0.05) \cdot \frac{2\rho - m_{up} - m_{low}}{m_{up} - m_{low}}\right] \tag{4.1}$$

where *erfc* is the complementary error function, ρ is the resistivity, and m_{up} and m_{low} are the resistivity values at which the function returns a weight of 0.975 and 0.025 respectively. The parameters m_{up} and m_{low} are identified during the inversion and they can be thought of as the resistivity limits for only clay / only sand. Fig. 4.2a illustrates the inversion of m_{up} and m_{low} for one data point. Before the inversion, the predicted CF_{rho} is in disagreement with the observed amount of clay CF_{log} described in the collocated borehole, resulting in a misfit. The inversion will adjust m_{up} and m_{low} to reduce the misfit.

The previous example shows the principle for one single data point, but the problem is correlated in 3D and aims at minimizing a global misfit. Two neighboring translator functions cannot have drastically different parameters. Fig. 4.2d shows for example how the translator function will

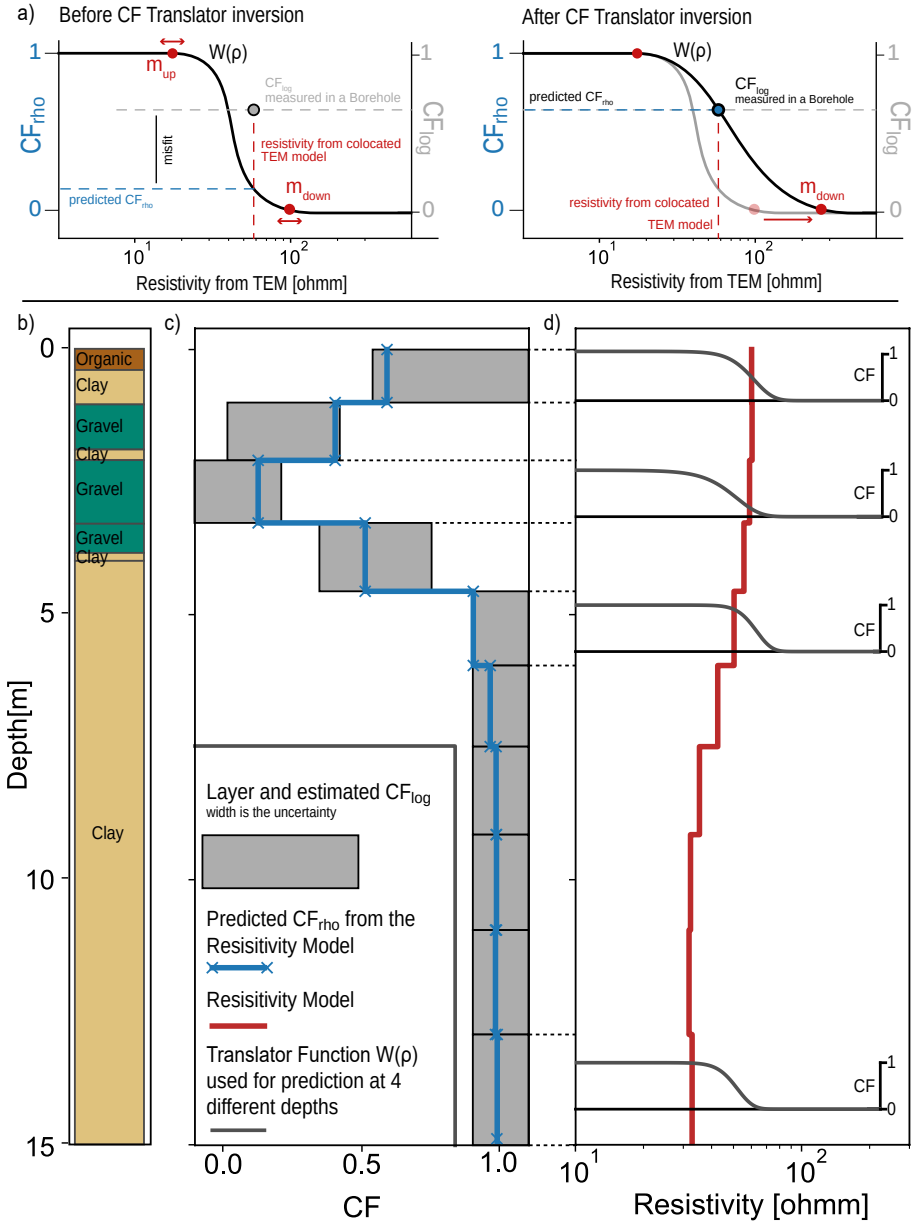

Figure 4.2: a) Illustration of the automatic fit of the translator function on one data point. b) Example of a USCS principal component lithological log. c) Associated estimated CF in a regular layered grid (CF_{log}), derived from the lithological description of the borehole. The line is the predicted CF using the resistivity models and the inverted translator functions. (CF_{rho}) d) Resistivity model from the smooth inversion of the tTEM data at the borehole position. For four layers, the inverted translator function used is displayed.

vary with depth for one single borehole. The function is outlined for four different depths in Fig. 4.2d, but is defined for all layers.

To obtain this type of result, the inversion procedure follows the following steps. First, the global data residual is defined as the average squared error between the predicted CF (CF^{rho}) and the measured CF in the boreholes (CF^{log}):

$$R_{dat} = \frac{1}{N} \cdot \sum_{i=1}^{N} \frac{(CF_i^{rho} - CF_i^{log})^2}{\sigma_i^2} \tag{4.2}$$

The data residual is normalized by σ^2 the combined variance of CF^{rho} and CF^{log}, and N the number of conditioning data. Regularization parameters are then added to ensure that neighboring points do not show sharp and unrealistic variations. The regularization will limit the spatial variation of m_{up} and m_{low}. The regularization is expressed as:

$$R_{con} = \sqrt{\frac{1}{N_{con}} \cdot \sum_{i=1}^{N_{con}} \frac{A_i^2}{ln(e_i)^2}} \tag{4.3}$$

where N_{con} is the number of pairs of parameters, A is the difference in log space between the two pairs of parameters investigated, and e_i is a distance-dependent factor. The further away two points are, the larger can their difference be. Finally, the complete objective function is:

$$Q = \sqrt{\frac{N \cdot R_{dat} + N_{con} \cdot R_{con}}{N + N_{con}}} \tag{4.4}$$

The optimal parameters (m_{up} and m_{low}) are obtained by minimizing Q using an iterative Gauss-Newton scheme with a Marquardt modification (Christiansen et al. 2014). The fit between the predicted CF_{rho} and CF_{log} is not perfect, since the translator function parameters are affected not only by the co-located data (borehole-resistivity models) but also by all the neighboring ones (via the regularization). The inversion looks for a global minimum.

In the example shown in 4.2, we can denote how the translator function adapts to identify structures that are not well represented in the resistivity model, such as the shallow clay layer.

At the end of the clay fraction estimation step, each resistivity log is associated with an estimated CF_rho log, obtained by applying the translator fonction with the optimal parameters m_{up} and m_{low}.

The final step is to propagate the uncertainty from the resistivity values to the CF. Indeed the resistivities were obtained by a geophysical inversion which is capable of estimating the uncertainty on the resistivities. The translator functions can then be applied not only to the resistivity data but also to their uncertainties, resulting in a CF_rho at the tTEM acquisition points with uncertainty. For the next simulation step, the CF_log will be used as conditioning data for the cells in which we have borehole information. CF_{rho} will be used elsewhere.

3D stochastic interpolation

To generate a full 3D model of CF and resistivity, the values obtained in the previous step need to be interpolated to cover the space where no geophysical or borehole data has been acquired. This part of the workflow corresponds to Step 2 in Fig. 4.1 and is the main novel part of the proposed workflow. We use the Direct Sampling MPS method (Mariethoz et al. 2010a) in this process. The main advantage of using an MPS approach is that it can learn automatically the spatial patterns of the regional structures from the very dense data set constructed in the previous step and use it to represent the spatial variability and uncertainty in the interpolated areas. However, the standard application of MPS techniques assume that the TI and the Hard Data (HD) are deterministic. Here, we propose a method to go a step further and account for uncertainty in these input data. We therefore include non-deterministic Hard Data (HD) and Training Images (TI) in the MPS algorithm. The overall method, as shown in Step 2 in Fig. 4.1, will consist in applying the MPS algorithm many times with different TI and different HD to generate an ensemble of interpolated 3D models of CF. From this ensemble of simulations, we will derive probability distributions for the CF and resistivity at any location in the 3D domain.

Direct sampling algorithm

The general principle of MPS algorithms is to fill a simulation grid iteratively while reproducing the patterns of the TI. The Direct Sampling

algorithm (Mariethoz et al. 2010a) is a versatile MPS algorithm based on a randomized and conditional re-sampling of the TI. In this chapter, we use the *Deesse* implementation (Straubhaar 2019) of the Direct Sampling. To generate one simulation, the 3D grid is first filled with the Hard Data available. The algorithm then randomly visits all the remaining locations in the grid. For simulating a value at a given location, the n closest HD and already simulated values are extracted to define a data pattern. The data pattern is then used to search for locations having similar patterns in the TI. During that search, the distance between the data pattern and the patterns found in the TI is computed. If the distance is below a threshold value (t), then the two patterns are considered similar and the pixel value at the missing location is taken from the TI and copied in the simulation grid. To accelerate the algorithm and avoid copying and pasting directly the TI, only a fraction (f) of the TI is scanned. The three parameters n, t, and f are chosen in advance by the user. Meerschman et al. (2013) offer some practical recommendations for the selection of those parameters. One strength of the Direct Sampling algorithm is that it can deal with continuous and multiple variable simultaneously. It means that one can provide a TI containing several variables, and the algorithm can simulate one or several variables conditioned to the other data available. This is described in detail in several papers (Mariethoz et al. 2012, 2010a). It can be used to describe the presence of trends in the training image and in the interpolation grid using auxiliary variables (Dall'Alba et al. 2020; Pirot et al. 2014).

To interpolate the CF and resistivity, we used DeeSse and designed a four variables simulation problem: two auxiliary variables and two main variables. Each of them has a specific number of neighbors n and threshold t, while the scanned fraction f is common to all variables. The two secondary variables are (1) the depth of the cell and (2) the northing of the cell. Their purpose is to describe our prior knowledge of spatial trends. This choice reflects our expectation to find certain patterns more preferably at a given depth or area in our interpolated domain. For example, the deepest structures in the TI will tend to be reproduced deepest in the simulation. However, the threshold values that we selected are high, and n is small constraining only midely the algorithm. We also chose to include only northing, since we expect to have most of the variations of the pattern along the valley (oriented roughly N-S). This assumption is based on a visual inspection of the resistivity models, geological knowledge, and boreholes. The main simulation variables are (3) the log-

arithmic transform of the resistivity and (4) the clay fraction. For these variables, we chose a higher number of neighbors n and a smaller threshold t as these are the variables of interest. All four of these variables are simulated. We also activate the Gaussian Pyramid option of the DeeSse code (Straubhaar 2019). By doing so, the spatial patterns in the data are analyzed and modeled at multiple scales. The algorithm uses Gaussian filters to create a pyramid of co-located coarser scale images that are used jointly for training and simulation (Straubhaar et al. 2020). This method improves the quality of the simulation between densely covered areas and sparser ones while being able to be sensitive to different scales of variation.

TI and HD Generation

To run the MPS algorithm and interpolate the CF, we need to provide a 3D TI and HD. The HD are simply the 3D punctual data derived from Step 1 in the boreholes and geophysical soundings. Because the spatial density of the tTEM data is very high (see the example application below), this data set is often very dense and it can be used directly as a TI. This situation corresponds to the so-called gap-filling problem, in which we need to interpolate only some parts of an already very dense data set. MPS and, in particular, direct sampling have proven to be very efficient for these problems (Mariethoz et al. 2012; Oriani et al. 2016). In these cases, the same data are used as HD and TI. The assumption behind this modeling decision is that the coverage of the data is sufficient to represent properly the spatial statistics of the variables that need to be interpolated.

In practice, the 3D HD points are placed in the 3D simulation grid upfront. These points are estimations of CF and resistivity values coming either from the boreholes or from the application of the translator function on the tTEM data. The auxiliary variables (depth and northing) are also computed and stored for every location in the grid. At this point of the methodology, we could simply apply the MPS algorithm and obtain an ensemble of stochastic simulations representing the uncertainty due to the interpolation. This is what is normally done when using MPS. The TI and the HD are deterministic.

But the CF and the resistivity values have an associated uncertainty that has already been estimated in the previous step of the workflow. To

account also for this source of uncertainty, a possibility would be to consider only the most certain points as HD, for example, the extremely low or high resistivities, which are almost certainly associated with entirely clay or non-clay points, respectively. But this idea has two main disadvantages. First, a choice has to be made to determine the upper and lower resistivity boundaries. This would go against the idea of implementing the most automated procedure possible. Second, the probability distribution of the values in the TI must be similar to the HD distribution. If they are different, the simulations may tend to over-represent the clay and non-clay points in the simulation. More generally, we would not be using all available HD, and we could assign a resistivity or a CF to a cell completely outside the uncertainty range derived from the field measurements. We, therefore, need a better way to account for the uncertainty on these data.

We overcome these challenges by a combined use of MPS and Gaussian Random Function (GRF) which allows us to perform MPS with a training image (TI) and a hard data set (HD) that are not deterministic. GRF models are well known (Chilès et al. 2009; Dietrich et al. 1993). The spatial variability is modeled using parametric multi-Gaussian distributions. These models are defined with a covariance or variogram model representing the spatial variability. Multiple realizations can be generated, with or without conditioning data. In the proposed methodology, the GRF model is used to represent the CF and resistivity data measurements uncertainty. In practice, for each MPS simulation of CF and resistivity, a different TI and HD are generated.

Figure 4.3 sketches the general principle of the generation of the TIs. The original HD (and TI) are slightly perturbed by adding some noise within the range of estimated uncertainty for these data points. The simulation of the noise is made using unconditional 3D GRF simulations. The GRF simulations have a mean of 0 and a variance of 1. To account for the possible spatial correlation of the noise a 3D anisotropic variogram is automatically fitted to the TI data. We use a Trust Region Reflective algorithm (Branch et al. 1999) to optimize the sill and the multi-direction ranges of two contributions (a Gaussian and an exponential). The sill is then normalized to obtain a variance of 1. The values of the CF and resistivity for the 3D TI (and 3D HD) at each iteration will then be defined by adding the HD with a rescaled noise:

$$\mathrm{TI}_i(x, y, z) = \mathrm{data}(x, y, z) + \sigma(x, y, z) \cdot \mathrm{GRF}_i(x, y, z) \qquad (4.5)$$

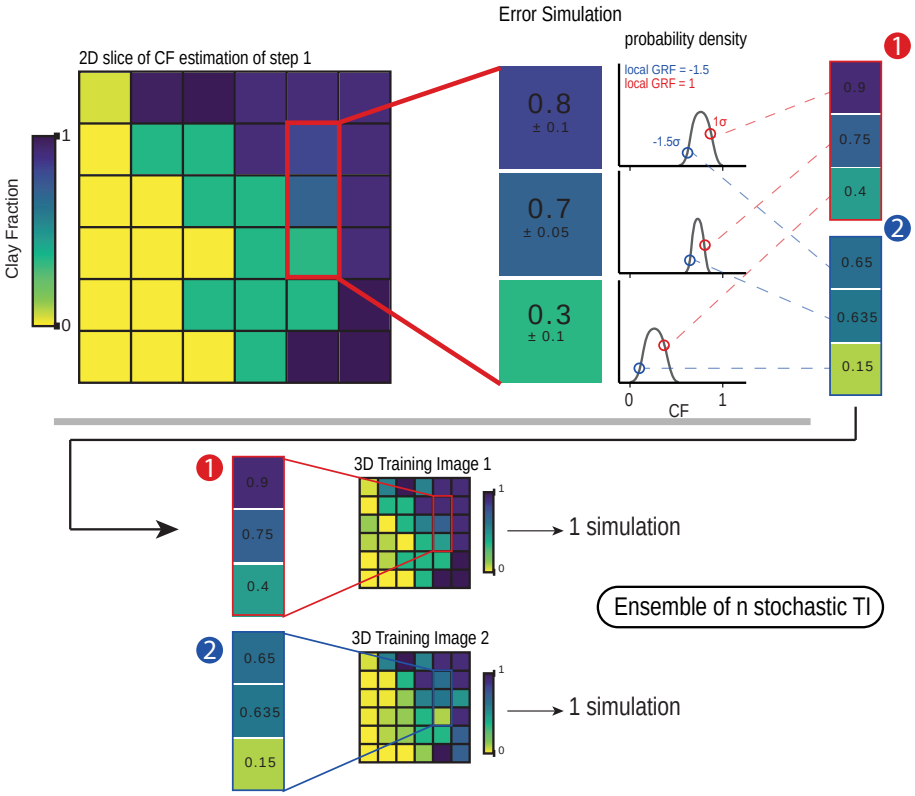

Figure 4.3: Principle for the generation of an ensemble of Hard Data and Training Images to account for the uncertainty on the HD.

with $\mathrm{TI}_i(x, y, z)$ being the TI value for the simulation i at position (x, y, z) in 3D, $\mathrm{data}(x, y, z)$ being the deterministic estimated value of the variable of interest at that location, $\sigma(x, y, z)$ the standard-deviation of the estimated uncertainty and $GRF_i(x, y, z)$ being the unconditional simulated GRF value at position (x, y, z) for iteration i. The final simulated value (CF for example) varies within the range of its uncertainty. In the example presented in Figure 4.3, the initial TI derived from the estimated CF is slightly modified by adding a correlated random noise that depends on the local uncertainties. The same operation is performed for the geophysical and borehole data sets using the same GRF model.

To summarize, one unique 3D TI and its corresponding 3D HD set are generated for each simulation and given as input to the Direct Sampling algorithm which will simulate the missing parts in the 3D grid. This will

result in an ensemble of 3D realizations. We can then stack all the simulations and calculate the mean and standard deviation for any location. The cells in which we have conditioning data will have the same mean and standard deviation as the original conditioning data.

Cross-Validation

To check the performance of the proposed methodology, we implemented a cross-validation step. Cross-Validation allows for quantifying the errors associated with a model. The principle is simple: a subset of data is created; it contains a random part of the original dataset. The stochastic interpolation is then applied using only the subset of data as HD, and the resulting simulation is compared with the excluded data. Various error indicators can be used (Juda et al. 2020). In our case, a random sampling of the dense CF data does not create sufficiently large gaps to produce representative error estimation. The missing points are too well constrained by the neighboring ones. To create a larger disruption in the data, we assigned each point a group, based on a 3D k-means clustering (Hartigan et al. 1979). The clustering is done using the spatial coordinates (x,y,z) of the model. The purpose of this step is to create multiple spatial groups of similar sizes that will be randomly excluded from the simulations. In our case, we defined 28 groups. The simulated value and the standard deviation of these zones are then compared to the real value. The error and the normalized error for each point in the model are defined as

$$\epsilon = \frac{1}{n}\sum_{i=1}^{n}(sim_i - true) \tag{4.6}$$

$$\epsilon_{norm} = \frac{1}{n}\sum_{i=1}^{n}\frac{\mid (sim_i - true) \mid}{\sigma} \tag{4.7}$$

with ϵ and ϵ_{norm} being respectively the error and the normalized error, n the total number of simulations, $true$ is the data point not included in the interpolated dataset, sim_i is the simulated value and $sigma$ is the standard deviation of sim_i over the n simulations that didn't include these data points. These two indicators are calculated point-wise. The normalized error is an important indicator since it shows how well we estimate the uncertainty, the objective being a ratio close to 1.

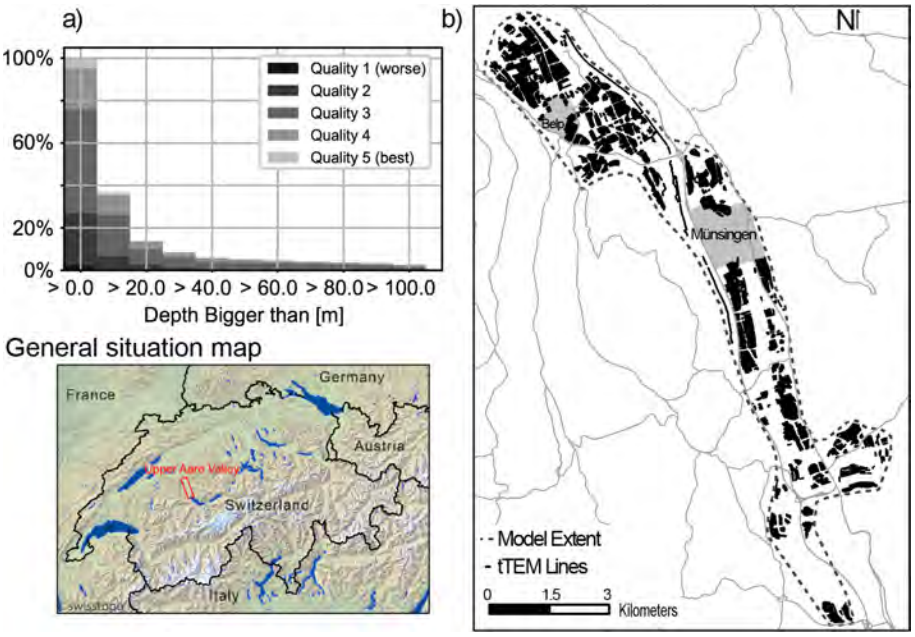

Figure 4.4: a) Distribution of layers quality and depth after normalization. b) Overview of the coverage of the tTEM data and general situation map (created with QGIS V3.22.9, qgis.org. Basemap is freely available from the Swiss Federal Office of Topography swisstopo).

4.3 Application on the Aare Valley

Geological Settings

The study area is located in central Switzerland and covers a section of approximately 20 km of the Upper Aare Valley between the cities of Thun and Bern. The Upper Aare Valley presents a typical Quaternary geology for alpine valleys: the basement is usually a few hundred meters deep and is covered with a complex interwinding of glacial, fluvio-glacial, and fluvial deposits (Kellerhals et al. 1981a; Volken et al. 2016). Multiple cycles of glacial advances and retreats have been identified in the Swiss basin, causing multiple changes in the deposition processes (Graf et al. 2016). In the particular case of the Upper Aare Valley, at least four glacial cycles have been identified in boreholes (Schlüchter 1989). However, a complete description of the lithostratigraphy based only on borehole descriptions is nearly impossible because similar deposits of different ages may be superimposed and intertwined. A surface aquifer constituted of

the most recent alluvial deposit is present all over the valley. The water table is between a few dozen centimeters up to 2 meters below the surface. General knowledge of the aquifer suggests that the thickness of the upper aquifer ranges from 4 to 20m depending on the area. In addition, a second deeper aquifer has been identified in some deep boreholes separated from the shallow one by a clay layer. Both aquifers' exact extend, connectivity, and thicknesses are not extensively known. However, in the area, several hundreds of pumping (shallow geothermal or drinking water) and injection wells (re-inject water after geothermal heat pump) are in use. In this context, the realization of a 3D model would be a great benefit for local authorities to evaluate the vulnerability of the upper aquifer. Because of the water table height, saturated conditions can be assumed in almost the entire height of the model.

Borehole Data

In the zone of interest, 1542 boreholes are lithologically described. The dataset used was provided by Swisstopo. The upper Aare Valley was also one of the test sites retained for the Geoquat project of the Swiss National Topographic Institute (Volken et al. 2016). In this context, they performed digitization and standardization of the borehole data. All boreholes descriptions were converted to standard USCS descriptions (Casagrande 1948). In addition, a QC value was added for each layer assessing the reliability of the geological information: a grade from one to five was assigned to each layer, depending on the precision of the description. one corresponds to a layer with only a basic description when five corresponds to a layer where lab measurements have been performed and a complete multi-phase lithological description is available. An estimation of the Clay Content in the boreholes was made using the USCS guidelines, and the uncertainty was scaled according to the grade. A poorly described clay layer (Quality 1) will have, for example, a CF value of 1 ± 0.5. We consider that up to 50% of the clay layer could be non-clay material. On the other hand, a well-described one (Quality 5) will have a CF value of 1 ± 0.08. Missing, artificial, or undescribed layers are not taken into account. Details of the classification procedure can be found in Volken et al. (2016).

Figure 4.4 a) shows the distribution of the quality of the layers and their associated depth. Since most of the borehole exploration is conducted for either shallow geothermal exploration or geotechnical purposes, we

observe that around 60% of the data are above 10m depth and that the proportion follows roughly a decreasing power-law distribution with depth.

tTEM data

In January 2020, a ground-based towed transient electromagnetic (tTEM) survey was conducted in the upper Aare valley. The data were inverted using different regularizations (sharp and smooth), and the resulting resistivity models are used in this study. All processing workflow, details about regularization, and data are described in Neven et al. (2021). The dataset covers about 1500 hectares, with a line spacing of 25 meters. The sampling frequency after processing is about 1 resistivity model every 10 meters along the lines. The inversion of the data was done using the *AarhusInv* (`https://hgg.au.dk/software/aarhusinv3`) inversion code (Auken et al. 2015). The average residual of the inversion is 0.52 for the sharp regularization, which means that we tend to have an excellent fit between the predicted data from our resistivity models and the field measurements. To estimate the depth of investigation, we used the Jacobian sensitivity matrix of the last iteration (Christiansen et al. 2012). By doing so, we can identify the exact depth at which each resistivity model is only poorly represented in the recorded data. Below this depth, the resistivity models are blinded. Again, for a complete description of the inversion of the data set and the quality check, the readers are referred to Neven et al. (2021).

The translator function inversion was performed on the *tTEM20AAR* sharp inversion dataset (Neven et al. 2021b). 57'862 30 layers resistivity models were taken into account and blinded at the standard depth of investigation. Figure 4.4 b) shows the coverage of geophysical data in the valley after processing. Most of the fields outside the cities are mapped. Cities, as well as the areas surrounding roads and train line, are clearly visible since the electromagnetic coupling in such environments forbids the use of an inductive method and are left uncovered. The uncertainty on the resistivity was estimated using the last iteration covariance matrix (Alumbaugh et al. 2000). Even if this method does not fully replace the uncertainty that can be derived from a stochastic inversion, it has the advantage of being able to reflect the relative uncertainty on the model's cells while being relatively fast to calculate.

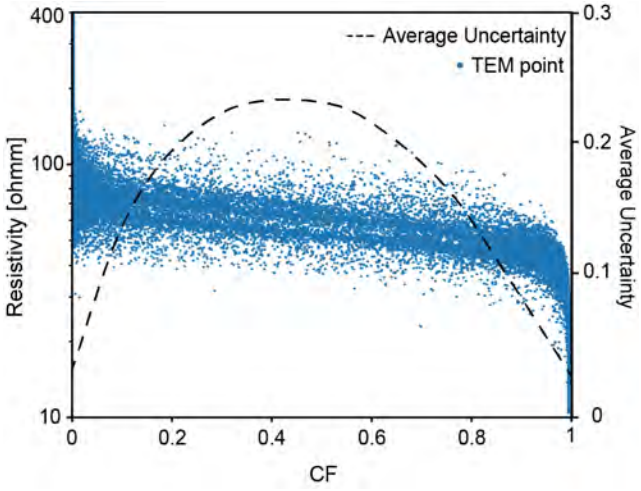

Figure 4.5: Results of the inverted translator function on 50'000 TEM points. In addition, the averaged uncertainty calculated with a moving mean of 0.05 CF width.

Clay fraction inversion

The average data-misfit (see Eqn.4.2), or residual, of the application of the translator function to the geophysical models is 0.38. It means that the difference between the predicted CF_{rho} and CF_{log} is more than two times within the uncertainty range. The main source of misfit is the resolution issue. Because of the footprint of the geophysical equipment and of the damped least-squares inversion performed on the TEM data, we tend to have a smoother transition in the geophysical data than the reality. Since we are comparing the results with borehole data, which tend to reflect these sharp transitions, they can be a source of misfit between the predicted CF_{rho} and the actual CF_{log}. Some thin layers of resistive materials in a large conductive layer will not be caught by the resolution of the equipment but will increase the residual value. However, the main advantage of the (Foged et al. 2014) methodology is that the translator function adapts, to counterbalance these effects. Figure 4.5 shows the results of the inverted translator function and the averaged uncertainty on 50'000 TEM points. Most of the uncertainty is concentrated on the central points that present an intermediate resistivity since small uncertainty on the resistivity in this area will affect strongly the predicted CF. The range of values in which the transition between clay to non-clay highest probability happens is between 90 and 40 Ohms. Such values are

coherent with the expected manual interpretation since Aare Quaternary Clay's have a resistivity between 10-50 Ohms. The models are blinded at the standard depth of investigation obtained from the resistivity models and are then passed to the geostatistical simulation step.

3D Clay Model

The results presented here are based on 200 simulation loops of 10 realizations each. For each loop, a new TI and HD set are generated with a GRF simulation, and a different random subset is drawn for the cross-validation. The groups were constituted using a K-Means algorithm of the spatial coordinates of the CF data. As mentioned before, the *Deesse* algorithm is controlled by three main parameters: the scanned fraction, the threshold, and the number of neighbors. The TI scan fraction is common to all variables and was set to 5%. The pattern will be compared to the TI until the scanned fraction is reached or until the threshold is respected on all variables. The threshold is set to 10% on three neighboring nodes for the auxiliary variable and 1% on 24 nodes for the main variables. If the threshold value is never reached, the best candidate is selected and the point is flagged to be re-simulated at the end of the simulation. The scanning path in the TI is random. The result is a set of 2000 simulations, based on 200 different TI. The resulting 3D model has a cell dimension of 50m by 50m, and a vertical resolution of 2m. The surface covered is 35 km^2, for a computational time of 10 hours on a CPU cluster. Vertical cross sections of the averaged model are displayed in Fig. 4.6. A clear trend in the model is present between the northern and the southern side of the model. Such variation was expected due to the variations in the shape of the valley and is corroborated by the boreholes. The uncertainty of the data is higher in the deeper cells of the model, where no boreholes and no geophysical data are present to constrain the simulations. They also denote the transition zones between clay and non-clay areas and reflect the uncertainty on the exact depth of transition. The shape of the underground structures is consistent with existing Quaternary deposition conceptual processes. However, the visual consistency is not sufficient to trust a model.

The error and the normalized error distributions (See Eqn. 4.6 and 4.7) calculated with cross-validation are displayed on Fig. 4.7. During the simulation, the two main variables are the resistivity and the clay fraction. The error estimation is performed during the cross-validation step

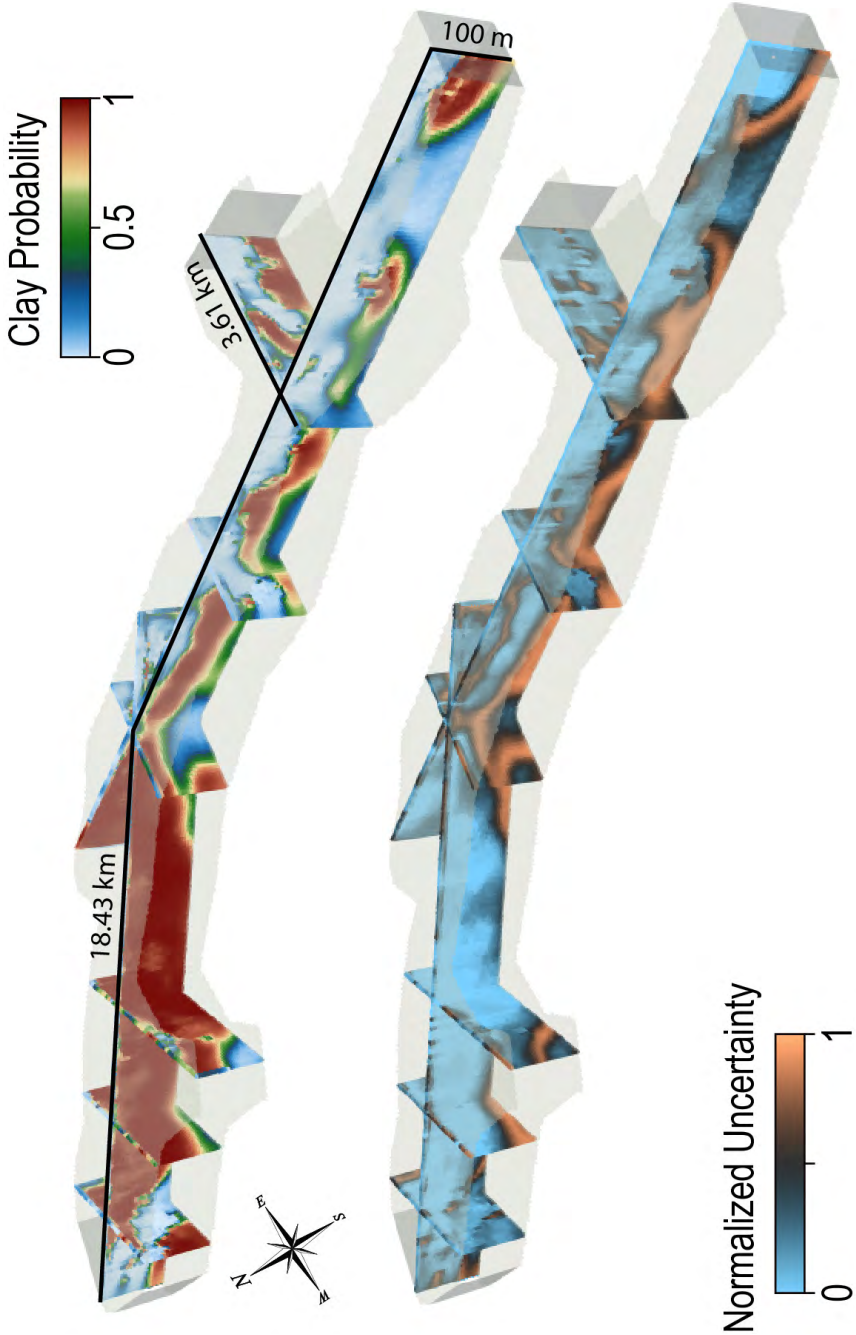

Figure 4.6: 3D view of slices through the clay fraction model, with the associated uncertainty. Z-scale is exaggerated.

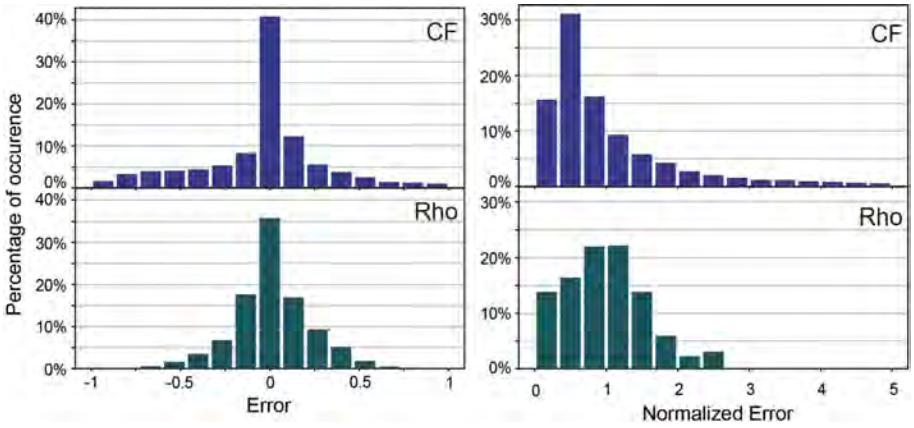

Figure 4.7: Histogram of the errors and the normalized error for the two main variables : Resistivity (Rho) and clay fraction (CF).

on both. The error is centered on zero for the two variables, suggesting that we tend to predict a correct value with no bias on average. Furthermore, in respectively 44 % and 37 % of the cases, we managed to predict a value that was in the 10% range of the real value. In terms of normalized error, the mean of the normalized error for resistivity is 0.98 ($\sigma = 0.56$). Such a distribution suggests that the standard deviations of the simulations are in the same order of magnitude as the real error. The uncertainty is therefore well predicted. On the other hand, the mean of the CF normalized error is 0.89 ($\sigma = 0.5$), meaning that we tend to slightly overestimate the error in the simulated data compared to the real one. But overall, the cross-validation indicates that on average the correct value is simulated with a standard deviation reflecting well the possible uncertainty. A strong result is that the resistivity is predicted accurately, showing that this method can be of great use even to only interpolate resistivity maps.

Comparison with Existing Model

In an effort to homogenize and digitize all geological data from the Quaternary formations in Switzerland, the *GeoQuat* project made a prototype and demonstration study on the Aare Valley. In this context, a deterministic geological model of the area was done using a manual interpretation of boreholes, of geophysical data, and geological knowledge through the use of geological cross sections (Volken et al. 2016). When

comparing the two models, we should keep in mind that new data are included in the stochastic model that was not available at the time of the construction of the deterministic model: even if the boreholes database is the same, the tTEM data were only acquired in 2019. After incorporating the boreholes, they manually correlated units and facies and used the nearest neighbor interpolation to extend the model to a full 3D volume (Volken et al. 2016).

Fig. 4.8 shows the comparison between a few cross-sections in the deterministic model and in the model generated with MPS. The deterministic geological model and the boreholes are displayed using their USCS primary components. The transparency of the MPS model reflects the uncertainty. Higher is the transparency, higher is the associated uncertainty. Both models are in agreement with the borehole data at least for large-scale structures. We can see that the MPS-generated geological model displays structures that are more realistic geologically speaking. Some comments can be made for the cross-sections:

- Sections 1 are in agreement. The model is well constrained by two deep boreholes, going all the way through the model. Even if the nearest neighbor interpolation of the deterministic model displays some unrealistic blocky shapes, the global structures are similar between the two models.
- Sections 2 are drastically different between the deterministic and the stochastic model. The absence of boreholes to constrain the deterministic model causes the propagation of a gravel layer from distant boreholes, that is most probably not present at this location. The inverted resistivity calculated from the tTEM data is low and is most probably associated with thick clay bodies. Such a section shows improvement in the models, because of the additional data.
- Section 3 illustrate the importance of uncertainty. No tTEM data has been acquired in the area, and both models rely on the same data. The upper few meters of the models are well identified as being the upper gravel body on both models. However, the use of the proposed method revealed a possible resistive (sand or gravel) body deeper. The uncertainty is high, as highlighted by the color range. The presence of this layer cannot be established for sure.

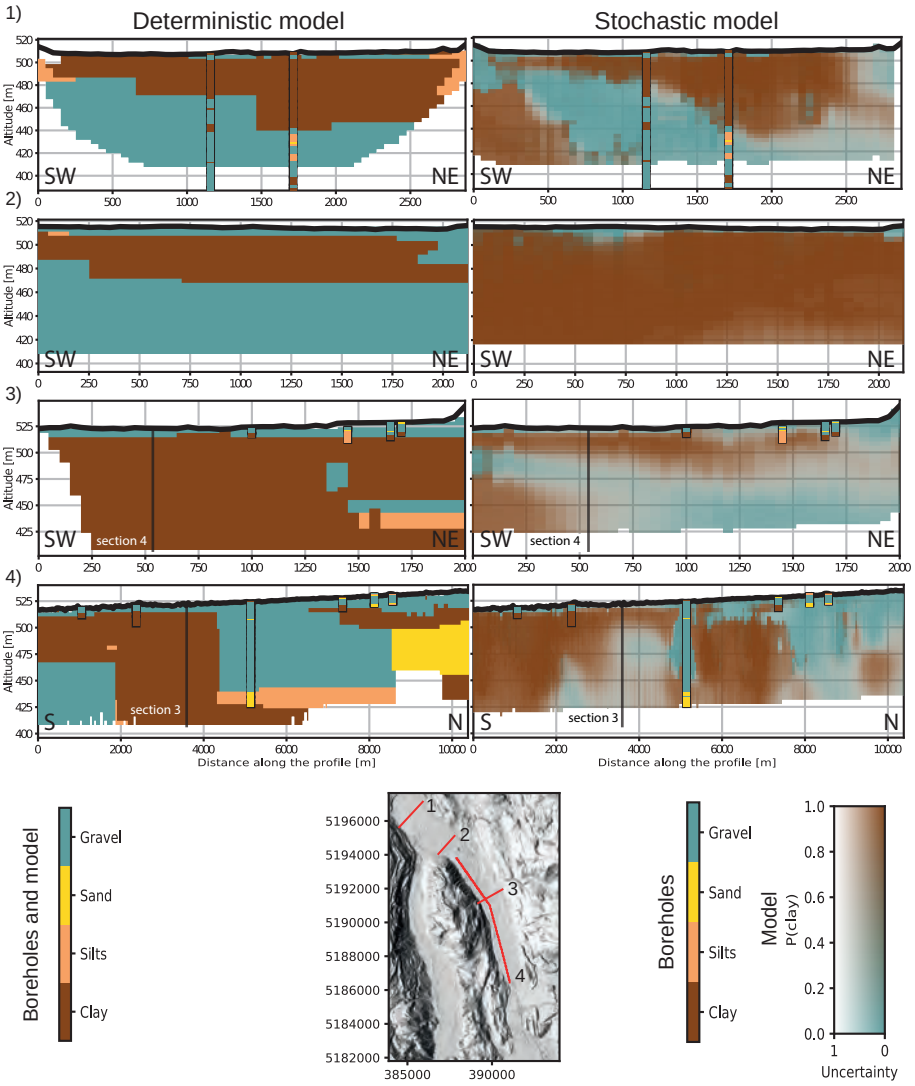

Figure 4.8: Comparison between existing geological model and this study model. The position of the sections 3&4 intersection is outlined.

Although it is likely a possibility that has to be considered when assessing the area.

- Sections 4 show how a local Quaternary structure can drastically affect the deterministic model. The 100m deep borehole in the center of the cross-section indicates a gravel layer that is abnormally thick compared to the other nearby boreholes in the section. This over-thickening is extremely local and is most probably due to a gravity mass movement a few dozen thousand years ago according to personal communication with a Quaternary geologist. Because of this special borehole, the deterministic geological model overestimated the thickness of the upper aquifer over a very wide area.

Finally, we could test the quality of the model by comparing the geology observed in several newly made boreholes that were not taken into account in any of the two models because they were drilled recently. The new shallow boreholes (max 10m) agree well with both models at 85% and 94% respectively for the deterministic and stochastic models. These results are not surprising since the upper gravel layer is more or less present in the whole domain. Most of the errors are due to some minor over- or underestimation of the depth of transition. However, a small number of deep boreholes were drilled too. In Fig. 4.9, a new borehole log is compared to the two models. The borehole is situated in the middle of the Wichtrach village, where no tTEM data can be acquired. The stochastic model reflects well the geology observed in the borehole. The stochastic model predicts the presence of the gravel aquifer on top and high resistivities like the deterministic model up to a depth of around 25m. But then the stochastic model predicts the presence of a massive clay layer up to a depth of 80 m before indicating a large uncertainty at a deeper depth. This prediction matches well the the geological observations well, which also show a massive clay layer. On the contrary, the deterministic model predicted a massive gravel layer just below a thin clay layer, followed by silts. This is not what has been observed in the borehole. Both models are constrained by the same boreholes data around, but while the stochastic model infers the regional trend by integrating all data of the area, the deterministic model is only constrained by nearby points and cross sections.

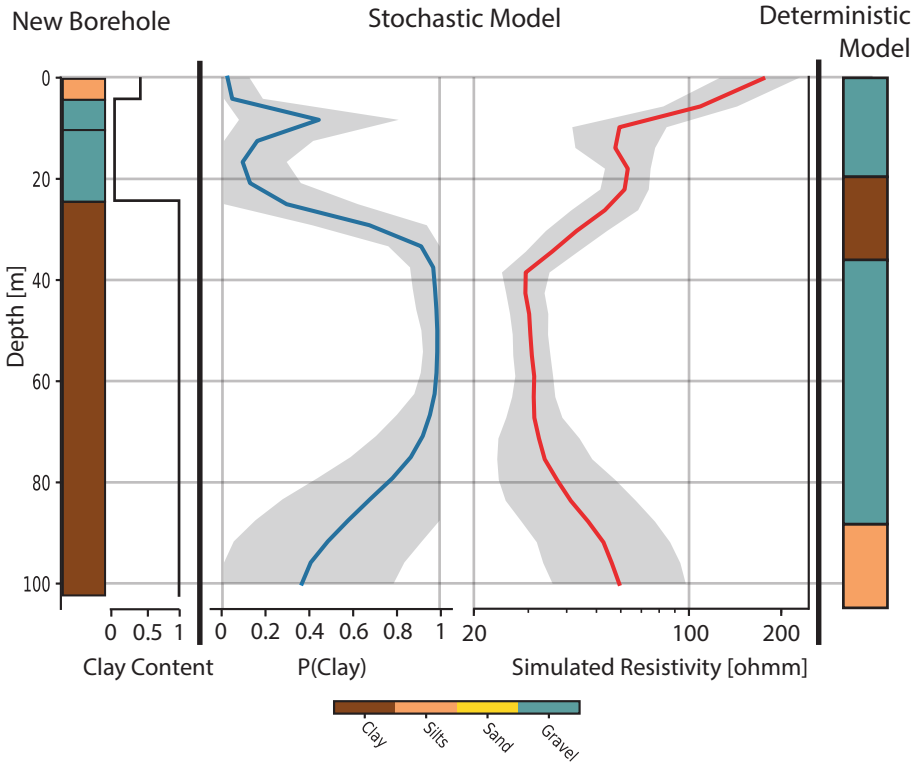

Figure 4.9: Comparison between a new borehole in the village of Wichtrach (not included in any of the two models), and the corresponding cell in the stochastic model and the deterministic geological model

4.4 Discussion

In this study, we presented a methodology that permits propagating the uncertainty all the way from the geophysical data and borehole logs to the final model of the clay fraction. In addition, all input data are used as hard data within their uncertainty range. A noisy data point in the field will lead to a more uncertain cell in the model. This feature is a key point to obtain a robust error estimation on the final model. Furthermore, being able to integrate uncertain boreholes due to poor or absent descriptions, unsure positioning, or missing data is a great advantage. The use of MPS makes it possible to generate models presenting complex patterns that are impossible to reproduce using approaches such as kriging. This method also has the advantage of relying on the data itself to deduce spatial patterns without any prior knowledge. Of course, this is

only applicable to regions where the data coverage is sufficient. However, in a region with sparser data coverage or to increase the quality of the model, prior knowledge can also be added through the use of auxiliary variables. After the data preparation, the generation of high-resolution valley scale models can be achieved in a few days compared to months for deterministic ones.

In terms of geophysical technique, the application of tTEM in the upper Aare valley has proven that the technique is reliable and capable of providing relevant information regarding the 3D structure of the underground. This is confirming the previous conclusion from Sandersen et al. (2021). But, in addition, we argue that a proper and efficient way to integrate these types of large 3D datasets with borehole data can only be achieved through automatic methods. We also expect that the proposed methodology could be very useful for the integration of dense airborne data sets.

However, the estimation of lithology from resistivity has also its limitations. First of all, if the area presents significant salinity variations, it may introduce nonuniqueness. The same resistivity can correspond to saturated resistive material or clay. We would then only rely on boreholes to identify the nature of the underground. The proposed method would need to be extended to include a 3D estimation of the salinity as an additional auxiliary variable. This may be possible using additional hydrogeological information. However, in the particular case of the upper Aare Valley, nothing suggests important salinity variations. Furthermore, the proposed method is efficient to distinguish between resistive layers (sand or gravel) and conductive layers (clay), but the resistivity difference between sand and gravel does not allow the method to distinguish them clearly. That is the main reason why we decided to focus on the identification of the clay layers only.

Even without simulating the CF value, our approach should also be considered when interpolating 1D or 2D resistivity models to 3D, especially when dealing with sharp inversion. Traditionally, this is done using Kriging. But kriging being the best unbiased linear estimator, it leads to a smooth interpolation between sharp models and inevitably adds artifacts in the 3D resistivity models. In the context of Quaternary deposits, the range of the variogram can sometimes be smaller or equivalent to the distance between two acquisition areas, leading to a wrong or incom-

plete interpolation. The proposed method has the advantage of relying on the 1D soundings directly instead of a grid made by kriging and it will interpolate between the models with a coherent sharpness. A similar methodology proposed by Vilhelmsen et al. (2019) relies on an indicator kriging on the cluster value with a short range to fill the gaps between the acquisition lines in the TI, causing possible over smoothing. The methodology presented here has the advantage of not relying on any pre-interpolation of the data before the simulation.

Of course, the inversion itself is already an interpretation of the data, and in addition, the uncertainty associated with the geophysical inverted models is often difficult to estimate exactly. A further improvement could be the development of an algorithm that would work on the geophysical measurements directly, without any inversion, instead of using inverted models. By doing so, we could be able to compare directly the errors between the real field data and the final CF model. It would also simplify the workflow by avoiding a double inversion, once for the resistivity model itself and once for the CF algorithm.

The comparison between the existing model and the automatic model (Fig. 4.8) shows the risk of using a deterministic approach. On cross-section 2, the deterministic model predicts a thick gravel body at about 40m depth. This body was placed there only relying on boreholes that are hundreds of meters away (close to Cross-section 1), and after including geophysical data we are now confident that there is no such body there. Of course, it was an interpretation based on the data available at that time, but no indication of uncertainty was available. In this context, taking decisions based on such kind of model is risky and should be avoided.

An increasing number of countries are developing centralized databases to host geological-related measurements. In this context, the use of an agile and reliable data aggregation algorithm is a promising approach. Being able to deduce spatial patterns from data without or with little prior knowledge prevents the infusion of structures from arbitrary choices, which is usually done during a manual interpretation of geophysical models. Standardization of description methods through different geological consulting companies and open-access data are the key points making this method application fast and more adaptable to new field areas.

An improvement of the method could be to forecast the permeability values. If permeability measurements are conducted, either by using packing or pumping tests, a new variable could be added to the simulation to directly generate permeability fields. Another possibility would be the definition of a coupled probability density function between resistivity, clay fraction, and permeability from a set of prior knowledge or field knowledge. We would then be able to directly generate parametric fields. Finally, as mentioned earlier, we think that the interpolation methodology introduced here could be applied to other 3D geophysical models, in order to fill partial maps. Often, due to field constrain (inaccessible areas, corrupted data, different acquisitions,...) some areas are less densely covered than others. Using non-deterministic MPS, with Gaussian Pyramids, complete and homogenous parametric models could be generated with proper uncertainty quantification.

4.5 Conclusion

In this paper, we showed that a novel workflow combining the clay fraction estimation algorithm (Foged et al. 2014) with our modified Multiple Points Statistic algorithm is a robust method for the automatic generation of a 3D clay fraction model. The resulting 3D models can be used by local authorities or project managers to better plan the development and underground use of the region. These models can be used for example to locate potentially highly permeable zones for groundwater exploitation, or geothermal developments, or to evaluate the potential presence of construction materials in the underground. Because the method provides uncertainty estimates it can also help design the acquisition of further data.

Our method has the advantage of being data-driven and not relying on manual interpretation of the structures. The workflow is automatic and includes: 1) the inversion of the translator function on the resistivity models and the borehole data, 2) the automated fit of the variogram used for the GRF model, and 3) the generation of different TI and HD for each MPS simulation. The user has only to select a few parameters such as the number n of neighbors, or the threshold t for the MPS simulations, but default values can be used. Furthermore, it reproduces structures that cannot be modeled using classical two-point interpolation such as kriging. By comparing the geology observed along newly made boreholes with our model, this study always predicts well the general trend as

shown for example in Fig. 4.9. The workflow can easily be incorporated with public databases, allowing the authorities to update automatically the 3D model regularly when new field data become available.

Finally, the paper also proves the efficiency of the towed Transient ElectroMagnetic (TEM) method. When most of the previous publications used airborne TEM data, we integrated towed TEM data. This allowed us to reach a resolution in space and depth that is unachievable with any other geophysical method at the moment. Such data acquisition should be increasingly considered, even for medium or small-scale problems. Easy, fast, and cheap integration of multiple, often existing, data types can only be beneficial whatever the purpose.

Additional information

Data availability The tTEM data used in this paper are available from the open access archive Zenodo at the following URL: `http://dx.doi.org/10.5281/ZENODO.4269887`. The data set is documented in Neven et al. (2021). The borehole data set is available from the Swiss Federal Office of Topography (Swisstopo) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from Swistopo upon reasonable request.

Accession codes The geostatistical codes used in this study are implemented in the python *Geone* package freely available at `https://github.com/randlab/geone`. To use all the functions of DeeSse, an academic license can be freely obtained. The CF model will be freely available.

Chapter 5

Stochastic multi-fidelity joint hydrogeophysical inversion of consistent geological models[3]

[3]This chapter was published as : Alexis Neven, Ludovic Schorpp, and Philippe Renard (Oct. 2022b). "Stochastic multi-fidelity joint hydrogeophysical inversion of consistent geological models". In: *Frontiers in Water* 4.

Abstract

In Quaternary deposits, the characterization of subsurface heterogeneity and its associated uncertainty is critical when dealing with groundwater resource management. The combination of different data types through joint inversion has proven to be an effective way to reduce final model uncertainty. Moreover, it allows the final model to be in agreement with a wider spectrum of data available on site. However, integrating them stochastically through an inversion is very time-consuming and resource expensive, due to the important number of forward simulations needed. The use of multi-fidelity models, by combining low-fidelity inexpensive and less accurate models with high-fidelity expensive and accurate models, allows one to reduce the time needed for inversion to converge. This multiscale logic can be applied for the generation of Quaternary models. Most Quaternary sedimentological models can be considered as geological units (large scale), populated with facies (medium scale), and finally completed by physical parameters (small scale). In this paper, both approaches are combined. A simple and fast time-domain EM 1D geophysical direct problem is used to first constrain a simplified stochastic geologically consistent model, where each stratigraphic unit is considered homogeneous in terms of facies and parameters. The ensemble smoother with multiple data assimilation (ES-MDA) algorithm allows generating an ensemble of plausible subsurface realizations. Fast identification of the large-scale structures is the main point of this step. Once plausible unit models are generated, high-fidelity transient groundwater flow models are incorporated. The low-fidelity models are populated stochastically with heterogeneous facies and their associated parameter distribution. ES-MDA is also used for this task by directly inferring the property values (hydraulic conductivity and resistivity) from the generated model. To preserve consistency, geophysical and hydrogeological data are inverted jointly. This workflow ensures that the models are geologically consistent and are therefore less subject to artifacts due to localized poor-quality data. It is able to robustly estimate the associated uncertainty with the final model. Finally, due to the simplification of both the direct problem and the geology during the low-fidelity part of the inversion, it greatly reduces the time required to converge to an ensemble of complex models while preserving consistency.

5.1 Introduction

Quaternary aquifers are frequently used for groundwater supply, but due to their high heterogeneity, they are difficult to characterize and model. A possible strategy is to combine geological knowledge, geophysical data, and hydraulic tests. Geophysical methods, especially electromagnetic (EM) ones, are inexpensive and can efficiently help when facing an under-constrained problem (Barfod et al. 2018; Christensen et al. 2017). They are usually sensitive to several petrophysical parameters, such as resistivity, but provide limited information regarding the hydraulic conductivities or the water storage capacity of the underground. Consequently, inverted geophysical models can be integrated into the first steps of the aquifer modeling workflow by delineating the geological structures from a manual interpretation and obtaining a so-called cognitive model (e.g. Høyer et al. 2015). This approach can be combined with stochastic models to populate the main stratigraphic units with lithologies and represent that level of heterogeneity. Furthermore, one can use geophysical inversion results and borehole data to estimate the probability of occurrence of several lithologies (e.g. clay or sand) and use these probabilities as soft information to generate stochastic realizations that are both constrained by some geological reasoning, borehole, and geophysical data (Carle et al. 2020; Jørgensen et al. 2015). This approach ensures consistency with available knowledge, it reproduces accurately the soft information in terms of probability, but nothing ensures that if the final models were used in a forward geophysical model they would reproduce the field measurements. Furthermore, it is likely that the overall uncertainty may be underestimated because it is rare that the aquifer geometry derived from the cognitive model is assumed uncertain.

To ensure consistency, we propose to reverse the methodology described in the previous paragraph and start by constructing a prior geological model that we will then use in geophysical and hydrogeological inversion. Therefore, the first key ingredient of our proposed methodology is the ArchPy hierarchical modeling approach developed recently (Schorpp et al. 2022). ArchPy decomposes the construction of the aquifer model in three main simulation steps: the stratigraphic units, the litho-facies, and the petrophysical parameters. The approach is automated and accounts for a geological concept described in a data structure called a stratigraphic pile as well as borehole data. ArchPy can quickly generate an ensemble of models compatible with the prior geological knowledge of the site. Each model can serve as input to any forward geophysical or hy-

drogeological models. Comparing the results of the calculated responses to data measured in the field allows one to estimate the likelihood of any proposed model and identify which one corresponds to the maximum a posteriori likelihood and generate realizations representing the posterior uncertainty. This type of Bayesian strategy combining geophysical and hydrogeological data has shown that it can produce consistent models and allowed to reduce the final uncertainty (e.g. Irving et al. 2010; Jardani et al. 2013). But these joint inverse problems are often solved using Markov chain Monte Carlo methods and are computationally challenging (Linde et al. 2016).

Ensemble Smoothers with Multiple Data Assimilation (ES-MDA) have shown to obtain solutions to complex nonlinear inverse problems more efficiently than Ensemble Smoothers (ES) (Emerick et al. 2013) and faster than Markov Chain Monte-Carlo (MCMC) methods (Juda et al. 2022). The method uses a Monte-Carlo approximation of the Kalman Filter (Kalman 1960) where the relations between the state variables and the parameters are estimated using an ensemble of models. ES-MDA and PESTPP-IES (White 2018) are both variants of the Ensemble Smoothers (ES) algorithm proposed by Leeuwen et al. (1996). The key aspect of ES-MDA is to perform iterative ES corrections of the parameters by assimilating the data of the previous iteration, when PESTPP-IES optimizes directly an objective function using a modified form of the Levenberg-Marquardt algorithm. Lam et al. (2020) showed a comparison of different Iterative Ensemble Smoothers, including PESTPP-IES and ES-MDA. It was shown that PESTPP-IES approach outperforms ES-MDA when the ensemble size is relatively small (200 in the study) but that ES-MDA tends to improve with an increase of the ensemble size, while PESTPP-IES does not. ES-MDA has been successfully applied in groundwater studies (Kang et al. 2019; Lam et al. 2020a; Xu et al. 2022). One important underlying assumption is that the state variables and parameters are normally distributed (and even multi-Gaussian). If not, one can apply a normal score transform to ensure that the marginal distributions are Gaussian (Zhou et al. 2011).

A recent study by Wang et al. (2022) proposed a hierarchical inversion, where the posterior distribution of global variables including for example hyper-parameters of the geostatistical models are first estimated using a machine learning approach. Following this, an ES algorithm that includes local reduction of dimension is applied to invert the field parameters.

Even if ES-MDA is known to be faster than the MCMC approaches, it can still be computationally heavy because a proper estimation of the covariance matrices used to estimate the Kalman gains requires running a large ensemble of models and, therefore, running a large ensemble of forward geophysical or hydrogeological models. Often, this step is the one that requires most of the computational power. Therefore, there is still a need to devise techniques to accelerate such approaches, and one is to use a multi-fidelity framework. By using a simple surrogate model, one can approximate the forward model and accelerate the inversion (Asher et al. 2015; Dagasan et al. 2020). This idea has been applied, for example, by Zheng et al. (2019), who trained a Gaussian process model to approximate the forward flow problem in a multi-fidelity ES-MDA algorithm.

In this paper, we employ the multi-fidelity principle and ES-MDA but with a different perspective. We do not try to build a surrogate model of the forward model, but instead, we use the fact that the computing times for the geophysical and hydrogeological forward models are very different and that ArchPy provides a hierarchy of levels of representation of the geological heterogeneity. In practice, we invert jointly the multiple data types in the same workflow using ES-MDA in two steps. To ensure that our aquifer models are geologically consistent, they are generated using ArchPy. To accelerate the inversion, we run first a fast low-fidelity ES-MDA inversion to obtain an initial representation of the main geological discontinuities with the fast geophysical forward only. The complete heterogeneous models are then generated from the low-fidelity ones, and used in a second ES-MDA high-fidelity inversion loop including both the geophysical and hydrogeological forward models. The present paper introduces this idea and demonstrates its applicability to two simple 2D synthetic cases of increasing complexity.

5.2 Methodology

In this study, we propose to combine the advantages of a stochastic ES-MDA inversion with a multi-fidelity approach for joint hydrogeophysical inversion. The model is inverted for both hydraulic conductivities and electrical resistivity. In this section, we first introduce the ArchPy modeling approach and its use for our approach for multi-fidelity geological models. We then briefly introduce the ES-MDA algorithm and present the test case used to benchmark our approach.

Geological model generation

The ArchPy approach

The first key tool in our methodology is the stochastic hierarchical geological modeling approach named ArchPy and proposed by Schorpp et al. (2022). The method is implemented in an open source python module and is capable of producing both Low-Fidelity (LFM) and High-Fidelity models (HFM). These models are consistent with the prior geological knowledge and capable of integrating geological information in a hierarchical manner. For a complete description of ArchPy capabilities and algorithms, the readers are referred to Schorpp et al. (2022) and to the repository of the code[4].

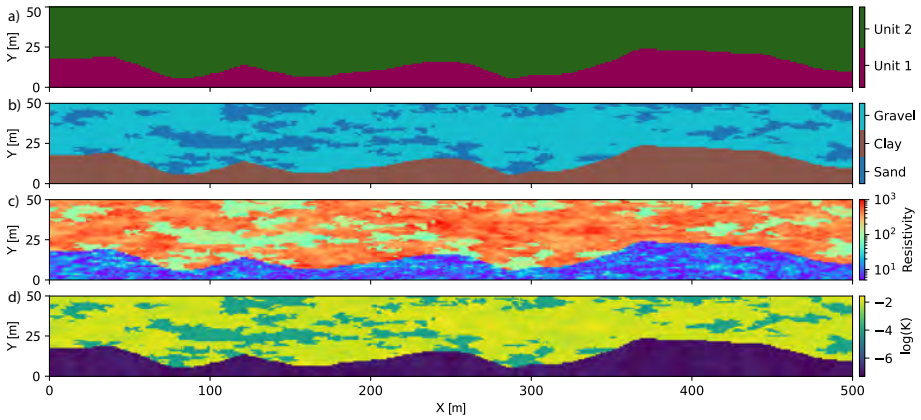

Figure 5.1: Two dimensional example of an ArchPy simulation with the 3 simulation steps: Units, Facies, and two petrophysical properties.

In short, ArchPy relies on the concept of Stratigraphic Pile (SP) which is used to formalize the existing geological knowledge for a given site. All the rules, information, and parameters required to generate the geological models are stored in the SP. For example, the SP contains the list of the stratigraphic units that must be simulated, the list of litho-facies to simulate in which units, and the different simulation parameters (covariance functions, Training images for MPS, etc.). The SP is defined by the user and represents his prior knowledge. Once the SP, is defined ArchPy constructs automatically the models in three main steps:

4 `http://www.github.com/randlab/ArchPy`

- The stratigraphic units are first simulated using 2D simulation of the surfaces bounding the units (Figure 5.1a). The user can select different geostatistical algorithms such as Multiple-point statistics (MPS) (Mariethoz et al. 2010a) or Sequential Gaussian Simulations (SGS) (Deutsch et al. 1992). This step handles erosional events where older and previously simulated surfaces are partially eroded by younger ones. It also allows some units not to be deposited (i.e. hiatus). Moreover, inequality data are used to account for incomplete information provided by a borehole that did not reach a certain unit or when there is a hiatus in a stratigraphic sequence.
- The units are then filled with litho-facies (e.g. gravel, silt, clay, etc.) using 3D categorical simulations methods (Figure 5.1b). Again, the user can chose among MPS or Sequential Indicator Simulation (SIS) (Journel 1983; Journel et al. 1984). The litho-facies models can be different for every unit and the simulations are conditioned by borehole data if available.
- Finally, the facies are populated by continuous petrophysical properties such as hydraulic conductivities, porosity, or electrical resistivity using SGS (Figure 5.1c,d). The parameters are defined separately for the different litho-facies.

An important feature of ArchPy is that the different hierarchical levels only depend on the higher ones (for example, facies only depend on the stratigraphic unit simulations). This feature allows to consider the geological representation of the underground at different level of fidelity.

Finally, ArchPy allows defining stratigraphic sub-units. This option is not used in the present paper but allows simulating complex stratigraphies when needed. By operating as described above, a large number of stochastic simulations can be obtained using ArchPy. They are all conditioned by the same borehole data and geological concept (the stratigraphic pile).

Low-Fidelity Models

The proposed inversion approach is divided in two main steps: low-fidelity and high-fidelity. The objective of using computationally inexpensive low-fidelity models is to reduce the dimension of the parameter

space in which we need to solve the inverse problem with the computationally costly high-fidelity models. Low-fidelity models consist of simplified models. LFMs neglect the small and medium-scale variability of the properties by assuming that the stratigraphic units are completely homogeneous with a unique property value for each unit. The value is drawn from a given uniform distribution in each layer. So, it can vary between the models, but not within one. This step is crucial as it prevents the inversion algorithm from being over-confident, in order to mitigate our assumption that the units are homogeneous. It also allows having a more complete exploration of the parameter space and plausible realizations. The LFMs are generated using the ArchPy package. They correspond to the first hierarchical modeling step (Figure 5.1a) and the units are homogeneously filled.

An important aspect to note is that it is not the resistivity field that is inverted at this stage. The parameters of interest are the altitude of the surface(s) delineating the main stratigraphic units. By doing so, we significantly reduce the number of unknowns while still working on a simplified geologically consistent model. Moreover, LFMs are solely evaluated on the geophysical data, as the geophysical forwards are much faster than the groundwater ones (around 15 times faster).

High-fidelity models

The high-fidelity models (HFMs) are also generated with ArchPy. These models depend on the LF ones. All the surfaces obtained via the LF inversion step are used to generate the next level of hierarchical simulations (heterogeneous litho-facies and properties. The HFMs are more realistic subsurface models than LFM and are closer to the geological concept. They present heterogeneous facies distributions, with high property contrasts even within a unit (see example on Figure 5.1c and 5.1c). These complete fields will then be inverted. The number of models is not necessarily the same in the two steps (LF and HF). This allows more flexibility, as it is expected that the two problems may not require the same number of models to converge properly. Generally, we expect that we will have more HF than LF models, since the second problem is more difficult and have to integrate more data. Therefore, the same surfaces can be used multiple times and be associated to different facies and parameter distributions.

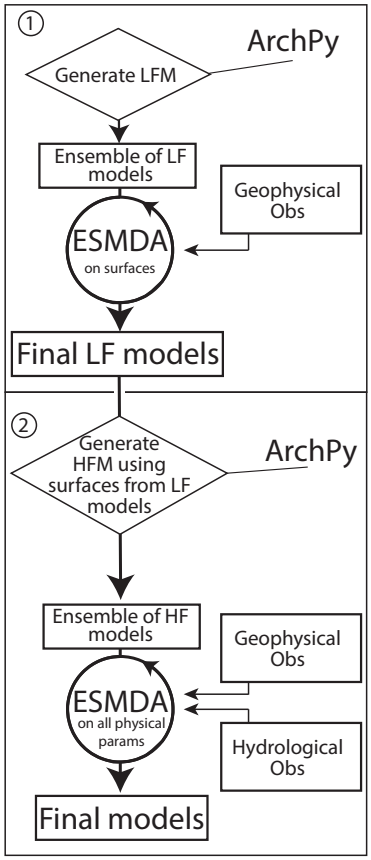

Figure 5.2: General workflow of our approach. The 2 main steps are differentiated: 1 for the low-fidelity part; 2 for the high-fidelity part.

Inversion algorithm

The general strategy for the multi-fidelity inversion is to employ the same ES-MDA algorithm (Emerick et al. 2013) successively for the low fidelity and high fidelity geological models (Figure 5.2) but on different parameters. For the low fidelity, the algorithm will update only the geometry of the surfaces bounding the stratigraphic units, while for the high fidelity models, it will update the complete model including geometry and parameter fields (resistivity and hydraulic conductivities). The two steps are coupled because the ensemble of surfaces obtained from the low fidelity inversion step are used to initialize the generation of the ensemble of high fidelity models used in the second step.

On both fidelities, the ES-MDA inversion algorithm is applied on a set of N members $\{m_i^{pr}, ..., m_N^{pr}\}$ representing the prior. To ensure that the parameters are multiGaussian, we applied a normal score transform on the prior models (Deutsch 2002, p.44-48). This step was not needed when performing the inversion on the low-fidelity models, since the surfaces are generated using Sequential Gaussian Simulation (SGS) and are by definition multiGaussian. The parameters are updated iteratively on the basis of the observations of the state variables to form a new conditional distribution of parameters. Updates are done, for each member i, according to the following equation :

$$m_i^{k+1} = m_i^k + K \cdot (d_{obs,i}^k - g(m_i^k)) \tag{5.1}$$

where k is the iteration step, K the Kalman matrix (or Kalman gain), and $d_{obs,i}^k - g(m_i^k)$ is the mismatch between the observed measurements and the predictions computed by the forward operator g using the current parameters. In order to stochastically account for the errors in the measurement, $d_{obs,i}^k$ is given by

$$d_{obs,i}^k = d_{obs} + \sqrt{\alpha_{k+1}} C_{err}^{1/2} z_{d,i} \text{ with } z_{d,i} \sim N(0,1) \tag{5.2}$$

with C_{err} being the expected error matrix and $z_{d,i}$ being drawn from a normal distribution. Compared to ES, the data will be assimilated multiple times in ES-MDA. For this reason, it will tend to overestimate the confidence given to the data Emerick et al. (2013). For this reason, a parameter $\alpha > 1$ is added to inflate the Gaussian noise. Because of that, the parameters covariance reduction will be limited at each iteration. The inflation factor needs to be fixed such as the following condition is satisfied:

$$\sum_{k=1}^{N_{iter}} \frac{1}{\alpha_k} = 1 \tag{5.3}$$

where N_{iter} is the number of iterations of ES-MDA. The alpha coefficient was kept unchanged through the iterations, equals to the number of iterations. Emerick (2016) has shown that varying them do not lead to a significant improvement in the convergence of the algorithm. The α factor is applied when calculating the Kalman gain, such as

$$K = (C_{MD}^k (C_{DD}^k + \alpha_{k+1} C_{err})^{-1}) LM \tag{5.4}$$

where C^k_{MD} is the cross-covariance matrix between the vector of parameters m and the approximated vector of predicted data $g(m)$ for the ensemble. C^k_{DD} is the autocovariance matrix of the predicted data $g(m)$. LM is the localization matrix. As shown in (Anderson 2001; Evensen 2009; Wen et al. 2005), with a limited amount of members, fortuitous long-range correlation can happen. It is necessary to filter them out before applying the Kalman update. To do so, the Kalman gain matrix is multiplied element-wise using a localization matrix (Chen et al. 2011). The purpose is to influence only the parameters located within a certain distance from the observation points during the update. The localization matrix has the same size as the Kalman matrix and its values vary between 0 and 1. It was only used in the HF step of the inversion, and was set as a uniform matrix of ones for the hydrogeological observation. We did that since the expected radius of parameters effect on a point was difficult to properly estimate. For the geophysical data, it was calculated using the correlation function proposed by Gaspari et al. (1999), with anisotropy in the distance calculation. Figure 5.3 shows an example of a localization matrix on resistivity parameters for an observation point. Further away is the observation point, weaker will be the contribution of the Kalman gain on the parameter.

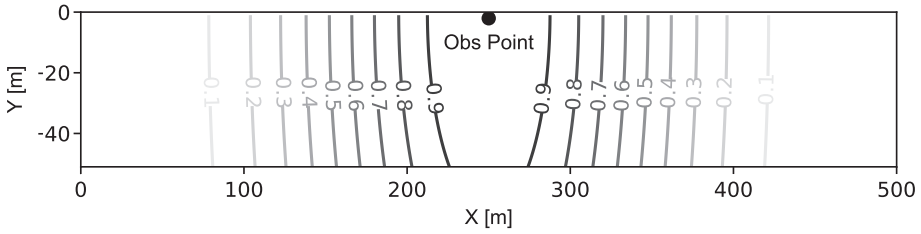

Figure 5.3: Example of localization matrix LM for one geophysical observation point using a length of 250 m.

Test Cases

Conceptual Model

To illustrate and test the proposed methodology, we only consider in this paper 2D vertical profiles. Note that the extension to 3D models is straightforward. The domain size is 500 m long for a depth of 50 m, with a cell size of 1 × 1 m. The resulting model consists of 25'000 cells.

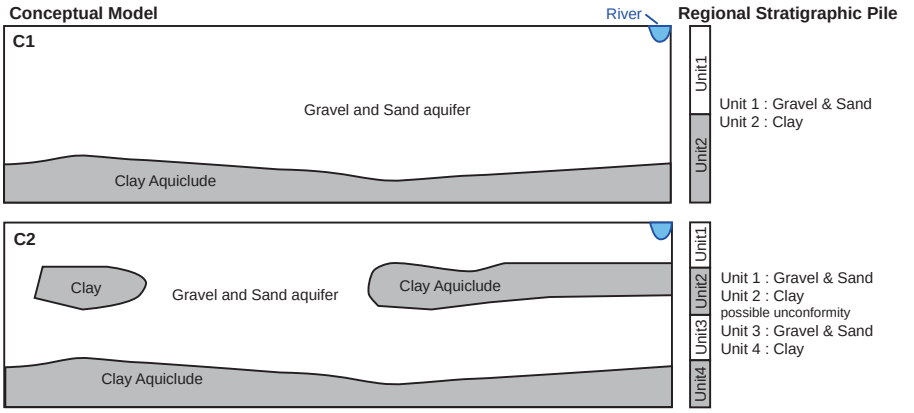

Figure 5.4: Schematic representation of the geological concepts C1 and C2, with their associated regional stratigraphic pile.

Two different cases, C1 and C2, were defined as illustrated in Figure 5.4:

- Case C1 is the simplest. It considers only two stratigraphic units: one aquifer unit (unit 1) and an aquitard (unit 2). In this situation, the most important feature to identify during the inversion process is the depth of the transition between the two units for each position x.
- The case C2 is more complex, as it includes a possibly discontinuous aquitard with a variable thickness within the main aquifer formation that can be divided into two sub-aquifers. This case can be modeled with four stratigraphic units in ArchPy (see the stratigraphic pile of C2 in Figure 5.4). The complexity of the problem is increased, as it is now also necessary to estimate, for each location x, the absence or presence of the intermediate aquiclude formation, as well as its depth and thickness in the latter case. A total of three surfaces must be inverted.

The global conceptual models are inspired from a realistic situation: the Upper Aare Valley in Switzerland. An upper fluvial deposit layer of a few dozen meters overlays a thick lacustrine clay layer. The geological context of the area is briefly described in Schorpp et al. (2022) or Graf et al. (2016). The upper fluvial deposit may show locally few superficial

clay layers. Our synthetic model is a simplified version of this real case. Concerning the geological simulations, for both cases, we assume that no borehole information is available. Thus, all geological simulations are unconditional.

For the low-fidelity models (cases 1 and 2), the surfaces are simulated with SGS and the stratigraphic unit domains are defined. Then, an electrical resistivity value is drawn uniformly between 100 and 400 Ωm in the log space for the aquifers and is taken constant and equal to 10 Ωm for the aquiclude. For cases 1 and 2, we generated an ensemble of 100 LFM to initiate the ES-MDA algorithm. For HFMs, property simulations required two additional steps. The surfaces delimitating the stratigraphic units were taken directly from the ensemble of results of the LFM inversion. The aquifer units (units 1 and 3 in Figure 5.4) were supposed to be composed only of gravel and sand with a proportion of 70 % of gravel and 30 % of sand. Sequential Indicator Simulations were used to simulate the position of these facies. The aquiclude units are assumed to be composed only of the clay facies. The different facies are then populated with the desired properties using SGS. The simulation parameters are given in Table 5.1. For both cases, we generated a total of 500 HFMs based on the 100 optimized LFMs. This implies that there are 500 members for the 2nd step ES-MDA.

Table 5.1: Geostatistical parameters used by ArchPy to generate the geological models. The variogram type used for all the litho-facies, resistivity, and hydraulic conductivities models are exponential.

Stratigraphic surface model

Unit	2	3	4
variogram type	spherical	cubic	cubic
mean [m]	-17	-20	-38
range x and y [m]	50	150	80
sill [m^2]	1	20	20

Litho-facies model

Unit	1	3
range x[m]	30	30
range z [m]	8	8
sill [m^2]	0.1	0.1

Resistivity Model

facies	mean $[\log_{10}[\Omega m]]$	range x [m]	range z [m]	sill $[\log_{10}[\Omega^2 m^2]]$
gravel	2.6	10	2	0.05
sand	2	10	2	0.05
clay	1	20	4	0.01

Hydraulic conductivity Model

facies	mean $[\log_{10}[m/s]]$	range x [m]	range z [m]	sill $[\log_{10}[m^2/s^2]]$
gravel	-2	10	2	0.05
sand	-4	20	4	0.05
clay	-7	15	15	0.01

Physical forwards

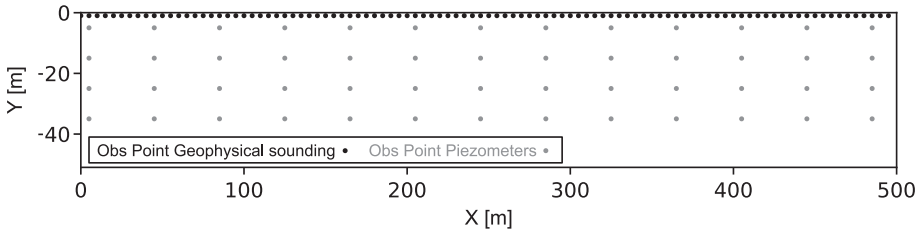

Figure 5.5: Location of the geophysical and hydrogeological observation points. The river is located at the position (500,0).

During the inversion, two distinct physical forward models are used: one geophysical forward and one hydrogeological forward. The geophysical one is a 1D Time Domain Electro-Magnetic forward. The code used is the software *AarhusInv* (Auken et al. 2015). It has the advantage of taking into account the complete geometry, waveform, and filter that correspond to the geophysical equipment of interest. We use a distinct ground-based source-receiver configuration. All the parameters of the emitters and receivers correspond to the Aarhus towed time domain EM equipment (tTEM) (Auken et al. 2019). Two distinct moments (high and low moments) are simulated in order to increase the sensitivity of the equipment to shallow variations of the underground while preserving the depth of investigation. When the propagation depth of the method goes beyond the limit of our resistivity model, the last layer is considered infinite. Gates, complete waveform, and exact geometry used can be found in Auken et al. (2019) and Neven et al. (2021). We consider one sounding every 5 meters (Figure 5.5), which corresponds to the acquisition rate of the instrument at normal driving speed.

The second physical forward is a transient groundwater model based on the MODFLOW 6 code (Hughes et al. 2017; Langevin et al. 2017) interfaced with the FloPy package (Bakker et al. 2016). We consider the propagation and attenuation of a periodic perturbation on the upper right corner of the model (Figure 5.4). The river level at his location changes daily following a sinusoidal curve. This can mimic either tidal or daily discharge variations. The river is assumed to be connected to the aquifer with a conductance of 10^{-2} $[m^2 s^{-1}]$. A second signal comes from a uniform recharge varying in time at the top of the model. We assume a no-flow boundary on the other sides of the model. Heads are recorded at

52 observation points in the model (Figure 5.5). The simulation lasts for 10 days, with a time step of 15 minutes. However, since most of the signal comes from the upper right corner, we expect a decreasing contribution of the observations to the inversion towards the left. Since the aquifer is supposed to be heterogeneous and of fluvial origin, we consider that the geological heterogeneity below the river bed can be modeled using the same parameters as the aquifer. Our synthetic model does not consider the river bed as a special compartment and does include any transient effect on the conductance of the river bed.

The computing time needed for the two forward models is drastically different. Geophysical forward takes about 2 seconds per model, when the transient hydraulic forward needs 30 seconds. The data used for the inversion are generated on a synthetic reference model and disturbed with Gaussian noise.

5.3 Results

To better understand how the proposed methodology performs, we first present the results of the intermediate step (low-fidelity step) before providing detailed results after the final high-fidelity step.

Low-Fidelity

For case C1, the number of ES-MDA iterations was set to 2 for the low fidelity step. The computing time was about 3 minutes on a personal computer. The mean residual decreased by 48% after one iteration and 62% after two iterations with respect to the unconditioned prior. Figure 5.6a shows the prior distribution for the position of the bottom of the aquifer for Model C1: it follows a normal distribution with an average depth of 38 meters and 95% (2σ) of the simulated surfaces are between -26.8m and -49.2m depth. This figure also shows the results of the inversion step: the average surface (solid black line) over the 100 members of the ensemble, the 95% confidence interval (gray shaded area), and the reference (in red). We see that even with a relatively small number of iterations and a small ensemble size, the ES-MDA algorithm converges rapidly to a plausible solution, even with the high simplification of the model. The reference surface is almost always within the predicted uncertainty range.

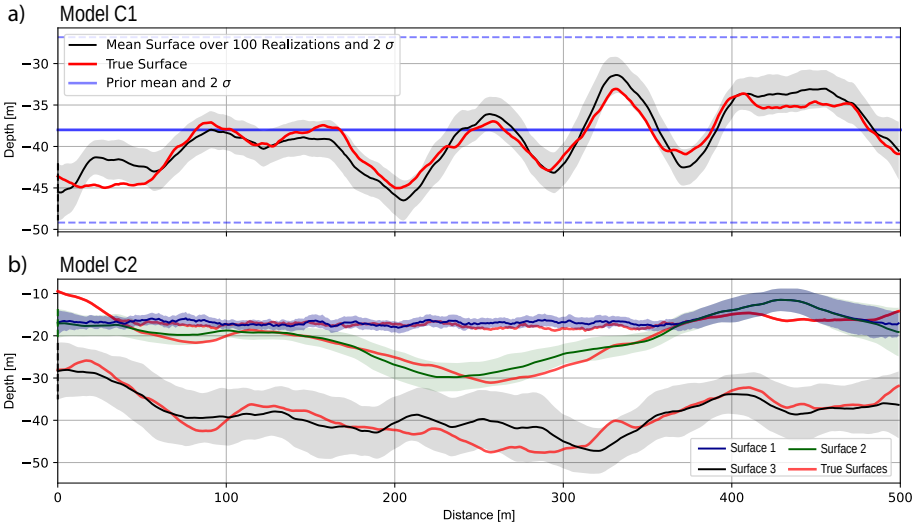

Figure 5.6: Low-Fidelity Inversion Results for Models C1 and C2 over 100 members and 2 iterations each. The model C1 has only one surface, and the model C2 has 3 surfaces. The shading areas correspond to 2 σ.

For case C2, Figure 5.6b shows the three average surfaces of 100 members. The computing time was about 3.5 min on a personal computer. In general, the inversion manages to correctly predict the depth of the transition and has identified the presence of the clay layer in the middle of the aquifer. If we compare surface 3 with its equivalent in the C1 model, we can denote that the corresponding uncertainty associated is much larger, even if the surfaces are exactly similar in terms of parameters and prior. This behavior is probably due to the presence of additional layers which imply some variations in the average resistivity above this surface that are much larger than in Case 1. Another interesting thing to denote are the fact that the surfaces 1 and 2 in the last 100m of the model are superimposed. Where the lower surface equals the upper one, this last then follows the lower surface and becomes one with it. This is a complex hierarchical principle, because a whole set of parameters suddenly have no effect on the model residual. ES-MDA as correctly identified this logic from the set of prior models. It illustrates that using ES-MDA with geological priors could help reaching complexity in the inversion not seen so far.

High-Fidelity

For cases C1 and C2, 15 iterations were performed. The total run time was about 8 hours on a personal computer, mainly due to solving the transient hydrogeological forward problem. Figure 5.7a shows the average electrical resistivity and hydraulic conductivity (Figure 5.7d) over the ensemble of models at the end of the high-fidelity step. The corresponding uncertainties are shown in Figures 5.7b and 5.7e, while the reference model is shown in Figures 5.7c and 5.7f. No weighting based on the residual is used for the calculation of the mean model. The lower clay layer shows little uncertainty for both parameters. This is probably due to its very low resistivity and hydraulic conductivity within the range of possible values, which makes them easily identifiable. Another factor is probably that this layer was already well resolved during the low-fidelity step, because of its monofacies characteristic. However, we denote an important uncertainty on the exact depth of the transition on both parameter fields. It is only slightly updated compared to the uncertainty estimated after the low-fidelity step.

The aquifer layer presents much higher variability for both parameters. First, we can denote that both fields seem spatially correlated. ES-MDA has correctly identified the correlation infused by the facies affiliation of the parameter field. More resistive zones are associated with more permeable areas, whereas less resistive zones are associated with less permeable sand. The uncertainty on the hydraulic conductivities is higher than the one associated with the resistivity, probably because the geophysical method is an active one, whereas our hydrological scenario is passive and extremely diffuse. Consequently, it is a more difficult problem to solve for the algorithm. This can also be denoted in Figures 5.7g and 5.7h.

Two High-Fidelity ES-MDA inversions were performed using one single dataset at a time. Not using the joint inversion approach has only a limited effect on the resistivity, compared to the joint inversion, even if some of the simulations can be marginally less noisy. On the other hand, inverting the hydraulic conductivities only shows a clear decrease in terms of the quality of the inversion. The mean simulation tends to converge to a hydraulic conductivities value intermediate between the sand and the gravel facies values. As mentioned above, this is expected because of the aquifer's high heterogeneity and the method's diffuse aspect. The only area clearly resolved is the upper right corner where the oscillating river limit is set. Since these cells control how the signal gets into the

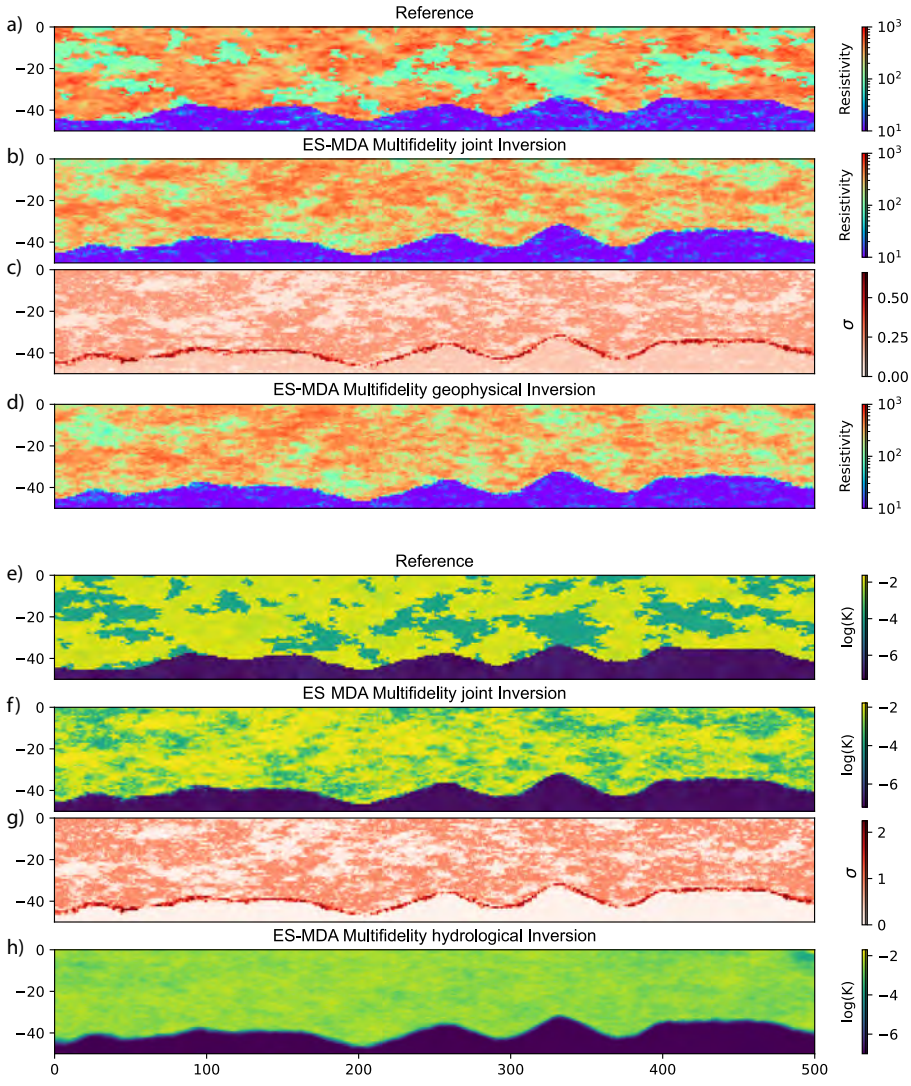

Figure 5.7: High-fidelity results for case C1. a) reference resistivity field , b) Mean model of resistivity, c) resistivity uncertainty, d) mean model for an application of the ES-MDA algorithm only using resistivity data, e) hydraulic conductivity reference, f) mean model of hydraulic conductivity, g) hydraulic conductivity uncertainty, and h) is the mean model for an application of the ES-MDA algorithm only using hydrological data.

aquifer, they have a strong influence on all the observation points and are consequently easier to resolve for the ES-MDA algorithm.

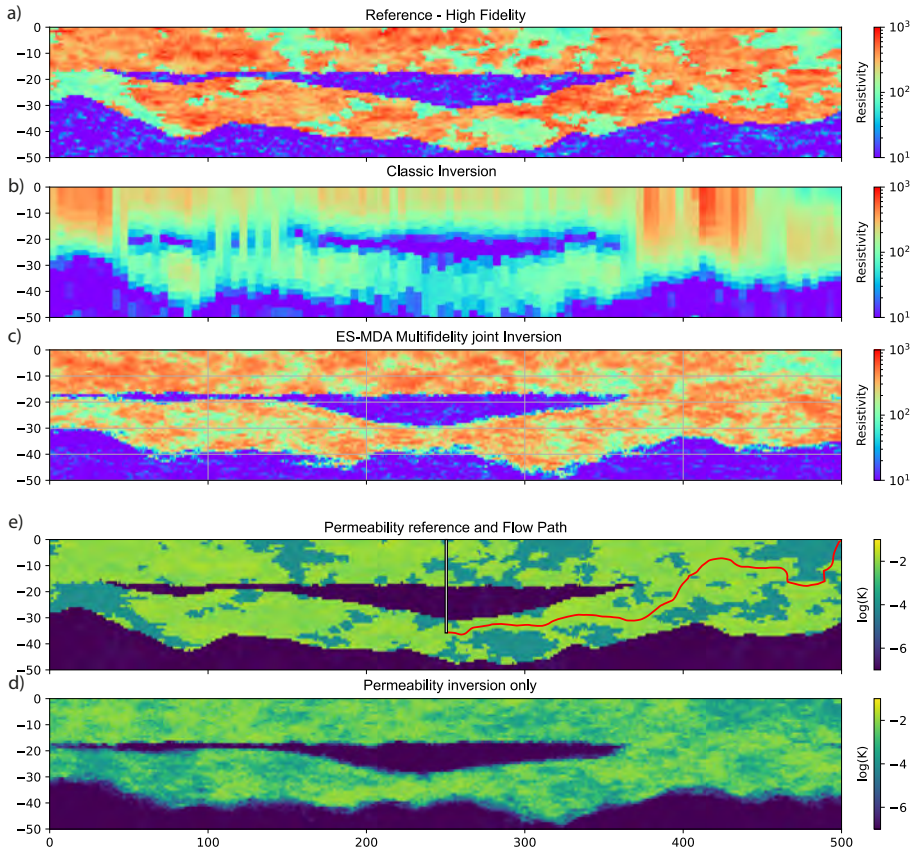

Figure 5.8: High-fidelity results for model C2. a) reference field for resistivity b) Resistivity field from a deterministic inversion c) ES-MDA inverted resistivity field d) reference field for hydraulic conductivities with are well and the simulated steady-state flow line e) mean model for an application of the ES-MDA algorithm on hydraulic conductivities only

Case C2 is more complex because the middle clay layer creates additional spatial variability and discontinuities. However, the same number of iterations as for case C1 were performed. Figure 5.8a shows the result of the geophysical data inversion only using a classical inversion based on a deterministic Newton-Gauss minimization using *AarhusInv* (Auken et al. 2015). Due to the diffusive nature of the geophysical method, the result-

ing model is significantly smoother than the reference (Figure 5.8c). The result of the ES-MDA inversion shows much sharper boundaries (Figure 5.8b) thanks to the embedding of the geological prior in the inversion method and the multi-fidelity approach.

Figure 5.8d shows the results of the ES-MDA inversion on the hydraulic conductivities only. Unlike model C1, this result shows a better identification of numerous discontinuities in the field of hydraulic conductivities. We interpret this difference as follows: the contrast of hydraulic conductivities being sharper in case C2, more information can be captured by the hydrogeological data. This result confirms that the poor identification of the spatial distribution of the hydraulic conductivities for case C1 was simply the result of a lack of information in the hydrogeological data set.

Another interesting result is the comparison between the classic ES-MDA "monofidelity" approach and the multifidelity approach. In a classic ES-MDA approach, data assimilation is done directly on the whole parameter fields, with no LF step. In other words, the classic ES-MDA inversion is only the HF, with the difference that the starting models are drawn in the full space of the prior and not in the reduced space constrained by the ensemble of LF models. Figures 5.9 a-c and d-f compare the results of the classic ES-MDA inversion, with the multifidelity. For both parameters, we can first denote that the simulations are visually more noisy than multifidelity simulations. The boundaries between the bodies are better defined in multifidelity. The continuity of the geological structures is also better established. The same number of iterations was performed on both. Figure 5.9 g) shows the mean residuals in relation to the computation time. The two iterations of the LF took only about 3.5 min. We can see a slight increase in the residual between the last LF iteration and the first HF iteration. But the absolute value is much lower than the monofidelity one. In terms of computing time, it is interesting to note that even if the absolute number of forward calls and ES-MDA loops is the same, the multifidelity is faster. The geophysical forward is almost not affected by the complexity of the model to simulate. However, the direct transient hydrological problem can show significant computing time variations depending on the hydraulic conductivity field. The first iterations are significantly slower, due to more complex models with high and abrupt hydraulic conductivity contrasts, for example. These results show that the use of the multifidelity approach tends to produce

more realistic models, geologically speaking, with a shorter or equivalent computing time.

Predictions

To further test the predictive capacity of the approach, we consider an additional hydrogeological scenario. We implement a well at position $x = 250$ m reaching a depth $z = -35$ m. The well is only screened along the last meter. Taking into account an hourly pumping rate of 36 $m^3/hour$, we calculated the path and travel time needed for the pollution introduced in the river to reach the well of water production. Such scenarios and questions are common for various applications in hydrogeology. Figure 5.8e) shows the reference path, going from the upper right corner river to the well in 37.5 days (advective time). The same computation was conducted for all prior and posterior models, as well as for all models obtained using the classical ESMDA joint inversion. The results of the flow path computations are shown in Figure 5.10a), 5.10b) and 5.10d). The reduction of the uncertainty on the envelope of the possible flow paths is clear between the prior distribution and the results of the two inversions. The two inversions provide similar results, with a proper estimation of the flow path compared to the reference. The classic ESMDA inversion tends to show a narrower uncertainty. The ES-MDA multi-fidelity ensemble predicts an arrival time of the pollutant between 35.75 and 87.1 days (10-90%interval) after injection. The prior gives a range between 21.5 and 10'000 days (10-68% interval) with 32% of the models predicting that the pollutant will not reach the well. The classical (mono-fidelity) ESMDA inversion ensemble predicts an arrival of the pollutant between 30.50 and 48.01 days (10-90%interval). Again, both inversions perform well, with a narrower time range for the classical ES-MDA inversion.

5.4 Discussion

The proposed multi-fidelity ES-MDA inversion has successfully identified sharp and complicated geological models.

When a Bayesian MCMC algorithm may take one or a few weeks to converge and generate an ensemble of realizations matching the data, the proposed approach only needs a few hours. Although several limitations

still exist. First of all, our reference model was generated using the same covariance models as the prior. As shown by Juda et al. (2022), the choice of an erroneous geostatistical prior can drastically decrease the realism and convergence rate of the inversion algorithm. This choice was straightforward in our synthetic approach but could be trickier when applying the method to real data sets. The solution requires properly analyzing all available field data to identify the required geostatistical parameters. If data are not sufficient to constrain the prior model, a possible solution would be to use published data from analog sites and generate a broad ensemble of models using different priors. The issue would then be to have a sufficient number of members to cover the whole prior parameter space and ensure that the ES-MDA algorithm would not create models that would be too far from reasonable geological models.

Another limitation of the inversion or data assimilation process is that we considered the noise in the data to be uncorrelated. This assumption is commonly used to treat each residual data point independently. In the synthetic case, this assumption is valid but could become problematic on real and strongly correlated data noise. Finally, the ES-MDA inversion (both in high- and low-fidelity) is not bounded to generate models that remain within the prior. It can be an advantage in the case of an uncertain prior, but it can also become a challenge if the algorithm generates physically impossible parameters.

The groundwater model used for the synthetic case C1 is not very informative as illustrated by the poor results obtained with hydrogeological data alone (Figure 5.7). Indeed, since the river is the only varying boundary condition applied to the model, only hydraulic conductivities close to the river are inferred with reduced uncertainty. This suggests that a large ensemble of hydraulic conductivity distributions is compatible with the data and that it is difficult for ES-MDA to approach the reference. Increasing the size of the ensemble could extend the possibilities and improve the results. However, the computational cost would also increase significantly. It is likely that using different hydrogeological situations, such as including pumping tests or tracer tests, would improve the identifiability of the hydraulic conductivity. Nevertheless, our results strengthen the advantage of using a joint inversion: the geophysical data help infer most of the subsurface structures and the spatial distribution of hydraulic conductivities. This contribution of geophysics is due to the

fact that the different physical parameters are spatially correlated via the underlying litho-facies.

The main contribution of this paper is to propose to split the ES-MDA algorithm in two steps using a multi-fidelity strategy. We show that starting with the LFMs accelerates the inversion procedure significantly. It allows one to quickly delineate the main structures of the subsurface and quantify uncertainty. However, as it relies on a simplified version of the model, it cannot reproduce the reference in some places (Figure 5.6). This is a consequence of using homogeneous units (in terms of properties); thus, it is not possible to account for some important local variations of facies, and these have a significant impact on the geophysical data. Using different homogeneous values among the different members (or models) of the initial ensemble mitigates this effect by enlarging the uncertainty and allows one to identify the depth of the true surface accurately almost everywhere (Figure 5.6). However, it is clear that the uncertainty that we propagate in the second step of the inversion has a significant impact on the results. For example, if we consider the high value of the surface at around 330 m for case C1 (Figure 5.6a). The true surface is barely contained within the range of uncertainty, it means that the majority of the models considered the surfaces higher than the reference. As a consequence in the second step, ES-MDA compensates by predicting mostly "gravel" (high resistivity/hydraulic conductivity) just above this location, where normally there should be a relatively large area of "sand" (low resistivity/hydraulic conductivity, Figure 5.7). We then understand that small initial errors can have major impacts and that we should be careful with the final models. However, it should also be mentioned that the generated geological models are totally unconstrained. It is certain that incorporating more geological knowledge (such as borehole data) into the models would have helped to detect and solve this kind of inconsistency.

Our hydrogeological scenario involves 13 multilevel piezometers uniformly distributed over 500-meter length. Even if some field sites show similar or even higher density of data, in a real application the density of piezometric information could be lower. The consequence will be a higher uncertainty associated with the hydraulic conductivity field. A preliminary sensitivity analysis on our synthetic models showed that removing two piezometers either close to the source of the signal or far from it have drastically different effects on the final uncertainty.

This is simply due to the difference in the amount of information carried by the two data series. In addition, in our synthetic example, we considered rather simple boundary conditions (such as a sinusoidal variation of the river level) without any uncertainty. In practice, it would be straightforward to cope with more complex boundary conditions, such as time-varying water levels. For the uncertainty on the boundary conditions, the proposed approach would be to include these boundary conditions as parameters within the ESMDA inversion, this would result in a higher level of uncertainty for the overall aquifer characterization.

Finally, we think that the approach proposed in this paper could be applied to a wide range of problems. We illustrate here the multi-fidelity idea using a simpler geological model and a faster forward operator at the same time. However, many other combinations could be tested: for example, the low-fidelity model could correspond to a steady-state hydrogeological model and the high-fidelity model could be the full transient model. Another possibility could be to use only a few geophysical observation points for the low-fidelity step and a complete detailed data set for the high-fidelity step. One could also group geological units to simplify the geological architecture during the low-fidelity step.

5.5 Conclusion

The results presented in this paper demonstrate that ES-MDA, multi-fidelity, and ArchPy, all together, can constitute a potentially powerful framework for performing geologically consistent inversions. The proposed approach allows integrating in a consistent and stochastic manner different types of data and thus reducing the global uncertainty on groundwater models. The use of the multi-fidelity approach on such a problem has proven to be more efficient in infusing prior geological knowledge into the inversion. This has resulted in more realistic geological models while being faster. Future work includes the application of the presented methodology to real 3D field data.

Several main conclusions can be drawn from this research:

- ES-MDA is an efficient tool to get an ensemble of plausible hydrogeological models in a multi-fidelity framework.

- Hierarchical multi-fidelity helps to keep models geologically consistent during the inversion process while improving the quality of the models.
- Hydrogeophyiscal joint inversions can be decomposed and improved within a multifidelity and hierarchical framework.
- ArchPy's models are useful priors to investigate subsurface uncertainty.

Additional information

Accession codes The geostatistical codes used in this study are implemented in the python *ArchPy* package freely available at `https://github.com/randlab/archpy`. The 1D forward code used is *AarhusInv* developped by the Aarhus University Hydrogeophysics group (Auken et al. 2015). The AarhusInv code is free to use for research purpose. The flow model code used in the open-source *ModFlow 6* code Hughes et al. 2017.

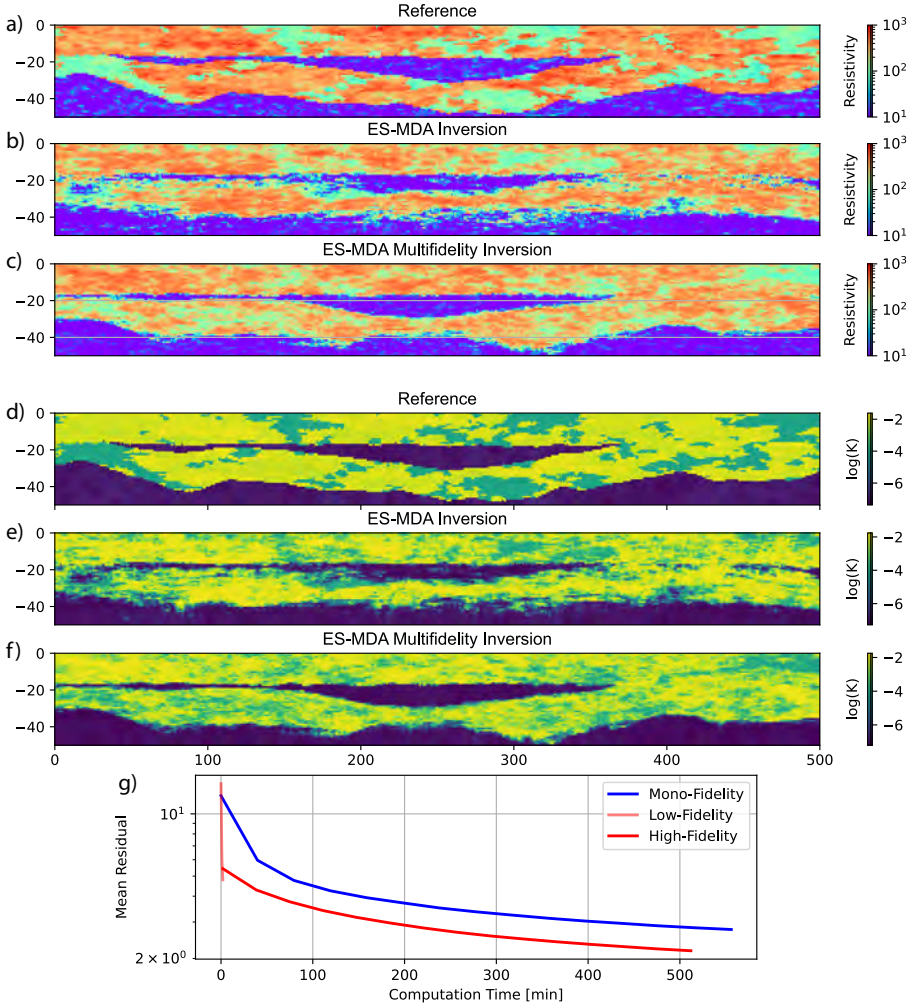

Figure 5.9: a to f) Comparison between the inversion model resulting from a classic ES-MDA inversion, the multifidelity ES-MDA inversion, and the reference model for both the hydraulic conductivities and resisitivity fields. g) The residuals versus the computation time for the classical ESMDA (monofidelity) and the multifidelity model.

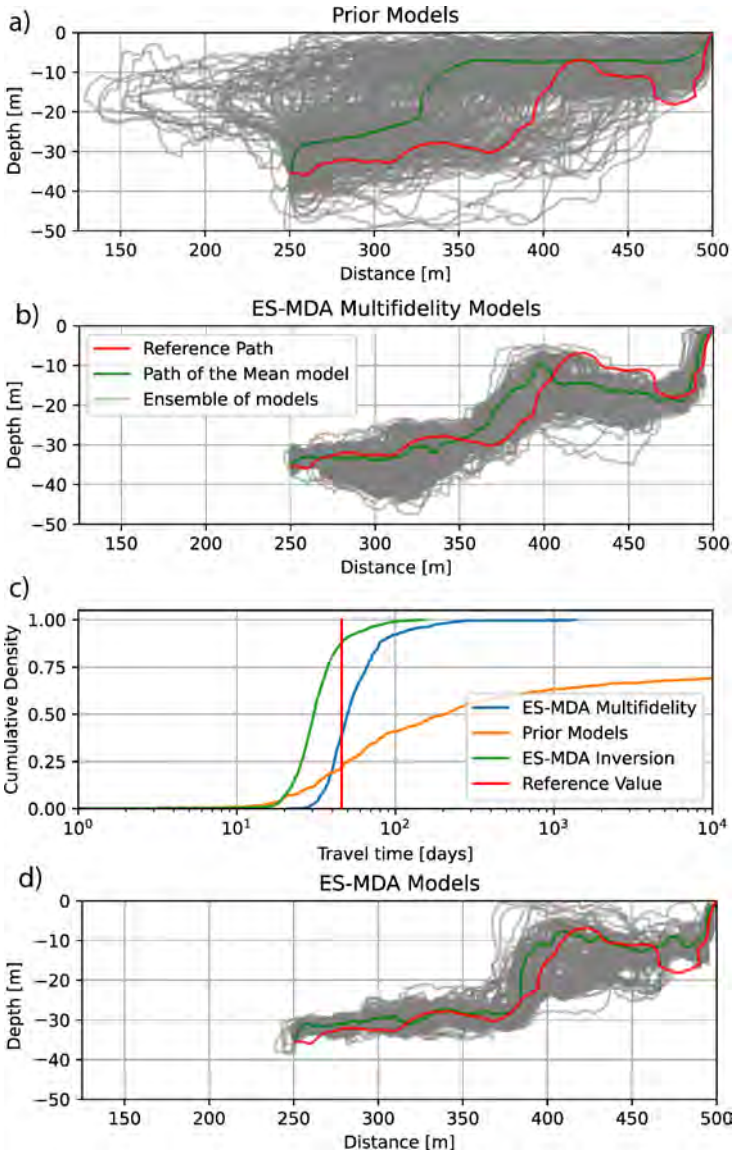

Figure 5.10: A and B): Flow lines for the prior and posterior multifidelity ensemble of model C2. C and D): Distribution of the travel time from the river to the well and Flow lines for the posterior monofidelity ensemble of model C2.

Chapter 6

A novel methodology for the stochastic integration of geophysical and hydrogeological data in geologically consistent models

Abstract

To address groundwater issues, it is often necessary to develop geological and hydrogeological models. Combining geological, geophysical and hydrogeological data available on a site to build such models is often a challenge. This paper presents a methodology to integrate such data within a geologically consistent model with robust error estimation. The methodology combines the Ensemble Smoother with Multiple Data Assimilation (ESMDA) algorithm with a hierarchical geological modeling approach (ArchPy). Geophysical and hydrogeological field data are jointly assimilated in a stochastic ESMDA framework. To speed up the inversion process, forward responses are computed in lower-dimensional spaces relevant to each physical problem. By doing so, the final models take into account multiple data sources and regional conceptual geological knowledge. This study illustrates the applicability of this novel approach using actual data from the upper Aare Valley, Switzerland. The results of cross-validation show that the combination of different data types, each sensitive to different spatial dimensions, enhances the quality of the model within a reasonable computing time. The proposed methodology allows the automatic generation of groundwater models with robust uncertainty estimation and could be applied to a wide variety of hydrogeological issues.

6.1 Introduction

Groundwater resources represent 99% of the total liquid fresh water on Earth and provides approximately 50% of the total volume of water withdrawn for domestic use in the world (United Nations Educational Scientific and Cultural Organization 2022). However, because the resource is at depth below the ground surface, the interest of the general public and policy makers tends to wane due to its invisible aspect (United Nations Educational Scientific and Cultural Organization 2022). In a changing climate, being able to understand, predict and model underground flows and resources is crucial. These questions can only be properly addressed by local to regional groundwater modeling.

Another issue owing to the underground position of groundwater is the difficulty in collecting data. Data are usually sparse and often represent only a tiny fraction of the total volume of the aquifer. The data can therefore miss most of the spatial heterogeneity. Conjectures based on such

groundwater models could lead to inadequate decisions. A recent paper by Adams et al. (2022) reviews the different existing remote sensing techniques (including geophysical techniques) and possible future approaches that could help fill this gap. They argue that satellite or airborne remote sensing could be key methods for understanding groundwater dynamics at various scales, which would have been impossible with classical spare in situ observations. One method mentioned in their paper is the electromagnetic induction techniques, and especially the Time Domain Electromagnetic (TEM). The TEM method is relatively inexpensive, fast and has a depth of investigation ranging from a few meters to a few hundred meters (Christiansen et al. 2009b; Fitterman 1987). Due to its inductive principle, the EM method will be mainly sensitive to underground electrical resistivity. Consequently, it will also be sensitive to all parameters that will affect this property, such as porosity, water saturation, clay content, and water electrical conductivity. EM was successfully applied to various hydrogeological situations, usually in combination with sparse borehole data (Barfod et al. 2018; Christensen et al. 2017; Dumont et al. 2018; Paine 2003, e.g.).

However, the use of electromagnetic (EM) data for hydrogeological interpretation also poses some challenges. Firstly, the method will provide only limited information regarding the water storage capacity or the hydraulic conductivity. Second, it is also sensitive to the geological composition of the underground. The amount of conductive clay will strongly affect the measured field and, therefore, the resulting inversion. Thirdly, due to the nature of the inverse problem and its high nonlinearity in the case of EM, the inverted resistivity models are nonunique (Tarantola 2005). As underlined by Adams et al. (2022), therefore a good understanding of the geological properties of the aquifer and their spatial variations is necessary in order to interpret geophysical data robustly. So far, the most popular workflow has been to independently invert EM data and then relate them to boreholes or other types of data. (Christensen et al. 2017; Dumont et al. 2018; Jørgensen et al. 2015, 2013; Kang et al. 2021; Knight et al. 2018; Neven et al. 2022a; Ringrose et al. 2015). Often, the inversion and some steps in the workflow are considered deterministic. Therefore, the final structural model, the so-called cognitive model, is the one that, according to the modeler, fits most of the comprehensive knowledge available on site (Henriksen et al. 2003; Høyer et al. 2015; Kollet et al. 2006; Lemieux et al. 2008). The calibration of the groundwater model parameter is then carried out while keeping the cog-

nitive model fixed. By doing so, there is no way to account for additional information about the geometry coming from the hydrogeological data, while the transmissivity and therefore the aquifer thickness clearly influence the head gradient or solute transport. If the geometry was fixed before inverting for hydrogeological parameters, and not correctly estimated, the hydrogeological parameters will likely have to be incorrectly identified during inversion in order to compensate for these initial errors. Furthermore, the use of a single cognitive model derived from geophysical data completely neglects structural uncertainty and, consequently, overall uncertainty in the aquifer characterization procedure.

To ensure consistency and the propagation of errors throughout the workflow, we propose to reverse the methodology described in the previous paragraph. Instead of going from the geophysical data to the structural model, we start by generating a set of prior plausible geological models and then updating these 3D models iteratively to fit the observed data. We adjusted the models jointly on the EM data acquired at the surface and the hydraulic heads measured in the aquifer. Since prior data are generated from the boreholes, we ensure that the final ensemble of models (posterior) agrees with them, even after parameter adjustment.

In this study, we generate geological models using the ArchPy hierarchical modeling approach (Schorpp et al. 2022). This method offers a formal description of a geological concept, the stratigraphic pile, that encompasses the succession of units, lithologies, and properties that must be modeled within a given domain. It also integrate knowledge from wells and conceptual understanding of the geological setup to create geostatistical realizations of potential aquifer geometries and properties. ArchPy utilizes a hierarchical approach that first simulates the main stratigraphic units, then the litho-facies, and finally the petrophysical parameters. The result is a complex and accurate representation of the aquifer structure. However, ArchPy cannot integrate other knowledge, such as geophysical or hydrological. The prior set of models needs then to be updated, in order to account for all the other types of data available on-site.

When updating models, the Markov chain Monte Carlo (MCMC) has been shown to be capable of producing consistent models and properly quantifying the final uncertainty (Irving et al. 2010; Jardani et al. 2013; Mosegaard et al. 1995; Reuschen et al. 2021). One downside of this method is the large number of forward calls required to converge and the

associated computational cost (Linde et al. 2016). Ensemble Smoother algorithms have proven to be a reasonable alternative to MCMC for high-dimensional problems. It managed to identify solutions for complex nonlinear inverse problems, with fewer computational resources and time than MCMC (Emerick et al. 2013; Juda et al. 2022). Ensemble Smoother with Multiple Data Assimilation (ESMDA) (Emerick et al. 2013) is a variant of the Ensemble Smoother (ES) algorithm proposed by Leeuwen et al. (1996). It approximates the relationship between the parameters and the data using their covariance computed using a finite prior ensemble of models. For large or extremely non-linear problems, the data can be assimilated multiple times iteratively (Chen et al. 2013). ESMDA was used successfully in various groundwater studies (Kang et al. 2019; Lam et al. 2020b; Li et al. 2015; Xu et al. 2021). Similar algorithms (ENKF, for example) were also successfully applied to jointly integrate different types of data, such as geophysical and hydrological (Bouzaglou et al. 2018). However, this application was adapted only on a relatively small-scale lab experiment.

In a previous study, we used a synthetic 2D data set and showed that joint hierarchical inversion combined with ESMDA can improve the identification of model parameters and the reliability of a prediction and its uncertainty, while integrating complex geological prior knowledge (Neven et al. 2022b). In the present study, we extend the method and show that this novel approach can be applied to real field data to integrate geophysical and hydrogeological data in 3D geologically consistent models. A new and important aspect of this paper is the development of a simple and consistent approach to identify the 3D parameter fields while computing the forward simulations in lower-dimensional spaces that can be different for each physical problem. To demonstrate its applicability, the new methodology is illustrated on a part ($\sim 15km^2$) of the upper Aare Valley, Switzerland. The area is densely sampled, with a towed Transient Electromagnetic survey (Neven et al. (2021)), several hundred boreholes, and 25 piezometers.

6.2 Methodology

The proposed methodology can be divided into three steps (Fig. 6.1): 1) the generation of prior geological models, 2) the forward simulations in a reduced dimensionality, and 3) the iterative update of the ensemble of models using ESMDA.

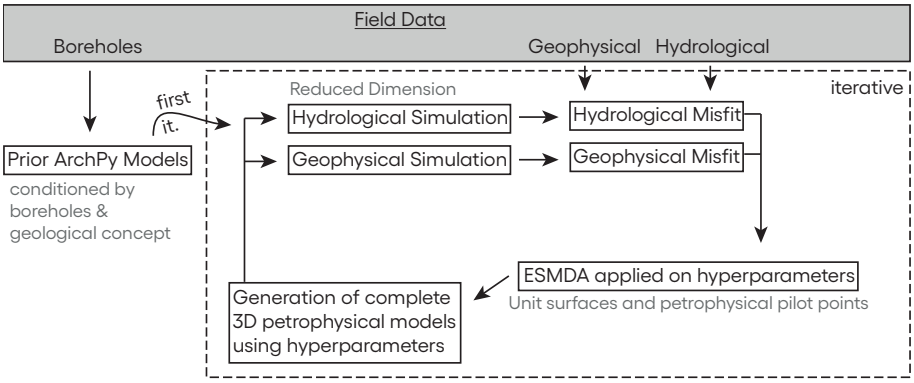

Figure 6.1: General sketch of the proposed data integration workflow

Generation of prior ArchPy models

We generate the prior ensemble of models using the open source stochastic hierarchical code, ArchPy[5] (Schorpp et al. 2022). This Python-based tool allows for a detailed and hierarchical depiction of the geological environment, starting from the simulation of the contacts of large syndepositional layers, using 2D geostatistical methods, to the distribution of categorical lithofacies within these layers, and ultimately the petrophysical properties associated to the lithologies. ArchPy relies on the concept of stratigraphic pile, which is a formal description of the known regional stratigraphy of a given area. A large ensemble of plausible geological models compatible with regional knowledge and borehole data can be easily generated. For an in-depth description of the ArchPy methodology and examples of its capabilities, the reader is referred to Schorpp et al. (2022).

In addition to the existing ArchPy approach, to address possible non-stationarity in the petrophysical parameter simulations, we added the possibility to employ petrophysical pilot points. The underlying assumption is that large areas of the same aquifer can be significantly more (or less) hydraulically permeable than others. The conductive clay content can also change and cause variations in terms of electrical resistivity within the same unit. Consequently, the local average used for the simulation cannot be assumed to be uniform in space in the whole domain. To account for this, pilot points are randomly placed in each layer using

[5] `http://www.github.com/randlab/ArchPy`

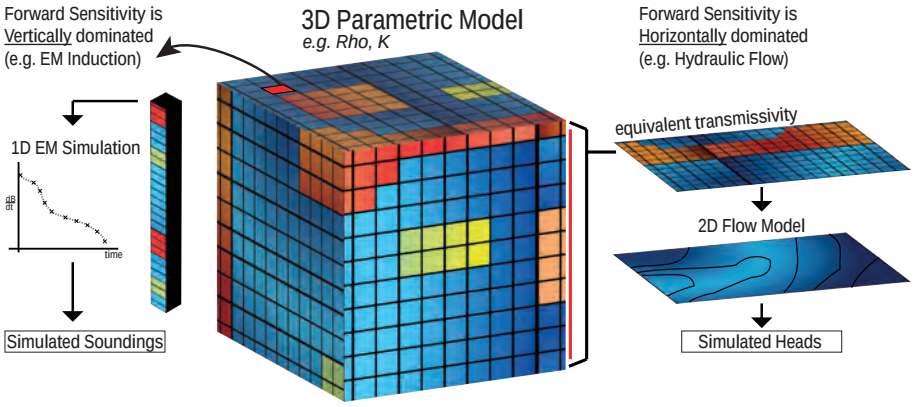

Figure 6.2: General sketch of the Forward in reduced dimensionality principle

a stratified sampling strategy. This ensures that the points are drawn randomly through the entire domain. The values of the pilot point are then drawn in a distribution of plausible range. Then, a linear variogram is used to krigge between the pilot points within each unit. This smooth map is then used as the local mean for the 3D petrophysical geostatistical simulation. The petrophysical parameters will vary following the variogram model provided around a spatially variable mean. A set of pilot points is defined for each layer and for each petrophysical parameter.

Forward simulations in reduced dimensionality

As pointed out, for example, by Linde et al. (2016), the inversion of high-dimensional models leads to a large number of iterations and a large number of forward model calls for solving the physical, possibly nonlinear, spatial and/or time-dependent equations. Forward calculations are therefore often the most computationally expensive part of an inversion. Consequently, it is common to reduce the number of parameters (Linde et al. 2016), or to use surrogate modeling (Asher et al. 2015), for example.

In this study, we propose applying the inverse procedure by correcting the 3D geological and parameter models, but before running the forward flow and geophysical calculations, we reduce the dimensionality of the model to accelerate the forward simulations. Computing time is drastically reduced with a marginal loss of information. Fig. 6.2 illustrates the

principle of this approach. The idea is to exploit the fact that different physical processes are sensitive to different subsets of the 3D parameter fields. For example, the propagation of the Electromagnetic Magnetic (EM) signal requires us to solve the Maxwell equations to simulate the secondary magnetic field decay. For a horizontal coil, the main sensitivity is oriented downward. For most applications, solving the complete 3D equations will only slightly affect the final results but will result in a high increase in computational complexity (Engebretsen et al. 2022). For this reason, the EM geophysical forward is only computed using multiple 1D models extracted along the Z axis. Each geophysical sounding is then computed in parallel using a fraction of the original model. The simulation time is greatly reduced, from several orders of magnitude. In this first example, the dimension is reduced by omitting some part of the model that is expected to have a limited effect on the predicted data.

However, flow within an aquifer is an example of horizontally dominated sensitivity. Instead of extracting some parts of the model, we use another approach here. Dimensional reduction is achieved by integrating spatial properties along the vertical direction. Because the overall thickness is small compared to the lateral extension and because we only have one hydraulic head measurement per piezometer, it is reasonable to solve the groundwater flow only in 2D. The complete 3D hydraulic conductivity model is then vertically integrated, and the horizontal transmissivity of the total aquifer is estimated at any location. The equation is the following :

$$T_{x,y} = \sum_{iz=1}^{nz} K_{x,y,iz} \cdot d_{iz} \tag{6.1}$$

where $T_{x,y}$ is the transmissivity in a cell x, y, nz is the number of cells in the model along the vertical axis, K is the hydraulic conductivity and d is the thickness of the cell. The equivalent transmissivity is used to to solve the 2D equivalent flow.

Of course, the simplified 1D and 2D forward simulations will not completely represent the actual 3D processes. The gain in terms of computing resources overcomes the marginal loss of information. For example, for the regional steady-state hydraulic heads simulation, the average gain in computing time between a full 3D and a 2D model is about 98% (51 s per model compared to 1 s per model), for a difference in response of a few centimeters (see Fig. 6.6 and discussion in section 6.3).

Data assimilation with ESMDA

ArchPy allows us to rapidly generate a large number of plausible models. This ensemble represents our prior knowledge; it includes the regional geological concept expressed through the stratigraphic pile and all the borehole data. However, the compatibility of each model with indirect field data (geophysical or hydrogeological) is not guaranteed. To adjust the models based on all available data, we implemented the Ensemble Smoother with Multiple Data Assimilation (ESMDA) algorithm introduced by Emerick et al. (2013). In this section only a brief introduction to ESMDA is provided; readers are directed to the reference paper (Emerick et al. 2013) for more details on the method.

This stochastic data assimilation algorithm performs iterative corrections of a finite ensemble of N models $\{m_i^{pr}, ..., m_N^{pr}\}$. At iteration 0, the models are conditioned solely by the boreholes. In the next iterations, the members of the ensemble are the corrected models of the previous iteration. At iteration k, the correction requires computing the Kalman gain matrix K from the experimental covariance matrix between the data and the model parameters C_{MD}^K combined with the data auto-covariance matrix C_{DD}^k and the expected uncertainty of the data C_{err}. C_{err} can be estimated using multiple stacking of the data in the field.

$$K = (C_{MD}^k (C_{DD}^k + \alpha_{k+1} C_{err})^{-1}) \tag{6.2}$$

The parameter α is the noise inflation ratio. This parameter is added to address one common issue with Ensemble Smoothers and Ensemble Methods in general: inbreeding (Houtekamer et al. 1998). This term is used to describe the underestimation of uncertainty due to the fact that the ensemble used for the calculation of the Kalman Gain K is the same as the one used to estimate the error. The individual models within the set are becoming more and more coupled after each iteration. Emerick et al. (2013) showed that this inbreeding tends to be reduced with an increased number of members, and proposed the use of an inflation factor α. This factor must respect the following relationship:

$$\sum_{k=1}^{N_{iter}} \frac{1}{\alpha_k} = 1 \tag{6.3}$$

with N_{iter} being the number of iterations. Changing α through iterations does not significantly increase the quality of assimilation (Emerick 2016). Consequently, we kept it unchanged at $\alpha = N_{iter}$ for all iterations. The

correction for each member at iteration k is determined as follows:

$$m_i^{k+1} = m_i^k + K \odot LM \cdot (d_{obs,i}^k - g(m_i^k)) \tag{6.4}$$

$$d_{obs,i}^k = d_{obs} + \sqrt{\alpha_{k+1}} C_{err}^{1/2} z_{d,i} \text{ with } z_{d,i} \sim N(0,1) \tag{6.5}$$

where $d_{obs,i}^k - g(m_i^k)$ is the Euclidean distance between the observed measurements and the predictions computed by the forward operator g using the current model parameters. To account for the imprecision in the measurement of the observed data, d_{obs} are perturbed with random Gaussian noise drawn from $N(0, \sqrt{\alpha_{k+1}} C_{err})$. The noise realizations are resampled for each iteration and ensemble member. As shown in Evensen (2018), each iteration of ES-MDA implicitly minimizes a different cost function. It combines on the one hand the difference between the parameters before and after the given iteration weighted by the inverse of the prior parameters covariance; and on the other hand, the misfit between the observations and the predictions. To reduce the risk of spurious correlations, a localization matrix LM is introduced. The localization matrix multiplies the Kalman gain elementwise. Spurious correlations occur mainly because ESMDA approximates the covariance matrix based on a finite number of models (Anderson 2001; Evensen 2009; Wen et al. 2005). Therefore, zero entries are difficult to reproduce in the covariance matrix, and spurious correlation can occur. One way to address this is either to increase indefinitely the number of members or to infuse prior knowledge in order to exclude correlations that are physically impossible. For example, we can expect no correlation between an EM sounding (data) and the resistivity value several kilometers away (parameter). The localization matrix varies between 0 and 1, and is inversely proportional to the spatial distance between a parameter and a data point. The Kalman Gain K will become negligible for a distant set of parameter-data. Several functions exist to scale the localization matrix with distance. In this study, we implemented the widely used correlation function introduced by Gaspari et al. (1999). They propose the following piecewise rational covariance function R_{GC} for data assimilation:

$$R_{GC}(r) = \begin{cases} 1 - \frac{5}{3}\left(\frac{r}{r_c}\right)^2 + \frac{5}{8}\left(\frac{r}{r_c}\right)^3 + \frac{1}{2}\left(\frac{r}{r_c}\right)^4 - \frac{1}{4}\left(\frac{r}{r_c}\right)^5 & 0 \le \frac{r}{r_c} < 1 \\ 4 - \frac{2}{3}\left(\frac{r}{r_c}\right)^{-1} - 5\left(\frac{r}{r_c}\right) - \frac{5}{3}\left(\frac{r}{r_c}\right)^2 + \frac{5}{8}\left(\frac{r}{r_c}\right)^3 + \frac{1}{2}\left(\frac{r}{r_c}\right)^4 + \frac{1}{12}\left(\frac{r}{r_c}\right)^5 & 1 \le \frac{r}{r_c} < 2 \\ 0 & \frac{r}{r_c} \ge 2 \end{cases} \tag{6.6}$$

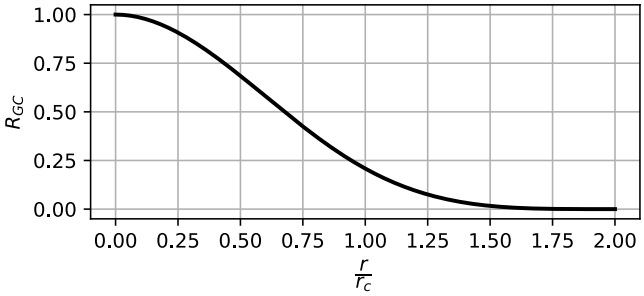

Figure 6.3: Gaspari Cohn correlation function (Gaspari et al. 1999) value versus the ratio between the distance r and the critical distance r_c.

where r is the distance matrix of the Euclidean distance between the data observation points and the parameters, and r_c is the critical distance beyond which the effect of a parameter on an observation is assumed to be negligible. Fig. 6.3 represents R_{GC} in 1D as a function of the ratio between r and r_c. The critical distance of the localization matrix should be chosen according to the expected correlation length. An analysis of the evolution of the correlation with distance for both observation types is available in the supplementary material (Figure S1). In our case, we set the critical distance of geophysical data to 150 meters, since a stabilization of correlation is visible around this range. No significant evolution in correlation with the parameters was observed for the hydraulic head measurements as a function of the distance. Therefore, the LM was not implemented for these observations and is set to 1 for all hydraulic heads measurements. The summary of the LM parameters is available in Supplementary Materials (Table S2).

6.3 Application to the Aare Valley

To illustrate the applicability of the proposed methodology on a real site, we used a test area in central Switzerland. The upper Aare valley is filled with Quaternary deposits, with a thickness ranging from a few meters on the sides of the valley to a few hundred in the center (Bandou et al. 2022). The investigation zone is located north of the upper Aare valley, just upstream of the city of Bern. It is a $12km^2$ area where a large amount of drinkable water is exploited. Several glacial cycles occurred, causing a complex interwinding of glacial and interglacial deposits (Bandou et al. 2022; Graf et al. 2016; Schorpp et al. 2022). At least two different levels

of aquifer have been identified in interglacial fluvial deposits (Schorpp et al. 2022). A shallow aquifer at the surface and a deeper one that has not been extensively studied. Despite this lack of knowledge, the upper Aare valley is widely used. In 60 km^2, we denote 4 quarries and more than 6000 wells (shallow geothermic or water production). Despite its importance, the exact dynamics of the aquifer and its extent on the regional scale are not well known. Due to its importance, this site has been extensively studied with different data types (see Fig. 6.5). It illustrates well how the abundance of data does not necessarily lead to an improved underground model due to the lack of an easy and applicable integration algorithm in hydrogeology. In the area, we denote the presence of 130 lithologically described boreholes, 35 hydraulic head measurements, and about 6500 tTEM soundings. While the hydraulic heads measurements are more or less homogeneously distributed in the domain, there is a clear spatial division for the other two data sets. The boreholes are drilled most of the time within the cities or around existing buildings (within a few hundreds of meters) in areas where the electromagnetic coupling prohibits the use of inductive methods such as EM. The two data sets then cover different areas and are, by nature, complementary.

ArchPy Model

To define the conceptual model of the area, we rely on expert knowledge, existing boreholes, and previous publications (Graf et al. 2016; Schorpp et al. 2022; Volken et al. 2016). Our zone of interest is the first 100 meters of the valley filling. The conceptual model of the area includes a very shallow aquifer followed by impermeable clay. Below the clay, a second aquifer can be present. Finally, a deeper clay layer is present everywhere. When the second aquifer is absent, the two clay layers become one, even if they are still considered two distinct units. The top surface of the first layer is the topography. Therefore, there are only three surfaces delimiting the four units. For each surface, a variogram model was derived from the analysis of the borehole data and expert knowledge. All details of the parameters are shown in Table 6.1. The surfaces are simulated as 2D Gaussian Random Fields with inequalities using the Sequential Gaussian Simulations (SGS) method (Chilès et al. 2009; Deutsch et al. 1992; Freulon et al. 1993). In each layer, we also define the variogram for the simulations of petrophysical parameters. Once the units are generated, the lithologies are assumed to be uniform within each unit: the aquifer units are assumed to be made of gravel to sandy-gravel, and the

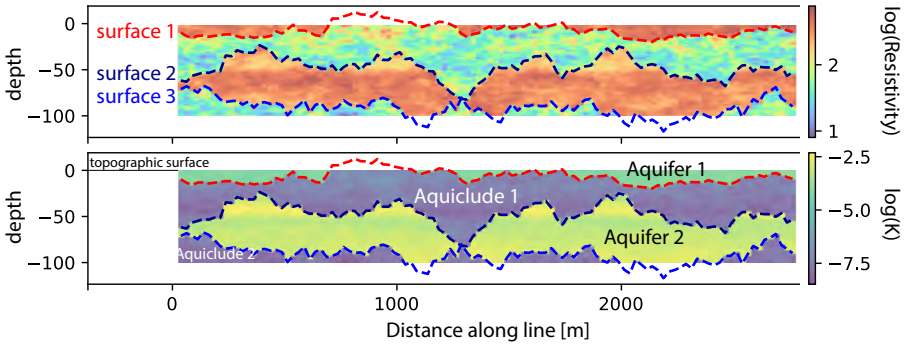

Figure 6.4: Cross section through one realization of a borehole-only ArchPy petrophysical model of electrical resistivity and hydraulic conductivity. In addition, the surfaces bounding the units are outlined. Note that there is a vertical exaggeration.

other units are assumed to be made of clay. We expect spatial variations in terms of petrophysical properties within each lithology as a result of sedimentological changes, while we do not expect major spatial variations of lithologies within each unit. The two petrophysical parameters of interest are electrical resistivity (ρ) and hydraulic conductivity (K). They are modeled within the units using 3D Gaussian Random Fields (GRF) (Chilès et al. 2009, e.g.) simulated with the circulant embedding spectral method (Lang et al. 2011). This method is faster for large 3D models compared to SGS when inequalities are not needed. In addition, as mentioned before, pilot points are implemented to represent a possible non-stationarity within the domain. They are used to parameterize the local spatial mean for the 3D GRF simulation during the data assimilation. Fig. 6.4 shows a cross-section through one realization of an ArchPy petrophysical model. The three simulated surfaces and the topographic surface define the four units. Two are considered to be mainly clay and 2 are considered aquifer. Thanks to the ArchPy hierarchical approach, we can simulate different parametric fields for each of the units, according to the prior range of values.

The borehole logs were digitized by the Swiss Geological Survey during the GeoQuat project (Volken et al. 2016). As discussed above, we distinguish only aquifer formations from aquicludes. Note that most of the boreholes were drilled for the exploitation of the shallow aquifer. Therefore, most of them stop before or when they reach the bottom of the

Table 6.1: Geostatistical parameters used by ArchPy to generate the geological models. The variogram type used for all the litho-facies, resistivity, and hydraulic conductivity models is exponential.

Stratigraphic surface model

Unit	2	3	4
variogram type	linear	linear	linear
range x and y [m]	1000	1000	1000
sill [m^2]	70	150	150

Litho-facies model

Unit	1	2	3	4
Type uniform facies	gravel-sand	clay	gravel-sand	clay

Resistivity Model

facies	mean [$\log_{10}[\Omega m]$]	range x [m]	range z [m]	sill [$\log_{10}[\Omega^2 m^2]$]
gravel-sand	2 – 2.7	200	20	0.05
clay	1.3 – 1.8	20	4	0.01

Hydraulic Conductivity Model

facies	mean [$\log_{10}[m/s]$]	range x [m]	range z [m]	sill [$\log_{10}[m^2/s^2]$]
gravel-sand	-5 – -1.5	200	20	0.01
clay	-9 – -6	200	20	0.05

upper aquifer. The vertical depths are then usually between 3 to 12 meters. Only 3 wells reach the lower aquifer at more than 40 m depth. All of these wells were incorporated into ArchPy and used to constrain the geostatistical simulations. An important feature of ArchPy is the integration of inequalities (Schorpp et al. 2022) during geostatistical simulations. A well that does not reach the top of a certain unit may still have an effect on the local uncertainty since it indicates that this surface must be absent at that location, or deeper than the bottom of the well. The complete ArchPy model used for the prior generation can be found in the

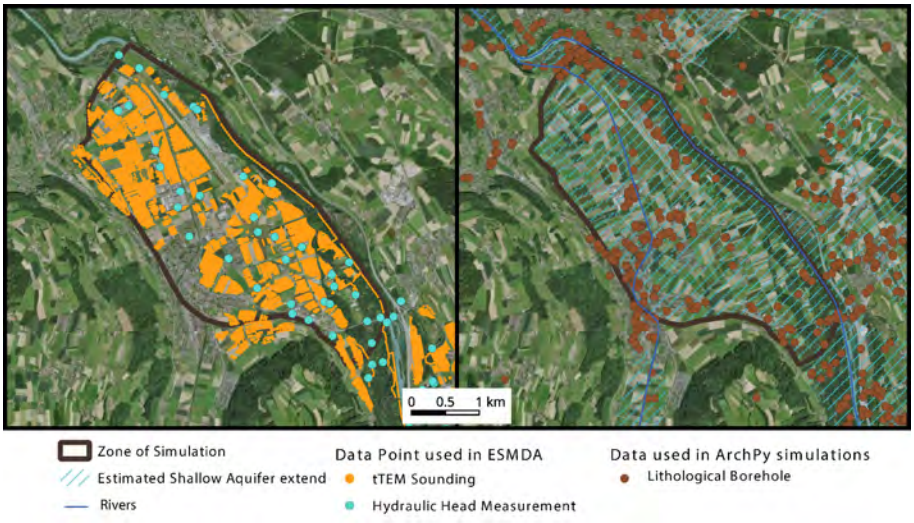

Figure 6.5: Extent of the simulation with the available data illustrated besides. The base map is published open access by the Swiss Topographic Office.

Zenodo repository, as a Python object (Neven et al. 2023). A summary of the parameters subject to assimilation and their prior distribution is available as supplementary materials (Table S1).

Geophysical Data

The tTEM data come from a large-scale acquisition carried out in 2020. Data are publicly available. All characteristics of the system, data processing, cleaning, and validation, were documented in Neven et al. (2021). Data are used in raw electromagnetic decay sounding form (dB/dt), and therefore are not affected by any inversion choice.The forward code used to simulate the models is a 1D Time-Domain EM implemented in Fortran in *AarhusInv* (Auken et al. 2015). The use of a 1D forward simulation code is common when dealing with EM technology (Auken et al. 2015; Farquharson et al. 1993; Scholl et al. 2009, e.g.). The relying assumption is that locally the subsurface can be considered in his vertical dimension only. The full system geometry is used for the calculation of the synthetic responses. These responses can then be compared with the raw electromagnetic decay measured on the field. As mentioned in the data descriptor, the uncertainty associated with the observations in the data

is derived from the multiple stacking of soundings during the acquisition, directly in the instrument. As mentioned in Auken et al. (2019), the tTEM instrument calculate the uncertainty by stacking few hundreds measurement. In the area of Belp, 6400 soundings of 29 gates each are integrated (see Fig. 6.5).

Hydrogeological Data

Hydraulic head measurements are from a large-scale piezometric campaign (Kellerhals et al. 1981b). At the same time, the heads were measured at hundreds of points in the valley, 35 of which are in our area of interest. The levels of the two rivers were also measured. Recent but less dense measurements acquired in 2020-2022 showed that those previous measurements are still representative of current aquifer conditions. Due to the lack of sufficient time series to compare with, we decided to use only a steady-state model in this paper. The forward groundwater model is two-dimensional and is based on ModFlow6 (Hughes et al. 2017). Fig. 6.6 shows the comparison between the heads simulated using a 2D transmissivity model and a complete 3D model for an ArchPy realization. We denote only a marginal loss of information with an average difference in heads of less than 12 centimeters. We can denote that the heads further away from the boundary conditions, which also present an intermediate head value on figure 6.6, are also those that present more variations between the 3D and the equivalent transmissivity simulation.

Recharge was estimated using precipitation measured at the site and was set spatially uniform to 652 mm of rain per year. River boundary conditions were imposed along the two stream paths (see Fig. 6.5) using the altitude profile of the streams and the hydraulic heads measured during a piezometric campaign in the river (Kellerhals et al. 1981b). A head is imposed with a stream conductance damping (Hughes et al. 2017). In addition, two constant head boundary conditions were assigned: one along the southeast boundary of our model and one along the north boundary. Their value was set to, respectively, 506 m and 516 m, following the isohead of multiple measurements. All known water pumping wells present in the area were added as flow boundary conditions, and their estimated flow is provided in Kellerhals et al. (1981). A conceptual representation of the groundwater flow model is available as supplementary material (Fig. S2). All Modflow 6 files are also available in the Zenodo repository Neven et al. (2023). During data assimilation, only the hydraulic con-

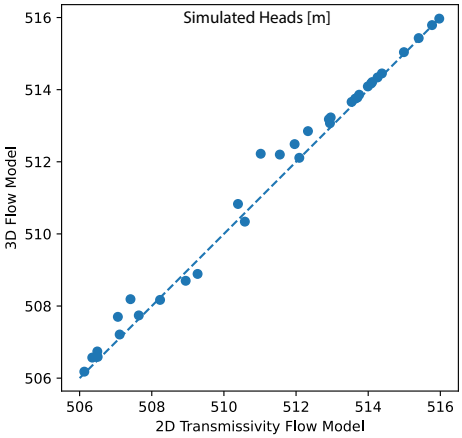

Figure 6.6: Comparison of the hydraulic heads in the 35 measuring stations computed using a complete 3D flow model and the 2D equivalent transmissivity model.

ductivity field and the geometry of the aquifers, and consequently the resulting transmissivity, are corrected. The other parameters are kept constant. They could be included in the data assimilation procedure in the future to improve the performance of the model.

Validation and Indicators

In applying data assimilation methodology to real data, unlike a synthetic test, the true geology, parameter fields, and uncertainty are unknown. To quantitatively measure the validity of our results, we implemented serveral indicators.

The first indicator used to benchmark the quality of the models is the data misfit. It benchmarks the agreement between the predicted data and the observations. This indicator is not used during the inversion since ES-MDA does not need an objective function. But, as mentioned before, ES-MDA will try to minimize a combination of terms that includes the misfit. It is used here to assess the posterior compatibility of the data. The residual at a given iteration is :

$$DataMisfit = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\frac{(d_{obs,i} - d_{frw,i})^2}{\sigma_{d,i}^2}} \tag{6.7}$$

where d_obs is the observed data, d_fwr is the simulated data, σ_d is the uncertainty of the observed data and N is the total number of data point. The indicator is dimensionless and allows us to analyze different types of data on a comparable basis.

To have a better representation of the permorfance of the parameters prediction, the next indicators rely on a cross-validation algorithm based on the borehole data. This involves selectively withholding a portion of the conditioning dataset from the simulation to create a basis for comparison and evaluate the prediction quality against these intentionally unaccounted-for boreholes. However, the spatial correlation inherent in the data means that a random selection of boreholes could lead to an overestimation of algorithm performance if the omitted boreholes are close to each other from a different fold. To counteract this, we employ the K-Means clustering algorithm (Hartigan et al. 1979) to the spatial coordinates of the boreholes, segregating them into 20 groups. This ensures boreholes in close proximity fall into the same group, providing a more accurate representation of the algorithm's performance.

From these 20 groups, we randomly assemble five similar-sized folds. For each fold, we generate a unique ArchPy prior ensemble, each time excluding the data of one fold, accounting for approximately 20% of the total data. Following this, we apply the data aggregation algorithm independently to all five sets, allowing for a comprehensive evaluation of its performance. The results of each set are then compared against their corresponding excluded boreholes to assess the prediction quality. These models are created only for the purpose of quality assessment and function independently from the complete models, which incorporate all available borehole data. This separation guarantees an unbiased evaluation of performance while ensuring that the final models yield the most accurate representation possible, given the available data.

Since our main objective is to illustrate and test the application of multiple data types assimilation to hydrogeology, we will also benchmark the performance of our models to outline the boundary of the aquifer. Therefore, we define three quality indicators for each borehole: the mean error, the mean normalized error, the Continuous Rank Probability Score (CRPS). The three indicators are applied on the predicition of the thickness of the upper aquifer, which corresponds to the altitude of the surface that defines the limit between the Unit 1 and the Unit 2. This limit is

also the one the most frequently defined in the boreholes. The mean error allows controlling that there is no constant positive or negative bias in the prediction and that the correct value is predicted on average. The mean error is defined as :

$$\epsilon_{x,y}^{k} = \frac{1}{N}\sum_{i=1}^{n}(z_{i,x,y}^{k} - z_{x,y}^{borehole}) \tag{6.8}$$

where $\epsilon_{x,y}^{k}$ is the mean error in iteration k at position x and y. N is the number of members, $z_{i,x,y}^{k}$ is the prediction of the thickness at the position x, y in the model i and $z_{x,y}^{borehole}$ is the true thickness of the aquifer from the description of the well. Therefore, the error is defined at each borehole position and should converge to zero on average. The second indicator is the normalized error. It is defined as the mean error divided point-to-point by the experimental variance of the simulation results. It shows if the predicted magnitude of the uncertainty is comparable with the actual ensemble errors, assuming a Gaussian distribution of errors. Normalized error is then defined as :

$$E_{x,y}^{k} = \frac{(\epsilon_{x,y}^{k})^2}{(\sigma_{x,y}^{k})^2} \tag{6.9}$$

where $E_{x,y}^{k}$ is the normalized error in iteration k at position x and y, and $\sigma_{x,y}^{k}$ is the experimental standard deviation calculated over the simulation results of the n members at position x and y. The normalized error should converge on average to 1, meaning that the uncertainty is in the same order of magnitude as the ensemble error. Finally, we used the Continuous Rank Probability Score (CRPS) (Gneiting et al. 2007). The CRPS is sensitive to both bias and uncertainty. It is defined as being the integral between the Cumulative Distribution Function (CDF) of the ensemble of the predictions at a given point and the step-function CDF of the discrete true value at the same point. It is expressed as :

$$CRPS_{x,y}^{k}(F, z_{x,y}^{borehole}) = \int_{-\infty}^{+\infty}\left(F(y) - \mathbb{1}(y > z_{x,y}^{borehole})\right)^2 dy \tag{6.10}$$

where $\mathbb{1}$ is the Heaviside function, so that $\mathbb{1}(y > z_{x,y}^{borehole}) = 1$ if $(y > z_{x,y}^{borehole})$ and 0 otherwise, and F is the Cumulative Density function of the ensemble of predictions at a given point. The minimum CRPS score is zero, and is reached if the prediction is correct with no uncertainty. If the prediction has a small uncertainty and is not biased, the score will be relatively small. On the contrary, as soon as the prediction is

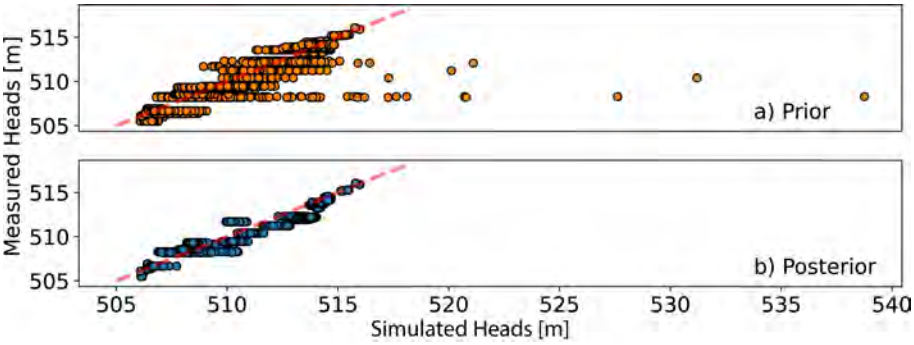

Figure 6.7: Simulated Heads versus Measured Heads in the 35 groundwater monitoring stations for the a) 500 Prior models and b) 500 posterior models. The dashed line is the perfect prediction.

biased or the uncertainty large, the CRPS increase. Unlike normalized error, CRPS has the advantage of not assuming a normal distribution of errors. All these indicators are point indicators, and will be calculated on all excluded boreholes positions for all five folds.

6.4 Results

500 initial members were generated using ArchPy. This set reflects the prior knowledge and will be referred to later as the prior. Figure 6.8 shows the evolution of the observation misfit thought the iterations. 10 ESMDA iterations were carried out. Iteration 0 corresponds to the prior observations. We denote a global decrease of the residual, with a mean total joint residual of the data (geophysical and hydrogeological) decreasing from 15.1 to 4.1. Of course, since the number of geophysical data points is much higher than the number of hydrological measurements, the total misfit closely follows the geophysical misfit. However, the hydrological misfit also shows a decrease in terms of variability and average value. We can see that the residual tends to greatly reduced in the first iterations, meaning that good results could probably be achieved in less than 10 iterations.

The computing time per iteration is about 4h20, for about 3 million active cells. On this total time, about 75% is dedicated to the forward calculation and the rest to the update of the model. The ensemble of models after data assimilation will be referred to as posterior. Fig. 6.7

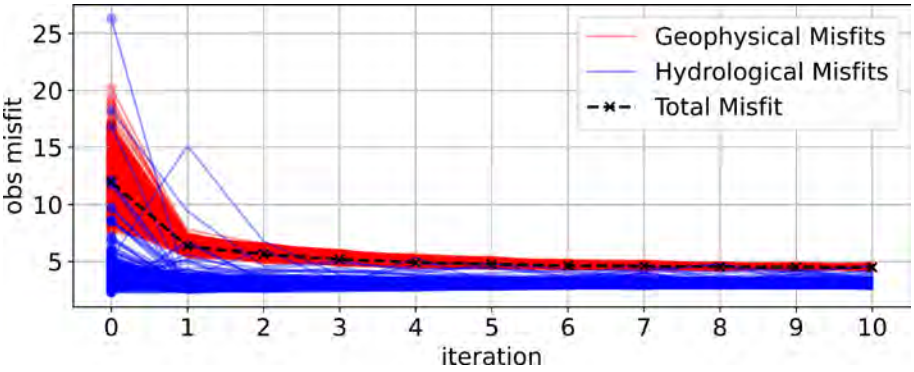

Figure 6.8: Evolution of misfit between the simulated observations and the measured observations. Iteration 0 corresponds to the prior models simulation

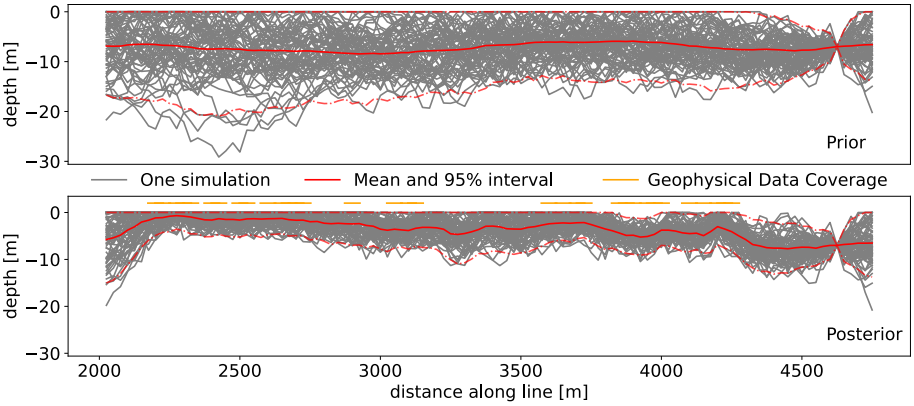

Figure 6.9: Example of a cross-section through the prior and posterior models showing the depth of the aquifer along a 2.5km line. The ensemble of simulations is shown in gray, and the arithmetic mean is shown in red with the 95% confidence interval envelope (2.5% and 97.5% quantile of the ensemble). The decreasing uncertainty at 4600 m is due to a conditioning borehole.

compares the heads measured in the valley with the heads predicted from the 500 models at the same position, both prior and posterior. A decrease in terms of variability of the heads can be denoted, and the final ensemble of prediction is closer to the perfect prediction line. It suggests that the posterior ensemble is capable of predicting heads in the domain fairly well.

Fig. 6.9 shows a cross-section through the model with all the boreholes taken into account. 75 members drawn out of the 500 of the prior and the posterior are compared in terms of the depth of the upper aquifer. The reduced uncertainty at the 4'600 m position is due to the presence of a conditioning borehole and is therefore consistently present in both prior and posterior sets. The locations of the geophysical EM data are outlined in yellow in the posterior cross-section. Only the points that are collocated are shown, but the effect of neighboring points can occur. We denote a clear reduction in uncertainty in the posterior compared to the prior, where geophysical data have been acquired. The confidence interval is significantly reduced. Where no geophysical data is present, the uncertainty logically tends to be equivalent to the one in the prior. When applying ESMDA, one common risk is the collapse of the posterior, where all members converge to a similar model, potentially leading to an underestimation of uncertainty. In the present case, despite 11 iterations, each model in the posterior ensemble exhibits notable differences, suggesting that the uncertainty estimation is reasonably accurate. However, it is important to acknowledge that validating the correctness of uncertainty estimates can be challenging.

One clear advantage of the presented methodology is that the final models are hierarchical. For example, for most applications in hydrology, the thickness and extent of the aquifer are of key interest. Unlike the 3D electrical resistivity field, in this workflow, the thickness of the aquifer is part of the model and can be easily extracted. Fig. 6.10 shows the mean depth of transition between the shallow aquifer and the aquiclude over the 500 members. The estimated uncertainty is also outlined. We can denote the zero uncertainty points in both the posterior and the prior, which correspond to the borehole positions where the depth is considered certain. The results appear to be spatially consistent. First, the variations in terms of aquifer depth show continuity over the range of a data point update in the Localization Matrix, and therefore beyond the effect of a geophysical point. The critical distance was set to 150 meters, which means that above twice this distance, for example, a geophysical measurement cannot have an effect during the update. Large areas with lower or higher depth are due to information that is actually brought by the data. It suggests that distant geophysical data are consistent and are successfully incorporated during the ESMDA inversion. Secondly, the continuous lower or higher depths are elongated in a SE-NW orientation, which is consistent with the deposition phases estimated at this site, de-

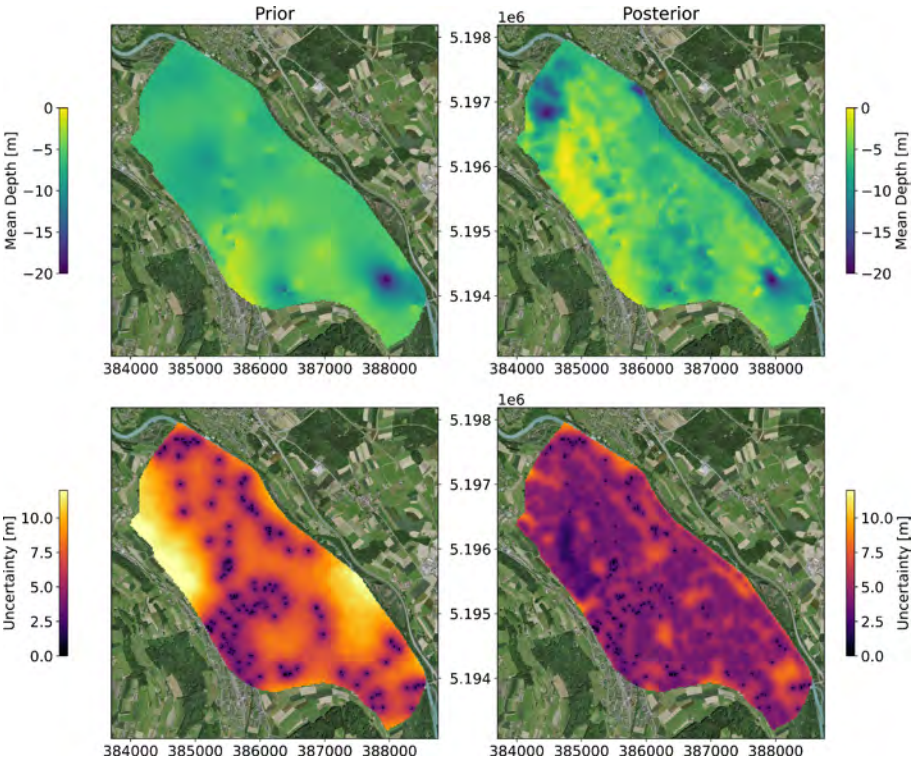

Figure 6.10: Mean depth of transition between the shallow aquifer and the aquiclude over the 500 posterior simulations (after data integration) and prior simulations (boreholes only), and associated uncertainty. The uncertainty at the location of the boreholes. The coordinates are in UTM32 (epsg: 32632). Basemap freely obtained from the Swiss Federal Office of Topography.

spite the fact that the variograms used for the generation of the prior are isotropic. Third, we note that the uncertainty is reduced where and only where geophysical data are present and kept unchanged far from them. Finally, the small artifacts that we can see mainly in the Prior are due to neighboring wells that carry a significant difference in terms of aquifer depth, probably unrealistic due to their close position from each other, and cause an abrupt change.

Cross-validation was performed using 5 distinct folds, 500 members, and 10 iterations each. The variable cross-validated here is the thickness of the first aquifer. The aggregated results are shown in Fig. 6.11. Each fold

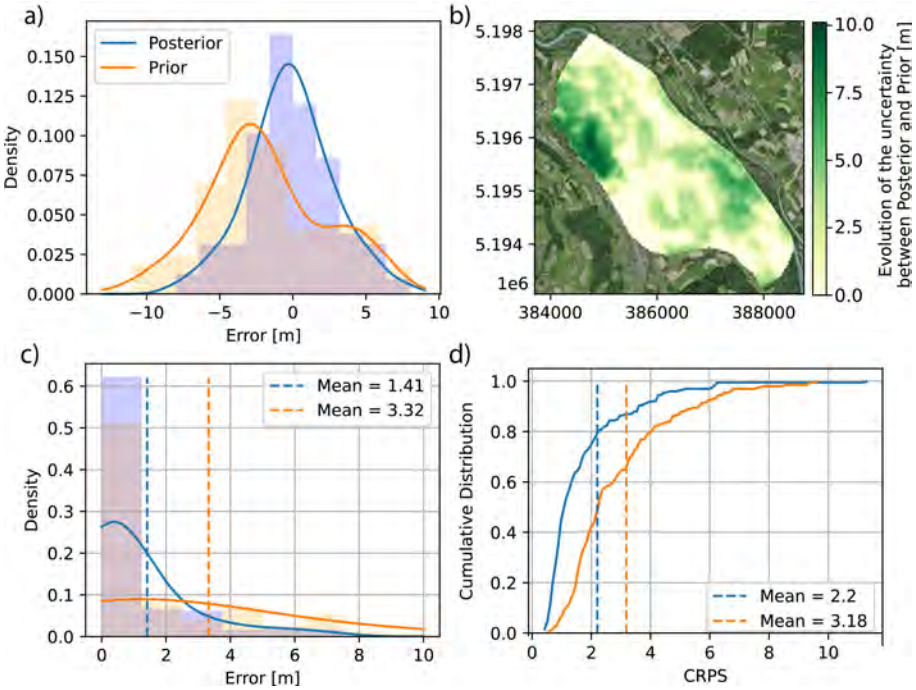

Figure 6.11: Results of the Cross-Validation: a) Error distribution for prior and posterior models b) Evolution of the uncertainty (2σ) between the posterior and the prior. A Positive value means a reduced uncertainty in the posterior, where a null value means that so significant changes has occurred. c) Distribution of the Normalized error d) Cumulative distribution of the Continuous Ranked Probability Score (CRPS).

is amputated of 20% of the borehole dataset. The purpose is not only to benchmark the results of the data integration algorithm, but also to compare them with the borehole-only approach. For this reason, cross-validation was also performed using only prior knowledge before data integration. Several results can be highlighted. First, compared to the prior models, the error range is smaller and is centered on zero (see Fig. 6.11a). This unbiased error is a critical point and suggests that, unlike the prior, we tend to equally over- and underestimate the thickness of the aquifer. This proves that this methodology can perform well even if the prior is uncertain or not totally correct. Fig. 6.11b shows that a global reduction in uncertainty is clearly visible almost everywhere in the domain,especially in places that are surrounded or covered by tTEM points. No negative change in uncertainty was denoted between the prior

and the posterior. Despite this global decrease in uncertainty, the normalized error is improved (Fig. 6.11c). The combination of these two indicators shows that the uncertainty in the posterior is smaller and is better scaled to the ensemble error. Furthermore, some remaining outliers points are in fact far from any measurements and consequently will only be slightly changed between prior and posterior. The distribution is asymmetric; with most of the values being lower than one; it suggests that in some locations the data assimilation algorithm still tends to slightly overestimate the uncertainty.

Unlike the normalized error, the CRPS shown in Fig. 6.11d) has a target value that is also the arithmetic minimum: 0. CRPS is used to compare a probabilistic prediction with a deterministic truth by integrating their cumulative density function. The CRPS has the advantage of not assuming a Gaussian distribution of the error and considers the probabilistic forecast as a whole. In our case, we denote that the CRPS is distinctly better in the posterior than in the prior, both in terms of global distribution or in terms of arithmetic mean. We can then assume that the posterior models give a better forecast in terms of aquifer thickness than a borehole-only-based model, without any manual integration of geophysical data.

In the above results, the thickness of the upper aquifer was used since it is the information that is the most present in the boreholes and can therefore be used for cross-validation and comparison. However, our data integration in the discrete geological model approach can also provide various other information. First, the distribution of the petrophysical parameters can be retrieved. Figure 6.13 shows an E-W cross section through one posterior realization (Fig. 6.13 a and c) and the mean petrophysical fields (Fig. 6.13 b and d). Since the petrophysical models are generated from the discrete unit model, where each cell is assigned a specific layer, we denote sharp transitions between the different facies. On the other hand, the mean realization is much smoother, since it is averaged over the members. For the resistivity mean cross section (Fig. 6.13d), we can see that averaged smooth petrophysical properties are closer to what can be achieved with classic deterministic inversion, compared to a single model (Fig. 6.13 c). Figure 6.13 e) compares the smooth inversion of the same tTEM data published by Neven et al. (2021). To facilitate the comparison of the structures, the mean resistivity field over the realiza-

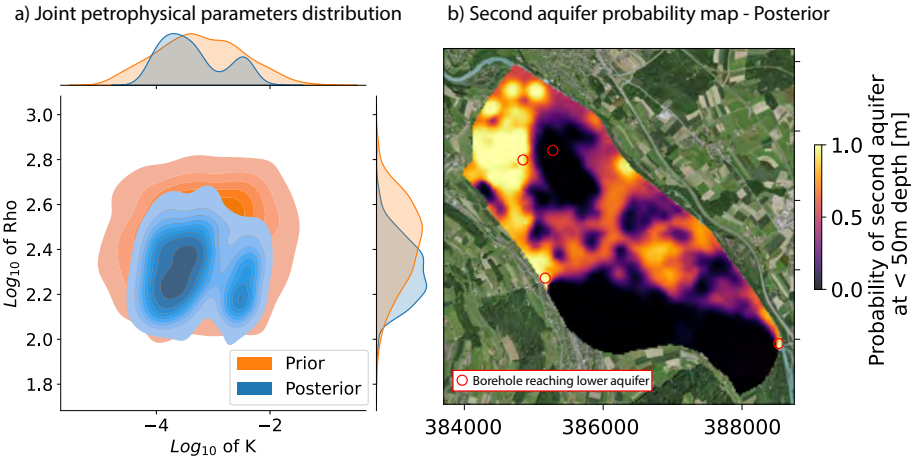

Figure 6.12: a) Joint Posterior Distribution of the two petrophysical parameters of interest in the aquifer: Electrical Resistivity (Rho) and Hydraulic Conductivity (K). b) Probability of the presence of the second aquifer at a depth less than 50 m with the position of the boreholes reaching the second aquifer.

tions is added in grayscale. The agreement between the mean resistivity field and the 1D resistivity models is satisfactory.

Figure 6.12 a) shows the joint distribution of resistivity versus hydraulic conductivity in the upper aquifer after inversion. As a reminder, in the prior both proprieties were assumed to be normally distributed around uniformly distributed random means drawn at the pilot points position. Recall that the uniform range for the pilot points was for resistivity and hydraulic conductivity, respectively, $100-500 (10^2-10^{2.7}) \Omega \cdot m$ and $10^{-5}-10^{-1.5}\ m \cdot s^{-1}$. The resulting prior petrophysical distributions for the upper aquifer are shown in orange in Fig. 6.12a. Logically, the maximum of the prior distribution of parameters corresponds roughly to the average of the initial pilot point values. The predicted value range in the posterior is smaller compared to the large prior range. Interestingly, the results of the data integration show two distinct populations not present in the prior (see Fig. 6.12a). We can denote that the two populations are clear in terms of hydraulic conductivity but overlap in terms of resistivity. The values for both resistivity and hydraulic conductivity are within the expected range in the area. Finally, Fig. 6.12b shows the probability of encountering the second aquifer above a certain depth. Since only

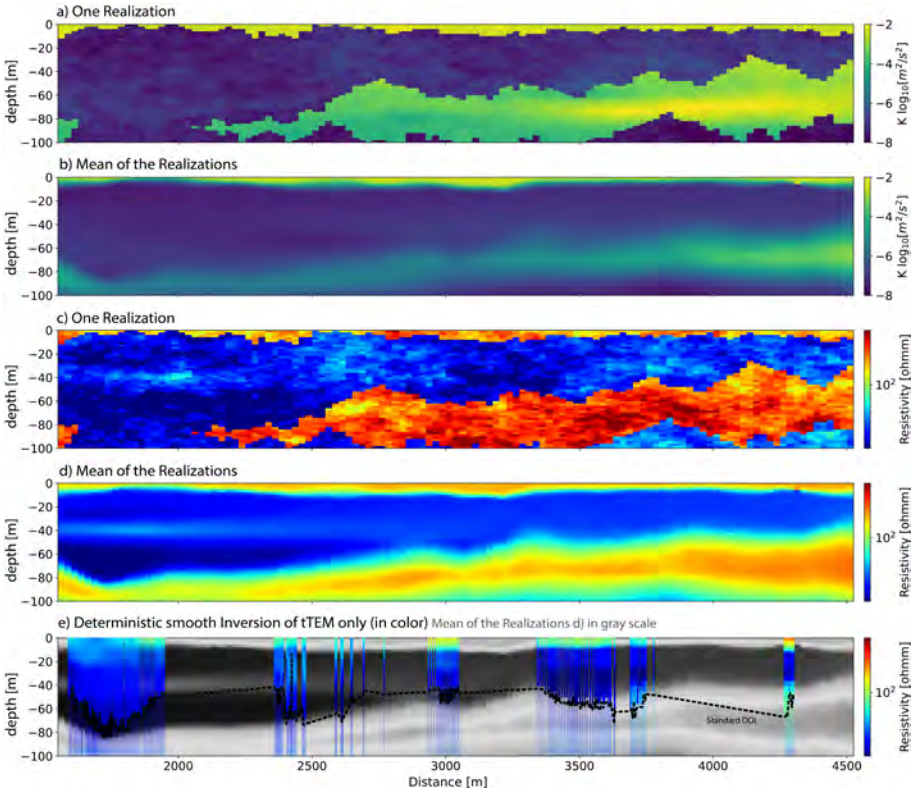

Figure 6.13: E-W Cross-Section (y = 5'194'568 m UTM32N) comparing the hydraulic conductivity field of one posterior model (a), the mean hydraulic conductivity over the posterior members (b), the resistivity field of one posterior model (c), the mean resistivity over the posterior members (d), and the 1D resistivity models from a deterministic smooth inversion of the tTEM data (e) from Neven et al. (2021). The dashed black line is the standard depth of investigation, retrieved from the inversion. The gray scale base map of (e) is the mean resistivity over the posterior members (d).

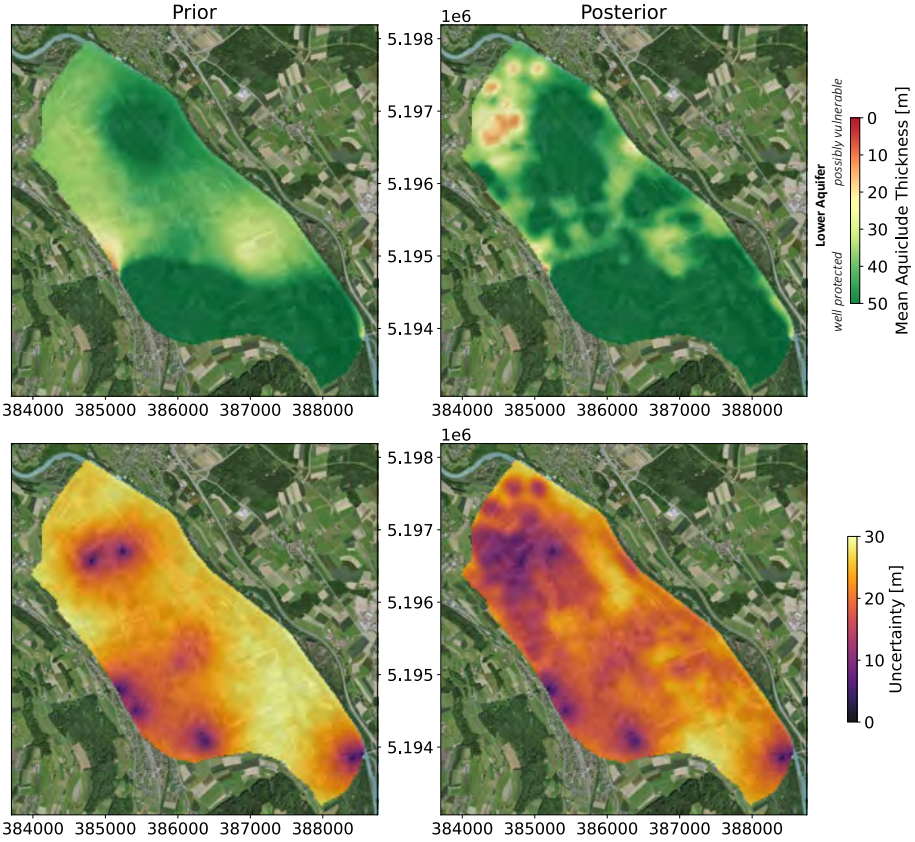

Figure 6.14: Thickness of the aquiclude, acting as the separation of the surface and shallow aquifers. Low thicknesses are correlated with a higher risk of pollutant contamination from the shallow aquifer to the lower protected aquifer. Zero-incertitude points are boreholes. The coordinates are in UTM32 (epsg: 32632). Basemap freely obtained from the Swiss Federal Office of Topography.

4 boreholes reach this second aquifer, the probability in the prior is almost uniform at 50%. The data aggregation algorithm has provided a significant amount of information, reducing the probability in some areas of the valley and predicting with high confidence the presence of the aquifer in other places.

Another interesting result is the possibility to assess the vulnerability of the second aquifer. Shallow aquifers in quaternary settings are generally more affected by pesticides, fertilizers, or permanent contaminants

(PFAS for example) than the deeper ones. In Switzerland, these problems lead some water authorities to start targeting deeper aquifers to produce drinking water. Quantifying the potential connection between the shallow and deeper aquifers is therefore an important practical issue. The modeling framework proposed in this paper allows estimating this thickness in every simulation and post-processing them to obtain some statistical estimates. Fig. 6.14, summarizes such results. The thinner the aquiclude, the higher the risk of communication between the upper and lower aquifer, and consequently the vulnerability of the second aquifer to surface pollutants. Fig. 6.14 shows that the uncertainty is relatively high in the set of prior simulations. It is low only around the seven borehole locations. Note that among these boreholes only three are touching the deeper aquifer and are used as conditioning points, while the others have not reached the second aquifer and are used as inequalities in the conditioning of the geostatistical simulations. The use of the EM data allows to constrain better the thickness map and reduce significantly the uncertainty (Posterior results in Fig. 6.14). But we observe that the posterior uncertainty still remains much higher than the uncertainty of the thickness of the shallow aquifer. The uncertainty reduction is lower, this is certainly due to the higher depth and the presence of a thick clay layer because the EM method is only poorly sensitive to structures at a large depth.

Finally, we show the effect of the data assimilation on different particle tracking scenarios to illustrate the impact of the method on possible solute transport predictions. The prior and posterior models are compared. Using the simulated heads of the groundwater model, the advective path is calculated for each of the members. We used the same hydrogeological forward model as the one used for inversion. As shown earlier (Fig. 6.7), the hydraulic heads are much less dispersed in the posterior simulations. Fig. 6.15 shows four situations with a different initial location for the contamination source and shows their respective prior and posterior advective paths. In all the cases, the uncertainty has been reduced by integrating the geophysical and groundwater data. We observe that all the posterior paths are contained within the prior. This suggests that a model based only on boreholes could be used to estimate the travel path uncertainty, but the uncertainty would likely be much too broad. For example, in the case of scenarios A and B (Fig. 6.15A and B), the river in which the pollutant may appear is not identified with confidence in the prior since it could reach either the Aare river on the east or the

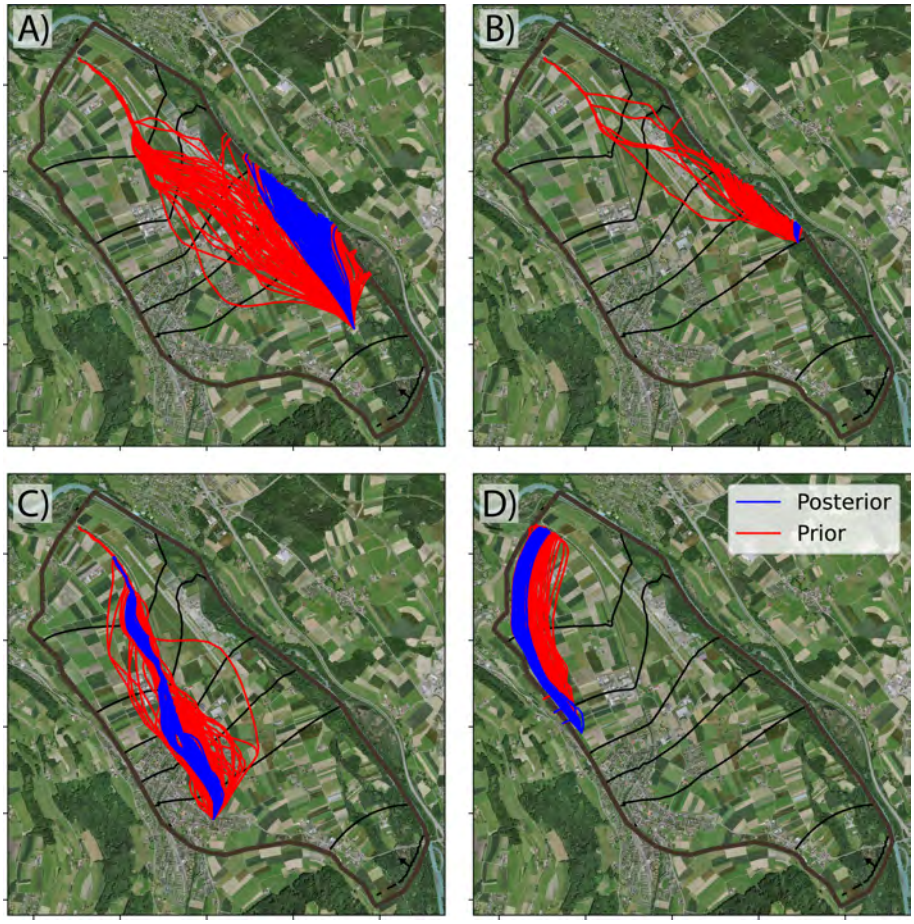

Figure 6.15: Advective particle path computed using the Prior and Posterior models for 4 different starting sites. For readability, only 100 members are represented.

Grube in the middle of the domain when the posterior distribution has clearly identified the most probable case. Fig. 6.15C-D illustrates how the thicker and more permeable aquifer areas identified through the assimilation of the data affect the path. The thicker aquifer northwest of the area (see Fig. 6.10) drains most of the particles along an impermeable boundary of the valley.

6.5 Discussion

The methodology and results presented in this study have proven the feasibility of integrating geological, geophysical, and hydrogeological field data to obtain consistent models at the scale of a part of a regional aquifer. The use of two different data types through joint inversion has proven to be a valuable approach. The two distinct hydraulic conductivity populations (Fig. 6.12a) probably correspond to two distinct subaquifer regional granulometry changes, which is typical of highly heterogeneous Quaternary environments. The two populations cannot be distinguished from resistivity only, since their respective ranges overlap, despite an order-of-magnitude difference in hydraulic conductivity. However, geophysical measurements constrain the thickness of the aquifer. Consequently, the space of uncertainty of the transmissivity is reduced, and the space of uncertainty for the hydraulic conductivity is limited. Without joint inversion, it is probable that these areas would have been considered to be a shallower aquifer section, instead of a less permeable one.

To accelerate the forward computations, we used a simplified 2D groundwater model. This is not a requirement for the proposed methodology, but it constitutes a reasonable simplification in the aquifer that we studied and with the available integrated head data. In a thicker aquifer with precise head measurements collected at different depths, this 2D approach would not be adequate to represent the system since the vertical flow components across the geological formations may have an important impact. In these situations, a complete 3D groundwater model would be required. Similarly, the use of 1D EM forward for inversion is widespread but significant differences can occur between 2D or 3D modeling in the case of complex geology such as seawater intrusion for example (Kang et al. 2015). However, our method lies in between since even if a 1D forward code is used, the underlying model is still 3D and has some constrains.

The proposed workflow has the advantage of being flexible and could easily be extended to account for more physical processes and parameters. The data assimilation method could include other types of geophysical or groundwater measurements and allow identifying additional model parameters such as storativity, porosity, or river conductance, as well as boundary conditions or source terms. Combining more information would require developing appropriate prior distributions for all these additional unknowns but could help to better quantify the overall uncertainty in groundwater models. The downside could be that adding more parameters, especially global ones, could lead to more spurious correlations in the ESMDA procedure and could also increase significantly the computing time required for the assimilation algorithm.

Compared to a classical geophysical inversion, the data residual of the proposed approach is much higher (4.1 versus less than 1 in the previous inversion by Neven et al. (2021)). Indeed, the classical deterministic inversion is only constrained by the regularization applied to the inversion algorithm. These fairly loose constraints allow the minimization scheme to fit the data almost perfectly. In contrast, our inversion scheme involves much more prior knowledge, such as the number of layers, the distribution of the parameters, and their spatial continuity. The final models are therefore the best models that fit both all the data and the conceptual knowledge of the area. For example, some local areas may show a thin additional layer of clay within the upper aquifer; this is not considered by the conceptual model and will result in a higher residual close to the point. Additionally, as shown in Figure 6.13, the mean of the realizations is very similar to the deterministic inversion. It is a good indicator that the space of uncertainty around this mean optimum is correctly represented by the ensemble of posterior models. It is also possible that the variogram models chosen for the parameter fields simulation are not suited, or that the number of pilot point should be increased. The balance between the representation of the detailed complexity and the large-scale structures always depends on the purpose of the final model, and one should be aware that the prior model will need to reflect the level of expected details. For hydrological observations, we denote only a marginal decrease of the residual. The fact that only a steady state model was used limits the amount of information which can be gathered from the hydrological data. The use of dynamic data such as tracer test or piezometric time series could probably increase the contribution of the hydraulic data.

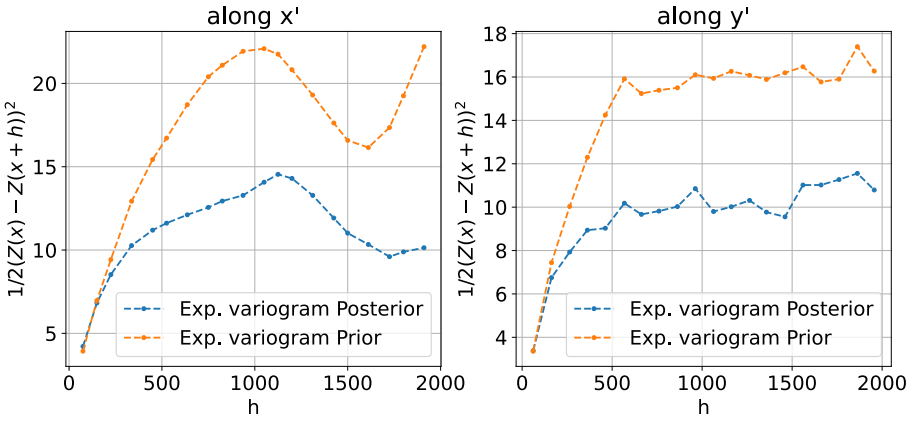

Figure 6.16: Experimentals variogram on the bottom of the aquifer surface (surface 1) along the X and Y axis for the posterior and the prior.

We do not assume a specific petrophysical relationship a priori. The geostatistical simulations for each petrophysical parameter are independent. The only prior knowledge that links resistivity and hydraulic conductivity comes from the prior distribution of the mean values for each layer (and rock types) drawn at the pilot point locations. This means that a permeable layer will always be associated with high resistivities; however, extremely high values within the high aquifer resistivity range will not necessarily be associated with high permeability within the aquifer values range.

The cross-validation performed on the prior models has revealed a bias of about -4 meters on average. Even if this error might seem negligible, for an aquifer that has a mean thickness of about 5-6 meters, it represents an important bias in terms of available resources. We interpret this result as a classical example of sampling bias and should be carefully considered when making geological models. The prior mean depth used for the geostatistical simulation is calculated by averaging the available thickness data from the boreholes. However, when the aquifer is deep, most drilled wells for geothermal heat pumps or irrigation do not reach the bottom of the aquifer. Therefore, the exact thickness of the aquifer is unknown at these locations. This leads to a sampling bias toward smaller thicknesses and consequently to a bias in the prior model. We note that the ESMDA approach successfully identified the bias and recenter the distribution. Choosing a correct prior when bias exists in the available data is a general issue when dealing with Bayesian approaches, and was previously

noted in the review by Linde et al. (2015). Properly capturing the heterogeneity and statistical parameters from the wells is straightforward with variogram-based methods and the simplicity of the approach makes the transcription of prior conceptual knowledge into statistics easier than with other more advanced object-based or multiple-point statistics modeling methods. Combining variogram-based geostatistical methods with inversion offers a good balance between simplicity and realism. If the prior is not totally statistically correct (such as the bias in the mean altitude of a geological interface in this study), data assimilation may correct it. The variogram of the posterior distribution of the base surface of the upper aquifer was compared to the original variogram model from the prior simulation in Fig. 6.16. It reveals that data aggregation has significantly changed the variogram, its sill has been reduced by a factor of almost two. This posterior variogram could help to scale accordingly similar areas that are not covered by geophysical measurement by assuming that the statistics are similar. Getting reliable and data-based variogram models in geology is always difficult, and applying this methodology to other environments could create potential databases of analogs.

As mentioned above, the total computing time was about 4h20 per iteration. The complete workflow was parallelized on 30 CPUs. Despite the reduced-dimensionality strategy, the forward simulations still represent 75% of the computing time. To accelerate even further the algorithm, one possible future research direction could be to employ non-physical surrogate modeling, such as, for example, machine learning (Bording et al. 2021; Juda et al. 2021). Using a less computationally expensive approach for the first iterations, or the complete inversion, could significantly speed up the computation.

6.6 Conclusion

This paper has shown the feasibility of integrating geological observations, an explicit geological concept, hydrogeological, and geophysical data on a shallow quaternary aquifer. Using the Ensemble Smoother for Multiple Data Assimilation (ESMDA) algorithm, the methodology is able to effectively integrate all these data and provide robust uncertainty estimation.

The approach consists in generating and updating a consistent ensemble of 3D geological models that are obtained using the ArchPy hierarchical approach. To accelerate the ESMDA computation, we propose to run the forward physical simulations in a reduced dimension that is different for the geophysical and hydrogeological problems. In addition, the nature of the ESMDA algorithm allows for benefiting heavily from parallel computing. The application of the methodology to the upper Aare Valley in Switzerland, with actual field data, demonstrates the applicability of the proposed workflow in real situations. The stochastic nature of the method and the existence of the underlying complete 3D geological model allows for estimating uncertainty, with the advantage of being able to retrieve not only petrophysical fields but also some underlying quantities derived from the model, such as the thickness of the aquifer, or the probability of occurrence of a protection layer above an aquifer. This feature can be used to deliver specific critical information for decision-makers which would not be simple to derive at the aquifer scale from geophysical data only.

To conclude, the results presented in this paper demonstrate that coupling the ArchPy methodology with ESMDA is a promising approach that could be applied to a wide variety of problems. All the components of this methodology are available as open-source Python-based software that can easily be modified. For example, any other forward simulator could be coupled with the existing code to extend it to a wide variety of situations to constrain the geological models with other types of geophysical observations or other hydrological processes such as heat or solute transport.

Additional information

Data availability The tTEM data are published in Neven et al. (2021). Most of the boreholes are distributed by the Canton of Bern (`http://www.agi.dij.be.ch`). Some part of it is distributed by Swisstopo, through the GeoQuat project (Volken et al. 2016), and can be obtained upon request for research purpose. To facilitate reproducibility, the ArchPy models (prior and posterior) and the flow model used for the flow simulation can be found in the zenodo repository (`https://zenodo.org/record/8047723`,Neven et al. (2023)).

Accession codes The geostatistical codes used in this study are implemented in the python *ArchPy* package freely available at `https://github.com/randlab/archpy`. The 1D forward code used is *AarhusInv* developped by the Aarhus University Hydrogeophysics group (Auken et al. 2015). The AarhusInv code is free to use for research purpose. The flow model code used in the open-source *ModFlow 6* code (Hughes et al. 2017).

Chapter 7

Complete Valley Inversion - Proof of Feasability

7.1 Introduction

The main objective of the work presented in the previous chapter is the inversion of hydrological and geophysical data within the same geologically consistent framework. Of course, the ultimate goal is to aggregate all the data available in the valley, and not only in the Belp area (see previous chapter). However, at the time of this thesis, the realization of the complete hydrological concept of the valley is still in development. Instead of using a hydrological model that is not well adapted to the site, we decided to work only on a subset of the area where the hydrological conceptual model was already defined. For this reason, the complete joint inversion study was only performed in the Belp area, towards the north of the valley, where the boundary conditions and the hydrological concept were better defined. The results were presented in the previous chapter. However, to prove the feasibility and scalability of the method, an inversion was carried out throughout the valley, using all the boreholes in the area and the tTEM soundings. The site of interest in the complete Upper Aare Valley, between the cities of Thun to the South, and Kehrsatz to the North. Hydrological measurements are not taken into account in this model. This short chapter only shows preliminary results of the geophysical inversion only, and briefly comment on them.

7.2 Preliminary Results

All the methodology described in the previous chapter is kept unchanged; for this reason, no methodology will be detailed here. The model is slightly coarser compared to the Belp area, with a resolution of 50 m x 50 m x 4 m. The parameters and the number of layers are kept similar. We expect that the conceptual geological conceptual model will not change drastically between the north and the south. The model includes 1430 wells and all geophysical soundings within the area of interest. However, the soundings were downsampled to the model resolution. Only one sounding is kept per model cell in this preliminary example. Although even if nearly all of the acquisition area has been considered, this results in the inclusion of only a small fraction of the available soundings. Specifically, fewer than 10,000 soundings are included out of a total of over 50,000. Figure 7.1 shows the spatial distribution of data within the valley. The model is 24 km long and ranges between 1 and 3 km in width. The total surface modeled is about $48 km^2$. As denoted before in the Belp area, there is a clear inverse correlation between the position

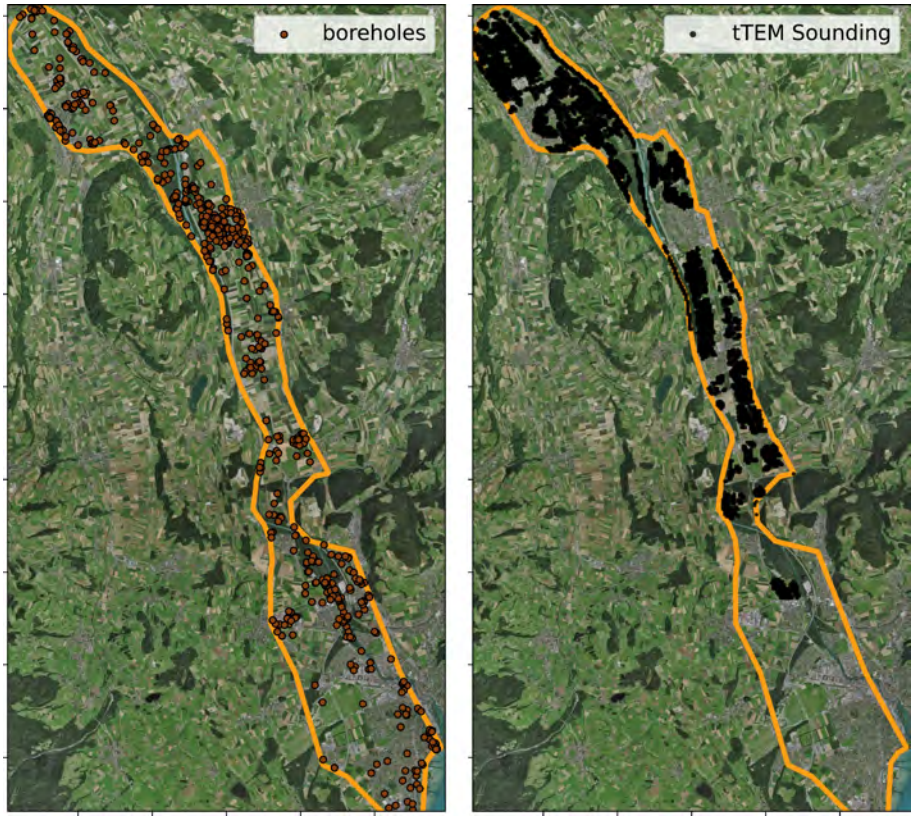

Figure 7.1: Aare Valley Modeling Area, with the Position of the Boreholes and the tTEM used for the inversion

of the tTEM and the boreholes. The city of Munsingen, for example, is clearly visible with a really high density of boreholes and no tTEM coverage. The tTEM aquisition stopped toward the south before the city of Thun, because most of the remaining area is military-owned. Since the model is larger, the inversion computing time was about 44 hours on a 30 CPU machine.

Figure 7.2 A) shows a cross section with 100 realizations after 10 iterations of ES-MDA, and the statistics of the bottom of the upper aquifer. The areas where geophysical data were acquired and the position of the cross section are outlined in 7.2 C). We can denote that, similarly to the Belp Area, the ES-MDA algorithm has successfully reduced the uncertainty where no data were available in prior. The prior model shows fairly uniform mean and standard deviation through the section, sug-

gesting that no conditionning boreholes are close, and that the ensemble distribution converges to the mean and variance of the statistical model. An interesting point to note in the section is that the posterior surface ensemble is mainly outside the prior distribution. This was also noted in the Belp area, but is even more clear in the present example. This suggests that the prior models calculated from the boreholes probably underestimate the thickness of the aquifer significantly beyond its uncertainty due to an unadapted prior. This is probably another example of sampling bias in the data. The areas where the thickness of the aquifer is the most important tend to be unmapped because the boreholes do not reach the bottom of it. Since the statistical model was fited at the depth derived from the boreholes, the same bias will propagate in the model. This issue is a common issue in statistical modeling, and no solution exists if there is additional prior hard knowledge or the use of informed assumptions.

Figure 7.2 B) shows the point-to-point probability of being in an aquifer, calculated over 510 members, compared to the prior knowledge. Probability is calculated by dividing the number of times a certain cell is simulated as being part of an aquifer by the total number of simulations. As mentioned above, the prior is poorly informed in the area, with almost no boreholes. The probability distribution converges to the statistical prior. This is why the bottom aquifer probability seems flat around the depth of -70m, which is the average depth at which boreholes hit the lower aquifer in the valley. The posterior shows significant changes, both in terms of the upper and lower aquifer distribution. The upper aquifer becomes extremely well resolved, where the lower aquifer position seems to be more probable on one side of the valley.

Figure 7.3 shows the average thickness of the upper aquifer in the upper Aare Valley, and the associated uncertainty. As expected, no changes in the model are visible on the southern side of the model. No geophysical data are present in these areas, and therefore no update was conducted between the prior and posterior. This proves that the use of a localization matrix as a way to reduce the potential spurious correlation was successful. The uncertainty is reduced where geophysical data is present. We denote a clear N-S trend in terms of the depth of the aquifer. This trend was established before using boreholes (see Fig. 7.4) but the integration of geophysical data gives an unprecended look at its extent. Figure 7.4 only shows the boreholes that reached the bottom of the aquifer (equal-

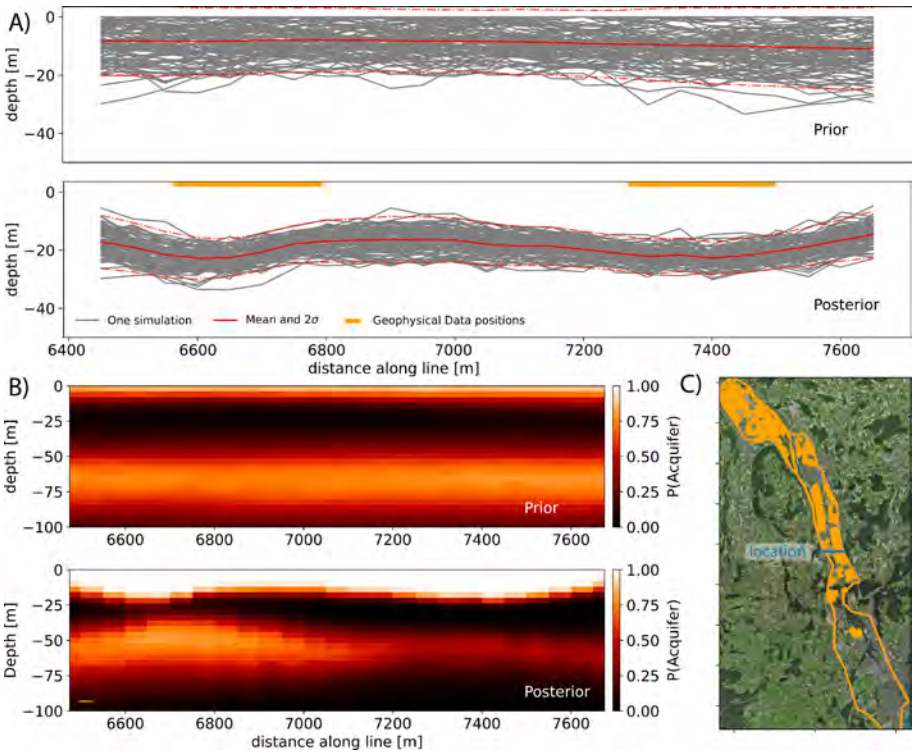

Figure 7.2: A) Mean upper aquifer depth and several realizations along a section in the Upper Aare Valley and comparison between the posterior and the prior. B) Probability of presence of the aquifer facies (Sand or Gravel) along a section in the Upper Aare Valley, and comparison between the posterior and the prior. C) Map of the area and cross section position (blue line) with the tTEM coverage in orange.

ity points) and their value. On the right side of the figure, the posterior model is added. As mentioned several times through this thesis, we observe that the boreholes are not regularly distributed in the area. We clearly see the Munsingen area, where many wells are concentrated, for example. However, we can also denote that deeper aquifer areas in the posterior tend to be unsampled by the boreholes. It is possible that this is an example of sampling bias. The assumption is that outside of the cities, the probability of a well being drilled is spatially roughly uniform, but the probability that this well reaches the bottom of the aquifer is inversely proportional to the aquifer thickness.

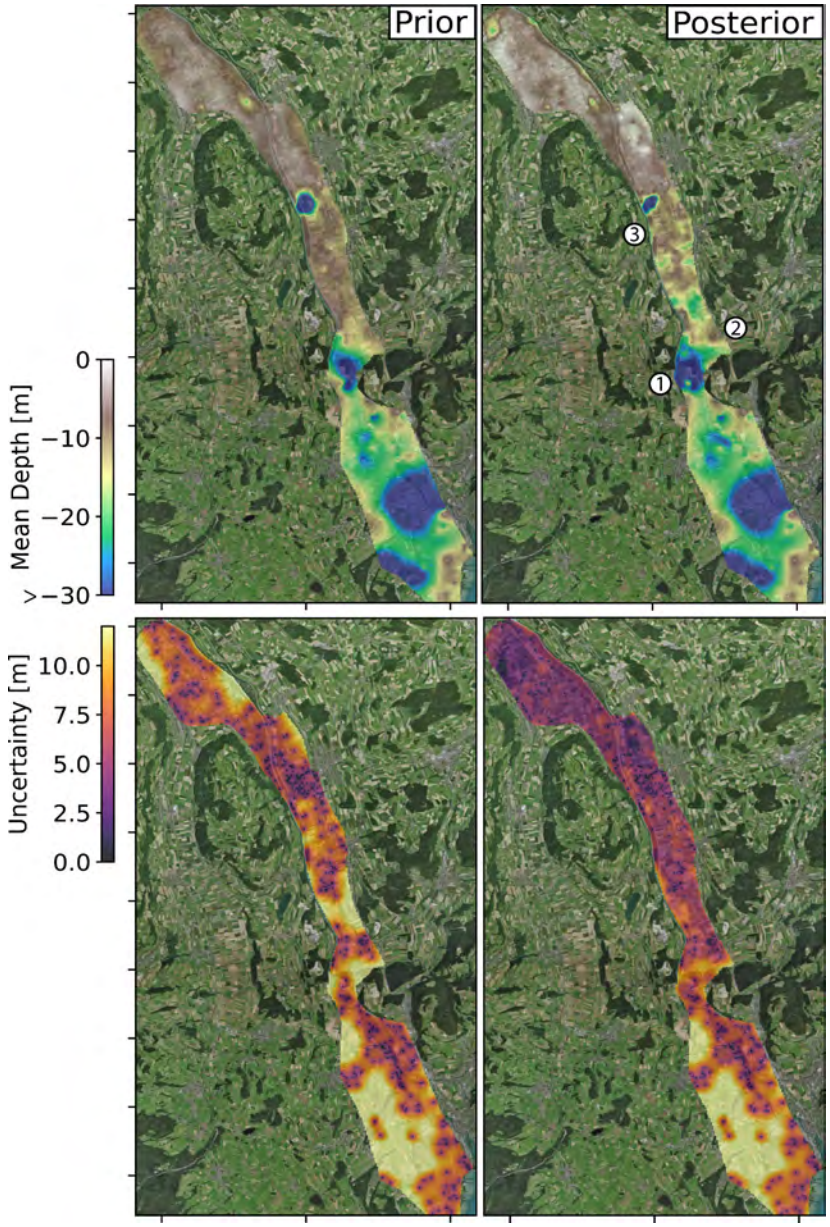

Figure 7.3: Mean depth of the upper aquifer in the complete Aare Valley, compared with a borehole-only-based model (prior) and their associated uncertainty

In terms of geological interpretation of the results, several points can be noted.

- Close to Heimberg, we see a clear narrowing of the valley (see Zone1 in Figure 7.3). Intrestingly, this narrowing also corresponds to a clear deepening of the surface aquifer, both in the prior and in the posterior. The posterior even slightly extends the over-deeping area. This over-deepening can be interpreted as a change in the deposition conditions due to the topographic lock.

- Just above this area (see Zone2 in Figure 7.3), at roughly the center of the model, we have the arrival of one of the main side valleys of the area of interest. The Oberdiessbach valley joins the upper Aare Valley from the East. The valley axis appears to prolong in the upper Aare Valley, perpendicular to the normal flow of the Aare, to the form of a significantly shallower area. The side valley is much steeper in terms of topography compared to the Aare valley. The gain in altitude in the side valley is greater at a distance of 100 m than the total topographic difference in the Aare Valley. This change in stream regime when entering the valley can cause the deposition of many finer sediments during glacial retreat, and potentially cause a shallower aquifer (Messan 1992).

- Between the Oberdiessbach valley and the city of Munsingen, several points can be denoted (see Zone3 in Figure 7.3). First, we can see discontinuities in terms of the thicknesses of the aquifer. Where the prior model estimated a more or less homogeneous aquifer, we see much more variations in the posterior. Then, we also see the large abnormally close to Munsingen. At this position, two boreholes have identified almost 100 meters of porous and saturated rocks. The anomaly is present in both models and its shape was refined by geophysical data. We believe that this could be gravitational deposits on the side of the valley, but more data will be needed to give a more precise answer.

The points above give an overview of the knowledge that can result from the data integration workflow on a large scale.

As mentioned above, the final ensemble is also associated with a relatively high geophysical residual when compared to the classic geophysical

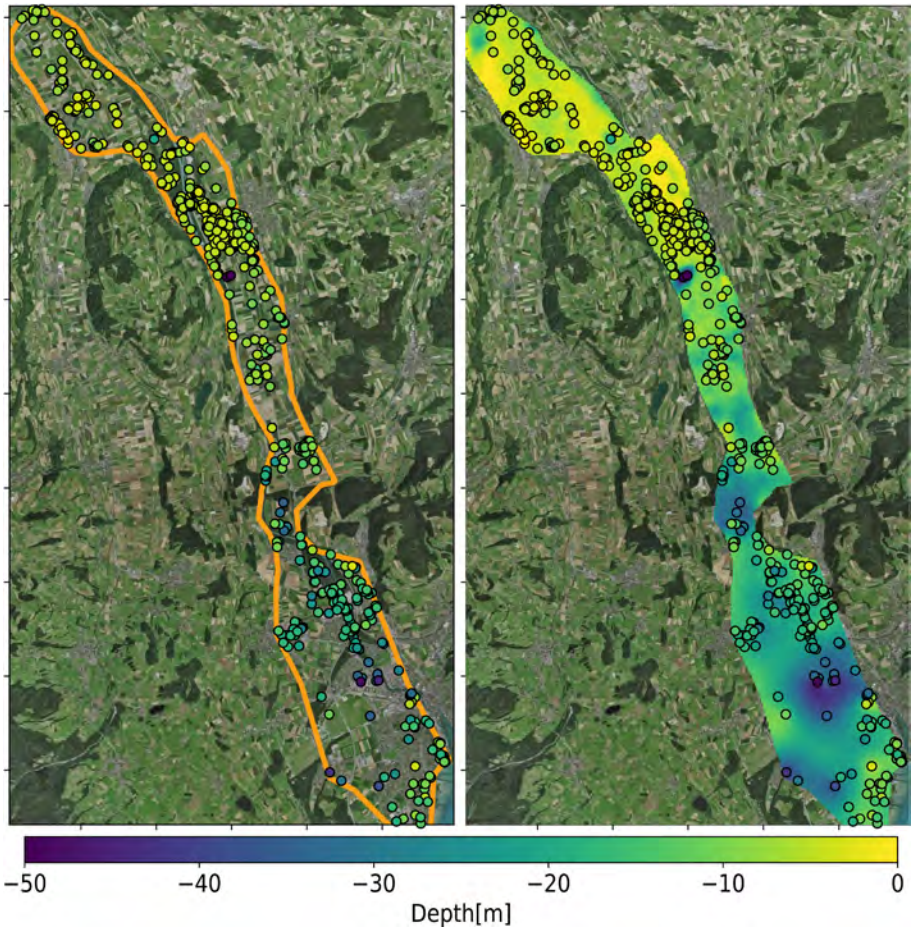

Figure 7.4: Posterior model of the depth of the aquifer compared with the depth retrieved from the boreholes. The boreholes were used in the prior and are therefore, by definition, all in agreement with the posterior model.

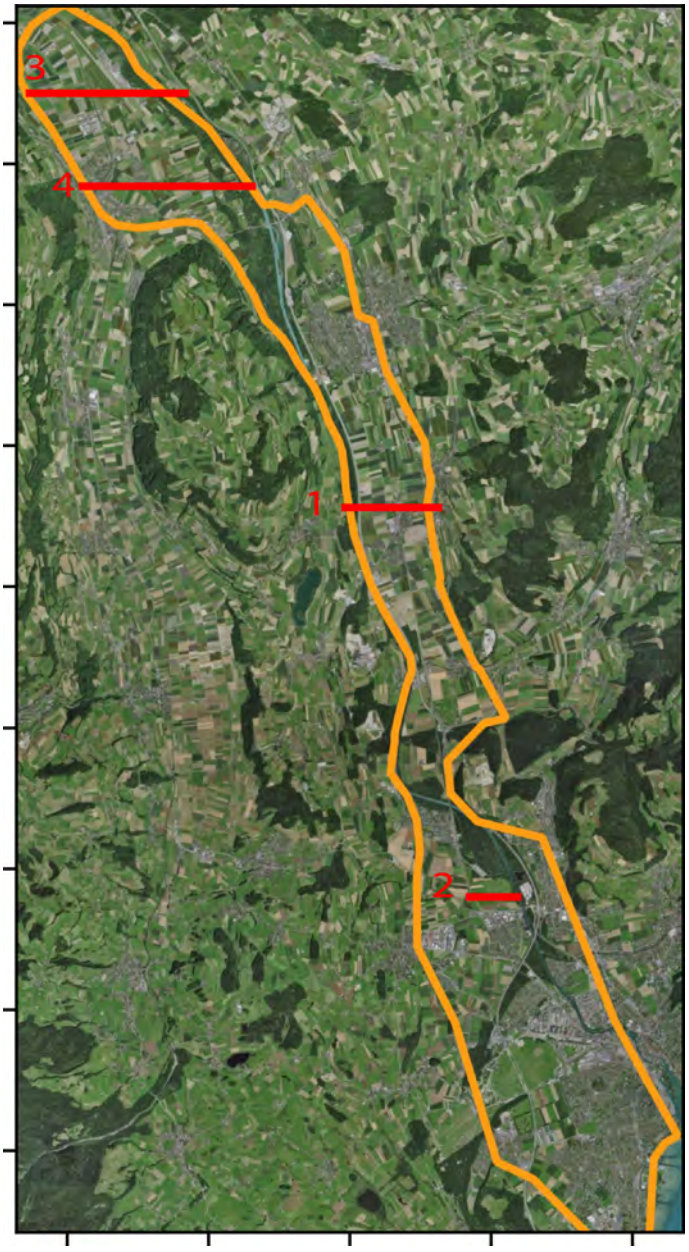

Figure 7.5: Map of the position of the 4 cross-sections visible on Fig. 7.6 and 7.7

inversion. The simulated soundings are several orders of magnitude outside of the uncertainty of the original data. Figures 7.6 and 7.7 compare the resistivity models recovered from the ES-MDA ensemble with the classic deterministic inversion results presented in Chapter 3. The positions of the cross sections are outlined in Figure 7.5. In general, we can denote that the 4 cross sections are visually similar and show consistent features. Cross Section 1 shows that the same soundings were inverted coherently in the two approaches. At position x = 380m, both models predicted a thicker and more conductive second layer. The same proper identification of aquifer thickness can be denoted in cross section 2. The resitivity values of the layer are different. The smooth inversion tends to converge to a larger resistivity contrast between the layers. Sections 3 and 4 focus on the identification of the second layer. The cross-section 3 is an area of high residual. The ES-MDA inversion does not manage to reduce the residual as much as the deterministic inversion. This cross section is close to a well that has successfully hit the lower aquifer at a depth of approximately 75m. In this case, we believe that the variogram used in the prior modeling is too constraining, and none of the prior models are able to fit the data. This leads to a possible poor correlation between the data and the model parameters and therefore to no correction. Cross section 4 on the other hand shows that only a few hundred meters away, the ES-MDA inversion successfully inverts a parametric model close to the one generated by deterministic inversion.

Figure 7.8 compares the same cross section in the original deterministic Geoquat model provided by the Swiss Geological Survey, the CF model presented in Chapter 4, the posterior ArchPy simuation, and the Smooth 1D inverted EM models introduced in Chapter 3. The cross section is the same as the one introduced in Figure 7.2. Overall, the main structures are quite coherent between the three methods and the inversion. It is important to note that the 3 models are binned in a regulard grid, when the inverted models are discrete point. For this reason, the spatial resolution of the resistivity line is higher, and the soundings may be not perfectly colocated with the grid, but are extremely close. All models agree with the presence of an aquifer at the top, but this was expected since it is present through the valley. However, there are several differences in terms of the thickness of the aquifer. The CF inversion is only guided by a relatively simple geological conceptual model, only assuming some continuity in the structures (regularization in the inversion) and is mainly based on boreholes to infer the translation function. In this

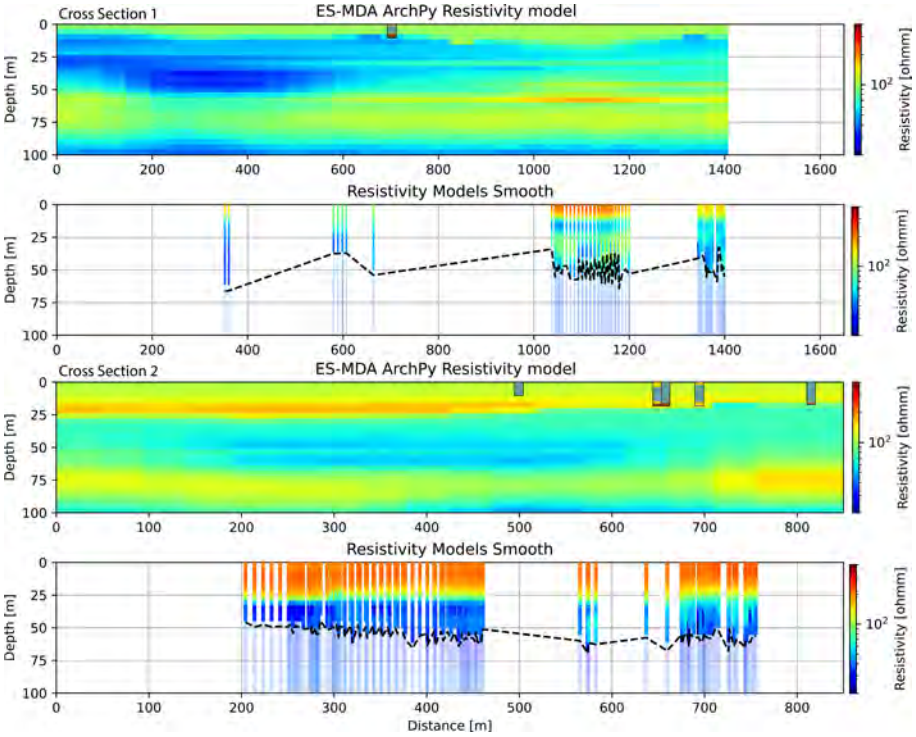

Figure 7.6: Cross-section showing the mean resistivity model retrieved from the ArchPy-ES-MDA posterior ensemble, and the inverted resistivity models using a deterministic smooth regularization. The models are blinded at the standard DOI estimation. Position of cross-section 1 and 2 are shown on Fig. 7.5

situation, the model predicts a thick aquifer, with a possible thickness of about 50 meters in some areas. The upper aquifer has on average a resistivity of about 400-500 ohmm, when the lacustrine clay showed resistivity usually below 60 ohmm. However, in some areas, the aquiclude has a low hydraulic conductivity but a lower content in conductive clay. In those areas, the resistivity of the aquiclude can be higher, up to 100 ohmm. As mentioned above, only a few wells are present in the area, possibly causing a poor constraint on the translator function in the area. Consequently, the model predicts a thick aquifer. On the other hand, the ArchPy and Geoquat models are based on conceptual knowledge of the area. Two aquifers, usually separated by a clay layer, of variable resistivity are expected to be present. Consequently, this layer is represented in the models.

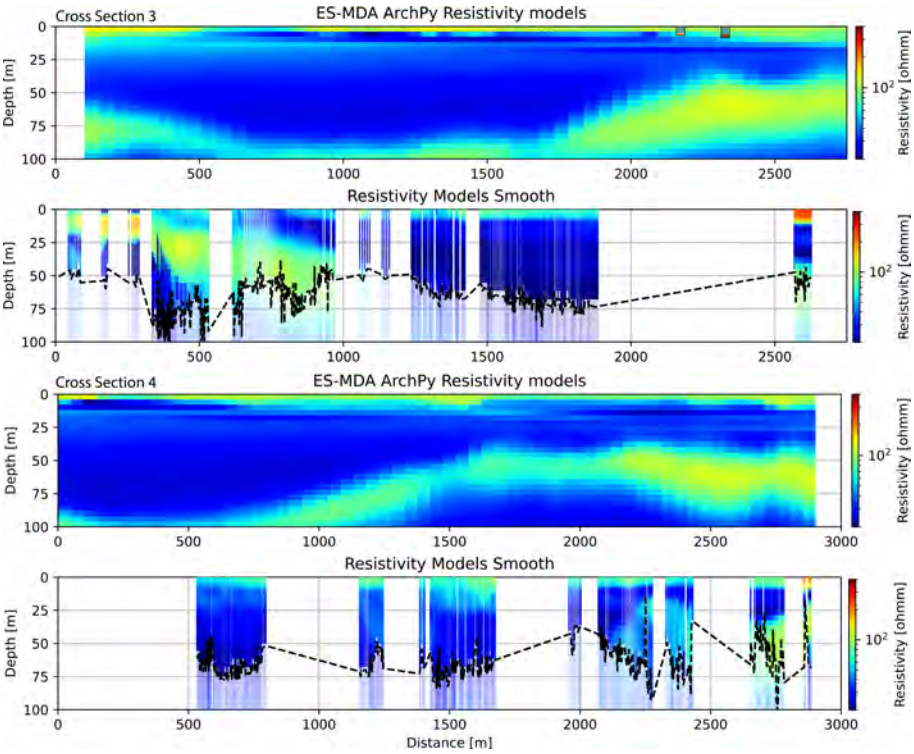

Figure 7.7: Cross-section showing the mean resistivity model retrieved from the ArchPy-ES-MDA posterior ensemble, and the inverted resistivity models using a deterministic smooth regularization. The models are blinded at the standard DOI estimation. Position of cross-section 3 and 4 are shown on Fig. 7.5

The final part of the results analyses the computing time of our inversion. As shown in Figure 7.9 and Table 7.1, one iteration of ES-MDA inversion took approximately 4 hours. The time is more or less evenly distributed between the forward and inverse. Most of the workflow is paralelized, and therefore, two times are provided. The duration per model per CPU is the computing time for one model on a single CPU. If several models are computed in parallel over different CPUs, the time per model can be reduced and the duration per model can be calculated. For example, the forward consists in solving the EM forward for the 10'000 soundings times 510 models. AarhusInv takes about 4:46 for one model, but since it is parallelized over 20 cores, the code solves parallelly 20 models and therefore needs 4:46 for 20 models. The same logic is used for all

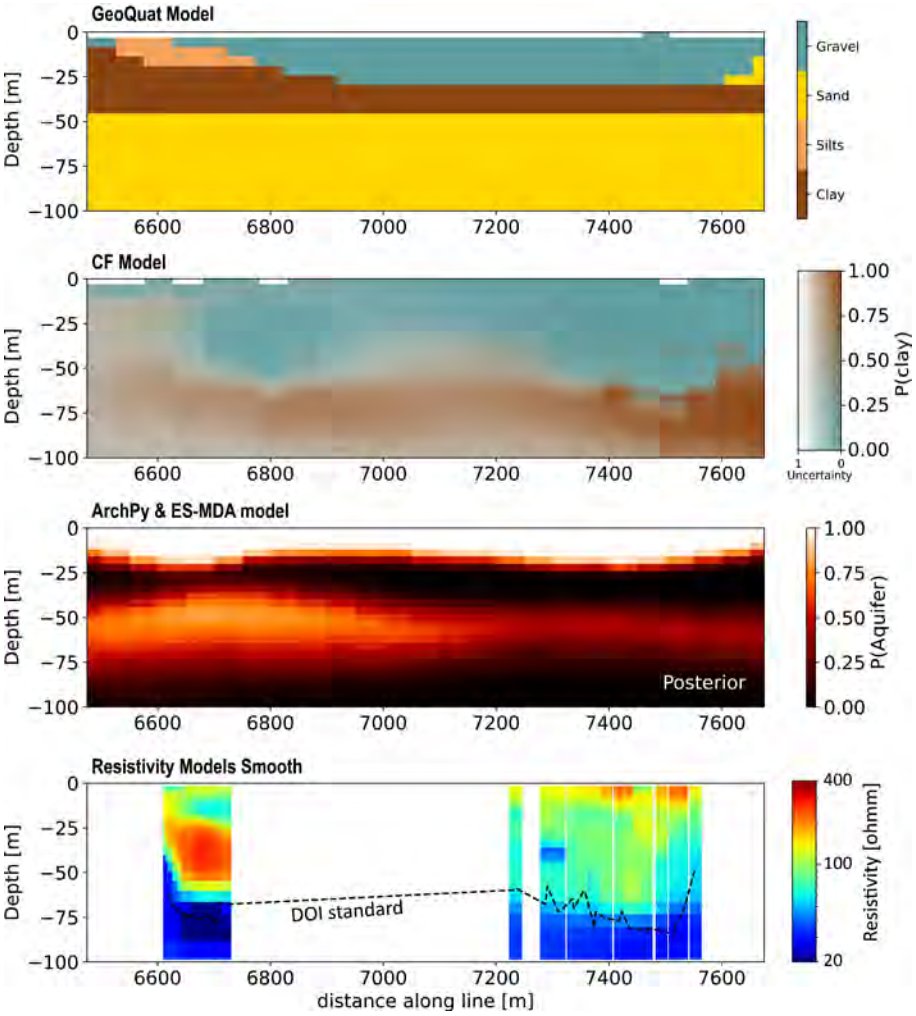

Figure 7.8: Comparison of the same cross section from the GeoQuat model, the CF model, the ArchPy after ES-MDA inversion, and the geophysical deterministic inversion. The Cross-section is a E-W section cross the valley, just at the north of the Kiesen village (x = 391'688 m in UTM32N)

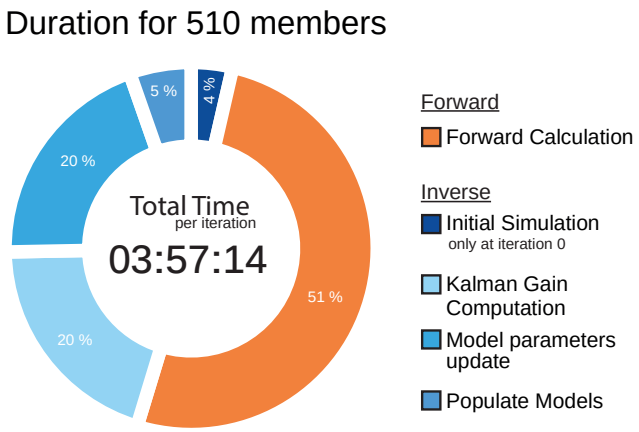

Figure 7.9: Computing times for the first iteration of the ES-MDA inversion on the full-valley scale

Table 7.1: Computation time for the first iteration of the Full Aare Valley inversion

510 members	Step	$\frac{Duration}{model}$	$\frac{Duration}{model \cdot CPU}$	CPUs
00:08:33	Initial Simulation	00:00:01	00:00:31	31
02:01:29	Forward Calculation	00:00:14	00:04:46	20
00:47:33	Kalman Gain Computation	00:00:06	00:02:53	31
00:47:24	Model Parameters Update	00:00:06	00:00:56	10
00:12:45	Populate Models	00:00:02	00:00:47	31
3:57:44	**TOTAL**	**0:00:28**	**00:09:53**	

steps. For efficiency, it is recommended to select the number of models such as $mod(nMembers, nCPUs) = 0$. Due to RAM usage, some part of the workflow runs on fewer CPUs than is available. For example, updating the model parameters involves large matrix multiplication and is demanding in terms of RAM. Increasing the available RAM could speed up these steps.

7.3 Discussion

The preliminary results presented in this chapter prove that our methodology can be applied successfully not only in the Belp area, but also in

the entire valley. When a reasonable groundwater model is available for the entire area, joint data integration could be performed. However, this chapter still proves that a geophysical-only inversion can still bring a wide amount of knowledge.

We saw in the results that the computing time is evenly separated between forward and inverse. In this workflow, all the steps that could take advantage of multithreading were optimized. If one wants to speed up the inversion without increasing the computational power, the use of simplified forward will be needed. This could be achieved by using machine learning-based forwards (e.g. (Bording et al. 2021)). In these cases, the forward does not need to solve complex physical equations but just evaluate the neural network. This could result in significant acceleration if training time is not taken into account or if the trained network can be reused for different sites. The use of a more efficient matrix inversion method can also be considered to speed up Kalman-Gain computation.

The high geophysical residual of the models are probably due to a combination of various causes. First, the mean resistivity value was only constrained by pilot points. Some continuity is expected and will not be able to fit the small-scale variations of resistivity. Secondly, an inadapted prior ensemble will not allow ES-MDA to converge properly, unlike a deterministic inversion where the prior is more diffuse. Increasing the diversity of the prior ensemble and adding more pilot points will certainly diminuish the residual, but increase the computing times.

An interesting point is the comparison between the different models (Fig. 7.8). As mentioned in the results, we denote differences between the CF model and the two other models. Beyond a simple difference in terms of aquifer thickness, we have here two different results that are due to two different approaches, relying on two different philosophies. On the one hand, we have two models that are dominated by prior knowledge and conceptual ideas (Geoquat and ArchPy), while on the other hand, we have a data-based model relying on almost no prior knowledge of the area (CF model). The two approaches showed good results, and both are plausible. In the cross section presented in Figure 7.8, it is possible that the two aquifers identified in the ArchPy inversion are in fact connected. But since the conceptual model of the area expects the presence of a separation between them, this scenario will probably be overrepresented in the posterior models. The CF model does not have this bias; still it relies

only on data. However, infusing prior knowledge can definitely improve the quality of the inversion, and this was proven throughout this thesis. But no uncertainty in the prior conceptual idea of the valley was considered. In this case, problems will arise if the data is integrated in areas that are not compatible with the prior knowledge. Even if ES-MDA has the capability to move outside of the prior, a complete change of conceptual idea is unlikely. In the presented case, the conceptual model is not changed and is uniform throughout the valley. It is possible that better results could be achieved by taking into account various conceptual models and combining them. Again, it proves that efficient and precise geological modeling can only be achieved by joint efforts that bring together geologists, geophysicists, and modelers.

7.4 Conclusion

In the presented section, we drafted some of the results of the complete valley inversion. Since no conceptual groundwater model is yet available, only the geophysical data were integrated with the boreholes using ArchPy and ES-MDA. It proves the applicability of the methodology that effectively integrates geological and geophysical data at the scale of a complete valley. The data are assimilated using Ensemble Smoother for Multiple Data Assimilation (ES-MDA), which leads to a robust error estimation. The methodology reduces computing times by using forward in a reduced-dimensionality and parallel computing. The methodology was applied to a real site in Switzerland The final ensemble of models is compatible with a significant number of geophysical soundings, demonstrating its potential for a wide range of applications.

Chapter 8

Conclusion and perspectives

This thesis has studied several approaches to generate large hydrogeological models automatically and applied them on a Pilot Site in the Upper Aare Valley, Switzerland. The data collection included a large 1500 hectares towed Transient Electro-Magnetic campaign that resulted in about 58'000 soundings after processing. Due to the large amount of data, an automatic interpretation technique had to be developed. The first technique is derived from the CF transfer function introduced in Christiansen et al. (2014) and Foged et al. (2014). This methodology uses boreholes to invert a transfer function that minimizes the mismatch between the predicted lithology derived from resistivity models and the observed lithology in the boreholes. The method was coupled with a stochastic MPS interpolation to propagate the uncertainty down the workflow, up to the final model.

Following the development of ArchPy within the PheniX project, the thesis focused on the integration of different data types within a geologically consistent framework. Given the high dimensionality of the problem and the will to robustly estimate uncertainties, the most determining choice was the inversion technique. Several methods were tested in the Markov chain Monte Carlo class, such as Metropolis-Hastings algorithm, iterative spatial resampling (ISR) (Mariethoz et al. 2010b) and the posterior population expansion (POPEX) (Jäggli et al. 2017). However, even for a 2D problem whose dimension was way below the targetted one, the number of iterations needed to converge was consequent and infeasible combined with the relatively long geophysical forward calulation times. Due to their ability to effectively address large-scale estimation issues in various fields, such as calibration of weather forecast models, ensemble methods for data assimilation were selected. We focus especially on the use of the Ensemble Smoother with Multiple Data Assimilation (ES-MDA) algorithm introduced by Emerick et al. 2013. The algorithm was tested in a synthetic 2D case and combined with a hierarchical inversion approach. This approach was chosen to take advantage of the ArchPy capabilities of hierarchical modeling. The inversion produced good results, and therefore the methodology was upscaled to 3D and applied to real data. The following sections summarize the main conclusions and limitations of the main aspects of this thesis and will include propositions of improvements and further research directions in the field.

8.1 Large Dataset Aquisition

This thesis successfully collected a large amount of data. To better understand the spatial heterogeneity of the valley, we collected a large and detailed time-domain electromagnetic (TDEM) data set in the upper Aare Valley, Switzerland. TDEM data were acquired in an area of 1500 hectares with a 20-meter spacing and are available for access at https://doi.org/10.5281/zenodo.4269887 (Neven et al. 2020). TDEM is an efficient and trustworthy method of measuring the magnetic field that is directly related to the underground resistivity. Inverted resistivity models derived from the acquisition of TDEM data present a range of investigation depth (DOI) between 40 and 120 meters with an average residual contained within the standard deviation of the data. Inverted resistivity models revealed important variations in terms of resistivity through the valley. The upper aquifer presents important variations in thickness, with some overthickened areas where the thickness exceeds the DOI (more than 100m) and some others where the clay aquiclude is only a few meters below ground. This illustrates why spatial heterogeneities have to be considered when the objective is to model medium- to large-scale areas. It also proved that the tTEM method developed by Auken et al. (2019) is an efficient method to map large areas of quaternary deposits.

However, as mentioned in Chapter 3, EM methods are also extremely sensitive to coupling. In a valley like the Upper Aare Valley, all the antropic infrastructures (train line, highway, power lines) are built at the bottom of the valley, where the topography is the flatest. This high density of disruption caused important couplings, and some important part of the original dataset had to be discarded during processing. Therefore, the applicability of this geophysical method could be limited in densely populated and anthropized areas.

8.2 Clay Fraction Inversion

In this chapter, we developped and applied a new method to generate 3D clay fraction models from boreholes and TEM data, using field data from the Upper Aare Valley. Our approach combines the clay fraction estimation algorithm of Foged et al. (2014) with a non-deterministic TI-based multiple-point statistic algorithm. The method provides uncertainty estimates that can inform the acquisition of additional data. The method

is data-driven and does not rely on manual interpretation of structures. It is also fully automatic, including the inversion of the translator function, the automatic fit of the variogram for the GRF error model, and the generation of different training images and conditioning data for each MPS simulation. The user needs to select only a few parameters, with default values available, but the process is largely automated. The method can also reproduce structures that cannot be modeled using traditional two-point interpolation methods. The method was benchmarked using cross-validation and was shown to be unbiased. It predicts uncertainty within the same range of magnitude as the actual error. The model was then compared with the existing GeoQuat deterministic model and presented significant improvements. Finally, the comparison of both models with geological observations along new boreholes not included in the database at the time of the modeling shows that the method accurately predicts general trends, even relatively far from the geophysical data. The workflow could be integrated with public databases to enable regular updating of the 3D clay model as new field data become available.

However, several limitations in this methodology must be underlined. First, the ACT algorithm is based on the colocation of the observations to infer the translator function. EM measurements will need to be at least close to the wells in order to have a robust objective function during the parameter inversion. However, in a lot of cases, the EM measurements cannot be exactly co-located with the EM measurements. In the methodology originally proposed by Christiansen et al. (2014) and applied in the present study, a small distance krigging is used to propagate the clay fraction from neighboring tTEM models to the position of the boreholes and calculate the misfit. This interpolation in environments that could show major spatial discontinuities or where the borehole density is low could be problematic. This need for colocation is one of the major limitations of the ACT method, but also what allows the method to work with only few prior knowledge.

Secondly, even if the uncertainty from the data can be propagated to the model, one source of uncertainty is not taken into account: the fit of the translator function itself. For example, in some areas, the final parameters of the translator function can better predict boreholes CF than in other areas. The causes can be diverse, such as errors in the lithological logging of the boreholes, important spatial variations of the resistivity-clay fraction relationship, inadapted regularization constraints on the

inversion function inversion, or a lower local density of the wells that causes a poorer constraint on the parameters. However, once the function has been optimized, its parameters are deterministic and assumed to be certain. Finally, for a wide variety of applications, the actual variable of interest for the end user is, of course, derived from the Clay Fraction, such as the depth of an aquifer, but not the CF itself. Translating the CF probability to an aquifer thickness for example can be tricky, and the choice at which probability of clay the transition between the aquifer and the aquiclude occurs is difficult. Even if this indicator is useful, it may be difficult for a large audience to adopt it.

8.3 MultiFidelity Stochastic inversion

In this chapter, we have shown, using synthetic tests, that ES-MDA and ArchPy can be combined to create a powerful framework for performing geologically consistent inversions. This approach enables the integration of different types of data in a consistent and stochastic manner, thus reducing global uncertainty on groundwater models. By incorporating the multi-fidelity approach, this study has demonstrated a more efficient infusion of prior geological knowledge into the inversion process, leading to the generation of more realistic geological models in a shorter time. The use of hierarchical multi-fidelity has been beneficial in ensuring that the models remain geologically consistent throughout the inversion process. This has improved the quality of the models generated. Furthermore, hydrogeophysical joint inversions have been effectively decomposed and enhanced within a multi-fidelity and hierarchical framework, demonstrating the versatility and usefulness of this approach in the solution of complex problems.

The authors have also highlighted the usefulness of ArchPy's models as priors for inverting complex 3D models. This chapter has shown that the proposed approach is a promising method for performing geologically consistent inversions that can reduce global uncertainty on groundwater models. In summary, several main conclusions can be drawn from this chapter. Firstly, ES-MDA is an efficient tool for generating an ensemble of plausible hydrogeological models in a multi-fidelity framework. Secondly, hierarchical multi-fidelity helps to maintain geologically consistent models during the inversion process while improving their quality. Thirdly, hydrogeophysical joint inversions can be decomposed and enhanced within a multi-fidelity and hierarchical framework. Finally,

ArchPy's models provide valuable priors for investigating subsurface uncertainty. However, one limitation is that the choice of an erroneous geostatistical prior within Archpy could decrease the realism and convergence rate of the inversion algorithm. This issue should be considered when applying the method to real data sets, and the solution requires proper analysis of all available field data and conceptual knowledge to identify the adapted geostatistical parameters. If the data are not sufficient to constrain the prior model, a possible solution would be to use published data from analog sites and generate a broad ensemble of models using different priors.

An important point to mention is that the groundwater model used for the synthetic case in the study is not very informative. Since the signal is only "naturally" induced and diffuse, the forcing on the permability is limited. It results in poor results obtained with hydrogeological data alone and without geophysical data. The use of different hydrogeological situations, such as including pumping tests or tracer tests, is likely to improve the identifiability of hydraulic conductivity.

8.4 Valley Scale Data integration

The results of this chapter demonstrate the effectiveness of ArchPy coupled with an ES-MDA framework in generating a plausible and robust ensemble of hydrogeological models. The case applied in the upper Aare Valley clearly illustrates the ability of this workflow to integrate different types of data into a geologically consistent framework. In addition to retrieving petrophysical fields, this approach can also estimate discrete hyperparameters, such as the depth of an aquifer. One of the most significant advantages of this methodology is its ability to combine different types of data, including wells, geophysical, and hydrological data, within a geological-driven framework. This approach allows for robust error estimation, ensuring that the resulting models are plausible and realistic, within a reasonable time frame. The benefits of this methodology are further supported by the results of cross-validation, which highlight the advantages of using this approach, particularly in cases where the prior is uncertain.

Overall, the results of this chapter suggest that ArchPy coupled with an ES-MDA framework is a promising tool for generating large-scale realistic and plausible hydrogeological models. The approach presented

in this study has a wide range of potential applications and could be used to address a variety of complex problems in geology, hydrology, and related fields.

8.5 Perspectives

This thesis has extensively explored stochastic data integration in large-scale geological models. We started by trying to interpret the inverted geophysical data in a similar way as a human would try to interpret a geophysical line through the use of the ACT algorithm. We then extend the work by trying to bypass the deterministic petrophysical inversion and directly fill the gap between the data and the geological model. Both methodologies showed good results for a computing time that is comparable. We believe that both can be used, depending on the data and knowledge available on a site. The CF approach has the advantage of relying on a lesser amount of conceptual knowledge than the ArchPy approach. The only prior is the assumption of a relative spatial continuity in terms of translator function parameters and starting values for the translator inversion. This last prior knowledge will especially have a determining effect far from borehole conditionning points. In these positions, the relative diffuse conceptual model will therefore produce a poorly constrained model. On the other hand, the ArchPy approach requires much more site knowledge, such as rough site caracteristics and typical parameter ranges. In our experience, these conceptual prior knowledges are often available, but can be tricky to formally describe. These more formalized priors will produce more complex geological models far from conditionning data points, and will include multiple layers, for example. However, for large-scale models, it is possible that the chosen conceptual model is not valid throughout the site. Therefore, it will lead to a realistic but inadequate model. In summary, both approaches have strengths and weaknesses. The use of one or the other will depend on the purpose and the amount of knowledge available. However, where possible, we believe that more complex prior and modeling methods such as ArchPy should be considered priorily.

In addition, several perspectives and research directions can be indicated. First, we must make an effort to standardize and centralize the data. This axis of research lies between political will and the research itself, but is the basis of all the data-based modeling. When we start dealing with large databases, unformity of description, positionning, or

unit identification has to be ensured. To date of this thesis, there is no guideline for the description of boreholes in Switzerland. Each 26 regions have their own policies and rules when it comes to data management, and each company has its own way of describing cuttings. In almost all cases, data transmission is done in Portable Document Format (PDF), and most of the information is not digitalized. There is no centralized database for either wells or geophysical data. Expanding the work presented here to other areas or even country-wide will not be possible without the data supporting it. Swiss authorities have to move towards systems that have existed for years in Danemark or Belgium for example, where the transmission of the data to the authorities is done through a defined format, with only a control of the data from the authorities. Public authorities will need to establish a standardization of the description of wells in collaboration with the universities and all actors in the field.

Secondly, a possible improvement of the ACT method would be the use of a stochastic inversion framework to tune the parameters of the translator function. In doing so, the uncertainty on the parameters of the translation function could be efficiently estimated and propagated. This would be even more critical if some part of the investigated site has no boreholes and is beyond the regularization range because the deterministic inversion would leave the parameters unchanged at the starting value. A stochastic inversion will also not change the parameters range in such areas, but will then just converge to the distribution of the prior, reflecing a high uncertainty.

The major part of this thesis focused on the use of ArchPy and the stochastic inverse algorithm. In this context, two main research directions could emerge. First, expanding the work presented in Chapter 6 to the entire Aare Valley with the inclusion of a hydrological model is a clear follow-up of this thesis. However, certain challenges will arise: In the case presented in Belp, the hyperparameters of the hydrological model are kept unchanged. These parameters, such as, for example, the recharge volume, the inflitration ratio, the river conductance or all of the boundary conditions, will probably affect much more the prediction of the complete valley size model, compared to the smaller example presented in this thesis. Indeed, the distance between the main north and south boundary conditions will be multiplied by more than 12. The coupling of all those parameters with the inversion will be necessary and will include a detailed analysis of which parameters should be included

and at which spatial scale. Another point that will be of interest and has not been extensively analyzed in this thesis is the balance between multiple data types during the ES-MDA inverison. Previous unpublished work from the author has shown that the stability of an objective function with data of different natures and different orders of magnitude can become unstable. The balance between the hydrological and geophysical data in Belp seems to be satisfactory, but extensive tests have not been performed within this thesis.

This thesis has shown that ES-MDA has succeeded in updating complex hierarchical models. The choice of ES-MDA was mainly due to the high number of parameters involved. To our knowledge, no other inversion scheme can be as efficient for large datasets as ES-MDA. However, this efficacity comes with several disadvantages, such as inbreeding. The inbreeding term was first introduced by Houtekamer et al. (1998) and describes the fact that iteration after iteration, the ES-MDA members become more and more coupled. This is mainly due to the fact that the same ensemble is used to calculate the Kalman Gain matrix and the misfit. The method used in this thesis to counterbalance this is the method proposed by Leeuwen et al. (1996) consisting of the use of an inflation factor alpha, avoiding the collapse of the ensemble to a single model. In addition, we took a number of ensemble larger than what is usually seen. However, despite this factor, ES-MDA has trouble efficiently scaling the uncertainty. In this paper, an extensive analysis was not performed and such an analysis will be needed. In addition, another disadvantage when the dataset is getting large is that the matrix inverse calculation during the Kalman gain can become computationally extremely expensive. As seen in Chapter 7, the computing time is almost equally distributed between the physical forwards and the inverse. In the valley inversion, for example, each pseudo-inverse took about an hour on a high-performance cluster. For this step, this calls for the use in the future of either a more efficient pseudo-inverse method, the implementation of singular value decomposition approaches or the implementation of local updating, where only a subset of the model is updated iteratively. For the forward part, the use of machine learning based forward (Bording et al. 2021, i.e.) could speed up the inversion. Finally, I want to point out the difference in scale between the models and the tTEM acquisition. In the present thesis, the tTEM has a resolution of around 10m along lines, and 20m cross lines. However, the desired resolution for a regional hydrological model is often in the range of 25 to 100 m cell size. We think each update should take

advantage of the best resolution available. The multifidelity presented in Chapter 5 could be extended, in a multiresolution framework. The first part of the update iteration would invert a unit model from the geophysical measurement, with a finer resolution (on the order of 10mx10m), and the petrophysical permeability simulation would be simulated using an upscaled version of the surfaces. In doing so, both methods could be combined at the peak of their efficiency.

This thesis represents a first step towards exploring the potential practical applications of the new methods presented. By demonstrating the feasibility and applicability of the approaches, we hope to inspire further research and development in this field. The data used in this study were carefully selected to ensure that they are not specific to any particular site. Consequently, the results obtained could easily be reproduced in various locations, demonstrating the scalability of the methods. One of the key contributions of this thesis is the improvement in uncertainty quantification, which is essential for practical applications. We focus specifically on directly quantifying the uncertainty of the variable of interest, such as the head, the probability of clay, the flow path, the permeability or the thickness of the aquifer. Significant improvements in their quantification were achieved. Overall, this thesis provides a solid foundation for future research and development in this field. The approaches and results presented here offer a glimpse of what could be achieved, and we believe that with further refinement and development, they could become an essential tool for a wide range of day-to-day applications.

Appendix A

Data Records File Description

Processed_Data.dat			
Column	Label	Unit	Description
1	RECORD		Global record number. Links the data to the resistivity model in in the *.inv files
2	LINE_NO		Line number (Line number 0 = data/model not tacked with a line number)
3	UTMX	(m)	UTMX coordinate, WGS 84 UTM zone 32N (epsg:32632)
4	UTMY	(m)	UTMY coordinate, WGS 84 UTM zone 32N (epsg:32632)
5	ELEVATION	(m)	Surface elevation
6	NUMDATA		Number of data points (gates) in-use for the segment/sounding
7	SEGMENT		Transmitter moment indicator. 1=Low moment, 2=High moment
8-37	DATA_#	(V/(Am4))	Processed z-component dB/dt data value for gate number #. 9999 values = data not in-use/not present
38-66	DATASTD_#	STD	Data uncertainty for DATA_#, stated as a relative STD in log space.

Table A.1: Structure of the .dat datafile

A.1 Processed data file

The table A.1 refers to chapter 3 and provide the details of the processed data file structure.

Smooth_Model.inv, Sharp_Model.inv			
Colum	**Label**	**Unit**	**Description**
1	RECORD		Global record number. Links the model the data in the *.inv files
2	LINE_NO		Line number (Line number 0 = data/model not tacked with a line number)
3	UTMX	(m)	UTMX coordinate, WGS 84 UTM zone 32N (epsg:32632)
4	UTMY	(m)	UTMY coordinate, WGS 84 UTM zone 32N (epsg:32632)
5	ELEVATION	(m)	Surface elevation
6	DATAFIT		Datafit (Data residual)
7-36	RHO_I_#	(Ohmm)	Resistivity of layer#.
37-65	THK_#	(m)	Thickness of layer #.
66	DOI_CONSERVATIVE	(m)	Estimated depth of investigation, conservative threshold value used
67	DOI_STANDARD	(m)	Estimated depth of investigation, standard threshold value used

Table A.2: Structure of the .inv datafile

A.2 Inversion Model File

The table A.2 refers to chapter 3 and provides the details of the inversion model file structure.

Forward_Data_Smooth.dat, Forward_Data_Sharp.dat			
Colum	**Label**	**Unit**	**Description**
1	RECORD		Global record number. Links the data to the resistivity model in in the *.inv files
2	LINE_NO		Line number (Line number 0 = data/model not tacked with a line number)
3	UTMX	(m)	UTMX coordinate, WGS 84 UTM zone 32N (epsg:32632)
4	UTMY	(m)	UTMY coordinate, WGS 84 UTM zone 32N (epsg:32632)
5	ELEVATION	(m)	Surface elevation
6	NUMDATA		Number of data points (gates) in-use for the segment/sounding
7	SEGMENT		Transmitter moment indicator. 1=Low moment, 2=High moment
8-37	DATA_#	(V/(Am4))	Model forward response, dB/dt, for gate number #. 9999 values = data not in-use/not present

Table A.3: Structure of the .syn datafile

A.3 Synthetic response file

The table A.3 refers to chapter 3 and provides the details of the synthetic response file structure.

Bibliography

Adams, Kyra H., John T. Reager, Paul Rosen, David N. Wiese, Tom G. Farr, Shanti Rao, Bruce J. Haines, Donald F. Argus, Zhen Liu, Ryan Smith, James S. Famiglietti, and Matthew Rodell (Oct. 2022). "Remote Sensing of Groundwater: Current Capabilities and Future Directions". In: *Water Resources Research* 58.10. DOI: 10.1029/2022wr032219.

Alumbaugh, David L. and Gregory A. Newman (Sept. 2000). "Image appraisal for 2-D and 3-D electromagnetic inversion". In: *Geophysics* 65.5, pp. 1455–1467. DOI: 10.1190/1.1444834.

Amaya, Macarena, Niklas Linde, and Eric Laloy (Aug. 2022). "Hydrogeological multiple-point statistics inversion by adaptive sequential Monte Carlo". In: *Advances in Water Resources* 166, p. 104252. DOI: 10.1016/j.advwatres.2022.104252.

Anderson, Jeffrey L. (Dec. 2001). "An Ensemble Adjustment Kalman Filter for Data Assimilation". In: *Monthly Weather Review* 129.12, pp. 2884–2903. DOI: 10.1175/1520-0493(2001)129<2884:aeakff>2.0.co;2.

Archie, G.E. (Dec. 1942). "The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics". en. In: *Transactions of the AIME* 146.01, pp. 54–62. ISSN: 0081-1696. DOI: 10.2118/942054-G.

Asher, Michael J, Barry FW Croke, Anthony J Jakeman, and Luk JM Peeters (2015). "A review of surrogate models and their application to groundwater modeling". In: *Water Resources Research* 51.8, pp. 5957–5973.

Auken, Esben, Anders Vest Christiansen, Casper Kirkegaard, Gianluca Fiandaca, Cyril Schamper, Ahmad Ali Behroozmand, Andrew Binley, Emil Nielsen, Flemming Effersø, Niels Bøie Christensen, Kurt Sørensen, Nikolaj Foged, and Giulio Vignoli (Sept. 2015). "An overview of a highly versatile forward and stable inverse algorithm for airborne, ground-based and borehole electromagnetic and electric data". In: *Exploration Geophysics* 46.3, pp. 223–235. DOI: 10.1071/eg13097.

Auken, Esben, Nikolaj Foged, Jakob Juul Larsen, Knud Valdemar Trøllund Lassen, Pradip Kumar Maurya, Søren Møller Dath, and Tore Tolstrup Eiskjær (Jan. 2019). "tTEM — A towed transient electromagnetic system for detailed 3D imaging of the top 70 m of the subsurface". In: *GEOPHYSICS* 84.1, E13–E22. DOI: 10.1190/geo2018-0355.1.

Bakker, Mark, V.+ Post, Christian D Langevin, Joseph D Hughes, Jeremy T White, JJ Starn, and Michael N Fienen (2016). "Scripting MODFLOW model development using Python and FloPy". In: *Groundwater* 54.5, pp. 733–739.

Bandou, D., F. Schlunegger, E. Kissling, U. Marti, M. Schwenk, P. Schläfli, G. Douillet, and D. Mair (Jan. 2022). "Three-dimensional gravity modelling of a Quaternary overdeepening fill in the Bern area of Switzerland discloses two stages of glacial carving". In: *Scientific Reports* 12.1. DOI: 10.1038/s41598-022-04830-x.

Barfod, Adrian A. S., Ingelise Møller, Anders V. Christiansen, Anne-Sophie Høyer, Júlio Hoffimann, Julien Straubhaar, and Jef Caers (June 2018). "Hydrostratigraphic modeling using multiple-point statistics and airborne transient electromagnetic methods". In: *Hydrology and Earth System Sciences* 22.6, pp. 3351–3373. DOI: 10.5194/hess-22-3351-2018.

Bording, Thue Sylvester, Muhammad Rizwan Asif, Adrian S. Barfod, Jakob Juul Larsen, Bo Zhang, Denys James Grombacher, Anders Vest Christiansen, Kim Wann Engebretsen, Jesper Bjergsted Pedersen, Pradip Kumar Maurya, and Esben Auken (Apr. 2021). "Machine learning based fast forward modelling of ground-based time-domain electromagnetic data". In: *Journal of Applied Geophysics* 187, p. 104290. DOI: 10.1016/j.jappgeo.2021.104290.

Bouzaglou, Véronique, Elena Crestani, Paolo Salandin, Erwan Gloaguen, and Matteo Camporese (Mar. 2018). "Ensemble Kalman Filter Assimilation of ERT Data for Numerical Modeling of Seawater Intrusion in a Laboratory Experiment". In: *Water* 10.4, p. 397. DOI: 10.3390/w10040397.

Branch, Mary Ann, Thomas F. Coleman, and Yuying Li (Jan. 1999). "A Subspace, Interior, and Conjugate Gradient Method for Large-Scale Bound-Constrained Minimization Problems". en. In: *SIAM Journal on Scientific Computing* 21.1, pp. 1–23. ISSN: 1064-8275, 1095-7197. DOI: 10.1137/S1064827595289108.

Carle, Steven F and Graham E Fogg (2020). "Integration of soft data into geostatistical simulation of categorical variables". In: *Frontiers in Earth Science* 8, p. 565707.

Carvalho, Paulo Roberto Moura de, João Felipe Coimbra Leite da Costa, Luiz Gustavo Rasera, and Luiz Eduardo Seabra Varella (2017). "Geostatistical facies simulation with geometric patterns of a petroleum reservoir". In: *Stochastic Environmental Research and Risk Assessment* 31.7, pp. 1805–1822. DOI: 10.1007/s00477-016-1243-5.

Casagrande, Arthur (1948). "Classification and identification of soils". In: *Transactions of the American Society of Civil Engineers* 113.1, pp. 901–930. DOI: 10.1061/TACEAT.0006109.

Chen, Yan and Dean S. Oliver (Dec. 2011). "Ensemble Randomized Maximum Likelihood Method as an Iterative Ensemble Smoother". In: *Mathematical Geosciences* 44.1, pp. 1–26. DOI: 10.1007/s11004-011-9376-z.

— (May 2013). "Levenberg–Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification". In: *Computational Geosciences* 17.4, pp. 689–703. DOI: 10.1007/s10596-013-9351-5.

Chilès, Jean-Paul and Pierre Delfiner (Sept. 2009). *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons. ISBN: 978-0-470-31783-9.

Christensen, N. K., B. J. Minsley, and S. Christensen (Feb. 2017). "Generation of 3-D hydrostratigraphic zones from dense airborne electromagnetic data to assess groundwater model prediction error". In: *Water Resources Research* 53.2, pp. 1019–1038. DOI: 10.1002/2016wr019141.

Christiansen, A.V., N. Foged, and E. Auken (Sept. 2014). "A concept for calculating accumulated clay thickness from borehole lithological logs and resistivity models for nitrate vulnerability assessment". en. In: *Journal of Applied Geophysics* 108, pp. 69–77. ISSN: 09269851. DOI: 10.1016/j.jappgeo.2014.06.010.

Christiansen, Anders Vest and Esben Auken (July 2012). "A global measure for depth of investigation". In: *GEOPHYSICS* 77.4, WB171–WB177. DOI: 10.1190/geo2011-0393.1.

Christiansen, Anders Vest, Esben Auken, and Kurt Sørensen (2009a). "The transient electromagnetic method". In: *Groundwater Geophysics: A Tool for Hydrogeology*. Ed. by Reinhard Kirsch. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 179–226. ISBN: 978-3-540-88405-7. DOI: 10.1007/978-3-540-88405-7_6.

— (2009b). "The transient electromagnetic method". In: *Groundwater Geophysics: A Tool for Hydrogeology*. Ed. by Reinhard Kirsch. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 179–226. ISBN: 978-3-540-88405-7. DOI: 10.1007/978-3-540-88405-7_6.

Dagasan, Yasin, Przemysław Juda, and Philippe Renard (May 2020). "Using Generative Adversarial Networks as a Fast Forward Operator for Hydrogeological Inverse Problems". In: *Groundwater* 58.6, pp. 938–950. DOI: 10.1111/gwat.13005.

Dall'Alba, Valentin, Philippe Renard, Julien Straubhaar, Benoit Issautier, Cédric Duvail, and Yvan Caballero (Oct. 2020). "3D multiple-point statistics simulations of the Roussillon Continental Pliocene aquifer using DeeSse". en. In: *Hydrology and Earth System Sciences* 24.10, pp. 4997–5013. ISSN: 1607-7938. DOI: 10.5194/hess-24-4997-2020.

De Marsily, Gh, Frédéric Delay, Julio Gonçalvès, Ph Renard, Vanessa Teles, and Sophie Violette (2005). "Dealing with spatial heterogeneity". In: *Hydrogeology Journal* 13.1, pp. 161–183. DOI: 10.1007/s10040-004-0432-3.

Deutsch, Clayton V and André G Journel (1992). *GSLIB. Geostatistical software library and user's guide*. New York: Oxford University Press.

Deutsch, Clayton V. (2002). *Geostatistical reservoir modeling*. Applied geostatistics series. Oxford ; New York: Oxford University Press.

Dietrich, C. R. and G. N. Newsam (Aug. 1993). "A fast and exact method for multidimensional gaussian stochastic simulations". en. In: *Water Resources Research* 29.8, pp. 2861–2869. ISSN: 00431397. DOI: 10.1029/93WR01070.

Dumont, M., P.A. Reninger, A. Pryet, G. Martelet, B. Aunay, and J.L. Join (Oct. 2018). "Agglomerative hierarchical clustering of airborne electromagnetic data for multi-scale geological studies". In: *Journal of Applied Geophysics* 157, pp. 1–9. DOI: 10.1016/j.jappgeo.2018.06.020.

Emerick, Alexandre A. (Mar. 2016). "Analysis of the performance of ensemble-based assimilation of production and seismic data". In: *Journal of Petroleum Science and Engineering* 139, pp. 219–239. DOI: 10.1016/j.petrol.2016.01.029.

Emerick, Alexandre A. and Albert C. Reynolds (June 2013). "Ensemble smoother with multiple data assimilation". In: *Computers & Geosciences* 55, pp. 3–15. DOI: 10.1016/j.cageo.2012.03.011.

Engebretsen, Kim Wann, Bo Zhang, Gianluca Fiandaca, Line Meldgaard Madsen, Esben Auken, and Anders Vest Christiansen (Feb. 2022). "Accelerated 2.5-D inversion of airborne transient electromagnetic data using reduced 3-D meshing". In: *Geophysical Journal International* 230.1, pp. 643–653. DOI: 10.1093/gji/ggac077.

Evensen, Geir (2009). *Data Assimilation*. Springer Berlin Heidelberg. DOI: 10.1007/978-3-642-03711-5.

— (Mar. 2018). "Analysis of iterative ensemble smoothers for solving inverse problems". In: *Computational Geosciences* 22.3, pp. 885–908. DOI: 10.1007/s10596-018-9731-y.

Farquharson, C. G. and D. W. Oldenburg (Sept. 1993). "Inversion of time-domain electromagnetic data for a horizontally layered Earth". In: *Geophysical Journal International* 114.3, pp. 433–442. DOI: 10.1111/j.1365-246x.1993.tb06977.x.

Fitterman, David V. (Nov. 1987). "Examples of Transient Sounding for Ground-Water Exploration in Sedimentary Aquifers". In: *Ground Water* 25.6, pp. 685–692. DOI: 10.1111/j.1745-6584.1987.tb02209.x.

Foged, N, P A Marker, A V Christansen, P Bauer-Gottwein, and F Jørgensen (2014). "Large-scale 3-D modeling by integration of resistivity models and borehole data through inversion". en. In: *Hydrol. Earth Syst. Sci.* 18, pp. 4349–4362. DOI: 10.5194/hess-18-4349-2014.

Foged, Nikolaj, Esben Auken, Anders Vest Christiansen, and Kurt Ingvard Sørensen (Mar. 2013). "Test-site calibration and validation of airborne and ground-based TEM systems". In: *GEOPHYSICS* 78.2, E95–E106. DOI: 10.1190/geo2012-0244.1.

Freulon, Xavier and Chantal de Fouquet (1993). "Conditioning a Gaussian model with inequalities". In: *Geostatistics Tróia'92: Volume 1*, pp. 201–212.

Gaspari, Gregory and Stephen E. Cohn (Jan. 1999). "Construction of correlation functions in two and three dimensions". In: *Quarterly Journal of the Royal Meteorological Society* 125.554, pp. 723–757. DOI: 10.1002/qj.49712555417.

Gneiting, Tilmann, Fadoua Balabdaoui, and Adrian E. Raftery (2007). "Probabilistic forecasts, calibration and sharpness". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 69.2, pp. 243–268.

Graf, Hans Rudolf and Reto Burkhalter (July 2016). "Quaternary deposits: concept for a stratigraphic classification and nomenclature—an example from northern Switzerland". In: *Swiss Journal of Geosciences* 109.2, pp. 137–147. DOI: 10.1007/s00015-016-0222-7.

Hansen, T. M. and C. C. Finlay (Oct. 2022). "Use of Machine Learning to Estimate Statistics of the Posterior Distribution in Probabilistic Inverse Problems—An Application to Airborne EM Data". In: *Journal of Geophysical Research: Solid Earth* 127.11. DOI: 10.1029/2022jb024703.

Hartigan, J. A. and M. A. Wong (1979). "Algorithm AS 136: A K-Means Clustering Algorithm". In: *Applied Statistics* 28.1, p. 100. ISSN: 00359254. DOI: 10.2307/2346830.

Henriksen, Hans Jørgen, Lars Troldborg, Per Nyegaard, Torben Obel Sonnenborg, Jens Christian Refsgaard, and Bjarne Madsen (2003). "Methodology for construction, calibration and validation of a national hydrological model for Denmark". In: *Journal of Hydrology* 280.1-4, pp. 52–71. DOI: 10.1016/S0022-1694(03)00186-0.

Houtekamer, P. L. and Herschel L. Mitchell (Mar. 1998). "Data Assimilation Using an Ensemble Kalman Filter Technique". In: *Monthly Weather Review* 126.3, pp. 796–811. DOI: 10.1175/1520-0493(1998)126<0796:dauaek>2.0.co;2.

Høyer, A-S, F Jørgensen, PBE Sandersen, A Viezzoli, and I Møller (2015). "3D geological modelling of a complex buried-valley network delineated from borehole and AEM data". In: *Journal of Applied Geophysics* 122, pp. 94–102.

Hughes, Joseph D., Christian D. Langevin, and Edward R. Banta (2017). *Documentation for the MODFLOW 6 framework*. DOI: 10.3133/tm6a57.

Ibele, Tobias (June 2011). "Tectonics of the Western Swiss molasse basin during Cenozoic times". In: *folia.unifr.ch*.

Irving, James and Kamini Singha (Nov. 2010). "Stochastic inversion of tracer test and electrical geophysical data to estimate hydraulic conductivities". In: *Water Resources Research* 46.11. DOI: 10.1029/2009wr008340.

Jäggli, C., J. Straubhaar, and P. Renard (Apr. 2017). "Posterior population expansion for solving inverse problems". In: *Water Resources Research* 53.4, pp. 2902–2916. DOI: 10.1002/2016wr019550.

Jardani, A., A. Revil, and J.P. Dupont (Feb. 2013). "Stochastic joint inversion of hydrogeophysical data for salt tracer test monitoring and hydraulic conductivity imaging". In: *Advances in Water Resources* 52, pp. 62–77. DOI: 10.1016/j.advwatres.2012.08.005.

Jørgensen, Flemming, Anne-Sophie Høyer, Peter B.E. Sandersen, Xiulan He, and Nikolaj Foged (Aug. 2015). "Combining 3D geological modelling techniques to address variations in geology, data type and density – An example from Southern Denmark". en. In: *Computers & Geosciences* 81, pp. 53–63. ISSN: 00983004. DOI: 10.1016/j.cageo.2015.04.010.

Jørgensen, Flemming, Rasmus Rønde Møller, Lars Nebel, Niels-Peter Jensen, Anders Vest Christiansen, and Peter B. E. Sandersen (Dec. 2013). "A method for cognitive 3D geological voxel modelling of AEM data". en. In:

Bulletin of Engineering Geology and the Environment 72.3-4, pp. 421–432. ISSN: 1435-9529, 1435-9537. DOI: 10.1007/s10064-013-0487-2.

Journel, A. G. (June 1983). "Nonparametric Estimation of Spatial Distributions". In: *Journal of the International Association for Mathematical Geology* 15.3, pp. 445–468. ISSN: 0020-5958, 1573-8868. DOI: 10.1007/BF01031292.

Journel, A. G. and E. H. Isaaks (Oct. 1984). "Conditional Indicator Simulation: Application to a Saskatchewan Uranium Deposit". In: *Journal of the International Association for Mathematical Geology* 16.7, pp. 685–718. ISSN: 1573-8868. DOI: 10.1007/BF01033030.

Juda, Przemysław and Philippe Renard (2021). "An attempt to boost posterior population expansion using fast machine learning algorithms". In: *Frontiers in Artificial Intelligence* 4, p. 624629.

Juda, Przemysław, Philippe Renard, and Julien Straubhaar (Aug. 2020). "A Framework for the Cross-Validation of Categorical Geostatistical Simulations". In: *Earth and Space Science* 7.8. DOI: 10.1029/2020ea001152.

Juda, Przemysław, Julien Straubhaar, and Philippe Renard (2022). "Comparison of three recent discrete stochastic inversion methods and influence of the prior choice". In: *Comptes Rendus. Géoscience* In review.

Kalman, Rudolph Emil (1960). "A new approach to linear filtering and prediction problems". In: *Transactions of the ASME–Journal of Basic Engineering (Series D)* 82, pp. 35–45.

Kang, S., R. Cockett, L. J. Heagy, and D. W. Oldenburg (Aug. 2015). "Moving between dimensions in electromagnetic inversions". In: *SEG Technical Program Expanded Abstracts 2015*. Society of Exploration Geophysicists. DOI: 10.1190/segam2015-5930379.1.

Kang, S., R. Knight, T. Greene, C. Buck, and G. Fogg (Oct. 2021). "Exploring the Model Space of Airborne Electromagnetic Data to Delineate Large-Scale Structure and Heterogeneity Within an Aquifer System". In: *Water Resources Research* 57.10. DOI: 10.1029/2021wr029699.

Kang, Xueyuan, Xiaoqing Shi, André Revil, Zhendan Cao, Liangping Li, Tian Lan, and Jichun Wu (2019). "Coupled hydrogeophysical inversion to identify non-Gaussian hydraulic conductivity field by jointly assimilating geochemical and time-lapse geophysical data". In: *Journal of Hydrology* 578, p. 124092.

Keller, Beat (2021). "Lake Lucerne and its spectacular landscape". In: *World Geomorphological Landscapes*. Cham: Springer International Publishing, pp. 305–323.

Kellerhals, P., Ch Haefeli, and B Tröhler (1981a). *Hydrogeologie Aaretal, zwischen Thun und Bern*. Wasser- u. Energiewirtschaftsamt des Kt. Bern (WEA). Bern: Schweizerische Geologische Dokumentationsstele, Landeshydrologie und -geologie.

— (1981b). *Grundlagen für Schutz und Bewirtschaftung der Grundwasser des Kantons Bern Hydrogeologie Aaretal, zwischen Thun und Bern*. Tech. rep. Bern: Wasser- u. Energiewirtschaftsamt des Kantons Bern, p. 144.

Khambhammettu, Prashanth, Philippe Renard, and John Doherty (Apr. 2020). "The Traveling Pilot Point method. A novel approach to parameterize the inverse problem for categorical fields". In: *Advances in Water Resources* 138, p. 103556. DOI: 10.1016/j.advwatres.2020.103556.

Kirkegaard, Casper, Kristoffer Andersen, Tue Boesen, Anders V. Christiansen, Esben Auken, and Gianluca Fiandaca (Dec. 2015). "Utilizing massively parallel co-processors in the AarhusInv 1D forward and inverse AEM modelling code". In: *ASEG Extended Abstracts* 2015.1, pp. 1–3. DOI: 10.1071/aseg2015ab125.

Knight, Rosemary, Ian Gottschalk, and Noah Dewar (2021). "Field-Scale Rock Physics for Near-Surface Applications". en. In: *Encyclopedia of Geology*. Elsevier, pp. 884–899. ISBN: 978-0-08-102909-1. DOI: 10.1016/B978-0-12-409548-9.12514-X.

Knight, Rosemary, Ryan Smith, Ted Asch, Jared Abraham, Jim Cannia, Andrea Viezzoli, and Graham Fogg (Nov. 2018). "Mapping Aquifer Systems with Airborne Electromagnetics in the Central Valley of California". en. In: *Groundwater* 56.6, pp. 893–908. ISSN: 0017-467X, 1745-6584. DOI: 10.1111/gwat.12656.

Kollet, Stefan J. and Reed M. Maxwell (July 2006). "Integrated surface groundwater flow modeling: A free-surface overland flow boundary condition in a parallel groundwater flow model". en. In: *Advances in Water Resources* 29.7, pp. 945–958. ISSN: 03091708. DOI: 10.1016/j.advwatres.2005.08.006.

Lam, Dan-Thuy, Jaouher Kerrou, Philippe Renard, Hakim Benabderrahmane, and Pierre Perrochet (2020a). "Conditioning Multi-Gaussian Groundwater Flow Parameters to Transient Hydraulic Head and Flowrate Data With Iterative Ensemble Smoothers: A Synthetic Case Study". In: *Frontiers in Earth Science* 8, p. 202.

Lam, Dan-Thuy, Philippe Renard, Julien Straubhaar, and Jaouher Kerrou (Feb. 2020b). "Multiresolution Approach to Condition Categorical Multiple-Point Realizations to Dynamic Data With Iterative Ensemble Smoothing". In: *Water Resources Research* 56.2. DOI: 10.1029/2019wr025875.

Lang, Annika and Jürgen Potthoff (Jan. 2011). "Fast simulation of Gaussian random fields". In: *Monte Carlo Methods and Applications* 17.3. DOI: 10.1515/mcma.2011.009.

Langevin, Christian D, Joseph D Hughes, Edward R Banta, Richard G Niswonger, Sorab Panday, and Alden M Provost (2017). *Documentation for the MODFLOW 6 groundwater flow model*. Tech. rep. US Geological Survey.

Leeuwen, Peter Jan van and Geir Evensen (Dec. 1996). "Data Assimilation and Inverse Methods in Terms of a Probabilistic Formulation". In: *Monthly Weather Review* 124.12, pp. 2898–2913. DOI: 10.1175/1520-0493(1996)124<2898:daaimi>2.0.co;2.

Lemieux, J.-M., E. A. Sudicky, W. R. Peltier, and L. Tarasov (Feb. 2008). "Dynamics of groundwater recharge and seepage over the Canadian landscape during the Wisconsinian glaciation". en. In: *Journal of Geophysical Research* 113.F1, F01011. ISSN: 0148-0227. DOI: 10.1029/2007JF000838.

Li, Ning, Dennis McLaughlin, Wolfgang Kinzelbach, WenPeng Li, and Xin-Guang Dong (2015). "Using an ensemble smoother to evaluate parameter uncertainty of an integrated hydrological model of Yanqi basin". In: *Journal of Hydrology* 529, pp. 146–158.

Linde, Niklas, Andrew Binley, Ari Tryggvason, Laust B. Pedersen, and André Revil (2006). "Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltime data". In: *Water Resources Research* 42.12. DOI: 10.1029/2006WR005131.

Linde, Niklas and Joseph Doetsch (Apr. 2016). "Joint Inversion in Hydrogeophysics and Near-Surface Geophysics". In: *Integrated Imaging of the Earth*. John Wiley & Sons, Inc, pp. 117–135. DOI: 10.1002/9781118929063.ch7.

Linde, Niklas, Philippe Renard, Tapan Mukerji, and Jef Caers (Dec. 2015). "Geological realism in hydrogeological and geophysical inverse modeling: A review". In: *Advances in Water Resources* 86, pp. 86–101. DOI: 10.1016/j.advwatres.2015.09.019.

Mariethoz, G and J Caers (2014). *Multiple-Point Geostatistics: stochastic modeling with training images*. Wiley-Blackwell. A John wiley and Sons, LTD, Publication.

Mariethoz, Gregoire, Matthew F McCabe, and Philippe Renard (2012). "Spatiotemporal reconstruction of gaps in multivariate fields using the direct sampling approach". In: *Water Resources Research* 48.10. DOI: 10.1029/2012WR012115.

Mariethoz, Gregoire, Philippe Renard, and Julien Straubhaar (2010a). "The direct sampling method to perform multiple-point geostatistical simulations". In: *Water Resources Research* 46.11, pp. 1–14. DOI: 10.1029/2008WR007621.

Mariethoz, Grégoire, Philippe Renard, and Jef Caers (Nov. 2010b). "Bayesian inverse problem and optimization with iterative spatial resampling". In: *Water Resources Research* 46.11. DOI: 10.1029/2010wr009274.

Matheron, G. (1963). "Principles of geostatistics". In: *Economic Geology* 58.8, pp. 1246–1266. DOI: 10.2113/gsecongeo.58.8.1246.

Meerschman, Eef, Guillaume Pirot, Gregoire Mariethoz, Julien Straubhaar, Marc Van Meirvenne, and Philippe Renard (2013). "A practical guide to performing multiple-point statistical simulations with the Direct Sampling algorithm". In: *Computers and Geosciences* 52. Publisher: Elsevier, pp. 307–324. DOI: 10.1016/j.cageo.2012.09.019.

Messan, Ange Kouami (1992). "Hydrogéologie de la vallée de l'Aar entre Berne et Thoune". MA thesis. Neuchatel: University of Neuchatel.

Mosegaard, Klaus and Albert Tarantola (1995). "Monte Carlo sampling of solutions to inverse problems". In: *Journal of Geophysical Research: Solid Earth* 100.B7, pp. 12431–12447.

NAQUA (2022). *État et évolution des eaux souterraines en Suisse*. Tech. rep. Office fédéral de l'environnement OFEV.

Neven, Alexis, Anders Vest Christiansen, and Philippe Renard (Oct. 2022a). "Automatic stochastic 3D clay fraction model from tTEM survey and bore-

hole data". In: *Scientific Reports* 12.1. DOI: 10.1038/s41598-022-21555-z.

Neven, Alexis, Valentin Dall'Alba, Przemysław Juda, Julien Straubhaar, and Philippe Renard (2021a). "Ice volume and basal topography estimation using geostatistical methods and ground-penetrating radar measurements: application to the Tsanfleuron and Scex Rouge glaciers, Swiss Alps". In: *The Cryosphere*, pp. 5169–5186. DOI: 10.5194/tc-15-5169-2021.

Neven, Alexis, Pradip K. Maurya, Anders Vest Christiansen, and Philippe Renard (Nov. 2020). *tTEM20AAR: a tTEM geophysical dataset*. type: dataset. DOI: 10.5281/ZENODO.4269887.

Neven, Alexis, Pradip Kumar Maurya, Anders Vest Christiansen, and Philippe Renard (June 2021b). "tTEM20AAR: a benchmark geophysical data set for unconsolidated fluvioglacial sediments". In: *Earth System Science Data* 13.6, pp. 2743–2752. DOI: 10.5194/essd-13-2743-2021.

Neven, Alexis and Philippe Renard (2023). *ArchPy / ES-MDA inverse Results*. DOI: 10.5281/ZENODO.8047723.

Neven, Alexis, Ludovic Schorpp, and Philippe Renard (Oct. 2022b). "Stochastic multi-fidelity joint hydrogeophysical inversion of consistent geological models". In: *Frontiers in Water* 4. DOI: 10.3389/frwa.2022.989440.

Oriani, Fabio, Andrea Borghi, Julien Straubhaar, Grégoire Mariethoz, and Philippe Renard (2016). "Missing data simulation inside flow rate time-series using multiple-point statistics". In: *Environmental Modelling & Software* 86, pp. 264–276. DOI: 10.1016/j.envsoft.2016.10.002.

Paine, Jeffrey G. (Mar. 2003). "Determining salinization extent, identifying salinity sources, and estimating chloride mass using surface, borehole, and airborne electromagnetic induction methods". In: *Water Resources Research* 39.3. DOI: 10.1029/2001wr000710.

Pirot, Guillaume, Julien Straubhaar, and Philippe Renard (2014). "Simulation of braided river elevation model time series with multiple-point statistics". In: *Geomorphology* 214, pp. 148–156.

Preusser, Frank, Hans Rudolf Graf, Oskar Keller, Edgar Krayss, and Christian Schlüchter (July 2011). "Quaternary glaciation history of northern Switzerland". In: *Quaternary Science Journal* 60.2/3, pp. 282–305. DOI: 10.3285/eg.60.2-3.06.

Preusser, Frank and Christian Schlüchter (Aug. 2004). "Dates from an important early Late Pleistocene ice advance in the Aare valley, Switzerland". In: *Eclogae Geologicae Helvetiae* 97.2, pp. 245–253. DOI: 10.1007/s00015-004-1119-4.

Pyrcz, Michael J and Clayton V Deutsch (2014). *Geostatistical reservoir modeling*. Oxford university press, p. 449.

Reuschen, Sebastian, Fabian Jobst, and Wolfgang Nowak (Aug. 2021). "Efficient Discretization-Independent Bayesian Inversion of High-Dimensional Multi-Gaussian Priors Using a Hybrid MCMC". In: *Water Resources Research* 57.8. DOI: 10.1029/2021wr030051.

Ringrose, Philip and Mark Bentley (2015). *Reservoir Model Design*. Springer Netherlands, p. 249. DOI: 10.1007/978-94-007-5497-3.

Sandersen, Peter BE, Anders J Kallesøe, Ingelise Møller, Anne-Sophie Høyer, Flemming Jørgensen, Jesper B Pedersen, and Anders V Christiansen (2021). "Utilizing the towed Transient ElectroMagnetic method (tTEM) for achieving unprecedented near-surface detail in geological mapping". In: *Engineering Geology* 288, p. 106125. DOI: 10.1016/j.enggeo.2021.106125.

Schlüchter, Christian (Jan. 1989). "The most complete quaternary record of the Swiss Alpine Foreland". In: *Palaeogeography, Palaeoclimatology, Palaeoecology* 72, pp. 141–146. DOI: 10.1016/0031-0182(89)90138-7.

Scholl, Carsten, Stefan L. Helwig, Bülent Tezkan, Mark Goldman, and Uri Kafri (Jan. 2009). "1-D multimodel joint inversion of TEM-data over multidimensional structures". In: *Geophysical Journal International* 176.1, pp. 81–94. DOI: 10.1111/j.1365-246x.2008.03973.x.

Schorpp, Ludovic, Julien Straubhaar, and Philippe Renard (May 2022). "Automated Hierarchical 3D Modeling of Quaternary Aquifers: The ArchPy Approach". In: *Frontiers in Earth Science* 10. DOI: 10.3389/feart.2022.884075.

Sophocleous, M.A, J.K Koelliker, R.S Govindaraju, T Birdie, S.R Ramireddygari, and S.P Perkins (Jan. 1999). "Integrated numerical modeling for basin-wide water management: The case of the Rattlesnake Creek basin in south-central Kansas". en. In: *Journal of Hydrology* 214.1-4, pp. 179–196. ISSN: 00221694. DOI: 10.1016/S0022-1694(98)00289-3.

Stafleu, J., D. Maljers, J.L. Gunnink, A. Menkovic, and F.S. Busschers (Dec. 2011). "3D modelling of the shallow subsurface of Zeeland, the Netherlands". In: *Netherlands Journal of Geosciences - Geologie en Mijnbouw* 90.4, pp. 293–310. DOI: 10.1017/s0016774600000597.

Straubhaar, J. (2019). *DeeSse user's guide*. Tech. rep. The Centre for Hydrogeology and Geothermics (CHYN), University of Neuchâtel: Neuchâtel, Switzerland.

Straubhaar, Julien, Philippe Renard, and Tatiana Chugunova (2020). "Multiple-point statistics using multi-resolution images". In: *Stochastic Environmental Research and Risk Assessment*. Publisher: Springer, pp. 1–23. DOI: 10.1007/s00477-020-01770-8.

Strebelle, Sebastien (2002). "Conditional simulation of complex geological structures using multiple-point statistics". In: *Mathematical geology* 34.1, pp. 1–21. DOI: 10.1023/A:1014009426274.

Strebelle, Sebastien, Karen Payrazyan, and Jef Caers (2002). "Modeling of a Deepwater Turbidite Reservoir Conditional to Seismic Data Using Multiple-Point Geostatistics". In: *SPE Annual Technical Conference and Exhibition*. DOI: 10.2118/77425-MS.

Tarantola, Albert (Jan. 2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. Society for Industrial and Applied Mathematics. DOI: 10.1137/1.9780898717921.

United Nations Educational Scientific and Cultural Organization (Apr. 2022). *The united nations world water development report 2022 the united nations world water development report 2022*. The United Nations World Water

Development Report. Paris, France: United Nations Educational Scientific and Cultural.

Vallat, Marie (Nov. 2022). "Aquifère supérieur de l'Aar : Analyse de la piézométrie et caractérisation de la conductance du lit de l'Aar". MA thesis. Neuchâtel, Switzerland: University of Neuchâtel.

Varga, Miguel de la, Alexander Schaaf, and Florian Wellmann (Jan. 2019). "GemPy 1.0: open-source stochastic geological modeling and inversion". In: *Geoscientific Model Development* 12.1, pp. 1–32. DOI: 10.5194/gmd-12-1-2019.

Viezzoli, Andrea, Anders Vest Christiansen, Esben Auken, and Kurt Sørensen (May 2008). "Quasi-3D modeling of airborne TEM data by spatially constrained inversion". In: *GEOPHYSICS* 73.3, F105–F113. DOI: 10.1190/1.2895521.

Vilhelmsen, Troels Norvin, Esben Auken, Anders Vest Christiansen, Adrian Sanchez Barfod, Pernille Aabye Marker, and Peter Bauer-Gottwein (July 2019). "Combining Clustering Methods With MPS to Estimate Structural Uncertainty for Hydrological Models". In: *Frontiers in Earth Science* 7, p. 181. ISSN: 2296-6463. DOI: 10.3389/feart.2019.00181.

Volken, Stefan, Giona Preisig, and Michael Gaehwiler (Oct. 2016). "GeoQuat: Developing a system for the sustainable management, 3D modelling and application of Quaternary deposit data". In: *Swiss Bulletin for Applied Geology* 21, pp. 3–16. DOI: 10.5169/seals-658182.

Wang, Lijing, Peter K. Kitanidis, and Jef Caers (Apr. 2022). "Hierarchical Bayesian Inversion of Global Variables and Large-Scale Spatial Fields". In: *Water Resources Research* 58.5. DOI: 10.1029/2021wr031610.

Wellmann, Florian and Guillaume Caumon (2018). "3-D Structural geological models: Concepts, methods, and uncertainties". In: *Advances in Geophysics*. Elsevier, pp. 1–121. DOI: 10.1016/bs.agph.2018.09.001.

Wellmann, J. Florian, Miguel de la Varga, Ruth E. Murdie, Klaus Gessner, and Mark Jessell (Oct. 2017). "Uncertainty estimation for a geological model of the Sandstone greenstone belt, Western Australia – insights from integrated geological and geophysical inversion in a Bayesian inference framework". In: *Geological Society, London, Special Publications* 453.1. Publisher: Geological Society of London, pp. 41–56. DOI: 10.1144/sp453.12.

Wen, X.-H. and W.H. Chen (2005). "Real-Time Reservoir Model Updating Using Ensemble Kalman Filter". In: *SPE Reservoir Simulation Symposium*. The Woodlands, Texas: SPE, SPE 92991. DOI: 10.2118/92991-ms.

White, Jeremy T. (Nov. 2018). "A model-independent iterative ensemble smoother for efficient history-matching and uncertainty quantification in very high dimensions". In: *Environmental Modelling & Software* 109, pp. 191–201. DOI: 10.1016/j.envsoft.2018.06.009.

Xu, Teng, J Jaime Gómez-Hernández, Zi Chen, and Chunhui Lu (2021). "A comparison between ES-MDA and restart EnKF for the purpose of the simultaneous identification of a contaminant source and hydraulic conductivity". In: *Journal of Hydrology* 595, p. 125681.

Xu, Teng, Wenjun Zhang, J Jaime Gómez-Hernández, Yifan Xie, Jie Yang, Zi Chen, and Chunhui Lu (2022). "Non-point contaminant source identification in an aquifer using the ensemble smoother with multiple data assimilation". In: *Journal of Hydrology* 606, p. 127405.

Zheng, Qiang, Jiangjiang Zhang, Wenjie Xu, Laosheng Wu, and Lingzao Zeng (2019). "Adaptive multifidelity data assimilation for nonlinear subsurface flow problems". In: *Water Resources Research* 55.1, pp. 203–217.

Zhou, Haiyan, J Jaime Gomez-Hernandez, Harrie-Jan Hendricks Franssen, and Liangping Li (2011). "An approach to handling non-Gaussianity of parameters and state variables in ensemble Kalman filtering". In: *Advances in Water Resources* 34.7, pp. 844–864.