



UNIVERSITÀ DI PARMA

UNIVERSITÀ DEGLI STUDI DI PARMA

Dottorato di Ricerca in Ingegneria Civile e Architettura
XXXVII Ciclo

Mapping the Invisible: Geostatistical Breakthroughs from Rainfall Analysis to Aquifer Characterization

Coordinatore:
Prof. Andrea Spagnoli

Tutor:
Prof. Andrea Zanini

Co-tutor:
Prof. Leonardo Azevedo

Dottoranda: Camilla Fagandini

Anni Accademici 2021/2022 – 2023/2024

Abstract

This thesis presents innovative methodologies to support groundwater management, focusing on geostatistical methods and inverse modeling procedures. It addresses three key research areas: gap-filling approaches for daily precipitation time series, inversion modeling techniques for subsurface applications, and experimental investigations in controlled laboratory environments.

The research begins by introducing a first-time application of geostatistical-based daily precipitation gap filling, tested against a well-established linear interpolation approach. The comparison was demonstrated in the province of Mantova (Italy) and in southern Portugal, but can be easily extended to any area of interest. By introducing monthly semi-variograms instead of daily ones, computational efficiency significantly improves without compromising the accuracy of precipitation estimates. The study also presents the strengths and weaknesses of the application of each procedure.

The work then presents three key distinct applications of inverse modeling techniques for subsurface characterization. These applications are demonstrated in synthetic cases through the integration of hydrological and geophysical data. By applying methods such as the Ensemble Smoother with Multiple Data Assimilation, the research demonstrates the power of inverse modeling in estimating complex subsurface properties, such as hydraulic conductivity and electrical resistivity. The method was evaluated (first application) and compared with another geostatistical technique (second application). The traditional geophysical forward model compo-

ment of the Ensemble Smoother with Multiple Data Assimilation was replaced by a Convolutional Neural Network to optimize processing time (third application). This innovation represents the chapter's key contribution. The studies reveal that coupling hydrological data with geophysical measurements improves the accuracy and reliability of subsurface models. This combined approach significantly reduces the uncertainty in the estimates, providing a powerful tool for more accurate and detailed aquifer characterization.

Finally, the thesis focuses on experimental sandbox investigations. In these experiments, electrical resistivity tomography data and concentration measurements were acquired to estimate subsurface properties in a controlled laboratory environment. The results from the sandbox experiments confirm the findings of the synthetic studies. This research constitutes a significant advance in the integration of geophysical and hydrological methodologies to improve the characterization of subsurface properties, marking the first application of this procedure to laboratory data. Furthermore, it establishes a robust foundation for future experimental studies and suggests several potential improvements in experimental design, data acquisition, and integration of diverse datasets for more accurate subsurface characterization.

In conclusion, the work makes significant methodological and practical contributions to support groundwater characterization. It demonstrates the effectiveness of advanced geostatistical, inverse modeling, and laboratory techniques to address key challenges in subsurface and hydrological modeling. The findings suggest several promising directions for future research.

Contents

Introduction	1
1 Comparative Analysis of Gap-Filling Methods for Daily Precipitation Time Series	5
1 Method	8
1.1 Food and Agriculture Organization (FAO) technique	9
1.2 Geostatistical Approach	10
2 Application	14
2.1 Case 1: Mantova province, Italy	17
2.2 Case 2: Guadiana River basin, Portugal	24
3 Concluding remarks	46
2 Inversion Modeling Techniques: From Theory to Practice in Sub-surface Applications	49
1 Material and Method	53
1.1 Fundamentals of groundwater hydraulics	53
1.2 Fundamentals of Electrical Resistivity Tomography (ERT) . .	54
1.2.1 Petrophysical relationship	58
1.3 Convolutional Neural Network	59
1.4 Ensemble Smoother with Multiple Data Assimilation (ES-MDA)	60
1.5 Empirical Bayes and Akaike's Bayesian Information Criterion	62
2 Applications	66

2.1	Evaluation of the Hydraulic Conductivity	66
2.1.1	Synthetic Case	67
2.1.2	Results	70
2.1.3	Concluding remarks	75
2.2	Comparative Analysis of Ensemble Smoother with Multiple Data Assimilation and Empirical Bayes PEST: Methodological Evaluation	77
2.2.1	Description of the methods	77
2.2.2	Synthetic Case	81
2.2.3	Results	83
2.2.4	Concluding remarks	88
2.3	Reconstruction of Contaminant Plumes through Coupled Hy- drogeophysical Inversion	91
2.3.1	Coupled Hydrogeophysical Inversion Scheme	92
2.3.2	Synthetic Case	94
2.3.3	Results	101
2.3.4	Concluding remarks	111
3	Experimental Sandbox Investigation Using Electrical Resistivity Tomography and Tracer Tests for Subsurface Characterization	113
1	Experimental installation	115
2	Data acquisition system	117
2.1	Electrical Resistivity Tomography (ERT)	117
2.2	Concentration	123
3	Inversion process	127
3.1	Case 1: Estimation of the Hydraulic Conductivity Field Using Concentration Data	130
3.2	Case 2: Estimation of the Resistivity Field Using Apparent Resistivity Data	133
4	Future developments	136

Overall conclusions	139
APPENDIX A: Evaluation Metrics	143
APPENDIX B: Arduino code for ERT measurements	145
APPENDIX C: MATLAB codes for ERT processing	151
Bibliography	160

List of Figures

1.1	Study area (UTM zone 32N-EPSSG: 32632). Yellow dots represent the locations of available precipitation stations, the diamond marks the station meeting FAO criteria, and blue dots indicate ERACLITO dataset stations. Elevation is shown using a color scale DEM.	18
1.2	Experimental and fitted semi-variogram models based on daily average rainfall data from the 27 monitoring stations.	21
1.3	Cumulative monthly rainfall for 1977 at the Saino di Pegognaga station: observed values and estimates based on FAO and OK methods.	22
1.4	Daily cross-validation results for 1977 at the Saino di Pegognaga station.	24
1.5	Study area (UTM zone 29N-EPSSG: 32629). Yellow dots represent the locations of available precipitation stations and the diamond marks the stations meeting FAO criteria. Elevation is shown using a colour scale DEM.	26
1.6	Experimental and fitted semi-variogram models based on daily average rainfall data from 60 monitoring stations.	29
1.7	Bar chart displaying the mean daily precipitation for each month from 1976 to 2005 in the study area.	31
1.8	Experimental and fitted semi-variogram models based on the elevation DEM.	32

1.9	Experimental cross-variograms (black dots) with the fitted cross-variogram models (solid line) and positive-definite condition curve (dashed line) based on the rainfall (i.e., the primary variable) and elevation data (i.e., the secondary variable).	33
1.10	Mean error computed using the applied approaches (FAO, OK, SCK) at precipitation stations, presented for daily (top) and monthly (bottom) estimation scales.	37
1.11	Mean absolute error computed using the applied approaches (FAO, OK, SCK) at precipitation stations, presented for daily (top) and monthly (bottom) estimation scales.	38
1.12	Root mean square error computed using the applied approaches (FAO, OK, SCK) at precipitation stations, presented for daily (top) and monthly (bottom) estimation scales.	39
1.13	Cumulative monthly rainfall for the year 1985, observed and estimated using FAO, OK, and SCK methods across the 12 precipitation stations.	41
1.14	Daily cross-validation results for 1985 at the selected stations.	43
1.15	Estimation of error variance for daily precipitation in January using OK (a), SCK (b), and the variance difference between OK and SCK (c). Crosses indicate the locations of the rainfall data, while circles represent the locations of elevation data (the secondary variable in cokriging).	45
2.1	Scheme of a conventional four-electrode array for measuring subsurface resistivity.	56
2.2	2D synthetic model of hydraulic conductivity.	68
2.3	Concentration monitoring points overlaid on the reference hydraulic conductivity model (top) and observed apparent resistivity (bottom).	69
2.4	Log10-estimated hydraulic conductivity fields for the three scenarios.	72
2.5	Standard deviation of the log10-estimated hydraulic conductivity fields for the three scenarios.	74

2.6	Workflow of the ES-MDA method.	79
2.7	Workflow of the ebaPEST method.	81
2.8	2D synthetic resistivity model. Black dots represent the locations of recorded apparent resistivity, while the red lines indicate the positions of the resistivity monitoring points.	82
2.9	Comparison of estimated results using the ES-MDA vs. ebaPEST.	84
2.10	Actual vs. estimated observations using ES-MDA and ebaPEST methods. The observed apparent resistivity values are represented by circles (black), while blue dots indicate the actual resistivity (observed at monitoring points).	87
2.11	Comparison of the logarithmic estimated standard deviation using the ES-MDA vs. ebaPEST.	88
2.12	Flowchart of the coupled hydrogeophysical inverse model	92
2.13	Hydraulic conductivity and concentration reference maps. The red grid cells represent the left and right boundary conditions. The cross indicates the injector position.	95
2.14	Reference resistivity model and observed pseudo-section. The cross indicates the position of the injector.	96
2.15	CNN Validation, the dashed line is the 1:1 line.	100
2.16	Estimated concentration map distribution for the five scenarios. The injector is marked by a cross. Red circles represent the locations of observation wells.	104
2.17	Scatter plot of the estimated vs. observed concentration for all scenarios.	105
2.18	Observed-Estimated apparent resistivity and concentration for scenarios 2-5.	107
2.19	Standard deviation maps of the estimated concentrations for the five scenarios. The cross denotes the injector. The observation wells are visually depicted by the red circles.	109

2.20	a) Estimated concentration distribution, Scenario 3 – SimPEG forward model. The injector is marked by a cross. Red circles represent the locations of observation wells. b) Differences between estimated concentrations (CNN vs SimPEG forward model).	111
3.1	Overview of the laboratory equipment.	116
3.2	ERT acquisition device.	119
3.3	Average resistance (above). Standard deviation resistance (bottom).	120
3.4	Scatter plot of apparent resistivity (top) and interpolated pseudo-section of apparent resistivity (bottom).	121
3.5	Sandbox filled with homogeneous sand with an impermeable, non-conductive barrier. The red rectangle identifies the ERT investigation area.	122
3.6	Scatter plot of observed resistance (top) and apparent resistivity (bottom).	123
3.7	Close-up view of the field sandbox.	124
3.8	Example of dimensionless, from blue (0) to yellow (1), concentration distribution at four time steps: 100 s, 250 s, 450 s, and 900 s.	126
3.9	Breakthrough curves for the four monitoring points located within the sandbox.	127
3.10	Log10 hydraulic conductivity (top) and resistivity (bottom) fields. The white cross in the hydraulic conductivity plot marks the tracer injection point, and the white circles indicate the locations of breakthrough monitoring points.	129
3.11	Average of the 500 realizations of the estimated hydraulic conductivity field (top) and respective standard deviation map (bottom), plotted in $\log_{10}(\text{cm/s})$. The white cross marks the location of the tracer injector, while the white circles denote the monitoring points where observations were collected.	131

3.12 Comparison between the observed and the estimated breakthrough curves for the four monitoring points located within the sandbox. . .	133
3.13 Observed apparent resistivity.	134
3.14 Average of the 500 realizations of the estimated resistivity fields (top) and respective standard deviation map (bottom).	135
3.15 Average of the 500 realizations of the estimated resistivity fields. . .	136

List of Tables

- 1.1 Theoretical semi-variogram models. C_0 is the nugget effect, θ is the sill value, l is the range, and h is the distance between pairs of points. 15
- 1.2 Station Id, Station Name, Data (%) and Elevation (m.a.s.l.) in the reference period (1976-2005). 19
- 1.3 Estimated parameters for the monthly semi-variograms of average daily precipitation, modeled using an exponential model. 20
- 1.4 Station Id, Data (%), and Elevation (m.a.s.l.) in the reference period (1976-2005). 27
- 1.5 Squared Pearson correlation coefficient (r^2) between each selected monitoring station and its best-correlated stations. 28
- 1.6 Estimated parameters for the monthly semi-variograms on the first variable (i.e., average daily precipitation) with an exponential model. 30
- 1.7 Estimated parameters for the semi-variogram on the second variable (i.e., elevation) with an exponential model. 31
- 1.8 Average correlation coefficients between the primary variable (precipitation) and secondary variable (elevation) for each month. 34
- 1.9 Estimated parameters for the monthly cross-variograms of average daily precipitation and elevation, modeled using an exponential function. 34

1.10	Evaluation metrics for the proposed methods in estimating daily rainfall (mm/d) in 1985 across the 12 monitoring stations. Values in bold indicate minimum <i>MAE</i> and <i>RMSE</i>	35
1.11	Evaluation metrics for the proposed methods in estimating monthly rainfall (mm/month) in 1985 across the 12 monitoring stations. Values in bold indicate minimum <i>MAE</i> and <i>RMSE</i>	36
2.1	Summary of the scenarios. Number of observations used.	70
2.2	Evaluation of the inversion approach performance assessed on the estimation of the hydraulic conductivity (cm/s) for each scenario. . .	72
2.3	Assessment of observation estimates performance for concentration (mg/L) and resistivity (Ohm·m) across each scenario.	73
2.4	Performance metrics for ES-MDA and ebaPEST	85
2.5	Comparison of <i>RMSE</i> for apparent resistivity (Ohm·m) and resistivity (Ohm·m).	86
2.6	Flow and transport model parameters.	95
2.7	Geophysical and petrophysical parameters.	96
2.8	Summary of the scenarios. Number of observations used.	97
2.9	CNN architecture.	99
2.10	Performance of the proposed approach assessed on the estimation of the contaminant plume (mg/L) for each scenario.	103
3.1	Properties of sand and gravel used in the laboratory experiment. . .	117

Introduction

Preserving water quality and quantity is essential for human well-being and environmental sustainability. However, rapid urbanization, industrial growth, and intensive agricultural practices have significantly impacted the surface and subsurface across the globe. The increasing demand for potable and irrigation water underscores the urgent necessity for sustainable management approaches that balance human needs with the preservation of water resources. Groundwater, a critical freshwater source in many regions, is particularly important. Safeguarding this resource is vital not only for current needs but also to ensure its availability for future generations.

Understanding geometry and properties of the aquifer is of main importance in evaluating the availability and quality of the resource. However, monitoring subsurface environments remains challenging due to their inherent complexity and the difficulty in acquiring direct observations. Traditional survey methods tend to be invasive, costly, and time-consuming and often yield limited, localized data. To overcome these limitations, numerical models, developed using techniques like geostatistical methods, can simulate subsurface behavior and provide dynamic subsurface insights, particularly in data-scarce situations. However, the accuracy of these models depends on the quality and completeness of their underlying data. Where data gaps exist, interpolation techniques and kriging-based methods are used to estimate missing information, ensuring that models yield meaningful insights into subsurface systems.

This work aims to explore and apply innovative aquifer investigation techniques,

with a focus on inverse procedures. The work is organized into three main sections, each examining complementary aspects to provide a methodological framework to address practical problems in aquifer resource management. The results and conclusions of the specific case studies are also presented.

The **first section** introduces a novel approach to fill gaps in daily precipitation time series, a crucial factor for aquifer recharge. Accurate and complete data are essential for hydrogeological models, as gaps or inaccuracies can lead to significant errors in model predictions and impact water resource management decisions. Geostatistical approaches often generate a semi-variogram for each day of missing data, leading to substantial computational costs for extended time series. Fagandini et al. (2024b) propose a more efficient procedure by employing monthly semi-variograms fitted to the average daily rainfall of all available meteorological stations for each month in the reference period. This approach allows the use of only 12 monthly semi-variograms instead of generating one for each day of the gap period, hence the reduction of the computational time. Missing daily precipitation is subsequently estimated using ordinary kriging and simple cokriging, with the semi-variograms corresponding to the month of interest and elevation data as the secondary variable. Based on investigations conducted in two distinct study areas—the province of Mantova in Italy and a River basin in Portugal—the results are compared with a simple linear interpolation approach. Simple cokriging was excluded from the analysis of the first region because its predominantly flat terrain minimizes the influence of elevation. By addressing key limitations of previous studies and providing a computationally efficient solution, this innovative methodology contributes to the advancement of gap-filling techniques and improves the accuracy of precipitation time series analysis, offering a valuable tool for hydrogeological modeling in diverse geographic contexts.

The **second section** focuses on the application of inverse modeling techniques, which are powerful tools for solving complex estimation problems by using observed measurements to infer subsurface properties. In recent years, hydrogeophysics has

emerged as a valuable, non-invasive, and cost-effective approach in hydrogeology, helping to address challenges associated with traditional survey methods. Geophysical data improves the understanding of hydrological processes and the geological structures that govern the subsurface. In this section, both hydrological and geophysical data are leveraged to characterize aquifer properties in three distinct applications. The first application serves as an initial test to understand and validate the Ensemble Smoother with Multiple Data Assimilation approach. In this case, hydraulic conductivity is estimated using hydrological and geophysical data, with results compared between different datasets. This work was carried out during a visiting PhD period at the Instituto Superior Técnico (IST) in Lisbon. The second application compared two methodologies—the Ensemble Smoother with Multiple Data Assimilation and the empirical Bayesian approach PEST—to estimate the electrical resistivity of the subsurface, highlighting differences in their performance (Fagandini et al., 2025). However, the third application represents the key innovation of this section. A new coupled hydrogeophysical methodology, proposed by Fagandini et al. (2024a), is applied to estimate a contaminant plume within the aquifer. This work replaces the traditional forward model with a Convolutional Neural Network to significantly speed up the solution process. Five scenarios are analyzed to assess the impact of multiple data sources, demonstrating the enhanced efficiency and potential of this novel approach.

The **third section** addresses laboratory experiments carried out at the Department of Engineering and Architecture of the University of Parma. One of the key challenges in groundwater hydrology is the limited availability of experimental data from controlled flow environments. To tackle this, a sandbox experiment is utilized, integrating electrical resistivity tomography data with concentration measurements at monitoring points. The primary goal is to estimate the hydraulic conductivity field of the sandbox by employing a coupled hydrogeophysical inversion approach (Fagandini et al., 2024a), specifically applying the Ensemble Smoother with Multiple Data Assimilation. This study marks a notable advancement, as it is the first to

apply this procedure to laboratory data. It illustrates the feasibility and potential of this method in controlled experimental settings for integrating hydrological and geophysical techniques. Furthermore, the study employs a custom-designed, non-commercial ERT system, specifically tailored for laboratory experiments (Delmonte et al., 2024). This innovation ensures flexibility and adaptability in data acquisition, enhancing the quality and relevance of the results. By combining a novel methodological application with purpose-built instrumentation, this section highlights an important step forward in the development of advanced hydrogeophysical techniques for groundwater research.

Finally, the thesis includes a comprehensive conclusion and recommendations for future research.

1

Comparative Analysis of Gap-Filling Methods for Daily Precipitation Time Series

This chapter reports the findings and applies the new procedure proposed by Fagandini et al. (2024b), designed to fill daily gaps in precipitation time series through computationally efficient geostatistical methods.

Gap-filling procedures are typically specific to the nature of the variable under study. Precipitation, due to its high spatio-temporal variability and the numerous interconnected variables involved, is one of the most challenging atmospheric variables to characterize, estimate, and forecast on a daily scale (Portuguez-Maurtua et al., 2022). Several interpolation methods have been developed to estimate missing values in precipitation time series. Thiessen Polygon (Thiessen, 1911 - THI) assigns the precipitation of the area of influence containing the unsampled location. Similarly, the Natural Neighbour (Sibson, 1981 - NN) estimates values at an unsampled point by averaging nearby data, with weights based on the overlap between

their areas of interest and that of the unsampled location. In contrast, the Nearest Neighbour (Brandsma and Buishand, 1998 - NNb) determines the closest point and directly assigns its value without performing an averaging calculation. Another approach, the Inverse Distance Weighted (U.S. National Weather Service, 1972 - IDW), calculates a value using a weighted average of the available data, where the weights are inversely proportional to the distance between the unknown and known data locations. Furthermore, the linear regression method proposed by the Food and Agriculture Organization (FAO) of the United Nations (Allen et al., 1998) fills gaps by using data from stations with a high correlation coefficient to the station of interest. This method is commonly applied in climate analysis (D’Oria et al., 2019; Huang et al., 2020).

Geostatistical modeling offers a preferred option for estimating missing precipitation data, as it incorporates spatial correlation of observations (Goovaerts, 1997) and allows for uncertainty assessment in the estimation. Kriging represents the most widely used geostatistical method for rainfall interpolation. It has been shown to provide better results than traditional gap-filling methods due to the incorporation of spatial continuity patterns in the predictions (Tabios and Salas, 1985; Bacchi and Kottegoda, 1995; Campling et al., 2001; Bostan et al., 2012). Geostatistical methods can also incorporate one or more secondary variables to improve estimation accuracy. Cokriging and kriging with an external drift (KED) are part of the multivariate geostatistical methods. Precipitation often correlates with elevation, especially in mountainous terrains (Hevesi et al., 1992), making elevation data from digital elevation models (DEMs) a common secondary variable to enhance rainfall estimation (Hevesi et al., 1992; Daly et al., 1994; Prudhomme and Reed, 1999; Goovaerts, 2000; Lloyd, 2005; Carrera-Hernández and Gaskin, 2007; Jacquin and Soto-Sandoval, 2013; Adhikary et al., 2017; Cheng et al., 2017).

Only a few studies have applied geostatistical methods to estimate daily rainfall data. Beek et al. (1992) applied kriging to interpolate precipitation data over four days in 1984 in North-Western Europe, aiming to investigate the spatial variability

of daily precipitation. Buytaert et al. (2006) analyzed the spatio-temporal patterns of daily precipitation series in the South of Ecuador, comparing THI and Universal kriging (UK) to point out differences in uncertainty assessment in mountainous terrains. Carrera-Hernández and Gaskin (2007) compared five different kriging methods (Ordinary kriging - OK, kriging with External Drift - KED, Block kriging with External Drift - BKED, Ordinary kriging in a local neighbourhood—OKI, and kriging with External Drift in a local neighbourhood - KEDI) to estimate three daily climatological variables (rainfall, minimum temperature, and maximum temperature). Elevation was considered the second variable in the external drift. Chen et al. (2010) compared five interpolation methods (ordinary nearest neighbour, local polynomial, radial basis function, IDW, and OK based on seasonal semi-variograms) to produce gridded daily precipitation in China. Ly et al. (2011) compared two deterministic algorithms (IDW and THI) and four geostatistical methods (OK, UK, KED and Ordinary cokriging -OCK; the last two techniques utilized elevation as the secondary variable) using a 30-year dataset. In addition, seven semi-variogram models were fitted for different density daily samples.

Alternative techniques have also been proposed for gap-filling methodologies and compared with kriging-based methods. Oriani et al. (2020) tested various models (IDW, various models based on OK, K-nearest neighbour and a new algorithm called vector sampling) to find an automatic method that requires minimal computational effort. Aguilera et al. (2020) compared three different methodologies (Spatio-temporal kriging - STK, multiple imputations by chained equations through predictive mean matching - PMM, and the random forest -RF- machine learning algorithm) for estimating missing precipitation data.

Most of the studies mentioned above have performed daily-scale estimations focusing on short periods or a limited number of days. Geostatistical methods typically compute a semi-variogram for the day under consideration before estimating precipitation at the desired location (Beek et al., 1992; Carrera-Hernández and Gaskin, 2007; Ly et al., 2011). One of the main problems with daily-scale analysis is the

insufficient data for evaluating semi-variograms, requiring the computation of as many semi-variograms as there are missing data days. This process can become computationally expensive for long time series that are more prone to contain large amounts of gaps. This study aims to identify a suitable and computationally efficient method for filling daily gaps in precipitation time series by comparing three different interpolation approaches: FAO, OK, and Simple cokriging (SCK). This comparison is necessary because no single estimation method can work well everywhere (Daly, 2006). The proposed method overcomes the aforementioned challenges by offering an innovative approach based on the use of monthly average semi-variograms (Fagandini et al., 2024b). Twelve monthly semi-variograms (one for each month) are built using all available data from the entire time series, rather than one for each day of the dataset. This approach reduces computational demand while accounting for seasonal variability, ensuring robust and reliable estimates. In addition, elevation is investigated to improve results as a secondary variable in SCK.

Two case studies are conducted to analyze and compare the proposed approaches. In the first application, only FAO and OK are considered. Specifically, the Mantova dataset is found to be unsuitable for applying SCK, as the region lacks the necessary elevation variation to benefit from this approach. Consequently, the analysis is expanded to a second case study in Portugal, where more complex environmental conditions, including varied terrain and the presence of mountains, provide a better setting for testing SCK alongside FAO and OK. This allows for a more comprehensive comparison of the methods under different geographical conditions.

1. Method

The methodology adopted in this work includes a linear regression method (FAO) and two kriging-based geostatistical techniques (OK and SCK). This section provides a brief overview of the approaches used. The reader is referred to Goovaerts (1997) and Kitanidis (1997) for extensive reviews of geostatistical theory.

1.1. Food and Agriculture Organization (FAO) technique

The linear regression approach, commonly referred to as the FAO method, fills gaps in time series by utilizing data information collected during the gap period from a consistent nearby station.

Given a dataset $\mathbf{Y}$ at a station with missing observations, its historical series is completed using observations from another dataset $\mathbf{X}$ from a nearby homogeneous station with similar characteristics. The method is characterized by a regression of $\mathbf{Y}$ on $\mathbf{X}$ performed for periods when data from both datasets are available:

$$y_i = b + a \cdot x_i, \quad (1.1)$$

where $a = \frac{\text{Cov}_{xy}}{\sigma_x^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$ and $b = \bar{y} - a \cdot \bar{x}$ are empirical regression parameters, and x_i and y_i are individual observations from datasets $\mathbf{X}$ and $\mathbf{Y}$, respectively. A crucial aspect of the FAO method is computing the correlation coefficient between pairs of stations, which highlights the degree of similarity and identifies the best-correlated stations. The Pearson correlation coefficient r is used and is defined as follows:

$$r = \frac{\text{Cov}_{xy}}{\sigma_x \sigma_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (1.2)$$

where $\bar{x}$ and $\bar{y}$ are the means of the two datasets computed from the concurrent observations, σ_x and σ_y are the standard deviations, and $\text{Cov}_{xy} = \frac{\sum_{i=1}^n (x_i \cdot y_i) - n \cdot \bar{x} \cdot \bar{y}}{n-1}$ represents the covariance between datasets $\mathbf{X}$ and $\mathbf{Y}$ at time i , with n being the number of simultaneous observations in the two time series. The FAO method can only be applied if the following three constraints are met:

1. the station to be filled must have at least 70% complete data;
2. the square of the Pearson correlation coefficient (r^2) between the station with missing data and the station used to fill the gaps must be greater than or equal to 0.7;

3. the regression coefficient (a) must be between 0.7 and 1.3.

These criteria serve as indicators of homogeneity, ensuring that the chosen nearby station is suitable for filling in the missing data. When these criteria are respected, the missing values for the period $k = n + 1, n + 2, \dots, m$ can be estimated using the regression equation characterized by parameters a and b from the best-correlated station, as follows:

$$\hat{y}_k = b + a \cdot x_k. \quad (1.3)$$

1.2. Geostatistical Approach

In geostatistics, the primary tool for accounting for the spatial distribution of natural phenomena is the semi-variogram. This function describes how the variance of increments increases with distance.

$$\gamma^*(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{j=1}^{N(\mathbf{h})} (Z(\mathbf{u}_j) - Z(\mathbf{u}_j + \mathbf{h}))^2, \quad (1.4)$$

where $\gamma^*(\mathbf{h})$ is the theoretical semi-variogram at lag $\mathbf{h}$, $N(\mathbf{h})$ is the number of pairs of locations separated by the vector $\mathbf{h}$, and $Z(\mathbf{u}_j)$ is the observed variable at the location $\mathbf{u}_j$ of the monitoring stations J .

Cokriging is a multivariate extension of kriging that utilizes secondary information to enhance the estimation accuracy of the primary variable. If the secondary variable has a strong correlation with the primary variable, incorporating it into the estimation process can reduce prediction error variance compared to using kriging alone (Goovaerts, 2000). The first step of the cokriging method is to develop a suitable model that captures the cross-continuity and dependency between the two variables involved. This positive correlation, known as cross-regionalization, or coregionalization (Goovaerts, 1997), can be quantified using the cross-variogram or

cross-covariogram between the primary and secondary variables, obtained by:

$$\gamma_{z_1 z_2}^*(\mathbf{h}) = \gamma_{z_2 z_1}^*(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{j=1}^{N(\mathbf{h})} [Z_1(\mathbf{u}_j) - Z_1(\mathbf{u}_j + \mathbf{h})] [Z_2(\mathbf{u}_j) - Z_2(\mathbf{u}_j + \mathbf{h})], \quad (1.5)$$

where $\gamma_{z_1 z_2}^*(\mathbf{h}) = \gamma_{z_2 z_1}^*(\mathbf{h})$ represents the experimental cross-variograms between the first observed variable z_1 and the second observed variable z_2 .

The semi-variogram must satisfy the positive-definite condition (PDC). In a univariate context, this condition is met by selecting a semi-variogram or covariogram from the admissible models (Christakos, 1984). Generally, a linear combination of these models can be considered to ensure that the matrix of kriging coefficients is invertible and the variance remains positive (Matheron, 1970). In multivariate geostatistics, considering two variables, a cross-variogram and two semi-variograms are needed. To have a valid model that respects the PDC, the linear model of coregionalization (LMC - Journel and Huijbregts, 1978) can be used. The use of a combination of permissible semi-variogram models, as well as each coregionalization matrix being positive semi-definite, are sufficient conditions for the LMC to be allowable (Goovaerts, 1999). The shape of the cross-variogram is an important consideration in the LMC, as it can affect the accuracy of the model. In general, the shape of the cross-variogram should be the same for all variables included in the LMC (Stein, 1999). This is because the LMC assumes that the variables are correlated with each other consistently across space, and a consistent cross-variogram shape helps ensure that this assumption is met. However, it is possible to use different cross-variogram shapes for different variables in the LMC if there is prior knowledge that the spatial correlation structure differs between variables. In such cases, a separate cross-variogram model can be fitted for each variable (Stein, 1999). To ensure that the cross-variogram model is positive-definite, the Cauchy-Schwarz inequality must be fulfilled for all values of h :

$$\gamma_{z_1 z_2}^*(\mathbf{h}) \leq [\gamma_{z_1}^*(\mathbf{h}) \gamma_{z_2}^*(\mathbf{h})]^{1/2}, \quad (1.6)$$

where $\gamma_{z_1}^*(\mathbf{h})$ and $\gamma_{z_2}^*(\mathbf{h})$ are the experimental semi-variograms on the first variable z_1 and on the second variable z_2 , respectively. A simpler graphical test of the PDC was suggested by Hevesi et al. (1992) by graphing the proposed model with the PDC curve:

$$\text{PDC}(\mathbf{h}) = [\gamma_{z_1}^*(\mathbf{h})\gamma_{z_2}^*(\mathbf{h})]^{1/2}. \quad (1.7)$$

The test is passed if the value of $\gamma_{z_1 z_2}^*(\mathbf{h})$ does not exceed the PDC value for all the values of $\mathbf{h}$, which means the cross semi-variogram is below the PDC curve.

Ordinary kriging

Ordinary kriging (OK) is the most widely used kriging method. It estimates a value at a point in a region with a known semi-variogram by using neighboring observations of the estimation location (Goovaerts, 1997). OK accounts for local mean fluctuations by restricting the domain of mean stationarity to the local neighborhood. OK estimate is provided by:

$$\hat{Z}(\mathbf{u}_0) = \sum_{j=1}^J \lambda_j Z(\mathbf{u}_j) \quad \text{with} \quad \sum_{j=1}^J \lambda_j = 1, \quad (1.8)$$

where $\hat{Z}(\mathbf{u}_0)$ is the estimated value at the target location $\mathbf{u}_0$, λ_j are the kriging weights, and $Z(\mathbf{u}_j)$ are the observed values at J monitoring stations. The OK weights can be obtained by solving the system:

$$\begin{cases} \sum_{j=1}^J \lambda_j \gamma_{ij} - \mu = \gamma_{i0} \\ \sum_{j=1}^J \lambda_j = 1 \end{cases} \quad \text{for } i = 1, \dots, J, \quad (1.9)$$

where γ_{ij} represents the semi-variogram values between sampling locations $\mathbf{u}_i$ and $\mathbf{u}_j$, γ_{i0} are the semi-variogram values between sampling location $\mathbf{u}_j$ and the target location $\mathbf{u}_0$, and μ is the Lagrange multiplier parameter. The constraint that the sum of the weights equals 1 ensures an unbiased estimate. The weights λ_j , obtained through this system, are then used in Eq.1.8 to make the estimation. OK also

provides a measure of uncertainty attached to the results to signify the reliability of the estimation (Goovaerts, 1997). This is achieved by calculating the estimation variance, which reflects the quality of the interpolation. The estimation variance is calculated as follows:

$$\sigma_{OK}^2 = \mathbb{E} \left[(\hat{Z}(\mathbf{u}) - Z(\mathbf{u}))^2 \right], \quad (1.10)$$

where $\mathbb{E}$ denotes the mathematical expectation.

Simple cokriging

Simple cokriging (SCK) assumes that the means of both the primary and secondary variables are known and constant within the study area. The method aims to estimate the primary variable as follows:

$$\hat{Z}_1(\mathbf{u}_0) = \left(\sum_{j=1}^J \lambda_{1j} [Z_1(\mathbf{u}_j) - m_1] + \sum_{k=1}^K \lambda_{2k} [Z_2(\mathbf{w}_k) - m_2] \right) + m_1, \quad (1.11)$$

where $\hat{Z}_1(\mathbf{u}_0)$ is the estimated value of the primary variable at the target location $\mathbf{u}_0$, $Z_1(\mathbf{u}_j)$ are the observed values of the primary variable at locations $\mathbf{u}_j$, $Z_2(\mathbf{w}_k)$ are the observed values of the secondary variable at K locations $\mathbf{w}_k$, λ_{1j} and λ_{2k} are the weights of the first and second variables, m_1 and m_2 are the stationary means of the primary and secondary variables, respectively. The weights are obtained by solving the following system:

$$\begin{cases} \sum_{j_1=1}^J \gamma_{z_1}(\mathbf{u}_{i_1} - \mathbf{u}_{j_1}) \lambda_{1i} + \sum_{j_2=1}^K \gamma_{z_1 z_2}(\mathbf{w}_{i_2} - \mathbf{u}_{j_1}) \lambda_{2i} = \gamma_{z_1}(\mathbf{u}_{j_1} - \mathbf{u}_0) \\ \text{for } i_1 = 1, \dots, J \\ \sum_{j_1=1}^J \gamma_{z_2 z_1}(\mathbf{u}_{i_1} - \mathbf{w}_{j_2}) \lambda_{1i} + \sum_{j_2=1}^K \gamma_{z_2}(\mathbf{w}_{i_2} - \mathbf{w}_{j_2}) \lambda_{2i} = \gamma_{z_1 z_2}(\mathbf{w}_{j_2} - \mathbf{u}_0) \\ \text{for } i_2 = 1, \dots, K \end{cases}, \quad (1.12)$$

where γ_{z_1} and γ_{z_2} are the theoretical semi-variograms of the first and second variables respectively, and $\gamma_{z_1 z_2} = \gamma_{z_2 z_1}$ are the theoretical cross-variograms between the two

variables. Unbiasedness is guaranteed by the equation:

$$\mathbb{E}[\hat{Z}_1(\mathbf{u}) - Z_1(\mathbf{u})] = 0. \quad (1.13)$$

To make the prediction, the weights obtained through the system of Eq.1.12 are used in Eq.1.11.

Two situations can be distinguished according to the sampling density of the primary and secondary variables (Goovaerts, 1999). The first is the equally sampled or isotopic case, where all variables are recorded at every sampled location. The second is the heterotopic case, where the primary variable is under-sampled relative to the secondary. Measurements of the attribute of interest are usually supplemented by more abundant data on attributes related to secondary variables, which generally require less sampling effort. The second case is the recommended one.

2. Application

In the provided application examples, instead of deriving a semi-variogram for every day of the time series (Beek et al., 1992; Carrera-Hernández and Gaskin, 2007; Ly et al., 2011), an average monthly semi-variogram is utilized. This is computed by taking the mean of the daily precipitation observed each month at each monitoring station for all the years of the analyzed period. The monthly-averaged semi-variogram facilitates faster calculations by reducing the number of semi-variograms to twelve instead of one for each day of the missing period. Additionally, it enables the consideration of seasonal variations during data interpolation. The procedure for determining the semi-variograms can be summarized as follows:

1. computing the average daily precipitation for each month at each monitoring station over the entire time series;
2. computing the monthly experimental semi-variogram using data grouped by month from all available gauge stations using Eq.1.4;

3. fitting a theoretical semi-variogram model, $\gamma(h)$, to the experimental semi-variogram. Exponential, Gaussian, and Spherical (Table 1.1) models are commonly used for kriging applications in hydrology (Adhikary et al., 2017). These can be combined with a Nugget model;
4. quantifying the suitability of the theoretical semi-variogram models to the experimental one by computing the sum of the residual squares (RSS):

$$\text{RSS} = \sum_{i=1}^{N_c} [\gamma^*(h_i) - \gamma(h_i)]^2, \quad (1.14)$$

where N_c is the number of intervals into which the data pairs are classified.

Table 1.1. Theoretical semi-variogram models. C_0 is the nugget effect, θ is the sill value, l is the range, and h is the distance between pairs of points.

Model	Equation
Nugget	$\gamma(h) = C_0$
Exponential	$\gamma(h) = \theta \left[1 - \exp\left(-\frac{h}{l}\right) \right]$
Gaussian	$\gamma(h) = \theta \left[1 - \exp\left(-\left(\frac{h}{l}\right)^2\right) \right]$
Spherical	$\gamma(h) = \theta \left[\frac{3}{2} \frac{h}{l} - \frac{1}{2} \left(\frac{h}{l}\right)^3 \right]$

To obtain the best-fitted model, the above procedure is repeated for different lag sizes, with parameters adjusted according to the least-squares methodology until the minimum RSS (Eq.1.14) is reached. The coefficients of the best model can then be used for kriging estimation. Cokriging methods fit semi-variogram and cross-variogram models as a linear combination of the same set of basic models listed in Table 1.1. The same procedure applied for OK is applied to SCK, ensuring that PDC (Eq.1.7) is respected, as described by Adhikary et al. (2017).

Semi-variogram functions could change not only with distance but with direction as well. In the literature, estimating spatial anisotropy to model natural phenomena is a common topic of discussion (e.g., Bernardi et al., 2018; Agou et al., 2019). A reliable determination of the existence and characteristics of anisotropy requires a large

dataset with sufficient variants in all possible directions. However, in many cases, including the present study, the available data is sparse or unevenly distributed, making reliable anisotropy estimation challenging. This work considers omnidirectional semi-variograms, which may be more reliable in limited or sparse data, simplifying the modeling process and reducing the risk of overfitting. Therefore, spatial homogeneity is assumed and anisotropy is neglected, as applied in the works of Goovaerts (2000) and Ly et al. (2011).

Since the FAO method is the most restrictive among the approaches tested, gap-filling is only performed on stations that met all three conditions specified by this technique (see **Subsection 1.1**). This ensures that each selected station has a corresponding twin station that also satisfies these criteria.

The performance of the proposed approaches is evaluated for each of the selected precipitation stations for a chosen year by:

1. removing the precipitation values of the selected year from the observations;
2. estimating the 365-day precipitation using the available data, i.e., the highest correlated station for the FAO method and data from the remaining stations for the kriging-based approaches;
3. computing the error statistics (ME , MAE , and $RMSE$ - For formulas, see **Appendix A**).

The mGstat tool (Hansen, 2022) implemented in MATLAB (The MathWorks Inc., 2023) is used for the application of geostatistical methods, with computation times for each precipitation station ranging from 2 minutes for OK to 2 hours for SCK. The FAO method is implemented using a MATLAB code, which took only a few seconds for each station's estimation. The accuracy of the three gap-filling methods (FAO, OK, and SCK) is evaluated through a cross-validation process and the comparison of the error statistics. In addition, the monthly cumulative values derived from the daily interpolations are examined to provide further insight into the performance of each method.

2.1. Case 1: Mantova province, Italy

The first case study focused on the province of Mantova in northern Italy, which covers 234144 ha (Fig. 1.1). Mantova borders Brescia to the north/north-west, Cremona to the west, the Emilia-Romagna region to the south, and the Veneto region to the east. The average maximum temperature is around 5.6°C in winter and 28.4°C in summer, while the average minimum temperature ranges from -0.1°C to 18.1°C. Rainfall is concentrated primarily in spring and autumn, with average annual totals ranging from 500 to 900 mm (<https://www.agenziapo.it/>).

Daily precipitation data are collected from 26 rain gauges for the period 1960-2020. Throughout the entire available time series, no station met the constraints imposed by the FAO. Therefore, a thirty-year reference period (1976-2005) is chosen to avoid lowering the constraints imposed and due to the significant data inconsistency at most stations. Moreover, the Casalmaggiore province of Cremona, covering 6.4 km² and located west of Viadana, is included to enhance the dataset. To further supplement the data, 33 stations from the Emilia-Romagna climate observatory (Eraclito - <https://dati.arpae.it/dataset/erg5-eraclito>) are used. The observatory has generated a daily climate dataset from 1960 to the present using spatial interpolation on a regular grid from the network of historical weather stations.

The region is predominantly flat, with elevations ranging from 11 m to 62 m above sea level (a.s.l.). An exception is the hilly zone in the north, formed by the morainic amphitheatre of Garda Lake, which slopes towards the Po Valley. Due to the area's flatness, cokriging is not considered for this application, as there would be insufficient correlation in the cross-variograms.

Only one station (Saino di Pegognaga) met the constraints imposed by the FAO and is therefore the only one included in the estimation process. Table 1.2 provides a summary of the monitoring network's characteristics, with the selected station for gap-filling highlighted in bold. Figure 1.1 illustrates the spatial distribution of the rain gauges in the study area.

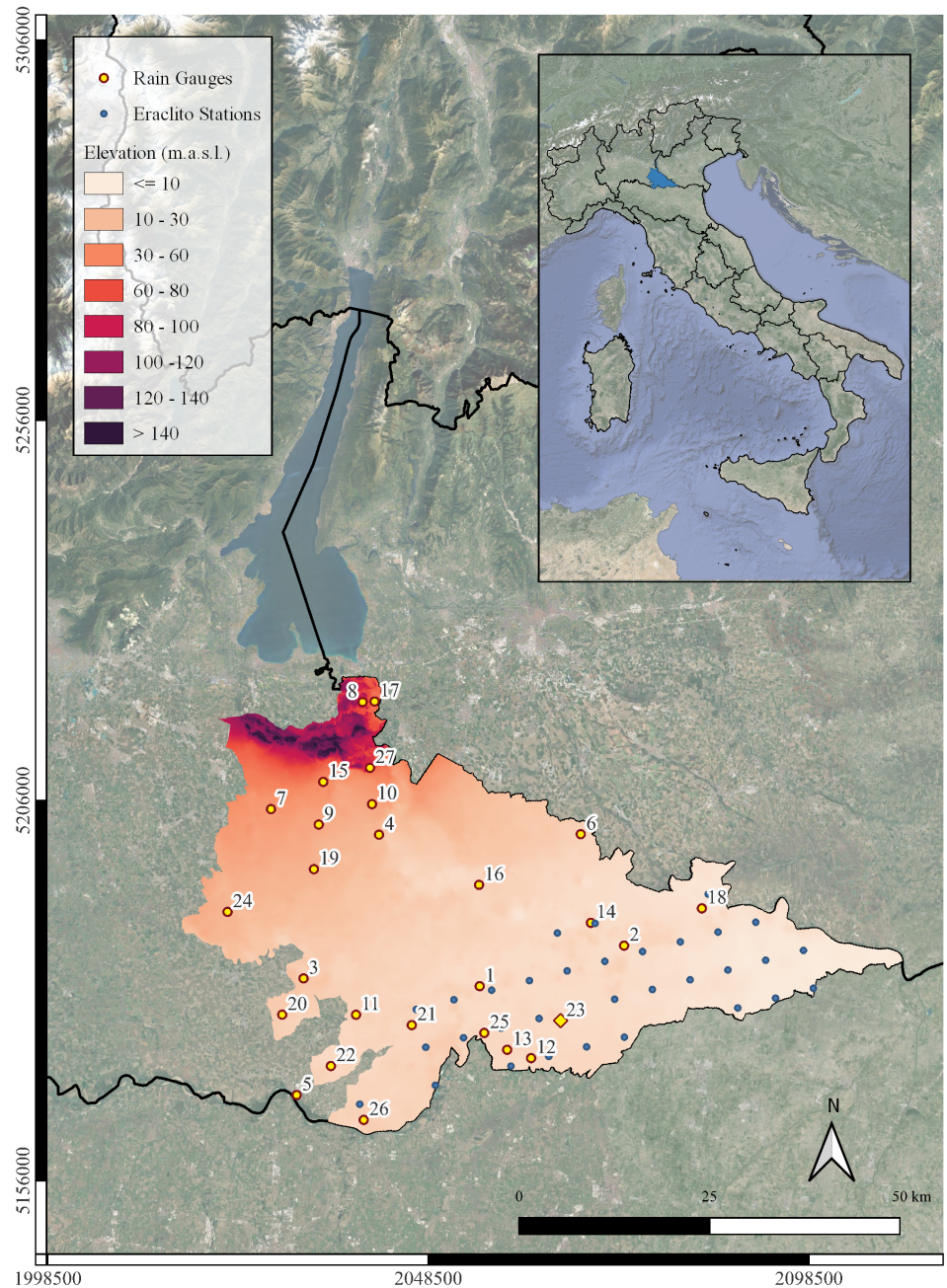


Figure 1.1. Study area (UTM zone 32N-EPG: 32632). Yellow dots represent the locations of available precipitation stations, the diamond marks the station meeting FAO criteria, and blue dots indicate ERACLITO dataset stations. Elevation is shown using a color scale DEM.

Table 1.2. Station Id, Station Name, Data (%) and Elevation (m.a.s.l.) in the reference period (1976-2005).

Station Id	Station Name	Data	Elev	Station Id	Station Name	Data	Elev
1	Borgoforte	19	18	15	Guidizzolo	28	61
2	Botte Sotto Secchia	60	16	16	Mantova	97	23
3	Bozzolo	83	30	17	Monzambano	87	88
4	Calliera di Goito	32	32	18	Ostiglia	79	13
5	Casalmaggiore	97	25	19	Piubega	60	40
6	Casteldario	86	24	20	Rivarolo Mantovano	39	26
7	Castelgoffredo	58	53	21	S-Matteo Delle Chiaviche	77	23
8	Castiglione delle Stiviere	50	109	22	Sabbioneta	9	18
9	Ceresara	31	44	23	Saino di Pegognaga	71	22
10	Cerlongo di Goito	16	45	24	Sorbara d'Asola	22	41
11	Gazzuolo	88	25	25	Suzzara	37	21
12	Gonzaga	6	22	26	Viadana	22	26
13	Gonzaga Palidano	0	20	27	Volta Mantovana	55	91
14	Governolo	66	17				

Semi-variograms computation

Monthly semi-variograms are generated by fitting the average daily precipitation data from the 27 rainfall stations to each month's observations across the entire study period. Isotropic semi-variograms are considered under the assumption of uniform spatial correlation in all directions. Based on this hypothesis, the semi-variograms are constructed following the procedure detailed in **Section 2**, using exponential and Gaussian models, both combined with a nugget effect (see Table 1.1). The resulting monthly semi-variograms are presented in Fig. 1.2. Although the fits of the two models are comparable, a combination of exponential semi-variograms and the nugget effect is selected. This decision is based on the fact that the Gaussian model implies strong continuity in the data, suggesting a high degree of spatial correlation even over short distances. However, such continuity is not observed because of the scarcity of data. The parameters for the selected semi-variogram models are presented in Table 1.3.

Table 1.3. Estimated parameters for the monthly semi-variograms of average daily precipitation, modeled using an exponential model.

Month	l	θ	C_0
Jan	70329	0.102	0.011
Feb	146205	0.155	0.022
Mar	77383	0.079	0.014
Apr	96700	0.050	0.026
May	428779	0.271	0.023
Jun	443216	0.473	0.019
Jul	499925	0.424	0.031
Aug	217513	0.099	0.028
Sep	399753	0.070	0.012
Oct	427527	0.296	0.012
Nov	146351	0.078	0.020
Dec	10007	0.023	0.008

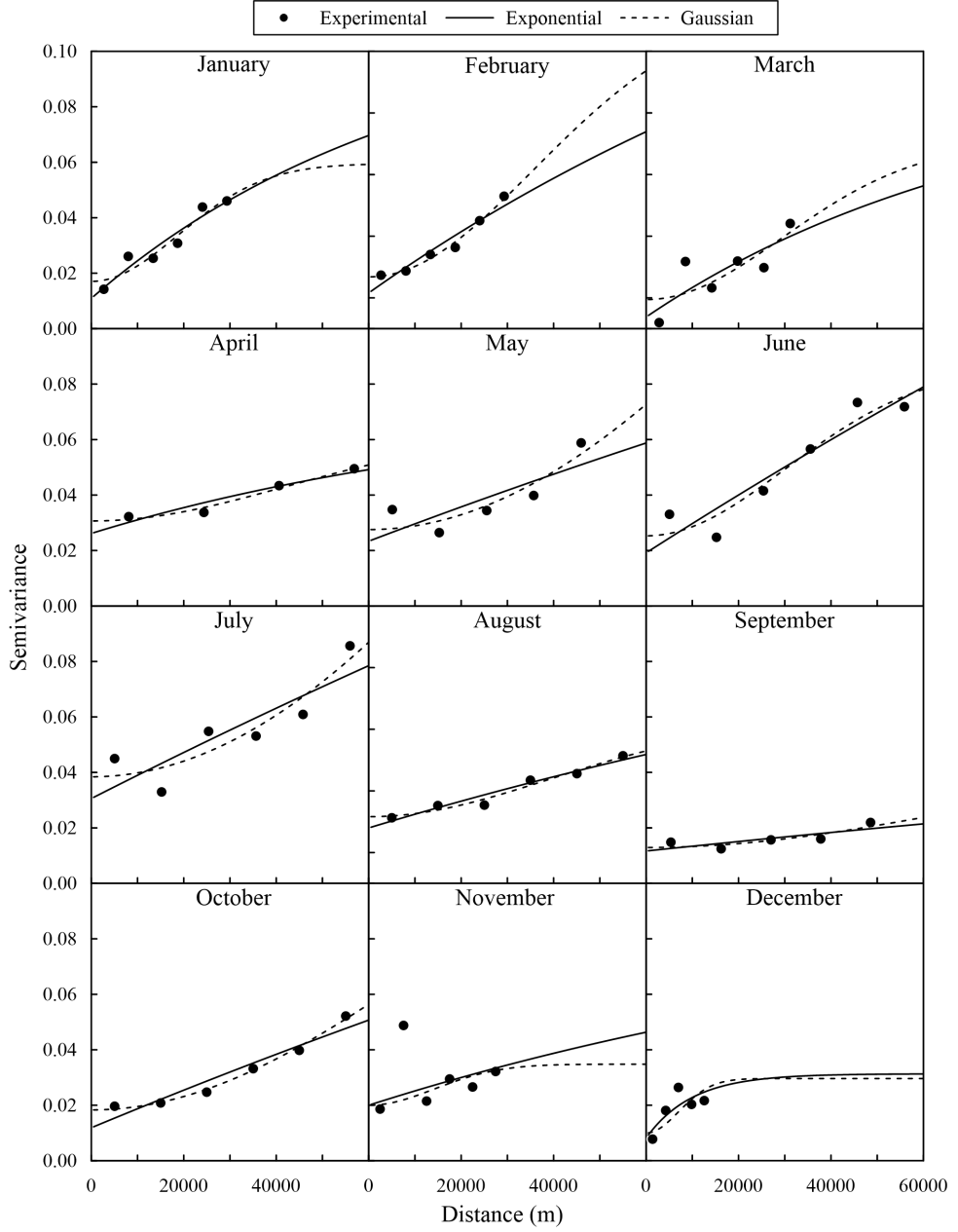


Figure 1.2. Experimental and fitted semi-variogram models based on daily average rainfall data from the 27 monitoring stations.

Results

The year 1977 is selected for testing the methods, as most stations have at least 70% of data. Estimation using kriging-based approaches involved a total of 60 rainfall stations.

The performance of the FAO and OK methods in estimating daily and monthly precipitation at the Saino di Pegognaga rainfall station is evaluated using ME , MAE , and $RMSE$. The comparison of the resulting cumulative monthly rainfall estimate is shown in Fig. 1.3.

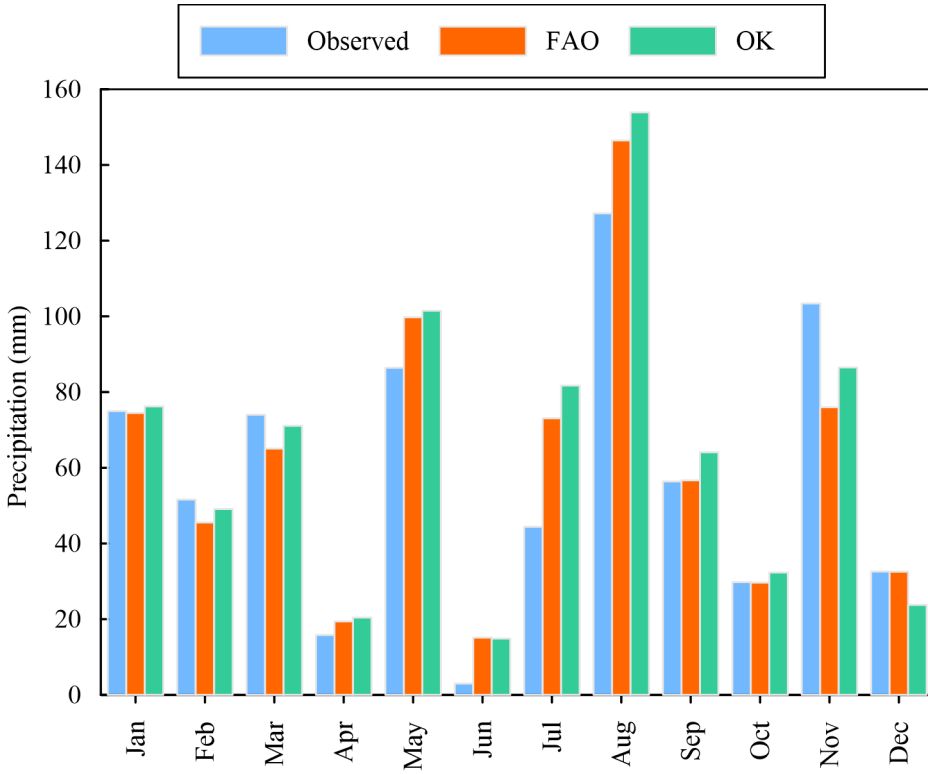


Figure 1.3. Cumulative monthly rainfall for 1977 at the Saino di Pegognaga station: observed values and estimates based on FAO and OK methods.

For daily precipitation estimates, the FAO method yielded ME of 0.094 mm/day, MAE of 1.001 mm/day, and $RMSE$ of 3.170 mm/day. In comparison, OK produced

ME of 0.208 mm/day, *MAE* of 1.094 mm/day, and *RMSE* of 3.459 mm/day. The slightly higher daily performance metrics of the OK method suggest a marginally greater propensity to overestimate or underestimate daily precipitation values. Instead, for monthly precipitation, the FAO method achieved *ME* of 2.857 mm/month, *MAE* of 10.041 mm/month, and *RMSE* of 14.150 mm/month. The OK method, on the other hand, resulted in *ME* of 6.324 mm/month, *MAE* of 11.510 mm/month, and *RMSE* of 15.648 mm/month. Similarly to the daily precipitation results, the FAO method showed a lower systematic error. Overall, the FAO method slightly outperformed the OK method on both daily and monthly time scales, as evidenced by lower *ME*, *MAE*, and *RMSE* values. Although the differences between the two methods are not substantial, the FAO approach appears to be more reliable for the estimation of precipitation in this specific case. This is likely due to the high correlation coefficient of the twin station (Suzzara station - $r^2 = 0.79$) in the FAO method. Additionally, the limited availability of reliable data causes challenges in determining suitable semi-variogram models in OK.

The coefficient of determination, (R^2) is used as the primary metric to evaluate the goodness of fit for both methods. For the estimation of daily precipitation, the FAO method achieved a R^2 value of 0.772, while the OK method achieved a R^2 value of 0.760. These values indicate that both FAO and OK provide good estimates of daily precipitation. For monthly precipitation estimation, both methods showed even stronger predictive capabilities. The FAO method achieved a R^2 value of 0.852, and the OK method achieved a slightly higher R^2 of 0.861. Both methods demonstrate a good ability to predict daily precipitation, with R^2 values close to 0.86. However, the difference in performance remains minor, with both methods demonstrating robust reliability for monthly predictions. Daily cross-validation (Fig. 1.4) is used to assess the predictive performance of both methods for rainfall estimation at the Saino di Pegognaga rainfall station. Cross-validation reveals that both FAO and OK methods perform well in estimating daily precipitation, as reflected by their high R^2 values.

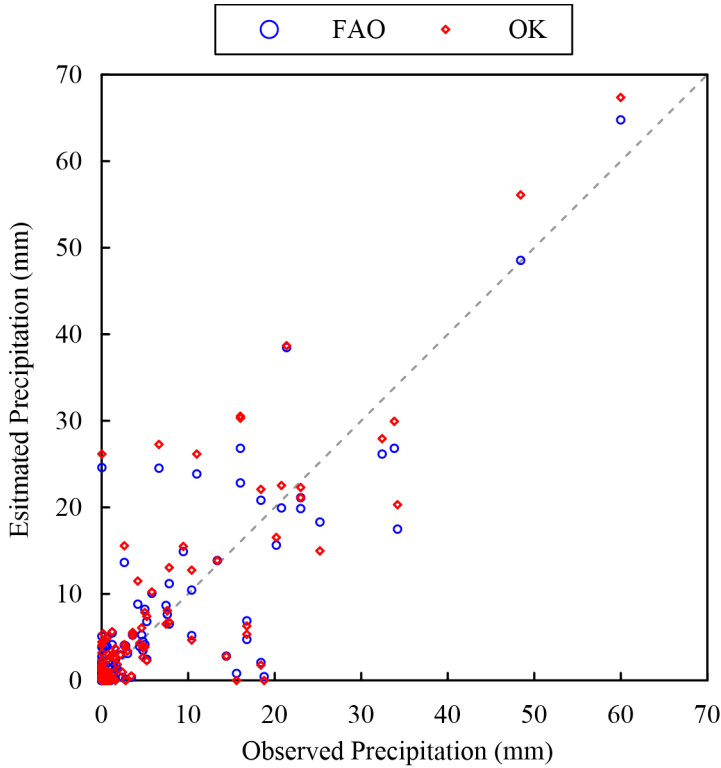


Figure 1.4. Daily cross-validation results for 1977 at the Saino di Pegognaga station.

2.2. Case 2: Guadiana River basin, Portugal

The second case study area, which forms part of the InTheMED PRIMA project, focuses on the Portuguese portion of the Guadiana Hydrographic Basin. This basin, shared by Spain and Portugal, spans a total area of 6713300 ha, with 5552800 ha in Spain and 1161100 ha in Portugal. The Guadiana River, stretching 820 km, traverses both countries. According to Todaro et al. (2022a), the Portuguese region recorded an average annual precipitation of 545 mm and an average temperature of 16.1°C between 1986 and 2005. These values align with the findings of Palop-Donat et al. (2020), who reported an average precipitation of 566 mm and an average temperature of 16.3°C for the period from 1980/81 to 2011/12.

Daily precipitation data are collected from 96 rain gauges between 1970 and

2021. Only stations with at least 70% data coverage for the 30-year reference period (1976–2005) are considered, resulting in the selection of 60 rain gauges. Elevation data for these stations, obtained from a Digital Elevation Model (DEM) (<https://data.europa.eu/data>), range from 3 to 499 m above sea level (a.s.l.).

A total of 12 stations fulfilled the constraints established by the FAO (**Subsection 1.1**). Table 1.4 summarizes the key characteristics of the monitoring network, with the selected stations for gap-filling highlighted in bold. Figure 1.5 illustrates the spatial distribution of these rain gauges across the study area.

Table 1.5 presents the squared Pearson correlation coefficients between the selected stations and their best-correlated counterparts.

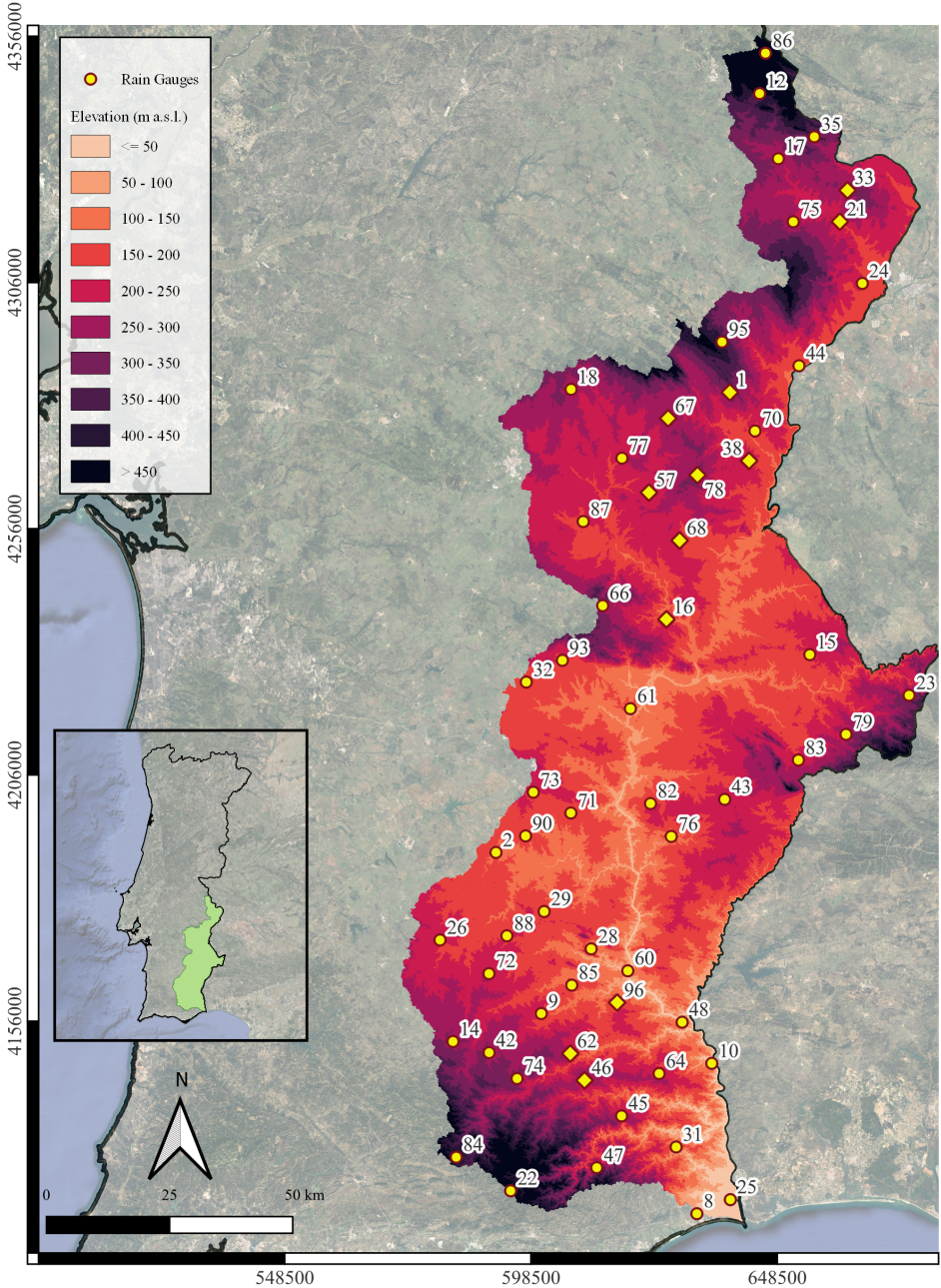


Figure 1.5. Study area (UTM zone 29N-EPSSG: 32629). Yellow dots represent the locations of available precipitation stations and the diamond marks the stations meeting FAO criteria. Elevation is shown using a colour scale DEM.

Table 1.4. Station Id, Data (%), and Elevation (m.a.s.l.) in the reference period (1976-2005).

Station Id	Data	Elev	Station Id	Data	Elev	Station Id	Data	Elev
1	97	302	32	99	180	71	100	185
2	88	133	33	81	269	72	87	239
8	98	34	35	82	341	73	86	195
9	98	173	38	81	212	74	83	328
10	100	32	42	85	295	75	81	254
12	83	449	43	98	223	76	84	145
14	99	286	44	96	198	77	90	208
15	100	204	45	85	253	78	97	308
16	84	195	46	100	294	79	100	271
17	94	257	47	99	392	82	100	209
18	97	270	48	85	143	83	83	223
21	85	217	57	75	236	84	100	450
22	97	467	60	83	29	85	74	155
23	91	265	61	98	128	86	82	499
24	82	201	62	83	273	87	99	190
25	83	3	64	100	250	88	100	182
26	100	217	66	97	302	90	100	172
28	87	142	67	78	300	93	100	203
29	71	163	68	100	218	95	99	417
31	73	185	70	81	247	96	86	158

Table 1.5. Squared Pearson correlation coefficient (r^2) between each selected monitoring station and its best-correlated stations.

Station to be estimated	r^2	Best-correlated station
POR001	0.753	POR067
POR016	0.746	POR068
POR021	0.766	POR033
POR033	0.766	POR021
POR038	0.766	POR078
POR046	0.739	POR096
POR057	0.716	POR078
POR062	0.736	POR046
POR067	0.753	POR001
POR068	0.746	POR016
POR078	0.766	POR038
POR096	0.739	POR046

Semi-variograms computation

The monthly averages of daily rainfall obtained from the monitoring stations throughout the reference period constitute the major variable, while elevation data serve as the secondary one. These elements are used to compute semi-variograms and cross-variograms.

Semi-variograms on the first variable are constructed for each month by fitting the average daily precipitation data from the available rainfall stations corresponding to that specific month across the entire study period. This process followed the procedure outlined in **Section 2** using the models in Table 1.1. The resulting monthly semi-variograms are shown in Fig. 1.6. The chosen models fit the data satisfactorily, and it is decided to estimate all 12 monthly semi-variograms using parameters (Table 1.6) using a combined exponential and nugget approach.

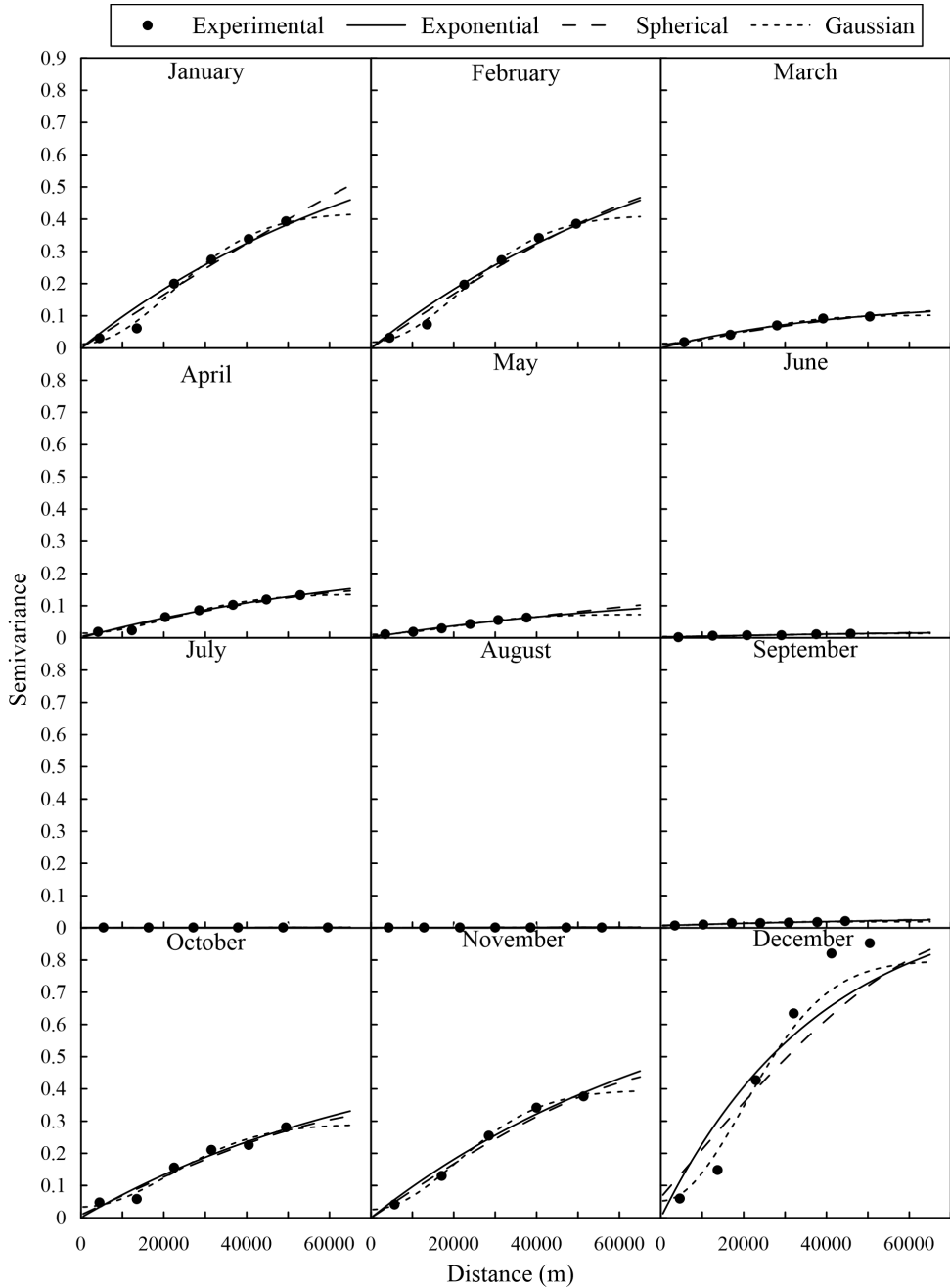


Figure 1.6. Experimental and fitted semi-variogram models based on daily average rainfall data from 60 monitoring stations.

Table 1.6. Estimated parameters for the monthly semi-variograms on the first variable (i.e., average daily precipitation) with an exponential model.

Month	l	θ	C_0
Jan	79999	0.826	1.00E-04
Feb	79999	0.824	1.00E-04
Mar	48579	0.156	1.86E-04
Apr	79911	0.277	1.00E-04
May	79869	0.161	1.84E-03
Jun	77803	0.024	1.98E-03
Jul	79337	0.002	5.28E-04
Aug	43149	0.002	6.20E-04
Sep	69940	0.029	6.73E-03
Oct	79985	0.599	1.00E-04
Nov	79974	0.819	1.00E-04
Dec	80000	1.848	1.00E-04

Figure 1.7 shows the average daily rainfall computed over the reference period for each month. It should be noted that the semi-variograms reflect the precipitation regime of the study area depicted in Fig. 1.5. Autumn and winter months (from October to February) are the wettest, reflected in the high variability of the semi-variograms (i.e., higher spatial variability). In spring months (i.e., March, April, and May), precipitation is lower than in winter, implying less variability in the semi-variograms. The summer months (from June to September) are the driest, with semi-variance tending towards zero.

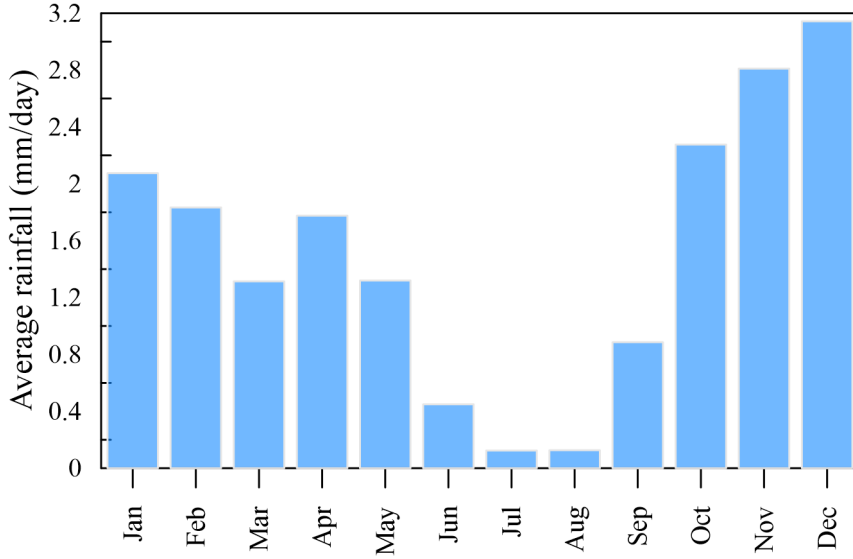


Figure 1.7. Bar chart displaying the mean daily precipitation for each month from 1976 to 2005 in the study area.

The experimental semi-variogram for the secondary variable is computed using approximately 180000 elevation data points extracted from a regular grid from the DEM. This approach is taken to avoid the isotopic case (i.e., using elevation data only at the locations of the precipitation stations) and to represent a greater spatial variability. The same assumptions applied to the precipitation semi-variograms are used.

Table 1.7. Estimated parameters for the semi-variogram on the second variable (i.e., elevation) with an exponential model.

Variable	l	θ	C_0
Elev	26365	7483.703	116.366

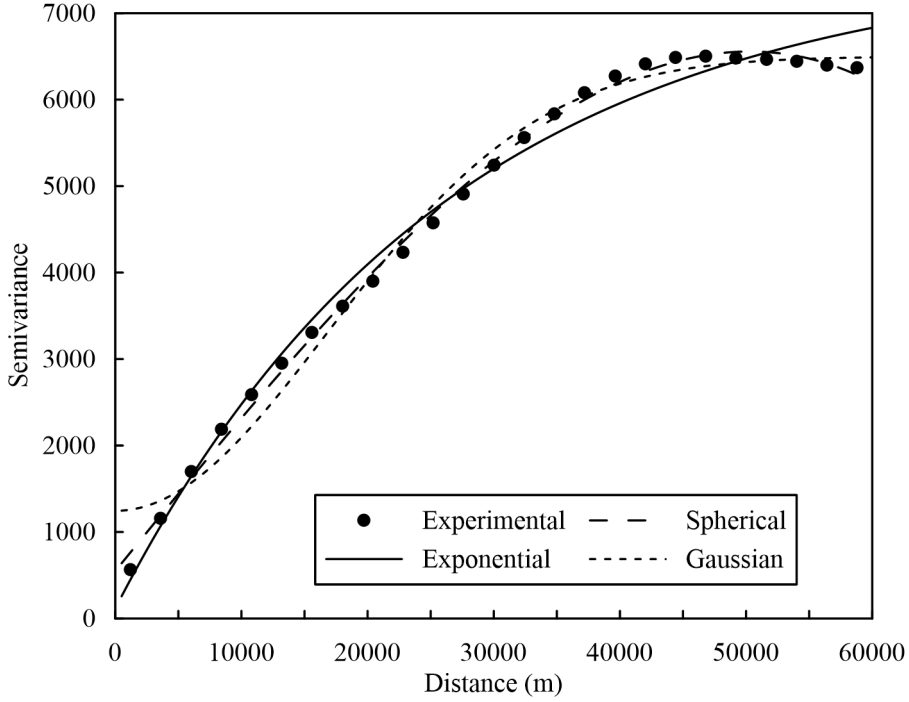


Figure 1.8. Experimental and fitted semi-variogram models based on the elevation DEM.

The cokriging analysis requires the estimation of the semi-variogram and cross-variogram models for precipitation and elevation simultaneously. Table 1.8 shows the correlation coefficients between the two variables. According to Asli and Marcotte (1995), using secondary information is justified only for correlations greater than 0.4. Given that the correlations between rainfall and elevation range from 0.4 to 0.77, it is considered worthwhile to carry out precipitation estimation using SCK. As with precipitation, isotropic experimental cross-variograms are computed. Figure 1.9 shows the cross-variograms between precipitation and elevation for each month. To meet the PDC condition, the same precipitation model is chosen for all the cross-variograms (Table 1.9). The PDC curve based on Eq.1.7 is also shown in Fig. 1.9 to graphically examine the criteria.

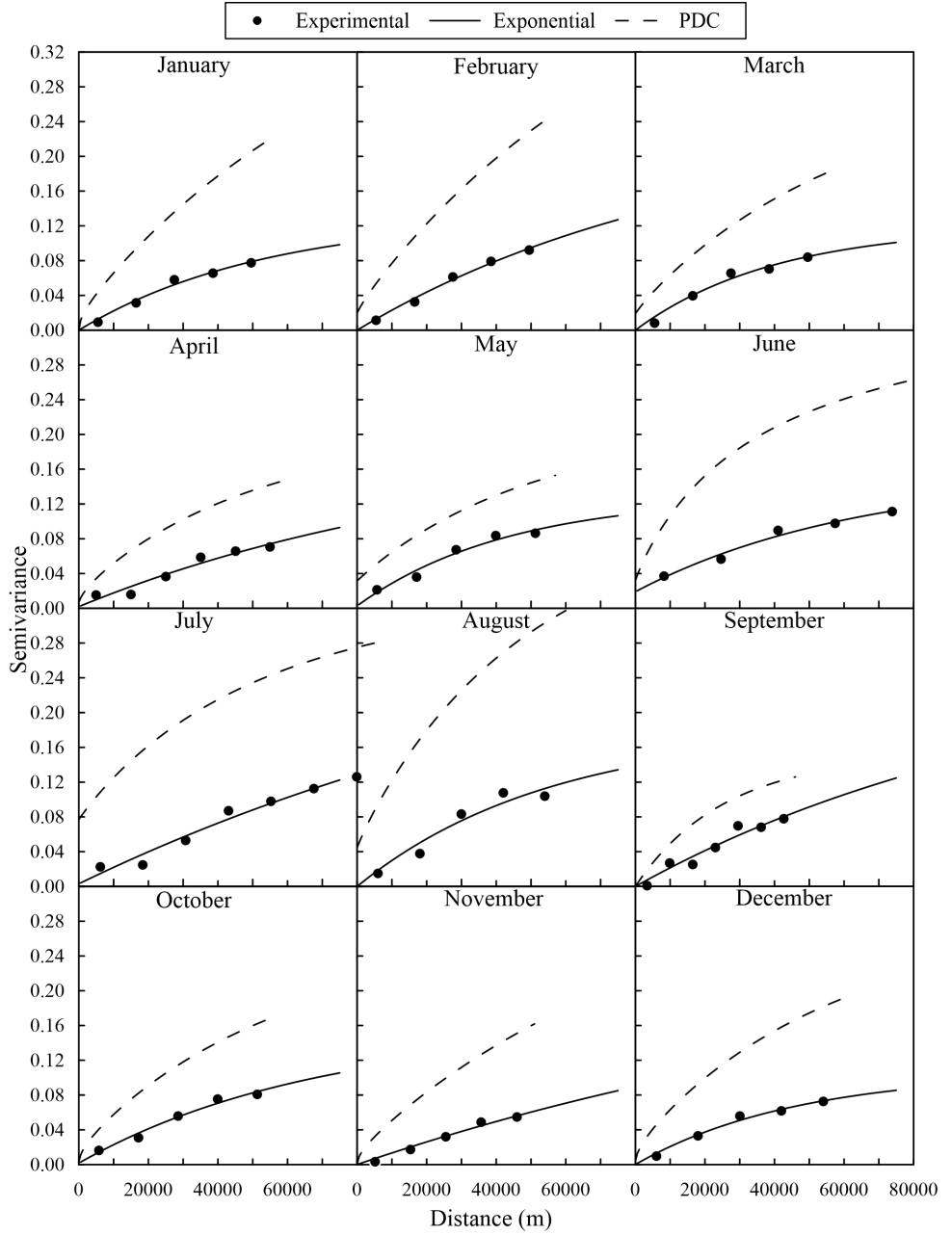


Figure 1.9. Experimental cross-variograms (black dots) with the fitted cross-variogram models (solid line) and positive-definite condition curve (dashed line) based on the rainfall (i.e., the primary variable) and elevation data (i.e., the secondary variable).

Table 1.8. Average correlation coefficients between the primary variable (precipitation) and secondary variable (elevation) for each month.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
r	0.75	0.75	0.75	0.71	0.76	0.45	0.40	0.53	0.73	0.77	0.62	0.63

Table 1.9. Estimated parameters for the monthly cross-variograms of average daily precipitation and elevation, modeled using an exponential function.

Month	l	θ	C_0
Jan	54068	0.131	1.00E-04
Feb	71034	0.184	1.72E-04
Mar	39763	0.118	1.02E-04
Apr	70344	0.133	3.19E-04
May	42978	0.125	2.70E-03
Jun	63256	0.135	1.88E-02
Jul	79938	0.195	1.00E-04
Aug	55644	0.181	1.00E-04
Sep	79833	0.191	1.00E-04
Oct	67541	0.155	1.69E-03
Nov	79836	0.124	1.00E-04
Dec	45354	0.105	1.00E-04

Results

The year 1985 is chosen for testing the method since none of the 12 selected stations has gaps in their period of the time series. The estimation using kriging-based approaches involved 60 rainfall stations and 110 elevation values as the primary and secondary variables, respectively. The number of elevation data points is reduced to limit computational time.

The metrics of the gap-filling methods used to estimate the daily and monthly cumulative precipitation in 1985 at the chosen stations are summarized in Tables 1.10–1.11, which are depicted in graphical form in Fig(s).1.10–1.12.

The kriging-based methods (OK and SCK) generally outperform the determin-

Table 1.10. Evaluation metrics for the proposed methods in estimating daily rainfall (mm/d) in 1985 across the 12 monitoring stations. Values in bold indicate minimum MAE and RMSE.

Station	<i>ME</i> (mm/d)			<i>MAE</i> (mm/d)			<i>RMSE</i> (mm/d)		
	FAO	OK	SCK	FAO	OK	SCK	FAO	OK	SCK
POR001	-0.336	-0.246	-0.252	0.824	0.544	0.542	2.369	1.598	1.605
POR016	-0.091	0.006	-0.014	0.685	0.495	0.484	2.225	1.655	1.611
POR021	-0.084	0.111	0.111	0.622	0.528	0.525	2.088	1.789	1.776
POR033	-0.199	0.012	-0.007	0.706	0.704	0.707	2.261	2.170	2.163
POR038	0.021	0.203	0.193	0.723	0.593	0.586	2.410	1.869	1.842
POR046	-0.275	0.061	0.063	0.597	0.726	0.724	1.919	2.095	2.094
POR057	-0.067	0.081	0.077	0.591	0.423	0.425	2.148	1.186	1.199
POR062	-0.154	-0.049	-0.045	0.689	0.736	0.727	2.306	2.187	2.174
POR067	-0.030	0.147	0.127	0.615	0.553	0.544	1.844	1.633	1.583
POR068	-0.230	0.041	0.027	0.727	0.503	0.497	2.315	1.529	1.528
POR078	-0.356	-0.243	-0.196	0.854	0.697	0.719	2.855	2.303	2.026
POR096	0.042	0.098	0.104	0.450	0.466	0.470	1.499	1.488	1.488

istic FAO method in the estimation of daily precipitation. The *ME* values (Table 6 and Fig. 1.10) indicate whether the estimates under - or overestimate the actual values. For daily *MAE* (Table 1.10 and Fig. 1.11), SCK achieves the best performance at six out of twelve stations, while OK leads in three cases, showing results similar to SCK. The FAO method exhibits marginally superior performance at three stations (POR046, POR062, and POR096) due to the robust correlation across the stations (Table 1.8). The *RMSE* analysis (Table 1.10 and Fig. 1.12) shows that the kriging-based approaches provide better estimates than the FAO, except for POR046. Among the evaluated methods, POR057 demonstrates the best daily statistics with *MAE* of 0.423 mm/d and *RMSE* of 1.186 mm/d for the OK method, whereas POR062 shows the poorest performance. Additionally, POR046 yields better results for the FAO method (*MAE* = 0.597 mm/d and *RMSE* = 1.919 mm/d) compared to the kriging-based approaches (*MAE* = 0.726 mm/d and *RMSE* = 2.095 mm/d for OK; *MAE* = 0.724 mm/d and *RMSE* = 2.094 mm/d for SCK). However, this station presents *ME* = -0.275 mm/d, suggesting an average underestimation of

Table 1.11. Evaluation metrics for the proposed methods in estimating monthly rainfall (mm/month) in 1985 across the 12 monitoring stations. Values in bold indicate minimum MAE and RMSE.

Station	ME (mm/month)			MAE (mm/month)			RMSE (mm/month)		
	FAO	OK	SCK	FAO	OK	SCK	FAO	OK	SCK
POR001	-10.234	-7.474	-7.661	15.282	9.562	9.740	19.694	12.165	12.464
POR016	-2.753	0.186	-0.419	6.624	6.429	6.166	10.141	8.904	8.465
POR021	-2.566	3.384	3.705	4.846	4.908	4.895	7.071	7.217	7.231
POR033	-6.066	0.364	-0.088	6.712	3.725	3.715	9.670	5.368	5.271
POR038	0.631	6.187	5.867	10.201	7.848	7.586	13.431	10.890	10.517
POR046	-8.360	1.869	1.913	9.855	5.435	5.437	16.167	8.466	8.512
POR057	-2.049	2.475	2.338	8.941	5.045	5.144	12.434	6.677	6.748
POR062	-4.697	-1.502	-1.362	6.974	4.727	4.573	10.307	7.339	7.288
POR067	-0.923	4.474	3.878	8.398	7.182	6.836	10.546	9.859	9.321
POR068	-7.007	1.244	0.819	7.906	5.442	5.264	17.716	8.067	8.303
POR078	-10.831	-7.390	-5.966	12.492	9.648	7.598	18.517	14.783	12.500
POR096	1.281	2.988	3.186	4.380	6.488	6.488	6.837	9.106	9.228

precipitation.

The actual cumulative monthly rainfall is compared with the one obtained from daily estimations (Fig. 1.13), and subsequently error statistics are calculated (Table 1.11 and Fig(s).1.10-1.12 bottom). This approach enables the evaluation of the accuracy of the daily estimates over an extended period and helps to identify any systematic errors present in the estimates. According to Table 1.11 and Fig. 1.11 bottom, kriging-based methods are comparable, although SCK showed better performance in terms of *MAE*, outperforming the other methods at seven out of twelve stations, while OK performed better at three stations. The FAO approach yielded slightly better results in two precipitation stations (POR021 and POR096). In terms of *RMSE* (Table 1.11 and Fig. 1.12 bottom), SCK achieved better results in six out of twelve stations, with OK performing better in four stations and FAO yielding similar results to *MAE* in the same precipitation stations.

The error metrics for daily and cumulative monthly precipitation differ because they consider different aspects. Daily rainfall provides information on the amount of

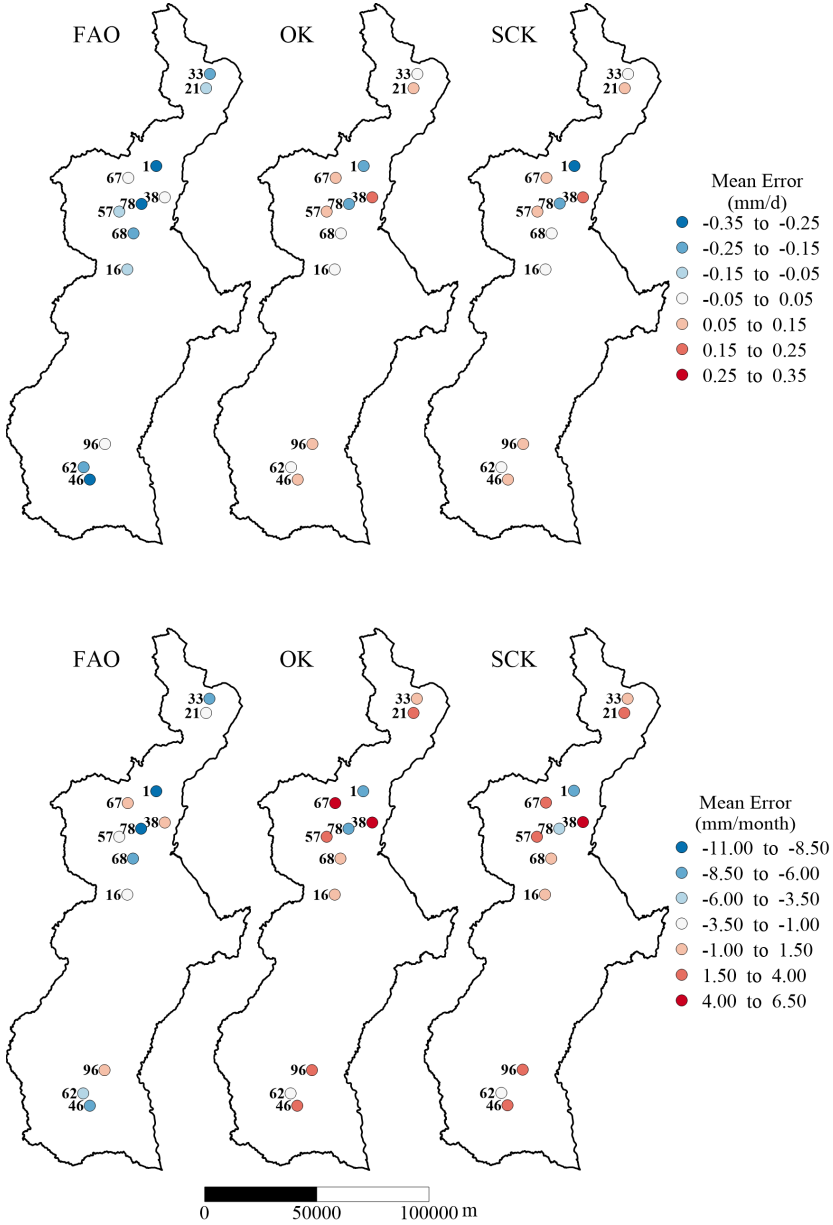


Figure 1.10. Mean error computed using the applied approaches (FAO, OK, SCK) at precipitation stations, presented for daily (top) and monthly (bottom) estimation scales.

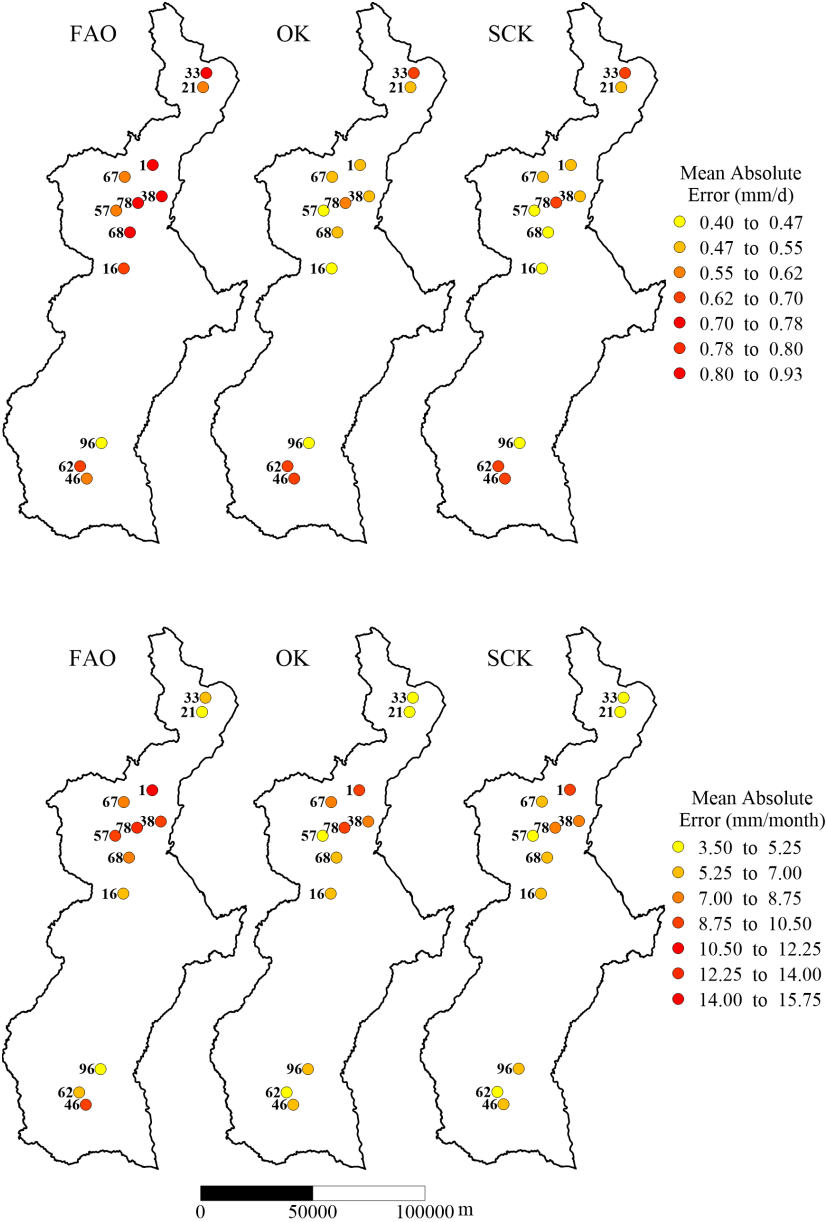


Figure 1.11. Mean absolute error computed using the applied approaches (FAO, OK, SCK) at precipitation stations, presented for daily (top) and monthly (bottom) estimation scales.

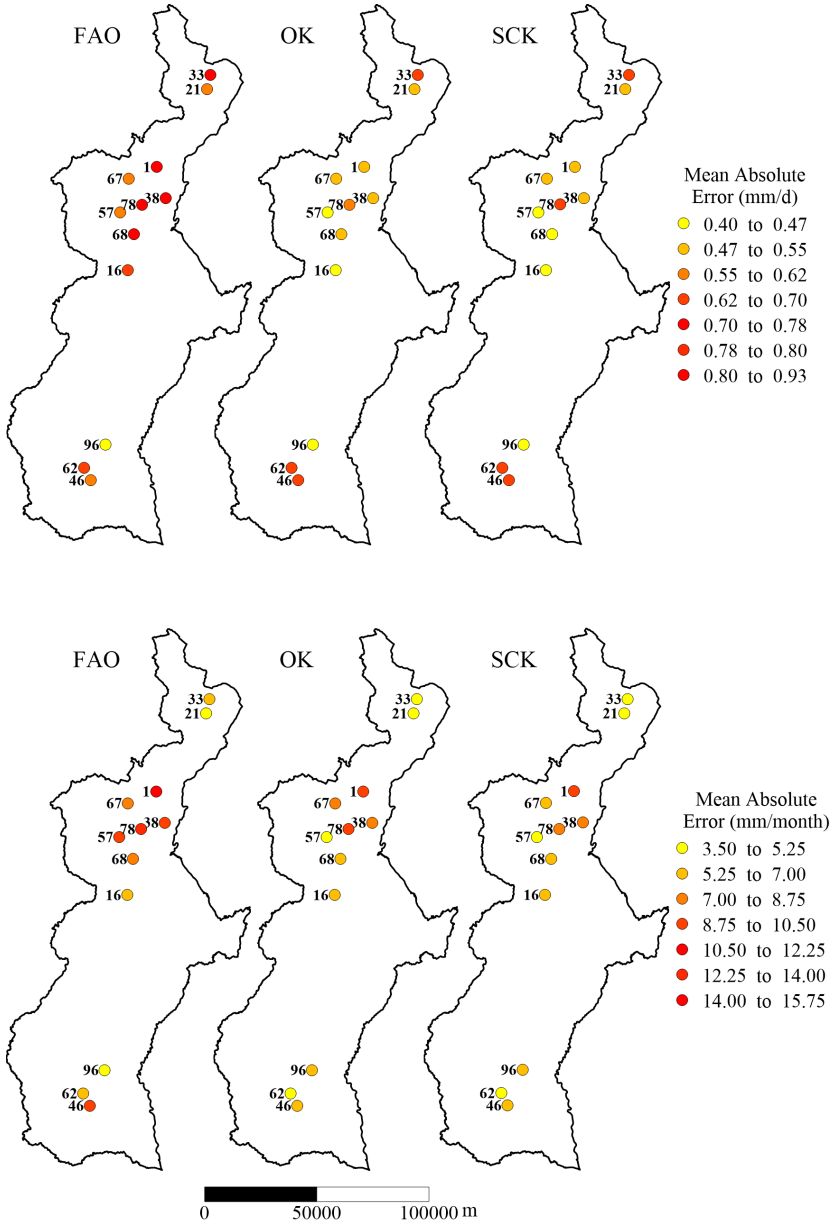


Figure 1.12. Root mean square error computed using the applied approaches (FAO, OK, SCK) at precipitation stations, presented for daily (top) and monthly (bottom) estimation scales.

precipitation on a specific day, while cumulative monthly precipitation data provide information on the total amount of precipitation in a given month. The differences in error metrics can be attributed to the high variability in daily precipitation, whereas monthly precipitation is generally smoother and more consistent.

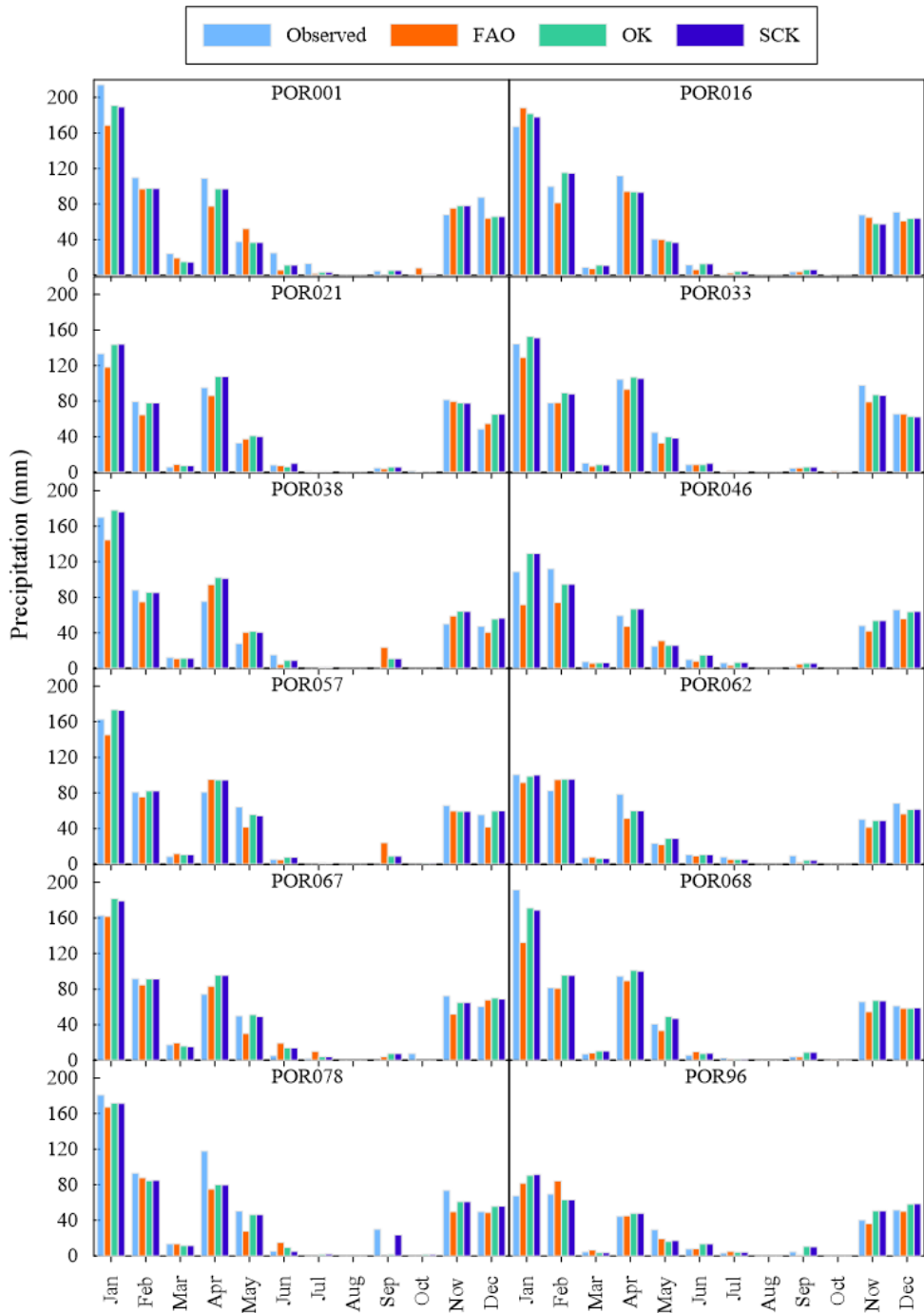


Figure 1.13. Cumulative monthly rainfall for the year 1985, observed and estimated using FAO, OK, and SCK methods across the 12 precipitation stations.

Considering all performance metrics (Tables 1.10–1.11 and Fig.(s) 1.10-1.12), the results of OK and SCK are comparable. Differences between estimates increase as the amount of secondary data exceeds that of primary data. A downside of the SCK method is that a large elevation dataset significantly prolongs the estimation process. Furthermore, there is no substantial improvement in using SCK despite the good correlations between the two variables (Table 1.8).

Cross-validation values of observed and estimated daily rainfall for 1985 at the 12 selected stations are plotted in Fig. 1.14. The 45° line indicates a perfect fit between observed and calculated values.

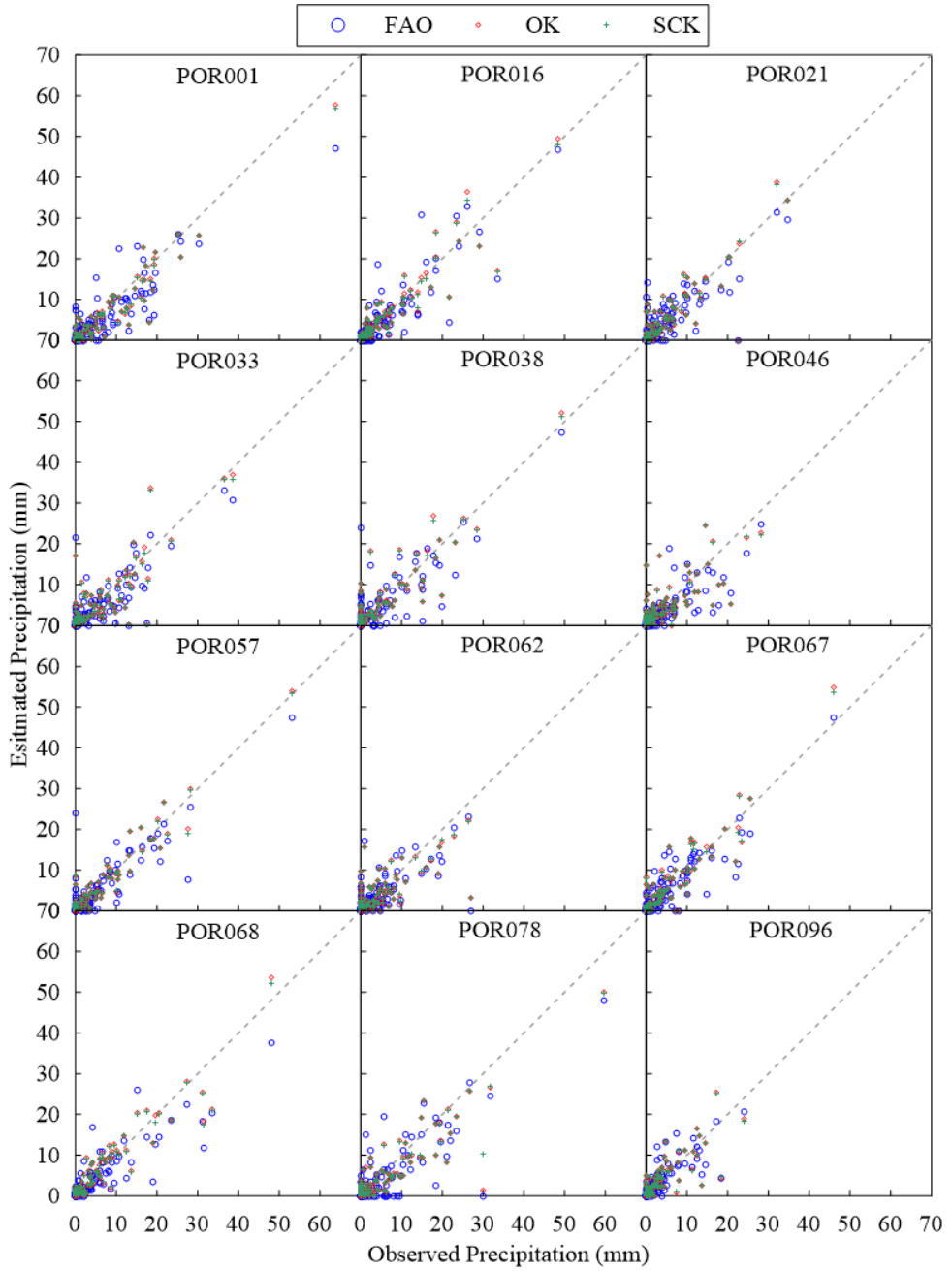


Figure 1.14. Daily cross-validation results for 1985 at the selected stations.

Regarding daily estimation across methods, kriging-based estimates are weighted on spatially distributed observations simultaneously (Eq.1.8), effectively identifying dry days but potentially smoothing localized intense rainfall events. The FAO method, on the other hand, relies heavily on the best-correlated station, replicating all weather events (both dry and intense) along the regression line (Eq.1.3).

One of the significant advantages of the kriging-based approaches over the FAO method is their ability to evaluate the uncertainty associated with the estimations. Figure 1.15 illustrates an example of error variance estimation for January, computed using OK (a) and SCK (b), along with the difference between them (c). The OK variance is calculated using information from 60 precipitation stations. In contrast, the SCK method utilized not only the precipitation data but also 110 elevation data points. These included 60 elevation data points collected at monitoring stations and 50 additional elevation data points retrieved from the DEM on a regular grid with a side length of 15 km.

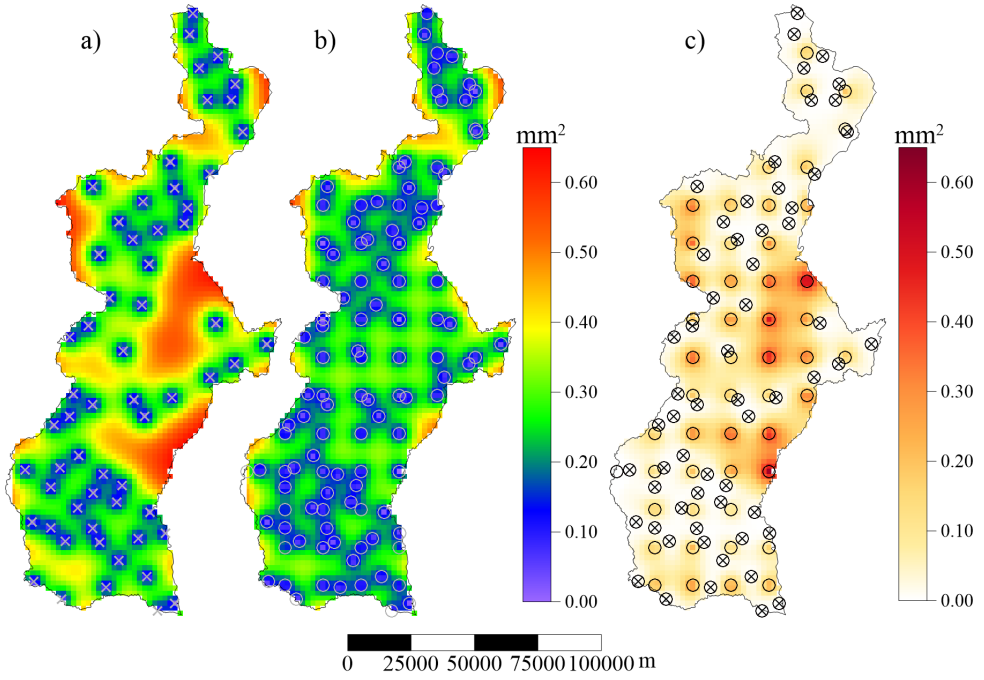


Figure 1.15. Estimation of error variance for daily precipitation in January using OK (a), SCK (b), and the variance difference between OK and SCK (c). Crosses indicate the locations of the rainfall data, while circles represent the locations of elevation data (the secondary variable in cokriging).

The error variance is zero at the data positions and increases away from them. In the OK variance map (Fig. 1.15-a), hotspots indicate areas where the estimation becomes more uncertain, resulting in higher error variance due to the lack of information. The use of auxiliary information with the SCK method can help reduce the estimation variance (Fig. 1.15-b). By incorporating additional information from the secondary variable, SCK provides a more accurate estimate of the primary variable, especially in areas lacking primary data points. Figure 1.15-c shows the difference between the error variances of OK and SCK. The areas with the greatest differences are those in which the SCK method significantly improves over OK. These areas likely have few primary data points, and the SCK method incorporates additional information from secondary variables.

The findings presented in Fig. 1.15 indicate that SCK can substantially enhance spatial interpolation accuracy, especially in regions with sparse primary data, where auxiliary information can improve estimates. Theoretically, the SCK estimator outperforms, as its error variance is consistently less than or equal to that of kriging. (Goovaerts, 1997).

3. Concluding remarks

An alternative gap-filling approach using kriging-based methods (Ordinary kriging - OK, and Simple cokriging - SCK) is presented and compared with a deterministic technique (FAO method). The proposed approach involves fitting kriging and cokriging semi-variograms to the monthly averages of the daily rainfall data over the entire time series for each station, resulting in the generation of 12 semi-variograms, one for each month. The objective is to estimate missing daily precipitation data at certain monitoring stations in two case study areas: the province of Mantova in Italy and the Portuguese portion of the Guadiana River basin.

According to the results obtained in both applications, despite its strong dependence on the imposed constraints, the FAO method is the least computationally demanding, as it is based on a simple linear regression between two correlated stations. Therefore, as long as these requirements are satisfied, it remains a viable option for gap-filling. However, unlike geostatistical approaches, it does not account for spatial correlation and does not provide a measure of the uncertainty associated with the results. On the other hand, geostatistical estimation requires a sufficient number of reliable observations to ensure accuracy and the computation of the semi-variograms that govern spatial variability.

In the first application, SCK was not utilized due to the flatness of the area. The distinctions between the FAO and OK approaches are minimal; nonetheless, the FAO approach has superior reliability in estimating precipitation at the designated station, evidenced by a robust correlation between the two stations. In

the second application, geostatistical techniques (OK and SCK) outperformed the deterministic FAO method. Cokriging is much more computationally demanding than kriging because it requires inferring and jointly modeling semi-variograms for each variable and cross semi-variograms, resulting in a large cokriging system to be solved. Although SCK is expected to provide better results than OK, it did not significantly improve the estimates in this study. Despite the monthly correlations between precipitation and elevations being greater than 0.4, SCK showed only slight improvements in estimation at both daily and monthly scales.

Generally, the use of monthly semi-variograms in the kriging methods yields satisfactory results. It significantly reduces computation time compared to the standard approach, which computes semi-variograms for each missing day. However, using this procedure may smooth local extreme rainfall events, as the monthly averaged semi-variograms may not accurately estimate their intensity and extent. In contrast, the impact of such extremes in the FAO approach depends on the correlation between the monitoring stations and the available data. In this case, an extreme event is estimated at a monitoring station only if a highly correlated station where it is detected exists.

2

Inversion Modeling Techniques: From Theory to Practice in Subsurface Applications

Over the last century, aquifers have faced increasingly severe environmental strains as a result of extensive industrial and agricultural development. Mitigating these pressures requires collaborative efforts, necessitating coordination with authorities and end-users to make informed decisions that prevent the depletion and contamination of aquifers. One of the most challenging problems in groundwater hydrology is identifying hydraulic parameters (e.g., transmissivity, hydraulic conductivity, specific storativity, etc.) and contaminant transport characteristics (e.g., identification of the source and associated release histories). This challenge demands a comprehensive understanding of the subsurface environment, whose complex spatial distribution is often difficult to characterize (Gómez-Hernández and Wen, 1994; Gómez-Hernández et al., 2003).

Recovering aquifer characteristics and groundwater contamination information

from direct and/or indirect measurements requires solving a complex inverse problem. Inverse modeling involves obtaining information about a model from observations of the modeled system (Carrera et al., 2005). This process includes two related ideas: model identification and parameter estimation. Model identification refers to techniques used to determine the characteristics of the model, such as the governing equations, boundary conditions, or patterns of heterogeneity. On the other hand, parameter estimation focuses on assigning specific values to the properties that define these characteristics.

Mathematically, the forward linear problem can be expressed as:

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \epsilon, \quad (2.1)$$

where $\mathbf{d}$ represents the observed data, $\mathbf{G}$ is the forward model, $\mathbf{m}$ represents the model parameters to be estimated, and ϵ is the observational error. In inverse modeling, the goal is to find $\mathbf{m}$ such that the misfit between the observed and predicted data is minimized, often under regularization constraints. Inverse modeling constitutes a crucial and indispensable component in hydrogeological research. Inverse problems are frequently ill-posed (Hadamard, 1902), indicating that a solution may be non-existent, non-unique, or may not vary continuously with the data. It is commonly managed by regularization and stochastic approaches that place further constraints or prior information on the solution. A comprehensive review of hydrogeology inverse methodologies is available in the works of McLaughlin and Townley (1996), Carrera et al. (2005), Hendricks Franssen et al. (2009), Zhou et al. (2014) and Gómez-Hernández and Xu (2022).

Several deterministic and stochastic methods have been proposed to address these challenges. Among stochastic inverse methods, the geostatistical approach (Kitanidis, 1995) offers an effective way to characterize spatial variability and infer properties at unsampled locations while quantifying the associated uncertainty (Michalak and Kitanidis, 2004; El Idrysy and De Smedt, 2007; Huysmans and Dassargues, 2009; Zhou et al., 2012; Butera et al., 2013; Cupola et al., 2015). Notably, the Empirical

Bayes solution and its implementation have shown significant promise within the geo-statistical framework (Zanini and Woodbury, 2016; Fernandez Visentini et al., 2020). Bayesian approaches provide a robust framework for handling uncertainty in inverse problems. By integrating prior knowledge and iteratively updating it based on observed data, these methods provide a posterior distribution that accounts for both prior beliefs and data-driven information. In addition, the ensemble Kalman filter (Evensen, 1994) and the ensemble smoother (Van Leeuwen and Evensen, 1996) have gained popularity in hydrogeology due to their adaptability and effectiveness (Chen and Zhang, 2006; Li et al., 2012; Crestani et al., 2012; Xu and Gómez-Hernández, 2016; Chen et al., 2018; Li et al., 2019). Specifically, Emerick and Reynolds (Emerick and Reynolds, 2012, 2013) introduced the ensemble smoother with multiple data assimilation (ES-MDA), which involves the iterative assimilation of the same data multiple times, enhancing the applicability and efficacy of the ensemble smoother (Todaro et al., 2019, 2021; D’Oria et al., 2022; Xu et al., 2021; Godoy et al., 2022; Chen et al., 2023).

Conventional survey methods often fail to capture the spatial distribution of sub-surfaces adequately. As a result, complementary observational techniques have been developed to address these survey-related challenges. Hydrogeophysics has emerged as a powerful tool in hydrogeology. Using geophysical data, it can gain insight into the hydrological processes and the underlying geology that govern the subsurface (Rubin and Hubbard, 2005; Vereecken et al., 2006). Methods such as electrical resistivity tomography (ERT), ground-penetrating radar (GPR), and seismic surveys enable subsurface imaging and detection of anomalies, which are unexpected or unusual features in the subsurface that stand out from the surrounding material. Furthermore, polluted groundwater typically exhibits increased electrical conductivity (Frohlich and Urish, 2002; Carpenter et al., 2012), making techniques that measure ground electrical conductivity, or its reciprocal electrical resistivity, particularly useful when combined with hydrological data. Consequently, ERT has become a widely used tool in hydrological studies (Page, 1969; Wilson et al., 2006; Pereira

et al., 2023). Hydrogeophysical inverse modeling has been utilized in several research studies to estimate hydraulic properties, such as hydraulic conductivity (Irving and Singha, 2010; Pollock and Cirpka, 2010, 2012), including Kalman-based techniques (Camporese et al., 2011; Crestani et al., 2015; Kang et al., 2019). Nonetheless, there has been a scarcity of studies that have focused on applying ERT measurements to predict groundwater contamination (Kang et al., 2018; Tso et al., 2020; Chen et al., 2023).

In this chapter, three inverse modeling applications are presented, each addressing key challenges in subsurface characterization. The first application serves as an initial test to understand and validate the ES-MDA for estimating hydraulic conductivity by integrating ERT data and measured concentration time series. By comparing results from different datasets, this study provides a basis for assessing the reliability and effectiveness of ES-MDA in hydrogeophysical inversion. Starting from the first application, the second performs a comparative analysis of two inversion methodologies, ES-MDA and ebaPEST, to estimate the electrical resistivity of the subsurface (Fagandini et al., 2025). This comparison highlights key differences in their performance, offering valuable insights into the strengths and limitations of each approach in real-world scenarios. The key innovation of this chapter lies in the third application, where a novel coupled hydrogeophysical methodology is introduced to estimate a contaminant plume within the aquifer. This approach, developed by Fagandini et al. (2024a), replaces the traditional forward model with a CNN, significantly speed-up the solution process. By analyzing five different scenarios with multiple data sources, this study demonstrates the enhanced efficiency and potential of machine learning in hydrogeophysical inversion, thereby facilitating quicker and more precise subsurface characterization.

These applications serve as a baseline for the inversion methodology discussed in **Section 3**, where real data from controlled sandbox experiments are used to test the proposed methodology.

1. Material and Method

This section outlines the methodologies and fundamentals behind the three inversion applications investigated in this chapter. The goal is to provide a comprehensive understanding of the tools, algorithms, and methods used in each case study in addressing the complexities of subsurface characterization and contaminant transport modeling.

1.1. Fundamentals of groundwater hydraulics

The fundamental equations¹ governing the flow and transport in groundwater through fully saturated soil are summarized in this section.

Flow

The continuity equation, representing the conservation of mass for a porous medium saturated with water, is as follows:

$$\text{div}(\mathbf{HK}\nabla h) = S_s \frac{\partial h}{\partial t} \pm q, \quad (2.2)$$

where $\mathbf{HK}$ is the symmetric tensor of the hydraulic conductivity, S_s represents the specific storativity, h is the piezometric head and q is a term of source/sink per unit of volume. Under steady state condition ($\frac{\partial h}{\partial t} = 0$) and without sources or sinks, Eq.2.2 results in:

$$\frac{\partial}{\partial x} \left(HK_{xx} \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(HK_{yy} \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(HK_{zz} \frac{\partial h}{\partial z} \right) = 0. \quad (2.3)$$

The hydraulic conductivity, which measures the ability of aquifer materials to transport water under hydraulic gradients, can be quantified as follows:

$$HK = \frac{k\gamma}{\mu} = \frac{kg}{\nu}, \quad (2.4)$$

¹This section is based on Bear Jacob (1972), Bear Jacob and Bachmat Y. (1990) and Domenico Patrick A. and Schwartz Franklin W. (1998).

where k is the intrinsic permeability of the porous matrix and γ/μ represents the influence of the fluid's properties.

Transport of contaminants

The partial differential equation describing the transport of a non-reactive contaminants of species r in three-dimensional, transient groundwater systems can be written as follows:

$$\frac{\partial (nC^r)}{\partial t} = \frac{\partial}{\partial x_i} \left(nD_{ij} \frac{\partial C^r}{\partial x_j} \right) - \frac{\partial}{\partial x_i} (nv_i C^r) + q_s C_s^r, \quad (2.5)$$

where C^r represents the dissolved concentration of species r ; t is time; x_i is the respective Cartesian coordinate; D_{ij} is the hydrodynamic dispersion coefficient tensor; v_i is the effective velocity; q_s is the volumetric flow rate per unit volume of aquifer representing fluid sources and sinks; C_s^r is the concentration of the source or sink flux for species r . $\frac{\partial}{\partial x_i} (nv_i C^r)$ represents the advection term and describes the transport of miscible contaminants at the same velocity as the groundwater.

1.2. Fundamentals of Electrical Resistivity Tomography (ERT)

This section summarizes the fundamental principles² of current flow in the ground, as applied in resistivity surveys, based on Ohm's Law.

The vector form of Ohm's Law in a continuous medium is expressed as:

$$\mathbf{J} = \sigma \mathbf{E} \quad , \quad (2.6)$$

where σ represents the electrical conductivity of the medium, $\mathbf{J}$ denotes the current density, and $\mathbf{E}$ signifies the electric field intensity. In practice, what is measured is the electric field potential ($\nabla\Phi$). The relationship between the electric potential and the electrical field intensity is described by:

$$\mathbf{E} = -\nabla\Phi. \quad (2.7)$$

²Based on Samouëlian et al. (2005) and Loke (2018)

Combining equations (2.6) and (2.7) yields:

$$\mathbf{J} = -\sigma \nabla \Phi. \quad (2.8)$$

In most surveys, where current sources are point sources, for an elementary volume ΔV surrounding a current source I , located at (x_s, y_s, z_s) , the relationship between current density and current is given by:

$$\nabla \cdot \mathbf{J} = \frac{I}{\Delta V} \delta(x - x_s) \delta(y - y_s) \delta(z - z_s), \quad (2.9)$$

where δ denotes the Dirac delta function. Consequently, Eq.2.8 can be rewritten as:

$$\nabla \cdot [\sigma \nabla \Phi(x, y, z)] = -\frac{I}{\Delta V} \delta(x - x_s) \delta(y - y_s) \delta(z - z_s). \quad (2.10)$$

This constitutes the “forward” modeling problem, i.e., determining the potential observed on a given subsurface structure. A wide range of approaches have been developed to solve Eq.2.10. It is important to note that in geophysical surveys, the resistivity of the medium ρ , which is the inverse of electrical conductivity ($\rho = \frac{1}{\sigma}$), is more commonly used. For the 1D problem, a single-point current source is placed on the ground surface. In this case, the current radiates outward from the source, and the potential decreases inversely with the distance from the current source. The potential in this case is expressed as:

$$\phi = \frac{\rho I}{2\pi r}, \quad (2.11)$$

where r is the distance of a point in the medium (including the ground surface) from the electrode. In practice, ϕ exhibits a symmetrical pattern about the vertical plane at the midpoint between two electrodes. The potential at a point in the medium

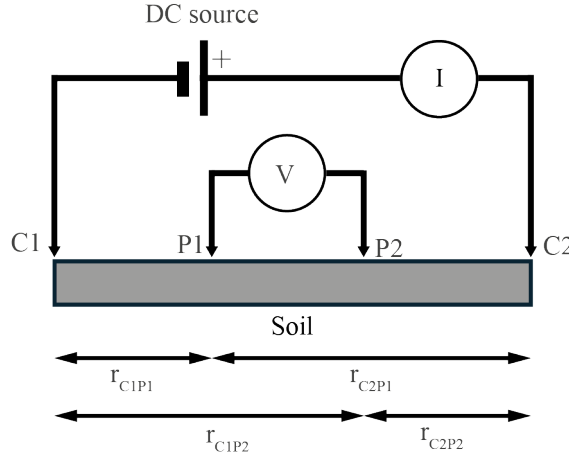


Figure 2.1. Scheme of a conventional four-electrode array for measuring subsurface resistivity.

generated by such a pair of electrodes is given by:

$$\phi = \frac{\rho I}{2\pi} \left(\frac{1}{r_{C1}} - \frac{1}{r_{C2}} \right), \quad (2.12)$$

where r_{C1} and r_{C2} are the distances from the point to the first and second current electrodes, respectively. In most surveys, the potential difference between two points (usually on the ground surface) is measured (Fig.2.1). The potential difference in this arrangement is expressed as:

$$\Delta\phi = \frac{\rho I}{2\pi} \left(\frac{1}{r_{C1P1}} - \frac{1}{r_{C2P1}} - \frac{1}{r_{C1P2}} + \frac{1}{r_{C2P2}} \right), \quad (2.13)$$

where r_{C1P1} is the distance between the current electrode $C1$ and the potential electrode $P1$, r_{C2P1} is the distance between the current electrode $C2$ and the potential electrode $P1$, r_{C1P2} is the distance between the current electrode $C1$ and the potential electrode $P2$, and r_{C2P2} is the distance between the current electrode $C2$ and the potential electrode $P2$. This equation represents the potential measured over a homogeneous half-space with a four-electrode array.

The resistivity measurements are performed by injecting current into the ground via two current electrodes and measuring the voltage difference between two potential electrodes. From the current and potential difference values, an apparent resistivity (ρ_a) is calculated using:

$$\rho_a = K \frac{\Delta\phi}{I}, \quad (2.14)$$

where:

$$K = 2\pi \left(\frac{1}{r_{C1P1}} - \frac{1}{r_{C2P1}} - \frac{1}{r_{C1P2}} + \frac{1}{r_{C2P2}} \right)^{-1} \quad (2.15)$$

is a geometric factor depending on the arrangement of the four electrodes. The calculated resistivity is not the true resistivity of the subsurface but an "apparent" value, representing the resistivity of a homogeneous ground that would produce the same resistance for the given electrode configuration. Determining the true subsurface resistivity from the apparent resistivity values is known as the "inverse" problem. To address any inverse problem, it is essential to first solve the corresponding forward problem, as this provides the necessary understanding of the relationship between the observed and the target quantities. In the forward modeling problem, the subsurface resistivity distribution is specified, and the objective is to compute the apparent resistivity that would be recorded by a geophysical survey over such a structure. Various open-source software and codes are available for this purpose in the literature, see e.g., [Pidlisecky and Knight, 2008](#); [Cockett et al., 2015](#).

Different array configurations have been defined based on the relative position of the potential and current electrodes, including pole-pole (PP), pole-dipole (PD), half-Wenner (HW), Wenner- α (WN), Wenner-Schlumberger (SC), dipole-dipole (DD), Wenner- β (WB), among others ([Dahlin and Zhou, 2004](#)). These arrays facilitate multiple simultaneous measurements for each current injection point, with each configuration offering unique advantages and limitations. In resistivity imaging, these configurations exhibit differences in their imaging capabilities. The electrode spacing of the array controls the depth of the ground section to be characterized: the

greater the distance between the current and potential electrodes, the deeper the investigation. The data are used to draw a 2D “pseudo-section” plot that gives a simultaneous display of both horizontal and vertical variations in resistivity. The horizontal location of the data point corresponds to the midpoint of the set of electrodes used to obtain the measurements. The vertical location of the plotting point is placed at a median depth related to the distance between the electrodes (Edwards, 1977). To ascertain the maximum depth, one must multiply the maximum electrode spacing, or the maximum array length, by the depth factor relevant to the specific investigation array.

One fundamental equation that describes the relationship between the resistivity of a porous rock and the fluid saturation factor is Archie’s Law (Archie, 1942). This law is applicable to certain types of rocks and sediments, especially those with low clay content, where electrical conduction is assumed to occur through the fluids filling the rock’s pores. Archie’s Law is formulated as follows:

$$\rho = a\phi^{-m}S^n\rho_w, \quad (2.16)$$

where ρ_w is the fluid resistivity, ϕ is the porosity (the fraction of the rock filled with fluid), S is the water saturation and a , m and n are empirical parameters known as the tortuosity factor, the cementation exponent, and the saturation exponent, respectively. Typically, for most rocks, a is approximately 1, m is around 2, and n depends on the type of pore fluid, normally around 2.

1.2.1 Petrophysiscal relationship

Petrophysical models are essential for connecting geophysical properties with hydrological models (Vereecken et al., 2006). The model developed by Pollock and Cirpka (2012) is applied to convert solute concentration into electrical conductivity using:

$$\sigma(t, x) = \sigma_0(x) + \sigma'(t, x), \quad (2.17)$$

where $\sigma(t, x)$ represents the bulk electrical conductivity at a given time t and location x , $\sigma_0(x)$ denotes the background bulk electrical conductivity (assumed constant over time), and $\sigma'(t, x)$ is a perturbation due to changes in solute concentration $c(t, x)$. The perturbation $\sigma'(t, x)$ is derived from Archie's law (2.16):

$$\sigma'(t, x) = \frac{\phi^m}{a} S^n \sigma_w c(t, x), \quad (2.18)$$

where σ_w is the electrical conductivity of the water.

1.3. Convolutional Neural Network

Convolutional Neural Networks (CNNs), first developed by Lecun et al. (1998), represent a class of machine learning models designed for processing and analyzing visual data, making them particularly effective for tasks involving images or spatially structured data. At their core, CNNs leverage convolutional filters: small learnable matrices that slide over the input image, capturing spatial hierarchies and local patterns. This allows CNNs to efficiently recognize complex patterns and spatial relationships within the data. Several review papers have been presented in the last few years, offering comprehensive overviews of CNN advancements and applications (see e.g., Gu et al., 2018; Khan et al., 2020; Alzubaidi et al., 2021). A CNN comprises an input layer, several hidden layers, and an output layer. The input layer receives the raw input data in the form of images or other grid-like data. Typically, CNN hidden layers consist of convolutional layers, activation functions, pooling layers, and possibly batch normalization. Convolutional layers apply filters to capture local features. The use of activation functions, such as rectified linear units (ReLU), introduces non-linearity to the model, enhancing its ability to capture intricate patterns. Pooling layers with specified pool sizes and strides downsample the spatial dimensions, reducing computational complexity. Batch normalization may be included to normalize input activations, enhancing training stability. The CNN architecture typically concludes with a fully connected (dense) layer, which

takes the features learned by the convolutional layers and combines them to make predictions. Dropout layers can also be included to mitigate overfitting by randomly deactivating a fraction of neurons during training. Ultimately, the output layer produces the final prediction. The training process involves iterative adjustment of the network weights using optimization algorithms, such as Adam optimizer (Zhang, 2018), to minimize the difference between the predicted and target values.

1.4. Ensemble Smoother with Multiple Data Assimilation (ES-MDA)

The Ensemble Smoother with Multiple Data Assimilation (ES-MDA), introduced by Emerick and Reynolds (2012, 2013), is an iterative data assimilation method that updates unknown parameters while maintaining consistency with observations. The ES-MDA scheme consists of an initialization phase followed by two main iterative steps: a forecast step and an update step. During the initialization phase, an initial ensemble of parameters is defined as $\in \mathbb{R}^{N_p \times N_e}$, where N_p represents the number of parameters to be estimated, and N_e denotes the ensemble size. This ensemble of realizations should be generated using all available information. In cases where prior data are not accessible, it is advisable to generate the ensemble semi-randomly based on expert knowledge. Additionally, this phase involves determining the number of iterations N , generating an ensemble of observed data measurement errors $\epsilon \in \mathbb{R}^{D \times N_e}$, where D indicates the number of observations, and defining the inflation coefficients $\{\alpha_i, i = 1, \dots, N\}$ necessary for the procedure. Observation errors are assumed to follow a Gaussian distribution with mean zero and covariance matrix $R \in \mathbb{R}^{D \times D}$. The coefficients α_i , which are applied to the measurement error and its covariance matrix at each iteration i , help prevent over-corrections. They must satisfy the condition:

$$\sum_{i=1}^N \frac{1}{\alpha_i} = 1, \quad (2.19)$$

which guarantees an exact equivalence between single and multiple data assimilation methods, at least for linear models. Evensen (2018) scheme is used to compute α_i , starting with any non-zero value for α_1^0 . The following α_i^0 are computed as:

$$\alpha_{i+1}^0 = \frac{\alpha_i^0}{\alpha_{\text{geo}}}, \quad (2.20)$$

where the constant α_{geo} controls the extent of the change of α_i from one iteration to the next. Finally, these values are scaled to obtain the final coefficients:

$$\alpha_i = \alpha_i^0 \left(\sum_{i=1}^N \frac{1}{\alpha_i^0} \right). \quad (2.21)$$

A simple choice is to set $\alpha_{\text{geo}} = 1$, leading to constant $\alpha_i = N$. A gradual reduction in α_i , achieved with $\alpha_{\text{geo}} > 1$, can improve the method's performance by reducing the magnitude of initial updates, where the misfit between observations and model predictions is usually larger.

During the forecast step, the iterative process begins. At each iteration i , for each realization j of the ensemble of parameters $\mathbf{X}_j$, where $i \in \mathbb{R}^{N_p}$, the forward model is run to obtain the model predictions. A subset $\mathbf{Y}_j$, where $i \in \mathbb{R}^D$, is extracted that coincides with the same locations and times as the observations $\in \mathbb{R}^D$.

Then, during the update step, parameters are updated for each ensemble realization j and iteration i according to the equation:

$$X_{j,i+1} = X_{j,i} + C_{XY}^i (C_{YY}^i + \alpha_i R)^{-1} (d + \sqrt{\alpha_i} \epsilon_j - Y_{j,i}), \quad (2.22)$$

where d is the vector of observations, C_{XY}^i is the cross-covariance matrix between parameters and predictions, C_{YY}^i is the auto-covariance matrix of predictions, and R is the auto-covariance matrix of the measurement errors, assumed to be uncorrelated.

At each iteration i , these matrices are computed from the ensemble as:

$$C_{XY}^i = \frac{1}{N_e - 1} \sum_{j=1}^{N_e} (X_{j,i} - \overline{X_i}) (Y_{j,i} - \overline{Y_i})^T, \quad (2.23)$$

$$C_{YY}^i = \frac{1}{N_e - 1} \sum_{j=1}^{N_e} (Y_{j,i} - \overline{Y_i}) (Y_{j,i} - \overline{Y_i})^T, \quad (2.24)$$

where $\overline{X_i}$ and $\overline{Y_i}$ are the ensemble means of parameters and predictions, respectively, at iteration i . The process then returns to the forecast step with $X_{j,i} = X_{j,i-1}$ and repeats until the final iteration.

1.5. Empirical Bayes and Akaike's Bayesian Information Criterion

In a Bayesian framework for the inverse problem, the observed data vector $\mathbf{d}^*$ and the unknown parameter vector $\mathbf{m}$ are connected through Bayes' theorem (Bayes, 1763):

$$p(\mathbf{m}|\mathbf{d}^*, I) = \frac{p(\mathbf{d}^*|\mathbf{m}, I)p(\mathbf{m}|I)}{p(\mathbf{d}^*|I)}, \quad (2.25)$$

where: $p(\mathbf{m}|I)$ is the prior probability density function (pdf) of model parameters given prior information I , $p(\mathbf{d}^*|\mathbf{m}, I)$ is the likelihood of observing $\mathbf{d}^*$ given $\mathbf{m}$ and the prior information I ; $p(\mathbf{d}^*|I) = \int p(\mathbf{d}^*|\mathbf{m}, I)p(\mathbf{m}|I)d\mathbf{m}$ serves as a normalizing constant.

The parameter vector $\mathbf{m}$ is treated as stochastic, characterized by a prior mean $\mathbf{s}$ and covariance $\mathbf{C}_m$, where $E[\mathbf{m}] = \mathbf{s}$ and $E[(\mathbf{m} - \mathbf{s})(\mathbf{m} - \mathbf{s})^T] = \mathbf{C}_m$. The noise vector $\boldsymbol{\nu}$ is random with zero mean and covariance $\mathbf{C}_d$. Assuming $\boldsymbol{\nu}$ is independent and identically distributed, $\mathbf{C}_d = \sigma_d^2 \mathbf{I}$, where σ_d quantifies the discrepancy between observations and model predictions. Under Gaussian assumptions, the posterior probability remains Gaussian (Tarantola, 2005).

The likelihood function is:

$$p(\mathbf{d}^*|\mathbf{m}, I) = (2\pi)^{-D/2} |\mathbf{C}_d|^{-1/2} \exp \left(-\frac{1}{2} (\mathbf{d}^* - \mathbf{Gm})^T \mathbf{C}_d^{-1} (\mathbf{d}^* - \mathbf{Gm}) \right), \quad (2.26)$$

and the prior pdf is:

$$p(\mathbf{m}|I) = (2\pi)^{-N_p/2} |\mathbf{C}_m|^{-1/2} \exp \left(-\frac{1}{2} (\mathbf{m} - \mathbf{s})^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{s}) \right). \quad (2.27)$$

The posterior pdf can be expressed as:

$$p(\mathbf{m}|\mathbf{d}^*, I) = (2\pi)^{-N_p/2} |\mathbf{C}_q|^{-1/2} \exp \left(-\frac{1}{2} (\mathbf{m} - \hat{\mathbf{m}})^T \mathbf{C}_q^{-1} (\mathbf{m} - \hat{\mathbf{m}}) \right), \quad (2.28)$$

where N_p is the number of parameters, D is the number of observations, $\hat{\mathbf{m}}$ and $\mathbf{C}_q$ are the mean and covariance of the posterior probability density function (pdf), computed as follows (Tarantola, 2005):

$$\hat{\mathbf{m}} = \mathbf{s} + \mathbf{C}_m \mathbf{G}^T \left(\mathbf{G} \mathbf{C}_m \mathbf{G}^T + \mathbf{C}_d \right)^{-1} (\mathbf{d}^* - \mathbf{Gs}), \quad (2.29)$$

$$\mathbf{C}_q = \mathbf{C}_m - \mathbf{C}_m \mathbf{G}^T \left(\mathbf{G} \mathbf{C}_m \mathbf{G}^T + \mathbf{C}_d \right)^{-1} \mathbf{G} \mathbf{C}_m. \quad (2.30)$$

The normalization term, assuming Gaussian distributions, can be written as (Mitsuhata, 2004):

$$p(\mathbf{d}^*|I) = (2\pi)^{-D/2} |\mathbf{C}_{dp}|^{-1/2} \exp \left(-\frac{1}{2} (\mathbf{d}^* - \mathbf{Gs})^T \mathbf{C}_{dp}^{-1} (\mathbf{d}^* - \mathbf{Gs}) \right), \quad (2.31)$$

with

$$\mathbf{C}_{dp} = \mathbf{G} \mathbf{C}_m \mathbf{G}^T + \mathbf{C}_d. \quad (2.32)$$

In empirical Bayes, a two-step procedure is adopted: (1) estimate the unknown model parameters $\mathbf{m}$ from observed data $\mathbf{d}^*$, and (2) estimate hyperparameters. The first step uses Eq.2.29, while various methods can be applied for the second step, including the Akaike's Bayesian Information Criterion (ABIC). ABIC has been ef-

fectively applied in geophysics (Woodbury and Ferguson, 2006), contaminant source identification (Zanini and Woodbury, 2016) and aquifer characterization (Zanini et al., 2017, 2020) problems.

The ABIC is computed in terms of Bayesian likelihood (Ulrych et al., 2001) and is defined as:

$$\text{ABIC} = -2 \ln p(\mathbf{d}^*|I) + 2N_h, \quad (2.33)$$

where N_h is the number of hyperparameters (i.e., structural covariance parameters). Combining Eq.2.31 and Eq.2.33 (Mitsuhata, 2004; Woodbury and Ferguson, 2006), the ABIC for the linear inverse problem becomes:

$$\text{ABIC} = N_p \ln(2\pi) + \ln |\mathbf{C}_{\mathbf{d}\mathbf{p}}| + (\mathbf{d}^* - \mathbf{G}\mathbf{s})^T \mathbf{C}_{\mathbf{d}\mathbf{p}}^{-1} (\mathbf{d}^* - \mathbf{G}\mathbf{s}) + 2N_h. \quad (2.34)$$

The goal is to identify hyperparameter values that minimize the ABIC.

For a general nonlinear relationship between the observations $\mathbf{d}^*$ and the model parameters $\mathbf{m}$, the measurement equation can be expressed as:

$$\mathbf{d}^* = \mathbf{g}(\mathbf{m}) + \epsilon, \quad (2.35)$$

where $\mathbf{g}(\mathbf{m})$ indicates the nonlinear transfer function. When the non-linearity is not too strong, the relationship between observations and parameters can be iteratively linearized around the current solution using a quasi-linear approximation (Tarantola, 2005). Given the current estimate of the parameters $\tilde{\mathbf{m}}$, the function $\mathbf{g}$ can be approximated by expanding it as:

$$\mathbf{g}(\mathbf{m}) \approx \mathbf{g}(\tilde{\mathbf{m}}) + \tilde{\mathbf{G}}(\mathbf{m} - \tilde{\mathbf{m}}), \quad (2.36)$$

where

$$\tilde{\mathbf{G}} = \left. \frac{\partial \mathbf{g}}{\partial \mathbf{m}} \right|_{\mathbf{m}=\tilde{\mathbf{m}}} \quad (2.37)$$

is the derivative of $\mathbf{g}$ with respect to $\mathbf{m}$ at $\tilde{\mathbf{m}}$, also known as the sensitivity matrix

or Jacobian. Substituting Eq.2.37 into Eq.2.36 gives:

$$\mathbf{d}^* - \mathbf{g}(\tilde{\mathbf{m}}) + \tilde{\mathbf{G}}\tilde{\mathbf{m}} = \tilde{\mathbf{G}}\mathbf{m} + \epsilon, \quad (2.38)$$

and by defining

$$\tilde{\mathbf{d}}^* = \mathbf{d}^* - \mathbf{g}(\tilde{\mathbf{m}}) + \tilde{\mathbf{G}}\tilde{\mathbf{m}}, \quad (2.39)$$

the linearized problem can be rewritten as:

$$\tilde{\mathbf{d}}^* = \tilde{\mathbf{G}}\mathbf{m} + \epsilon, \quad (2.40)$$

which is formally equivalent to the linear problem described in Eq.2.1. The non-linearity between the observations and the parameters means that both parameters and hyperparameters must be estimated iteratively.

2. Applications

The first application aims to estimate hydraulic conductivity by integrating ERT data with time-lapse concentration measurements using the ES-MDA method. The effectiveness of various datasets is assessed through comparative analysis. In the second application, the ES-MDA and ebaPEST approaches (Zanini et al., 2017) are used to estimate subsurface electrical resistivity, and their performance is compared. The third application presents an innovative coupled hydrogeophysical technique (Fagandini et al., 2024a) to compute a contaminant plume within an aquifer. Five scenarios are examined to evaluate the impact of multiple data sources, using a surrogate model to speed up forward modeling and the solution process. In this study, the flow equation (Eq.2.2) is solved using the numerical model MODFLOW (Harbaugh, 2005), and the transport equation (Eq.2.5) with MT3DMS (Zheng and Wang, 1999). The geophysical forward models applied are those proposed by Pidlisecky and Knight (2008) and Cockett et al. (2015). The software package genES-MDA developed by Todaro et al. (2022b) is used to apply the ES-MDA procedure, while the open-source software Empirical Bayesian Approach within the PEST (ebaPEST) developed by Zanini et al. (2020) is used for the application of the empirical Bayes approach.

2.1. Evaluation of the Hydraulic Conductivity

Accurate hydraulic conductivity estimates are crucial for understanding and predicting flow and transport in porous media, making this an important aspect for hydrogeological modeling.

This section focuses on the application of the Ensemble Smoother with Multiple Data Assimilation (ES-MDA - see **Subsection 1.4**) to estimate the spatial distribution of a hydraulic conductivity field. In this study, the estimation process is improved by integrating electrical resistivity tomography (ERT) data with time-varying pollutant concentration measurements. This combination allows for a more

comprehensive understanding of subsurface properties, highlighting the potential of hydrogeophysics in groundwater applications. The inversion methodology is evaluated using a two-dimensional synthetic case study, designed to simulate realistic groundwater flow conditions. Three scenarios are explored, each using different types of observation data. The analysis evaluates how these varying observations impact the estimation of the hydraulic conductivity field.

2.1.1 Synthetic Case

Set up of the Test Case

The effectiveness of the proposed methodology is evaluated through its application to a two-dimensional synthetic model, which simulates a vertical cross-section of a heterogeneous, fully saturated, unconfined aquifer. The layout of the model is illustrated in Fig. 2.2. The aquifer measures 96 cm by 20 cm and is discretized into a grid composed of 96 by 20 cells, where each cell measures 1 cm x 1 cm. The porosity is assumed to be constant at 0.37 throughout the aquifer. The boundary conditions include an impermeable base, a phreatic surface at the top, and fixed hydraulic heads upstream (left in Figure 2.2) and downstream (right in Figure 2.2). This configuration creates a head loss of 2 cm, driving groundwater flow from left to right throughout the model. The reference hydraulic conductivity solution has three clear homogeneous zones as seen in Fig. 2.2: 0.17 cm/s (-0.77 in log10) in dark blue, 0.69 cm/s (-0.16 in log10) in light blue and 10.4 cm/s (1.02 in log10) in yellow. The hydraulic conductivity parameters to be estimated correspond to the value in each model grid cell ($Np = 1920$).

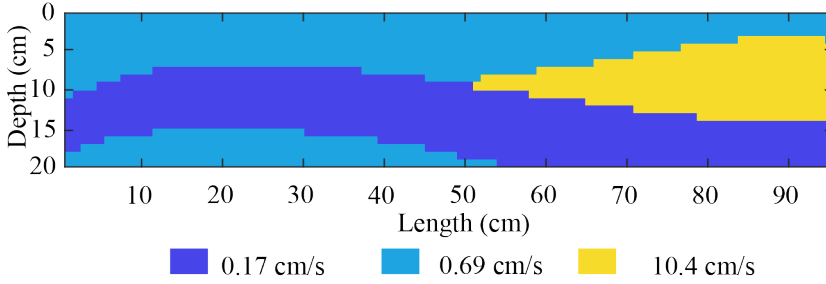


Figure 2.2. 2D synthetic model of hydraulic conductivity.

Three different datasets are examined for their impact on inversion results. These datasets are designed to demonstrate the advantages of employing specific combinations of apparent resistivity measurements (D_1) and time-varying concentration data (D_2). The two observed variables are shown in Fig. 2.3.

In the first case, time-varying concentration data are used. The observations utilized in the inverse modeling procedure are derived from a simulation executed with MODFLOW 2005 (Harbaugh, 2005) for groundwater flow and MT3DMS (Zheng and Wang, 1999) for mass transport processes. A tracer with a mass loading of 20 mg/s is continuously injected across the full vertical left side of the model. The dispersivity parameters are set with a longitudinal dispersivity of 0.16 cm and a transverse dispersivity of 0.016 cm. The simulation lasts for 5000 seconds, with concentration measurements recorded every 500 seconds at 198 monitoring locations evenly distributed 2 cm apart in both directions (Fig. 2.3 - top). This results in a total of 1980 concentration measurements ($D_1 = 1980$).

In the second case, only apparent resistivity values are used as observations. The relationship between the intrinsic permeability, k , and the saturated electrical conductivity, σ_s , is explored empirically and through derivation from models (Brace et al., 1968; Walsh and Brace, 1984; Mawer et al., 2015). It was found that k is related to σ_s by a power relation:

$$k = a\sigma_s^b, \quad (2.41)$$

where a is a proportionality constant and b is typically between 1 and 3. Knowing k from Eq.2.4 and combining it in Eq.2.41 the reference resistivity model under saturated conditions is calculated. The observed values are then derived from the geophysical forward model (SimPEG framework - Cockett et al., 2015) based on the reference resistivity field, using a Wenner-Schlumberger acquisition array with 32 electrodes spaced 3 cm apart. This approach yields 225 apparent resistivity measurements ($D_2 = 225$).

The third case involves the integration of both concentration data and apparent resistivity measurements ($D_1 = 1980$ and $D_2 = 225$). This combined approach aims to leverage the strengths of each dataset to improve the accuracy of hydraulic conductivity estimations. By synthesizing information from both sources, the model seeks to improve the estimation of the hydraulic conductivity field by highlighting the benefits of using more resources.

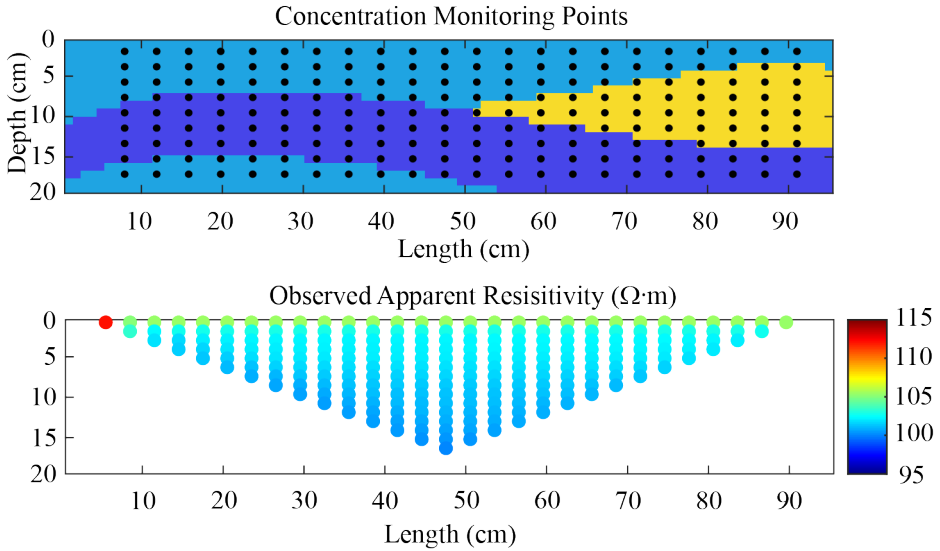


Figure 2.3. Concentration monitoring points overlaid on the reference hydraulic conductivity model (top) and observed apparent resistivity (bottom).

A summary of the three scenarios is provided in Table 2.1.

Inverse model set up

Table 2.1. Summary of the scenarios. Number of observations used.

Scenario	1	2	3
ERT	-	225	225
Concentration	1980	-	1980

The ES-MDA algorithm is performed over six iterations using homogeneous ensembles of 500 realizations, each drawn randomly within the interval $[0.1, 3]$ cm/s. The parameter vector undergoes transformation before the update step and is then reverted to its original parameter space afterward. Specifically, the parameters are logarithmically transformed for the update and subsequently back-transformed into a bounded range of $[0.05, 15]$ cm/s. The observation errors are normally distributed with zero mean in all scenarios but with different variances: 10^{-4} (cm/s)² for Scenario 1, 1 (Ohm·m)² for Scenario 2 and 0.1 (cm/s)² and 0.1 (Ohm·m)² for Scenario 3. A spatial covariance localization is applied with a correlation length of 30 cm and covariance inflation is used with a factor of 1.01. A decreasing α set with values equal to $[364.0; 121.3; 40.4; 13.5; 4.5; 1.5]$ is used.

2.1.2 Results

This section evaluates the performance of the ES-MDA aimed at reconstructing the hydraulic conductivity field based on various data inputs. The true field, shown in Fig. 2.2, serves as a reference model to evaluate the accuracy of the three different inversion scenarios (Table 2.1). The performance of each Scenario is quantified using the evaluation metrics described in **Appendix A**. The results are summarized in Table 2.2 for parameter estimation and Table 2.3 for observation estimation. Figure 2.4 shows the resulting log10-hydraulic conductivity field for scenarios 1 to 3.

In the first Scenario, where only time-varying concentration data were used, the inversion yielded relatively high error metrics ($ME = -1.215$ cm/s, $MAE = 1.997$ cm/s, and $RMSE = 3.898$ cm/s). Although this Scenario provides a general understanding of the field's spatial distribution, it lacks precision in capturing hydraulic

conductivity values, despite the large number of observations. This suggests that relying solely on concentration data may limit the inversion's accuracy. The second scenario, which used exclusively apparent resistivity data, showed modest improvements in the estimation metrics ($ME = 0.531$ cm/s, $MAE = 1.728$ cm/s, and $RMSE = 3.566$ cm/s). The results reflect a slightly better alignment compared to the first Scenario. Visual inspection reveals an overestimation in the region of maximum permeability on the right, likely due to the absence of resistivity data acquisition in that area, resulting in propagated errors during the inversion process. The estimated field highlights the utility of apparent resistivity data, which, though limited, provide better resolution of hydraulic conductivity variations compared to concentration data alone. However, it cannot accurately reconstruct the field's spatial distribution. The final Scenario, which combined in-time solute concentration and apparent resistivity data, produced the most accurate inversion results with the lowest error metrics ($ME = -0.802$ cm/s, $MAE = 1.471$ cm/s, and $RMSE = 2.925$ cm/s). This combined approach led to a more precise reconstruction of the hydraulic conductivity field, particularly in capturing both large-scale patterns and finer spatial variations. The improvement in performance suggests that integrating multiple types of data increases the amount of information available for inversion, thereby enhancing the field's characterization. Overall, the comparison demonstrates that while each data type individually contributes to the inversion, their integration significantly improves the accuracy and detail of the reconstructed hydraulic conductivity field.

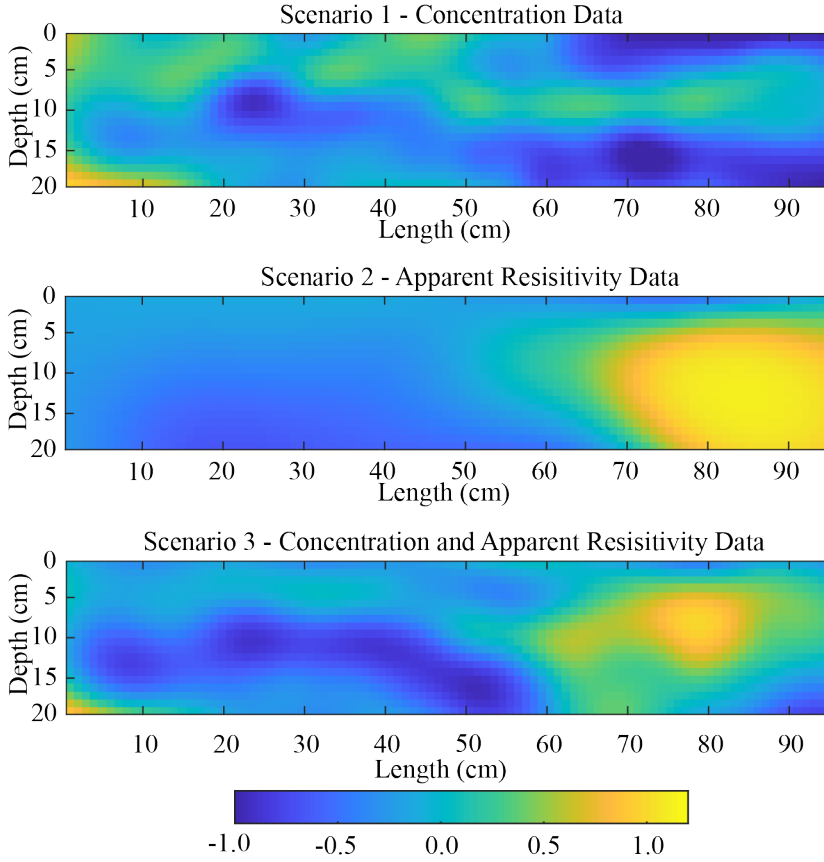


Figure 2.4. Log10-estimated hydraulic conductivity fields for the three scenarios.

Table 2.2. Evaluation of the inversion approach performance assessed on the estimation of the hydraulic conductivity (cm/s) for each scenario.

Scenario	ME	MAE	RMSE
1	-1.215	1.997	3.898
2	0.531	1.728	3.566
3	-0.802	1.471	2.925

Considering the observation estimation, the performance metrics for each Scenario confirm the results obtained in parameter estimation. In the first Scenario

employing time-concentration data, the model demonstrated a fairly accurate alignment between observed and predicted concentration values (Table 2.3). However, the errors suggest that the model struggled to capture finer details, affecting the accuracy of the field estimation. In contrast, the second Scenario, which involved the use of apparent resistivity, demonstrated significantly improved performance metrics (Table 2.3). The results indicate a remarkable alignment with the observed resistivity data. The third test incorporated both solute concentration and apparent resistivity data for a combined estimation approach. The performance metrics revealed mixed results (Table 2.3); while the resistivity metrics remained high, reflecting excellent accuracy, the concentration metrics were not as favourable as in the first test. This indicates the importance of choosing good ES-MDA input settings.

Table 2.3. Assessment of observation estimates performance for concentration (mg/L) and resistivity (Ohm·m) across each scenario.

Scenario	Concentration				Resistivity			
	<i>ME</i>	<i>MAE</i>	<i>RMSE</i>	<i>R²</i>	<i>ME</i>	<i>MAE</i>	<i>RMSE</i>	<i>R²</i>
1	0.168	0.901	1.889	0.957	-	-	-	-
2	-	-	-	-	0.008	0.008	0.014	0.999
3	0.434	1.426	3.191	0.881	0.021	0.024	0.038	0.999

Figure 2.5 shows the standard deviation maps for the three scenarios, providing insight into the uncertainty of the estimates, reflecting the impact of each dataset in the inversion process. The color scale ranges from blue (low standard deviation, higher estimation accuracy) to yellow (high standard deviation, greater uncertainty). In Scenario 1, when only concentration data are used for inversion (Fig. 2.5, above), the standard deviation remains low throughout the domain, as indicated by the predominant blue coloration, which can be attributed to the large number of concentration data points (1980 observations). In contrast, Scenario 2 (Fig. 2.5, middle) exhibits higher standard deviation, especially at greater depths. This is likely due to the indirect relationship between resistivity and hydraulic conductivity, as well as the smaller dataset (225 observations), which is limited to the central region

and not distributed over time like concentration in Scenario 1. Finally, in Scenario 3 (Fig. 2.5, bottom), integrating concentration and apparent resistivity data produces a standard deviation distribution similar to Scenario 1 but with lower values (deeper blue tones), indicating reduced estimation uncertainty. This is primarily due to the dominant influence of the large dataset (1980 concentration + 225 apparent resistivity), which ensures a well-constrained estimation. The findings highlight the importance of dataset selection and data density in geophysical and hydrogeological inversion applications.

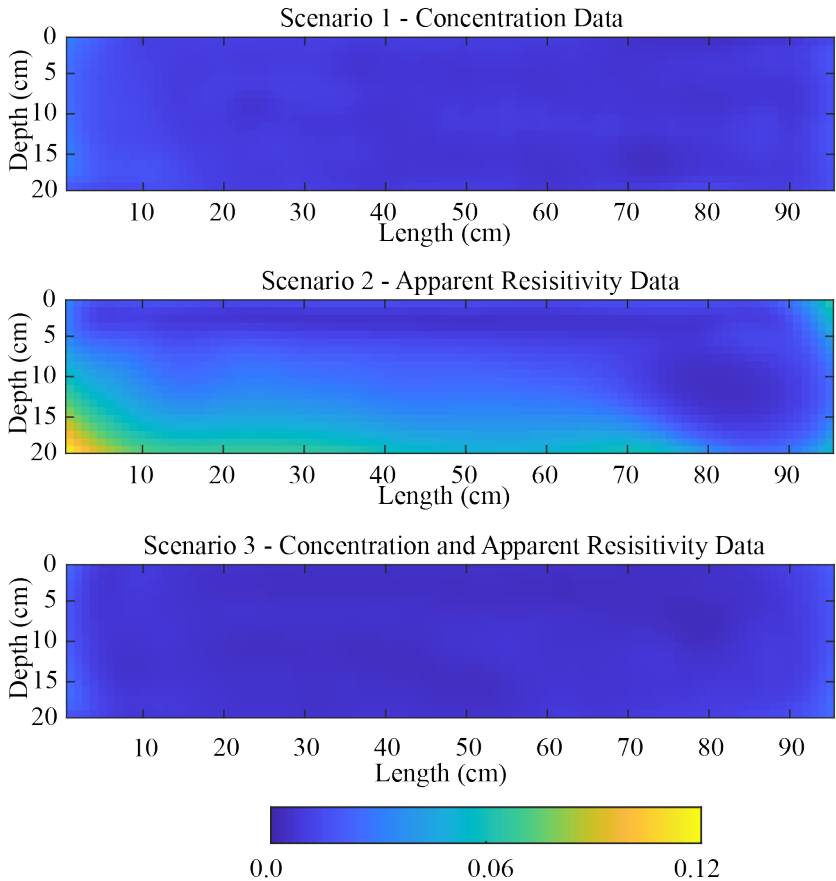


Figure 2.5. Standard deviation of the $\log_{10}$ -estimated hydraulic conductivity fields for the three scenarios.

2.1.3 Concluding remarks

The primary objective of this application was to reconstruct a synthetic hydraulic conductivity field using time-varying solute concentration and apparent resistivity data from ERT measurements. The study utilized the ES-MDA methodology to assess the impact of various data configurations in order to accomplish this goal. Scenario 1 utilized only solute concentration data, Scenario 2 relied on apparent resistivity data, and Scenario 3 combined both datasets. The results indicate that the combined approach in Scenario 3 significantly outperformed the other tests, yielding the lowest errors and highest R^2 values. This demonstrates the advantages of integrating multiple data types to enhance inversion accuracy. Scenario 1 provided a broad understanding of field patterns but lacked precision in capturing the finer details of the hydraulic conductivity values. However, its large dataset (1980 concentration measurements) contributed to a lower standard deviation, indicating a more stable and constrained inversion. Scenario 2 showed improved performance, particularly in terms of observation data accuracy, though it still faced challenges in fully resolving the complex features of the true field. The higher standard deviation observed in Scenario 2, particularly at deeper levels, result from the smaller data set of 225 apparent resistivity measurements that are less uniformly distributed in space and lack temporal coverage, unlike the concentrations in Scenario 1. Scenario 3, which integrated both datasets (1980 concentration + 225 apparent resistivity), resulted in the lowest standard deviation, confirming that the combination of multiple data types, along with a higher number of observations, enhances inversion accuracy while reducing uncertainty.

The choice of the true hydraulic conductivity field posed a challenging benchmark for inversion methods, as accurately reconstructing both sharp boundaries and gradual transitions between conductive and less conductive areas is crucial for successful inversion. Furthermore, the use of heterogeneous ensemble field realizations could further improve the accuracy of the estimates, as incorporating spatial variability may provide a more realistic representation of subsurface conditions and enhance

the inversion process.

Future work will focus on refining the ES-MDA technique with particular emphasis on addressing issues related to underestimation and improving computational efficiency. Moreover, exploring different petrophysical relationships that link geophysical and hydrological data with hydraulic conductivity estimation could offer significant improvements to the estimation.

2.2. Comparative Analysis of Ensemble Smoother with Multiple Data Assimilation and Empirical Bayes PEST: Methodological Evaluation

Geophysics plays a vital role in subsurface characterization, enabling researchers to infer key physical properties of the Earth's interior without direct access. Resistivity data are useful for delineating differences in subsurface materials, including groundwater and geological formations, which are relevant for environmental and resource research.

This section focuses on estimating a synthetic resistivity field using apparent resistivity data and observed field measurements of resistivity. The study evaluates the performance of two approaches, Ensemble Smoother with Multiple Data Assimilation (ES-MDA) and Empirical Bayes PEST (ebaPEST) combined with Akaike's bayesian information criterion (ABIC), in assimilating geophysical data to estimate subsurface resistivity distributions. The research aims to assess the accuracy, computational efficiency, and handling of uncertainty by comparing the two methods, offering valuable insights for geophysical model calibration and resistivity field estimation.

The application presented shows the preliminary results of Fagandini et al. (2025) and shared at the geoENV conference in Chania, Greece, held from 19-21 June 2024.

2.2.1 Description of the methods

The ES-MDA procedure (see **Subsection 1.4**), optimized for the problem under investigation and illustrated in the flowchart in Fig. 2.6, is detailed below:

1. **Initialization:** the first step involves generating the initial parameter realizations, which are random resistivity fields representing different potential subsurface scenarios. During this phase, the number of iterations and observation errors are defined, ensuring that uncertainties in the data are properly accounted for;

2. **Forecast:** in the second phase, upon entering the loop, each realization from the ensemble is input into the forward model (Pidlisecky and Knight, 2008) to generate predictions. Additionally, point resistivity values are extracted from boreholes along the field, providing supplementary observational data;
3. **Update:** in the third step, using Eq.2.22, the resistivity fields are updated based on discrepancies between predictions and observations. This iterative update process is applied to every realization in the ensemble. Once all realizations are updated, the next iteration begins with the new adjusted ensemble. The cycle of prediction and updating continues until the final iteration is completed, progressively refining the resistivity model.

The findings of the final iteration, which reflect the average of the ensemble of parameter realizations, are analyzed using the evaluation metrics (**Appendix A**). The workflow is implemented using a Python open-source package (Todaro et al., 2022b), which efficiently manages the ensemble updating process and integrates the necessary components of ES-MDA.

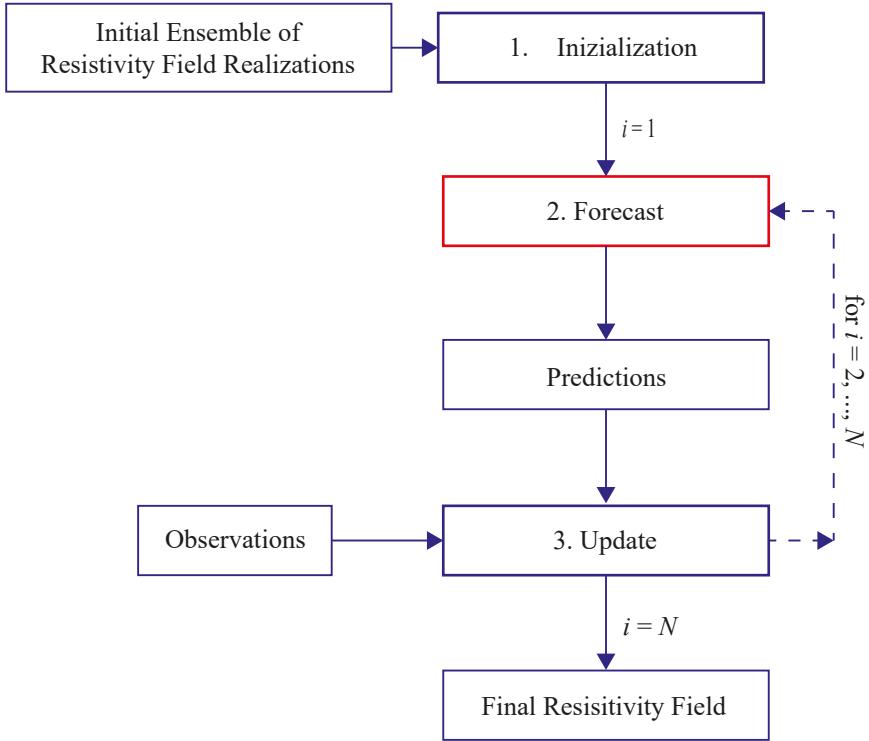


Figure 2.6. Workflow of the ES-MDA method.

The ebaPEST technique (see **Subsection 1.5**), optimized for the problem under investigation and illustrated in the flowchart in Fig. 2.7, is detailed as follows:

1. **Initialization:** an initial estimate of the resistivity fields $\tilde{\mathbf{m}}$ is provided, along with the initial hyperparameters for the covariance model and the error variance σ_d^2 ;
2. **Forward Run:** use the forward model (Pidlisecky and Knight, 2008) to compute predictions $\mathbf{g}(\tilde{\mathbf{m}})$ and calculate the Kernel function $\tilde{\mathbf{G}}$ at the current parameter estimate $\tilde{\mathbf{m}}$ using a finite difference approach;
3. **Updating Parameters:** compute the modified data vector $\tilde{\mathbf{d}}^*$ and form the covariance matrices for the model $\mathbf{C}_m$ and the data $\mathbf{C}_d$. Then, update the

model parameters $\tilde{\mathbf{m}}$ using the equation:

$$\tilde{\mathbf{m}} \approx \mathbf{s} + \mathbf{C}_m \tilde{\mathbf{G}}^T \left(\tilde{\mathbf{G}} \mathbf{C}_m \tilde{\mathbf{G}}^T + \mathbf{C}_d \right)^{-1} \left(\tilde{\mathbf{d}}^* - \tilde{\mathbf{G}} \mathbf{s} \right)$$

and compute the residual between the observed and computed data;

4. **Objective Function Check:** evaluate the objective function (defined as the *RMSE* in this application) of the difference between the observed and predicted data, normalized by the range of the observations. If the objective function falls below a specified tolerance, exit the loop; otherwise, return to step 2 for further updates;
5. **Minimize the ABIC:** form the covariance matrix for the data predictions, $\mathbf{C}_{dp} \approx \tilde{\mathbf{G}} \mathbf{C}_m \tilde{\mathbf{G}}^T + \mathbf{C}_d$, compute the natural logarithm of its determinant, and calculate $\left(\tilde{\mathbf{d}}^* - \tilde{\mathbf{G}} \mathbf{s} \right)^T \mathbf{C}_{dp}^{-1} \left(\tilde{\mathbf{d}}^* - \tilde{\mathbf{G}} \mathbf{s} \right)$. Use this to minimize the ABIC (Eq.2.33) with respect to the hyperparameters;
6. **Hyperparameter Convergence Check:** check for convergence of the hyperparameters by evaluating the improvement in the objective function between two successive iterations using the formula:

$$2 \cdot \frac{|\text{ABIC}_i - \text{ABIC}_{i-1}|}{|\text{ABIC}_i| + |\text{ABIC}_{i-1}|}$$

If this improvement falls below a predefined tolerance, exit the loop; otherwise, return to step 2.

Once convergence is achieved, compute the final resistivity field, the corresponding covariance matrix, and the final ABIC value, completing the estimation process.

The inverse procedure outlined above has been implemented through the open-source software Empirical Bayesian Approach within the PEST (Zanini et al., 2020) (ebaPEST), developed for highly parameterized inversion. This software operates independently of any specific forward model, as long as it adheres to the PEST framework.

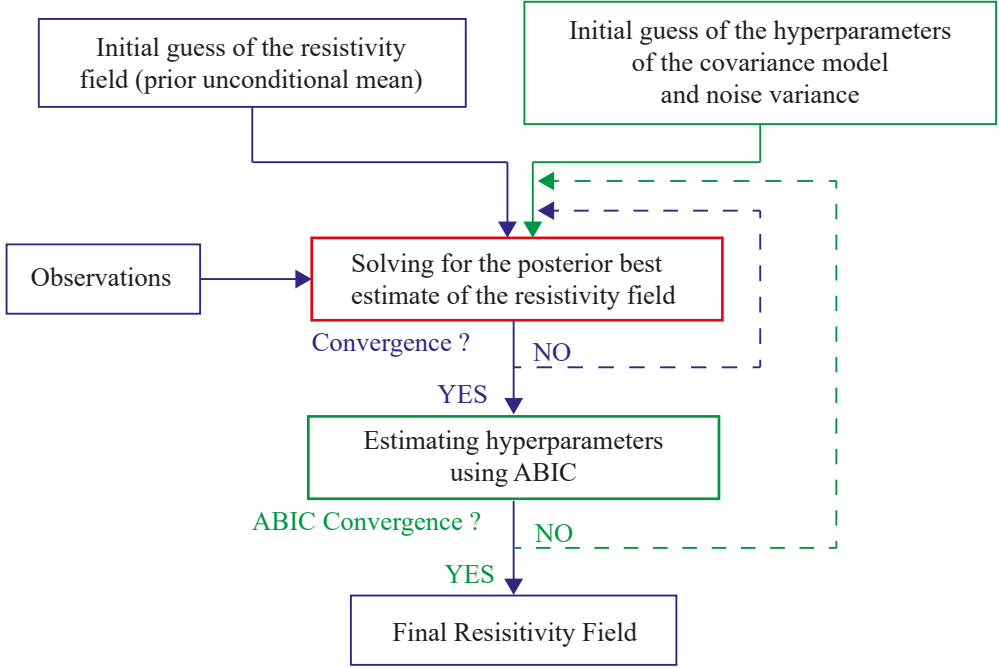


Figure 2.7. Workflow of the ebaPEST method.

2.2.2 Synthetic Case

Set up of the Test Case

The synthetic resistivity field (Fig. 2.8) illustrates a two-dimensional distribution of resistivity values, visually represented by distinct colors: 50 Ohm·m in dark blue, 250 Ohm·m in light blue, and 500 Ohm·m in green. These significantly different values were deliberately chosen to clearly highlight variations in the properties within the model. The model has dimensions of 96 cm x 20 cm, discretized into a grid of 48 by 10 cells, each cell measuring 2 cm x 2 cm. The parameters to be estimated correspond to the resistivity in each model grid cell ($Np = 480$). The model assumes fully saturated conditions, with two boreholes positioned vertically at $x = 23$ and $x = 47$ cm. Each borehole contains 10 observation points, totaling 20 observation locations. Apparent resistivity values, shown in the scatter plot overlaid on the model in Fig. 2.8, were obtained by running the synthetic ERT field through the forward model.

The observation points mark the locations where apparent resistivity measurements were recorded using a Wenner-Schlumberger array composed of 32 electrodes spaced at 3 cm intervals, generating 225 observations.

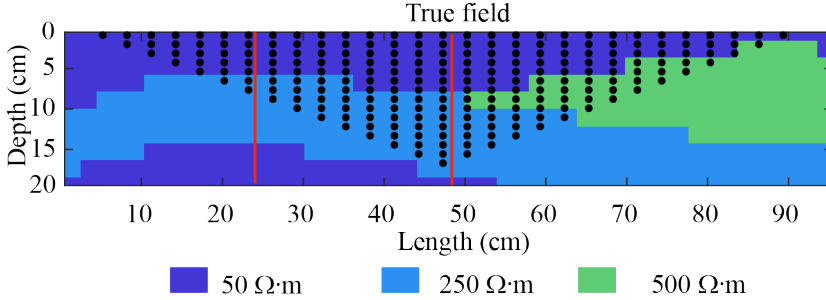


Figure 2.8. 2D synthetic resistivity model. Black dots represent the locations of recorded apparent resistivity, while the red lines indicate the positions of the resistivity monitoring points.

Inverse model set up

The EBA and ES-MDA are compared to estimate the same number of parameters ($Np = 480$) using an identical set of observations ($D = 245$).

The ES-MDA is executed with six iterations and an ensemble size of 500. The initial ensemble of parameters is generated using stochastic sequential simulations (Pyrzcz et al., 2021). This process generates Gaussian random fields in logarithmic space, which are subsequently back-transformed into the resistivity space. Each realization is characterized by an anisotropic exponential semi-variogram, with the azimuth of maximum continuity randomly chosen within the range of $[80, 100]$ degrees. The mean log-resistivity values for the initial ensemble are sampled from a uniform distribution within the interval $[5, 5.5]$, with a standard deviation ranging between $[0.5, 1]$. The vertical correlation range is randomly selected from a uniform distribution with values between $[2, 8]$ cm, while the anisotropy ratio is sampled from the interval $[7, 10]$. Observation errors are assumed to follow a normal distribution with a zero mean and a variance of $1 (\text{Ohm} \cdot m)^2$. A decreasing α set with values equal to $[364.0; 121.3; 40.4; 13.5; 4.5; 1.5]$ is used. Additionally, spatial covariance lo-

calization is applied with a correlation length of 20 cm to enhance the assimilation process and limit the influence of distant observations, ensuring the preservation of localized correlations. In this case, no covariance inflation is applied.

In the ebaPEST setup, the inversion process begins with a uniform initial resistivity field, representing a single homogeneous field derived from the average resistivity of the observations (140 Ohm·m). This serves as a neutral starting point, from which updates are progressively made based on the observed data and model responses. To ensure non-negativity, the resistivity field is estimated in the logarithmic space. The prior covariance function is modeled using an exponential model. The variance of the observation errors is assumed to be known and set to $0.2 \text{ (Ohm}\cdot\text{m)}^2$. The correlation length is set to 100 cm. This choice of a long correlation length and low variance emphasizes a smooth inversion process, leading to gradual and continuous updates to the model during the assimilation of geophysical data. As the inversion proceeds, ebaPEST adapts the model parameters, refining the semi-variogram and covariance structure as additional data is integrated. This iterative refinement ensures that the model converges towards a solution that fits the observed resistivity values while balancing smoothness and accuracy. The emphasis on regularization in ebaPEST helps avoid overfitting while preserving larger-scale resistivity structures. Throughout the parameter estimation process, a Jacobian sensitivity matrix ($\tilde{\mathbf{G}}$) must be calculated, requiring one model run per parameter.

2.2.3 Results

The results of the inversion are presented by comparing the logarithmic reference resistivity field (Fig. 2.8 - top) with the estimated fields obtained using ES-MDA (Fig. 2.9 - middle) and ebaPEST (Fig. 2.9 - bottom). For the ES-MDA, the mean of the ensemble realizations at the final iteration is presented.

The estimated ES-MDA field (Fig. 2.9 - middle) shows a more gradual transition compared to the true field (Fig. 2.9 - top). Although the model effectively represents the overall distribution of the resistivity field, it lacks distinct boundaries between

resistivity zones, resulting in more diffuse variations. In the estimated ebaPEST field (Fig. 2.9 - bottom), the reconstruction appears to better delineate resistivity transitions and maintain some sharper contrasts, especially in the high-resistivity region near the right-hand side of the domain. However, it also introduces some localized overestimations of resistivity, particularly in the deeper regions. Both reconstructions closely align with the true resistivity distribution. However, the EBA provides smoother transitions between layers influenced by the underlying covariance function, whereas the ES-MDA method better captures the block-like properties of the true field.

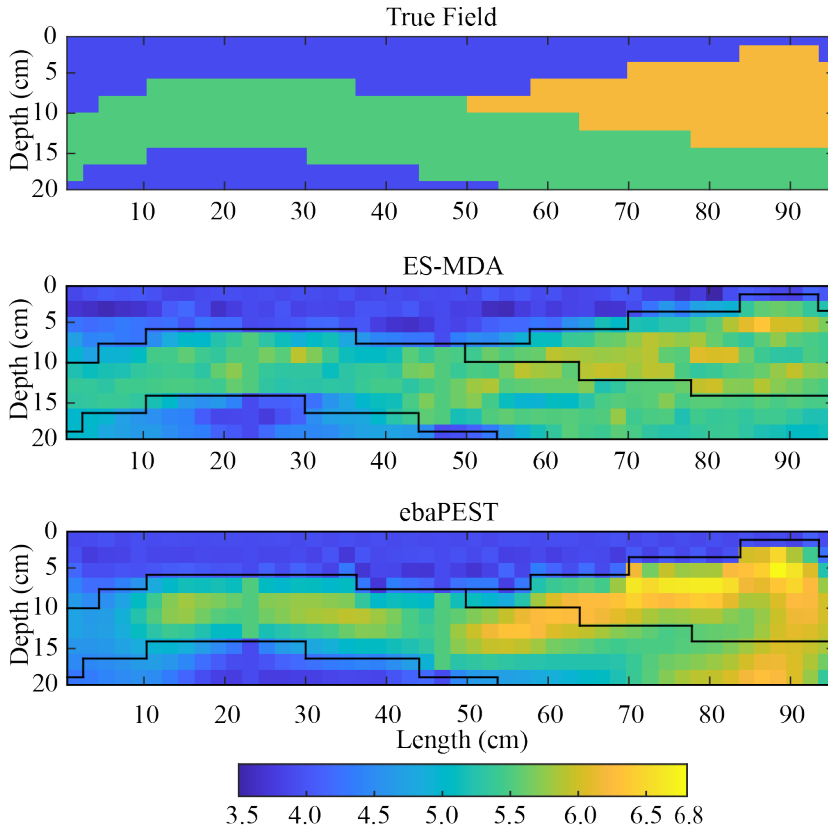


Figure 2.9. Comparison of estimated results using the ES-MDA vs. ebaPEST.

The effectiveness of each method can be quantified by evaluating the error metrics

(**Appendix A**) alongside the estimated minimum and maximum resistivity values. The overall performance is summarized in Table 2.4 for the resistivity field and Table 2.5 for the observations.

Analyzing performance for the resistivity field, the *RMSE* for ebaPEST is slightly higher than that for ES-MDA (Table 2.4) suggesting that ES-MDA may offer a marginally better fit overall. However, the negative *ME* in the ES-MDA (-14.77 Ohm·m) indicates that this method tends to underestimate the resistivity values more significantly than ebaPEST ($ME = 2.03$ Ohm·m). The *MAE* also follows a similar trend, with ebaPEST showing a slightly higher value (65.72 Ohm·m compared to 59.47 Ohm·m). The minimum and maximum resistivity values show a broader range in the ebaPEST result (40 to 885 Ohm·m) compared to the narrower range in the ES-MDA result (33 to 631 Ohm·m) considering the actual minimum of 50 Ohm·m and a maximum of 500 Ohm·m.

Table 2.4. Performance metrics for ES-MDA and ebaPEST

	<i>ME</i>	<i>MAE</i>	<i>RMSE</i>	<i>Min</i>	<i>Max</i>
ES-MDA	-14.77	59.47	90.64	33	631
ebaPEST	2.03	65.72	102.43	40	885

The ebaPEST demonstrates superior performance in the comparison of observation estimates, achieving much lower *RMSE* values for both apparent resistivity (0.12 Ohm·m) and observed resistivity (10^{-3} Ohm·m). This improved performance can be attributed to the ebaPEST tendency to minimize errors associated with observations directly during the estimation process. In contrast, ES-MDA shows significantly higher *RMSE* values, with 2.91 Ohm·m for apparent resistivity and 0.16 Ohm·m for observed resistivity. These results indicate that, while ES-MDA may capture fine-scale variations, ebaPEST is more accurate in matching both observation values.

Figure 2.10 presents the comparison between actual and estimated observation values using the ES-MDA and ebaPEST methods. The plots show two distinct

Table 2.5. Comparison of *RMSE* for apparent resistivity (*Ohm-m*) and resistivity (*Ohm-m*).

	Apparent Resistivity	Resistivity
ES-MDA	2.91	0.16
ebaPEST	0.12	0.01

types of data: apparent electrical resistivity and electrical resistivity observed from the boreholes. In both methods, the estimated apparent resistivity follows a strong linear correlation with the observed values, indicating the effectiveness of both inversion approaches. Nevertheless, the ES-MDA shows slightly greater variability in the higher resistivity range, as indicated by the wider spread of scatter points. In contrast, the ebaPEST method maintains a closer match across the resistivity spectrum. Both methods estimate the electrical resistivity reasonably well, confirming the *RMSE* values presented in Table 2.5.

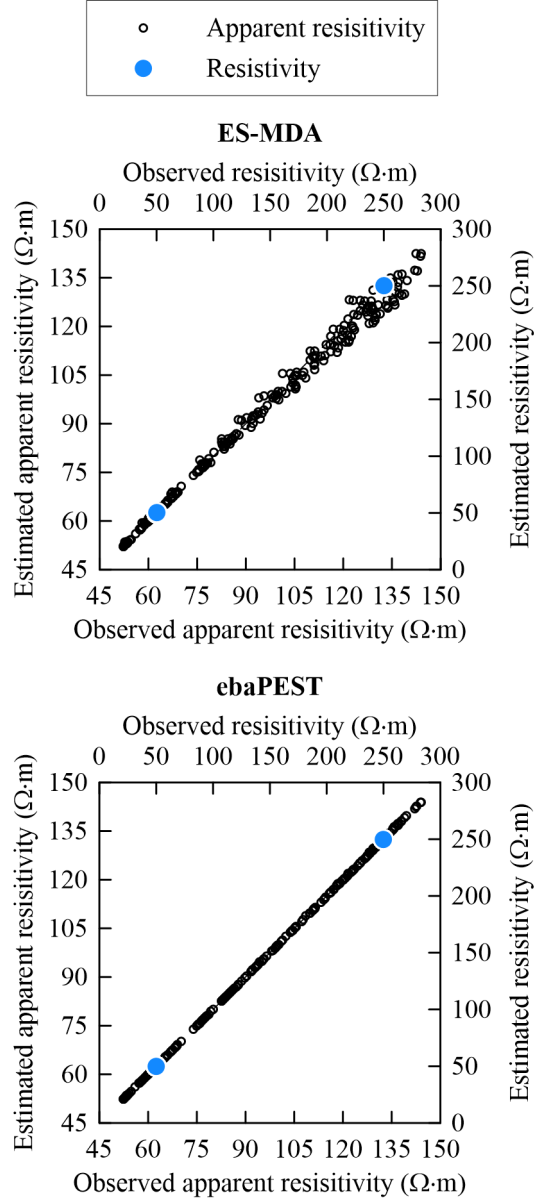


Figure 2.10. Actual vs. estimated observations using ES-MDA and ebaPEST methods. The observed apparent resistivity values are represented by circles (black), while blue dots indicate the actual resistivity (observed at monitoring points).

Figure 2.11 presents the logarithmic standard deviation maps of the resistivity

fields obtained from the two methodologies, offering insight into the uncertainty associated with these estimates and highlighting the reliability of each method within the inversion process. ES-MDA (Fig. 2.11, top) shows low standard deviation across the entire domain. Most areas have values below 0.3, with marginally higher values found at greater depths. On the other hand, ebaPEST (Fig. 2.11, bottom) shows increased standard deviation, especially at deeper layers where values approximate 0.6, indicating a more pronounced uncertainty in the calculated resistivity. The strong vertical bands of near-zero standard deviation in both cases indicate regions where observations tightly constrain the inversion, reinforcing the idea that data availability strongly influences uncertainty.

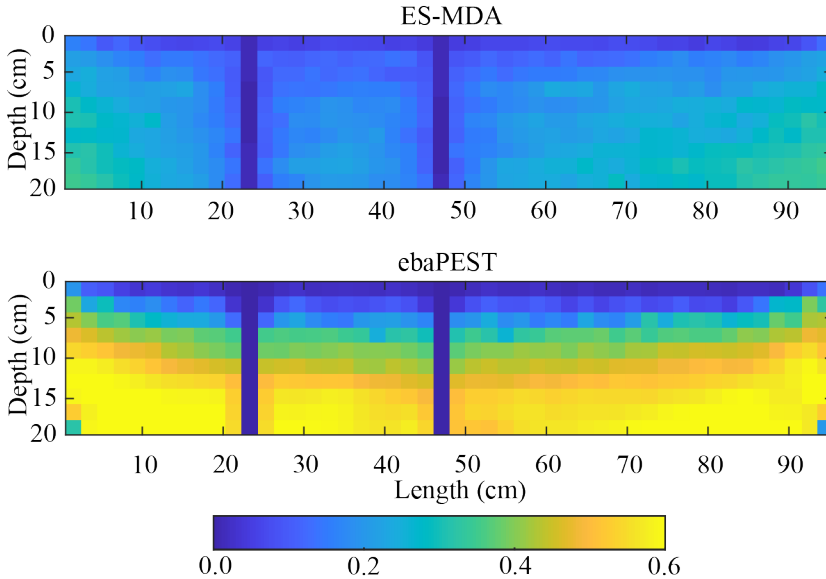


Figure 2.11. Comparison of the logarithmic estimated standard deviation using the ES-MDA vs. ebaPEST.

2.2.4 Concluding remarks

This section focused on estimating a synthetic resistivity model using apparent resistivity data and actual measurements of resistivity from the field. The primary

aim was to assess and compare the performance of two inversion techniques, the Ensemble Smoother with Multiple Data Assimilation and the Empirical Bayes PEST.

Although the approaches differ, both methods demonstrated their ability to assimilate geophysical data and reconstruct subsurface resistivity distributions with comparable accuracy, as indicated by similar error metrics. ES-MDA effectively captures block-like features, aligning better with the real resistivity distribution; however, it does not distinctly define boundaries. In contrast, ebaPEST produces smoother transitions due to the influence of its covariance function, exhibiting localized overestimations, suggesting a potential sensitivity to noise or parameterization constraints. The evaluation metrics provided further insights into the strengths and limitations of both methods. ES-MDA achieved a lower overall *RMSE*, indicating a better global fit to the reference resistivity field. However, its negative *ME* suggests an underestimation of the resistivity values. On the other hand, ebaPEST produced a broader range of estimated resistivity values, closer to the extremes of the reference field, despite a slightly higher *RMSE*. When analyzing observation data, ebaPEST demonstrated superior accuracy, achieving significantly lower *RMSE* values for both apparent resistivity and resistivity from the boreholes. This suggests that ebaPEST is more effective in matching observed values, likely due to its direct optimization strategy. ES-MDA, while capable of capturing fine-scale variations, exhibited greater deviations in observed resistivity, particularly in high apparent resistivity regions. In terms of uncertainty, standard deviation analysis shows that the ES-MDA consistently has lower uncertainty, with values remaining below 0.3. In contrast, the ebaPEST exhibits slightly higher variability, especially in deeper regions without information.

The methodologies also differ in their computational demands. ES-MDA does not require explicit and time-consuming calculation of the sensitivity matrices to solve the inverse problem, as these are incorporated into the covariance matrices of the set. For this reason, the ES-MDA computation time is approximately 2 hours, compared to about 4 hours for ebaPEST.

It is also important to emphasize the complexity of the synthetic case study presented. The subsurface consists of only three distinct blocks of material with significantly different values. Nevertheless, although the ES-MDA can better capture block-like features, it continues to struggle with managing sharp discontinuities. Moreover, these fields present challenges for approximation with traditional covariance functions, such as those employed in ebaPEST, due to the continuous structures they enforce even when structural parameters are optimized.

Potential improvements could be achieved by carefully choosing covariance functions, which would better reflect the underlying material properties, or by integrating the methods with more advanced sampling techniques, such as Markov Chain Monte Carlo (MCMC), to increase heterogeneity in the resistivity field estimation. Such refinements could help bridge the gap between the continuous nature of the estimation methods and the discrete characteristics of the subsurface materials.

2.3. Reconstruction of Contaminant Plumes through Coupled Hydrogeophysical Inversion

One major challenge in inverse modeling is the complexity of the underlying forward models, which are often computationally expensive or analytically unsolvable. Surrogate models offer a viable solution to overcome these issues (Asher et al., 2015; Jamshidi et al., 2020; Secci et al., 2022). In recent years, neural networks have emerged as a promising tool for replacing full forward models and reducing computational demand. A well-known neural network is the convolutional neural network (CNN) introduced by Lecun et al. (1998). CNNs specialize in processing grid-based data, such as images, and exhibit an inherent capacity to capture and hierarchically represent spatial features in data. Thus, CNN technology is widely employed in various fields, including groundwater spatial modeling or in geophysical context (Das et al., 2019; Lähivaara et al., 2019; Panahi et al., 2020).

The results presented in this section are based on the methodology detailed by Fagandini et al. (2024a). This work focuses on combining hydrological and ERT data to accurately estimate the spatial distribution of a contaminant within a groundwater system. By coupling the ES-MDA inverse procedure with CNN, the study aims to harness the potential of hydrogeophysics to address environmental pollution while reducing the computational cost of the inversion process. ES-MDA is applied to estimate the contaminant plume distribution using a robust geophysical forward model and assimilating both ERT data and sparse concentration measurements from monitoring points. To improve efficiency, a CNN is integrated to replace the forward model component that transforms the electrical resistivity of materials into apparent resistivity from ERT surveys. The methodology is evaluated through a two-dimensional synthetic case study simulating a tracer test in an unconfined aquifer, exploring various scenarios to assess the influence of multiple data sources and their availability.

2.3.1 Coupled Hydrogeophysical Inversion Scheme

This section outlines the proposed coupled hydrogeophysical inversion methodology based on ES-MDA (**Subsection 1.4**), which aims to estimate the spatial distribution of a tracer plume by integrating available observations, i.e. ERT data and concentration values at monitoring points. The steps of the methodology are detailed below (Fig. 2.12).

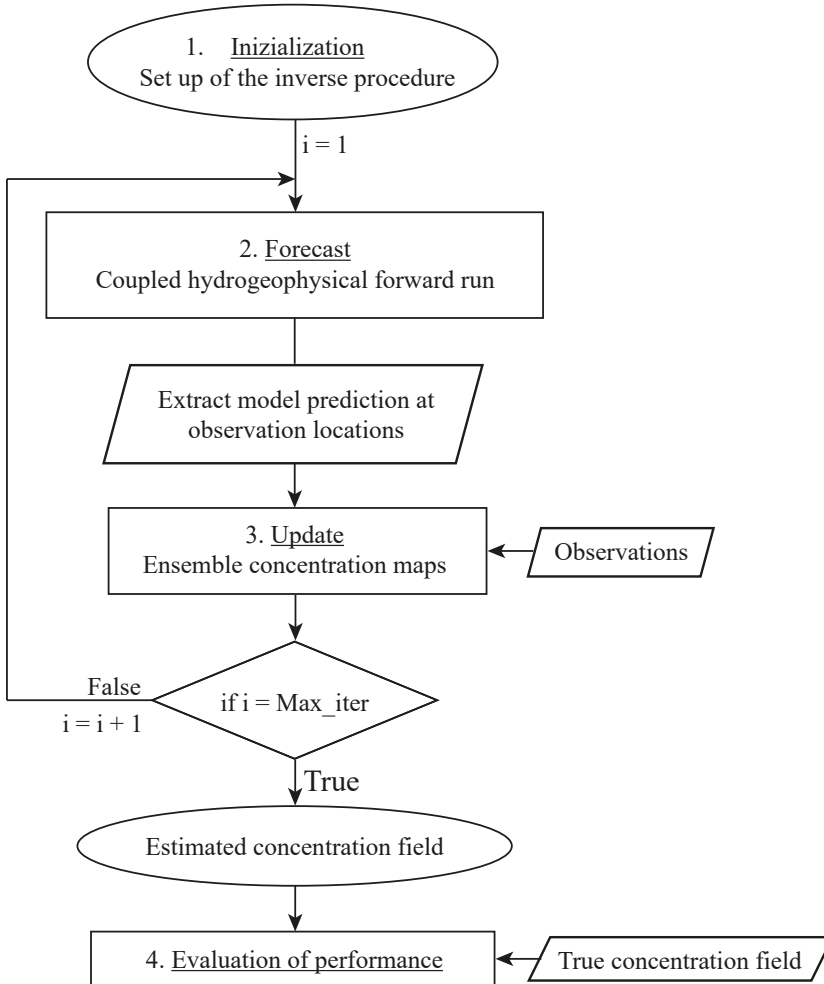


Figure 2.12. Flowchart of the coupled hydrogeophysical inverse model

Step 1. Initialization

The first step involves generating initial parameter realizations corresponding to different concentration fields. These realizations incorporate a priori information and represent the specific problem under consideration. The initial concentration maps can be generated through various approaches to comprehensively explore subsurface conditions:

1. assume homogeneity across all parameters: each realization has a distinct constant value drawn from a uniform distribution. This straightforward method is feasible when no prior information is available;
2. run stochastic sequential simulations: generate fields using a semi-variogram model. The semi-variogram can be fitted to existing field data if available, or variogram parameters can be selected randomly based on prior knowledge. This approach captures the spatial correlations in the reference dataset, ensuring the initial ensemble reflects expected concentration map patterns;
3. run a numerical transport model: generate diverse realizations by simulating injection from different locations within the model domain and varying tracer concentrations randomly selected from predefined ranges. This ensures each realization realistically represents contaminant distribution in the subsurface.

This step also defines the number of iterations N , observation errors, coefficients α_i , and the CNN training.

Step 2. Forecast: CNN-based Forward Model

For each iteration and realization, the petrophysical relationship, described in **Subsection 1.2.1**, transforms concentration maps into electrical resistivity maps. These maps are used as input in the trained CNN to compute apparent electrical resistivity values. A filtering function extracts the subset of prediction data at the observation locations.

Step 3. Update

At each iteration, the prediction vector updates the concentration map following

Eq.2.22. After updating the concentration, the next iteration begins with the updated parameter ensemble. Steps 2 and 3 are repeated until the final iteration.

Step 4. Analysis and Interpretation of Results

The results are analyzed in terms of the mean and standard deviation computed from the ensemble, associating parameter estimation with their uncertainty. If a reference solution is available, typically in synthetic case studies, a thorough comparison is made between the estimated and actual concentration values. The results are assessed using the *ME*, *MAE*, *RMSE*, and R^2 (described in **Appendix A**).

2.3.2 Synthetic Case

Set up of the Test Case

The validity of the proposed methodology is demonstrated by its application to a two-dimensional synthetic model representing the vertical cross section of a heterogeneous unconfined aquifer under fully saturated conditions, where a contaminant plume is present. This model resembles the sandbox developed at the University of Parma's Hydraulic Laboratory, extensively used in experimental and computational studies (Citarella et al., 2015; Cupola et al., 2015; Chen et al., 2018, 2021; Pereira et al., 2023; Todaro et al., 2021, 2023). Figure 2.13 offers a schematic representation of the synthetic model. It is discretized into 96 by 1 by 20 cells, each measuring 1 by 10 by 1 cm. The hydraulic conductivity varies in space with three well-defined homogeneous zones differing by two orders of magnitude (Fig. 2.13 and Table 2.6), and a uniform porosity equal to 0.37. The boundary conditions are impermeable at the bottom, phreatic surface at the top, and fixed heads at the left and right sides. This setup generates a head loss of 1 cm, inducing flow from left to right. The initial condition is zero concentration everywhere. A continuous injection of a conservative non-reactive tracer, with a concentration of 20 mg/L, is introduced into the model from a designated injector point at location ($X = 12.5$ cm, $Z = 10.5$ cm). Longitudinal and transverse dispersivity values are assumed to be 0.16 cm and 0.016 cm, respectively. The reference solution is derived from a simulation carried out using

MODFLOW (Harbaugh, 2005) and MT3DMS (Zheng and Wang, 1999) to model the groundwater flow and mass transport processes, respectively. Table 2.6 summarizes the model parameters. The simulation extends for a total duration of 3600 seconds to achieve a well-developed plume, with the concentration map at the final time step serving as the reference map. The parameters to be estimated correspond to the concentration at each model grid cell ($N_p = 1920$).

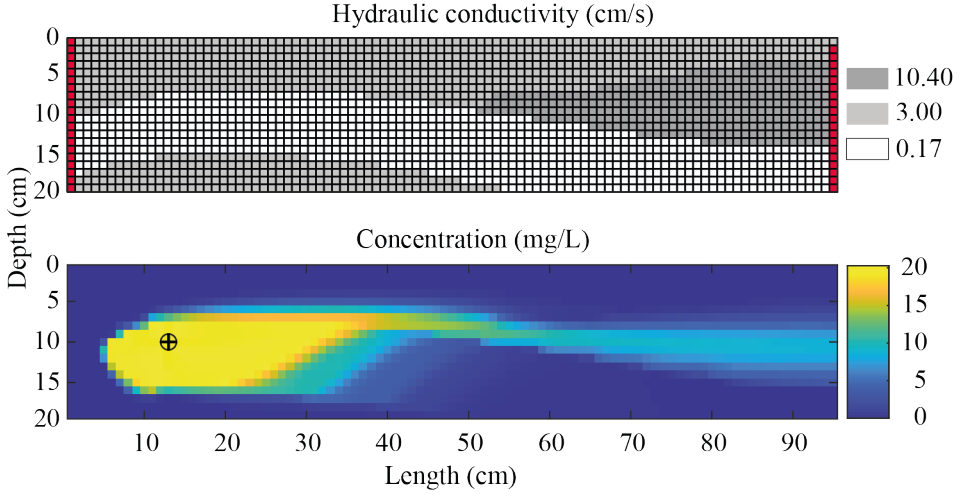


Figure 2.13. Hydraulic conductivity and concentration reference maps. The red grid cells represent the left and right boundary conditions. The cross indicates the injector position.

Table 2.6. Flow and transport model parameters.

Parameter	Value
Hydraulic Conductivity 1 (cm/s)	0.17
Hydraulic Conductivity 2 (cm/s)	3.00
Hydraulic Conductivity 3 (cm/s)	10.40
Porosity	0.37
Longitudinal dispersivity (cm)	0.16
Transverse dispersivity (cm)	0.016
Injected concentration (mg/L)	20

The reference electrical resistivity map (Fig. 2.14) is obtained by applying the

petrophysical model described in **Subsection 1.2.1**. Then, the SimPEG (Cockett et al., 2015) package processes the resulting map to derive the apparent resistivity at 225 locations, representing the observations to be used in the inverse procedure. This estimate is made solving Eq.2.10 and considering a Wenner-Schlumberger acquisition array, which consists of 32 electrodes spaced at 3 cm intervals. Table 2.7 summarizes the geophysical and petrophysical parameters used.

Table 2.7. Geophysical and petrophysical parameters.

Parameter	Value
Number of electrodes	32
Electrode spacing (cm)	3
m	1.3*
n	2
a	1
S_w	1
$\sigma_w(\mu S/cm)$	357
* (Mavko et al., 2009)	

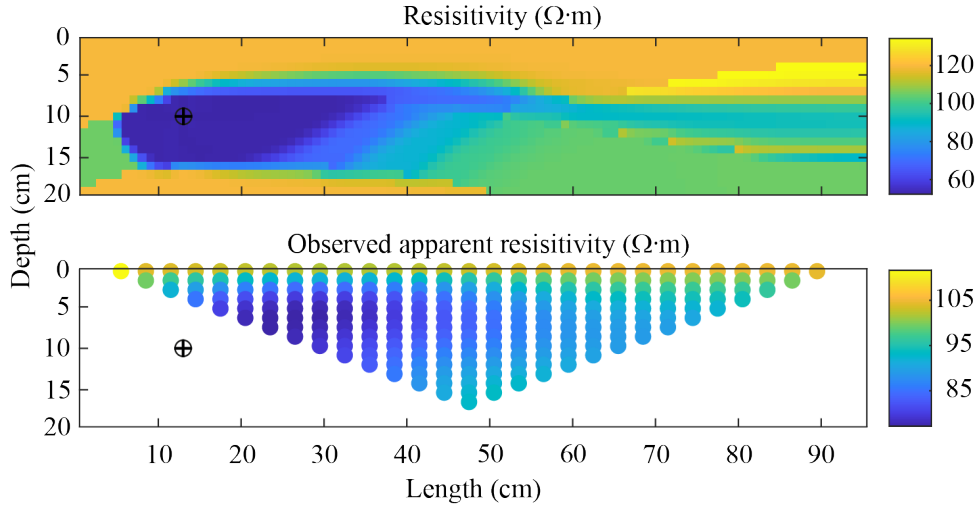


Figure 2.14. Reference resistivity model and observed pseudo-section. The cross indicates the position of the injector.

Investigated Scenarios

The idea of the work arises from the need to visualize the spread of the plume into aquifers. One possibility is the interpolation of observed concentrations in the field if they are available. Normally, these data are few and spatially sparse. Therefore, the introduction of ERT measurements, which are spatially exhaustive, is considered. In order to assess the capabilities of the proposed approach, five distinct scenarios considering different datasets are developed. Each dataset aims to emphasize the advantages of employing specific combinations of apparent resistivity (D_1) and concentration measurements (D_2).

Three monitoring wells are placed vertically at $x = 23.5, 47.5$, and 71.5 cm, each with five equidistant observation points spaced at 3 cm intervals, for a total of 15 observation points. In Scenario 1, a limited dataset, comprising only the 15 concentration values, is used to interpolate the concentration map. This map is generated using a kriging-based interpolation method, with the variogram model computed using the 15 concentration observations ($D_1 = 0, D_2 = 15$). The intent is to demonstrate how difficult it is to obtain a good estimate using a spatially sparse dataset. In the other scenarios, parameters are estimated in each cell of the model grid using the ES-MDA hydrogeophysical inversion, with the number of observations varying according to the specific case under examination. In Scenario 2 ($D_1 = 225, D_2 = 0$), only ERT data are employed. In Scenario 3 ($D_1 = 225, D_2 = 15$), the ERT data are combined with 15 concentration values. In Scenario 4 ($D_1 = 225, D_2 = 9$), the ERT data are combined with 9 concentration values. And, in Scenario 5 ($D_1 = 225, D_2 = 3$), the ERT data are combined with only 3 concentration values. A summary of the five scenarios is provided in Table 2.8.

Table 2.8. Summary of the scenarios. Number of observations used.

Scenario	1	2	3	4	5
ERT	-	225	225	225	225
Concentration	15	-	15	9	3

CNN's set up

To speed up the execution of the forward model, a CNN is implemented to replace the SimPEG package that converts electrical resistivity into apparent electrical resistivity data (i.e., pseudo-electrical resistivity sections). To train the network, a dataset including 7000 realizations obtained with SimPEG is considered. This input dataset undergoes preprocessing involving the normalization of input and output data and is then split into training (70%), validation (15%), and test (15%) sets. The CNN architecture is outlined in Table 2.9. The model is optimized using the Adam optimizer with a learning rate of 0.001, and the mean squared error between predicted and target apparent resistivity values is used as the loss function. The training is performed with a batch size of 18 over 300 epochs. After training, the model is evaluated on the validation set, and predictions are inverse transformed to the original scale. The complete CNN training and validation process took approximately 3 hours utilizing a computer equipped with an Intel i9-10920X CPU at 3.5GHz, and 32 GB RAM.

Table 2.9. CNN architecture.

Layer	Number of filters	Size of each filter	Stride	Padding	Batch Normalization	Activation	Output size
Input image	-	-	-	-	-	-	20x96x1
Convolutional	8	5x5x1	1x1	Same	True	ReLU	20x96x8
Pooling (Average)	-	2x2	2x2	0	False	None	10x48x8
Convolutional	16	5x5x8	1x1	Same	True	ReLU	10x48x16
Pooling (Average)	-	2x2	2x2	0	False	None	5x24x16
Convolutional	32	5x5x16	1x1	Same	True	ReLU	5x24x32
Convolutional	64	5x5x32	1x1	Same	True	ReLU	5x24x64
Dropout (50%)	-	-	-	-	False	None	5x24x64
Fully connected	-	-	-	-	False	Linear	1x225

Figure 2.15 reports the results of the validation set. The good agreement between the true and computed apparent resistivity is clear. The computational time of the CNN was compared to that of the 2.5D forward ERT model, showing a substantial reduction for each realization from 2.3 seconds to approximately 20 milliseconds using the computational infrastructure described above. Furthermore, to validate the reliability of the presented approach, the inverse problem in Scenario 3 is solved by using both the CNN model and the full forward model (SimPEG), comparing their performance.

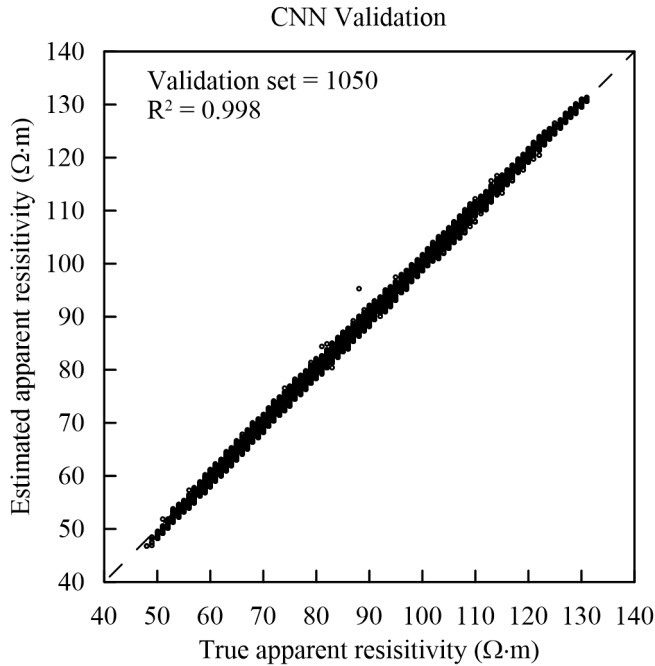


Figure 2.15. CNN Validation, the dashed line is the 1:1 line.

Inverse model set up

For Scenarios 2-5, the ES-MDA is performed with six iterations and an ensemble size of 500. In this study, the initial ensemble of parameters is generated following the approach described in **Subsection 2.3.1** (Step 1) using the Python package GeostatsPy (Pyrz et al., 2021), which interfaces the Geostatistical Software Library

(GSLIB) with Python. It is employed to generate Gaussian random fields in logarithmic space to prevent negative values, which are subsequently back-transformed into the concentration space. Each realization is based on an anisotropic exponential variogram, with the azimuth of maximum continuity randomly chosen within the range of $[80, 100]$ degrees. The mean log-concentration is randomly selected from a uniform distribution within the interval $[-2, 2]$, while the standard deviation is equal to 1.1. The correlation range in the vertical direction is randomly selected from a uniform distribution with ranges $[10, 20]$ cm, while the anisotropy ratio is sampled within the range $[7, 10]$ cm. The observation errors are normally distributed with a zero mean and variance equal to $10^{-4}(\Omega m)^2$ for the apparent electrical resistivity and $0.01 (mg/L)^2$ for the concentrations. A decreasing α set with values equal to $[364.0; 121.3; 40.4; 13.5; 4.5; 1.5]$ is used. A spatial covariance localization is applied considering a space correlation length of 30 cm. A covariance inflation is applied with a factor equal to 1.01 (refer to Todaro et al. (2023) for a detailed explanation of the ES-MDA set up).

2.3.3 Results

The comparative analysis of the five scenarios offers insights into the efficacy of ES-MDA in estimating the distribution and values of the concentration plume for a given release history. The results, shown in Fig. 2.16, present the estimated concentration for Scenarios 2-5, depicted by the ensemble mean. Table 2.10 provides the evaluation metrics, alongside the maximum estimated concentration for comparison with the actual value of 20 mg/L. It also includes an assessment of the estimate uncertainty indicated by the standard deviation.

In the first scenario (Fig. 2.16.a), the concentration map is obtained through kriging interpolation using 15 concentration values, providing a baseline for performance evaluation. Scenario 2 (Fig. 2.16.b) uses ES-MDA with only apparent resistivity as observations. The results exhibit poorer accuracy in estimating the concentration map compared to the previous one. Although the *RMSE* of 3.69 mg/L is comparable

to that of Scenario 1, there is a significant overestimation of the injected concentration, resulting in a higher ME (-0.48 mg/L) and MAE (2.64 mg/L). Particularly, the maximum estimated concentration reaches around 32 mg/L, whereas the actual concentration is 20 mg/L. This highlights the importance of incorporating concentration data for a more robust estimation. In Scenario 3 (Fig. 2.16.c), which combines apparent resistivity data and 15 concentration values, the integrated approach outperforms the previous scenarios with $ME = 0.06$ mg/L; $MAE = 1.56$ mg/L; $RMSE = 2.74$ mg/L; $R^2 = 0.82$, and the maximum estimated concentration is 22 mg/L, close to the actual injected value. This combination improves the model's ability to capture plume distribution. Scenario 4 (Fig. 2.16.d), similar to Scenario 3 but with only 9 concentration observations, reveals a trade-off between data quantity and model accuracy. Although the reduction in concentration data slightly affects accuracy, the overall performance remains good ($ME = -0.01$ mg/L; $MAE = 1.97$ mg/L; $RMSE = 3.01$ mg/L; $R^2 = 0.79$). The limitations of the sparse concentration information become more pronounced in the final scenario (Fig. 2.16.e), where only three concentration data points are used in conjunction with apparent resistivity data. Despite the model's adaptability, the reduced dataset compromises the accuracy of the estimated concentration map, as indicated by $ME = -0.38$ mg/L; $MAE = 2.21$ mg/L; $RMSE = 3.17$ mg/L; Max Concentration = 28.62 mg/L; $R^2 = 0.76$.

Fig. 2.17 shows the scatter plot between true and estimated concentrations for all scenarios, indicating dispersion among data points and overall satisfactory predictive performance, as shown by good R^2 values (Table 2.10). The inversion procedure captures a significant portion of the variation in the true concentration field, though some limits are identified. Comparing the true contaminant distribution (Fig. 2.13) with the estimated one (Fig. 2.16a) in Scenario 1, overestimation occurs upstream of the source due to kriging interpolation. In Scenarios 2 to 5, the procedure better estimates lower values but presents a large uncertainty on the maximum concentration. Adding concentration information in Scenario 3 improves the results significantly. After a thorough examination of the results and associated metrics, the third sce-

nario, utilizing both apparent resistivity data and concentration values, emerges as the optimal configuration in terms of estimation accuracy and pattern distribution. This comprehensive evaluation underscores the importance of integrating diverse datasets in hydrological studies to achieve more accurate and reliable estimations of contaminant plume distribution.

Table 2.10. Performance of the proposed approach assessed on the estimation of the contaminant plume (mg/L) for each scenario.

Scenario	<i>ME</i>	<i>MAE</i>	<i>RMSE</i>	<i>Max Conc.</i>
1	-0.08	1.99	3.69	22.34
2	-0.48	2.64	3.69	32.59
3	0.06	1.56	2.74	22.00
4	-0.01	1.97	3.01	24.04
5	-0.38	2.21	3.17	28.62

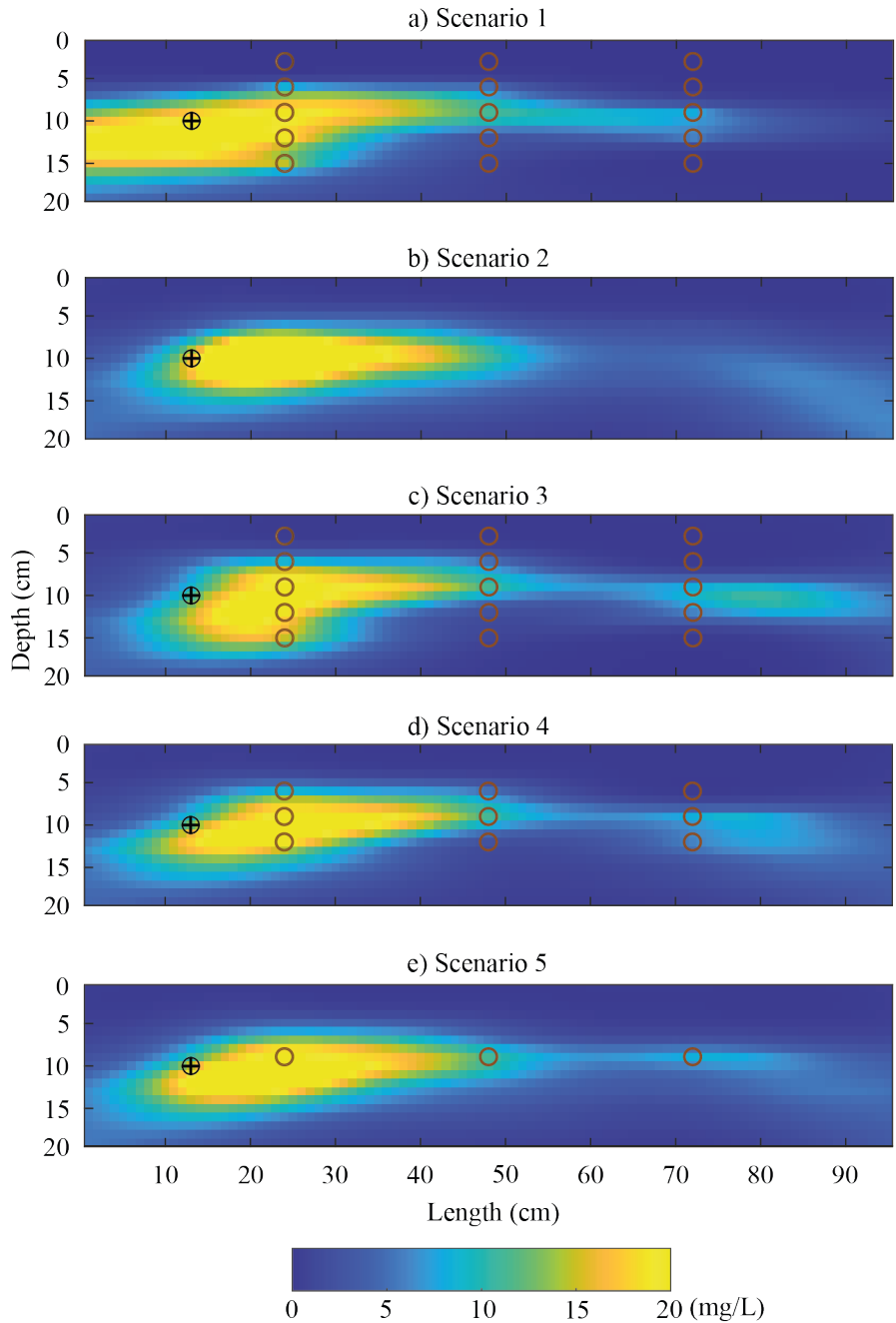


Figure 2.16. Estimated concentration map distribution for the five scenarios. The injector is marked by a cross. Red circles represent the locations of observation wells.

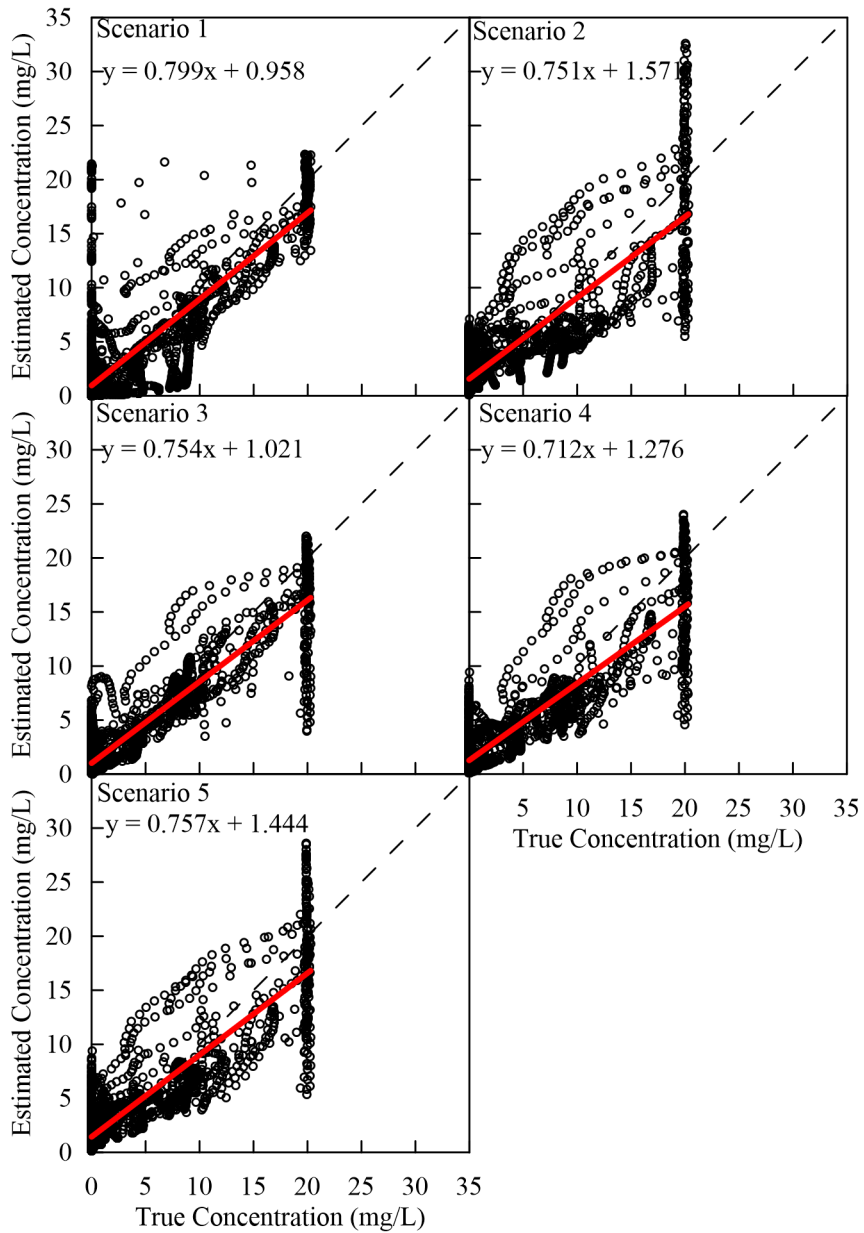


Figure 2.17. Scatter plot of the estimated vs. observed concentration for all scenarios.

Fig. 2.18 represents the agreement between the observed values and the corre-

sponding predictions for Scenarios 2 to 5, with a 45° line indicating a perfect fit. The proximity of data points to this line signifies model accuracy. Scenario 1 points align along the 45° line due to kriging being an exact estimator.

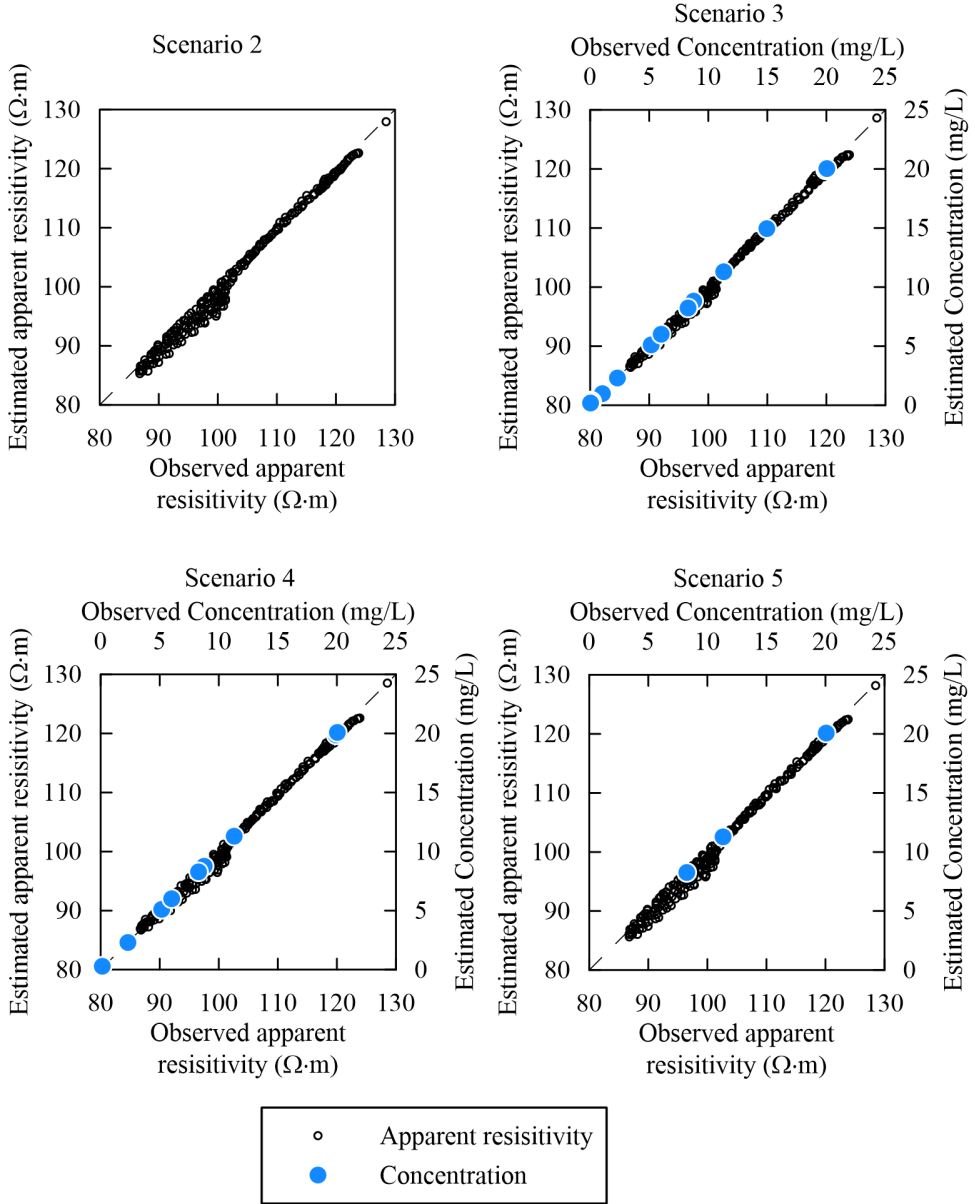


Figure 2.18. Observed-Estimated apparent resistivity and concentration for scenarios 2-5.

Assessing the uncertainty in the estimated concentration map is crucial for a comprehensive understanding of the reliability and robustness of the proposed ap-

proach. Figure 2.19 and Table 2.10 show the standard deviation maps for the five scenarios. Scenario 1, which employs kriging, exhibits a standard deviation of zero at observation points. The standard deviation increases with distance from these points, peaking at 25.82 mg/L at the model's boundaries. The average standard deviation across the entire map is 12.51 mg/L. The remaining scenarios compute standard deviation from the ensemble of concentration maps. Scenarios 2 and 5 show a high standard deviation near the source due to a lack of concentration values. Scenario 2 exhibits a mean value of 2.42 mg/L and a maximum value of 13.62 mg/L. These values are comparable to those in Scenario 5, which has a mean value of 2.10 mg/L and a maximum value of 13.21 mg/L. Scenario 3 demonstrates the smallest standard deviation, with an average value of 1.16 mg/L and a maximum value of 6.18 mg/L. Similarly, Scenario 4 has mean and maximum values of 1.43 mg/L and 6.59 mg/L, respectively, closely matching those of Scenario 3. For ease of comparison, the color bar for all scenarios is standardized, limited to 15 mg/L to optimize the display for Scenarios 3 and 4. Notably, the maximum standard deviation in Scenario 1 exceeds 25 mg/L, whereas Scenarios 3 and 4 remain below 7 mg/L. This discrepancy is attributed to the significantly higher values in the border area of Scenario 1, where no concentration information was available. Scenario 3 has the smallest standard deviation, indicating the best performance. These results confirm that combining ERT and concentration data provides reliable estimation of concentration distribution in an aquifer. While more concentration data improves results, even three observations yield acceptable outcomes.

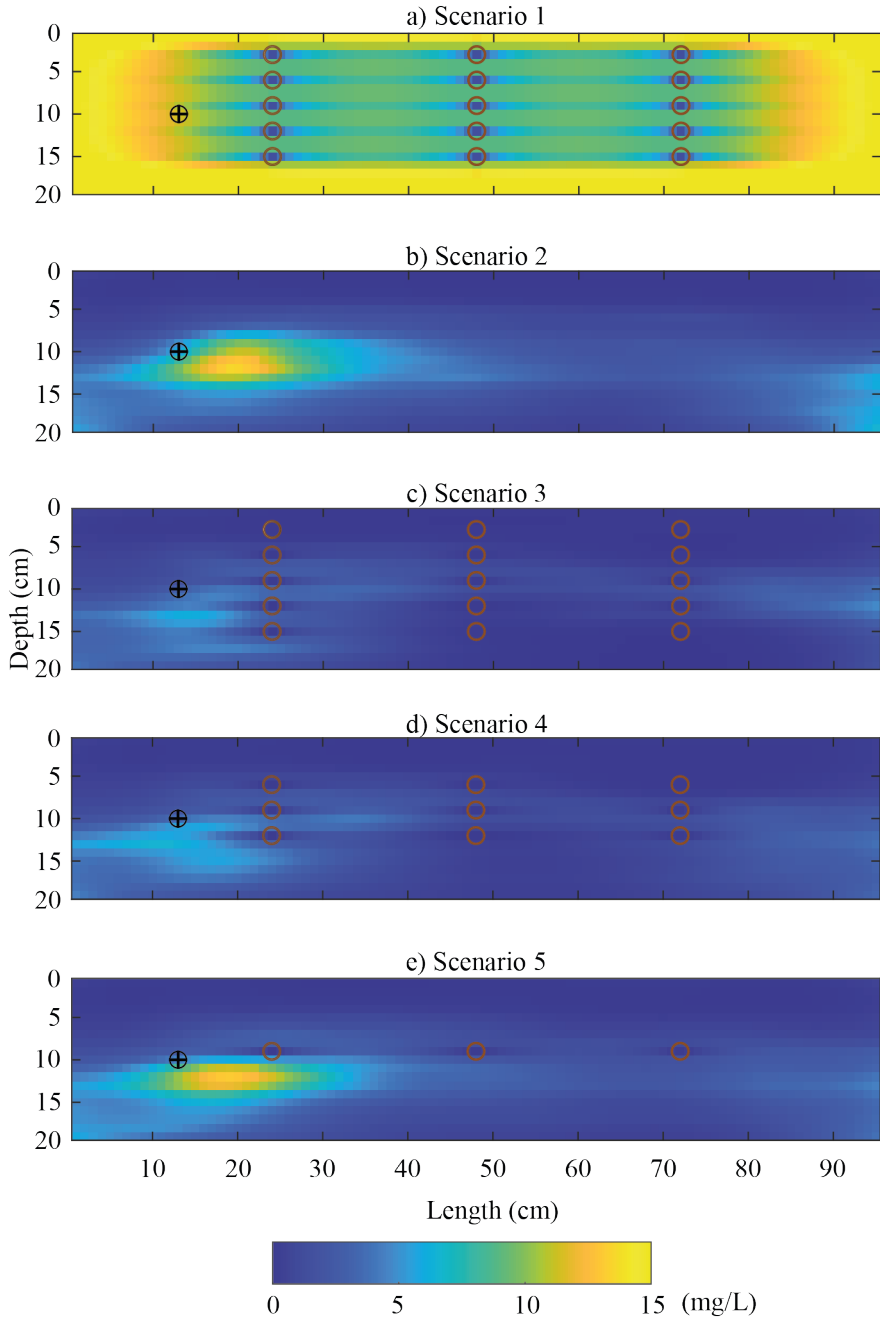


Figure 2.19. Standard deviation maps of the estimated concentrations for the five scenarios. The cross denotes the injector. The observation wells are visually depicted by the red circles.

Full Forward Model vs CNN

The validity of the proposed inversion approach is further investigated by solving Scenario 3 using the full forward model (SimPEG) instead of the CNN. Figure 2.20a shows the estimated plume using SimPEG, and Fig. 2.20b illustrates differences between the two approaches. Differences are negligible except for a small area beneath the source location. Both models demonstrate comparable performance across several metrics, with the CNN model slightly outperforming the full forward model in some areas. Specifically, the CNN model achieves a lower $RMSE$ (2.75 mg/L) compared to the full forward model (2.93 mg/L), and a higher R^2 value (0.83 vs. 0.81). When examining ME and MAE metrics, which gauge the average magnitude of prediction errors, the full forward model records an ME of 0.32 mg/L and an MAE of 1.68 mg/L, while the CNN model exhibits an ME of 0.07 mg/L and an MAE of 1.56 mg/L. In terms of mean and maximum standard deviation, the CNN model again shows a slight advantage with lower values for both mean standard deviation (1.11 mg/L vs 1.68 mg/L) and maximum standard deviation (11.39 mg/L vs. 16.41 mg/L). Finally, both models exhibit similar maximum concentration values, with the CNN model displaying a slightly higher value at 22.35 mg/L compared to 20.37 mg/L for the full forward model. Notably, the SimPEG model takes approximately 2 hours to run, whereas the CNN completes the task in about 5 minutes on an Intel i9-10920X 3.5GHz system with 32 GB RAM.

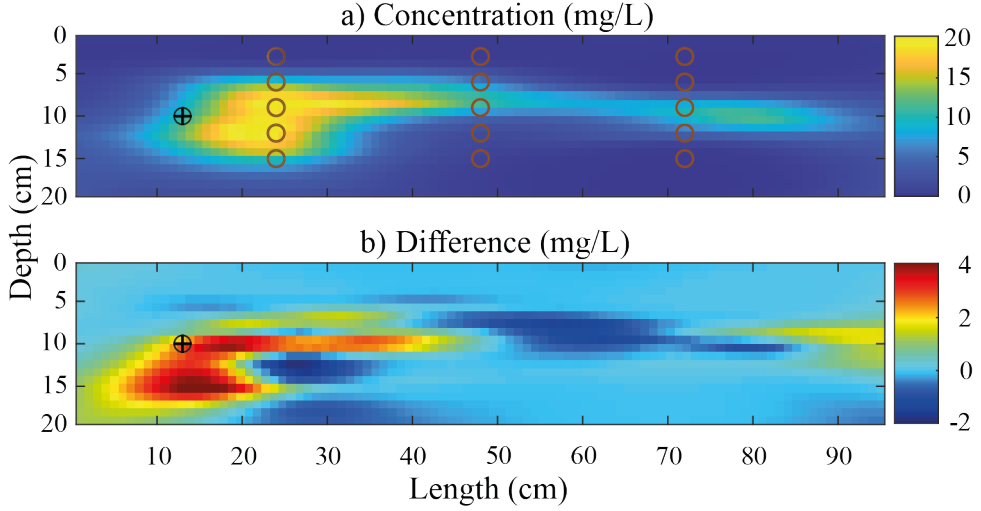


Figure 2.20. a) Estimated concentration distribution, Scenario 3 – SimPEG forward model. The injector is marked by a cross. Red circles represent the locations of observation wells. b) Differences between estimated concentrations (CNN vs SimPEG forward model).

2.3.4 Concluding remarks

This work investigates the effectiveness of the Ensemble Smoother with Multiple Data Assimilation (ES-MDA) model in addressing the complex challenge of accurately estimating the spatial distribution of a concentration plume. This is achieved through the simultaneous assimilation of observed electrical resistivity tomography (ERT) data and scattered concentration values from monitoring wells. A distinguishing feature of this approach is the integration of a convolutional neural network (CNN) to accelerate the forward model.

The study compares five different datasets to evaluate the performance of the proposed approach. These scenarios allow for a thorough examination of the advantages of combining data from multiple sources (Linde and Doetsch, 2016), highlighting the effects of different observation datasets on the accuracy of plume distribution assessments. The first scenario employs a kriging-based approach to interpolate 15 concentration values, while subsequent scenarios evaluate the capability of the proposed inverse hydrogeophysical approach. The second scenario uses only apparent

resistivity data as observations in the ES-MDA; the third to fifth scenarios combine apparent resistivity data with different subsets of concentration values: 15, 9, and 3, respectively. The third scenario, which combines apparent resistivity with 15 concentration values, emerges as the most promising configuration in terms of accuracy and precision. The least accurate estimates are observed in the cases of kriging interpolation (Scenario 1) and ES-MDA utilizing only apparent resistivity data (Scenario 2). A pertinent point to mention, based on the comparison of these results, is the inherent difficulty in relying solely on 15 concentration values derived from a survey for interpolation purposes. This challenge becomes even more pronounced with the use of 9 or 3 values, which are insufficient for constructing the variogram in the case of kriging. These findings suggest that such a limited dataset may not provide sufficient information to accurately capture the spatial variability of subsurface concentration maps, emphasizing the importance of combining multiple data sources.

Additionally, the comparison between the full ERT forward model (i.e., SimPEG) and the CNN showcases significant enhancements in computational efficiency using the surrogate model while maintaining robust predictive performance. The overall results demonstrate the efficacy of the proposed inverse methodology in accurately capturing and predicting the plume concentration's distribution and values, providing a quick tool for supporting optimal strategies for contaminated site remediation.

Considering the factors that influence the accuracy of the results, one must keep in mind the petrophysical relationships that play a key role in determining the reliability of concentration estimates. These models may face some uncertainties that might impact the inversion results (Linde et al., 2017). Furthermore, in this work, simplifications have been made in the geophysical properties of the electrical model that must be considered in real cases.

3

Experimental Sandbox Investigation Using Electrical Resistivity Tomography and Tracer Tests for Subsurface Characterization

Aquifers play a fundamental role in water supply, making it essential to safeguard their quality and quantity. When these crucial resources are threatened, preservation or restoration actions become critical to ensuring their continued existence. In this context, accurately determining aquifer properties and understanding their spatial variability are among the most significant challenges in groundwater flow and transport modeling. The complexity and extent of an aquifer are directly influenced by the properties of the porous medium through which water moves.

A major challenge in groundwater hydrology is the scarcity of experimental data in controlled flow environments. To better comprehend the factors influencing aquifer behavior, researchers often rely on laboratory-based simulations. Controlled

laboratory conditions allow for the systematic study of groundwater flow and transport under replicable conditions (Ursino et al., 2001; Sternberg, 2004; Maina et al., 2018). Recently, experimental setups such as sandboxes have gained popularity for examining two-dimensional flow and transport in heterogeneous porous media (Jose et al., 2004; Berg and Illman, 2011; Citarella et al., 2015; Cupola et al., 2015; Chen et al., 2018; Todaro et al., 2021; Zhao et al., 2022; Todaro et al., 2023). These laboratory models provide well-characterized environments with clearly defined boundary conditions, enabling precise measurements of hydraulic and transport parameters.

In recent years, the integration of geophysical techniques into laboratory setups, such as sandboxes, has further enhanced the study of aquifers. Techniques like electrical resistivity tomography (ERT) and ground-penetrating radar (GPR) are now frequently incorporated into sandbox experiments, offering non-invasive methods to monitor subsurface processes in real-time (Irving and Singha, 2010; Pollock and Cirpka, 2012; Kang et al., 2018, 2019; Chen et al., 2023). This integration allows for the observation of spatial and temporal changes in flow and transport, providing a more detailed and accurate representation of aquifer heterogeneity. By combining traditional hydraulic measurements with geophysical monitoring, researchers can significantly improve the reliability of groundwater models and gain deeper insights into the behavior of aquifers under controlled yet realistic conditions.

The main objective of this study is to acquire both ERT and concentration measurements to identify the hydraulic conductivity field of a sandbox using the fully coupled hydrogeophysical approach proposed by Fagandini et al. (2024a). This work represents the first application of this procedure to laboratory data, marking an important step in advancing the integration of geophysical and hydrological methods for enhanced subsurface characterization. None of the previously cited works have combined the use of ERT and concentration data to jointly estimate hydraulic and geophysical parameters using the Ensemble Smoother with Multiple Data Assimilation (ES-MDA) based on laboratory measurements. It is also worth emphasizing that this study employs a non-commercial ERT system, tailored specifically for lab-

oratory experiments. Before applying the approach to laboratory data, the method was first validated using synthetic cases (see **Chapter 2**). The successful application of ES-MDA in the synthetic scenario provided confidence in the technique's ability and therefore allowed its use with real experimental data.

1. Experimental installation

During 2007, at the hydraulic laboratory of the Civil and Environmental Engineering Department of the University of Parma, a two-dimensional aquifer model on the vertical plane was set up. The model represents a phreatic aquifer aimed at reproducing different phenomena for teaching purposes.

The experimental tests were conducted to replicate field conditions and assess the hydrogeological properties of the system. The setup of the device involved the use of the sandbox model to collect data through tracer tests and electrical resistivity tomography (ERT). Figure 3.1 provides an overview of laboratory equipment utilized during the experiments.

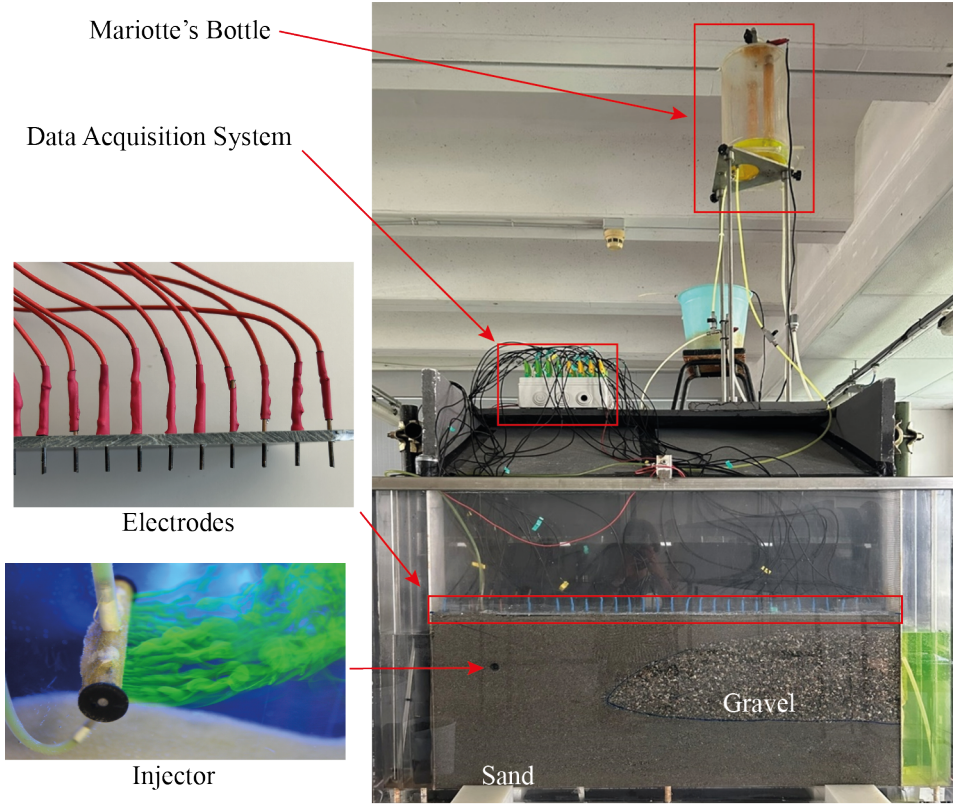


Figure 3.1. Overview of the laboratory equipment.

The sandbox, constructed from transparent polymethylmethacrylate (PMMA), consists of side panels with a thickness of 2 cm and a bottom panel of 3 cm. With dimensions of 120 cm in width, 73 cm in height, and 14 cm in thickness. The sandbox is divided into three sections: upstream and downstream tanks and an internal chamber measuring 95 cm x 10 cm x 70 cm. The central chamber is filled with porous media, consisting of 40 cm of fine sand along with a mixed gravel lens. To prevent air entrapment and minimize material segregation, the porous media are packed layer by layer under fully saturated conditions. The tanks and the porous medium are separated using perforated iron plates with brass wire gauze. Hydraulic properties, including porosity and hydraulic conductivity of both sand and gravel, are deter-

mined through laboratory experiments (refer to Table 3.1 for details). Porosity is measured using standard gravimetric methods, where the dry weight and saturated weight of the materials are used to calculate void ratios. Hydraulic conductivity is measured using a constant-head permeameter. A tank and a centrifugal pump located beneath the sandbox facilitate water circulation, allowing water to be discharged into the tank or directly into the sewage system. The resistivity values of both materials are presented in Table 3.1. These values are calculated using Archie’s law (Eq.2.16), based on the porosity of the two materials. The calculations assume a cementation exponent $m = 1.3$, a saturation exponent $n = 2$, and a formation water salinity $S_w = 0.95$, with a measured water conductivity $\sigma_w = 400 \mu S/cm$.

Table 3.1. Properties of sand and gravel used in the laboratory experiment.

	Sand	Gravel
Diameters (mm)	1–4	4–6
Porosity (-)	0.36	0.42
Hydraulic conductivity (cm/s)	0.01	6.56
Resistivity (Ohm·m)	106	87

2. Data acquisition system

Two methods—ERT and tracer testing—are employed to capture spatial and temporal variations within the laboratory sandbox. To achieve reliable and high-resolution measurements, a custom-built data acquisition system is designed specifically for these experiments. In the following subsections, the specifics of the ERT setup and tracer concentration measurement tests are outlined.

2.1. Electrical Resistivity Tomography (ERT)

ERT is a geophysical imaging technique used to investigate subsurface properties by measuring the electrical resistivity of the ground. The method involves injecting

electrical current into the ground via a pair of electrodes and measuring the voltage difference using other electrode pairs. A detailed resistivity profile or image of the subsurface is generated by systematically moving these electrodes across the survey line. For further details, see **Chapter 2 Subsection 1.2**.

For the sandbox experiments, a linear array of 32 copper electrodes, spaced 3 cm apart, is placed within the outermost unsaturated sand layer. Each electrode sinks 1 cm into the sand to ensure proper contact. The ERT data acquisition system (Delmonte et al., 2024), custom-developed at the University of Parma (Fig. 3.2), integrates an Arduino Due microcontroller, a signal conditioning shield to convert analog inputs into digital form, and multiplexing shield boards designed to select any quadripole configuration from the 32-electrode array. This setup offers several advantages, including reduced cost, weight, and volume, largely due to the lower current requirements of laboratory-scale experiments. The system’s compact design and affordability make it particularly well-suited for use in controlled laboratory settings. Additionally, the flexibility to modify the firmware allows users to implement various algorithms to control the sequence of ERT measurements, providing a significant edge over conventional commercial systems. The programming process begins by initializing key parameters such as the number of measurements, repetitions, and sampling frequency. After configuring the necessary pins, the electrode sequences (C1, C2, P1, and P2) are defined. During the experiment, the multiplexers activate the system and bias the current electrodes (C1 and C2). A brief delay is introduced to allow the system to stabilize and minimize transient effects from inductive or capacitive interactions. Voltages from the electrodes are then recorded using the analog-to-digital converter. To mitigate electrochemical interference, the biasing of C1 and C2 is reversed. The potential values are stored in a text file and displayed on the serial monitor in real time. After each cycle, the acquisition process repeats. (The complete firmware script is provided in **Appendix B**). The system operates on a 24V DC power supply, with data transfer and post-processing facilitated through a USB connection to a PC.

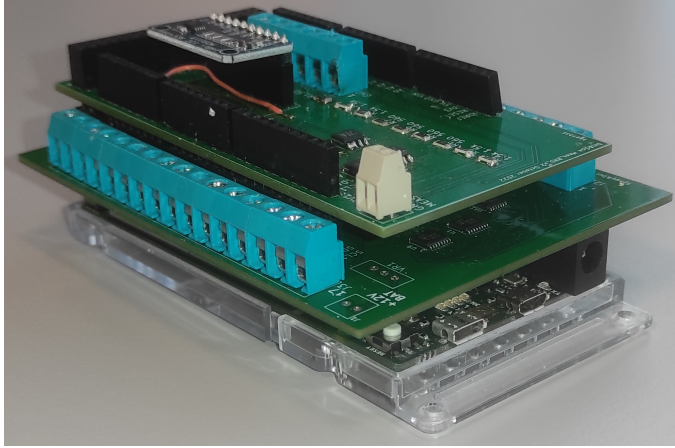


Figure 3.2. ERT acquisition device.

Moreover, a MATLAB script is developed to process the raw ERT data. Once the key parameters of the firmware are loaded, the script handles data correction by identifying outliers and averaging the data to derive accurate resistivity values. It also calculates the variance to evaluate measurement reliability. Using the geometric factor K (Eq. 2.15), the script computes the apparent resistivity (Eq. 2.14). Finally, it generates visual representations of the results, including scatter plots and pseudo-section maps. The post-processing script is detailed in **Appendix C** for further reference.

Two applications of ERT within the sandbox are presented, both employing the Wenner-Schlumberger array configuration. In this arrangement, the current electrodes are placed at the ends of the electrode line, while the potential electrodes are positioned between them. By maintaining a constant spacing between the potential electrodes and systematically increasing the distance between the current and potential electrodes, this array enables greater depth penetration into the subsurface. The input is incrementally transmitted to the array along the sandbox to obtain a complete pseudo-section profile, capturing 225-point measurements.

The first application measured the resistivity in the homogeneous sand and gravel system shown in Fig. 3.1. The average resistance values and their corresponding

standard deviation are presented in Fig. 3.3, with the electrode positions marked as a reference. Notably, the highest standard deviation is observed in the region containing gravel, reflecting the increased variability in resistivity in that area.

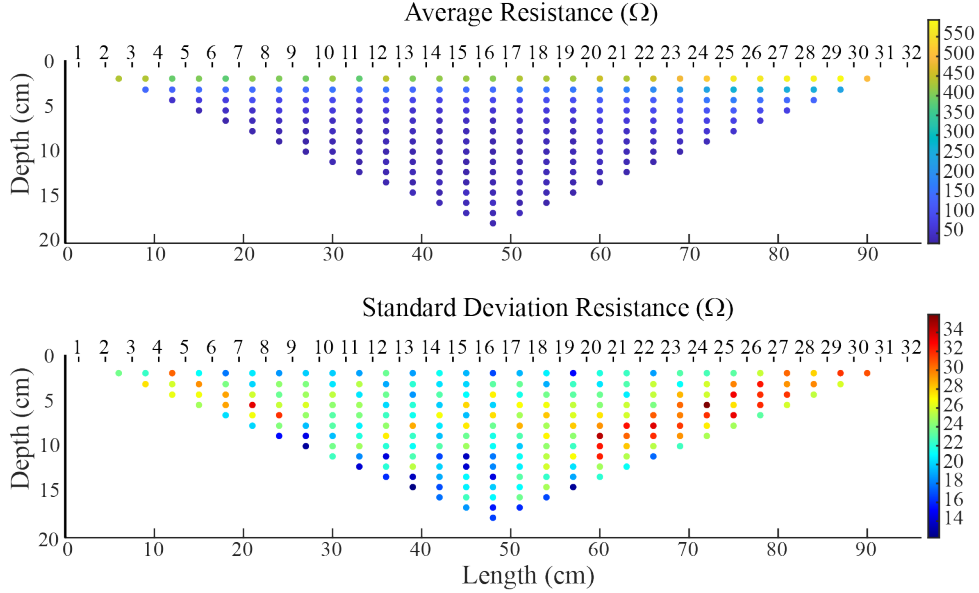


Figure 3.3. Average resistance (above). Standard deviation resistance (bottom).

The apparent resistivity is then computed using the post-processing MATLAB script, assigning the values to their respective points in the sandbox. To enhance visualization of the resistivity distribution, point measurements are interpolated, generating a pseudo-section that provides a clearer view of resistivity variations. Fig. 3.4 presents the scatter plot of apparent resistivity (top) and the interpolated pseudo-section (bottom), with the grayscale background representing the materials in the sandbox (light gray for sand, dark gray for gravel). Analysis of the measured data revealed that the instrument detected two distinct materials. However, the resistivity values for sand and gravel did not align with those expected under ideal, fully saturated conditions, as described in the literature (Table 3.1). According to Archie’s law, sand should have a higher resistivity than gravel. In contrast to

this, the measurements showed higher resistivity values for the gravel. Currently, no comparative data from other instruments is available to determine whether this anomaly is valid or due to experimental limitations. This discrepancy could be due to inaccuracies in the resistivity measurements, potentially caused by small-scale setup inconsistencies (e.g., low current levels, unstable cable connections, metal plates inducing preferential current paths, or boundary effects) or environmental factors (e.g., external noise or air micro-bubbles in the materials).

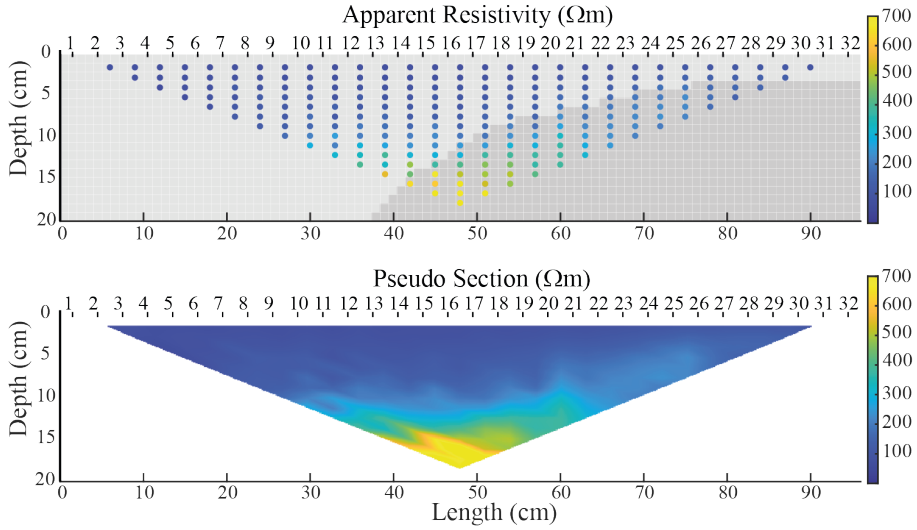


Figure 3.4. Scatter plot of apparent resistivity (top) and interpolated pseudo-section of apparent resistivity (bottom).

For the second test case, the sandbox is filled with 40 cm of fine sand (Table 3.1) and to simulate a material contrast, a 1 cm thick impermeable, non-conductive plastic barrier is introduced between electrodes 22 and 23. The experimental setup is shown in Fig. 3.5, with the red rectangle highlighting the area investigated by the ERT.

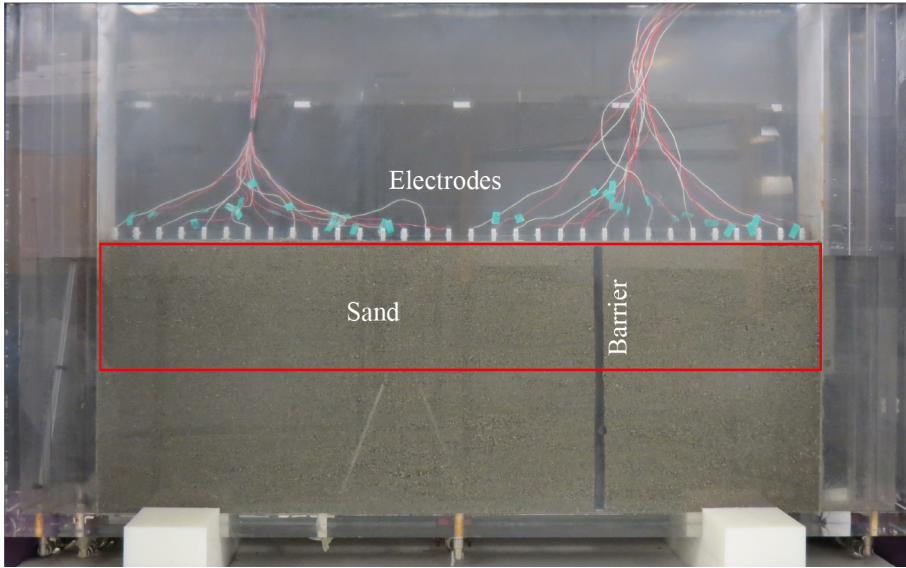


Figure 3.5. Sandbox filled with homogeneous sand with an impermeable, non-conductive barrier. The red rectangle identifies the ERT investigation area.

Fig. 3.6 shows the resistance measured for the second test case. The significantly higher resistance between electrodes 22 and 23, where the barrier is located, is clearly evident. Similar results are observed in terms of apparent resistivity (Fig. 3.6, bottom), with the background indicating the actual materials in the sandbox: light gray for sand and dark gray for the barrier.

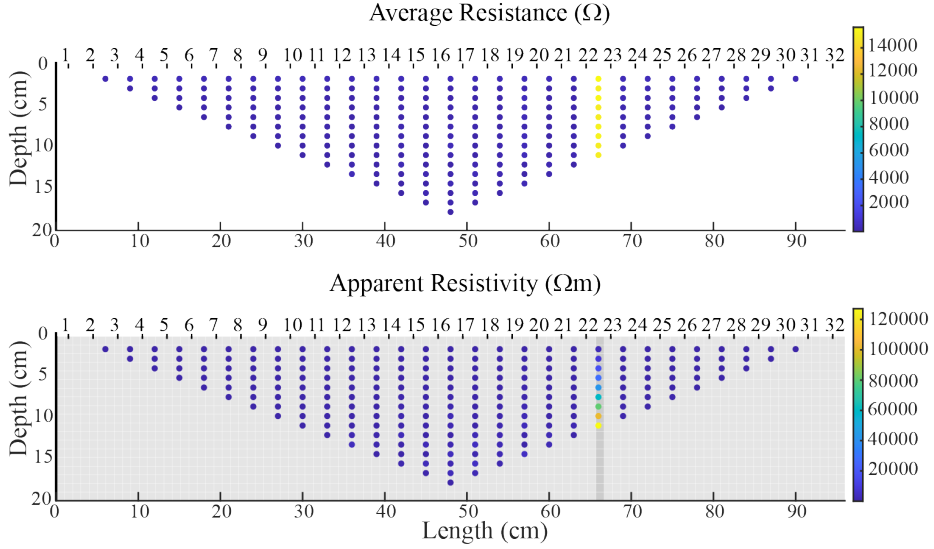


Figure 3.6. Scatter plot of observed resistance (top) and apparent resistivity (bottom).

2.2. Concentration

A tracer test is conducted to analyze flow patterns and transport dynamics within the sandbox. Constant water levels are maintained at both the upstream (38.6 cm) and downstream (29.9 cm) ends to control flow. Fluorescein sodium salt, a non-toxic tracer, is injected at a concentration of 40 mg/L at a constant rate of 4 mL/s. The tracer injection point is located upstream at coordinates $X = 13$ cm and $Z = 12$ cm, covering the full thickness of the sandbox. The injector is supplied with a pipe (1 mm diameter) connected to a Mariotte's bottle that maintains a constant hydraulic head, ensuring a stable injection rate throughout the test. Tracer visualization is achieved using a photographic acquisition technique developed by Citarella et al. (2015). Behind the sandbox, a custom-built dark room is positioned to optimize contrast and minimize light interference during the experiments. A digital camera, placed outside the dark room, is programmed to capture images at regular intervals. The camera is connected to a computer that runs a LabVIEW script that controls the time of image capture. To enhance the visibility of the tracer, the fluorescein

sodium is excited using blue light LEDs placed around the sandbox. The fluorescent properties of fluorescein, activated by blue light, allowed the tracer to glow distinctly, facilitating its detection through digital images. The images are captured every 5 seconds, creating a time-lapse dataset of the tracer's movement. Initially, the concentration of the tracer is set to zero across the entire sandbox, ensuring that the test started under known conditions. The experiment lasted approximately 1000 seconds, with tracer injection beginning 40 seconds after the start of the data acquisition. Throughout the test, the movement of the tracer through the porous media is continuously monitored. Image processing techniques are used to analyze the captured data. Specifically, each pixel's luminosity, which correlates with the concentration of fluorescein, is converted into concentration values. This conversion process is executed using a custom MATLAB script, which allowed for precise quantification of the tracer concentration at different points in the sandbox over time. For additional details on the procedure, see Citarella et al. (2015) and Cupola et al. (2015). Figure 3.7 shows the reference field traversed by the tracer.



Figure 3.7. Close-up view of the field sandbox.

Figure 3.8 presents a series of spatial concentration maps at four time steps: 100 s, 250 s, 450 s, and 900 s. These snapshots illustrate the evolution of the tracer plume within the sandbox, with increasing spread and intensity over time. At early times ($T = 100$ s), the plume is barely visible. By $T = 250$ s, the tracer began to propagate, reaching the interface between the sand and the gravel lens.

The plume continues to spread and intensify by $T = 450$ s. By $T = 900$ s, close to the end of the experiment, the plume has almost reached its full extent. The red circles indicate the positions where the concentration data are collected, which correspond to the locations used for the breakthrough curves in Fig. 3.9. Figure 3.9 shows the normalized luminosity over time for the four monitoring points (M1, M2, M3, and M4). The curves exhibit a distinct time lag at each monitoring point influenced by varying hydraulic conductivity and porosity for each material. Once the concentration tracer stabilizes, the curves reach a plateau.

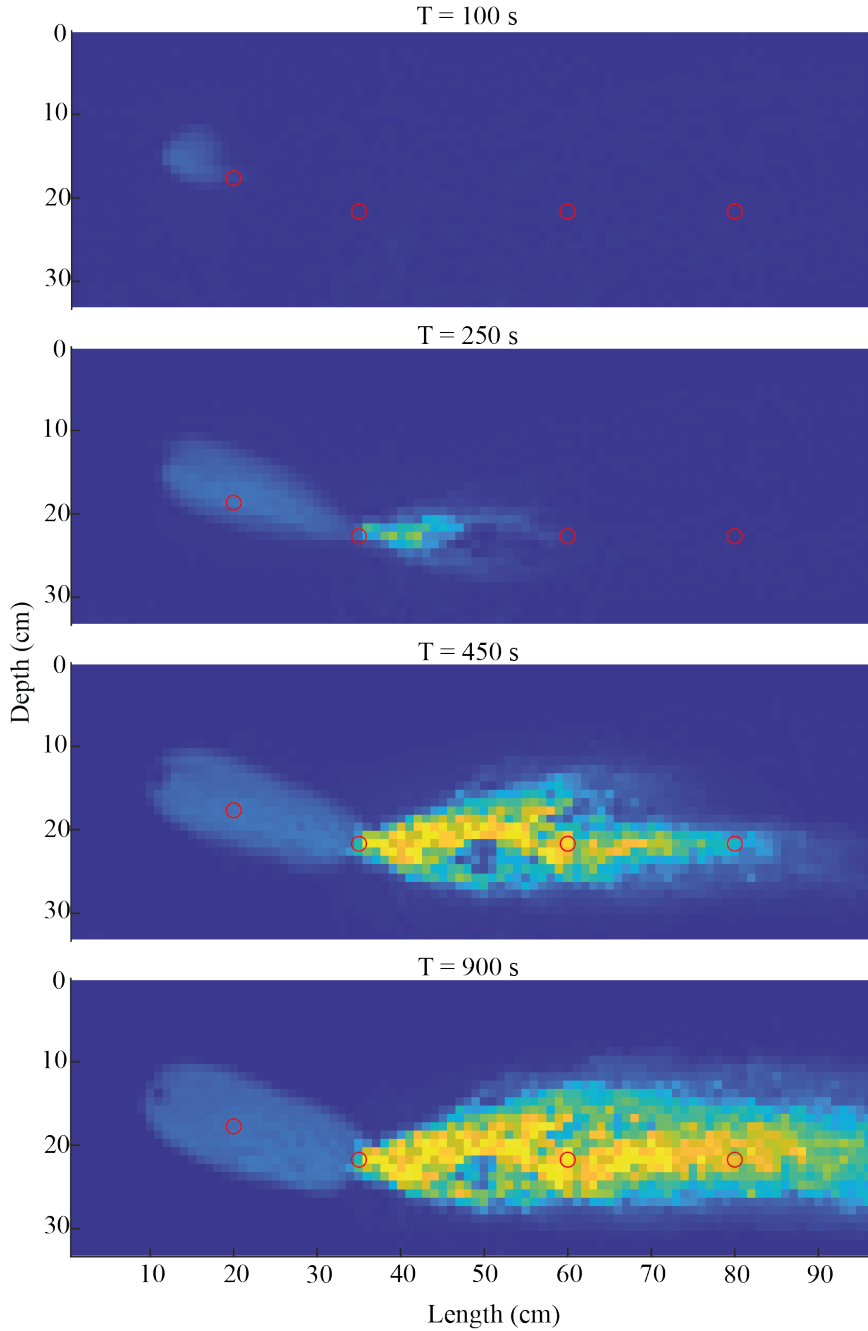


Figure 3.8. Example of dimensionless, from blue (0) to yellow (1), concentration distribution at four time steps: 100 s, 250 s, 450 s, and 900 s.

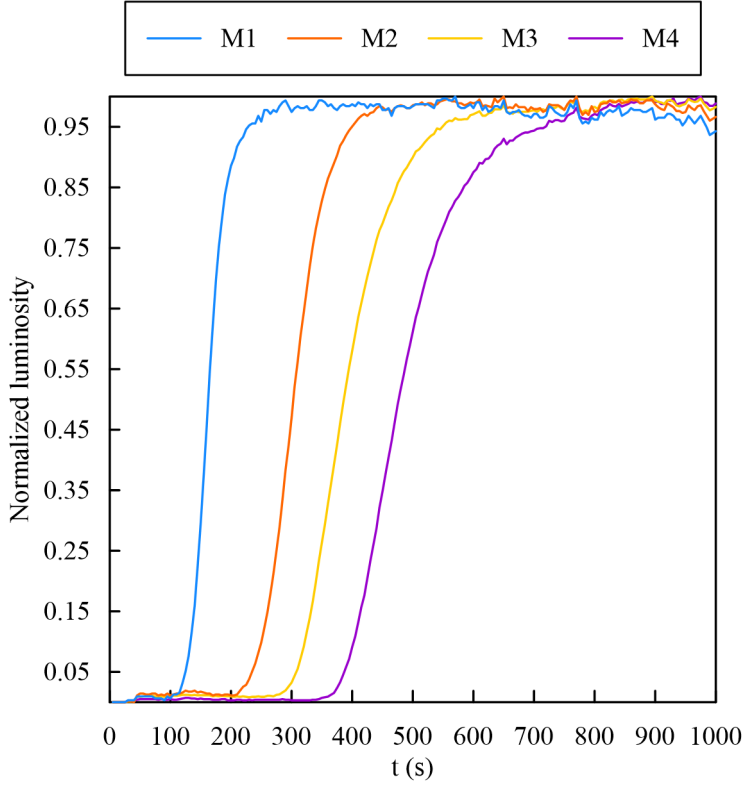


Figure 3.9. Breakthrough curves for the four monitoring points located within the sandbox.

3. Inversion process

Two case studies are conducted under different conditions to evaluate the performance of the inversion process using the ES-MDA. The models of both case studies are shown in Fig. 3.10. In the first case (Fig. 3.10-top), the objective is to reconstruct the hydraulic conductivity field based on concentration measurements from the four monitoring points. In the second case (Fig. 3.10-bottom), the aim is to estimate the resistivity field using apparent resistivity as the observation data. Both models are designed to replicate the geometry of the sandbox, as described in **Sub-**

section 1. The model domain is discretized into a grid of $96 \times 1 \times 39$ cells, each with dimensions of 1 cm x 1 cm x 1 cm, resulting in a total of 3744 parameters. The accuracy of the estimated fields is assessed using $RMSE$, while the coefficient of determination (R^2) is used to compare the estimated values to the true observations (**Appendix A**).

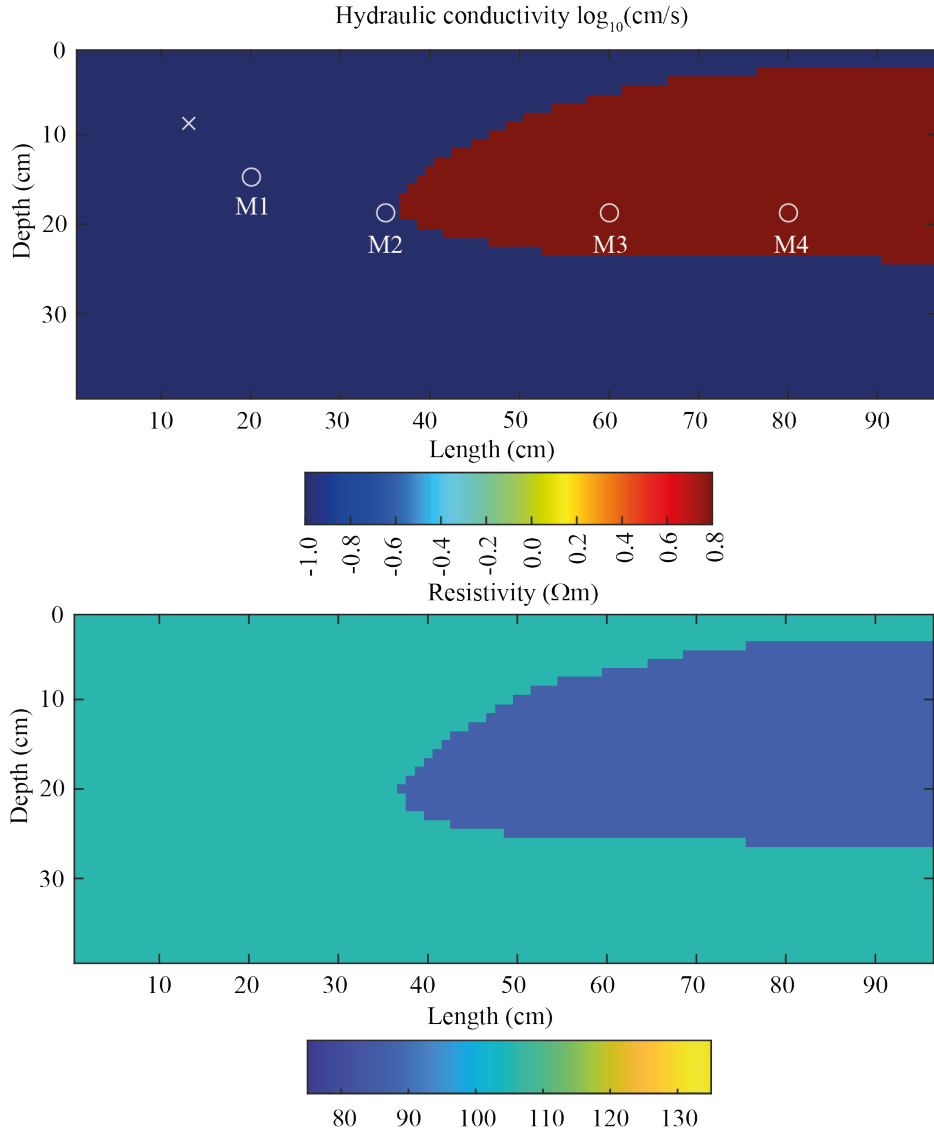


Figure 3.10. Log10 hydraulic conductivity (top) and resistivity (bottom) fields. The white cross in the hydraulic conductivity plot marks the tracer injection point, and the white circles indicate the locations of breakthrough monitoring points.

3.1. Case 1: Estimation of the Hydraulic Conductivity Field Using Concentration Data

In Case 1, the hydraulic conductivity field is estimated based on concentration data collected from the laboratory experiment. The converted breakthrough curves from the four monitoring points are used to guide the inversion process. Each monitoring point provided 200 observations, resulting in 800 data points (Fig. 3.9). Forward modeling is conducted with MODFLOW to simulate groundwater flow, while MT3DMS is used to simulate mass transport.

The ES-MDA is performed over six iterations, using an ensemble size of 500. The initial ensemble of parameters is homogeneous, with values randomly selected from the range [0.1–6] mg/L. Observation errors are normally distributed, with a zero mean and a variance of 0.01 (mg/L)². Spatial covariance localization is applied, considering a correlation length of 40 cm, and covariance inflation is used with a factor of 1.01. A decreasing α set with values equal to [364.0; 121.3; 40.4; 13.5; 4.5; 1.5] is used.

The result, plotted in Fig. 3.11 (top) on a logarithmic scale to highlight differences, demonstrates that the inversion process effectively captured the general distribution of hydraulic conductivity ($RMSE = 2.35$ cm/s). The figure clearly reveals a low hydraulic conductivity zone on the left side of the domain, which is marked by the dark blue region. This zone corresponds to an area of lower hydraulic conductivity material, likely sand, as observed from laboratory experiments. In contrast, the right side of the domain is characterized by higher hydraulic conductivity, represented by the yellow to red colors, which likely correspond to more permeable materials like gravel. Figure 3.11 (bottom) presents the log10 standard deviation corresponding to the 500 realizations of the estimated hydraulic conductivity. The accuracy of the hydraulic conductivity estimation is highest near the monitoring points, where more data are available to guide the inversion process. Further from these points, the uncertainty in the estimated values increases.

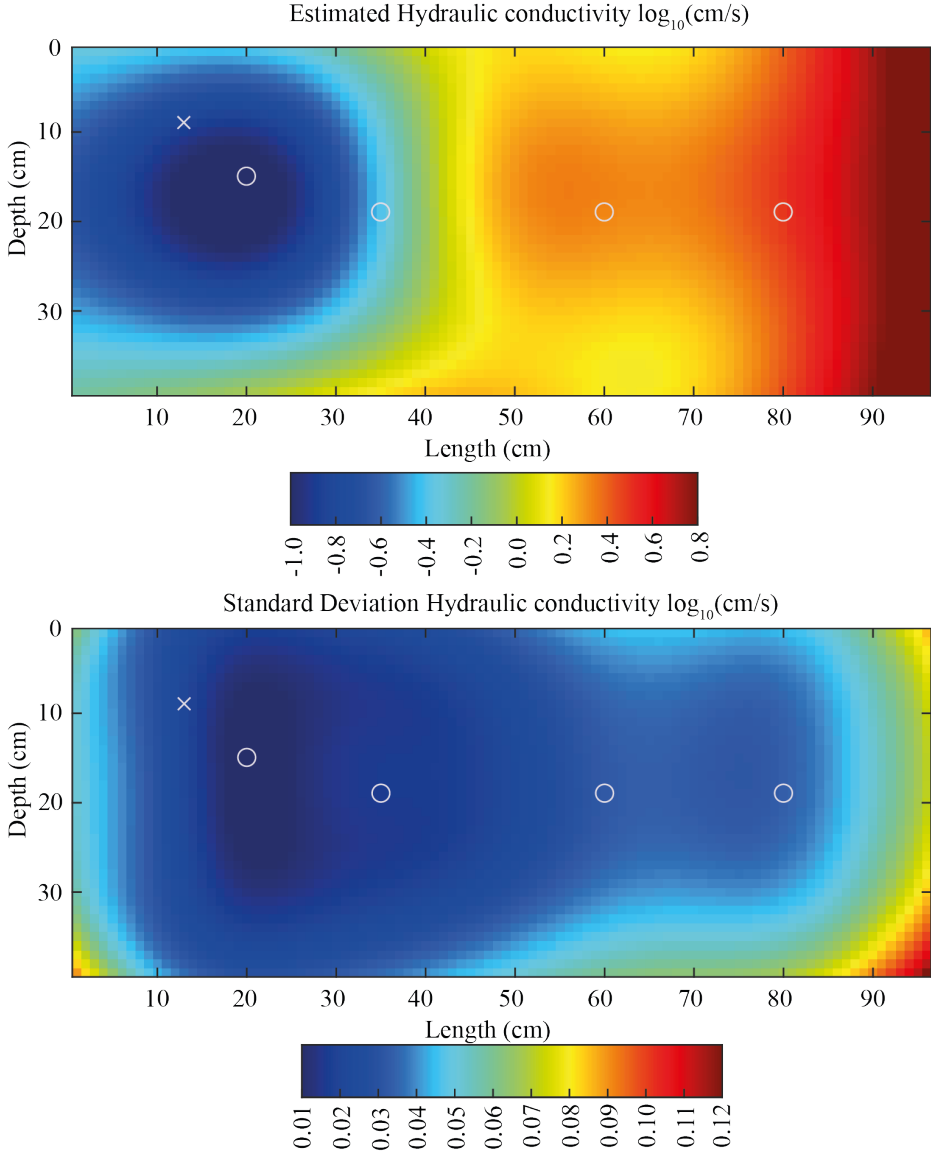


Figure 3.11. Average of the 500 realizations of the estimated hydraulic conductivity field (top) and respective standard deviation map (bottom), plotted in $\log_{10}(\text{cm/s})$. The white cross marks the location of the tracer injector, while the white circles denote the monitoring points where observations were collected.

Figure 3.12 presents a comparison between observed (blue dots) and estimated

(black line) concentrations over time for the four monitoring points (M1, M2, M3, M4). The M1 estimated curve closely follows the observed data, showing a strong correlation between model predictions and experimental results ($R^2 = 0.998$). The M2 curve shows a delayed increase in concentration but accurately captures the values both right before the rise and on the plateau ($R^2 = 0.988$). The M3 breakthrough curve also exhibits a later increase in concentration, with a slight overestimation on the plateau ($R^2 = 0.988$). This could suggest minor model inaccuracies in representing tracer transport in this sandbox region. Similar to M2 and M3 curves, M4 occurs later but the estimated curve matches well with the observed data ($R^2 = 0.988$). There is a slight lag in the estimated curve reaching the plateau, but overall the model accurately represents the behavior of the tracer.

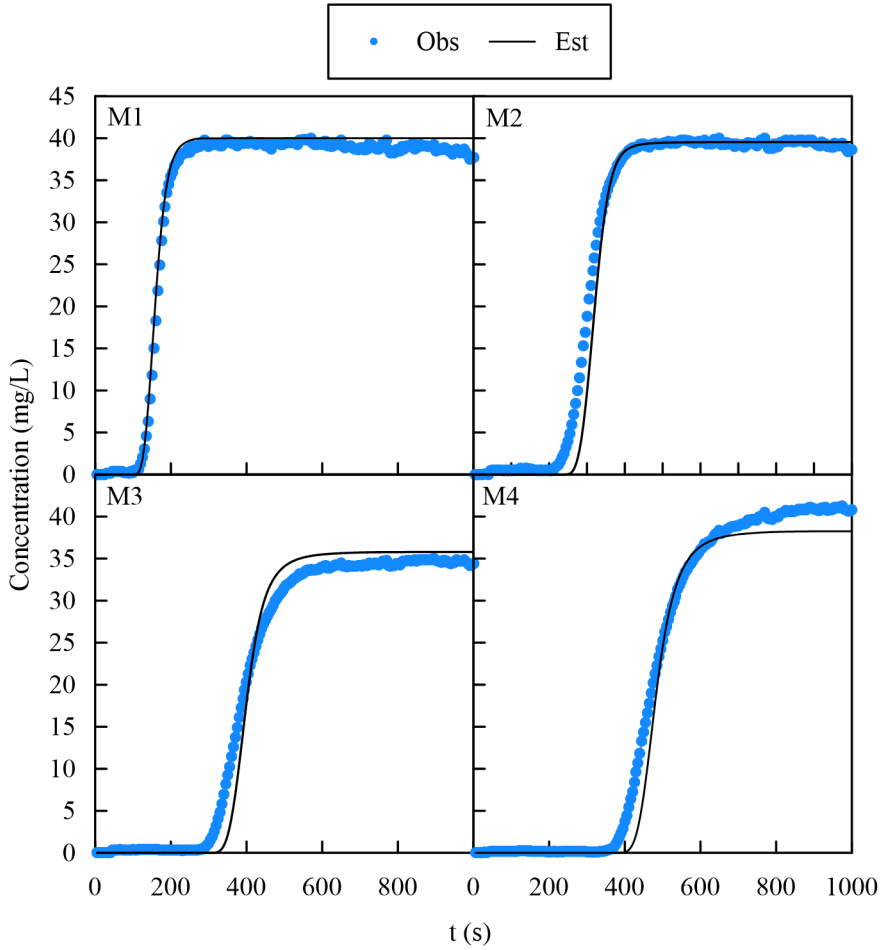


Figure 3.12. Comparison between the observed and the estimated breakthrough curves for the four monitoring points located within the sandbox.

3.2. Case 2: Estimation of the Resistivity Field Using Apparent Resistivity Data

The goal of Case 2 is to estimate the resistivity field based on apparent resistivity measurements obtained from laboratory experiments. Since these measured values required further analysis for accurate interpretation, the approach is first validated

using a synthetic model. This model is constructed using the resistivity values of the two materials, as listed in Table 3.1. The synthetic model generated 225 apparent resistivity values using a forward model (Cockett et al., 2015) that simulated the ERT sandbox system. These generated values served as observation data for the inversion process. The apparent resistivity distribution observed is presented in Fig. 3.13.

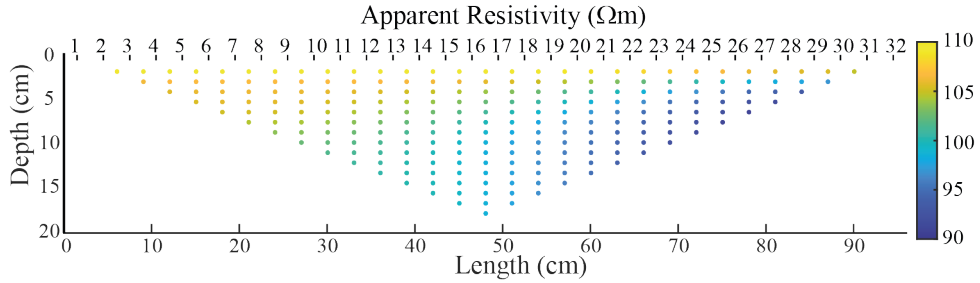


Figure 3.13. Observed apparent resistivity.

The ES-MDA is performed, running over six iterations with an ensemble size of 500. As in Case 1, each realization of the parameter ensemble is initialized with homogeneous values randomly picked from the range of $[10\text{--}500]$ Ohm-m. Observation errors are assumed to follow a normal distribution with a zero mean and a variance of $0.1 (\text{Ohm}\cdot\text{m})^2$. Spatial covariance localization is applied with a correlation length of 40 cm. A decreasing α set with values equal to $[364.0; 121.3; 40.4; 13.5; 4.5; 1.5]$ is used.

The results, shown in Fig. 3.14 (top), indicate that the inversion process successfully captured the general resistivity distribution, achieving a $RMSE = 4.63$ Ohm-m across the field. However, difficulties in distinguishing the interfaces between subsurface materials, especially at the sand and gravel transition zones, are noted. Furthermore, Fig. 3.14 (bottom) presents the standard deviation map of the estimated resistivity. The greatest uncertainty, represented by yellow shades, occurs near the edges, while the central region exhibits very low standard deviation values resulting from the presence of the observations.

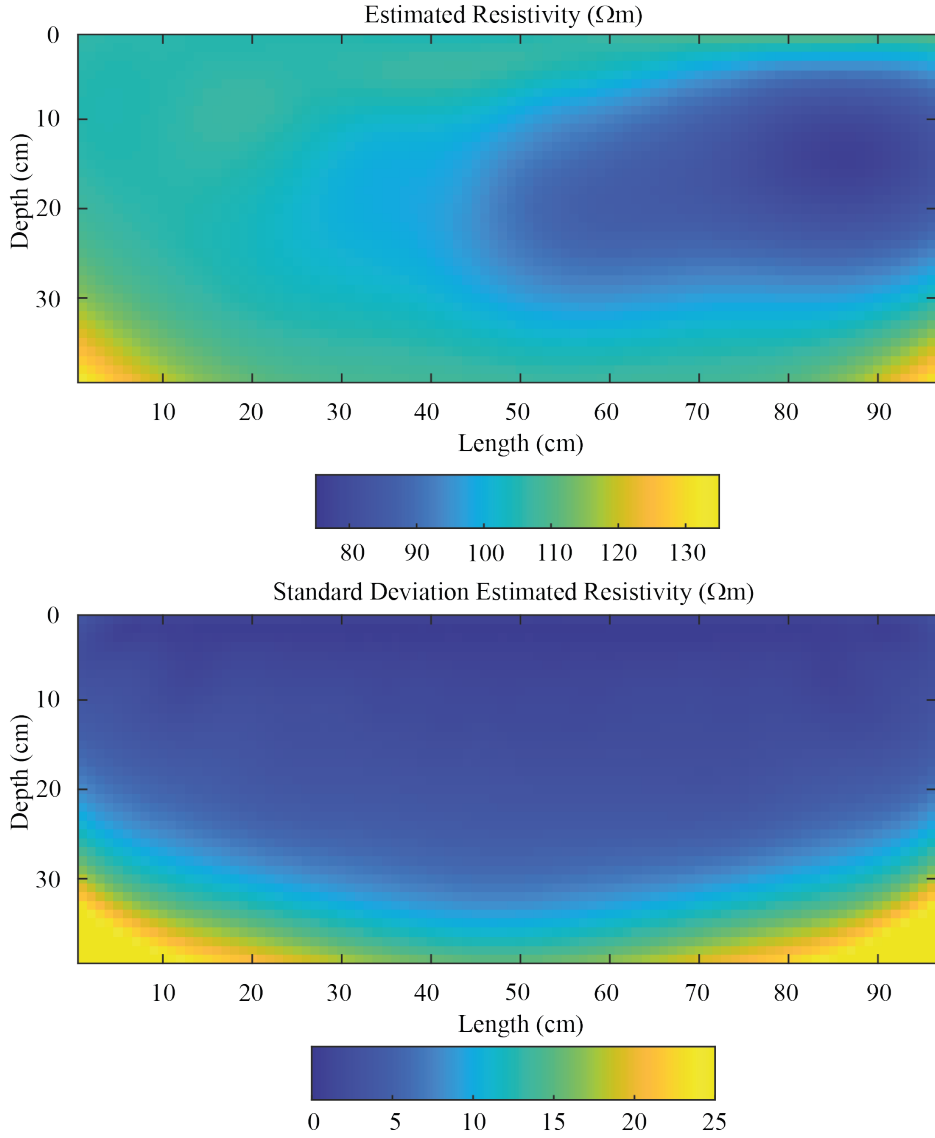


Figure 3.14. Average of the 500 realizations of the estimated resistivity fields (top) and respective standard deviation map (bottom).

Figure 3.15 presents the scatterplot that compares the observed and estimated apparent resistivity. The points are closely aligned along a diagonal trend, indicating a strong correlation between the observed and estimated values, further confirmed

by a high determination coefficient of $R^2 = 0.998$. This alignment suggests that the estimation method accurately reflects the observed resistivity values.

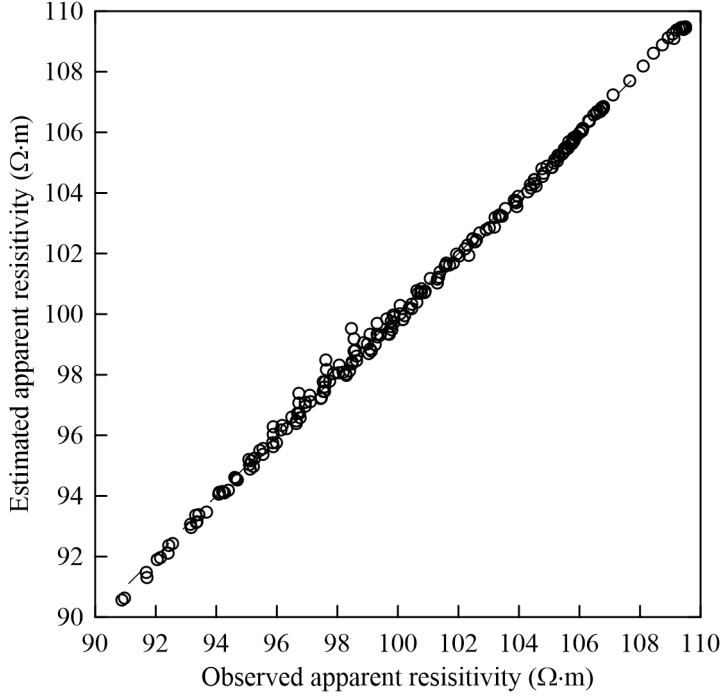


Figure 3.15. Average of the 500 realizations of the estimated resistivity fields.

4. Future developments

The study demonstrates a hydrogeophysical approach using the Ensemble Smoother with Multiple Data Assimilation to laboratory data, combining ERT and concentration measurements to estimate the hydraulic conductivity field of a sandbox. Concentration measurements have proven to be highly reliable and accurate in the transition to laboratory data. However, improvements in the ERT data acquisition system are still necessary to fully optimize their integration into the estimation process.

A key direction for future work will involve integrating measured ERT data with concentration-time measurements, as highlighted in the research conducted by Fagandini et al. (2024a). This integration is expected to improve the accuracy of hydraulic conductivity estimations during inversion procedures. As observed in the current study, relying only on concentration data limits the estimation to regions near the observation points. By combining ERT and concentration datasets, more detailed and reliable reconstructions of subsurface heterogeneities can be achieved, leading to an overall enhancement of the model's performance. Another critical area for development is the establishment of an accurate petrophysical relationship, essential to link electrical properties with hydraulic characteristics. Additionally, the use of heterogeneous random fields generated through stochastic Gaussian simulation could further improve the inversion process by better representing subsurface variability, leading to more realistic and accurate results.

The ERT acquisition system requires enhancements, especially in distinguishing materials with similar electrical properties. Furthermore, once validated, the acquisition system can be taken into consideration for field data collection.

In summary, although the proposed approach is very promising, further developments in ERT acquisition are needed to realize the method's full potential in subsurface characterization. Integrating ERT data with other types of data, such as concentration measurements, will be a key step in obtaining more accurate and robust estimates of hydraulic conductivity.

Overall conclusions

The aim of this PhD thesis is to introduce innovative methodologies for groundwater management, highlighting the application of geostatistics and inverse modeling techniques. The work is divided into three main sections, each addressing complementary aspects that contribute to a methodological framework useful for solving practical applications in aquifer water resource management. Specifically, it covers three interconnected domains: gap-filling methods for precipitation time series, inversion techniques for subsurface delineation, and laboratory-based experimental studies for data acquisition.

A key aspect in groundwater applications is identifying aquifer recharge, which is mainly related to precipitation. However, to effectively support the characterization and analysis of water resources, rainfall data sets must be complete and continuous. Consequently, gap-filling methodologies play a crucial role in addressing data gaps. The first section of this thesis demonstrates the effectiveness of kriging-based methods for filling gaps in daily precipitation time series. A novel approach that uses monthly semi-variograms and cross semi-variograms, rather than daily ones, was developed to reduce computational demands while maintaining accuracy. The study applied ordinary kriging and simple cokriging techniques, along with the traditional FAO method, demonstrating how spatial correlations can enhance precipitation estimates while accounting for seasonality and reducing computational costs. As a result of the study, the FAO method showed advantages in flat terrain, particularly at stations with high correlation. In contrast, geostatistical approaches, especially

those that incorporate secondary variables (i.e., elevation), enhance precipitation estimates in complex and variable terrain. These results underscore the versatility and robustness of the proposed methodologies, providing a valuable framework for improving precipitation data quality across diverse geographical settings.

Once external factors affecting groundwater resources are identified, characterization of the subsurface becomes possible. This includes the determination of hydrological properties (e.g., hydraulic conductivity, transmissivity and specific storage), geophysical characteristics (e.g., electrical conductivity or resistivity), as well as tracing the origin and history of potential contamination sources. The aim of the second section was to demonstrate how the integration of hydrological and geophysical data improves the understanding of aquifer properties and subsurface processes. Each application showed the effectiveness of the proposed methodology and highlighted how the combination of multiple data sources improved property estimates. Moreover, the importance of selecting appropriate petrophysical relationships, those that link hydrological and geophysical variables, is emphasized. Another aspect to be taken into account is the uncertainty of the prediction model. Since the inverse methodology requires a numerical model that can accurately describe forward processes, errors in the model structure may add to the measurement noise. Therefore, a correct and calibrated prediction model is crucial for obtaining a reliable inverse solution, and careful verification of the most uncertain model parameters is advisable.

Finally, after a synthetic model has been developed that accurately represents the characteristics of our system, real data acquisitions can be integrated. The third section of the thesis focused on investigations in a controlled laboratory environment. The successful application of the ES-MDA to the synthetic application provided confidence in the technique's ability and therefore allowed its use with real experimental data. Electrical resistivity tomography data and concentration measurements were acquired to identify the field characteristics of the sandbox using the fully coupled hydrogeophysical approach. This study employed a non-commercial electrical resistivity tomography system, specifically tailored for laboratory experiments.

Concentration measurements proved to be reliable and precise in the transition to laboratory data. However, further enhancements in the electrical resistivity tomography data acquisition system are still needed to fully optimize their integration into the estimation process.

This thesis makes significant methodological and practical contributions to the characterization of the aquifers. It demonstrated the effectiveness of advanced geostatistical, inversion modeling and laboratory techniques for addressing key challenges in subsurface and hydrological modeling. The findings suggested several promising directions for future research. First, even though kriging-based methods have proven effective, exploring and developing alternative approaches could further improve data interpretation and applicability. Second, as datasets grow in size and complexity, enhancing the computational efficiency of inversion techniques will be essential. Third, to address the limitations identified in the Electrical Resistivity Tomography system, future work could focus on designing more sensitive and accurate instruments tailored for both laboratory and field applications. Additionally, expanding field validation of these techniques across diverse environmental settings will refine the models and increase their robustness for practical applications in aquifer and water resource management.

APPENDIX A: Evaluation Metrics

This study employs the following evaluation metrics.

The Mean Error (ME) is the average of all errors. It quantifies the error in terms of underestimation if the result is negative and overestimation if the result is positive.

$$ME = \frac{1}{N} \sum_{k=1}^N (\hat{y}_k - y_k), \quad (\text{A.1})$$

where N is the number of observations, y_k is the observed value, and $\hat{y}_k$ is the estimated value.

The Mean Absolute Error (MAE) measures the magnitude of the errors.

$$MAE = \frac{1}{N} \sum_{k=1}^N |\hat{y}_k - y_k|. \quad (\text{A.2})$$

The Root Mean Square Error ($RMSE$) is a common measurement of accuracy performance that is frequently used as a magnitude error evaluation. $RMSE$ is never negative, and a value of 0 indicates a perfect fit with the data.

$$RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^N (\hat{y}_k - y_k)^2}. \quad (\text{A.3})$$

The coefficient of determination, R^2 , indicates how well the estimated values align with the observed data, with values closer to 1 indicating stronger predictive per-

formance.

$$R^2 = 1 - \frac{\sum_{k=1}^N (\hat{y}_k - y_k)^2}{\sum_{k=1}^N (y_k - \bar{y})^2}, \quad (\text{A.4})$$

where $\bar{y}$ is the mean of the observed values.

APPENDIX B: Arduino code for ERT measurements

```
#include "ADS1X15.h"

// 32 electrodes with Wenner-Schlumberger configuration
// Modified resistors, voltage 24V, 220 Ohm resistors
// Repeated measurements (A swapped with B, N swapped with M):
// odd are direct and even are inverse

ADS1115 ADS(0x48); // object for ADC operations

int A, M, N, B;
int i, j, k;
int na, nb, nm, nn;
float va, vb, vm, vn, f;

const int num_ERT_config = 225; // Number of ERT measurements
const int num_misure = 50;      // Number of repeated measurements
const int frequenza = 20;       // Sampling frequency in Hz

char out[250];
int arA[] = {1,2,3,4,...};      // Values for A
```

```
int arB[] = {4,5,6,7,...};      // Values for B
int arN[] = {2,3,4,5,...};      // Values for N
int arM[] = {3,4,5,6,...};      // Values for M

int pA, pB, pN, pM; // Electrode positions

void setup() {
    ADS.begin(); // Initialize ADS
    ADS.setGain(0); // Set gain to default
    Serial.begin(115200); // Start serial communication

    // Configure output pins
    for (i = 2; i < 20; i++) {
        pinMode(i, OUTPUT);
    }
    for (j = 22; j < 54; j++) {
        pinMode(j, OUTPUT);
    }
}

void loop() {
    Serial.println("START MEASUREMENT");

    for (j = 0; j <= num_ERT_config-1; j++) {
        for (i = 0; i <= num_misure-1; i++) {
            // Direct measurement
            pA = arA[j] - 1;
            pB = arB[j] - 1;
            pM = arM[j] - 1;
```

```
pN = arN[j] - 1;

digitalWrite(51, 1); // Enable multiplexer

// Configure electrodes
configureElectrodes(pA, pB, pM, pN);

delay(10); // Wait for ground capacitance to charge

// Read ADC values
na = ADS.readADC(0);
nb = ADS.readADC(1);
nn = ADS.readADC(2);
nm = ADS.readADC(3);
f = ADS.toVoltage(1);

// Disable multiplexer
digitalWrite(51, 0);

// Calculate voltages
calculateVoltages(na, nb, nn, nm, f);

// Inverted measurement
pA = arB[j] - 1;
pB = arA[j] - 1;
pM = arN[j] - 1;
pN = arM[j] - 1;

digitalWrite(51, 1); // Enable multiplexer again
```

```
        configureElectrodes(pA, pB, pM, pN);
        delay(10); // Wait for ground capacitance

        // Read inverted ADC values
        readInvertedADC();

        delay(1000/frequenza); // Wait between measurements
    }
    delay(5000); // Wait before next measurement cycle
}
Serial.println("MEASUREMENT COMPLETED");
}

void configureElectrodes(int pA, int pB, int pM, int pN) {
    // Configure electrodes for the multiplexer
    digitalWrite(4, bitRead(pA, 0));
    digitalWrite(3, bitRead(pA, 1));
    digitalWrite(2, bitRead(pA, 2));
    digitalWrite(49, bitRead(pA, 3));
    digitalWrite(53, bitRead(pA, 4));

    digitalWrite(19, bitRead(pM, 0));
    digitalWrite(18, bitRead(pM, 1));
    digitalWrite(17, bitRead(pM, 2));
    digitalWrite(16, bitRead(pM, 3));
    digitalWrite(15, bitRead(pM, 4));

    digitalWrite(14, bitRead(pN, 0));
```



```
digitalWrite(13, bitRead(pN, 1));
digitalWrite(12, bitRead(pN, 2));
digitalWrite(11, bitRead(pN, 3));
digitalWrite(10, bitRead(pN, 4));

digitalWrite(9, bitRead(pB, 0));
digitalWrite(8, bitRead(pB, 1));
digitalWrite(7, bitRead(pB, 2));
digitalWrite(6, bitRead(pB, 3));
digitalWrite(5, bitRead(pB, 4));
}

void calculateVoltages(int na, int nb, int nn, int nm, float f) {
    va = na * 7.85 * f;
    vb = nb * 7.85 * f;
    vn = nn * 7.85 * f;
    vm = nm * 7.85 * f;
    sprintf(out, "%d %d %.7f %.7f %.7f %.7f %d %d %d %d",
        j + 1, i + 1, va, vb, vn, vm, pA + 1, pN + 1, pM + 1, pB + 1);
    Serial.println(out);
}

void readInvertedADC() {
    nb = ADS.readADC(0);
    na = ADS.readADC(1);
    nm = ADS.readADC(2);
    nn = ADS.readADC(3);
    f = ADS.toVoltage(1);
```

```
digitalWrite(51, 0); // Disable multiplexer

calculateVoltages(na, nb, nn, nm, f);
}
```

APPENDIX C: MATLAB codes for ERT processing

Listing C.1. ERT data acquisition code

```
N_ERT = 225; %Number of ERT measurements
N_rep = 50*2; %Number of repeated measurements

rows = N_ERT * N_rep + 3;

filename = 'test_/_/.txt'; % Recorded file name
if isfile(filename)
    disp('Attention:_the_file_already_exists!')
    return
else
    fileout = fopen(filename, 'w');
end

arduinoObj = serialport("COM3", 115200); % Select the serial port
configureTerminator(arduinoObj, "CR/LF");
flush(arduinoObj);
arduinoObj.UserData = struct("Data", [], "Count", 1);
```

```
%Save data in .txt file
i = 1;
while i < rows
    data = readline(arduinoObj);
    fprintf(fileout, '%s\n', data);
    disp(data);
    %disp(i);
    i = i + 1;
end

fclose(fileout);
```

Listing C.2. Reading data from text files and constructing the pseudosection

```
clear, clc
close all

n_ele = 32; % Number of electrodes
step = 0.03; % Electrode spacing
N_ERT = 225;
N_rep = 50;
file_OUT = 'Appres_test_/_/.xlsx'; %File out name
if isfile(file_OUT)
    disp('Attention:_the_file_already_exists!')
    return
else
    file_IN = 'test_/_/.txt';
end
```

% Reading data from direct and inverse measurement files

```
data = load(file_IN);  
even = 2:2:N_ERT * N_rep * 2;  
odd = 1:2:N_ERT * N_rep * 2;
```

*% Voltages and A B M N positions of direct and
inverse measurements*

```
V_A_dir = data(odd,3);  
V_A_inv = data(even,3);  
V_B_dir = data(odd,4);  
V_B_inv = data(even,4);  
V_M_dir = data(odd,6);  
V_M_inv = data(even,6);  
V_N_dir = data(odd,5);  
V_N_inv = data(even,5);  
A_dir = data(odd,7);  
A_inv = data(even,7);  
B_dir = data(odd,10);  
B_inv = data(even,10);  
M_dir = data(odd,9);  
M_inv = data(even,9);  
N_dir = data(odd,8);  
N_inv = data(even,8);  
clear data even odd
```

```
Direct_Current = abs(V_A_dir - V_B_dir) / 220;  
Direct_Res = abs(V_M_dir - V_N_dir) ./ Direct_Current;  
[Direct_Res_mean, Direct_Res_var] =
```

```
Evaluate_M_V(Direct_Res , N_rep, N_ERT);
Direct_Res_mean = Direct_Res_mean';
Direct_Res_std = sqrt(Direct_Res_var)';

Inverse_Current = abs(V_A_inv - V_B_inv) / 220;
Inverse_Res= abs(V_M_inv - V_N_inv) ./ Inverse_Current;
[Inverse_Res_mean , Inverse_Res_var] =
Evaluate_M_V(Inverse_Res , N_rep, N_ERT);
Inverse_Res_mean = Inverse_Res_mean';
Inverse_Res_std = sqrt(Inverse_Res_var)';
```

figure

```
title('Direct-inverse_resistivity')
plotregression(Direct_Res_mean, Inverse_Res_mean)
```

figure

```
t = tiledlayout(1, 2);
h1 = nexttile;
errorbar(1:1:N_ERT, Direct_Res_mean, Direct_Res_std, '.r')
title('Standard_Deviation_direct_Resistivity')
h2 = nexttile;
errorbar(1:1:N_ERT, Inverse_Res_mean, Inverse_Res_std, '.r')
title('Standard_Deviation_inverse_Resistivity')
```

```
% Electrode identification
```

```
A = A_dir(1:N_rep:N_ERT * N_rep);
B = B_dir(1:N_rep:N_ERT * N_rep);
M = M_dir(1:N_rep:N_ERT * N_rep);
N = N_dir(1:N_rep:N_ERT * N_rep);
```

```
%% Apparent resistivity
% computation of distances between electrodes
AM = abs(M - A) * step; % in meters
AN = abs(N - A) * step; % in meters
BM = abs(M - B) * step; % in meters
BN = abs(N - B) * step; % in meters
AB = abs(B - A) * step; % in meters
MN = abs(N - M) * step; % in meters

H = 1 ./ AN - 1 ./ AM - 1 ./ BN + 1 ./ BM;
K = 2 * pi ./ H; % geometric factor (m)

App_Direct_Res = K .* Direct_Res_mean;
App_Inverse_Res = K .* Inverse_Res_mean;

% Measurements position
X = ((M + N) / 2 - 1) * step + 0.015; % (m)
X = X';

% Measurement depth positioning
n = abs(M - A); % Identification of the measurement layer
Z = zeros(N_ERT, 1);
ele_dist = zeros(N_ERT, 1);

for i = 1:N_ERT
    ele_dist(i) = max(AB(i), MN(i));
end
```

```
for i = 1:N_ERT
    if n(i) == 1
        Z(i) = -0.173 * ele_dist(i); % (m)
    elseif n(i) == 2
        Z(i) = -0.186 * ele_dist(i); % (m)
    elseif n(i) == 3
        Z(i) = -0.189 * ele_dist(i); % (m)
    elseif n(i) == 4
        Z(i) = -0.190 * ele_dist(i); % (m)
    elseif n(i) == 5
        Z(i) = -0.190 * ele_dist(i); % (m)
    elseif n(i) == 6
        Z(i) = -0.191 * ele_dist(i); % (m)
    else
        Z(i) = -0.191 * ele_dist(i); % (m)
    end % if
end

% Grid construction
delta = 0.01; % Grid step (cm)
x_gr = 0.:delta:96; z_gr = 0.:delta:20; % (cm)
[X_gr, Z_gr] = meshgrid(x_gr, z_gr);
X_gr = X_gr / 100; Z_gr = -Z_gr / 100; % (m)

figure
image = griddata(X, Z, App_Direct_Res, X_gr, Z_gr);
pcolor(X_gr * 100, Z_gr * 100, image)
shading interp
axis equal
```

```

xlabel ( 'cm' ), ylabel( 'cm' ), colorbar
xlim([0 96]), ylim ([-20 1])

% Make data labels
text(X * 100, Z * 100, sprintfc('_%d', 1:N_ERT));

% Electrodes
ele = 1:1:n_ele; ele = ele * step - 0.015;
z_ele = zeros(n_ele, 1);
hold on
plot(X * 100, Z * 100, 'ko') % measurement position
scatter(ele * 100, z_ele * 100, 'k|'); % electrode position
text(ele * 100, z_ele * 100 + 1, sprintfc('_%d', 1:n_ele))

figure
scatter(X * 100, Z * 100, [], App_Direct_Res,
'filled', 'sizeData', 50)
hold on
scatter(ele * 100, z_ele * 100, 'k|');
text(ele * 100, z_ele * 100 + 1, sprintfc('_%d', 1:n_ele))
axis equal
title('Apparent_resistivity_(Ohm\cdotm)')
xlabel ( 'cm' ), ylabel( 'cm' ), colorbar

figure
scatter(X * 100, Z * 100, [], Direct_Res_mean,
'filled', 'sizeData', 50)
hold on
scatter(ele * 100, z_ele * 100, 'k|');

```

```

text(ele * 100, z_ele * 100 + 1, sprintf('_%d', 1:n_ele))
axis equal
title('Resistivity_(Ohm)')
xlabel('cm'), ylabel('cm'), colorbar

figure
scatter(X * 100, Z * 100, [], Direct_Res_std,
'filled', 'sizeData', 100)
hold on
scatter(ele * 100, z_ele * 100, 'k|');
axis equal
title('Standard_Deviation_Resistivity_(Ohm)')
xlabel('cm'), ylabel('cm'), colorbar
text(X * 100, Z * 100, sprintf('_%d', 1:N_ERT))

```

Listing C.3. Function: computation of mean and variance of measurements

```

function [V_mean, V_var] = Evaluate_M_V(V, n_r, N_m)
% Evaluate_M_V: Evaluate the mean and variance of
voltage measurements
%
% Input:
% V      - Voltage measurements vector
% n_r    - Number of repetitions per measure
% N_m    - Total number of measurements
%
% Output:
% V_mean - Mean of measurements for each electrode
% V_var  - Variance of measurements for each electrode

```

```
% Reshape of the measurement vector
V_r = reshape(V, n_r, N_m);

%Displaying the measurement boxplot
figure
boxplot(V_r)

%Identification of outliers
[TF] = isoutlier(V_r, 'mean');

%Removal of outliers and replacement with NaN
V_a = V_r;
V_a(TF == 1) = NaN;
V_a(V_a == 0) = NaN;

%Mean of the measurements, ignoring NaN
V_mean = mean(V_a, "omitnan");

%Variance of the measurements, ignoring NaN
V_var = var(V_a, "omitnan");

end
```


Bibliography

- Adhikary, S. K., Muttill, N., and Yilmaz, A. G. (2017). Cokriging for enhanced spatial interpolation of rainfall in two Australian catchments. *Hydrological Processes*, 31(12):2143–2161.
- Agou, V. D., Varouchakis, E. A., and Hristopulos, D. T. (2019). Geostatistical analysis of precipitation in the island of Crete (Greece) based on a sparse monitoring network. *Environmental Monitoring and Assessment*, 191(6):353.
- Aguilera, H., Guardiola-Albert, C., and Serrano-Hidalgo, C. (2020). Estimating extremely large amounts of missing precipitation data. *Journal of Hydroinformatics*, 22(3):578–592.
- Allen, R. G., Pereira, L. S., Raes, D., and Smith, M. (1998). *Crop evapotranspiration - Guidelines for computing crop water requirements - FAO Irrigation and drainage paper 56*. FAO - Food and Agriculture Organization of the United Nations.
- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., and Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1).
- Archie, G. E. (1942). The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics. *Open Journal of Geology*.
- Asher, M. J., Croke, B. F., Jakeman, A. J., and Peeters, L. J. (2015). A review of

- surrogate models and their application to groundwater modeling. *Water Resources Research*, 51(8):5957–5973.
- Asli, M. and Marcotte, D. (1995). Comparison of approaches to spatial estimation in a bivariate context. *Mathematical Geology*, 27(5):641–658.
- Bacchi, B. and Kottegoda, N. T. (1995). Identification and calibration of spatial correlation patterns of rainfall. *Journal of Hydrology*, 165(1-4):311–348.
- Bear Jacob (1972). *Dynamics of Fluids in Porous Media*. American Elsevier Publishing Company, Inc., New York.
- Bear Jacob and Bachmat Y. (1990). *Introduction to Modeling of Transport Phenomena in Porous Media*. Kluwer Academic Publisher, The Netherlands.
- Beek, E. G., Stein, A., and Janssen, L. L. F. (1992). Spatial variability and interpolation of daily precipitation amount. *Stochastic Hydrology and Hydraulics*, 6(3):209–221.
- Berg, S. J. and Illman, W. A. (2011). Capturing aquifer heterogeneity: Comparison of approaches through controlled sandbox experiments. *Water Resources Research*, 47(9).
- Bernardi, M. S., Carey, M., Ramsay, J. O., and Sangalli, L. M. (2018). Modeling spatial anisotropy via regression with partial differential regularization. *Journal of Multivariate Analysis*, 167:15–30.
- Bostan, P. A., Heuvelink, G. B. M., and Akyurek, S. Z. (2012). Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. *International Journal of Applied Earth Observation and Geoinformation*, 19:115–126.
- Brace, W. F., Walsh, J. B., and Frangos, W. T. (1968). Permeability of granite under high pressure. *Journal of Geophysical Research*, 73(6):2225–2236.

- Brandsma, T. and Buishand, T. A. (1998). Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling. *Hydrology and Earth System Sciences*, 2(2/3):195–209.
- Butera, I., Tanda, M. G., and Zanini, A. (2013). Simultaneous identification of the pollutant release history and the source location in groundwater by means of a geostatistical approach. *Stochastic Environmental Research and Risk Assessment*, 27(5):1269–1280.
- Buytaert, W., Celleri, R., Willems, P., Bièvre, B. D., and Wyseure, G. (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of Hydrology*, 329(3-4):413–421.
- Campling, P., Gobin, A., and Feyen, J. (2001). Temporal and spatial rainfall analysis across a humid tropical catchment. *Hydrological Processes*, 15(3):359–375.
- Camporese, M., Cassiani, G., Deiana, R., and Salandin, P. (2011). Assessment of local hydraulic properties from electrical resistivity tomography monitoring of a three-dimensional synthetic tracer test experiment. *Water Resources Research*, 47(12).
- Carpenter, P. J., Ding, A., and Cheng, L. (2012). Identifying groundwater contamination using resistivity surveys at a landfill near Maoming. *Nature Education Knowledge*, 3(7).
- Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., and Slooten, L. J. (2005). Inverse problem in hydrogeology.
- Carrera-Hernández, J. and Gaskin, S. (2007). Spatio temporal analysis of daily precipitation and temperature in the Basin of Mexico. *Journal of Hydrology*, 336(3-4):231–249.
- Chen, D., Ou, T., Gong, L., Xu, C.-Y., Li, W., Ho, C.-H., and Qian, W. (2010). Spatial interpolation of daily precipitation in China: 1951–2005. *Advances in Atmospheric Sciences*, 27(6):1221–1232.

- Chen, Y. and Zhang, D. (2006). Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Advances in Water Resources*, 29(8):1107–1122.
- Chen, Z., Gómez-Hernández, J. J., Xu, T., and Zanini, A. (2018). Joint identification of contaminant source and aquifer geometry in a sandbox experiment with the restart ensemble Kalman filter. *Journal of Hydrology*, 564:1074–1084.
- Chen, Z., Xu, T., Gómez-Hernández, J. J., and Zanini, A. (2021). Contaminant Spill in a Sandbox with Non-Gaussian Conductivities: Simultaneous Identification by the Restart Normal-Score Ensemble Kalman Filter. *Mathematical Geosciences*, 53(7):1587–1615.
- Chen, Z., Zong, L., Gómez-Hernández, J. J., Xu, T., Jiang, Y., Zhou, Q., Yang, H., Jia, Z., and Mei, S. (2023). Contaminant source and aquifer characterization: An application of ES-MDA demonstrating the assimilation of geophysical data. *Advances in Water Resources*, 181.
- Cheng, M., Wang, Y., Engel, B., Zhang, W., Peng, H., Chen, X., and Xia, H. (2017). Performance Assessment of Spatial Interpolation of Precipitation for Hydrological Process Simulation in the Three Gorges Basin. *Water*, 9(11):838.
- Christakos, G. (1984). On the Problem of Permissible Covariance and Variogram Models. *Water Resources Research*, 20(2):251–265.
- Citarella, D., Cupola, F., Tanda, M. G., and Zanini, A. (2015). Evaluation of dispersivity coefficients by means of a laboratory image analysis. *Journal of Contaminant Hydrology*, 172:10–23.
- Cockett, R., Kang, S., Heagy, L. J., Pidlisecky, A., and Oldenburg, D. W. (2015). SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications.

- Crestani, E., Camporese, M., Bá U, D., and Salandin, P. (2012). EnKF vs. ES for tracer test data assimilation Hydrology and Earth System Sciences Discussions Ensemble Kalman filter versus ensemble smoother for assessing hydraulic conductivity via tracer test data assimilation EnKF vs. ES for tracer test data assimilation. *Hydrol. Earth Syst. Sci. Discuss*, 9:13083–13115.
- Crestani, E., Camporese, M., and Salandin, P. (2015). Assessment of hydraulic conductivity distributions through assimilation of travel time data from ERT-monitored tracer tests. *Advances in Water Resources*, 84:23–36.
- Cupola, F., Tanda, M. G., and Zanini, A. (2015). Laboratory sandbox validation of pollutant source location methods. *Stochastic Environmental Research and Risk Assessment*, 29(1):169–182.
- Dahlin, T. and Zhou, B. (2004). A numerical comparison of 2D resistivity imaging with 10 electrode arrays. *Geophysical Prospecting*, 52(5):379–398.
- Daly, C. (2006). Guidelines for assessing the suitability of spatial climate data sets. *International Journal of Climatology*, 26(6):707–721.
- Daly, C., Neilson, R. P., and Phillips, D. L. (1994). A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *Journal of Applied Meteorology*, 33(2):140–158.
- Das, V., Pollack, A., Wollner, U., and Mukerji, T. (2019). Geophysics 1 Convolutional Neural Network for seismic impedance inversion. *Geophysics*.
- Delmonte, N., Spaggiari, D., Chiorboli, G., Fagandini, C., and Zanini, A. (2024). An Electrical Resistivity Tomography System for Imaging at Laboratory Scale (Manuscript submitted for publication). *Measurement*.
- Domenico Patrick A. and Schwartz Franklin W. (1998). *Physical and Chemical Hydrogeology*, volume 1. John Wiley & Sons, University of Michigan.

- D’Oria, M., Ferraresi, M., and Tanda, M. G. (2019). Quantifying the impacts of climate change on water resources in northern Tuscany, Italy, using high-resolution regional projections. *Hydrological Processes*, 33(6):978–993.
- D’Oria, M., Mignosa, P., Tanda, M. G., and Todaro, V. (2022). Estimation of levee breach discharge hydrographs: comparison of inverse approaches. *Hydrological Sciences Journal*, 67(1):54–64.
- Edwards, L. S. (1977). A modified pseudosection for resistivity and IP. *Geophysics*, 42(5):1020–1036.
- El Idrysy, E. H. and De Smedt, F. (2007). A comparative study of hydraulic conductivity estimations using geostatistics. *Hydrogeology Journal*, 15(3):459–470.
- Emerick, A. A. and Reynolds, A. C. (2012). History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations.
- Emerick, A. A. and Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Computers and Geosciences*, 55:3–15.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5).
- Evensen, G. (2018). Analysis of iterative ensemble smoothers for solving inverse problems. *Computational Geosciences*, 22(3):885–908.
- Fagandini, C., D’Oria, M., Todaro, V., and Zanini, A. (2025). A comparison between Empirical Bayes combined with Akaike’s Bayesian Information Criterion and Ensemble Smoother with Multiple Data Assimilation to evaluate geophysical data (Manuscript submitted for publication). *Mathematical Geosciences*.
- Fagandini, C., Todaro, V., Escada, C., Azevedo, L., Gómez-Hernández, J. J., and Zanini, A. (2024a). Coupled hydrogeophysical inversion through ensemble smoother with multiple data assimilation and convolutional neural network for

- contaminant plume reconstruction. *Stochastic Environmental Research and Risk Assessment*.
- Fagandini, C., Todaro, V., Tanda, M., Pereira, J., Azevedo, L., and Zanini, A. (2024b). Missing Rainfall Daily Data: A Comparison Among Gap-Filling Approaches. *Mathematical Geosciences*, 56(2).
- Fernandez Visentini, A., Linde, N., Borgne, T. L., and Dentz, M. (2020). Inferring geostatistical properties of hydraulic conductivity fields from saline tracer tests and equivalent electrical conductivity time-series. *Advances in Water Resources*, 146.
- Frohlich, R. K. and Urish, D. W. (2002). The use of geoelectrics and test wells for the assessment of groundwater quality of a coastal industrial site. *Journal of Applied Geophysics*.
- Godoy, V. A., Napa-García, G. F., and Gómez-Hernández, J. J. (2022). Ensemble Smoother with Multiple Data Assimilation as a Tool for Curve Fitting and Parameter Uncertainty Characterization: Example Applications to Fit Nonlinear Sorption Isotherms. *Mathematical Geosciences*, 54(4):807–825.
- Gómez-Hernández, J. J., Hendricks Franssen, H. J., and Sahuquillo, A. (2003). Stochastic conditional inverse modeling of subsurface mass transport: A brief review and the self-calibrating method. *Stochastic Environmental Research and Risk Assessment*, 17(5):319–328.
- Gómez-Hernández, J. J. and Wen, X. H. (1994). Probabilistic assessment of travel times in groundwater modeling. *Stochastic Hydrology and Hydraulics*, 8(1):19–55.
- Gómez-Hernández, J. J. and Xu, T. (2022). Contaminant Source Identification in Aquifers: A Critical View. *Mathematical Geosciences*, 54(2):437–458.
- Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*. Oxford University Press Inc.

- Goovaerts, P. (1999). Geostatistics in soil science: state-of-the-art and perspectives. *Geoderma*, 89:1–45.
- Goovaerts, P. (2000). Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. *Journal of Hydrology*, 228(1-2):113–129.
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., and Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, 77:354–377.
- Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13:49–52.
- Hansen, T. M. (2022). mGstat: A Geostatistical Matlab toolbox.
- Harbaugh, A. W. (2005). MODFLOW-2005, the U.S. Geological Survey modular ground-water model – the Ground-Water Flow Process. Technical report, U.S. Geological Survey, Reston, Virginia.
- Hendricks Franssen, H. J., Alcolea, A., Riva, M., Bakr, M., van der Wiel, N., Stauffer, F., and Guadagnini, A. (2009). A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. *Advances in Water Resources*, 32(6):851–872.
- Hevesi, J. A., Istok, J. D., and Flint, A. L. (1992). Precipitation Estimation in Mountainous Terrain Using Multivariate Geostatistics. Part I: Structural Analysis. *Journal of Applied Meteorology*, 31(7):661–676.
- Huang, Y., Franssen, H. H., Herbst, M., Hirschi, M., Michel, D., Seneviratne, S. I., Teuling, A. J., Vogt, R., Detlef, S., Pütz, T., and Vereecken, H. (2020). Evaluation of different methods for gap filling of long-term actual evapotranspiration time series measured by lysimeters. *Vadose Zone Journal*, 19(1).

- Huysmans, M. and Dassargues, A. (2009). Application of multiple-point geostatistics on modelling groundwater flow and transport in a cross-bedded aquifer (Belgium). *Hydrogeology Journal*, 17(8):1901–1911.
- Irving, J. and Singha, K. (2010). Stochastic inversion of tracer test and electrical geophysical data to estimate hydraulic conductivities. *Water Resources Research*, 46(11).
- Jacquín, A. P. and Soto-Sandoval, J. C. (2013). Interpolation of monthly precipitation amounts in mountainous catchments with sparse precipitation networks. *Chilean journal of agricultural research*, 73(4):406–413.
- Jamshidi, A., Samani, J. M. V., Samani, H. M. V., Zanini, A., Tanda, M. G., and Mazaheri, M. (2020). Solving inverse problems of unknown contaminant source in groundwater-river integrated systems using a surrogate transport model based optimization. *Water (Switzerland)*, 12(9).
- Jose, S. C., Rahman, M. A., and Cirpka, O. A. (2004). Large-scale sandbox experiment on longitudinal effective dispersion in heterogeneous porous media. *Water Resources Research*, 40(12):1–13.
- Journel, A. G. and Huijbregts, C. J. (1978). *Mining Geostatistics*. Academic Press, New York.
- Kang, X., Shi, X., Deng, Y., Revil, A., Xu, H., and Wu, J. (2018). Coupled hydro-geophysical inversion of DNAPL source zone architecture and permeability field in a 3D heterogeneous sandbox by assimilation time-lapse cross-borehole electrical resistivity data via ensemble Kalman filtering. *Journal of Hydrology*, 567:149–164.
- Kang, X., Shi, X., Revil, A., Cao, Z., Li, L., Lan, T., and Wu, J. (2019). Coupled hydrogeophysical inversion to identify non-Gaussian hydraulic conductivity field by jointly assimilating geochemical and time-lapse geophysical data. *Journal of Hydrology*, 578.

- Khan, A., Sohail, A., Zahoor, U., and Qureshi, A. S. (2020). A survey of the recent architectures of deep convolutional neural networks. *Artificial Intelligence Review*, 53(8):5455–5516.
- Kitanidis, P. K. (1995). Quasi-Linear Geostatistical Theory for Inversing. *Water Resources Research*, 31(10):2411–2419.
- Kitanidis, P. K. (1997). *Introduction to Geostatistics: Applications in Hydrogeology*. Cambridge University Press, New York, NY.
- Lähivaara, T., Malehmir, A., Pasanen, A., Kärkkäinen, L., Huttunen, J. M., and Hesthaven, J. S. (2019). Estimation of groundwater storage from seismic data using deep learning. *Geophysical Prospecting*, 67(8):2115–2126.
- Lecun, Y., Bottou, E., Bengio, Y., and Haffner, P. (1998). Gradient-Based Learning Applied to Document Recognition. *Institute of Electrical and Electronics Engineers*.
- Li, J., Lu, W., Wang, H., and Fan, Y. (2019). Identification of groundwater contamination sources using a statistical algorithm based on an improved Kalman filter and simulation optimization. *Hydrogeology Journal*, 27(8):2919–2931.
- Li, L., Zhou, H., Hendricks Franssen, H. J., and Gómez-Hernández, J. J. (2012). Modeling transient groundwater flow by coupling ensemble Kalman filtering and upscaling. *Water Resources Research*, 48(1).
- Linde, N. and Doetsch, J. (2016). Joint Inversion in Hydrogeophysics and Near-Surface Geophysics. In *Integrated Imaging of the Earth*, chapter 7, pages 117–135. American Geophysical Union (AGU).
- Linde, N., Ginsbourger, D., Irving, J., Nobile, F., and Doucet, A. (2017). On uncertainty quantification in hydrogeology and hydrogeophysics. *Advances in Water Resources*, 110:166–181.

- Lloyd, C. (2005). Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. *Journal of Hydrology*, 308(1-4):128–150.
- Loke, M. H. (2018). Tutorial: 2-D and 3-D Electrical Imaging Surveys. Technical report, Universiti Sains Malaysia (USM).
- Ly, S., Charles, C., and Degré, A. (2011). Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrology and Earth System Sciences*, 15(7):2259–2274.
- Maina, F. H., Ackerer, P., Younes, A., Guadagnini, A., and Berkowitz, B. (2018). Benchmarking numerical codes for tracer transport with the aid of laboratory-scale experiments in 2D heterogeneous porous media. *Journal of Contaminant Hydrology*, 212:55–64.
- Matheron, G. (1970). La théorie des variables régionalisées et ses applications. In *Les cahiers du Centre de Morphologie Mathématique*, chapter Fascicule, page 212. ENSMP, Fontainebleau, France.
- Mavko, G., Mukerji, T., and Dvorkin, J. (2009). *The Rock Physics Handbook*. Cambridge University Press.
- Mawer, C., Knight, R., and Kitanidis, P. K. (2015). Relating relative hydraulic and electrical conductivity in the unsaturated zone. *Water Resources Research*, 51(1):599–618.
- McLaughlin, D. and Townley, L. R. (1996). A reassessment of the groundwater inverse problem.
- Michalak, A. M. and Kitanidis, P. K. (2004). Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling. *Water Resources Research*, 40(8).

- Mitsuhata, Y. (2004). Adjustment of regularization in ill-posed linear inverse problems by the empirical Bayes approach. *Geophysical Prospecting*, 52(3):213–239.
- Oriani, F., Stisen, S., Demirel, M. C., and Mariethoz, G. (2020). Missing Data Imputation for Multisite Rainfall Networks: A Comparison between Geostatistical Interpolation and Pattern-Based Estimation on Different Terrain Types. *Journal of Hydrometeorology*, 21(10):2325–2341.
- Page, L. M. (1969). The use of the geoelectric method for investigating geologic and hydrologic conditions in Santa Clara County, California. *Journal of Hydrology*, 7(2):167–177.
- Palop-Donat, C., Paredes-Arquiola, J., Solera, A., and Andreu, J. (2020). Comparing performance indicators to characterize the water supply to the demands of the Guadiana River basin (Spain). *Hydrological Sciences Journal*, 65(7):1060–1074.
- Panahi, M., Sadhasivam, N., Pourghasemi, H. R., Rezaie, F., and Lee, S. (2020). Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR). *Journal of Hydrology*, 588.
- Pereira, J. L., Gómez-Hernández, J. J., Zanini, A., Varouchakis, E. A., and Azevedo, L. (2023). Iterative geostatistical electrical resistivity tomography inversion. *Hydrogeology Journal*, 31(6):1627–1645.
- Pidlisecky, A. and Knight, R. (2008). FW2_5D: A MATLAB 2.5-D electrical resistivity modeling code. *Computers and Geosciences*, 34(12):1645–1654.
- Pollock, D. and Cirpka, O. A. (2010). Fully coupled hydrogeophysical inversion of synthetic salt tracer experiments. *Water Resources Research*, 46(7).
- Pollock, D. and Cirpka, O. A. (2012). Fully coupled hydrogeophysical inversion of a laboratory salt tracer experiment monitored by electrical resistivity tomography. *Water Resources Research*, 48(1).

- Portuguez-Maurtua, M., Arumi, J. L., Lagos, O., Stehr, A., and Montalvo Arquiniño, N. (2022). Filling Gaps in Daily Precipitation Series Using Regression and Machine Learning in Inter-Andean Watersheds. *Water*, 14(11):1799.
- Prudhomme, C. and Reed, D. W. (1999). Mapping extreme rainfall in a mountainous region using geostatistical techniques: a case study in Scotland. *International Journal of Climatology*, 19(12):1337–1356.
- Pyrz, M. J., Jo., H., Kuppenko, A., Liu, W., Gigliotti, A. E., Salomaki, T., and Santos, J. (2021). GeostatsPy Python Package, PyPI, Python Package Index.
- Rubin, Y. and Hubbard, S. (2005). *Hydrogeophysics*, volume 50 of *Water Science and Technology Library*. Springer Netherlands, Dordrecht.
- Samouëlian, A., Cousin, I., Tabbagh, A., Bruand, A., and Richard, G. (2005). Electrical resistivity survey in soil science: A review.
- Secchi, D., Molino, L., and Zanini, A. (2022). Contaminant source identification in groundwater by means of artificial neural network. *Journal of Hydrology*, 611.
- Sibson, R. (1981). A Brief Description of Natural Neighbor Interpolation. In Barnett, V., editor, *Interpreting Multivariate Data*, pages 21–36. John Wiley & Sons.
- Stein, M. L. (1999). *Interpolation of Spatial Data*. Springer Series in Statistics. Springer New York, New York, NY.
- Sternberg, S. P. K. (2004). Dispersion Measurements in Highly Heterogeneous Laboratory Scale Porous Media. *Transport in Porous Media*, 54:107–124.
- Tabios, G. Q. and Salas, J. D. (1985). A comparative analysis of techniques for spatial interpolation of precipitation. *JAWRA Journal of the American Water Resources Association*, 21(3):365–380.
- Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Philadelphia, PA.

- The MathWorks Inc. (2023). MATLAB Version 23.2.0 (R2023b).
- Thiessen, A. H. (1911). Precipitation averages for large areas. *Monthly Weather Review*, 39(7):1082–1089.
- Todaro, V., D’Oria, M., Secci, D., Zanini, A., and Tanda, M. G. (2022a). Climate Change over the Mediterranean Region: Local Temperature and Precipitation Variations at Five Pilot Sites. *Water*, 14(16):2499.
- Todaro, V., D’Oria, M., Tanda, M. G., and Gómez-Hernández, J. J. (2019). Ensemble smoother with multiple data assimilation for reverse flow routing. *Computers and Geosciences*, 131:32–40.
- Todaro, V., D’Oria, M., Tanda, M. G., and Gómez-Hernández, J. J. (2021). Ensemble smoother with multiple data assimilation to simultaneously estimate the source location and the release history of a contaminant spill in an aquifer. *Journal of Hydrology*, 598.
- Todaro, V., D’Oria, M., Tanda, M. G., and Gómez-Hernández, J. J. (2022b). genES-MDA: A generic open-source software package to solve inverse problems via the Ensemble Smoother with Multiple Data Assimilation. *Computers and Geosciences*, 167.
- Todaro, V., D’Oria, M., Zanini, A., Gómez-Hernández, J. J., and Tanda, M. G. (2023). Experimental sandbox tracer tests to characterize a two-facies aquifer via an ensemble smoother. *Hydrogeology Journal*, 31(6):1665–1678.
- Tso, C. H. M., Johnson, T. C., Song, X., Chen, X., Kuras, O., Wilkinson, P., Uhlemann, S., Chambers, J., and Binley, A. (2020). Integrated hydrogeophysical modelling and data assimilation for geoelectrical leak detection. *Journal of Contaminant Hydrology*, 234.
- Ulrych, T. J., Sacchi, M. D., and Woodbury, A. (2001). A Bayes tour of inversion: a tutorial. *Geophysics*, 66(1):55–69.

- Ursino, N., Gimmi, T., and Flühler, H. (2001). Combined effects of heterogeneity, anisotropy, and saturation on steady state flow and transport: A laboratory sand tank experiment. *Water Resources Research*, 37(2):201–208.
- Van Leeuwen, P. J. and Evensen, G. (1996). Data assimilation and inverse methods in terms of a probabilistic formulation. *Monthly Weather Review*, 124(12):2898–2913.
- Vereecken, H., Binley, A., Cassiani, G., Revil, A., and Titov, K. (2006). Applied hydrogeophysics. In Vereecken, H., Binley, A., Cassiani, G., Revil, A., and Titov, K., editors, *Applied Hydrogeophysics*, Dordrecht. Springer Netherlands.
- Walsh, J. B. and Brace, W. F. (1984). The effect of pressure on porosity and the transport properties of rock. *Journal of Geophysical Research*, 89(B11):9425–9431.
- Wilson, S. R., Ingham, M., and McConchie, J. A. (2006). The applicability of earth resistivity methods for saline interface definition. *Journal of Hydrology*, 316(1-4):301–312.
- Woodbury, A. D. and Ferguson, G. (2006). Ground surface paleotemperature reconstruction using information measures and empirical Bayes. *Geophysical Research Letters*, 33(6).
- Xu, T. and Gómez-Hernández, J. J. (2016). Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble Kalman filtering. *Water Resources Research*, 52(8):6587–6595.
- Xu, T., Gómez-Hernández, J. J., Chen, Z., and Lu, C. (2021). A comparison between ES-MDA and restart EnKF for the purpose of the simultaneous identification of a contaminant source and hydraulic conductivity. *Journal of Hydrology*, 595.
- Zanini, A., D’Oria, M., Tanda, M. G., and Woodbury, A. D. (2020). Coupling Empirical Bayes and Akaike’s Bayesian Information Criterion to Estimate Aquifer Transmissivity Fields. *Mathematical Geosciences*, 52(3):425–441.

- Zanini, A., Tanda, M. G., and Woodbury, A. D. (2017). Identification of transmissivity fields using a Bayesian strategy and perturbative approach. *Advances in Water Resources*, 108:69–82.
- Zanini, A. and Woodbury, A. D. (2016). Contaminant source reconstruction by empirical Bayes and Akaike’s Bayesian Information Criterion. *Journal of Contaminant Hydrology*, 185-186:74–86.
- Zhang, Z. (2018). Improved Adam Optimizer for Deep Neural Networks. In *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, pages 1–2.
- Zhao, Z., Berg, S. J., Illman, W. A., and Qi, Y. (2022). Improving predictions of solute transport in a laboratory sandbox aquifer through high-resolution characterization with hydraulic tomography. *Journal of Hydrology*, 615.
- Zheng, C. and Wang, P. P. (1999). MT3DMS: A modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems. Technical report, In Documentation and User’s Guide; U.S. Army Engineer Research and Development Center No. SERDP-99-1, Vicksburg, MS, USA.
- Zhou, H., Gómez-Hernández, J. J., and Li, L. (2012). A pattern-search-based inverse method. *Water Resources Research*, 48(3).
- Zhou, H., Gómez-Hernández, J. J., and Li, L. (2014). Inverse methods in hydrogeology: Evolution and recent trends.