



POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Data-Driven Enhancement of Ordinary Kriging for Uncertainty-Aware Aquifer Characterization in Thermal Energy Storage Systems

TESI DI LAUREA MAGISTRALE IN
ENERGY ENGINEERING-INGEGNERIA ENERGETICA

Author: Milad Mansouri

Student ID:	10699616
Advisor:	Dr. Ehsan Ranaee
Co-advisor:	Prof. Alberto Guadagnini
Academic Year:	2026-27

Abstract

Accurate characterization of hydraulic conductivity is essential for reliable simulation and design of Aquifer Thermal Energy Storage (ATES) systems, since groundwater flow and heat transport are strongly influenced by subsurface heterogeneity. This study proposes a data-driven framework combined with conventional Ordinary Kriging (OK) techniques aiming to enhance characterization of highly heterogeneous rock properties for numerical reservoir models and evaluates its impact on performance of ATES simulations.

Four techniques of generating hydraulic conductivity fields were investigated: (i) a homogeneous (reference) field, (ii) a classic OK generated field, and (iii) a hybrid field obtained by integrating Ordinary Kriging with Random Forest (OK–RF) and (iv) a hybrid field obtained by combining Ordinary Kriging with Gaussian Process Regression (OK–GPR). These models were incorporated into a coupled groundwater flow and heat transport simulation model developed in MATLAB Reservoir Simulation Toolbox (MRST). A low-temperature ATES system was then simulated for four mentioned reservoir models over a five-year operational period under identical operating conditions.

Proposed methodology was formulated within a maximum likelihood framework by assuming that OK prediction provides an estimator of unknown hydraulic conductivity. Under this assumption, cross-validation residuals were interpreted as estimation errors and subsequently modelled using machine-learning techniques to improve hydraulic conductivity characterization.

Results demonstrate that geological heterogeneity has a significant influence on groundwater flow, thermal plume evolution, well hydraulic behavior, and ATES energy recovery. Compared with the homogeneous model, heterogeneous hydraulic conductivity fields produced more complex thermo-hydraulic responses. Hybrid OK–RF and OK–GPR approaches further account for the representation of local-scale heterogeneity, among the investigated heterogeneous fields, OK–GPR representation produced the largest changes in the simulated thermo-hydraulic responses relative to the homogeneous reference model.

Our initiative framework demonstrates potential of combining geostatistical and machine-learning techniques to improve characterization of hydraulic conductivity fields and enhance the reliability of numerical reservoir simulation practices.

Key-words: Aquifer Thermal Energy Storage; Hydraulic Conductivity; Ordinary Kriging; Gaussian Process Regression; Machine Learning; Groundwater Flow and Heat Transport; MRST.

Abstract in lingua italiana

La caratterizzazione accurata della conducibilità idraulica è essenziale per una simulazione e una progettazione affidabili dei sistemi di Accumulo di Energia Termica in Acquifero (Aquifer Thermal Energy Storage, ATES), poiché il flusso delle acque sotterranee e il trasporto di calore sono fortemente influenzati dall'eterogeneità del sottosuolo. Questo studio propone un framework basato sui dati, combinato con le tecniche convenzionali di Kriging Ordinario (OK), con l'obiettivo di migliorare la caratterizzazione delle proprietà petrofisiche altamente eterogenee per i modelli numerici di serbatoio e di valutarne l'impatto sulle prestazioni delle simulazioni ATES.

Sono state investigate quattro tecniche per la generazione dei campi di conducibilità idraulica: (i) un campo omogeneo (di riferimento), (ii) un campo generato mediante Kriging Ordinario classico (OK), (iii) un campo ibrido ottenuto integrando il Kriging Ordinario con Random Forest (OK-RF) e (iv) un campo ibrido ottenuto combinando il Kriging Ordinario con Gaussian Process Regression (OK-GPR). Tali modelli sono stati implementati in un modello di simulazione accoppiato del flusso delle acque sotterranee e del trasporto di calore sviluppato nel MATLAB Reservoir Simulation Toolbox (MRST). Successivamente, un sistema ATES a bassa temperatura è stato simulato per i quattro modelli di serbatoio sopra citati lungo un periodo operativo di cinque anni, mantenendo condizioni operative identiche.

La metodologia proposta è stata formulata nell'ambito di un framework di massima verosimiglianza, assumendo che la previsione ottenuta mediante OK fornisca uno stimatore della conducibilità idraulica incognita. Sulla base di tale ipotesi, i residui della validazione incrociata sono stati interpretati come errori di stima e successivamente modellati mediante tecniche di apprendimento automatico al fine di migliorare la caratterizzazione della conducibilità idraulica.

I risultati dimostrano che l'eterogeneità geologica esercita un'influenza significativa sul flusso delle acque sotterranee, sull'evoluzione del pennacchio termico, sul comportamento idraulico dei pozzi e sul recupero energetico dei sistemi ATES. Rispetto al modello omogeneo, i campi eterogenei di conducibilità idraulica hanno prodotto risposte termo-idrauliche più complesse. Gli approcci ibridi OK-RF e OK-GPR hanno inoltre migliorato la rappresentazione dell'eterogeneità su scala locale; tra i campi eterogenei analizzati, la rappresentazione OK-GPR ha prodotto le variazioni più marcate nelle risposte termo-idrauliche simulate rispetto al modello omogeneo di riferimento.

Il framework proposto dimostra il potenziale della combinazione di tecniche geostatistiche e di apprendimento automatico per migliorare la caratterizzazione dei campi di conducibilità idraulica e aumentare l'affidabilità delle pratiche di simulazione numerica dei serbatoi.

Parole chiave: Aquifer Thermal Energy Storage (ATES); Conducibilità idraulica; Ordinary Kriging; Gaussian Process Regression; Apprendimento automatico; Flusso delle acque sotterranee e trasporto di calore; MRST.

Contents

Abstract.....	i
Abstract in lingua italiana	iv
Contents	1
1. Introduction	5
1.1 Background and Motivation	5
2. State of The Art.....	8
2.1 Large-Scale Energy Storage in Geological Formations	8
2.2 Rock Properties in Geological Reservoir Models.....	11
2.3 Numerical Characterization of Reservoir Rock Properties	13
2.3.1 Direct Field Measurements	13
2.3.2 Geostatistical Approaches.....	15
2.3.3 Machine Learning Approaches	17
2.3.4 Hybrid Approaches.....	20
2.4 Research Gap and Motivation	22
3. Methodology	25
3.1 Materials and Methods	25
3.2 Governing Equations of Flow and Heat Transport in Porous Media.....	26
3.2.1 Conservation of Mass	27
3.2.2 Conservation of Momentum: Darcy's Law	28
3.2.3 Conservation of Energy	29
3.2.4 Role of Hydraulic Conductivity in ATEs Systems	31
3.3 Numerical Implementation of ATEs Simulations in MRST.....	31
3.4 Variogram Modeling and Anisotropy Characterization	34

3.4.1	Spatial Dependence and Semivariogram.....	34
3.4.2	Theoretical Semivariogram.....	34
3.4.3	Experimental Semivariogram.....	35
3.4.4	Theoretical Variogram Models.....	36
3.4.5	Variogram Parameters.....	37
3.4.6	Anisotropy Characterization	37
3.5	OK for Hydraulic Conductivity Estimation	38
3.5.1	Introduction to OK Technique	38
3.5.2	Assumptions of OK Technique	39
3.5.3	OK Estimator.....	40
3.5.4	Unbiasedness Constraint.....	41
3.5.5	Minimization of Estimation Variance	41
3.5.6	OK System	42
3.5.7	Best Linear Unbiased Predictor.....	43
3.5.8	Kriging Variance and Prediction Uncertainty.....	43
3.6	Leave-One-Out Cross-Validation Technique	44
3.7	Maximum-Likelihood Interpretation of LOOCV Estimates	45
3.8	Discrepancy Decomposition	46
3.9	Spatial Modeling of Estimation Discrepancy	48
3.9.1	Gaussian Process Discrepancy Model.....	49
3.9.2	RF Discrepancy Model.....	52
3.10	Hybrid Prediction Framework	54
3.11	Uncertainty Quantification.....	56
3.12	Validation Criteria.....	57
3.12.1	Root Mean Squared Error	57
3.12.2	Mean Absolute Error	58
3.12.3	Coefficient of Determination	58
3.12.4	Kling–Gupta Efficiency	58
3.13	Methodological Contribution	59

4. Case Study	61
4.1 Introduction.....	61
4.2 Data Acquisition	62
4.3 Characterization of Hydraulic Conductivity Fields	67
4.3.1 Variogram Characterization	67
4.3.2 OK Results	70
4.3.3 Leave-One-Out Cross-Validation Residual Analysis.....	71
4.3.4 Residual Modelling Using Machine Learning	73
4.3.5 Hybrid Prediction and Final Evaluation.....	74
5. ATES Simulations.....	79
5.1 Introduction.....	79
5.2 ATES System Design and Operating Strategy	80
5.2.1 Operating Schedule.....	81
5.2.2 Well Configuration.....	82
5.2.3 Well Spacing Design Criteria.....	84
5.2.4 Well Operation.....	87
5.3 Hydraulic Conductivity Fields	88
5.3.1 Homogeneous Hydraulic Conductivity Field.....	89
5.3.2 Heterogeneous Hydraulic Conductivity Fields.....	90
5.4 Summary of Simulation Scenarios	92
5.5 Numerical Results	93
5.6 Thermal Plume Evolution	93
5.6.1 Thermal plume evolution during the first operating cycle.....	94
5.6.2 Thermal plume evolution after five operating cycles	95
5.7 Energy Recovery	97
5.7.1 Heat Recovery	97
5.7.2 Cold Recovery.....	99
5.8 Well Temperature Evolution	100
5.8.1 Hot Well Temperature Evolution	100

5.8.2	Cold Well Temperature Evolution.....	101
5.9	Well Hydraulic Performance	103
5.9.1	Hot Well Flow Rate Evolution.....	104
5.9.1	Cold Well Flow Rate Evolution.....	105
5.10	Bottom-Hole Pressure Evolution.....	107
5.10.1	Hot Well Bottom-Hole Pressure.....	107
5.10.2	Cold Well Bottom-Hole Pressure.....	108
5.11	Comparative Performance of the Hydraulic Conductivity Models.....	110
6.	CONCLUSIONS AND FUTURE WORK.....	112
6.1	Conclusion	112
6.2	Study limitations.....	113
6.3	Future Work	115
	Bibliography.....	117
	List of Figures.....	121
	List of Tables.....	125
2.	List of Symbols.....	127
3.	Acknowledgements	133

1. Introduction

1.1 Background and Motivation

Global transition toward sustainable and low-carbon energy systems has substantially increased demand for large-scale and long-duration energy storage technologies [1], [2]. Intermittent nature of renewable energy sources, such as solar and wind power, introduces significant temporal mismatches between energy production and energy demand. Consequently, energy storage technologies capable of storing large quantities of energy over seasonal timescales have become essential components of future energy infrastructures.

Among the various energy storage alternatives, underground geological formations have emerged as particularly attractive solutions because of their enormous storage capacities and their ability to accommodate energy over extended periods [3]. Geological formations are increasingly considered for underground hydrogen storage, carbon capture and storage, natural gas storage, geothermal systems, and Aquifer Thermal Energy Storage (ATES). The successful implementation of these technologies relies heavily on an adequate understanding of subsurface conditions and the accurate characterization of subsurface properties.

Despite their strategic importance, the characterization of geological formations remains one of the major sources of uncertainty in subsurface energy systems. Unlike engineered surface facilities, direct access to underground formations is inherently limited. Subsurface properties controlling subsurface flow and transport processes can only be inferred through measurements collected at discrete locations, such as boreholes and field investigations. Because these measurements are expensive and time-consuming to acquire, they are generally sparse relative to the scale and complexity of the subsurface systems under investigation. As a result, significant uncertainty remains regarding the spatial distribution of Subsurface properties.

Among the various reservoir properties, hydraulic conductivity plays a particularly important role because it governs groundwater movement and strongly influences advective transport of heat within aquifer systems. Consequently, uncertainty

associated with hydraulic conductivity may substantially affect predictions of fluid flow, thermal plume evolution, energy recovery efficiency, and the overall performance of subsurface energy storage systems. The reliable characterization of hydraulic conductivity fields therefore represents a fundamental requirement for the development of robust groundwater and energy storage models [4].

Traditional geostatistical approaches have long been employed to address this challenge. Ordinary Kriging (OK) has become one of the most widely used techniques for estimating hydraulic properties at unsampled locations owing to its ability to exploit spatial dependence and provide measures of prediction uncertainty [4], [5], [6]. However, Predictive performance of OK may be affected by the complexity and heterogeneity of geological formations. Highly heterogeneous aquifer systems often exhibit spatial patterns and localized behaviors that are difficult to represent fully through conventional geostatistical assumptions.

Recent advances in statistical learning and machine learning have created new opportunities for improving subsurface property characterization for numerical simulations. Data-driven techniques, particularly, offer capability to improve identifying complex structures and nonlinear relationships embedded within observed field data measurements and therefore provide potential avenues for complementing and enhancing traditional geostatistical predictions [7]. Nevertheless, their applicability to the study of subsurface systems requires careful consideration because geological variables are inherently spatial and associated with significant uncertainties.

Present study focuses on improving characterization of hydraulic conductivity fields for numerical models in highly heterogeneous aquifer systems, where sparse observations and unresolved geological complexity constitute major sources of predictive uncertainties. The central question addressed in this work is how different hydraulic conductivity characterization strategies influence the numerical representation of subsurface properties and how these differences propagate into predictions of ATEs performance. To investigate this question, OK is employed as primary estimator for spatial prediction of logarithmic-transformed hydraulic conductivity measurements collected from a highly heterogeneous aquifer system.

Within this study, OK is interpreted within a Maximum Likelihood framework by assuming that the OK prediction provides an estimator of the unknown hydraulic conductivity field. Under this assumption, the differences between the observed hydraulic conductivity values and their Leave-One-Out Cross-Validation (LOOCV) predictions are interpreted as estimation errors. These estimation errors are subsequently modelled using machine-learning techniques to develop alternative hydraulic conductivity representations while preserving the large-scale spatial

structure estimated by OK. In addition, OK is interpreted from a probabilistic perspective by viewing the hydraulic conductivity field as a realization of a Gaussian random field. Under this interpretation, OK can be regarded as a conditional estimator within a Gaussian Process (GP) framework, providing a statistical basis for modelling spatial variability and estimation uncertainty [8].

Finally, implications of reservoir characterization uncertainty are investigated within the context of ATEs systems. Since hydraulic conductivity governs both groundwater flow and advective heat transport, uncertainties in its characterization may propagate directly into predictions of thermal plume evolution, energy recovery efficiency, and overall system performance [3]. Consequently, different representations of hydraulic conductivity fields may lead to substantially different predictions of groundwater flow, thermal plume evolution, energy recovery efficiency, and overall ATEs behavior. Evaluating how model characterization uncertainty propagates through ATEs simulations is therefore essential for understanding reliability of predictive models and supporting informed decision-making in subsurface energy applications.

Rest of this thesis is organized as follows. Chapter 2 reviews current state of the art concerning geological energy storage, reservoir characterization techniques, geostatistical approaches, machine learning methods, and hybrid frameworks. Chapter 3 presents methodology adopted in this study, including geostatistical characterization procedures, probabilistic interpretation of OK, and advanced data-driven approaches investigated for improving hydraulic conductivity predictions. Chapter 4 introduces a case study considered in this work and datasets employed for analysis. Chapter 5 investigates implications of numerical reservoir characterization for ATEs simulations and discusses its influence on system performance. Finally, Chapter 6 summarizes principal findings of the study, highlights its limitations, and provides recommendations for future developments of the research.

2. State of The Art

2.1 Large-Scale Energy Storage in Geological Formations

Among the various technologies proposed for large-scale and long-duration energy storage, geological reservoirs have emerged as particularly promising solutions owing to their ability to store substantial amounts of energy over extended periods. Unlike conventional surface-based storage technologies, subsurface formations offer large storage capacities, seasonal storage capabilities, and potential to utilize existing geological resources. Consequently, geological energy storage has attracted increasing attention as a key component of future low-carbon energy systems [9], [10].

Conventional surface-based storage technologies, including electrochemical batteries, are generally well suited for short-duration storage applications. However, their economic and technical feasibility may become limited when very large storage capacities or seasonal storage durations are required. In contrast, geological reservoirs have emerged as promising alternatives because they can accommodate substantially larger quantities of stored energy over extended periods. Buursink et al. emphasized that underground geological formations are capable of storing energy at scales and durations that are difficult to achieve using conventional battery technologies, thereby making them attractive options for supporting ongoing energy transition [9].

Geological energy storage technologies can be broadly categorized according to form of energy being stored. Buursink et al. classified geological energy storage into three principal categories: chemical storage, mechanical storage, and thermal storage. Chemical storage includes underground storage of natural gas, methane, and hydrogen. Mechanical storage encompasses technologies such as compressed air energy storage and underground pumped hydro storage. Thermal storage involves storage and recovery of heat or cold within subsurface formations for subsequent use

in heating, cooling, industrial applications, or electricity generation. These storage technologies can be implemented in a variety of geological settings, including depleted oil and gas reservoirs, salt caverns, abandoned mines, boreholes, and freshwater or saline aquifers. Figure 2.1 presents a schematic overview of main geological energy storage technologies and their associated geological settings [9].

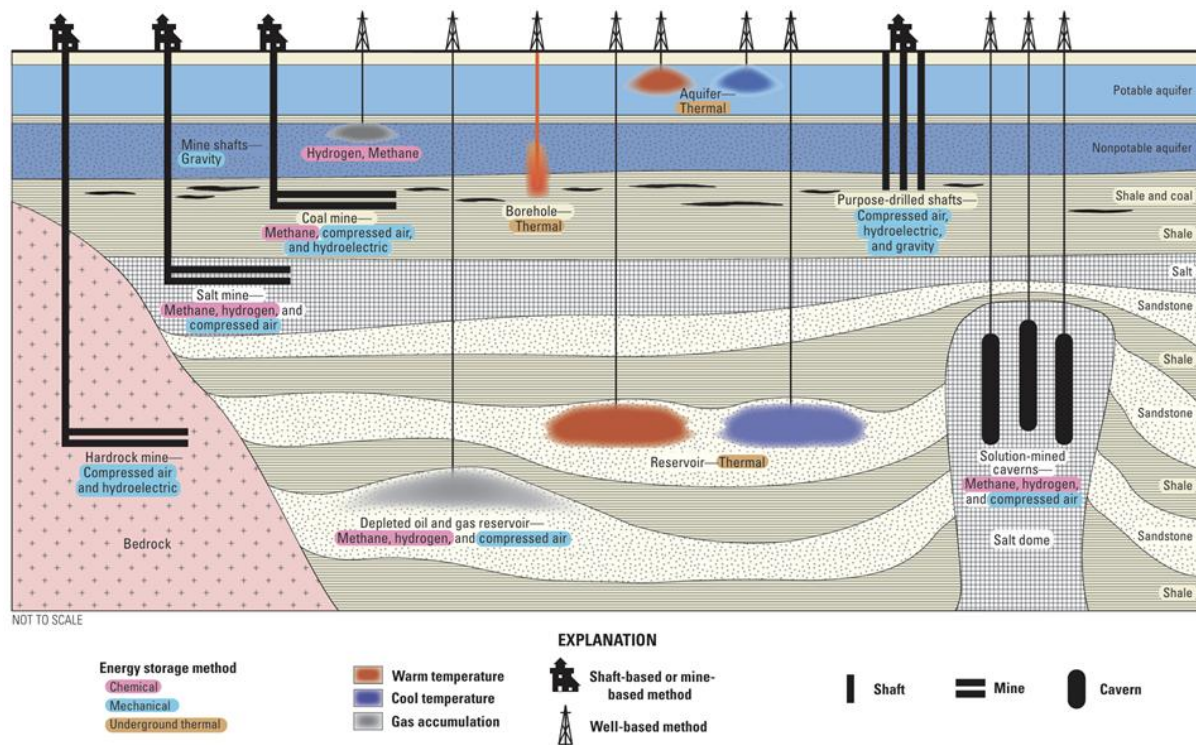


Figure 2.1 Geological energy storage technologies and their associated geological settings (adapted from Buursink et al.) [9].

Among the emerging applications of geological energy storage, underground hydrogen storage has attracted considerable attention owing to increasing recognition of hydrogen as a potential clean energy carrier. Hydrogen is increasingly recognized as a key energy carrier and has an important role in decarbonizing sectors that are difficult to electrify directly. Nevertheless, large storage capacities required for future hydrogen economies cannot be efficiently provided by surface facilities alone. Chen et al. noted that underground geological formations provide practical alternatives for long-duration and large-scale hydrogen storage [11]. Potential storage sites include depleted hydrocarbon reservoirs, salt caverns, and saline aquifers, although their suitability depends strongly on subsurface characteristics such as porosity, permeability, caprock integrity, storage depth, pressure conditions, and possible geochemical interactions.

Geological reservoirs are also fundamental components of several other energy and environmental applications. Carbon Capture and Storage (CCS) technologies rely on suitable subsurface formations to securely inject and retain carbon dioxide over geological timescales. Underground natural gas storage systems provide strategic reserves and help balance seasonal fluctuations in energy demand [9], [12]. Geothermal energy systems exploit natural thermal energy of the subsurface for heating and electricity generation. Similarly, Geological Thermal Energy Storage (GeoTES) and Aquifer ATEs systems utilize subsurface formations to store thermal energy during periods of surplus availability and recover it when required [10].

Particularly relevant to the present study is growing importance of geological thermal energy storage. Witter et al. identified GeoTES as a promising long-duration and grid-scale energy storage solution because subsurface reservoirs provide large storage volumes, seasonal storage capability, and operational flexibility [10]. Unlike many conventional storage technologies, geological thermal storage systems can often increase storage duration without proportional increases in surface infrastructure costs. Consequently, these systems are increasingly being considered as viable technologies for supporting renewable-based heating and cooling networks and enhancing the resilience of future energy systems.

Despite these advantages, successful implementation of geological energy storage technologies depends critically on accurate characterization of subsurface properties. Reservoir performance, injectivity, containment efficiency, thermal recovery, operational safety, and long-term sustainability are all strongly influenced by geological heterogeneity and hydrogeological parameters such as permeability, hydraulic conductivity, porosity, and thermal properties. Uncertainty associated with these parameters may significantly affect predictions of fluid flow, pressure propagation, gas migration, and thermal plume evolution [6], [13], [14].

Therefore, while geological reservoirs offer unique opportunities for large-scale and long-duration energy storage, their effective utilization requires reliable characterization of subsurface properties and their spatial variability. Developing robust methods to estimate and quantify uncertainty in these properties has consequently become a central topic in geological reservoir modeling and hydrogeological investigations [13], [14]. This necessity motivates the following section, which reviews the principal approaches used to characterize rock properties and develop geological reservoir models for energy storage applications.

2.2 Rock Properties in Geological Reservoir Models

Successful numerical simulations of geological energy storage systems depend fundamentally on the accurate characterization of subsurface properties governing fluid flow, mass transport, and heat transfer. The spatial distribution of these properties strongly influences predicted performance of geological energy storage systems, as well as their operational safety and long-term sustainability. Consequently, reliable characterization of rock properties constitutes a fundamental component of geological reservoir modelling.

Natural geological formations are inherently heterogeneous. Their properties result from complex depositional environments, diagenetic alterations, tectonic processes, weathering mechanisms, and fracture development occurring over geological timescales. As a consequence, reservoir properties often exhibit significant spatial variability across multiple scales, ranging from pore-scale heterogeneity to large-scale stratigraphic variations. Such heterogeneity influences movement of fluids and transfer of heat within subsurface and introduces substantial uncertainty into numerical predictions of reservoir behavior. Failure to adequately represent this variability may therefore lead to inaccurate assessments of storage capacity, injectivity, fluid migration pathways, pressure propagation, and thermal recovery performance [13], [14].

Among the various rock properties controlling reservoir performance, permeability and hydraulic conductivity are of particular importance because they directly govern fluid flow through porous media. Permeability is an intrinsic property of geological medium that describes its capacity to transmit fluids independently of fluid characteristics, whereas hydraulic conductivity additionally incorporates fluid properties such as density and viscosity. These parameters determine groundwater flow velocities, pressure responses to injection and extraction, and advective transport of dissolved constituents and thermal energy. Consequently, accurate characterization of permeability and hydraulic conductivity fields is essential for predicting behavior of geological energy storage systems and evaluating their operational performance [13], [15].

In aquifer-based energy storage systems, hydraulic conductivity assumes an especially critical role. In Aquifer Thermal Energy Storage systems, groundwater movement controls migration of thermal plumes and directly affects thermal recovery efficiency and long-term storage effectiveness. Similar considerations apply to other subsurface storage systems, where permeability influences injectivity, deliverability, and fluid migration. Therefore, uncertainty in hydraulic conductivity

may significantly affect predictions of energy recovery, storage efficiency, and operational reliability.

Porosity constitutes another fundamental reservoir property because it determines the amount of void space available for fluid storage within geological formation. Effective porosity, which represents interconnected pore volume contributing to fluid movement, influences both storage capacity and recoverability. In carbon capture and storage applications, porosity partly determines volume available for carbon dioxide sequestration. Likewise, in hydrogen and natural gas storage systems, porosity affects quantity of gas that can be accommodated within the reservoir. Consequently, accurate estimation of porosity distributions is essential for evaluating feasibility and capacity of geological storage projects [15], [16].

Thermal properties also play a crucial role in characterization of geological reservoirs, particularly in geothermal and thermal energy storage applications. Thermal conductivity governs ability of geological materials to conduct heat, while volumetric heat capacity determines the amount of thermal energy that can be stored within subsurface. Together, these properties control propagation of thermal fronts, development of thermal plumes, and the efficiency of heat recovery processes. In ATES systems, the interaction between groundwater flow and heat transport further emphasizes the importance of simultaneously characterizing both hydraulic and thermal properties to accurately predict subsurface thermal behavior [10], [17].

Reservoir integrity characteristics, including caprock permeability, fracture networks, and geomechanical stability, are also important in applications requiring long-term containment. These factors influence ability of geological formations to retain injected fluids and prevent leakage. Their consideration is particularly relevant for carbon sequestration and underground gas storage applications [16], [18].

Despite their importance, characterization of reservoir properties remains a major challenge because direct measurements are often limited in both number and spatial coverage. Information derived from core analyses, laboratory testing, pumping tests, slug tests, well logs, and field observations typically provides only sparse samples of highly heterogeneous systems. Furthermore, measurements obtained at one scale may not be directly transferable to another scale because of the scale-dependent nature of geological heterogeneity. Consequently, substantial uncertainty frequently accompanies the estimation of subsurface properties, complicating the development of reliable reservoir models [13], [14].

The need to account for geological heterogeneity while integrating limited observational data has therefore made reservoir characterization one of central challenges in hydrogeology and geological engineering. Robust approaches capable

of estimating spatially distributed rock properties and quantifying associated uncertainties are required to improve predictions of fluid flow, mass transport, and heat transfer in subsurface. Among various properties influencing reservoir behavior, hydraulic conductivity emerges as particularly important because it governs groundwater movement and advective heat transport within aquifer systems. Accordingly, accurate characterization of hydraulic conductivity fields represents a prerequisite for reliable assessment and optimization of geological energy storage systems. This necessity motivates the following section, which reviews principal methods used to estimate the spatial distribution of reservoir properties in heterogeneous geological environments.

2.3 Numerical Characterization of Reservoir Rock Properties

Characterization of reservoir properties represents a fundamental component of geological reservoir modeling because the performance of energy storage systems depends strongly on accurate representation of subsurface heterogeneity. However, direct observations of geological properties are typically limited in number and spatial extent, while the properties themselves often exhibit substantial variability across multiple scales [13], [14]. Consequently, a variety of approaches have been developed to estimate spatial distribution of reservoir properties and quantify associated uncertainty. These approaches range from direct field and laboratory measurements to geostatistical interpolation techniques and, more recently, data-driven methodologies. Each method provides distinct advantages and limitations with respect to data requirements, predictive capability, uncertainty representation, and computational complexity [6], [14]. The following subsections review principal approaches used for reservoir characterization, including direct measurement techniques, geostatistical methods, and emerging data-driven approaches, with particular emphasis on estimation of hydraulic conductivity fields in heterogeneous aquifer systems.

2.3.1 Direct Field Measurements

Direct measurements constitute primary source of information for characterization of reservoir properties and provide the observational basis for development, calibration, and validation of geological reservoir models. These methods aim to determine hydrogeological parameters directly from field observations or borehole investigations and are essential for quantifying properties such as hydraulic conductivity, transmissivity, storativity, and porosity. Because these parameters

govern fluid flow and transport processes in subsurface, their accurate estimation is fundamental for evaluating performance and safety of energy storage systems in geological formations.

Hydraulic testing techniques are among the most widely used direct methods in hydrogeological investigations. According to Woessner et al. hydraulic testing estimates hydrogeological parameters from observed hydraulic responses generated by controlled extraction or injection of water and represents a foundational component of groundwater characterization. In particular, hydraulic conductivity and transmissivity obtained from such tests are considered key parameters controlling groundwater flow behavior [19].

Among field-based hydraulic tests, pumping tests and slug tests are commonly employed. Pumping tests involve sustained groundwater withdrawal or injection while monitoring hydraulic head responses in one or more observation wells. Because they interrogate relatively large subsurface volumes, pumping tests provide field-scale estimates that are often considered representative of aquifer behavior. Slug tests, in contrast, consist of instantaneous addition or removal of a known water volume from a well and subsequent monitoring of water-level recovery. Owing to their simplicity, low operational cost, and minimal water requirements, slug tests are frequently used when pumping tests are impractical. However, because slug tests investigate relatively small portions of subsurface, they primarily reflect local hydraulic conditions and may be strongly influenced by small-scale heterogeneity [20]. The authors further noted that slug tests offer logistical and economic advantages and can resolve a higher degree of aquifer heterogeneity due to their limited sampling scale.

Additional characterization information can be obtained through geophysical well logging techniques. Unlike discrete hydraulic tests, well logs provide continuous measurements along borehole profiles and can be used to infer hydrogeological properties indirectly. Recent studies have demonstrated potential of geophysical logs for estimating hydraulic conductivity and improving parameterization of hydrogeological models, particularly in regions where conventional hydraulic data are scarce. Lora-Ariza et al. showed that well-log-derived hydraulic conductivity estimates can provide valuable information for characterization of heterogeneous multilayer aquifer systems and may constitute a practical alternative in data-scarce environments [21].

Despite their importance, direct measurement techniques exhibit several limitations. Drilling operations, field experiments, and laboratory analyses are often expensive and time-consuming, resulting in sparse and uneven spatial data coverage. Moreover, hydraulic parameter estimates are inherently scale-dependent, as different

measurement methods investigate different volumes of subsurface. Consequently, direct measurements alone are generally insufficient to describe full spatial variability and heterogeneity of geological reservoirs. Lora-Ariza et al. emphasized that most characterization methods provide information only at local scales, whereas hydrogeological models typically require continuous representations of reservoir properties over much larger domains [21].

Therefore, although direct measurements provide indispensable information regarding subsurface properties, challenge of extending these sparse observations throughout heterogeneous geological formations has motivated development of spatial characterization approaches. These methods seek to integrate available measurements to estimate reservoir properties at unsampled locations while accounting for spatial variability and uncertainty. principal spatial characterization approaches developed to address this challenge are reviewed in the following sections.

2.3.2 Geostatistical Approaches

Characterization of reservoir properties is often challenged by limited availability and uneven spatial distribution of direct measurements. Because acquiring hydrogeological observations through field campaigns and borehole investigations is generally expensive and time-consuming, measurements are commonly sparse relative to scale of subsurface systems being investigated. Consequently, methods capable of predicting reservoir properties at unsampled locations while accounting for spatial uncertainty are essential for development of reliable geological reservoir models.

Among available approaches, geostatistical methods have become fundamental tools for reservoir characterization because they explicitly exploit spatial dependence exhibited by geological variables. Underlying principle of geostatistics is that observations located closer to one another tend to be more similar than observations separated by larger distances. By quantifying this spatial continuity, geostatistical techniques enable estimation of subsurface properties throughout domain of interest while providing a framework for uncertainty assessment.

Geostatistical methods have been extensively applied to characterization of hydraulic conductivity, permeability, porosity, groundwater levels, and groundwater quality indicators in hydrogeological and reservoir studies. Panagiotou et al. (2022) emphasized that geostatistics provides a comprehensive framework for analyzing spatial variability, interpolating values at unsampled locations, integrating multiple sources of information, and propagating uncertainty in groundwater investigations.

Their review demonstrated the widespread use of kriging and geostatistical simulation techniques in addressing complex hydrogeological problems, highlighting their importance in groundwater management and characterization [22].

A key component of geostatistical analysis is characterization of spatial continuity through variogram modeling. Variograms describe degree of similarity between observations as a function of separation distance and direction and thereby provide basis for spatial prediction. Appropriate representation of spatial dependence allows geostatistical methods to reproduce important features of geological heterogeneity, including anisotropy and scale-dependent variability, which are often neglected by purely deterministic interpolation techniques.

Several geostatistical approaches have been developed for reservoir characterization. Among kriging-based methods, Ordinary Kriging is arguably the most widely applied approach in hydrogeological investigations. It has been extensively used for estimation of hydraulic conductivity, transmissivity, groundwater levels, and groundwater quality indicators because it provides statistically robust predictions while accounting for spatial dependence inferred from observed data. In addition, Ordinary Kriging provides estimation variance as a measure of prediction uncertainty, making it particularly valuable for risk-informed decision-making and subsurface characterization under sparse sampling conditions [4], [5], [6], [14], [22]. Other kriging variants, such as Simple Kriging, Universal Kriging, and Co-Kriging, have also been employed depending on assumptions regarding the mean structure of variable of interest and availability of auxiliary information.

In contrast to interpolation approaches that provide a single best estimate, geostatistical simulation techniques generate multiple equally probable realizations of subsurface. These realizations preserve short-scale variability and allow uncertainty associated with spatial predictions to be propagated through subsequent analyses. As a result, simulation-based approaches have become increasingly important in reservoir studies where uncertainty quantification and risk assessment are required.

Recent studies further demonstrate continuing importance of geostatistical approaches in subsurface investigations. Heße et al. highlighted that subsurface characterization is inherently affected by data scarcity and uncertainty and emphasized value of geostatistical information, such as variogram structures and correlation scales, for improving the representation of hydraulic properties and supporting uncertainty-informed analyses [23]. Similarly, Lino Pereira et al. showed that geostatistical techniques remain essential for groundwater characterization and uncertainty quantification, enabling construction of alternative geological scenarios and supporting informed decision-making in complex hydrogeological systems [24].

Despite their widespread use and rigorous statistical foundations, geostatistical approaches are subject to several limitations. Their performance depends strongly on specification of spatial continuity model, and assumptions such as stationarity may not always hold in highly heterogeneous geological environments. Furthermore, interpolation-based methods may smooth local variability and may not fully capture complex nonlinear relationships present in heterogeneous geological environments. These limitations have motivated increasing exploration of complementary data-driven approaches capable of identifying complex nonlinear relationships and residual structures that may not be fully represented by traditional geostatistical models. Such developments have contributed to the emergence of machine learning and hybrid approaches for reservoir characterization, which are discussed in the following sections.

2.3.3 Machine Learning Approaches

The increasing complexity and heterogeneity of geological systems have motivated adoption of machine learning (ML) techniques for reservoir characterization. Traditional approaches often rely on simplifying assumptions regarding the relationships among reservoir properties or statistical characteristics of data. However, subsurface environments are governed by intricate interactions among geological, hydrological, and environmental factors that frequently exhibit nonlinear behavior. In recent years, advances in computational power, data availability, and algorithm development have accelerated application of ML techniques in hydrogeology, geotechnical engineering, and reservoir studies. These approaches offer flexible frameworks capable of extracting hidden patterns from complex datasets and generating predictive models without requiring explicit assumptions regarding underlying functional relationships.

Machine learning methods are particularly attractive for subsurface investigations because direct measurements of reservoir properties are often sparse, expensive, and time-consuming to obtain. As a result, there is increasing interest in developing data-driven models capable of predicting reservoir properties from limited observations and readily available auxiliary information. Sheil et al. emphasized that artificial intelligence is transforming geotechnical and subsurface engineering by enabling extraction of knowledge from increasingly large and diverse datasets, thereby enhancing efficiency, reliability, and decision-making processes in engineering practice [25].

A wide variety of ML algorithms have been employed for characterization of geological reservoirs. Artificial Neural Networks (ANNs) are among the earliest and most widely used techniques. Inspired by the structure of biological neural systems,

ANNs are capable of approximating highly nonlinear relationships between input and output variables. Their flexibility has led to successful applications in prediction of soil hydraulic properties, pedotransfer function development, and other hydrogeological estimation problems. Agyare et al. noted that ANNs can effectively model complex input–output relationships and provide improved predictive performance compared with conventional empirical approaches when estimating saturated hydraulic conductivity [26].

Support Vector Machines (SVMs) have also been extensively used in hydrogeological applications because of their strong generalization capabilities and effectiveness when training datasets are relatively limited. Farasati et al. compared Random Forest (RF), SVM, and Least Squares Support Vector Machines (LSSVM) for predicting soil hydraulic conductivity and demonstrated that machine learning approaches can provide accurate estimates of hydraulic properties using readily available soil characteristics. Their study further indicated that RF yielded the highest predictive performance among evaluated models [27].

Among existing ML techniques, RF has emerged as one of the most popular algorithms for reservoir characterization. RF is an ensemble learning method that combines predictions of multiple decision trees constructed from bootstrap samples of training data. This strategy reduces overfitting and enhances predictive robustness while simultaneously providing measures of variable importance. Kalumba et al. employed RF and several alternative machine learning techniques to estimate hydraulic properties across Zambezi River Basin and showed that RF-based models effectively captured spatial variability using environmental covariates derived from climate data, topography, land use, and remote sensing products [28]. Similarly, Rezaei et al. demonstrated that RF outperformed several tree-based algorithms for predicting saturated hydraulic conductivity using limited soil observations combined with remotely sensed information, highlighting capability of RF to generate reliable spatial predictions from sparse datasets [29].

Another promising ML technique is Gaussian Process Regression (GPR), a Bayesian nonparametric approach that defines a probability distribution over functions [8]. Unlike many deterministic algorithms, GPR provides both predictive estimates and associated uncertainty intervals, making it particularly attractive for applications involving risk assessment and uncertainty-informed decision-making. Devi et al. compared ANN, Adaptive Neuro-Fuzzy Inference Systems (ANFIS), RF, and GPR for hydraulic conductivity prediction and reported that GPR achieved superior predictive accuracy among evaluated methods. Probabilistic nature of GPR further enables explicit quantification of predictive uncertainty, which is highly desirable in hydrogeological investigations [30].

Recent developments have expanded application of ML beyond the prediction of hydraulic properties alone. Tilahun and Korus (2023) utilized RF to construct high-resolution three-dimensional models of hydrostratigraphic units and hydraulic conductivity by integrating borehole records with airborne electromagnetic data. Their results demonstrated that supervised learning algorithms can successfully transform large geophysical datasets into hydrogeologically meaningful information and improve representation of heterogeneous aquifer systems [31]. In addition, Elmorsy et al. applied transfer learning and convolutional neural networks to predict effective thermal conductivity in digital porous media, illustrating growing role of deep learning techniques in geothermal and subsurface energy applications [32].

The growing popularity of ML approaches in reservoir characterization can be attributed to several advantages. First, these methods can capture complex nonlinear relationships that are difficult to describe using conventional statistical models. Second, they can simultaneously integrate large numbers of predictor variables derived from field observations, laboratory measurements, remote sensing products, and geophysical surveys. Third, algorithms such as RF provide information regarding predictor importance, thereby improving understanding of factors controlling reservoir behavior. Finally, ML methods can substantially reduce the need for extensive field campaigns by exploiting existing datasets to generate predictions over large spatial domains.

Despite these advantages, machine learning approaches also exhibit important limitations. Most algorithms are designed primarily to optimize predictive accuracy and therefore do not explicitly account for spatial autocorrelation, which is a defining characteristic of geological variables. Consequently, purely data-driven models may neglect important information related to spatial continuity and geological structure. Furthermore, some ML algorithms require relatively large training datasets, while highly complex models may suffer from reduced interpretability and increased computational demands. Uncertainty quantification also remains challenging for many ML methods, although probabilistic approaches such as GPR offer promising solutions.

These limitations suggest that machine learning methods should not necessarily be viewed as replacements for traditional geostatistical approaches. Rather, they may be regarded as complementary tools capable of identifying nonlinear patterns and extracting information that may not be fully represented by spatial interpolation techniques alone. Integration of geostatistics and machine learning has therefore emerged as an increasingly important direction in reservoir characterization research, combining strengths of both paradigms to improve predictive performance and uncertainty assessment in heterogeneous subsurface environments.

2.3.4 Hybrid Approaches

Increasing complexity and heterogeneity of subsurface systems have highlighted the limitations of relying exclusively on either geostatistical methods or machine learning techniques for reservoir characterization. Geostatistical approaches provide rigorous frameworks for representing spatial dependence and quantifying uncertainty, but their predictive capabilities may be constrained by assumptions regarding stationarity and specification of covariance structures. Conversely, machine learning methods excel at capturing nonlinear relationships from large and heterogeneous datasets but often neglect the spatial autocorrelation that is fundamental to geological variables. Consequently, recent research has increasingly focused on developing hybrid approaches that integrate strengths of both paradigms to improve predictive performance and uncertainty representation in subsurface investigations.

Hybrid geostatistical-machine learning methods seek to combine explicit representation of spatial dependence provided by geostatistics with ability of machine learning algorithms to capture complex nonlinear relationships from heterogeneous datasets. In these approaches, machine learning models are often employed to capture deterministic trends and nonlinear relationships among variables, while geostatistical techniques are subsequently used to model residual spatial dependence that remains unexplained by deterministic components. Such integration enables simultaneous consideration of complex predictor interactions and spatial continuity, thereby overcoming limitations associated with individual methodologies.

One of the earliest and most widely adopted hybrid frameworks is Regression Kriging (RK) [33]. In RK, a deterministic model is first developed using auxiliary variables, and geostatistical interpolation is then applied to the residuals of the deterministic predictions. This strategy exploits both explanatory power of environmental covariates and spatial structure contained within the residual component. The effectiveness of this approach has been demonstrated in numerous soil and groundwater studies, particularly where auxiliary information is available but direct measurements remain sparse.

Recent studies have extended this concept by replacing conventional regression models with more sophisticated machine learning algorithms. Kalumba et al. evaluated several digital soil mapping approaches for estimating soil hydraulic properties across Zambezi River Basin and complemented the deterministic predictions with residual kriging whenever spatial autocorrelation remained in the model residuals. Their results showed that combination of Random Forest and residual kriging consistently produced the highest predictive performance among

the evaluated methods. The authors concluded that integrating machine learning with geostatistical correction significantly improved spatial prediction of hydraulic properties while preserving valuable information related to residual spatial dependence [28].

Similarly, Mohammadpour et al. proposed a hybrid framework integrating Least Squares Support Vector Regression (LSSVR) with kriging for petrophysical property estimation. The methodology consisted of first constructing an LSSVR model to capture nonlinear relationships between auxiliary variables and the target variable, followed by kriging interpolation of the residuals to account for spatial variability. The resulting hybrid model outperformed standalone LSSVR, linear regression, and regression kriging approaches, demonstrating that simultaneous consideration of nonlinear dependencies and spatial autocorrelation can substantially enhance predictive accuracy. Importantly, the authors emphasized that conventional machine learning models are generally not designed to account for spatial autocorrelation and that ignoring this characteristic may lead to overly optimistic performance assessments [34].

Beyond residual kriging frameworks, researchers have explored alternative strategies for integrating geostatistics and machine learning. Fouedjio and Talebi introduced a geostatistical semi-supervised learning approach designed for situations where observations of the target variable are limited, while auxiliary spatial data are abundantly available. Their framework leveraged geostatistical conditional simulation to generate pseudo-labels at unlabeled locations based on spatial autocorrelation. These pseudo-labels were subsequently incorporated into machine learning training datasets, enabling simultaneous exploitation of labeled and unlabeled spatial information. The study demonstrated that combining geostatistical concepts with semi-supervised learning could effectively improve spatial prediction while accounting for uncertainty in the pseudo-labeling process [35].

The integration of deep learning and geostatistics has also emerged as an active research area. Yao et al. combined Deep Feedforward Neural Networks (DFNNs) with multipoint geostatistics (MPS) to improve reservoir facies modeling. Conventional two-point geostatistical approaches, which rely primarily on variograms, often struggle to reproduce complex geological structures characterized by higher-order spatial relationships. By incorporating neural networks capable of representing nonlinear dependencies into multipoint geostatistical framework, the authors achieved improved simulation of sedimentary facies exhibiting complex spatial patterns. Their findings illustrate potential of hybrid approaches to simultaneously preserve geological realism and exploit pattern-recognition capabilities of modern artificial intelligence algorithms [36].

Recent developments have further extended hybrid methodologies toward integration of geostatistics with advanced uncertainty quantification frameworks. Levy et al. proposed an approach for conditioning multiple-point geostatistical simulations to indirect geophysical observations, thereby enabling generation of geological realizations that simultaneously honor prior geological knowledge and indirect measurement information. Their work illustrates the growing emphasis on combining multiple information sources to improve characterization of complex subsurface systems while preserving uncertainty representations [37].

Despite these advances, several challenges remain. Many existing hybrid approaches rely on residual correction derived from deterministic predictions and may therefore be sensitive to the quality of the underlying models. In addition, the integration of spatial dependence, nonlinear predictive capabilities, and rigorous uncertainty quantification within a unified framework remains an active area of research. The diversity of existing methodologies indicates that no universally applicable solution has yet emerged for characterization of highly heterogeneous subsurface systems. Consequently, hybrid geostatistical–machine learning approaches continue to evolve rapidly and represent one of the most promising directions in contemporary reservoir characterization research.

2.4 Research Gap and Motivation

The literature reviewed in this chapter demonstrates substantial progress in characterization of reservoir properties using direct measurements, geostatistical techniques, machine learning algorithms, and hybrid methodologies. Direct measurement methods provide essential information regarding subsurface conditions but are inherently limited by their high cost, sparse spatial coverage, and inability to fully capture heterogeneity of geological formations. Geostatistical approaches, particularly kriging-based methods, have become widely accepted tools for estimating spatially distributed reservoir properties and quantifying uncertainty. Machine learning techniques have further expanded capabilities of reservoir characterization by enabling identification of complex nonlinear relationships from large and heterogeneous datasets. More recently, hybrid approaches integrating geostatistics and machine learning have shown promising results by combining spatial modeling capabilities with data-driven predictive power.

Despite these advances, several challenges remain unresolved. Hydraulic conductivity, one of the most influential parameters controlling groundwater flow and heat transport in aquifer systems, commonly exhibits strong spatial heterogeneity, nonlinearity, and scale-dependent variability. Traditional geostatistical approaches rely heavily on assumptions regarding spatial stationarity

and variogram structure, and their predictive performance may deteriorate when these assumptions are violated. Furthermore, interpolation-based methods often smooth local variability and may not adequately reproduce complex patterns associated with heterogeneous geological environments.

Machine learning approaches offer greater flexibility for capturing nonlinear relationships present in hydraulic conductivity data. However, most machine learning algorithms do not explicitly account for spatial autocorrelation, despite spatial dependence being a defining characteristic of hydrogeological variables. Consequently, purely data-driven models may fail to fully exploit spatial information contained within the observed data and may produce spatially inconsistent predictions. In addition, uncertainty quantification remains less straightforward for many machine learning techniques compared with geostatistical methods.

Although hybrid geostatistical-machine learning frameworks have emerged as promising alternatives, their application to hydraulic conductivity characterization remains relatively limited. Existing studies have largely focused on soil properties, petrophysical variables, or generalized spatial prediction problems, while comparatively fewer investigations have systematically evaluated hybrid approaches for spatial characterization of hydraulic conductivity in heterogeneous aquifer systems. Furthermore, there remains a need for comparative assessments that quantify the relative advantages of geostatistical, machine learning, and hybrid methods using same hydrogeological dataset and evaluation framework.

Therefore, a clear research gap exists in development and evaluation of integrated methodologies capable of simultaneously capturing nonlinear relationships, preserving spatial dependence, and reducing prediction errors under the adopted validation framework for hydraulic conductivity estimation in heterogeneous aquifers. Addressing this gap is particularly important for geological energy storage applications, where reliable characterization of hydraulic conductivity fields is essential for predicting groundwater flow, heat transport, storage performance, and long-term operational reliability.

Motivated by these challenges, present study develops and evaluates two hybrid residual-modelling frameworks that integrate Ordinary Kriging with Gaussian Process and RF models. The proposed OK-GPR and OK-RF approaches are designed to capture systematic patterns remaining in the residual component of OK predictions while preserving the underlying spatial structure of hydraulic conductivity fields. By comparing performance of OK, OK-GPR, and OK-RF, the study aims to evaluate how machine-learning-based residual modelling modifies hydraulic conductivity characterization relative to conventional OK and how these

alternative hydraulic conductivity representations influence predicted thermo-hydraulic response of numerical ATEs simulations.

3. Methodology

3.1 Materials and Methods

Principal objective of this chapter is to develop a data-driven framework for improving spatial characterization of logarithm of hydraulic conductivity, i.e., $\log K$, field for a highly heterogeneous aquifer systems through a hybrid geostatistical-machine learning framework.

The proposed methodology is motivated by fundamental role of hydraulic conductivity in governing groundwater flow and heat transport processes relevant to ATEs systems.

We first introduce governing equations describing saturated flow and heat transport in porous media, through principles of conservation of mass, conservation of momentum, and conservation of energy. Hydraulic conductivity is identified as a key parameter controlling groundwater movement and advective transport of thermal energy. Consequently, reliable estimation of hydraulic conductivity fields is essential for accurate prediction of subsurface thermal behavior and ATEs performance.

To characterize spatial variability of hydraulic conductivity, classical OK technique is employed as primary geostatistical interpolation method. Variogram analysis and anisotropy characterization are used to quantify spatial dependence structure of observed data and enable prediction at unsampled locations.

To evaluate predictive performance under genuine out-of-sample conditions, LOOCV is applied to OK estimator. Resulting prediction residuals are interpreted as empirical realizations of a discrepancy process representing systematic deviations that are not fully captured by the primary geostatistical model.

We treat OK estimation of $\log K$ values at LOOCV locations being interpreted within a Maximum-Likelihood framework by viewing $\log K$ field as a realization of a Gaussian random field. Under this interpretation, OK predictor corresponds to the

conditional mean estimator, or Best Linear Unbiased Predictor (BLUP), of Gaussian process (GP). Accordingly, Maximum-Likelihood framework provides a probabilistic basis for covariance parameter estimation and calibration and establishes connection between classical geostatistics and statistical inference.

structured component of ML estimation residuals is then modeled using two alternative discrepancy modeling approaches: (i) GPR to provide probabilistic discrepancy estimates together with associated uncertainty quantification, and (ii) RF technique to capture potentially complex nonlinear relationships embedded within the residual patterns without imposing explicit distributional assumptions.

Finally, estimated discrepancy fields are incorporated into original OK predictions to obtain bias-corrected representations of the logK field. The predictive performances of original and corrected models are subsequently evaluated using independent (so called test) validation borehole measurements and standard statistical performance metrics.

Our proposed methodology integrates governing physical principles, ML geostatistical estimation, cross-validation-based discrepancy identification, and machine learning correction within a unified framework for improved characterization of heterogeneous aquifer systems

3.2 Governing Equations of Flow and Heat Transport in Porous Media

Performance of ATEs systems is governed by coupled processes of groundwater flow and heat transport occurring within porous geological formations. These processes control movement of injected and extracted groundwater, evolution of thermal plumes, and ultimately efficiency of heat storage and recovery. Their mathematical formulation is based on fundamental principles of conservation of mass, conservation of momentum, and conservation of energy.

At the scale of practical groundwater and reservoir applications, porous media are commonly described using continuum approach. Under this assumption, heterogeneous geological formation is represented by an equivalent continuum characterized by averaged hydraulic and thermal properties. The validity of this approach relies on concept of Representative Elementary Volume (REV), defined as the smallest volume over which spatially averaged properties can be considered representative of porous medium. Once this condition is satisfied, groundwater flow and heat transport can be described using continuum-scale partial differential equations expressed in terms of effective medium properties [13], [38].

3.2.1 Conservation of Mass

The principle of conservation of mass states that rate of mass accumulation within a control volume must be equal to the net mass flux entering control volume plus any external sources or sinks. For saturated groundwater flow, integral form of continuity equation can be expressed as

$$\frac{\partial}{\partial t} \int_{\Omega} \rho \phi d\Omega + \int_{\partial\Omega} \rho \mathbf{v} \cdot \mathbf{n} dS = \int_{\Omega} \rho q d\Omega \quad (3.1)$$

where Ω denotes control volume, $\partial\Omega$ is the boundary of control volume, ρ is fluid density, ϕ is porosity, $\mathbf{v}$ is Darcy velocity, $\mathbf{n}$ is the outward unit normal vector, and q represents volumetric source or sink term.

Applying Gauss' divergence theorem to flux term gives

$$\int_{\partial\Omega} \rho \mathbf{v} \cdot \mathbf{n} dS = \int_{\Omega} \nabla \cdot (\rho \mathbf{v}) d\Omega \quad (3.2)$$

Since this relation must hold for any arbitrary control volume Ω , differential form of mass conservation equation is

$$\frac{\partial(\rho\phi)}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = \rho q \quad (3.3)$$

For slightly compressible fluids and porous media, changes in density and porosity resulting from pressure variations can be described using compressibility coefficients. Total compressibility of system is defined as

$$\beta_t = \beta_r + \beta_f \quad (3.4)$$

where β_r is rock compressibility and β_f is fluid compressibility. Combining continuity equation with Darcy's law,

$$\mathbf{v} = -\frac{k}{\mu} (\nabla p - \rho g \nabla z) \quad (3.5)$$

leads to pressure diffusion equation

$$\beta_t \phi \frac{\partial p}{\partial t} - \nabla \cdot \left[\frac{\rho k}{\mu} (\nabla p - \rho g \nabla z) \right] = \rho q \quad (3.6)$$

The above equation constitutes governing equation for single-phase saturated groundwater flow in porous media. It describes temporal evolution of pressure resulting from combined effects of fluid storage, pressure-driven flow, gravitational

forces, and source or sink terms, while highlighting fundamental role of permeability and hydraulic conductivity in controlling groundwater movement. Since hydraulic conductivity constitutes target variable of this study, its accurate spatial characterization is essential for reliable prediction of groundwater flow and associated heat transport processes.

3.2.2 Conservation of Momentum: Darcy's Law

The movement of groundwater through porous media is commonly described by Darcy's law, which serves as constitutive relationship linking fluid flow to driving hydraulic forces. At continuum scale, Darcy's law represents momentum balance for creeping flow conditions, where inertial effects are negligible compared with viscous forces. Under these assumptions, Darcy velocity is proportional to hydraulic potential gradient and inversely proportional to fluid viscosity [13], [39].

In terms of pressure, Darcy's law may be written as

$$\vec{v} = -\frac{k}{\mu}(\nabla p - \rho g \nabla z) \quad (3.7)$$

where v denotes specific discharge or Darcy velocity, k is intrinsic permeability tensor, μ represents dynamic viscosity of fluid, p is fluid pressure, ρ is fluid density, g denotes acceleration due to gravity, and z is elevation (vertical coordinate). In three-dimensional media, permeability is generally represented by a second-order tensor:

$$k = \begin{bmatrix} k_{xx} & k_{xy} & k_{xz} \\ k_{yx} & k_{yy} & k_{yz} \\ k_{zx} & k_{zy} & k_{zz} \end{bmatrix} \quad (3.8)$$

This tensor formulation allows anisotropic hydraulic behavior to be represented, reflecting directional dependence of permeability and groundwater flow within heterogeneous geological formations.

Hydraulic conductivity, which is more commonly used in hydrogeological applications, is related to intrinsic permeability through

$$K = \frac{k\rho g}{\mu} \quad (3.9)$$

where K is hydraulic conductivity, k is intrinsic permeability (or permeability tensor in anisotropic media), ρ is fluid density, g is acceleration due to gravity, and μ is dynamic viscosity of fluid.

The relationship between hydraulic conductivity and intrinsic permeability demonstrates that hydraulic conductivity depends both on geological properties of porous medium and on physical properties of fluid. While intrinsic permeability is solely a property of porous matrix, hydraulic conductivity also varies with fluid density and viscosity.

Consequently, permeability enters groundwater flow equations through hydraulic conductivity and exerts direct control over groundwater velocities. Accurate characterization of hydraulic conductivity fields is therefore essential for reliable prediction of groundwater flow and associated heat transport processes in aquifer systems.

3.2.3 Conservation of Energy

Since primary objective of ATEs systems is storage and recovery of thermal energy within subsurface aquifers, accurate representation of heat transport processes is essential. Conservation of energy principle states that rate of thermal energy accumulation within a control volume must be equal to net thermal energy entering control volume plus any internal heat sources or sinks. In saturated porous media, heat is transported by both advective, associated with groundwater movement, and conduction, driven by temperature gradients.

The general energy balance equation for saturated porous media can be expressed as [38]:

$$\frac{\partial}{\partial t}(M_f u_f + M_r u_r) + \nabla \cdot (\vec{F}_f h_f) + \nabla \cdot (\vec{H}_f + \vec{H}_r) = Q_f h_f \quad (3.10)$$

where M_f and M_r denote masses of fluid and solid phases, respectively, u_f and u_r represent internal energy per unit mass of fluid and rock, F_f is fluid mass flux, h_f is specific enthalpy of fluid, H_f and H_r denote conductive heat fluxes in fluid and solid phases, respectively, and Q_f represents fluid source and sink terms associated with injection and extraction processes.

The general energy balance equation indicates that change in thermal energy stored within porous medium results from combined effects of advective heat transport by groundwater flow, conductive heat transfer through fluid and rock matrix, and energy exchange associated with sources and sinks.

The internal energy of each phase may be approximated as

$$u_i = c_{p,i} T, \quad i \in \{f, r\} \quad (3.11)$$

where $c_{p,i}$ is specific heat capacity of phase i and T is temperature. Assuming local thermal equilibrium between solid and fluid phases, both phases are considered to share the same temperature at a given location.

The fluid enthalpy is related to internal energy by

$$h_f = u_f + \frac{p}{\rho_f} \quad (3.12)$$

where p is fluid pressure and ρ_f is fluid density.

Heat conduction within porous medium is commonly described using Fourier's law. Accordingly, conductive heat fluxes in rock and fluid phases are given by

$$\vec{H}_r = -k_r \nabla T \quad (3.13)$$

Similarly, conductive heat flux in fluid phase is given by

$$\vec{H}_f = -k_f \nabla T \quad (3.14)$$

where k_r and k_f are thermal conductivities of rock matrix and fluid, respectively.

Under assumptions of local thermal equilibrium and effective thermal properties, energy conservation equation can be simplified into advection-conduction form commonly employed in groundwater heat transport studies (Bear, 1972; Nield and Bejan, 2017):

$$(\rho c)_e \frac{\partial T}{\partial t} + \rho_f c_f \mathbf{v} \cdot \nabla T = \nabla \cdot (\lambda_e \nabla T) + Q_T \quad (3.15)$$

where $(\rho c)_e$ is effective volumetric heat capacity of saturated porous medium, ρ_f and c_f denote density and specific heat capacity of fluid, respectively, λ_e is effective thermal conductivity of porous medium, and Q_T represents thermal source or sink terms.

The simplified advection–conduction equation illustrates that groundwater velocity directly governs advective heat transport within the aquifer. Since groundwater velocity is determined by Darcy's law, hydraulic conductivity plays a fundamental role in controlling migration of thermal plumes and the overall thermal performance of ATEs systems. Consequently, accurate characterization of hydraulic conductivity fields is essential for reliable prediction of subsurface heat transport and thermal energy storage behavior.

3.2.4 Role of Hydraulic Conductivity in ATEs Systems

The governing equations presented in the previous sections highlight the importance of hydraulic conductivity in simulation and performance assessment of ATEs systems. As a key parameter describing the hydraulic properties of the subsurface, hydraulic conductivity influences groundwater flow patterns and consequently affects transport and distribution of thermal energy within the aquifer. Since heat advection is directly linked to groundwater movement, spatial variability of hydraulic conductivity exerts a strong influence on evolution of thermal plumes, timing of thermal breakthrough, and efficiency of energy storage and recovery.

Because natural aquifers are inherently heterogeneous, inaccurate characterization of hydraulic conductivity may result in substantial uncertainties in numerical predictions of ATEs performance. Reliable estimation of hydraulic conductivity fields is therefore essential for developing robust groundwater flow and heat transport models capable of supporting the design, evaluation, and optimization of ATEs systems.

Motivated by this critical dependence, the following sections present the methodology adopted to provide an alternative spatial characterization of hydraulic conductivity through integration of geostatistical techniques and machine learning approaches in heterogeneous aquifer systems.

3.3 Numerical Implementation of ATEs Simulations in MRST

In this study, MATLAB Reservoir Simulation Toolbox (MRST) was employed to simulate groundwater flow and heat transport within ATEs system. MRST provides a numerical framework for solving governing equations presented in Section 3.2 and for evaluating influence of spatially varying hydraulic conductivity on ATEs performance.

This open-source numerical tool offers several modules useful for geological applications, in particular for geothermal modelling in which flow and heat transport equations are incorporated, discretized and solved over simulation period. It also supports complex grid geometries, control and operational strategies for injection and production scenarios, in addition to efficient linear and nonlinear solvers [40].

MRST employs finite-volume discretization method (FVM), two-point flux approximation (TPFA), and implicit Euler scheme to solve governing equations introduced in the previous section using automatic differentiation (AD) techniques (Lie, 2016). Moreover, to obtain faster simulations, MRST offers MEX accelerated AD

backend and a compiled iterative Krylov solver with constrained pressure residual (CPR) and algebraic multigrid (AMG) preconditioning [40].

Using the implicit backward Euler discretization in time, governing equations presented in Section 3.2 are transformed into a system of discrete equations. Discretized mass conservation and Darcy flow equations are expressed, respectively, as:

$$\frac{(\phi\rho)^{n+1} - (\phi\rho)^n}{\Delta t^n} + \text{div}(\rho v)^{n+1} = q^{n+1} \quad (3.16)$$

and

$$v^{n+1} = -\frac{k}{\mu^{n+1}} [\text{grad}(p^{n+1}) - g\rho^{n+1}\text{grad}(z)] \quad (3.17)$$

where div and grad are divergence and gradient discrete operators respectively.

The numerical tool assumes homogeneous Neumann conditions for no flow external boundaries. Presence of a natural flow can be reflected by considering a hydrostatic condition on two parallel facies in simulation prototype, known as a Dirichlet boundary condition:

$$p(\vec{x}) = p_a(\vec{x}, t) \quad \text{for } \vec{x} \in \Gamma_a \subset \partial\Omega \quad (3.18)$$

where Γ_a represents the portion of the domain boundary on which the prescribed pressure condition is applied.

The exchange of fluid between aquifer and wells is represented using well model implemented in MRST. For a fully penetrating vertical well, well-flow rate may be expressed as

$$q = \frac{2\pi h c k_e}{\mu B \left[\ln\left(\frac{r_e}{r_w}\right) - 0.75 + S \right]} (p_R - p_{bh}) \quad (3.19)$$

where q denotes well flow rate, c is completion factor representing open fraction of wellbore, $k_e = \sqrt{k_x k_y}$ is effective permeability, subsequently defined in terms of principal permeability components k_x and k_y , h is reservoir thickness, μ is fluid viscosity, B is formation volume factor, r_e is equivalent drainage radius, r_w is well radius, S is skin factor, p_R is reservoir pressure, and p_{bh} is bottom-hole pressure.

For anisotropic permeability fields, Peaceman (1983) reformulated equivalent radius to account for directional permeability variations. The equivalent radius is expressed as

$$r_e = 0.28 \frac{\left(\sqrt{k_y/k_x} \Delta x^2 + \sqrt{k_x/k_y} \Delta y^2 \right)^{\frac{1}{2}}}{\left(k_y/k_x \right)^{\frac{1}{4}} + \left(k_x/k_y \right)^{\frac{1}{4}}} \quad (3.20)$$

where k_x and k_y denote principal permeability components, and Δx and Δy represent grid dimensions in horizontal directions [41].

Incorporating anisotropic equivalent radius into well model yields:

$$q = \frac{2\pi h c k_e}{\mu B \left[\ln \left(\frac{r_e}{r_w} \right) + S \right]} (p_R - p_{bh} - \rho(z - z_{bh})g) \quad (3.21)$$

where z denotes elevation (vertical coordinate), z_{bh} is elevation of bottom-hole pressure reference point, and $\rho g(z - z_{bh})$ represents hydrostatic pressure difference between the reservoir cell and well reference depth.

In addition, implicit backward Euler discretization in time method is employed for stability in simulations that are time dependent [40]. Applying this method to energy equation (3.10), latter can be represented in residual form as follows:

$$R^{n+1} = \frac{1}{\Delta t^n} \left([M_f u_f + M_r u_r]^{n+1} - [M_f u_f + M_r u_r]^n \right) + \nabla \cdot [\vec{F}_f h_f]^{n+1} + \nabla \cdot [\vec{H}_f + \vec{H}_r]^{n+1} - \nabla \cdot [Q_f h_f]^{n+1} = 0 \quad (3.22)$$

where the superscript n denotes time t^n , and timestep referred to as $\Delta t^n = t^{n+1} - t^n$.

Finally, it is important to note that default units in MRST are SI units.

The numerical framework described above requires a spatially distributed hydraulic conductivity field as input. Because hydraulic conductivity measurements were available only at discrete borehole locations, geostatistical and machine-learning approaches were employed to estimate the three-dimensional hydraulic conductivity distribution. The methodology adopted for hydraulic conductivity field generation is presented in following sections.

3.4 Variogram Modeling and Anisotropy Characterization

3.4.1 Spatial Dependence and Semivariogram

A fundamental assumption underlying geostatistical analysis is that observations located close to one another in space tend to exhibit greater similarity than observations separated by larger distances. This concept, commonly referred to as spatial dependence or spatial autocorrelation, forms the basis of spatial prediction methods such as OK [5].

In hydrogeological systems, spatial distribution of hydraulic properties is controlled by geological processes, depositional environments, and sedimentary structures. Consequently, hydraulic conductivity is rarely distributed randomly in space. Instead, neighboring locations often exhibit similar values because they are influenced by the same geological conditions. Quantifying this spatial continuity is therefore essential for reliable interpolation and uncertainty assessment.

The semivariogram is principal tool used in geostatistics to characterize spatial dependence. It describes how dissimilarity between observations changes as the separation distance increases and provides mathematical foundation for calculating kriging weights and prediction uncertainty [6], [42].

In the present study, semivariogram was employed to characterize spatial continuity of logK and to provide the basis for OK interpolation procedure.

3.4.2 Theoretical Semivariogram

Spatial dependence of a regionalized variable can be characterized through semivariogram, which describes how the similarity between observations changes as the separation distance increases. For a regionalized variable $Z(x)$, theoretical semivariogram is defined as

$$\gamma(h) = \frac{1}{2} E\{[Z(x+h) - Z(x)]^2\} \quad (3.23)$$

where $\gamma(h)$ denotes semivariance associated with separation vector h , $E\{.\}$ represents expectation operator, and $Z(x)$ and $Z(x+h)$ denote the values of the regionalized variable at two locations separated by vector h .

Theoretical semivariogram defines semivariance as one-half of expected squared difference between observations separated by a given distance and direction.

Consequently, semivariogram provides a quantitative measure of spatial continuity. Small semivariance values indicate that neighboring observations tend to be similar and therefore strongly correlated, whereas large semivariance values imply weaker spatial dependence and greater dissimilarity between observations.

Under intrinsic hypothesis, semivariogram depends only on separation vector (h) and is independent of absolute spatial positions of the observations. This assumption enables spatial structure of the regionalized variable to be described using distance-based relationships and forms theoretical foundation for variogram modeling and kriging interpolation (Journel and Huijbregts, 1978).

3.4.3 Experimental Semivariogram

Although theoretical semivariogram provides a mathematical description of spatial dependence, it cannot be evaluated directly because true probability distribution of the regionalized variable is generally unknown. Therefore, semivariogram must be estimated empirically from the available observations [43]. Resulting estimator is referred to as experimental semivariogram and is expressed as

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (3.24)$$

where $\hat{\gamma}(h)$ denotes estimated semivariance corresponding to lag distance h , $N(h)$ is number of observation pairs belonging to lag class, and $Z(x_i)$ and $Z(x_i + h)$ represent paired observations separated by lag vector h .

Experimental semivariogram estimator calculates the average half-squared difference between pairs of observations separated by similar distances. By repeating this calculation for a sequence of lag intervals, an empirical representation of the spatial continuity structure of the regionalized variable is obtained. In practice, exact separation distances are rarely encountered because sampling locations are generally irregularly distributed. Consequently, pairs of observations are grouped into predefined lag classes using specified distance tolerances.

In accordance with structural analysis approach described by Chilès and Delfiner (1999), experimental semivariogram was constructed by grouping point pairs according to both separation distance and direction. Because hydraulic conductivity field may exhibit directional dependence, horizontal and vertical semivariograms were computed separately. Horizontal semivariograms were estimated from observation pairs located at approximately similar elevations, whereas vertical semivariograms were calculated using observation pairs belonging to the same borehole. All valid within-borehole pairs were considered and subsequently

combined across the entire borehole dataset. This approach minimizes horizontal separation in vertical variogram and allows estimated semivariance to primarily reflect vertical variability of logK field. Number of pairs $N(h)$ in each lag class was also recorded, since the reliability of each experimental semivariogram point depends on number of contributing pairs. Variogram estimates supported by a larger number of pairs are generally considered more statistically reliable than estimates based on only a few observations. Resulting horizontal and vertical experimental semivariograms were subsequently used for theoretical variogram fitting and anisotropy characterization [44].

3.4.4 Theoretical Variogram Models

Because experimental semivariogram consists of discrete empirical estimates, it cannot be directly employed within OK equations. Kriging requires a continuous and mathematically valid function capable of describing spatial dependence structure of the regionalized variable over entire range of separation distances. Therefore, a theoretical variogram model must be fitted to experimental semivariogram.

Several theoretical models have been proposed in geostatistical literature, among which spherical, exponential, and Gaussian models are the most commonly used [6], [42], [43]. Selection of an appropriate model depends on characteristics of experimental semivariogram and ability of the theoretical function to reproduce observed pattern of spatial continuity.

For the present study, spatial continuity was represented using exponential variogram model, which may be expressed as

$$\gamma(h) = N_0 + \sigma^2(1 - e^{-h/l}) \quad (3.25)$$

where $\gamma(h)$ is fitted theoretical semivariogram, h is lag distance, N_0 is nugget effect, σ^2 is partial sill (structured variance contribution), and l represents exponential length scale controlling rate of decay of spatial correlation. sill of variogram is given by $N_0 + \sigma^2$.

Unlike spherical models, exponential variogram approaches sill asymptotically and therefore do not possess a finite theoretical range. Consequently, practical range is commonly approximated as three times exponential length scale, corresponding to distance at which approximately 95% of sill is reached [42].

Regardless of mathematical form adopted, theoretical variogram models are generally characterized by three principal parameters: nugget, sill, and range. These parameters summarize magnitude and scale of spatial dependence and directly influence kriging weights and prediction uncertainty obtained during interpolation.

In the present study, exponential variogram models were adopted to characterize both horizontal and vertical spatial continuity and to derive anisotropy parameters used in the three-dimensional OK implementation. Corresponding variogram parameters were estimated by fitting exponential models to experimental semivariograms. Separate exponential variogram models were fitted to the horizontal and vertical experimental semivariograms, yielding directional length scales l_H and l_V used for anisotropy characterization. The principal characteristics of the fitted variogram models are described by nugget, sill, and range parameters discussed in the following section.

3.4.5 Variogram Parameters

Theoretical variogram models are commonly characterized by three principal parameters: nugget, sill, and range. These parameters describe magnitude and scale of spatial dependence and directly influence kriging weights and prediction uncertainty obtained during interpolation.

The nugget effect represents semivariance observed at separation distances approaching zero and may arise from measurement errors, microscale variability, or geological heterogeneity occurring at scales smaller than the sampling interval. Sill corresponds to value at which the variogram stabilizes and reaches a plateau, approximating overall variance of the regionalized variable. Range defines distance beyond which observations become effectively uncorrelated. Consequently, observations separated by distances shorter than range tend to exhibit spatial dependence, whereas those separated by larger distances may be regarded as effectively uncorrelated.

Together, these parameters determine covariance structure employed by OK estimator and therefore play a fundamental role in characterization of spatial continuity.

3.4.6 Anisotropy Characterization

Anisotropy refers to condition in which spatial continuity varies according to direction. In hydrogeological systems, anisotropic behavior is frequently observed because geological processes often generate formations with preferential orientations and layered structures. As a result, degree of spatial correlation may differ substantially between directions [42].

In the present study, anisotropy was evaluated through directional variogram analysis. Separate experimental semivariograms were computed for horizontal and vertical directions and subsequently fitted using exponential variogram models. For

exponential variogram model, practical range is conventionally approximated as three times exponential length scale.

The practical ranges associated with the fitted models were determined as

$$Range_H = 3l_H \quad (3.26)$$

and

$$Range_V = 3l_V \quad (3.27)$$

where l_H and l_V denote horizontal and vertical exponential length scales, respectively.

The degree of anisotropy was quantified using anisotropy ratio defined as

$$AR = \frac{Range_H}{Range_V} \quad (3.28)$$

where AR denotes anisotropy ratio, $Range_H$ represents range associated with horizontal direction, and $Range_V$ denotes range associated with vertical direction.

The resulting anisotropy ratio was subsequently incorporated into three-dimensional OK implementation through coordinate transformation. Specifically, vertical coordinate was scaled according to

$$z_{scaled} = z \times AR \quad (3.29)$$

where z denotes original vertical coordinate. When horizontal range exceeds vertical range, scaling increases vertical distances relative to horizontal distances, thereby reducing influence of observations separated by large vertical intervals. This transformation modifies distance calculations used during kriging interpolation so that difference between horizontal and vertical correlation lengths is properly represented. Consequently, three-dimensional interpolation procedure accounts for anisotropic spatial structure of the hydraulic conductivity field.

3.5 OK for Hydraulic Conductivity Estimation

3.5.1 Introduction to OK Technique

Characterizing spatial distribution of hydraulic conductivity is a fundamental challenge in hydrogeology because measurements are generally available only at discrete sampling locations, such as boreholes and monitoring wells. However, groundwater flow and heat transport simulations require continuous representations

of subsurface properties over entire domain of interest. Consequently, reliable interpolation methods are required to estimate hydraulic conductivity at unsampled locations while preserving spatial structure of the observed data.

Geostatistics provides a mathematical framework for addressing this problem by incorporating both magnitude and spatial correlation of observations into estimation process. Among the various geostatistical interpolation techniques, OK is one of the most widely applied methods in hydrogeological investigations because it provides unbiased predictions with minimum estimation variance [5], [6].

Unlike deterministic interpolation techniques, such as inverse distance weighting, OK explicitly accounts for spatial autocorrelation through the variogram model. Interpolation weights are therefore determined not only by distance between observations and estimation location but also by spatial continuity structure of the variable. As a result, OK generally produces more realistic estimates in heterogeneous geological environments [42], [43].

OK belongs to the family of linear geostatistical estimators and assumes that the variable under investigation can be represented as a regionalized variable exhibiting spatial dependence. The method estimates the value of the variable at an unsampled location using neighboring observations while minimizing the associated prediction uncertainty.

3.5.2 Assumptions of OK Technique

The formulation of OK relies on several assumptions regarding statistical properties of the regionalized variable. Let spatial variable be denoted by $Z(x)$

where (x) represents spatial coordinates of a location within study domain. OK assumes that the regionalized variable can be decomposed into two components according to

$$Z(x) = \mu_{OK} + \varepsilon(x) \quad (3.30)$$

where μ_{OK} represents an unknown but locally constant mean and $\varepsilon(x)$ denotes a zero-mean random component characterized by spatial dependence. This formulation implies that, although exact mean value of variable is not known, it may reasonably be assumed to remain constant within local neighborhood used for interpolation.

Under this framework, dependence between observations is assumed to be governed solely by their spatial separation rather than their absolute positions within the domain. Consequently, covariance structure is considered stationary within

neighborhood of estimation, and the regionalized variable is assumed to possess finite variance. These assumptions enable characterization of spatial continuity through variogram and provide theoretical basis for deriving OK estimator [6], [42].

Although these assumptions are generally appropriate for many hydrogeological applications, natural geological systems rarely satisfy them perfectly. In particular, highly heterogeneous formations may exhibit non-stationary behavior resulting from complex depositional environments, abrupt lithological transitions, or large-scale trends. Nevertheless, when interpolation is restricted to relatively homogeneous local neighborhoods, assumptions of OK often provide a reasonable approximation of underlying spatial structure and support reliable estimation.

3.5.3 OK Estimator

Objective of OK is to estimate value of a regionalized variable at an unsampled location using information obtained from neighboring observations. Consider an estimation location denoted by x_0 . OK estimator is expressed as

$$\hat{Z}_{OK}(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (3.31)$$

OK estimator is derived as the Best Linear Unbiased Estimator under the assumption of a locally constant mean [6], where $\hat{Z}_{OK}(x_0)$ denotes estimated value at location x_0 , $Z(x_i)$ represents the observed value at neighboring location x_i , λ_i is kriging weight assigned to observation i , and n corresponds to number of observations used in interpolation process.

OK estimator indicates that prediction is obtained as a weighted average of neighbouring observations. However, unlike deterministic interpolation techniques in which weights are prescribed according to simple distance relationships, kriging weights are determined through an optimization procedure that explicitly accounts for spatial dependence structure of the regionalized variable. Consequently, contribution of each observation depends not only on its proximity to estimation location but also on degree of redundancy and correlation among neighboring observations. This characteristic enables OK to exploit spatial information more efficiently and generally leads to more reliable predictions in heterogeneous environments.

3.5.4 Unbiasedness Constraint

A fundamental requirement of OK is that the estimator must be unbiased. This condition implies that the expected value of estimation error is equal to zero.

The estimation error may be expressed as

$$e(x_0) = \hat{Z}_{OK}(x_0) - Z(x_0) \quad (3.32)$$

and unbiasedness condition requires

$$E[e(x_0)] = 0 \quad (3.33)$$

Substituting OK estimator into above expression yields

$$E[\hat{Z}_{OK}(x_0) - Z(x_0)] = 0 \quad (3.34)$$

Under the assumption of a locally constant mean, unbiasedness condition leads to following constraint on the kriging weights:

$$\sum_{i=1}^n \lambda_i = 1 \quad (3.35)$$

This condition guarantees that estimator neither systematically overestimates nor underestimates true value of the variable [6].

3.5.5 Minimization of Estimation Variance

Among all linear unbiased estimators, OK seeks the one associated with minimum estimation variance.

The kriging variance is defined as

$$\sigma_{OK}^2(x_0) = Var[\hat{Z}_{OK}(x_0) - Z(x_0)] \quad (3.36)$$

Objective of OK procedure is therefore to determine set of weights that minimizes $\sigma_{OK}^2(x_0)$, while simultaneously satisfying unbiasedness constraint.

Minimizing estimation variance is particularly important in hydrogeological applications because it ensures that resulting predictions are associated with lowest possible uncertainty under assumptions of the adopted variogram model [6].

3.5.6 OK System

The constrained optimization problem is solved using method of Lagrange multipliers. By introducing Lagrange multiplier μ_K , objective function can be written as

$$L = \sigma_{OK}^2(x_0) + \mu_K \left(\sum_{i=1}^n \lambda_i - 1 \right) \quad (3.37)$$

where second term enforces unbiasedness constraint of OK estimator. Minimization of Lagrangian with respect to kriging weights, while simultaneously satisfying this constraint, leads to OK system:

$$\sum_{j=1}^n \lambda_j \gamma(x_i, x_j) + \mu_K = \gamma(x_i, x_0), \quad i = 1, \dots, n \quad (3.38)$$

and

$$\sum_{i=1}^n \lambda_i = 1 \quad (3.39)$$

In the above equations, $\gamma(x_i, x_j)$ denotes semivariogram value between observations located at x_i and x_j , $\gamma(x_i, x_0)$ represents semivariogram value between observation x_i and estimation location x_0 , and μ_K is Lagrange multiplier introduced to ensure that resulting estimator remains unbiased.

The semivariogram values appearing in the kriging system were computed using anisotropy-adjusted coordinates as described in Section 3.3.6.

The solution of this system of linear equations provides optimal kriging weights used to estimate unknown value at location x_0 (Journel and Huijbregts, 1978). These weights depend on spatial configuration of the observations and variogram model describing the spatial continuity of the regionalized variable. Consequently, observations located closest to estimation point do not necessarily receive largest weights, since neighboring observations may contain redundant information. By accounting simultaneously for proximity and spatial correlation, OK can provide predictions associated with minimum estimation variance under assumptions of the adopted variogram model.

3.5.7 Best Linear Unbiased Predictor

OK is commonly referred to as BLUP [5], [6]. Term "best" indicates that the estimator possesses minimum variance among all linear unbiased estimators. Term "linear" reflects that predictions are obtained through weighted sums of observations.

The term "unbiased" implies that:

$$E[\hat{Z}_{OK}(x_0)] = E[Z(x_0)] \quad (3.40)$$

The BLUP property explains widespread use of OK in hydrogeological investigations, mining applications, petroleum engineering, and environmental sciences.

Nevertheless, optimality of OK is conditional upon validity of assumptions embedded within the adopted variogram model. Violations of stationarity assumptions or inadequately specified variograms may reduce predictive performance.

3.5.8 Kriging Variance and Prediction Uncertainty

One of major advantages of OK over deterministic interpolation methods is its ability to quantify prediction uncertainty.

The OK variance is computed as

$$\sigma_{OK}^2(x_0) = \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \mu_K \quad (3.41)$$

Unlike residual errors, which can only be evaluated when observations are available, kriging variance provides an estimate of prediction uncertainty at every location within interpolation domain. The magnitude of kriging variance generally decreases in vicinity of observation points and increases in sparsely sampled regions. It is influenced by geometry of the sampling network and adopted variogram model but does not depend directly on observed values themselves. Consequently, OK not only provides estimates of regionalized variable but also quantifies confidence associated with those estimates [42].

In summary, OK constitutes a robust geostatistical framework for spatial interpolation. By combining unbiasedness, minimum estimation variance, and explicit consideration of spatial dependence, it offers a theoretically sound approach for estimating spatially distributed hydrogeological variables and forms foundation

for the subsequent cross-validation and hybrid geostatistical-machine learning framework developed in this study.

3.6 Leave-One-Out Cross-Validation Technique

Predictive capability of OK estimator is evaluated under genuine out-of-sample conditions to determine its ability to generalize beyond the observed data. In the proposed framework, LOOCV is employed to generate out-of-sample prediction Residuals that subsequently provide empirical basis for assessing predictive performance and identifying differences between OK predictions and observed values. By repeatedly withholding individual observations during model construction, LOOCV enables differences between geostatistical predictions and observed values to be identified under genuine out-of-sample conditions.

LOOCV provides a rigorous assessment of predictive performance by repeatedly withholding one observation from dataset, constructing OK model using the remaining observations, and predicting the omitted value under the same modeling assumptions [45].

For each observation located at x_i , the leave-one-out predictor is denoted by

$$\hat{Z}_{OK}^{(-i)}(x_i) \quad (3.42)$$

where the superscript $(-i)$ indicates that observation i was excluded from the calibration dataset during prediction.

Corresponding LOOCV residual is defined as

$$r(x_i) = Z_{obs}(x_i) - \hat{Z}_{OK}^{(-i)}(x_i) \quad (3.43)$$

where $Z_{obs}(x_i)$ represents observed value at location x_i . This residual quantifies difference between the observed value and the prediction obtained under genuine out-of-sample conditions.

Unlike apparent prediction errors obtained by evaluating model using same observations employed during calibration, LOOCV residuals provide a more reliable representation of predictive performance because each observation is predicted using a model that excludes that observation from calibration dataset [45]. Consequently, these residuals reflect ability of OK model to generalize unseen data.

LOOCV predictions and their associated residuals constitute basis for the subsequent probabilistic interpretation of OK estimator within a maximum-likelihood framework. Resulting residuals are treated as empirical realizations of the prediction

discrepancy process and provide foundation for distinguishing structured prediction Residuals from unstructured random fluctuations in the subsequent discrepancy modeling stage.

3.7 Maximum-Likelihood Interpretation of LOOCV Estimates

LOOCV-based OK estimates introduced in previous section may be interpreted within a probabilistic framework by viewing log-transformed hydraulic conductivity field as a realization of a Gaussian process. Under this interpretation, OK corresponds to a statistical estimator whose predictions can be expressed through conditional moments of a GP. This perspective establishes connection between classical geostatistics and statistical inference [5], [6], [8].

Assume that the spatial field of log-transformed hydraulic conductivity is a realization of a GP:

$$Z(x) \sim GP(m(x), C_\theta(x, x')) \quad (3.44)$$

where $m(x)$ denotes mean function and $C_\theta(x, x')$ represents covariance function describing spatial dependence between locations x and x' [8]. Covariance function is parameterized by vector of covariance parameters θ . In OK, mean function is assumed to be unknown but locally constant within estimation neighborhood.

Under Gaussian-process formulation, observed data are assumed to follow a multivariate normal distribution whose covariance structure is governed by parameter vector θ . In a Maximum-Likelihood framework, covariance parameters are estimated by maximizing likelihood of the observed data. Although present study employs variogram-based parameter estimation rather than direct likelihood optimization, Gaussian-process interpretation remains useful because it provides a probabilistic framework for interpreting kriging predictions and their associated uncertainty.

Within this probabilistic interpretation, covariance parameters may be viewed as hyperparameters governing GP. In present study, these parameters were inferred through variogram modeling under second-order stationarity assumptions and include quantities associated with nugget effect, sill, range, and anisotropy.

Given observations

$$D = \{Z(x_i)\}_{i=1}^n \quad (3.45)$$

OK predictor can be expressed as conditional expectation

$$\hat{Z}_{OK}(x) = E[Z(x) | D, \hat{\theta}] \quad (3.46)$$

where $\hat{\theta}$ denotes estimated covariance parameters [46].

In the present study, the Ordinary Kriging prediction is assumed to represent an estimator of the unknown hydraulic conductivity. Under this assumption, the LOOCV residuals quantify the estimation error associated with the Ordinary Kriging estimator and provide the basis for the subsequent discrepancy-learning stage.

Under GP formulation, this conditional expectation coincides with BLUP derived from classical kriging equations (Matheron, 1963; Journel and Huijbregts, 1978; Yildirim et al., 2026). Consequently, OK can be interpreted as posterior mean of the GP conditioned on the observed data.

Uncertainty associated with prediction is quantified through the conditional variance:

$$\sigma_{OK}^2(x) = \text{Var}[Z(x) | D, \hat{\theta}] \quad (3.47)$$

This quantity represents posterior uncertainty of the GP conditioned on available observations and estimated covariance structure. Larger values of kriging variance indicate greater uncertainty in regions characterized by sparse observations or limited spatial information [8].

Consequently, OK estimator may be interpreted as a probabilistic predictor derived from conditional distribution of a Gaussian process. This interpretation provides theoretical basis for viewing LOOCV residuals as realizations of a discrepancy process and motivates subsequent modeling of their structured component using Gaussian Process and RF approaches. Similar likelihood-based perspectives on OK have also been discussed by Yildirim et al. (2026) [46].

3.8 Discrepancy Decomposition

LOOCV residuals introduced in Section 3.5 and interpreted probabilistically in Section 3.6 provide empirical evidence of limitations of the primary geostatistical model. Classical OK implicitly assumes that adopted covariance structure adequately characterizes spatial variability of the regionalized variable. Under this assumption, deviations between predictions and observations are treated as random uncertainty quantified through kriging variance. However, heterogeneous geological systems often exhibit behaviors that cannot be fully represented by selected variogram

model. Consequently, systematic deviations may exist between kriging predictions and true underlying spatial field [47].

To account for these limitations, true spatial field is decomposed into two components:

$$Z(x) = \hat{Z}_{OK}(x) + \delta(x) \quad (3.48)$$

where $Z(x)$ denotes true value of regionalized variable at location x , $\hat{Z}_{OK}(x)$ represents OK prediction, and $\delta(x)$ denotes a model-discrepancy field describing systematic deviations between kriging model and the observed hydraulic conductivity measurements[47].

Equation (3.48) represents a conceptual decomposition in which discrepancy field captures systematic deviations between OK prediction and underlying spatial process.

In practice, true discrepancy field cannot be observed directly because true spatial field is unknown at unsampled locations. Accordingly, LOOCV residuals generated in Section 3.5 are interpreted as empirical realizations of discrepancy process and may be expressed as:

$$r(x_i) = \delta(x_i) + \varepsilon_i \quad (3.49)$$

Under this formulation, structured component $\delta(x_i)$ represents information that is potentially predictable from spatial coordinates and auxiliary variables, whereas ε_i is assumed to represent irreducible random variability.

Discrepancy field may equivalently be defined as:

$$\delta(x) = Z(x) - \hat{Z}_{OK}(x) \quad (3.50)$$

Discrepancy decomposition indicates that discrepancy process represents portion of spatial variability that remains unexplained by OK model. Unlike purely random noise, discrepancy component may exhibit its own spatial structure and therefore contain predictable information that can be modeled and used to improve prediction accuracy [48].

Several factors may contribute to existence of a discrepancy field in hydrogeological applications. These include departures from stationarity assumptions, misspecification of adopted variogram model, unresolved small-scale heterogeneity, anisotropy effects that are not fully represented by covariance structure, and limitations arising from sparse sampling. Therefore, LOOCV residuals generated by

OK may contain systematic patterns rather than representing independent random fluctuations.

Discrepancy decomposition provides conceptual foundation of the proposed methodology. Rather than assuming that OK fully describes spatial variability of the system, the framework explicitly acknowledges existence of model inadequacy through discrepancy component. This discrepancy field is subsequently approximated using out-of-sample prediction residuals and modeled separately to obtain bias-corrected predictions [47], [48].

Accordingly, following section focuses on modeling structured discrepancy component using GPR and RF approaches to obtain bias-corrected predictions of the log-transformed hydraulic conductivity field.

3.9 Spatial Modeling of Estimation Discrepancy

LOOCV residuals introduced in previous section are interpreted as empirical realizations of discrepancy process. Because these residuals may exhibit spatial organization rather than purely random behavior, they can be modeled to capture systematic deviations that are not explained by primary OK model. Objective of discrepancy modeling is therefore to characterize latent process responsible for structured component of LOOCV residuals and subsequently use this information to improve prediction accuracy [47].

Let set of LOOCV residuals be denoted by

$$\{r(x_i)\}_{i=1}^n \quad (3.51)$$

where $r(x_i)$ represents residual associated with observation location x_i . These residuals are treated as observations of an underlying discrepancy field describing systematic difference between the true spatial field and the OK predictions. These residual observations constitute training dataset used to infer structured discrepancy component through either GPR or RF modeling.

One approach consists of representing discrepancy field as a GP. Under this formulation,

$$\delta(x) \sim GP(\mu_\delta(x), C_\delta(x, x')) \quad (3.52)$$

where μ_δ denotes mean function of the discrepancy process and $C_\delta(x, x')$ represents covariance function describing spatial dependence between discrepancy values at locations x and x' . Covariance structure of discrepancy field may be represented

through an appropriate covariance kernel inferred from LOOCV residual observations.

Alternatively, discrepancy field may be modeled using a machine learning approach. Rather than explicitly specifying a covariance structure, a supervised learning algorithm can be trained to learn relationship between discrepancy values and a set of explanatory variables derived from geostatistical model and spatial configuration of the observations. In the present study, RF is adopted owing to its ability to capture nonlinear relationships without requiring strong assumptions regarding underlying data distribution [49]. Discrepancy estimate obtained using RF may be expressed as:

$$\hat{\delta}_{RF}(x) = f_{RF}(v(x)) \quad (3.53)$$

where $v(x)$ denotes the vector of explanatory variables associated with location x , and $f_{RF}(\cdot)$ denotes RF discrepancy model.

Discrepancy observations used for training correspond to LOOCV residuals defined in Equation (3.43). For both GPR and RF modeling, a common set of explanatory variables was constructed from geostatistical predictions and spatial configuration of the observations. These variables included spatial coordinates, OK prediction, OK standard deviation, the distance to the nearest training observation, local point density, relative depth, the mean distance to neighboring observations, and anisotropy-based distance measures.

In the present study, discrepancy field was modeled independently using GPR and RF, resulting in two alternative discrepancy-correction frameworks. Former assumes that residual behavior can be described through an explicit covariance structure and provides a probabilistic characterization of uncertainty, whereas latter allows potentially complex and nonlinear relationships to be learned directly from data [47], [48]. Regardless of modeling strategy adopted, ultimate objective remains the same: to estimate structured component of LOOCV residuals and use this information to correct primary OK predictions [48].

Accordingly, following subsections present GP and RF formulations adopted to model structured discrepancy component prior to their incorporation into corrected prediction framework.

3.9.1 Gaussian Process Discrepancy Model

GPR is a non-parametric Bayesian machine learning method that provides a flexible framework for regression, interpolation, and uncertainty quantification. Unlike conventional regression approaches that assume a predefined functional relationship between predictors and responses, Gaussian Processes define a probability

distribution over a space of possible functions and update this distribution using available observations. As a result, method provides both predictions and associated measures of uncertainty. These characteristics make GPR particularly suitable for hydrogeological applications characterized by sparse observations, heterogeneous geological conditions, and spatially correlated variables. The theoretical background presented here is adapted specifically for discrepancy modeling using LOOCV residuals generated from OK [8].

A GP may be interpreted as a collection of random variables for which every finite subset follows a joint multivariate Gaussian distribution. Method assumes that observations located closer in feature space tend to exhibit stronger similarity than observations located farther apart. This dependence structure is described through a covariance function, commonly referred to as a kernel function. Within the proposed framework, discrepancy field is modeled as

$$\delta(x) \sim GP(\mu_\delta(x), C_\delta(x, x')) \quad (3.54)$$

where $\mu_\delta(x)$ denotes mean function of the discrepancy process and $C_\delta(x, x')$ represents covariance function describing the relationship between discrepancy values at locations x and x' [8].

Mean function is defined as

$$\mu_\delta(x) = E[\delta(x)] \quad (3.55)$$

whereas covariance function is given by

$$C_\delta(x, x') = E[(\delta(x) - \mu_\delta(x))(\delta(x') - \mu_\delta(x'))] \quad (3.56)$$

Let training dataset be defined by

$$D = \{X, r\} \quad (3.57)$$

where X represents matrix of predictor variables and

$$r = [r(x_1), r(x_2), \dots, r(x_n)]^T \quad (3.58)$$

denotes vector of LOOCV residuals.

In the present study, each row of predictor matrix X corresponds to a feature vector associated with a specific observation location. The predictor variables included spatial coordinates, OK prediction, OK standard deviation, distance to the nearest training observation, local point density, relative depth, mean distance to neighboring observations, and anisotropy-based distance measures. Consequently,

gaussian process kernel operates in resulting multidimensional feature space rather than on spatial coordinates alone.

Given a new feature vector $v(x_*)$ associated with prediction location x_* , predictive mean of the discrepancy process is expressed as follows.

$$\hat{\delta}_{GP}(x_*) = c_*^T (C + \sigma_n^2 I)^{-1} r \quad (3.59)$$

where C is covariance matrix among training observations, c_* is covariance vector between prediction location and training observations, σ_n^2 denotes the observation noise variance, I represents the identity matrix, and r denotes vector of LOOCV residuals.

Corresponding predictive variance is given by the following expression.

$$\sigma_{GP}^2(x_*) = c(x_*, x_*) - c_*^T (C + \sigma_n^2 I)^{-1} c_* \quad (3.60)$$

where $\sigma_{GP}^2(x_*)$ represents predictive variance associated with the discrepancy estimate at location x_* .

Predictive variance quantifies uncertainty associated with the discrepancy prediction. Locations surrounded by numerous observations generally exhibit smaller uncertainty, whereas regions situated far from sampled data tend to display larger uncertainty. This characteristic provides an important advantage over deterministic correction methods because it allows confidence associated with the discrepancy estimates to be quantified explicitly.

Covariance kernel determines behavior of GP and defines degree of similarity between observations. Commonly used kernels include Radial Basis Function (RBF), Matérn, and Rational Quadratic kernels. For illustration, RBF kernel may be expressed as follows.

$$C(v, v') = \sigma_f^2 \exp\left(-\frac{\|v - v'\|^2}{2l_k^2}\right) \quad (3.61)$$

where v and v' denote feature vectors and σ_f^2 denotes signal variance and l_k is characteristic length scale controlling rate at which correlation decreases with distance. Matérn kernel provides additional flexibility through a smoothness parameter, while Rational Quadratic kernel may accommodate variability occurring over multiple spatial scales. Choice of kernel should therefore reflect characteristics of discrepancy process inferred from LOOCV residuals.

Parameters governing covariance function, commonly referred to as hyperparameters, are estimated automatically during model training through the

maximization of marginal log-likelihood (Rasmussen and Williams, 2006). This procedure allows GP model to adapt its covariance structure to observed discrepancy field without requiring subjective parameter specification.

GP discrepancy modeling offers several advantages within the proposed framework. First, it preserves probabilistic nature of the geostatistical methodology by providing both discrepancy predictions and uncertainty estimates. Second, covariance function naturally incorporates spatial dependence, allowing structured residual patterns to be represented explicitly. Finally, non-parametric formulation enables nonlinear relationships to be modeled without assuming a predefined functional form.

Consequently, GPR provides an effective approach for modeling structured component of OK residuals and generating discrepancy corrections that account for both predictive uncertainty and complex relationships embedded within feature space [47].

3.9.2 RF Discrepancy Model

RF is a supervised machine learning algorithm introduced by Breiman (2001) that combines multiple decision trees to improve predictive accuracy and reduce overfitting. The algorithm belongs to family of ensemble learning methods, in which several individual models are combined to produce a more robust predictor than any single model alone. Owing to its ability to model complex nonlinear relationships and interactions among variables, RF has been widely applied in environmental sciences, hydrology, geology, and groundwater studies [50].

Within the proposed framework, RF is employed to model discrepancy process represented by LOOCV residuals obtained from OK. Objective is to learn relationship between residual values and a set of explanatory variables describing local prediction environment. In doing so, method seeks to estimate structured component of LOOCV residuals that cannot be adequately represented by primary geostatistical model.

Fundamental building block of a RF model is decision tree. A decision tree recursively partitions predictor space into smaller regions according to a series of decision rules. For regression problems, each terminal node contains a predicted value corresponding to average of the observations assigned to that node. Prediction generated by an individual decision tree is expressed as follows.

$$T_t(v(x)) = f_t(v(x)) \quad (3.62)$$

where $T_t(v(x))$ is prediction produced by the t -th decision tree, $f_t(\cdot)$ represents decision-tree model, and $v(x)$ denotes feature vector associated with prediction location x .

Although individual decision trees are capable of modeling nonlinear relationships, they are highly sensitive to variations in training data and often suffer from overfitting. RF addresses this limitation by constructing a large number of decision trees and combining their predictions through an ensemble strategy.

First key component of RF is bootstrap aggregation, commonly referred to as bagging [50].

Training dataset used for RF discrepancy modeling is defined as follows.

$$D = \{(v(x_i), r_i)\}_{i=1}^n \quad (3.63)$$

where $v(x_i)$ denotes feature vector associated with observation i , r_i represents the corresponding LOOCV residual, and n is total number of observations used for training. Feature vectors used in RF were identical to those employed by GPR discrepancy model and consisted of spatial coordinates, OK prediction metrics, neighborhood statistics, relative depth, and anisotropy-related measures. These explanatory variables provide information regarding both reliability of OK estimates and local conditions under which discrepancies occur.

Bootstrap datasets are generated by randomly sampling observations with replacement from original training dataset. Each bootstrap sample is subsequently used to construct an independent decision tree. As a result, individual trees are exposed to slightly different datasets and therefore learn different patterns from data. Because each tree is trained using a different bootstrap realization of data and a different subset of predictor variables, the ensemble is able to reduce prediction variance while preserving nonlinear relationships present in the discrepancy observations.

Second component of the algorithm is random feature selection. During construction of each tree, only a random subset of predictor variables is considered at every split [50]. This additional source of randomness reduces correlation among trees and improves overall predictive performance of ensemble. Consequently, RF produces models that are more stable and less prone to overfitting than individual decision trees.

Suppose that RF model consists of N_T regression trees. discrepancy prediction obtained from the ensemble is expressed as follows.

$$\hat{\delta}_{RF}(x) = \frac{1}{N_T} \sum_{t=1}^{N_T} T_t(v(x)) \quad (3.64)$$

where $\hat{\delta}_{RF}(x)$ denotes predicted discrepancy at location x , N_T represents total number of trees in the ensemble, and $T_t(v(x))$ is prediction generated by the t -th decision tree.

RF ensemble prediction indicates that final prediction is obtained by averaging outputs of all trees within the ensemble. Averaging process reduces prediction variance and increases model robustness [50]. Unlike a single decision tree, which may be strongly influenced by specific observations, RF combines information from many trees and therefore produces more stable and reliable predictions.

Unlike GP discrepancy modeling, RF does not explicitly define a covariance structure governing the residual field. Instead, the method identifies relationships directly from training data through recursive partitioning. Consequently, RF can effectively capture nonlinear interactions and threshold effects that may be difficult to represent using conventional covariance-based approaches.

RF offers several advantages for discrepancy modeling within heterogeneous aquifer systems. Algorithm can accommodate nonlinear relationships, interactions among variables, noisy datasets, and high-dimensional predictor spaces without requiring assumptions regarding data distributions. Furthermore, RF is relatively insensitive to outliers and multicollinearity and can effectively model complex environmental systems characterized by strong heterogeneity [49].

Despite these advantages, RF does not directly quantify prediction uncertainty in the same probabilistic manner as GPR. Therefore, within the proposed framework, RF and GPR represent two complementary approaches for modeling structured component of the LOOCV residuals. Former emphasizes predictive performance through data-driven learning, whereas latter provides explicit uncertainty quantification through a probabilistic covariance structure.

Consequently, RF discrepancy modeling provides a flexible and powerful mechanism for correcting OK predictions and improving characterization of hydraulic conductivity heterogeneity within the investigated aquifer system [47], [49].

3.10 Hybrid Prediction Framework

Discrepancy models developed in previous section provide estimates of structured component of LOOCV residuals associated with OK. These discrepancy estimates

can subsequently be incorporated into original geostatistical predictions to obtain bias-corrected estimates of spatial field. Objective of corrected prediction framework is therefore to combine strengths of geostatistical interpolation and machine learning discrepancy modeling within a unified methodology.

Let $\hat{Z}_{OK}(x)$ denote OK prediction at location x , and let $\hat{\delta}(x)$ represent estimated discrepancy obtained from a discrepancy model. Corrected prediction is expressed as follows.

$$\hat{Z}_{corr}(x) = \hat{Z}_{OK}(x) + \hat{\delta}(x) \quad (3.65)$$

where $\hat{Z}_{corr}(x)$ denotes corrected prediction at location x , $\hat{Z}_{OK}(x)$ represents OK estimate, and $\hat{\delta}(x)$ is estimated discrepancy field [47]. Equation (3.58) directly follows from discrepancy decomposition introduced in Section 3.7, where true spatial field was represented as sum of the OK prediction and a discrepancy component.

Corrected prediction consists of two complementary components. First component captures large-scale spatial structure inferred from observed data through OK, whereas second component represents systematic deviations that cannot be adequately described by primary geostatistical model.

When discrepancy field is modeled using GPR, the corrected prediction is expressed as follows.

$$\hat{Z}_{corr}^{GP}(x) = \hat{Z}_{OK}(x) + \hat{\delta}_{GP}(x) \quad (3.66)$$

where $\hat{Z}_{corr}^{GP}(x)$ denotes GP-corrected prediction and $\hat{\delta}_{GP}(x)$ represents discrepancy estimate obtained from GP discrepancy model.

GP-corrected framework preserves ability of OK to characterize spatial continuity while additionally incorporating residual structures through a probabilistic discrepancy model. Consequently, resulting predictions benefit from both geostatistical interpolation and probabilistic discrepancy modeling (Rasmussen and Williams, 2006).

Alternatively, when discrepancy field is modeled using RF, corrected prediction is given by the following expression.

$$\hat{Z}_{corr}^{RF}(x) = \hat{Z}_{OK}(x) + \hat{\delta}_{RF}(x) \quad (3.67)$$

where $\hat{Z}_{corr}^{RF}(x)$ denotes RF-corrected prediction and $\hat{\delta}_{RF}(x)$ represents discrepancy estimate obtained from the RF discrepancy model.

RF-corrected framework combines large-scale spatial representation provided by OK with ability of ensemble learning techniques to identify nonlinear residual patterns

embedded within data. This approach enables systematic patterns contained within LOOCV residuals to be reduced without imposing explicit assumptions regarding covariance structure of discrepancy field [49].

Corrected prediction framework proposed in this study therefore does not replace underlying geostatistical methodology. Instead, it augments OK through an additional discrepancy-correction stage designed to account for unresolved heterogeneity, variogram misspecification, sparse sampling effects, and departures from assumptions underlying classical geostatistical interpolation.

Accordingly, GP-corrected and RF-corrected predictions represent two alternative implementations of proposed hybrid methodology. Their predictive performances are subsequently evaluated and compared to assess extent to which discrepancy correction improves characterization of hydraulic conductivity heterogeneity and enhances predictive capability of the overall framework [47], [49].

3.11 Uncertainty Quantification

An important advantage of geostatistical approaches is their ability to quantify prediction uncertainty. In the proposed framework, uncertainty quantification is most naturally implemented for GP discrepancy modeling. First source is uncertainty associated with OK predictions, while second source originates from discrepancy model used to correct kriging estimates. Combined effect of these two components provides an estimate of uncertainty associated with the corrected predictions.

Let uncertainty associated with discrepancy process be defined as follows.

$$\sigma_{\delta}^2(x) = \text{Var}[\delta(x) | D] \quad (3.68)$$

Where $\sigma_{\delta}^2(x)$ denotes discrepancy uncertainty at location x , $\delta(x)$ represents discrepancy field, and D denotes training dataset used to construct discrepancy model [8].

Assuming weak dependence between uncertainty associated with OK and uncertainty associated with discrepancy model, total uncertainty of the corrected prediction may be approximated as follows.

$$\sigma_{corr}^2(x) \approx \sigma_{OK}^2(x) + \sigma_{\delta}^2(x) \quad (3.69)$$

where $\sigma_{corr}^2(x)$ denotes variance associated with corrected prediction, $\sigma_{OK}^2(x)$ is OK variance, and $\sigma_{\delta}^2(x)$ represents discrepancy variance.

Combined uncertainty approximation indicates that uncertainty associated with corrected prediction can be interpreted as the sum of geostatistical interpolation

uncertainty and uncertainty arising from discrepancy model. Assumption of weak dependence implies that covariance between these two uncertainty sources is sufficiently small to be neglected.

For GP discrepancy modeling, the discrepancy variance is obtained directly from predictive variance of GP model [8]. In contrast, RF does not naturally provide probabilistic uncertainty estimates [50]. Consequently, approximation presented above is most readily applicable within GP correction framework, whereas additional techniques would be required to quantify uncertainty associated with RF predictions.

The uncertainty quantification framework provides valuable information regarding confidence associated with corrected predictions and facilitates identification of regions characterized by greater predictive uncertainty. Such information is particularly important in hydrogeological applications, where decision-making often requires an explicit assessment of prediction reliability.

3.12 Validation Criteria

Predictive performance of proposed framework was evaluated using seven boreholes reserved for independent validation. Objective of validation procedure was to assess whether discrepancy correction improved predictive capability of OK. Performances of OK, GP-corrected OK, and RF-corrected OK were compared using four statistical indicators commonly employed in regression, interpolation, and hydrological modeling studies: the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), the coefficient of determination (R^2), and the Kling–Gupta Efficiency (KGE) [51], [52], [53].

3.12.1 Root Mean Squared Error

Root Mean Squared Error is defined as follows.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [Z_{obs}(x_i) - Z_{pred}(x_i)]^2} \quad (3.70)$$

where N is total number of validation observations, $Z_{obs}(x_i)$ denotes observed value at location x_i , and $Z_{pred}(x_i)$ represents corresponding predicted value.

RMSE measures overall magnitude of the prediction errors and assigns greater weight to larger deviations. Lower RMSE values indicate improved predictive performance [53].

3.12.2 Mean Absolute Error

Mean Absolute Error is expressed as follows.

$$MAE = \frac{1}{N} \sum_{i=1}^N |Z_{obs}(x_i) - Z_{pred}(x_i)| \quad (3.71)$$

where N is total number of validation observations, $Z_{obs}(x_i)$ denotes observed value at location x_i , and $Z_{pred}(x_i)$ represents corresponding predicted value.

MAE quantifies average magnitude of the prediction errors without considering their direction. Unlike RMSE, it does not assign greater weight to larger errors. Therefore, smaller MAE values indicate improved predictive accuracy [53].

3.12.3 Coefficient of Determination

Coefficient of determination is calculated as follows.

$$R^2 = 1 - \frac{\sum_{i=1}^N (Z_{obs}(x_i) - Z_{pred}(x_i))^2}{\sum_{i=1}^N (Z_{obs}(x_i) - \bar{Z}_{obs})^2} \quad (3.72)$$

where $Z_{obs}(x_i)$ denotes observed value at location x_i , $Z_{pred}(x_i)$ represents corresponding predicted value, $\bar{Z}_{obs}$ is mean of the observed values, and N is total number of validation observations.

Coefficient of determination quantifies proportion of variability in the observed data explained by the predictions. Values of R^2 approaching unity indicate stronger agreement between observed and predicted values, whereas values closer to zero indicate limited explanatory capability [52].

3.12.4 Kling–Gupta Efficiency

In addition to RMSE, MAE, and R^2 , model performance was further evaluated using Kling–Gupta Efficiency (KGE), which provides a composite assessment of prediction accuracy by simultaneously considering correlation, variability, and bias between observed and predicted values [51].

KGE is defined as

$$KGE = 1 - \sqrt{(r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2} \quad (3.73)$$

where r denotes Pearson correlation coefficient, α is variability ratio, and β is bias ratio [51].

Variability ratio is defined as

$$\alpha = \frac{\sigma_{pred}}{\sigma_{obs}} \quad (3.74)$$

where σ_{pred} and σ_{obs} denote standard deviations of the predicted and observed values, respectively.

Bias ratio is expressed as

$$\beta = \frac{\mu_{pred}}{\mu_{obs}} \quad (3.75)$$

where μ_{pred} and μ_{obs} denote means of the predicted and observed values, respectively.

KGE metric attains an optimal value of 1 when predictions perfectly reproduce observations. Values closer to unity therefore indicate improved predictive performance, whereas lower values indicate deficiencies in correlation, variability representation, bias, or a combination thereof. Because KGE simultaneously evaluates multiple aspects of model performance, it is particularly useful for assessing ability of hydraulic conductivity characterization approaches to reproduce both spatial variability and statistical properties of the observed dataset.

Predictive performances of OK, GP-corrected OK, and RF-corrected OK were compared using these statistical indicators. Improvements resulting from discrepancy correction were assessed through reductions in RMSE and MAE together with increases in R^2 and KGE. This validation procedure provides a quantitative basis for evaluating the effectiveness of proposed hybrid framework in improving prediction of hydraulic conductivity fields.

3.13 Methodological Contribution

This study extends classical OK framework by introducing a data-driven spatial discrepancy modeling approach based on LOOCV residuals [47]. Rather than assuming that OK fully captures spatial variability of hydraulic conductivity, proposed methodology explicitly models the systematic deviations revealed through LOOCV residuals generated from geostatistical predictions.

Our currently developed framework combines strengths of geostatistical interpolation, statistical inference, and machine learning techniques within a unified hybrid approach. OK is employed to characterize large-scale spatial continuity of hydraulic conductivity field through variogram-based interpolation, whereas

LOOCV is used to evaluate predictive performance under genuine out-of-sample conditions and generate prediction residuals.

A distinctive feature of the proposed methodology is probabilistic interpretation of LOOCV-based OK estimator within a maximum-likelihood framework. By viewing log-transformed hydraulic conductivity field as a realization of a Gaussian random field, OK predictor is interpreted as conditional mean estimator, or BLUP, of corresponding GP. This interpretation establishes connection between classical geostatistics and statistical inference and provides theoretical basis for distinguishing structured prediction errors from random uncertainty.

Structured component of LOOCV residuals is subsequently modeled using discrepancy-learning approaches. Two alternative discrepancy modeling strategies are investigated. First employs GPR to provide probabilistic discrepancy estimates together with associated uncertainty quantification through an explicit covariance structure [8]. Second utilizes RF to capture complex nonlinear relationships embedded within residual patterns without requiring assumptions regarding underlying data distribution [49], [50].

Corrected prediction framework subsequently integrates OK estimates with discrepancy predictions to generate bias-corrected representations of hydraulic conductivity field [47], [48]. In addition, uncertainty associated with corrected predictions is explicitly considered within GP correction framework through combination of kriging uncertainty and discrepancy uncertainty under assumption of weak dependence.

Consequently, the proposed methodology moves beyond conventional assumption that geostatistical interpolation alone is sufficient to characterize subsurface heterogeneity. By incorporating probabilistic interpretation, cross-validation-based discrepancy identification, and data-driven residual correction into prediction process, framework provides an enhanced representation of heterogeneous aquifer systems and offers a flexible strategy for improving predictive performance in hydrogeological applications.

Accordingly, this study establishes a hybrid geostatistical-machine learning framework that integrates governing physical principles, probabilistic spatial estimation, statistical inference, cross-validation-based discrepancy identification, and data-driven error correction for improved characterization of heterogeneous aquifer systems.

4. Case Study

4.1 Introduction

This chapter presents application of methodology developed in Chapter 3 to the selected case study. Objective is to evaluate capability of proposed geostatistical and machine-learning-based discrepancy-correction framework to characterize and predict spatial distribution of log-transformed hydraulic conductivity within a highly heterogeneous aquifer system.

Chapter begins by introducing study area and describing available boreholes the dataset. The geological and hydrogeological setting of investigated aquifer is then presented, with particular emphasis on the hydrostratigraphic unit selected for modeling. Subsequently, preprocessing of the hydraulic conductivity data, separation of training and validation datasets, and implementation of OK, GPR, and RF approaches are described.

Predictive performances of OK, GP-corrected OK, and RF-corrected OK models are then evaluated using validation framework introduced in Chapter 3. Particular attention is given to ability of the discrepancy-correction approaches to improve characterization of spatial heterogeneity and reduce prediction errors relative to baseline geostatistical model.

Finally, the resulting three-dimensional hydraulic conductivity fields are analysed and compared in terms of their spatial distribution and predictive performance. The hydraulic conductivity fields generated in this chapter provide the input for ATEs simulations presented in Chapter 5, where the influence of different hydraulic conductivity representations on the predicted groundwater flow and heat transport behaviour is investigated.

4.2 Data Acquisition

Study area is located in Milan, northern Italy, around Parco Sempione. This area belongs to broader hydrogeological setting of Po Plain, one of the most important alluvial groundwater systems in Europe. Po Plain is characterized by Quaternary fluvio-glacial and alluvial deposits derived from Alpine and Apennine source areas, forming complex aquifer systems with pronounced spatial variability in grain size, lithology, and hydraulic properties. Such depositional environments commonly exhibit marked heterogeneity, making the spatial characterization of hydraulic conductivity particularly challenging.

Selected Parco Sempione dataset was extracted from a broader Milan hydrogeological database and was chosen because of relatively high spatial density of boreholes within a limited area. Dataset consists of 35 boreholes distributed around Parco Sempione. Borehole locations were extracted and processed using QGIS in the UTM Zone 32 coordinate system. The spatial distribution of study area and selected boreholes is shown in Figure 4.1.

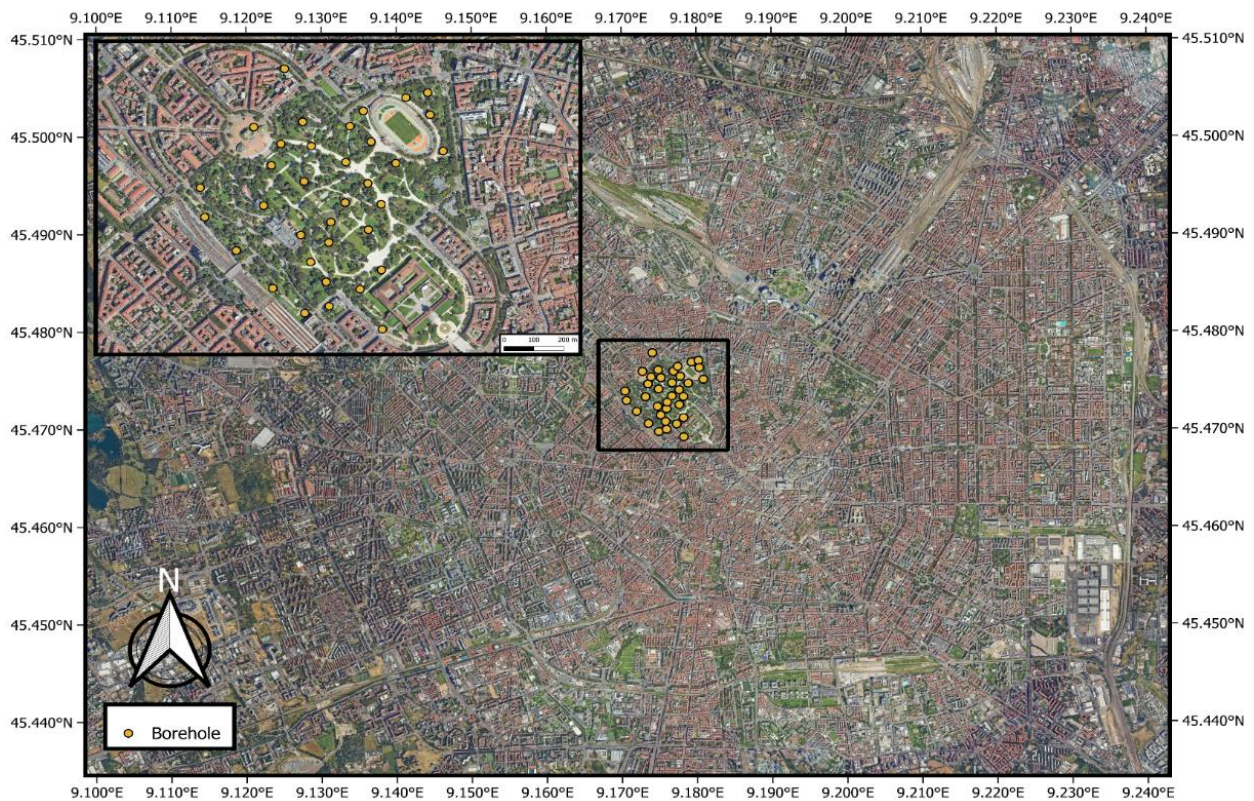


Figure 4.1 Study area and borehole locations.

Original lithostratigraphic information was vertically resampled at regular intervals of 0.5 m along each borehole. This preprocessing step was necessary to obtain

regularly spaced observations suitable for three-dimensional variogram analysis and geostatistical modeling. Each resampled interval contains spatial coordinates, elevation information, log-transformed hydraulic conductivity values, and estimated percentages of principal lithological classes, including gravel, sand, silt, clay, and rock. Hydraulic conductivity values are expressed as $\log_{10}(K)$, where K is given in m/s.

Subsurface properties in study area were characterized using multiple sources of information, including grain-size distribution obtained from sieve analyses, pumping and drawdown tests, and Lefranc tests. Since direct hydraulic measurements are not continuously available along boreholes, lithological descriptions were also used to derive hydrogeological information. Dataset employed in this study was provided by a research team at University of Milano-Bicocca and had been previously processed using the regional hydrofacies characterization framework described by Previati et al. [54]. Consequently, the present work utilizes resulting hydraulic conductivity dataset and associated hydrogeological information without performing the hydrofacies classification procedure directly.

From a hydrogeological perspective, selected domain comprises a shallow phreatic aquifer and a deeper semi-confined aquifer. Present study focuses on Aquifer B, which corresponds to semi-confined aquifer located beneath shallow phreatic zone. Definition of modeled hydrostratigraphic unit was based on both regional hydrogeological information and local borehole observations. Figure 4.2 combines these two sources of information: regional cross-section illustrates occurrence of major aquifer systems beneath Milan, whereas vertical $\log K$ profiles provide local evidence for identifying boundaries of Aquifer B within study area.

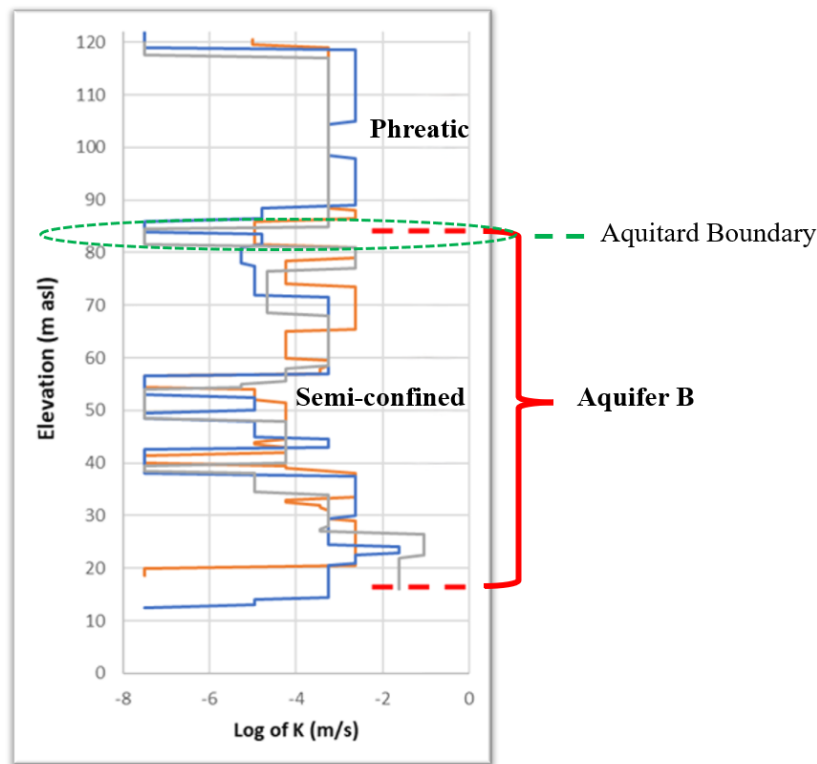


Figure 4.2 Hydrogeological setting of Aquifer B.

Based on available geological information and vertical logK profiles, aquitard separating shallow aquifer from Aquifer B was identified at approximately 82–86 m above sea level. Accordingly, top boundary of Aquifer B was defined at 83.5 m a.s.l. Interpreted bottom of semi-confined aquifer occurs approximately between 15 and 20 m a.s.l. However, because not all boreholes extend to these depths, full thickness of Aquifer B could not be modeled reliably. Consequently, lower boundary of the numerical domain was established at 29.5 m a.s.l., resulting in a modeled aquifer thickness of approximately 54 m. Principal hydrostratigraphic parameters adopted in this study are summarized in Table 4.1.

Parameter	Unit	Value
Surface elevation	m a.s.l.	124.5
Top of Aquifer B (aquitard boundary)	m a.s.l.	83.5
Interpreted bottom of Aquifer B	m a.s.l.	15–20
Modelled lower boundary of Aquifer B	m a.s.l.	29.5
Thickness of modelled Aquifer B	m	54

Table 4.1 Main parameters of the modeled Aquifer B.

Lower modeling boundary was placed above interpreted bottom of Aquifer B because several boreholes did not extend to greater depths, preventing a reliable characterization of deeper portion of the aquifer.

Vertical logK profiles further indicate that Aquifer B cannot be considered homogeneous. Hydraulic conductivity varies substantially with depth, and several less conductive layers are embedded within more permeable sediments. This pronounced vertical variability confirms heterogeneous nature of the aquifer and highlights need for three-dimensional characterization approaches rather than simplified homogeneous parameterization.

Prior to geostatistical analysis, coordinates were transformed into a local reference system, and vertical coordinate was converted from elevation to depth below ground surface. Figure 4.3 illustrates three-dimensional distribution of measured logK values after this transformation. color scale highlights pronounced variability of hydraulic conductivity both laterally and vertically, providing a first indication of complex spatial heterogeneity of Aquifer B.

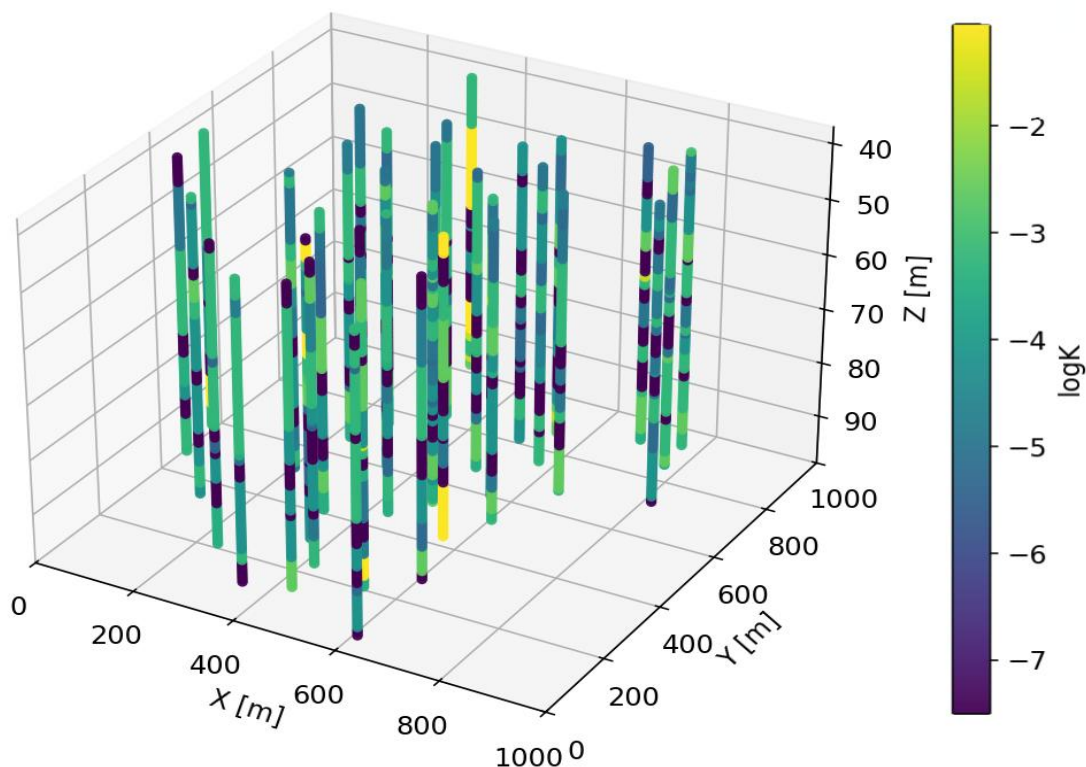


Figure 4.3 Three-dimensional distribution of measured logK values.

To evaluate predictive performance objectively, the 35 boreholes were divided into independent training and testing subsets. Twenty-eight boreholes (80% of the

dataset) were used for variogram fitting, OK modeling, LOOCV residual generation, and the development of the GPR and RF discrepancy models, whereas the remaining seven boreholes (20%) were reserved exclusively for independent validation. Separation was performed at borehole level rather than at individual sample level to avoid information leakage between training and testing datasets and to ensure a rigorous assessment of predictive performance. The spatial distribution of training and testing boreholes is presented in Figure 4.4.

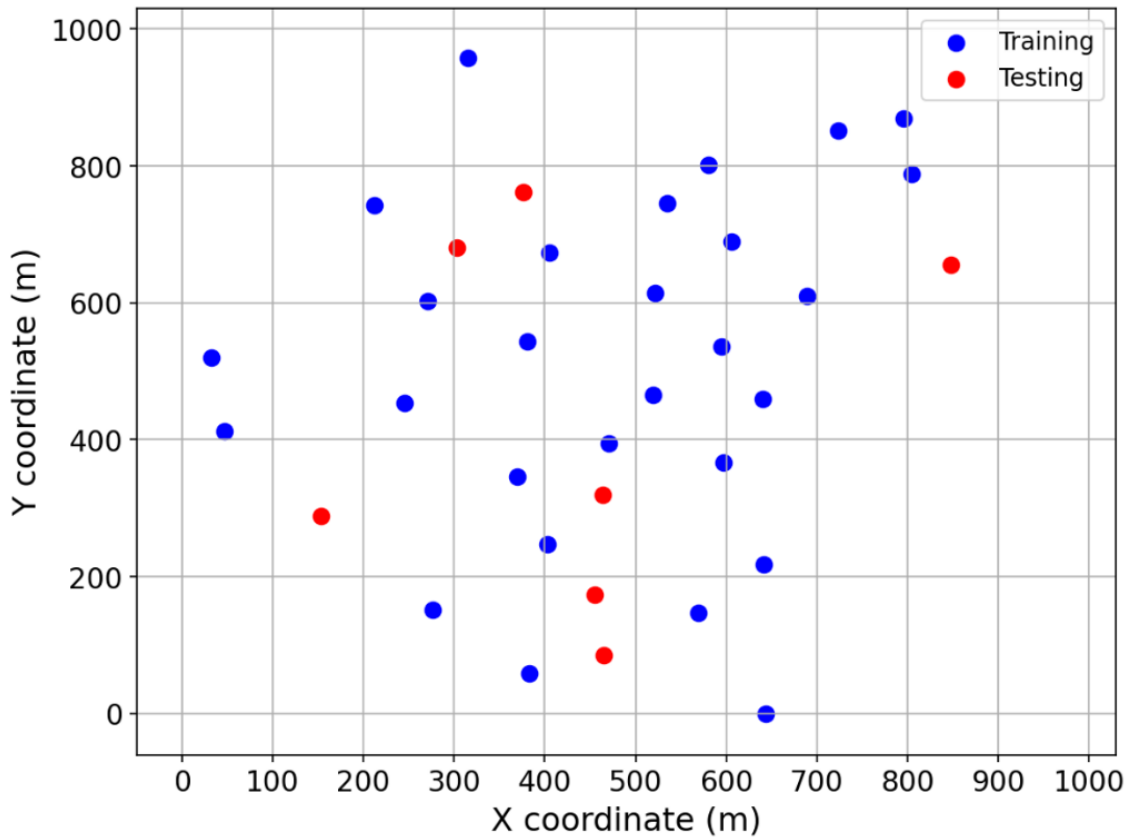


Figure 4.4 Spatial distribution of training and testing boreholes.

This training–testing strategy ensures that predictive capability of characterization models is assessed using observations that were not involved in model development. Testing boreholes therefore provide an independent basis for comparing predicted and observed logK values and for evaluating generalization ability of the proposed approaches.

Training boreholes were subsequently used for variogram modeling, LOOCV analysis, discrepancy modeling, and hydraulic conductivity field generation following framework developed in Chapter 3.

Owing to sparse sampling density, strong vertical variability, and heterogeneous aquifer structure, Parco Sempione dataset provides an appropriate benchmark for evaluating ability of geostatistical and data-driven approaches to predict hydraulic conductivity fields in complex subsurface environments. These characteristics make case study particularly suitable for investigating how different characterization strategies represent subsurface heterogeneity and uncertainty.

4.3 Characterization of Hydraulic Conductivity Fields

Following training–testing separation described in previous section, characterization framework developed in Chapter 3 was applied to training dataset. Workflow consisted of variogram characterization, OK interpolation, LOOCV analysis, and discrepancy modeling using GPR and RF approaches. The resulting hydraulic conductivity fields and associated predictions were subsequently evaluated using independent testing boreholes.

Following subsections present implementation and results of each stage of characterization workflow.

4.3.1 Variogram Characterization

Characterization of spatial continuity constitutes a fundamental step in geostatistical modelling, as predictive performance of OK depends directly on the adequacy of adopted variogram model. Therefore, prior to performing OK interpolation, spatial dependence structure of the log-transformed hydraulic conductivity values was investigated using training dataset.

Experimental semi-variograms were computed using measurements obtained from 28 training boreholes. Because hydraulic conductivity in sedimentary aquifers commonly exhibits different degrees of continuity in horizontal and vertical directions, an anisotropic analysis was conducted. Separate experimental variograms were therefore calculated for horizontal and vertical directions in order to capture directional characteristics of the spatial variability within Aquifer B.

An exponential variogram model was selected to represent observed spatial continuity and was calibrated against experimental variograms obtained from training data. Figure 4.5 presents the resulting experimental variograms together with fitted exponential models for both directions.

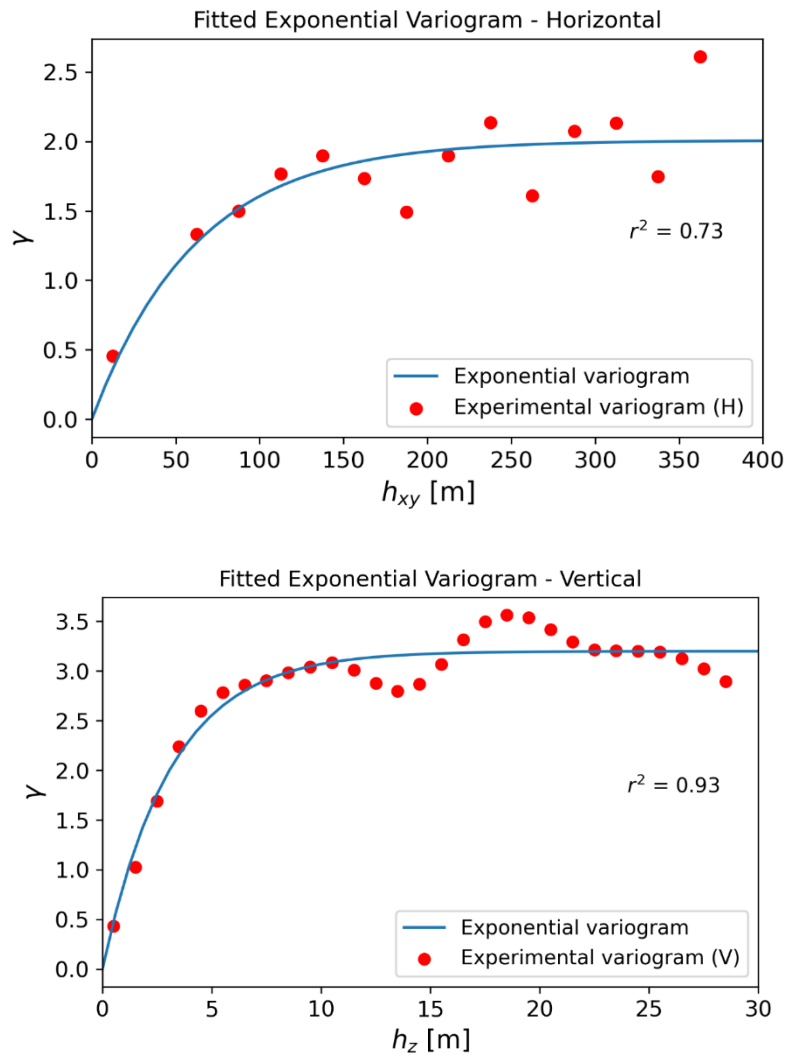


Figure 4.5 Experimental and fitted exponential variograms in vertical and horizontal directions derived from training boreholes.

Fitted exponential models provided a satisfactory representation of experimental variograms, yielding goodness-of-fit values of 0.93 and 0.73 for vertical and horizontal directions, respectively. Stronger agreement observed in vertical direction suggests that vertical continuity structure was more consistently represented by exponential model, whereas larger scatter observed in horizontal variogram reflects greater complexity of lateral heterogeneity within study area.

Calibrated variogram parameters are summarized in Table 4.2.

Parameter	Horizontal	Vertical	Description
Lag	25	1	Distance interval h of pair points at which $\gamma(h)$ is calculated.
Nugget	0.00	0.00	Ordinate at lag 0 representing micro-scale variability or measurement error.
Tolerance	12.5	0.5	Acceptable deviation around each bin of a single lag h .
Range	186.5	9.35	Distance beyond which spatial correlation between data points becomes negligible.
sill	2	3.2	Total variance of spatial data once spatial correlation becomes negligible.
Anisotropy	19.94		Ratio between horizontal and vertical variogram ranges used for anisotropy scaling.

Table 4.2 Calibrated parameters of anisotropic exponential variogram model used for OK.

Results reveal pronounced anisotropic behaviour in hydraulic conductivity field. Estimated horizontal range (186.5 m) was substantially greater than vertical range (9.35 m), indicating that hydraulic conductivity values remain spatially correlated over considerably larger distances in horizontal direction than along vertical axis. Using anisotropy ratio defined in Equation (3.28), a value of approximately 20 was obtained, indicating substantially stronger spatial continuity in the horizontal direction than in vertical direction.

This pronounced anisotropy is consistent with depositional nature of alluvial aquifer systems, where sedimentary layering promotes relatively continuous hydraulic properties laterally while abrupt lithological changes occur over shorter vertical distances. Much smaller vertical range therefore suggests that hydraulic conductivity varies more rapidly with depth, reflecting heterogeneous stratification identified previously from the borehole profiles.

Furthermore, relatively large sill values indicate substantial variability within log-transformed hydraulic conductivity measurements, supporting interpretation that Aquifer B is characterized by a high degree of heterogeneity. Such conditions present considerable challenges for interpolation-based approaches and motivate subsequent evaluation of OK performance and investigation of advanced approaches aimed at improving predictive accuracy.

Calibrated anisotropic variogram model obtained in this analysis was subsequently adopted for OK predictions and LOOCV procedures implemented according to framework described in Chapter 3.

4.3.2 OK Results

Predictive capability of calibrated OK model was evaluated using seven independent testing boreholes that were excluded from variogram fitting and model development. Figure 4.6 compares observed log-transformed hydraulic conductivity values within testing boreholes with corresponding OK estimates.

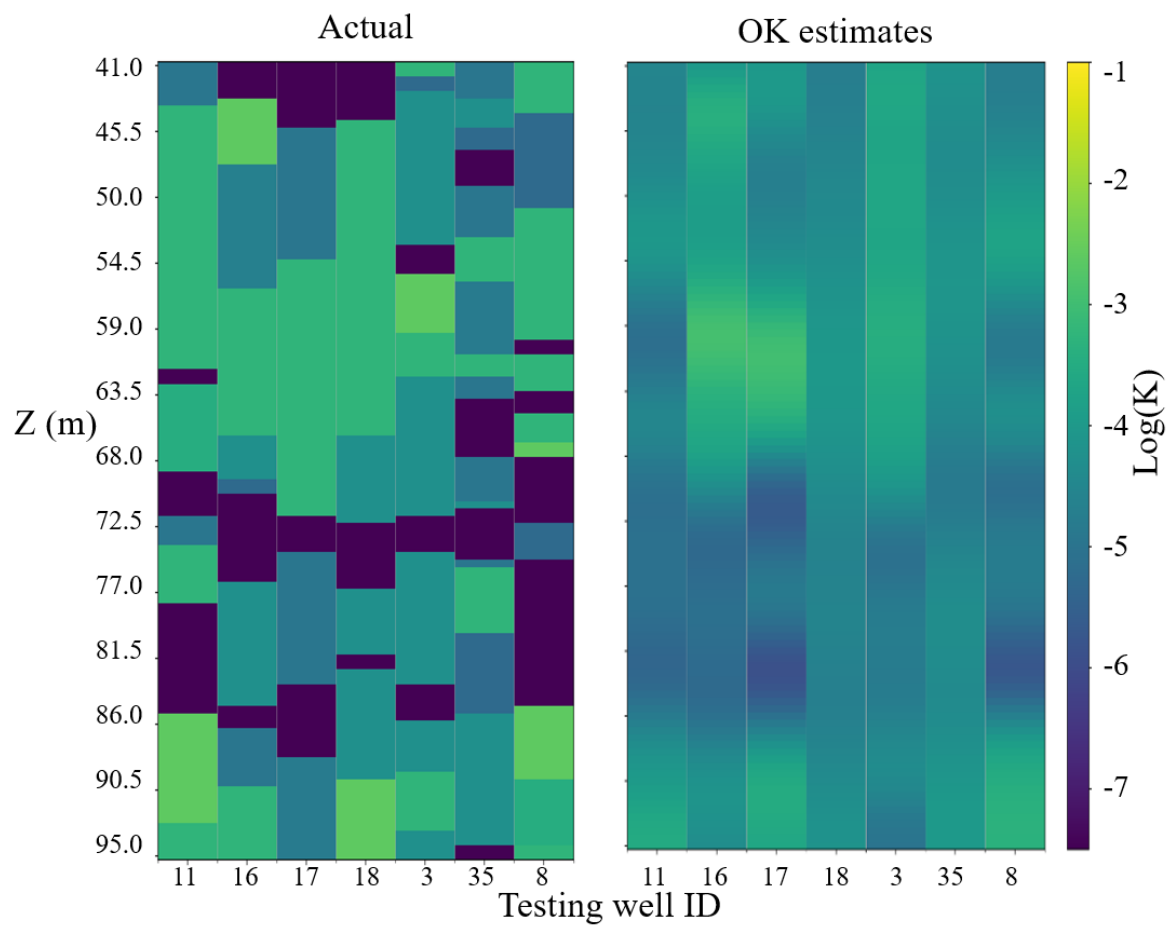


Figure 4.6 Comparison between observed and OK predicted logK values for independent testing boreholes.

Observed data exhibit pronounced vertical heterogeneity, characterized by abrupt transitions between high- and low-conductivity intervals. In contrast, OK predictions reproduce dominant trends but provide substantially smoother profiles. Although

interpolated estimates capture general distribution of hydraulic conductivity, several local extremes and short-scale variations remain unresolved.

Quantitative performance of OK on independent testing dataset is summarized in Table 4.3.

Dataset	RMSE	R^2	KGE	r	α	β
Test	1.49	0.15	0.15	0.43	0.36	0.94

Table 4.3 Performance metrics of OK evaluated using testing boreholes.

RMSE value of 1.49 indicates a moderate average prediction error and reflects difficulty of reproducing highly heterogeneous hydraulic conductivity distribution observed within testing boreholes.

Results indicate that OK provides only a moderate representation of hydraulic conductivity field. The relatively low coefficient of determination ($R^2 = 0.15$) and correlation coefficient ($r = 0.43$) suggest limited ability to reproduce detailed variability observed within testing boreholes. Similarly, Kling–Gupta Efficiency value of 0.15 reflects combined influence of limited correlation and poor reproduction of the observed variability.

Variability ratio $\alpha = 0.36$ further indicates that predicted hydraulic conductivity field is substantially less variable than observations, confirming smoothing effect commonly associated with kriging estimators. Conversely, bias ratio ($\beta = 0.94$) remains close to unity, suggesting that OK does not introduce substantial systematic overestimation or underestimation of hydraulic conductivity.

Overall, these findings demonstrate that, while OK successfully captures broad spatial trends of the aquifer, it is less effective at reproducing pronounced local heterogeneity characteristic of study area. Persistence of these discrepancies suggests that a portion of hydraulic conductivity variability remains unresolved by geostatistical model. This observation motivated a more detailed investigation of prediction residuals using LOOCV framework described in Chapter 3, which forms basis for the discrepancy-correction approaches presented in following sections.

4.3.3 Leave-One-Out Cross-Validation Residual Analysis

After evaluating predictive performance of OK using independent testing boreholes, next step focused on identifying component of variability that remained unexplained by geostatistical estimator. Although test dataset provided an unbiased assessment of predictive capability of OK, it could not be used to generate residual information

required for development of subsequent advanced models, since these data had to remain completely independent for final validation purposes.

To address this limitation, a LOOCV strategy was implemented using only 28 training boreholes. In this procedure, one complete training borehole was temporarily removed from dataset, while remaining 27 boreholes were used to recalibrate OK model. This approach ensured that all observations belonging to the excluded borehole remained unseen during prediction. Recalibrated model was then used to predict hydraulic conductivity values at excluded borehole.

Difference between observed and predicted values at excluded borehole was subsequently calculated to quantify prediction residual:

$$r_i = \log K_{obs,i} - \log K_{OK,i} \quad (4.1)$$

where r_i is residual associated with observation i , $\log K_{obs,i}$ denotes observed log-transformed hydraulic conductivity value, and $\log K_{OK,i}$ represents corresponding OK estimate.

This procedure was repeated sequentially until each of 28 training boreholes had been excluded once and predicted using a model calibrated without its own information. Consequently, a residual value was obtained for every observation contained within training dataset.

Resulting residual dataset represents portion of hydraulic conductivity variability that could not be captured by OK despite accounting for anisotropic spatial structure of data. Importantly, these residuals should not be interpreted as validation errors, since their primary purpose was not to assess model performance. Instead, they were used to investigate whether systematic patterns remained within unexplained variability and whether these patterns could be learned through advanced data-driven techniques.

Residuals generated through LOOCV procedure therefore constituted basis for the subsequent modelling stage, in which machine learning algorithms were employed to predict and correct residual component of OK estimates. By doing so, study aimed to determine whether the integration of data-driven approaches could improve characterization of highly heterogeneous hydraulic conductivity fields beyond what could be achieved through conventional geostatistical estimation alone.

In addition to residual values themselves, a series of descriptive variables was derived for each observation to serve as predictor variables in subsequent GPR and RF discrepancy models. These variables included both geostatistical information obtained from OK framework and descriptors of the local spatial environment. Geostatistical variables comprised quantities directly related to kriging system,

including OK prediction, associated kriging variance, anisotropy-adjusted distance measures, and indicators describing local density of conditioning observations. Furthermore, spatial context variables, such as Cartesian coordinates of observations, their relative depth within aquifer, mean distance to neighbouring samples, and distance to the nearest conditioning point, were also extracted.

Inclusion of these variables was motivated by assumption that magnitude and structure of OK residuals may depend not only on predicted hydraulic conductivity values themselves but also on local spatial configuration of available data. Consequently, final residual dataset consisted of LOOCV residuals together with their associated geostatistical and spatial descriptors. This enriched dataset was subsequently used as input for the advanced data-driven models investigated in this study.

LOOCV procedure produced a complete residual dataset describing component of hydraulic conductivity variability not reproduced by OK model. These residuals were subsequently enriched with geostatistical and spatial descriptors and used in the development of advanced residual prediction models, while seven testing boreholes remained untouched for final independent evaluation. Resulting dataset therefore provided foundation for the residual-modelling approaches investigated in the following section.

4.3.4 Residual Modelling Using Machine Learning

After generating residual dataset through LOOCV procedure and extracting corresponding descriptive variables, two advanced modelling approaches were investigated: RF and GPR.

Objective of this stage was to determine whether variability that remained unexplained by OK contained systematic patterns that could be learned through data-driven methods. Unlike geostatistical estimator, which relies primarily on spatial correlation structure represented by variogram, RF and GPR are capable of identifying complex relationships between predictor variables and residual component of the predictions.

Both RF and GPR models were trained using LOOCV residuals as target values. Predictor variables consisted of geostatistical and spatial-context features described in previous section, including OK estimate, kriging variance, anisotropy-scaled distance measures, neighbourhood characteristics, relative depth, nearest-neighbour information, and spatial coordinates of the observations.

Trained models were subsequently applied to predict residuals at unsampled locations. These predicted residuals represent corrections to OK estimates and

quantify component of hydraulic conductivity variability that could not be reproduced by geostatistical model alone.

By modelling unexplained component of OK predictions, the proposed hybrid framework seeks to improve characterization of highly heterogeneous hydraulic conductivity fields. predicted residuals obtained from each model were combined with corresponding OK estimates to generate corrected hydraulic conductivity predictions. performance of these hybrid approaches was subsequently evaluated using independent testing boreholes and compared with results obtained using OK alone.

4.3.5 Hybrid Prediction and Final Evaluation

After training the Random Forest and Gaussian Process models using the LOOCV residuals, the resulting machine-learning models were used to estimate the structured discrepancy field. The final corrected predictions were then obtained according to Eq. (3.65). This procedure preserves the large-scale spatial structure captured by Ordinary Kriging while compensating for local variability that cannot be reproduced by the geostatistical estimator alone.

To evaluate whether the machine-learning models successfully captured the structured component of the Ordinary Kriging residuals, the distributions of the predicted residuals were compared with the LOOCV residuals used for model training.

Figure 4.7 compares distribution of the true Ordinary Kriging residuals with residuals predicted by RF and GPR models for training dataset. Blue histograms represent true OK residuals obtained from LOOCV procedure, while green and purple histograms represent residuals predicted by RF and GP, respectively. Horizontal bars below each histogram show interquartile range, from 25th to 75th percentile, and vertical black marker indicates median.

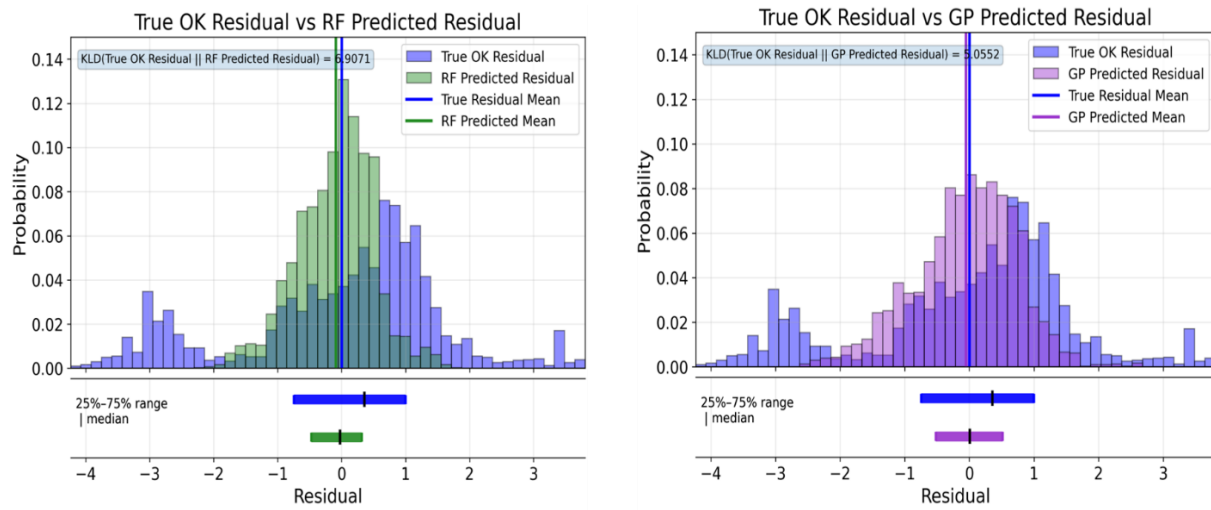


Figure 4.7 Comparison between true and predicted Ordinary Kriging residual distributions for training dataset.

For training dataset, both machine-learning models reproduce central tendency of the residuals, with predicted medians located close to zero. However, RF-predicted residuals are more concentrated around centre and do not fully reproduce broader spread and tails of true residual distribution. In contrast, GP model provides a closer match to overall shape of the true residual distribution. This is confirmed by lower Kullback–Leibler divergence value obtained for GP compared with RF, indicating that GP better captures statistical structure of the OK residuals.

Figure 4.8 presents same comparison for independent testing boreholes. Since these boreholes were not used during model calibration, this analysis provides a more rigorous assessment of generalization capability of the residual models. RF model again produces a narrower residual distribution, suggesting a tendency to smooth residual component and underestimate full variability of the true OK errors. GP model shows a closer overlap with true residual histogram and a more similar interquartile range, indicating improved reproduction of both central tendency and variability of the residuals.

Overall, these results show that both RF and GP are able to learn part of residual structure left unresolved by Ordinary Kriging. However, GP provides a better representation of residual distribution, especially in terms of spread and statistical similarity. This suggests that GP-based residual correction is more effective at capturing uncertainty and variability associated with unexplained component of the OK predictions, whereas RF provides a more localized but smoother correction.

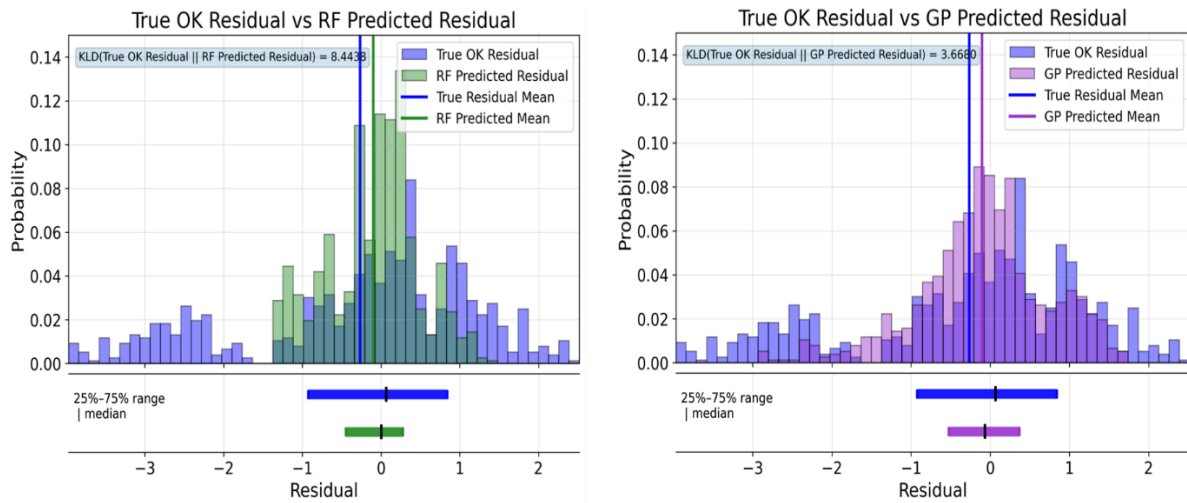


Figure 4.8 Comparison between true and predicted Ordinary Kriging residual distributions for testing dataset.

Impact of residual correction on final hydraulic conductivity estimates is illustrated in Figure 4.9. Observed hydraulic conductivity field within testing boreholes is compared with predictions obtained using Ordinary Kriging, RF hybrid model, and Gaussian Process hybrid model.

Ordinary Kriging reproduced the broad spatial trends of hydraulic conductivity but failed to capture several localized variations associated with highly heterogeneous nature of the aquifer. The hybrid approaches introduced additional spatial details by correcting OK estimates through residual modelling. Visually, Gaussian Process hybrid appeared to better reproduce the observed spatial patterns, whereas RF hybrid provided localized corrections in regions characterized by larger discrepancies.

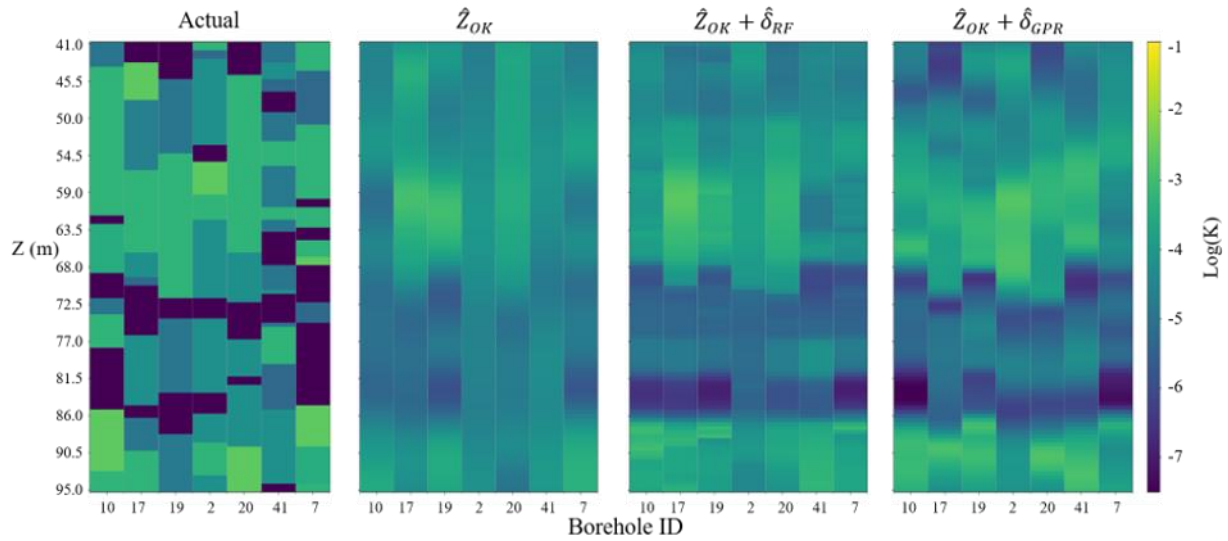


Figure 4.9 Comparison between observed hydraulic conductivity values and predictions obtained using Ordinary Kriging, OK-RF, and OK-GP for testing boreholes.

Finally, predictive performance of different approaches was quantitatively evaluated using independent testing dataset. Figure 4.10 provides a visual comparison of principal performance metrics, while Table 4.4 summarizes values of RMSE, coefficient of determination (R^2), Kling-Gupta Efficiency (KGE), correlation coefficient (r), variability ratio α , and bias ratio β obtained for Ordinary Kriging and two hybrid approaches.

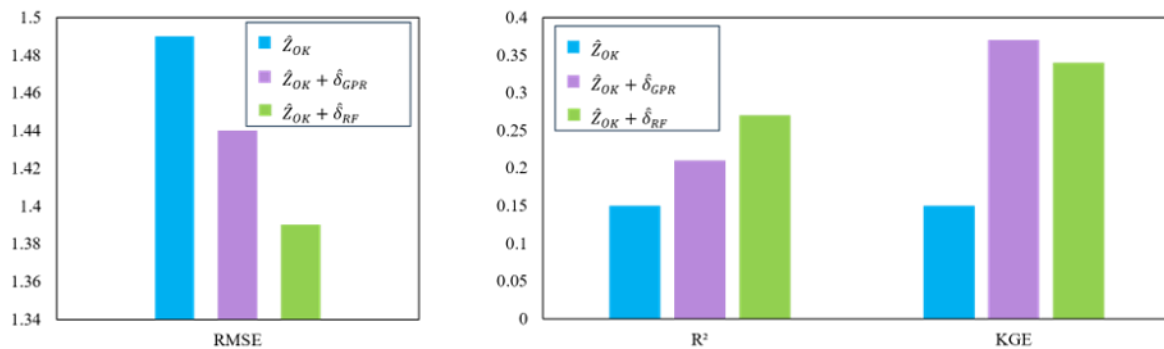


Figure 4.10 Comparison of RMSE, R^2 , and KGE obtained using Ordinary Kriging (OK), OK-RF, and OK-GP on independent testing dataset.

Results demonstrate that both hybrid approaches improved predictive performance of Ordinary Kriging. Relative to baseline OK model, both GPR and RF reduced prediction error and increased explanatory capability of model. RF-corrected approach achieved the lowest RMSE (RMSE=1.39) and the highest coefficient of

determination ($R^2 = 0.27$), indicating the greatest improvement in overall predictive accuracy. In contrast, GPR-corrected approach achieved the highest Kling–Gupta Efficiency (KGE = 0.37) and the largest variability ratio ($\alpha = 0.63$), suggesting a superior ability to reproduce statistical characteristics and variability of observed hydraulic conductivity field.

Dataset	RMSE	R^2	KGE	r	α	β
$\hat{Z}_{OK}$	1.49	0.15	0.15	0.43	0.36	0.94
$\hat{Z}_{OK} + \hat{\delta}_{GP}$	1.44	0.21	0.37	0.49	0.63	0.96
$\hat{Z}_{OK} + \hat{\delta}_{RF}$	1.39	0.27	0.34	0.53	0.53	0.96

Table 4.4 Quantitative comparison of Ordinary Kriging and hybrid models using independent testing dataset.

Correlation coefficient increased from 0.43 for Ordinary Kriging to 0.49 and 0.53 for GPR- and RF-corrected models, respectively, indicating an improved correspondence between observed and predicted values. Furthermore, both hybrid approaches maintained bias ratios close to unity ($\beta = 0.96$), demonstrating that residual-correction procedure did not introduce substantial systematic bias.

Overall, the results indicate that machine-learning-based residual correction enhances hydraulic conductivity prediction in highly heterogeneous aquifers. While RF-corrected model provided the highest predictive accuracy, GPR-corrected model more effectively reproduced observed variability and statistical structure of hydraulic conductivity field.

5. ATES Simulations

5.1 Introduction

Hydraulic conductivity characterization approaches presented in Chapter 4 demonstrated that different modelling techniques produce distinct representations of aquifer heterogeneity. While statistical validation provides a quantitative assessment of predictive accuracy of these approaches, it does not directly indicate how hydraulic conductivity fields influence groundwater flow, heat transport, and overall performance of an ATES system. Consequently, practical implications of different characterization methods must be evaluated within a numerical simulation framework that reproduces coupled hydraulic and thermal processes governing seasonal energy storage.

This chapter presents design and implementation of numerical ATES model developed for Parco Sempione study area using MRST (REFERENCE). Model simulates coupled groundwater flow and heat transport under seasonal operating conditions representative of ATES system. A three-dimensional finite-volume formulation is employed to evaluate long-term evolution of pressure and temperature distributions, thermal plume migration, and energy storage behaviour within the aquifer.

To isolate influence of subsurface characterization on ATES performance, identical numerical settings were adopted for all simulations. Operating schedule, well configuration, injection temperatures, pumping conditions, boundary conditions, and thermophysical properties remained unchanged throughout the study. Consequently, the hydraulic conductivity field represents only varying input parameter among the investigated scenarios.

Four hydraulic conductivity representations were considered. First consists of a homogeneous hydraulic conductivity field, which serves as a reference case by neglecting geological heterogeneity. Remaining three fields were obtained using characterization approaches developed in Chapter 4, namely deterministic OK field

and two hybrid conductivity fields corrected using RF (OK–RF) and GPR (OK–GPR). Comparing these four representations enables a systematic assessment of how different descriptions of subsurface heterogeneity influence hydraulic and thermal response of ATES system.

Chapter first describes ATES operating strategy and well configuration adopted for numerical simulations. Subsequently, numerical model setup and governing simulation parameters are presented. Finally, hydraulic conductivity fields used as inputs to numerical simulations are presented. Simulation results obtained using these different hydraulic conductivity representations are presented and discussed in Chapter 6.

5.2 ATES System Design and Operating Strategy

Design of ATES system was established based on hydrogeological characteristics of Aquifer B and numerical framework described in previous chapters. A three-dimensional coupled groundwater flow and heat transport model was developed using MRST to simulate long-term ATES system under representative seasonal conditions [3], [55].

To enable a consistent evaluation of influence of subsurface heterogeneity, ATES system was operated using identical well configurations, operating schedules, hydraulic controls, thermal boundary conditions, and fluid and rock properties throughout all numerical simulations. Consequently, any differences observed in groundwater flow, thermal plume evolution, energy storage, or recovery efficiency can be attributed exclusively to differences in the hydraulic conductivity distribution used in each simulation.

Prior to ATES operation, numerical model was initialized through a long-term pre-equilibrium simulation to establish stable hydraulic and thermal conditions within aquifer. Following this initialization stage, ATES system was operated under a cyclic seasonal schedule consisting of alternating heat-storage and cold-storage seasons separated by intermediate resting periods. This operating sequence was repeated over a five-year simulation period in order to investigate the long-term evolution of pressure, groundwater flow, temperature distribution, thermal plume migration, and energy recovery [38], [40].

Following sections describe operating schedule, well configuration, design criteria, and operational parameters adopted in numerical simulations.

5.2.1 Operating Schedule

Seasonal operating strategy adopted for ATES system was designed to reflect the climatic conditions of Milan metropolitan area, which is characterized by a humid continental climate with warm summers and relatively cold winters. Consequently, annual thermal energy demand alternates between cooling in summer and heating in winter, making low-temperature Aquifer Thermal Energy Storage (LT-ATES) an appropriate technology for seasonal thermal energy management [3], [55].

Prior to cyclic operation of ATES system, a 50-year pre-equilibrium simulation was performed to establish stable groundwater flow and thermally equilibrated initial conditions within aquifer. This initialization step eliminates artificial transients associated with arbitrary initial conditions and ensures that subsequent ATES simulations begin from a realistic hydraulic and thermal equilibrium state [38], [40].

Following the pre-equilibrium simulation, ATES system was simulated over a five-year operational period, during which same annual operating cycle was repeated. As shown in Figure 5.1, adopted operating schedule consists of four successive phases:

- Warm-water storage period (May–August): Warm water is injected into warm storage wells, while cold groundwater is simultaneously extracted from cold storage wells to satisfy summer cooling demand.
- Recovery period (September–October): All wells remain inactive, allowing groundwater flow together with heat advection, thermal conduction, and thermal dispersion to govern heat redistribution within aquifer.
- Cold-water storage period (November–February): Cold water is injected into cold storage wells, while warm groundwater is extracted from warm storage wells to supply winter heating demand.
- Recovery period (March–April): A second inactive period is introduced before beginning of the next annual cycle, allowing further redistribution of thermal energy within aquifer prior to the subsequent storage phase.

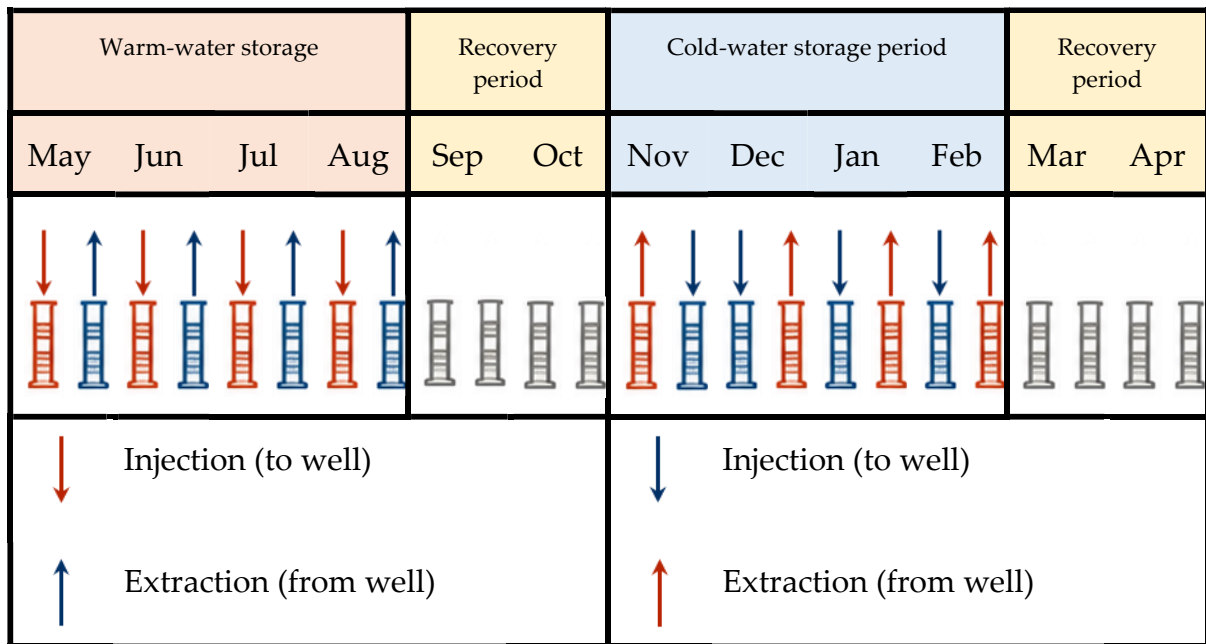


Figure 5.1 Annual operating schedule adopted for simulated ATEs system.

Annual operating sequence was repeated identically throughout five-year simulation period. By maintaining identical operating schedules, well configurations, injection temperatures, pumping conditions, and boundary conditions throughout all simulations, any differences in hydraulic and thermal behaviour can be attributed exclusively to hydraulic conductivity field employed in each numerical model.

5.2.2 Well Configuration

Performance of an ATEs system is strongly influenced by spatial arrangement of injection and production wells. In practical ATEs systems, warm and cold wells are generally positioned to maximize their separation and minimize thermal interference between the corresponding thermal plumes. Objective of the present study, however, is not to optimize ATEs performance. Instead, the adopted five-spot well configuration was intentionally selected to increase the sensitivity of the numerical simulations to spatial variations in hydraulic conductivity. This configuration provides a suitable framework for evaluating how different hydraulic conductivity representations influence the predicted thermo-hydraulic behaviour of ATEs system.

Unlike conventional ATEs systems, which are commonly based on a separated warm and cold well doublets, numerical model developed in this study employs a customized five-spot multi-well configuration consisting of thirteen fully penetrating wells, including four warm wells and nine cold wells.

Well configuration adopted in numerical simulations is illustrated in Figure 5.2.

The warm wells are arranged in a 2×2 array near centre of the model domain, while cold wells are distributed in a surrounding 3×3 array, forming an approximately symmetric storage configuration. adopted layout consists of a central group of warm wells surrounded by a peripheral group of cold wells, providing spatial separation between two storage regions while defining thermal storage domain used in numerical simulations.

Throughout simulations, each well retained its thermal designation (warm or cold), whereas its hydraulic operating mode changed according to the seasonal operating schedule. During summer storage period, warm water was injected into four warm wells while cold groundwater was simultaneously extracted from nine cold wells to satisfy cooling demand. During winter storage period, operating modes were reversed: cold water was injected into cold wells, whereas warm groundwater was extracted from warm wells to supply heating. During intermediate recovery periods, all wells remained inactive, allowing groundwater flow together with conductive, advective, and dispersive heat transport to redistribute thermal energy within aquifer.

Selected well layout was adopted consistently throughout all numerical simulations. Because identical well locations, operating schedules, injection temperatures, pumping rates, and boundary conditions were used in every simulation, hydraulic conductivity field represents the only varying input parameter among the investigated scenarios. Consequently, any differences observed in simulated hydraulic and thermal responses can be attributed exclusively to differences in representation of subsurface heterogeneity.

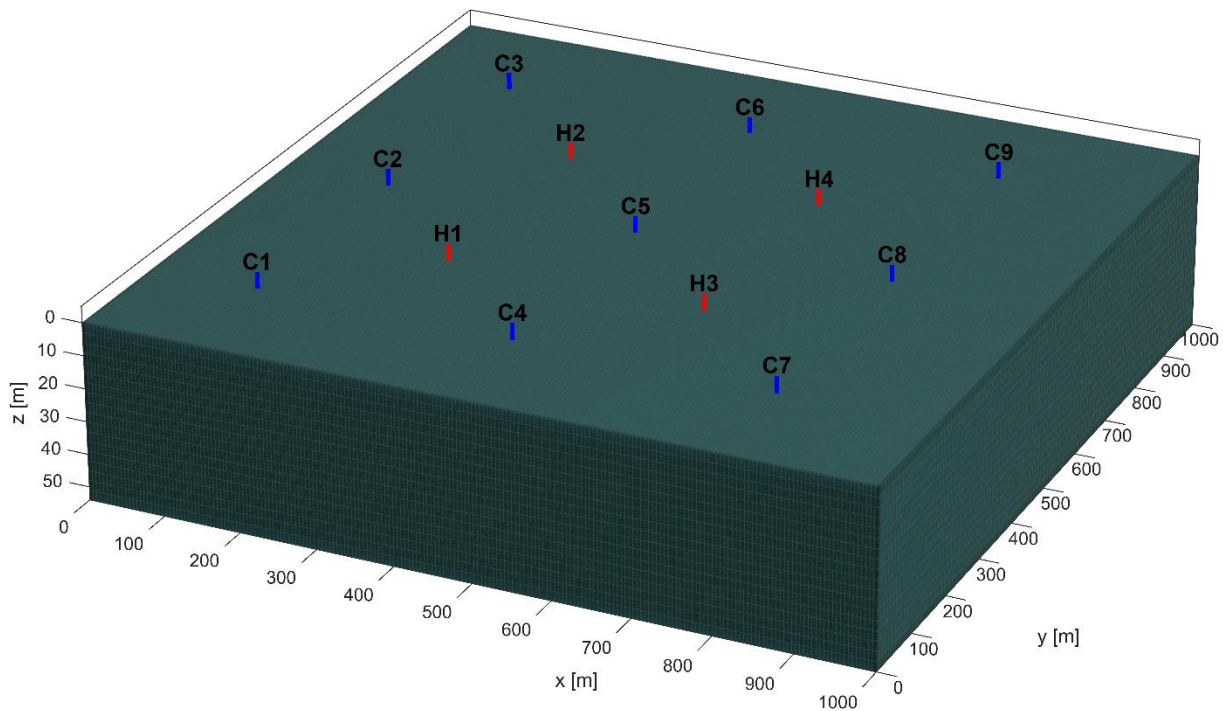


Figure 5.2 Spatial arrangement of warm and cold wells used in numerical ATES simulations.

5.2.3 Well Spacing Design Criteria

Spacing between injection and production wells is one of the most important design parameters in ATES systems, as it directly influences thermal recovery efficiency, hydraulic interaction, and long-term system performance. Adequate well spacing limits thermal interference between warm and cold storage zones, thereby reducing risk of premature thermal breakthrough and improving efficiency of seasonal heat storage and recovery [3], [55]. Conversely, insufficient separation may result in early mixing of thermal plumes, reducing amount of recoverable thermal energy and compromising long-term system operation.

Several analytical and empirical approaches have been proposed to support preliminary design of ATES systems. Among the most widely adopted methods is concept of thermal radius, introduced by Sommer et al., which estimates the maximum radial extent of thermal storage volume surrounding an injection well under idealized conditions [56]. This analytical formulation assumes homogeneous aquifer properties and neglects influence of regional groundwater flow, thermal conduction, and thermal dispersion. Building upon this concept [55], demonstrated that, in aquifers affected by significant regional groundwater flow, well spacing should also account for downstream displacement of thermal plume caused by groundwater advection [55].

Although these analytical relationships provide valuable guidance during preliminary design stage, they are primarily intended for simplified or site-specific optimization studies. In the present work, objective is not to optimize ATES layout but to investigate influence of different hydraulic conductivity representations on hydraulic and thermal behaviour of storage system. Consequently, the well geometry and operational strategy were intentionally maintained unchanged throughout all simulations so that the influence of different hydraulic conductivity representations could be evaluated independently of changes in system design.

Based on previous ATES design recommendations and considering dimensions of the computational domain, well spacing adopted in numerical model was selected to provide adequate separation between warm and cold storage zones while ensuring that all wells could be accommodated within simulation domain. Principal geometric and hydrogeological parameters used for well-spacing design are summarized in Table 5.1.

To verify that adopted well spacing is appropriate under regional groundwater flow conditions of study area, groundwater seepage velocity was estimated using Darcy's law [15]. hydraulic gradient was assumed to be 0.001, which is representative of regional groundwater flow in study area. Darcy flux was therefore calculated as

$$q = -K \frac{dh}{dx} \quad (6.1)$$

where q is Darcy flux (m/s), K is hydraulic conductivity (m/s), and dh/dx is regional hydraulic gradient.

Using average hydraulic conductivity of Aquifer B adopted for preliminary design calculations,

$$K = 1.44 \times 10^{-4} \text{ (m/s)} \quad (6.2)$$

together with assumed regional hydraulic gradient,

$$\frac{dh}{dx} = 0.001 \quad (6.3)$$

Resulting Darcy flux is

$$q = 1.44 \times 10^{-7} \text{ (m/s)} \quad (6.4)$$

Since Darcy flux represents volumetric discharge per unit cross-sectional area rather than actual groundwater velocity, seepage velocity was computed as

$$v = \frac{q}{\phi_e} \quad (6.5)$$

where ϕ_e denotes effective porosity of aquifer.

Assuming a constant porosity throughout numerical domain, $\phi_e = 0.13$ groundwater seepage velocity was estimated to be approximately

$$v = \frac{1.44 \times 10^{-7}}{0.13} \times 365 \times 24 \times 3600 = 35 \text{ (m/year)} \quad (6.6)$$

Estimated seepage velocity of approximately 35 m/year indicates that groundwater advection is sufficiently significant to influence thermal plume migration. Estimated seepage velocity indicates that groundwater advection is sufficiently significant to influence thermal plume migration. Adopted well spacing was therefore evaluated with respect to these groundwater-flow conditions and subsequently maintained unchanged throughout all numerical simulations.

Principal design parameters adopted for numerical simulations are summarized in Table 5.2.

Parameter	Value	Unit
Hydraulic conductivity	1.44×10^{-4}	m/s
hydraulic gradient	0.001	m/m
Specific discharge	1.44×10^{-7}	m/s
Seepage velocity	35	m/year
Distance between wells of different temperature	150	m
Distance between wells of same temperature	36	m

Table 5.1 ATEs system design parameters accounting for seepage velocity of the groundwater.

Adopted well spacing and well geometry were maintained unchanged throughout all numerical simulations. Likewise, operating schedule, injection temperatures, pumping rates, and boundary conditions were kept identical for every simulation. Consequently, hydraulic conductivity field remained the only varying input parameter among the investigated scenarios, namely homogeneous field, OK field, OK–RF field, and OK–GPR field. This strategy enables a direct assessment of the influence of subsurface heterogeneity on ATEs performance while eliminating effects of operational or geometric variability.

5.2.4 Well Operation

The operation of the ATES system followed annual schedule described in Section 5.2.1. During each operating cycle, the wells alternated between injection and production modes according to seasonal heating and cooling demands. The operating strategy remained identical throughout all numerical simulations to ensure that any differences in system performance could be attributed exclusively to hydraulic conductivity field.

Numerical model represents ATES system. Based on average natural groundwater temperature of approximately 15 °C in study area, warm and cold water were injected at temperatures above and below ambient groundwater temperature, respectively. Selected injection temperatures are consistent with values commonly reported for ATES systems and provide sufficient thermal contrast while limiting undesirable hydrogeochemical effects [3].

During storage phases, each injection well was operated at a constant injection rate of 500 m³/day (5.79×10^{-3} m³/s), while production wells were controlled using a constant bottom-hole pressure (BHP) constraint. This operating strategy ensures stable groundwater circulation throughout storage system while allowing production rates to adjust naturally to pressure variations generated by injection process. Production BHP was prescribed at 0.1 bar below initial reservoir pressure, providing a slight pressure sink that promotes continuous groundwater circulation while minimizing excessive drawdown, over-pressurization, and numerical instability within aquifer.

Same operating conditions were applied throughout all numerical simulations, including homogeneous hydraulic conductivity field, OK field, OK–RF field, and OK–GPR field. Consequently, hydraulic conductivity field remained only varying input parameter among investigated scenarios, allowing influence of subsurface heterogeneity on ATES performance to be evaluated independently of operational conditions.

Principal operating and thermophysical parameters adopted in numerical simulations are summarized in Table 5.3.

Parameter	Value	Unit
Specific heat capacity of water	4200	$J/Kg.K$
Specific heat capacity of solid	1000	$J/Kg.K$
Screen length	23	m
Volumetric heat capacity of water	4200000	$J/m^3.K$
Volumetric heat capacity of solid	1600000	$J/m^3.K$
Volumetric heat capacity of aquifer	1938000	$J/m^3.K$
Total volume injected per cycle (either warm or cold)	246968	m^3
Injection time per cycle	2920	hours
Injection rate (all injection wells)	0.00579	m^3/s
Bottom-hole pressure	2.26	barsa
Hot water injection temperature	22	$^{\circ}C$
Cold water injection temperature	8	$^{\circ}C$

Table 5.2 Operating parameters used in numerical simulations.

5.3 Hydraulic Conductivity Fields

The numerical simulations presented in this study were performed using four different representations of hydraulic conductivity field. These hydraulic conductivity fields represent progressively increasing levels of geological complexity and were selected to investigate influence of subsurface heterogeneity on groundwater flow, heat transport, and overall performance of ATEs system. While the operating strategy, well configuration, numerical grid, and boundary conditions remained identical for all simulations, hydraulic conductivity distribution was varied to evaluate sensitivity of the simulated thermal behaviour to different levels of geological complexity. This approach allows influence of geological heterogeneity to be evaluated independently of operational and numerical modelling parameters.

Four hydraulic conductivity fields considered in the numerical simulations are:

- a homogeneous hydraulic conductivity field;
- a deterministic field generated using OK
- a hybrid field obtained by correcting OK predictions using RF (OK–RF);
- a hybrid field obtained by correcting OK predictions using GPR (OK–GPR).

Numerical simulations were performed on a structured Cartesian grid generated using MRST. Computational domain represents a volume of $1000 \times 1000 \times 54$ m discretized into $205 \times 205 \times 18$ grid cells, corresponding to a horizontal resolution of approximately 4.89 m and a vertical resolution of 3 m. Principal characteristics of computational grid are summarized in Table 5.3

Parameter	Domain size [m]	Number of grid cells	Grid resolution [m]
x-direction	1000	205	4.89
y-direction	1000	205	4.89
z-direction	54	18	3

Table 5.3 Computational domain and structured grid characteristics used in MRST simulations.

Following subsections describe generation of each hydraulic conductivity field and its implementation within numerical ATEs model.

5.3.1 Homogeneous Hydraulic Conductivity Field

A homogeneous hydraulic conductivity field was considered as reference case to evaluate influence of geological heterogeneity on hydraulic and thermal performance of ATEs system. In this representation, hydraulic conductivity is assumed to be spatially uniform throughout computational domain, thereby neglecting all lateral and vertical variations in aquifer properties. Although this assumption does not reflect natural complexity of sedimentary aquifers, homogeneous models are widely adopted as baseline scenarios for assessing influence of subsurface heterogeneity on groundwater flow and heat transport[3], [55].

Homogeneous hydraulic conductivity was defined as the arithmetic mean of measured hydraulic conductivity values available within study area. Prior to numerical simulations, hydraulic conductivity field was converted to intrinsic permeability using Equation (3.9), as required by MRST simulator. Resulting permeability was then assigned uniformly to every computational cell, producing a spatially homogeneous representation of aquifer.

Resulting homogeneous hydraulic conductivity field is illustrated in Figure 5.3. As expected, field exhibits no spatial variability, and all grid cells possess identical hydraulic properties. This homogeneous representation serves as reference case against which OK, OK–RF, and OK–GPR hydraulic conductivity fields are subsequently compared.

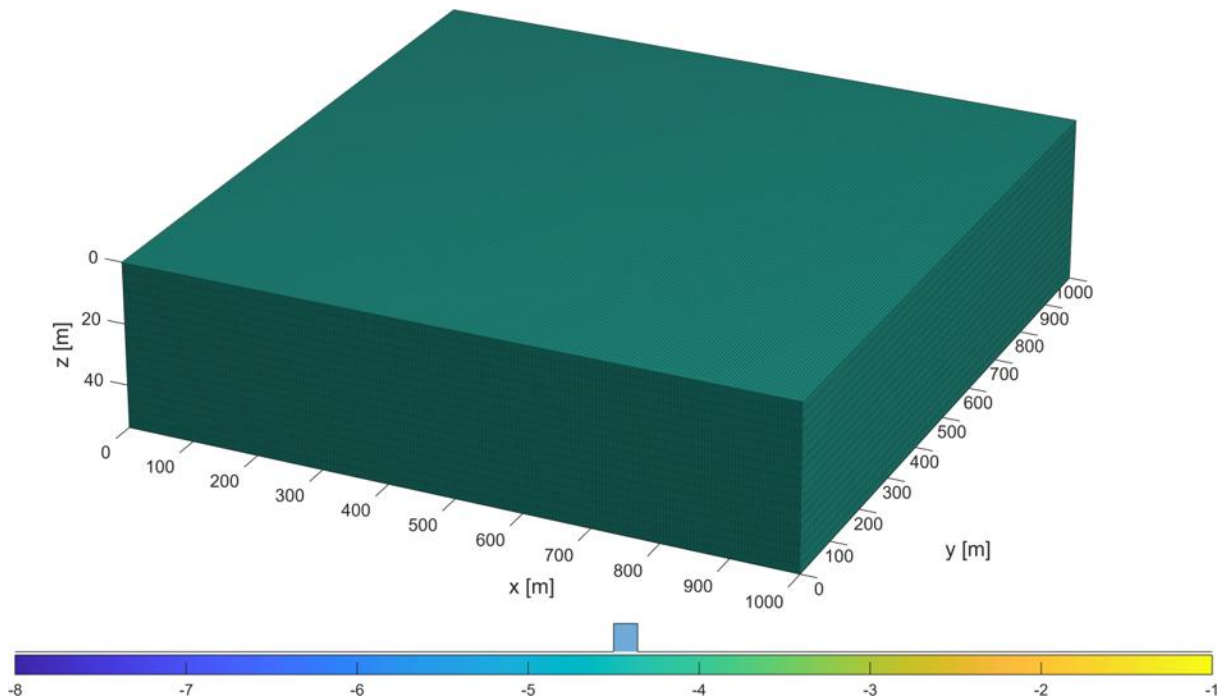


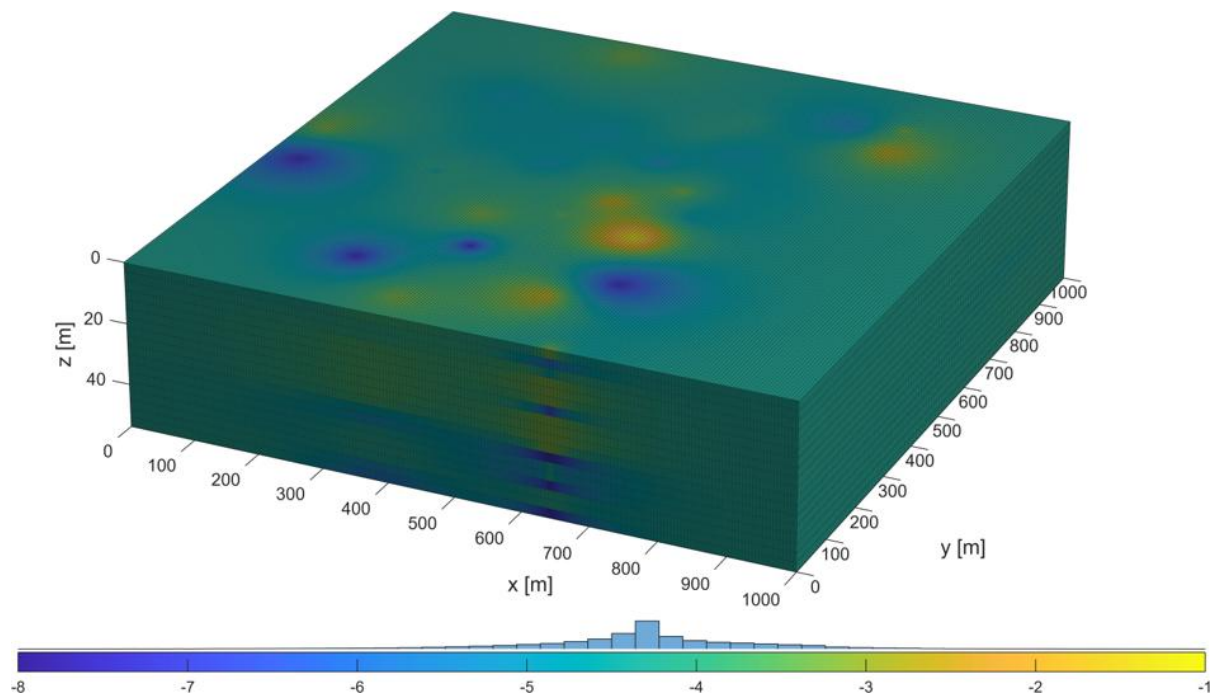
Figure 5.3 Homogeneous logarithmic hydraulic conductivity field adopted as the reference case for numerical ATEs simulations.

5.3.2 Heterogeneous Hydraulic Conductivity Fields

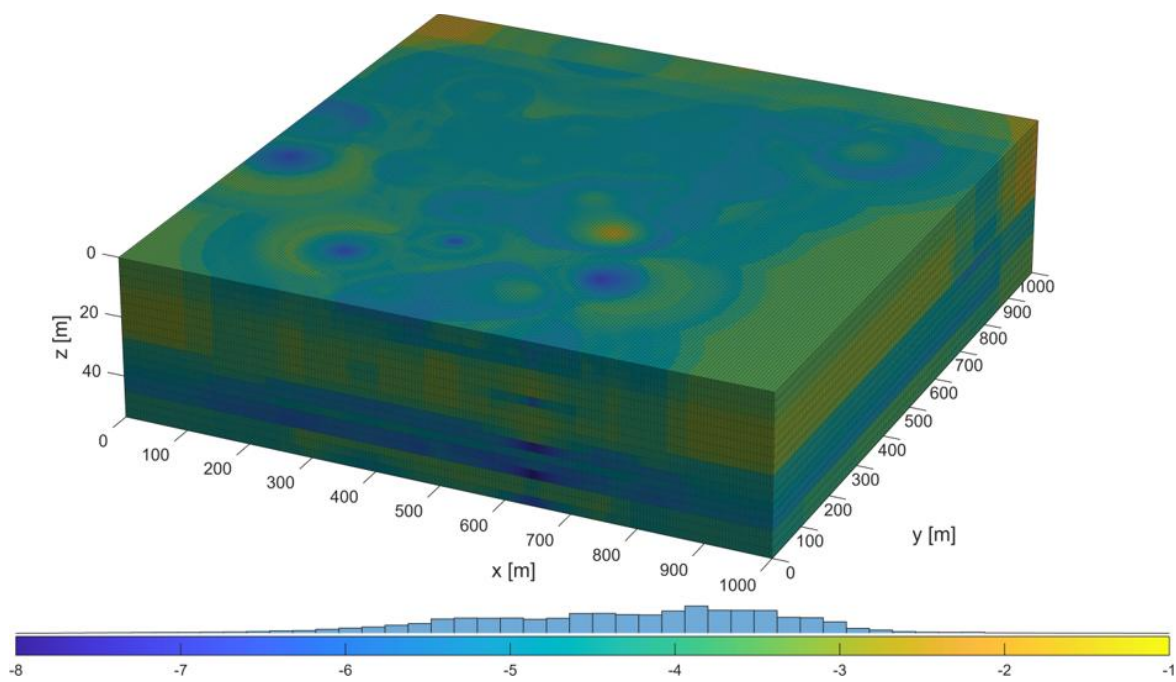
Three heterogeneous hydraulic conductivity fields were used in numerical simulations to evaluate influence of different subsurface characterization approaches on ATEs performance. These fields were developed in Chapter 4 using geostatistical and hybrid geostatistical–machine learning methods and subsequently mapped onto MRST computational grid for numerical simulation.

First heterogeneous representation corresponds to OK field, which provides a deterministic geostatistical reconstruction of spatial distribution of logarithmic hydraulic conductivity based on calibrated variogram model.

Second and third heterogeneous representations correspond to hybrid OK–RF and OK–GPR fields. These conductivity fields were generated by applying Random Forest and GPR residual-correction techniques to deterministic Ordinary Kriging field, as described in Chapter 4. By incorporating machine-learning-based residual corrections, these approaches provide a more detailed representation of local-scale geological variability while preserving large-scale spatial structure estimated by Ordinary Kriging.



(a) OK



(b) OK-RF

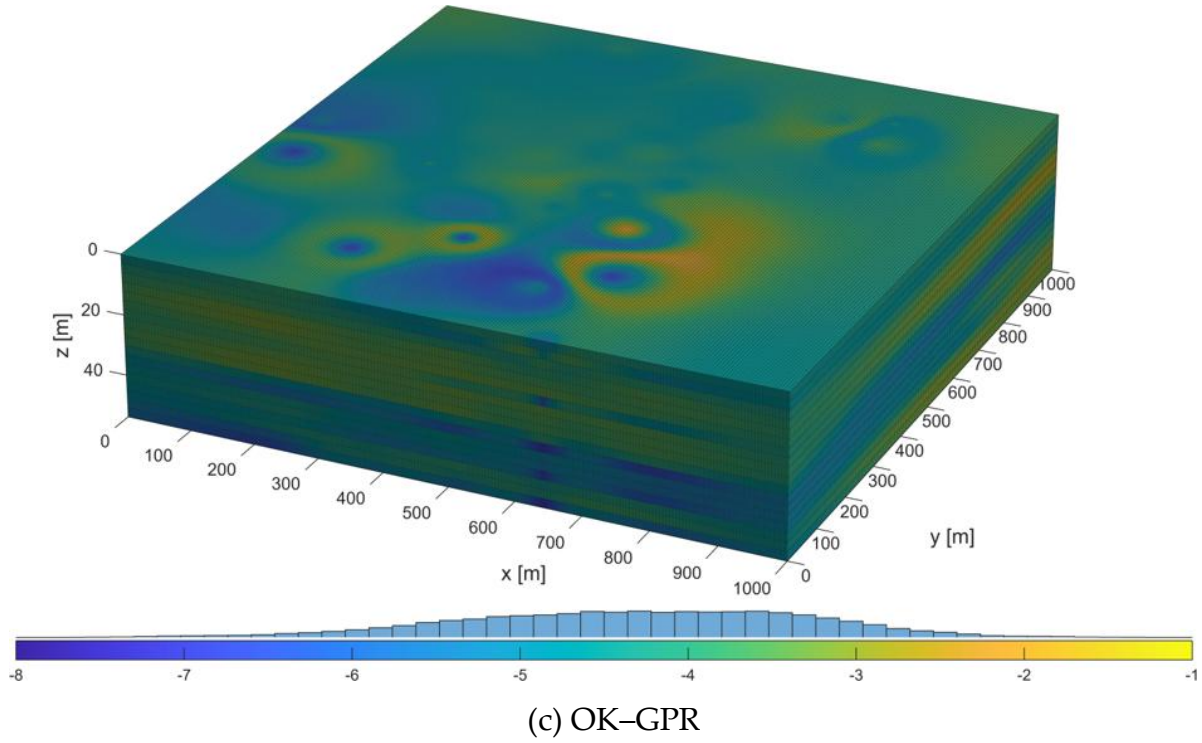


Figure 5.4 Heterogeneous logarithmic hydraulic conductivity fields used in numerical ATES simulations: (a) Ordinary Kriging (OK), (b) OK-RF, and (c) OK-GPR.

Figure 5.4 compares three heterogeneous logarithmic hydraulic conductivity fields used in numerical simulations. While OK field represents large-scale spatial variability of aquifer, OK-RF and OK-GPR fields exhibit progressively finer spatial heterogeneity resulting from machine-learning residual corrections.

Together with homogeneous reference case presented in Section 5.3.1, these heterogeneous hydraulic conductivity fields constitute complete set of subsurface models used in numerical ATES simulations.

5.4 Summary of Simulation Scenarios

This chapter presented the design and implementation of the numerical ATES model developed for Parco Sempione study area. Operating schedule, well configuration, operating conditions, and hydraulic conductivity adopted in the numerical simulations were described in detail. Four hydraulic conductivity fields were considered: a homogeneous reference field, a deterministic OK field, and two hybrid fields generated using OK-RF and OK-GPR.

Throughout all simulations, identical well locations, operating schedules, injection temperatures, pumping conditions, thermophysical properties, boundary conditions, and numerical settings were maintained. Consequently, the hydraulic conductivity representation remained the only varying input parameter. The homogeneous hydraulic conductivity field serves as the reference simulation throughout the subsequent analyses, and all differences are interpreted relative to this reference case.

5.5 Numerical Results

Numerical ATES model described in the preceding sections was used to investigate the influence of the different hydraulic conductivity representations on the predicted thermo-hydraulic response of the system. Objective of these simulations is not to determine which hydraulic conductivity representation is more realistic or superior, but rather to evaluate how alternative hydraulic conductivity representations influence the predicted groundwater flow, heat transport, and well responses relative to the homogeneous reference simulation.

Results first examine the evolution of simulated thermal plumes to illustrate how hydraulic conductivity influences the spatial distribution of thermal energy. The seasonal heat and cold recovery factors are then analysed, followed by the thermal and hydraulic responses of the individual wells, including production temperatures, well flow rates, and bottom-hole pressures. Finally, results are synthesised to evaluate the overall influence of the different hydraulic conductivity representations on the predicted thermo-hydraulic response of the numerical ATES model.

5.6 Thermal Plume Evolution

Evolution of the simulated thermal plumes illustrates how different hydraulic conductivity representations influence the spatial distribution of thermal energy during ATES operation. Since identical operating conditions, well configuration, boundary conditions, pumping schedules, and thermophysical properties were adopted for all simulations, the observed differences are solely attributable to the hydraulic conductivity field. The homogeneous hydraulic conductivity field is considered the reference simulation, and the remaining scenarios are interpreted relative to this reference case.

To evaluate both the short-term and long-term influence of hydraulic conductivity heterogeneity, the thermal plumes are examined after the first and fifth operational cycles.

5.6.1 Thermal plume evolution during the first operating cycle

Figures 5.5 and 5.6 present the simulated thermal plumes at the end of hot-water (Month 4) and cold-water (Month 10) injection periods during the first operational cycle. At this stage, thermal transport is primarily controlled by the initial propagation of the injected thermal fronts.

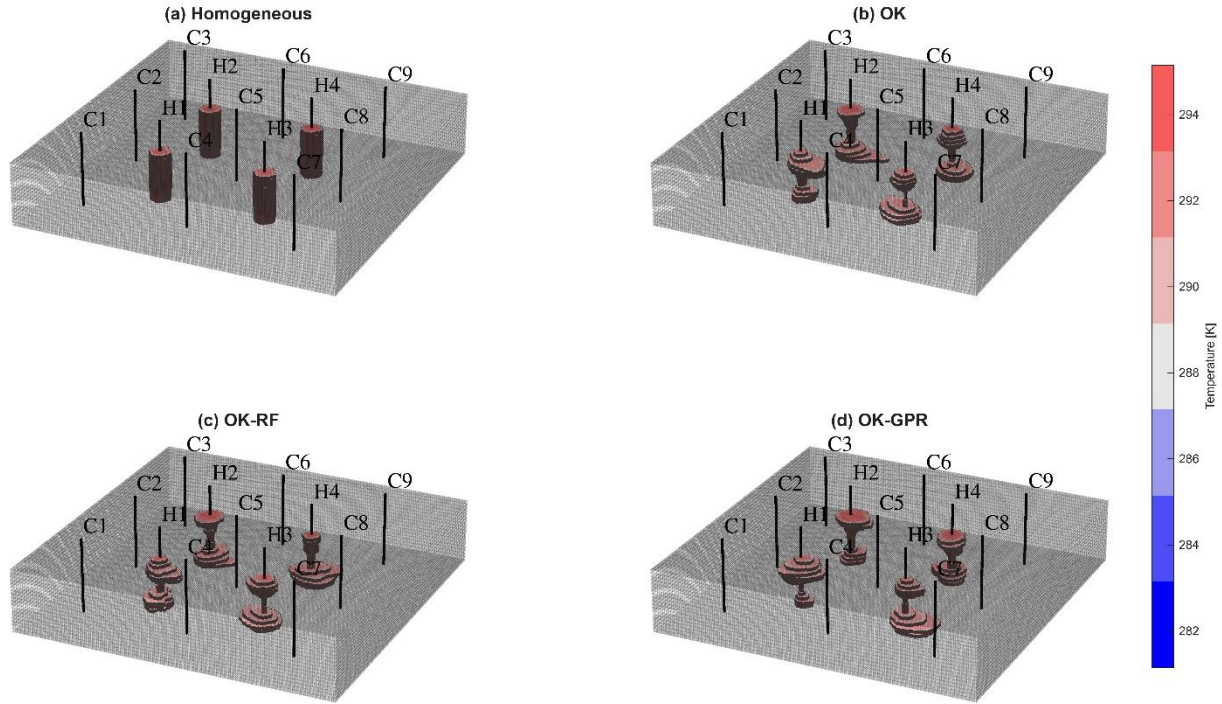


Figure 5.5 Comparison of the simulated thermal plumes at the end of hot-water injection period (Month 4) of the first operating cycle for the four hydraulic conductivity models: (a) Homogeneous, (b) OK, (c) OK-RF, and (d) OK-GPR.

Compared with the homogeneous reference simulation, all heterogeneous hydraulic conductivity fields produce asymmetric plume geometries. While the homogeneous model generates nearly circular thermal plumes because of the spatially uniform hydraulic conductivity, the heterogeneous fields produce localized elongation and contraction of the thermal fronts, reflecting the influence of spatial variations in hydraulic conductivity on groundwater flow and heat transport.

Differences among the three heterogeneous representations remain relatively limited during the first operating cycle. OK-RF simulation generally exhibits more irregular plume boundaries than the deterministic OK simulation, whereas OK-GPR simulation produces comparatively smoother plume geometries. Nevertheless, the overall differences among the heterogeneous fields remain modest because thermal transport has occurred over only one seasonal operating cycle.

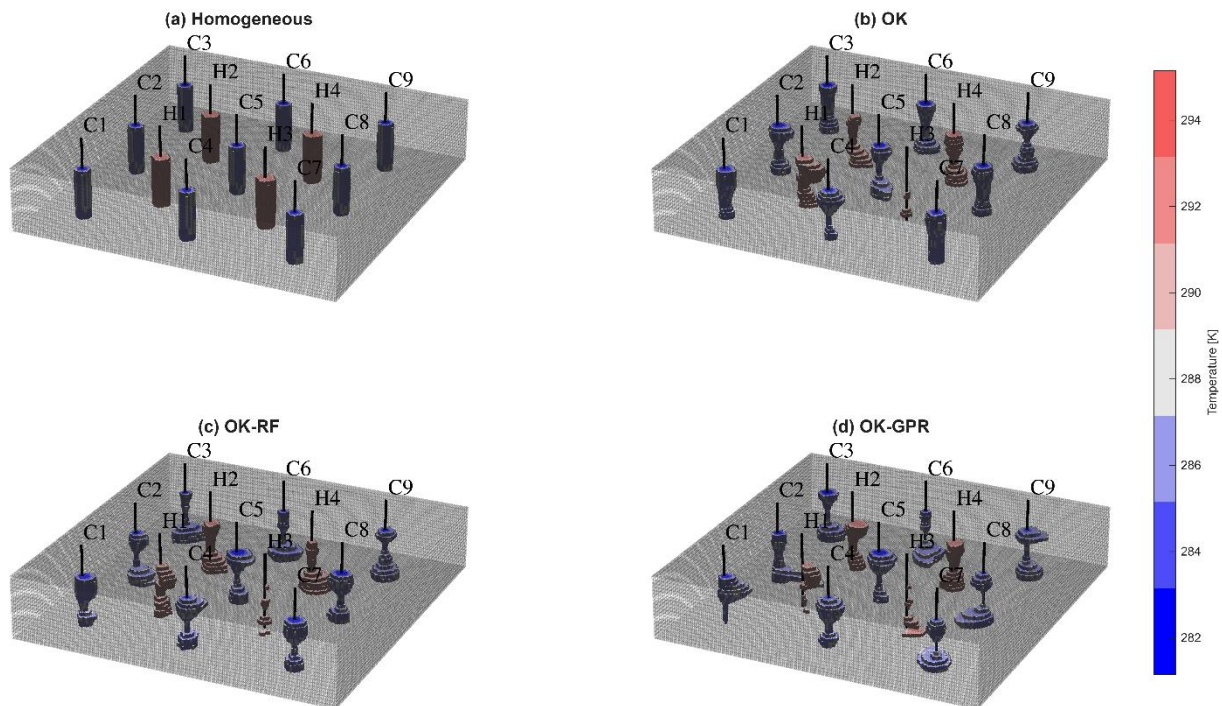


Figure 5.6 Comparison of the simulated thermal plumes at the end of cold-water injection period (Month 10) of the first operating cycle for the four hydraulic conductivity models: (a) Homogeneous, (b) OK, (c) OK-RF, and (d) OK-GPR.

5.6.2 Thermal plume evolution after five operating cycles

Influence of hydraulic conductivity representation becomes more evident after five years of seasonal operation (Figures 5.7 and 5.8). Repeated injection and production cycles progressively redistribute thermal energy throughout the aquifer, allowing the cumulative influence of groundwater flow to become more apparent.

Relative to the first operating cycle, the heterogeneous hydraulic conductivity fields produce increasingly asymmetric plume geometries and larger differences in thermal plume extent. In contrast, the homogeneous reference simulation continues to generate nearly symmetric plume patterns because groundwater flow remains spatially uniform throughout the domain.

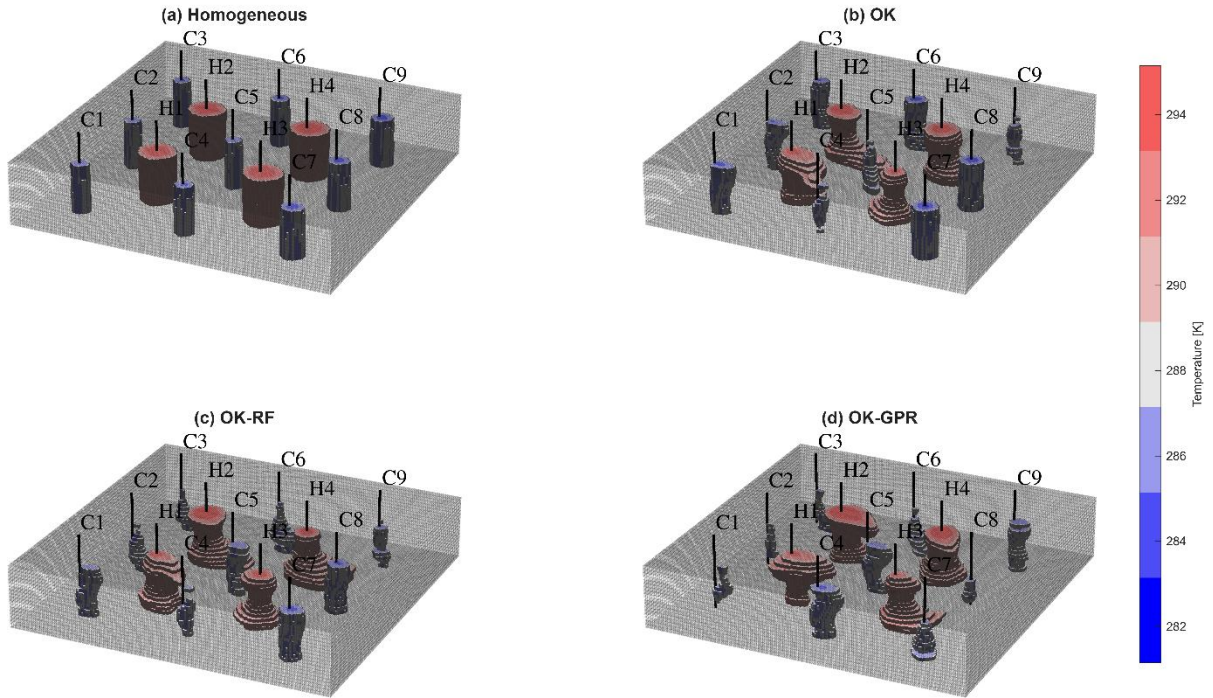


Figure 5.7 Comparison of the simulated thermal plumes at the end of hot-water injection period (Month 4) of the fifth operating cycle for the four hydraulic conductivity models: (a) Homogeneous, (b) OK, (c) OK-RF, and (d) OK-GPR.

Although the three heterogeneous simulations display different plume geometries, they all deviate from the homogeneous reference model by producing non-uniform thermal plume expansion. OK-RF simulation generally exhibits greater local irregularity, whereas OK-GPR simulation produces smoother but still asymmetric plume boundaries. These differences indicate that the predicted long-term thermal response of ATEs system is sensitive to the adopted hydraulic conductivity representation.

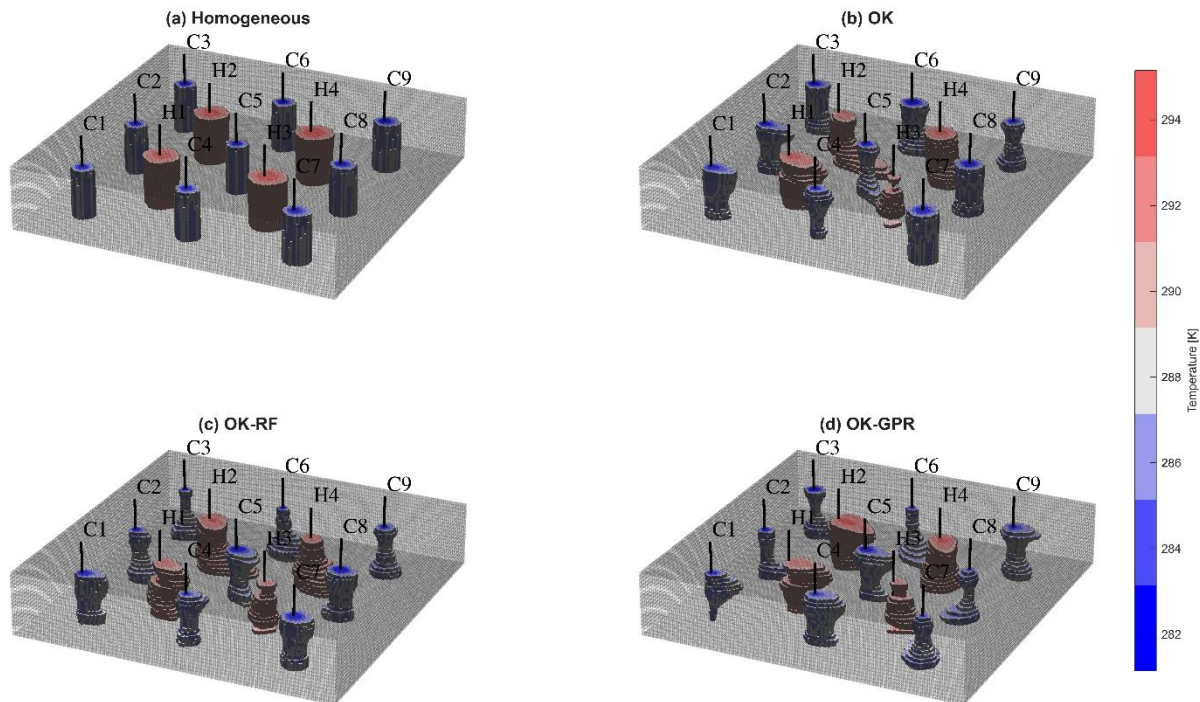


Figure 5.8 Comparison of the simulated thermal plumes at the end of cold-water injection period (Month 10) of the fifth operating cycle for the four hydraulic conductivity models: (a) Homogeneous, (b) OK, (c) OK-RF, and (d) OK-GPR.

Comparison of the four hydraulic conductivity representations clearly demonstrates that geological heterogeneity exerts an increasingly important control on long-term ATES performance. Whereas the homogeneous model predicts simplified and highly symmetric thermal behaviour, the heterogeneous fields generate realistic preferential groundwater circulation that significantly alters the migration and distribution of stored thermal energy. These differences become progressively larger as seasonal operating cycles accumulate.

5.7 Energy Recovery

5.7.1 Heat Recovery

Thermal performance of ATES system was first evaluated using the heat recovery factor, defined as the ratio between the thermal energy recovered during the production phase and the thermal energy injected during the preceding storage period. Larger recovery factors therefore indicate more efficient utilization of the stored thermal energy.

Figure 5.9 presents the evolution of heat recovery factor during the five-year simulation period for the four hydraulic conductivity models considered in this study

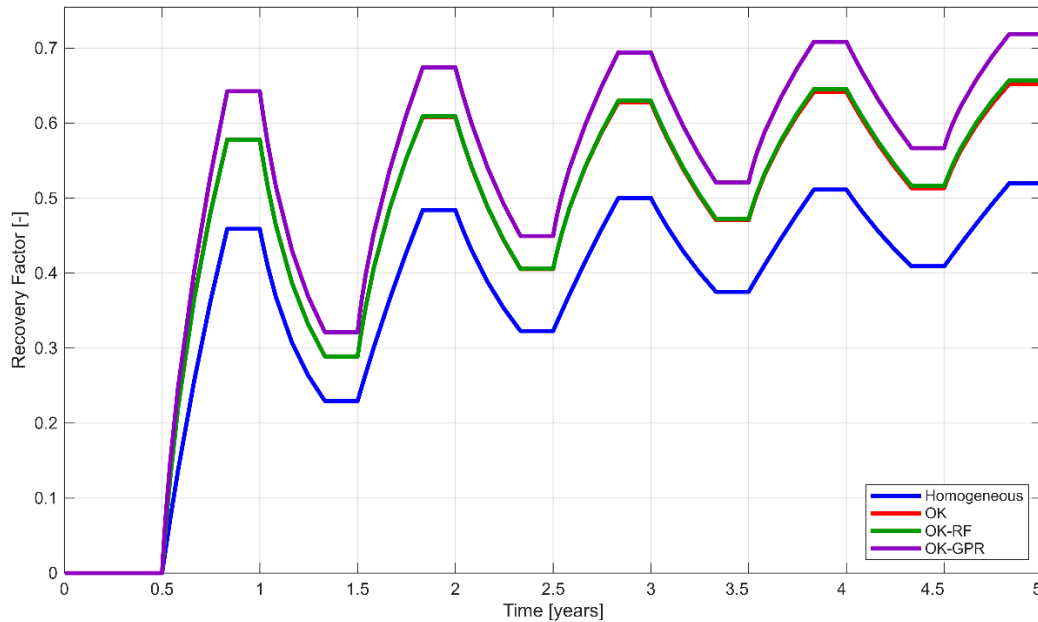


Figure 5.9 Evolution of the heat recovery factor over the five-year simulation period for the homogeneous, OK, OK–RF, and OK–GPR hydraulic conductivity models.

All four simulations exhibit the same cyclic behaviour imposed by the prescribed seasonal operating schedule. The following discussion therefore focuses on the differences among hydraulic conductivity representations.

Relative to the homogeneous reference simulation, all heterogeneous hydraulic conductivity fields produce higher simulated heat recovery throughout the five-year operational period. Deterministic OK simulation consistently predicts greater heat recovery than the homogeneous model, indicating that the predicted thermal response is sensitive to the spatial distribution of hydraulic conductivity.

Two hybrid hydraulic conductivity representations further modify the simulated heat recovery. OK–RF simulation produces heat recovery values that remain close to those predicted by the deterministic OK field throughout the simulation period. In contrast, OK–GPR simulation consistently predicts the highest heat recovery during all operational cycles. Although the differences among the heterogeneous representations remain relatively small, they become progressively more apparent as seasonal operation continues.

Overall, the results indicate that the predicted heat recovery depends on the adopted hydraulic conductivity representation. Relative to the homogeneous reference

simulation, all heterogeneous hydraulic conductivity fields produce higher simulated heat recovery, with OK–GPR representation yielding the largest heat recovery under the investigated operating conditions.

5.7.2 Cold Recovery

The seasonal performance of the ATES system was also evaluated using cold recovery factor, defined as the ratio between the thermal energy recovered during cold-water production phase and the thermal energy injected during the preceding cold-water storage period.

Figure 5.10 presents the evolution of cold recovery factor during the five-year simulation period for the four hydraulic conductivity representations.

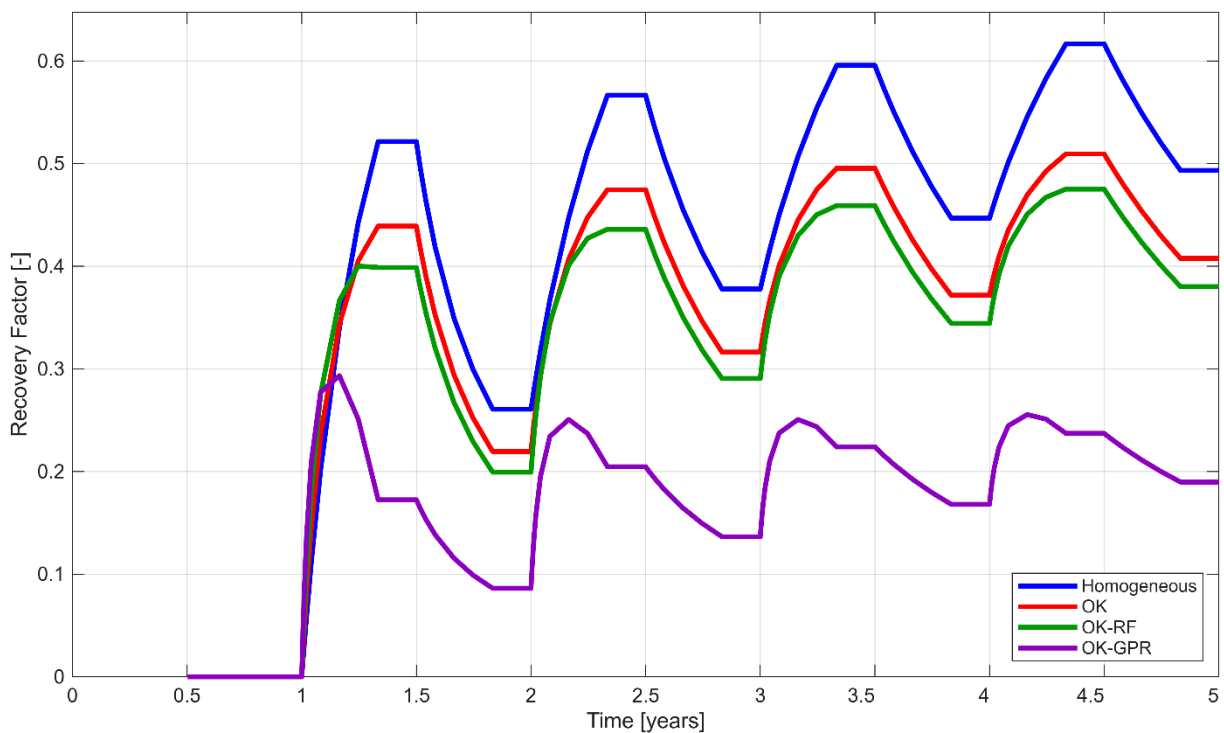


Figure 5.10 Evolution of cold recovery factor over the five-year simulation period for the homogeneous, OK, OK–RF, and OK–GPR hydraulic conductivity models.

Similar to heat recovery results, all four simulations exhibit same cyclic behaviour associated with the prescribed seasonal operating schedule. The discussion therefore focuses on the differences among hydraulic conductivity representations.

In contrast to heat recovery results, the homogeneous reference simulation consistently predicts the highest cold recovery throughout the five-year simulation period, whereas all heterogeneous hydraulic conductivity fields produce lower

recovery values. Among the heterogeneous representations, deterministic OK field generally predicts the highest cold recovery, followed by OK–RF simulation, while OK–GPR simulation consistently predicts the lowest cold recovery.

As discussed in Section 5.6, the heterogeneous hydraulic conductivity fields generate different thermal plume geometries relative to the homogeneous reference simulation. These differences in the simulated thermal response are accompanied by corresponding differences in the predicted cold recovery during long-term operation.

Overall, the results demonstrate that different hydraulic conductivity representations influence the predicted heat and cold recovery differently under identical operating conditions. While OK–GPR simulation predicts the highest heat recovery, it simultaneously predicts the lowest cold recovery among the investigated hydraulic conductivity representations. These findings indicate that the predicted seasonal energy recovery of the numerical ATES model is sensitive to the adopted hydraulic conductivity representation and that changes observed in one performance indicator are not necessarily accompanied by similar changes in other indicators.

5.8 Well Temperature Evolution

Evolution of production-well temperatures provides insight into the local thermal response of the numerical ATES model during seasonal operation. Unlike heat and cold recovery factors, which describe the overall thermal performance of the system, well temperatures reflect the thermal conditions at individual production wells. Because all operating conditions, well configuration, and numerical parameters remained identical throughout the simulations, differences in well temperatures are interpreted relative to the homogeneous reference simulation and reflect the influence of the adopted hydraulic conductivity representation.

5.8.1 Hot Well Temperature Evolution

Temporal evolution of production temperatures for the four hot wells (H1–H4) under the four hydraulic conductivity representations is shown in Figure 5.11.

All simulations exhibit the same cyclic temperature behaviour imposed by the prescribed seasonal operating schedule. The discussion therefore focuses on the differences among the hydraulic conductivity representations.

Relative to the homogeneous reference simulation, heterogeneous hydraulic conductivity fields generally predict lower production temperatures during the extraction periods. Among the heterogeneous representations, the OK–GPR simulation frequently predicts the lowest production temperatures, particularly for

wells H1 and H3, whereas the deterministic OK simulation generally produces intermediate values. OK–RF simulation typically lies between the OK and OK–GPR results. Differences among the hydraulic conductivity representations become more noticeable during the later operating cycles as seasonal heat injection and extraction continue.

Thermal response also varies among individual wells. Wells H1 and H3 exhibit larger temperature reductions than H2 and H4, indicating that the predicted thermal response depends on the local hydraulic conductivity distribution surrounding each well rather than on the overall operating conditions.

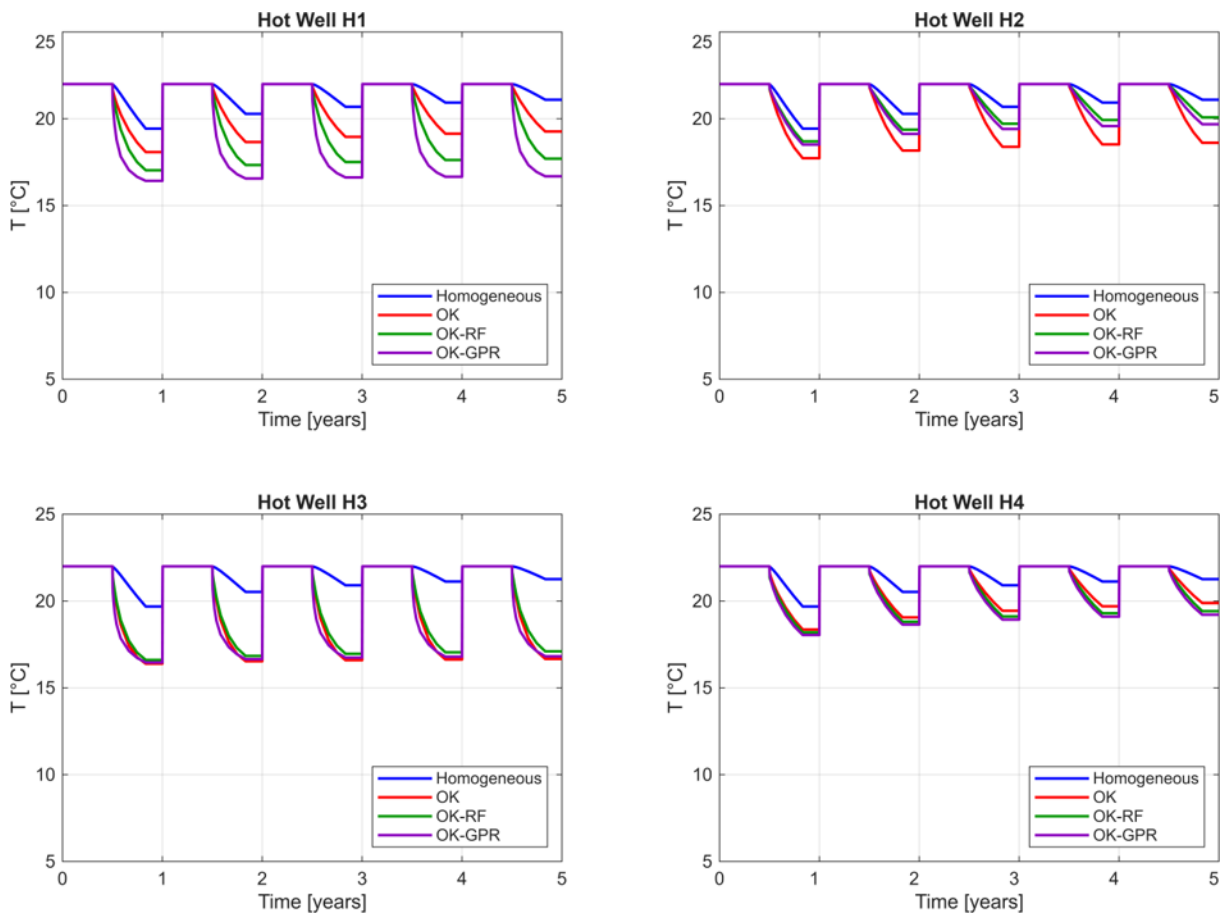


Figure 5.11 Temperature evolution of the four hot production wells for the homogeneous, OK, OK–RF, and OK–GPR hydraulic conductivity fields.

5.8.2 Cold Well Temperature Evolution

Figure 5.12 presents the production temperatures of nine cold wells (C1–C9) for four hydraulic conductivity representations during the five-year simulation period.

Compared with hot wells, cold wells exhibit larger variability among both the individual wells and the hydraulic conductivity representations. Relative to the homogeneous reference simulation, the heterogeneous hydraulic conductivity fields generally predict greater temperature changes during production, although the magnitude of these differences varies considerably from one well to another.

Influence of hydraulic conductivity representation is not uniform across all wells. For several wells, including C7 and C8, OK-GPR simulation predicts higher production temperatures than the other hydraulic conductivity representations, whereas wells such as C2 and C5 exhibit only minor differences among the heterogeneous simulations. These results indicate that the predicted thermal response depends strongly on the local hydraulic conductivity distribution surrounding each well.

Cold wells exhibit greater variability than hot wells, indicating that their predicted thermal response is more sensitive to the adopted hydraulic conductivity representation. When considered together with the thermal plume evolution and energy recovery results, these observations demonstrate that different hydraulic conductivity representations produce different local thermal responses within the numerical ATES model under identical operating conditions.

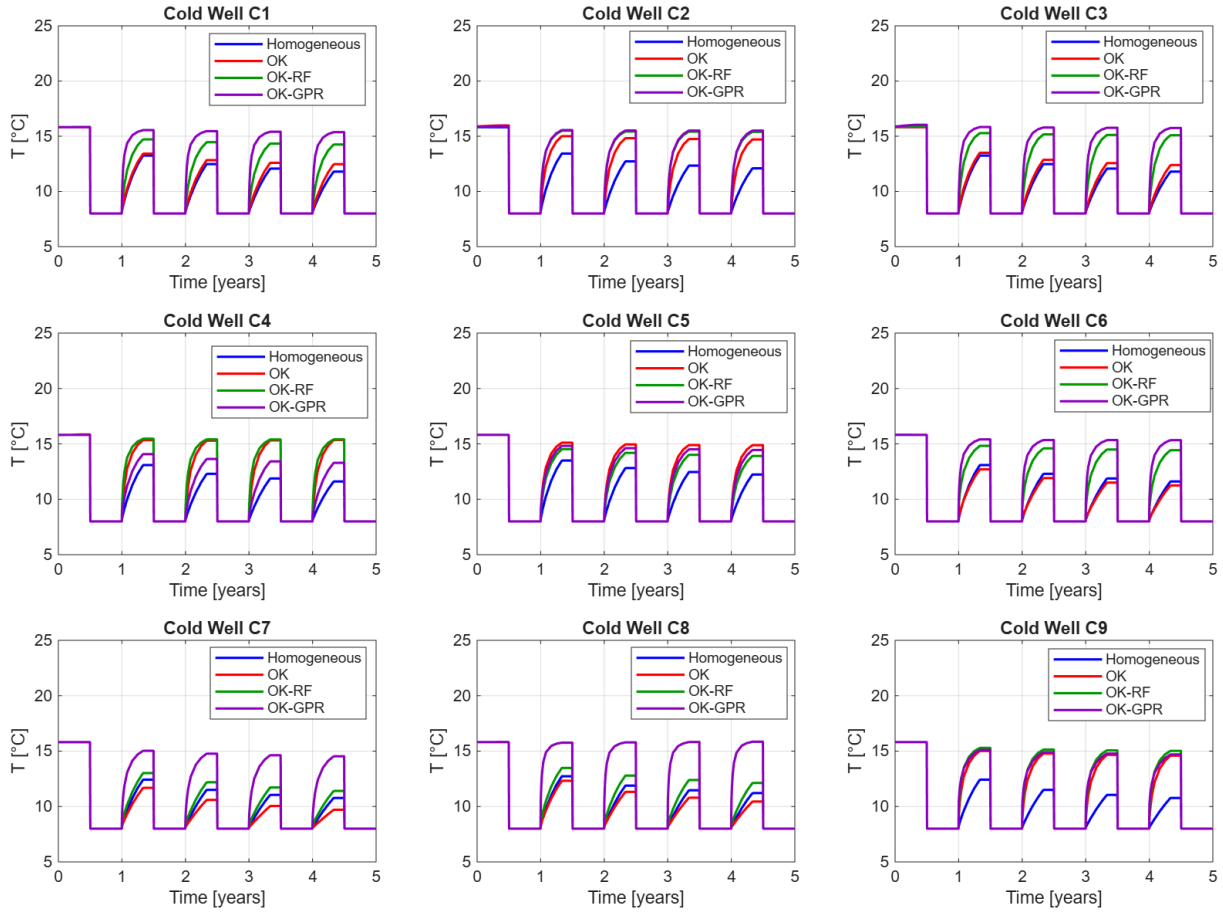


Figure 5.12 Temperature evolution of the nine cold production wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields over the five-year simulation period.

5.9 Well Hydraulic Performance

Hydraulic response of the numerical ATEs model was evaluated by comparing the evolution of well flow rates for the four hydraulic conductivity representations. Because identical operating schedules, well controls, and numerical settings were adopted in all simulations, differences in well flow rates are interpreted relative to the homogeneous reference simulation and reflect the influence of the hydraulic conductivity field. Compared with the well-temperature responses presented in the previous section, the well flow rates exhibit substantially larger variations among the investigated hydraulic conductivity representations, indicating that the predicted hydraulic response is more sensitive to the adopted hydraulic conductivity field than the thermal response.

5.9.1 Hot Well Flow Rate Evolution

Figure 5.13 presents the production and injection flow rates of four hot wells during the five-year simulation period.

All simulations follow the prescribed seasonal operating schedule, and the injection rates remain similar because identical operating controls were applied throughout the simulations. Consequently, the discussion focuses on the differences observed during the production periods.

Relative to the homogeneous reference simulation, the heterogeneous hydraulic conductivity fields generally predict larger variations in production flow rates. The differences are most pronounced for wells H1 and H3, whereas wells H2 and H4 exhibit comparatively similar responses among the four hydraulic conductivity representations. Among the heterogeneous simulations, OK–GPR field generally predicts the largest deviations in production flow rates, while the deterministic OK and OK–RF simulations produce intermediate responses.

Variability among the individual wells indicates that the predicted hydraulic response depends strongly on the local hydraulic conductivity distribution surrounding each well. Overall, hot-well results demonstrate that the simulated production flow rates are sensitive to the adopted hydraulic conductivity representation, with larger differences than those observed for the corresponding well temperatures.

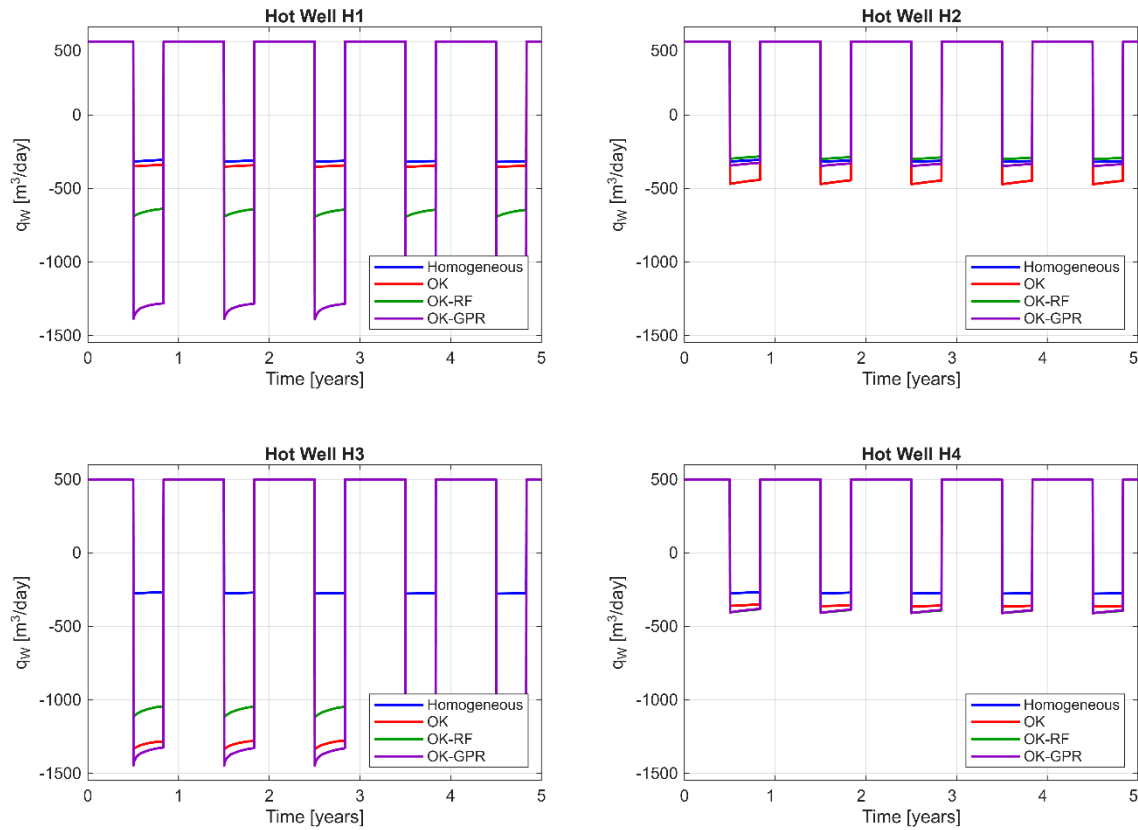


Figure 5.13 Flow-rate evolution of the four hot wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields during the five-year simulation period.

5.9.1 Cold Well Flow Rate Evolution

Compared with hot wells, cold wells exhibit substantially larger variability among the four hydraulic conductivity representations. Although all simulations follow the same operating schedule, noticeable differences are observed in both production and injection flow rates, indicating that the predicted hydraulic response varies considerably from one well to another.

Production and injection flow rates of the nine cold wells over the five-year simulation period are shown in Figure 5.14.

Relative to the homogeneous reference simulation, the heterogeneous hydraulic conductivity fields produce different flow-rate responses depending on the local hydraulic conductivity distribution. For several wells, including C1, C3, C4, C6, and C8, OK-GPR simulation predicts larger production and injection flow rates than the other hydraulic conductivity representations, whereas wells C5 and C7 exhibit only minor differences among the heterogeneous simulations. These results indicate that

the influence of the hydraulic conductivity representation is strongly location dependent.

An important observation is that the ranking of the heterogeneous hydraulic conductivity representations is not consistent across all wells. A hydraulic conductivity representation producing the largest flow rate for one well does not necessarily produce the largest flow rate for another. This behaviour indicates that the predicted hydraulic response is primarily controlled by the local hydraulic conductivity distribution surrounding each well rather than by the average hydraulic conductivity of the entire model domain.

Cold wells exhibit greater variability than hot wells, indicating that their predicted hydraulic response is more sensitive to the adopted hydraulic conductivity representation. Together with the thermal plume, energy recovery, and temperature analyses presented in the previous sections, these results demonstrate that different hydraulic conductivity representations produce different hydraulic responses under identical operating conditions.

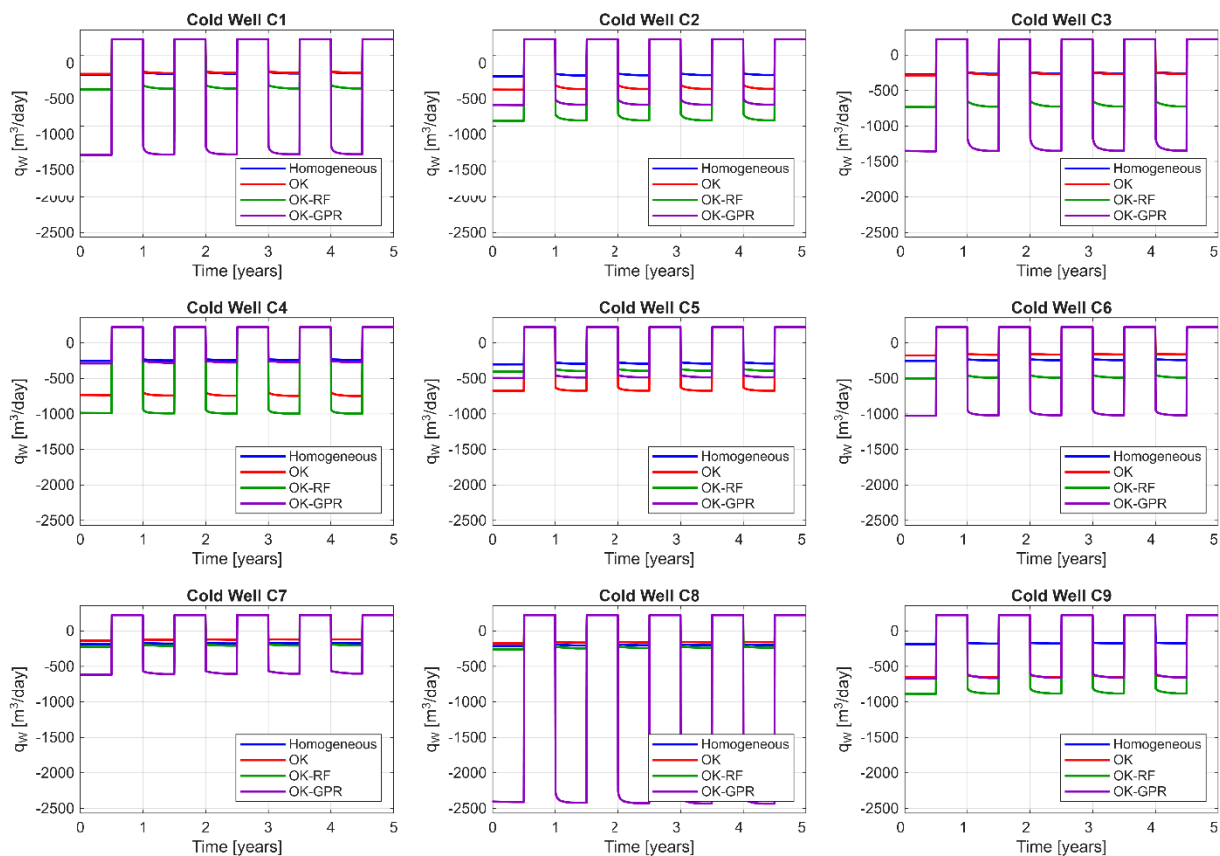


Figure 5.14 Flow-rate evolution of the nine cold wells for the homogeneous, OK, OK–RF, and OK–GPR hydraulic conductivity fields during the five-year simulation period.

5.10 Bottom-Hole Pressure Evolution

BHP was analyzed to evaluate the hydraulic response of the numerical ATES model during seasonal operation. Unlike well flow rates, which quantify groundwater production and injection, BHP represents the pressure required to sustain the prescribed operating conditions. Because all simulations were performed using identical well controls, operating schedules, and numerical settings, differences in BHP are interpreted relative to the homogeneous reference simulation and reflect the influence of the adopted hydraulic conductivity representation.

5.10.1 Hot Well Bottom-Hole Pressure

Figure 5.15 presents the bottom-hole pressure evolution of four hot wells throughout the five-year simulation period.

All simulations exhibit the same cyclic behaviour imposed by the seasonal operating schedule. Consequently, the discussion focuses on the differences among the hydraulic conductivity representations.

Relative to the homogeneous reference simulation, all heterogeneous hydraulic conductivity fields generally predict lower bottom-hole pressures throughout the simulation period. Among the heterogeneous representations, OK–GPR simulation frequently predicts the lowest pressures, followed by OK–RF and deterministic OK simulations. The largest differences are observed for wells H1 and H3, whereas wells H2 and H4 exhibit comparatively similar responses among the heterogeneous simulations.

Compared with the well flow-rate results presented in Section 5.9, BHP responses display a more consistent ranking among the four hydraulic conductivity representations. Overall, the results indicate that the predicted operating pressures are sensitive to the adopted hydraulic conductivity representation, with the largest differences occurring for wells H1 and H3.

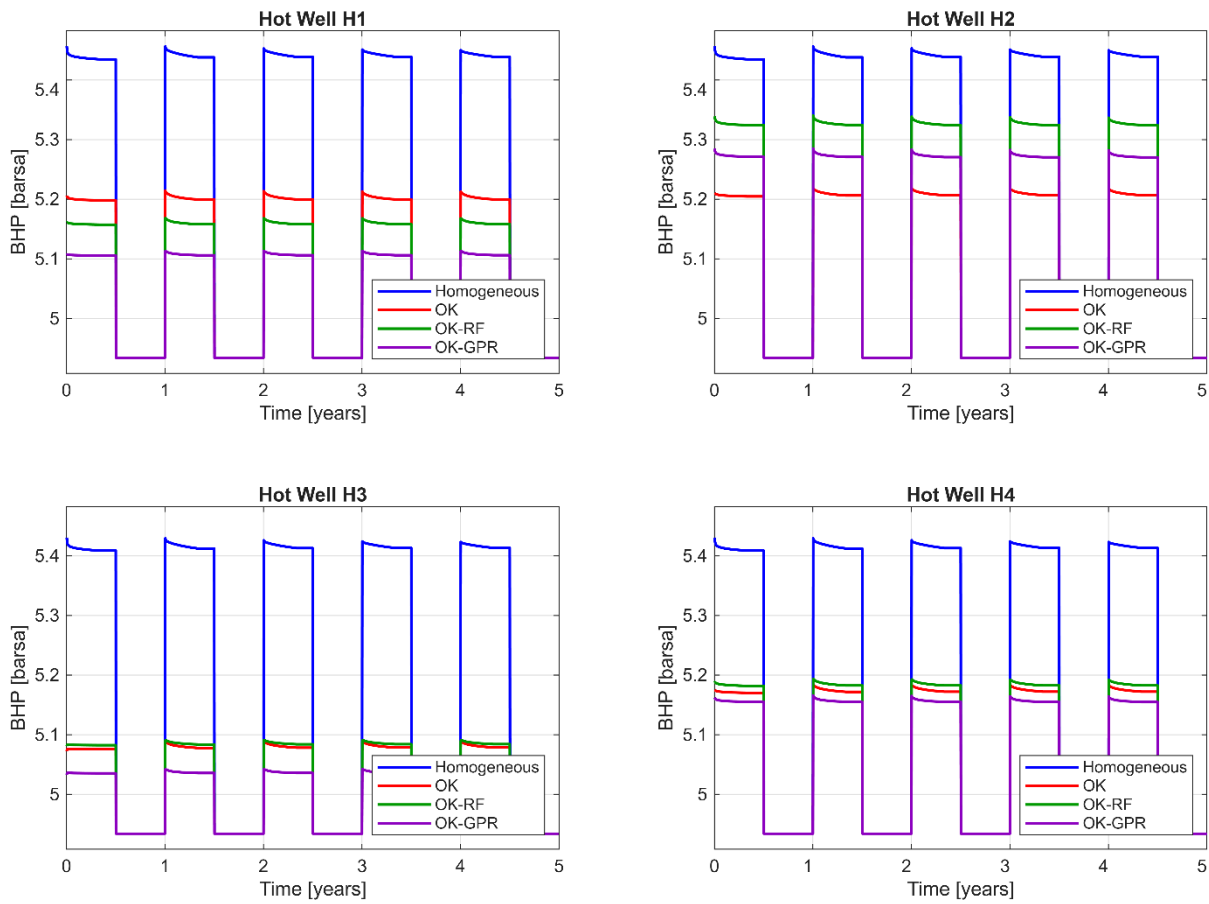


Figure 5.15 Bottom-hole pressure evolution of the four hot wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields during the five-year simulation period.

5.10.2 Cold Well Bottom-Hole Pressure

Figure 5.16 presents the bottom-hole pressure evolution of the nine cold wells during the five-year simulation period.

Compared with hot wells, cold wells exhibit greater variability among the four hydraulic conductivity representations. Relative to the homogeneous reference simulation, the heterogeneous hydraulic conductivity fields generally predict lower bottom-hole pressures, although the magnitude of the differences varies considerably from one well to another.

Unlike hot wells, no consistent ranking is observed among the heterogeneous hydraulic conductivity representations. For several wells, including C1, C2, C3, C6, C7, and C8, OK-GPR simulation predicts the lowest operating pressures, whereas for wells C4 and C5 the OK-RF and OK-GPR simulations produce similar responses.

Well C9 exhibits only minor differences among the heterogeneous hydraulic conductivity representations.

The variation observed among the individual wells indicates that the predicted pressure response depends strongly on the local hydraulic conductivity distribution surrounding each well. Consequently, the relative influence of the hydraulic conductivity representation differs from one well to another.

Relative to the homogeneous reference simulation, the heterogeneous hydraulic conductivity fields generally predict lower operating pressures, although the magnitude of the differences depends on the individual well and the adopted hydraulic conductivity representation.

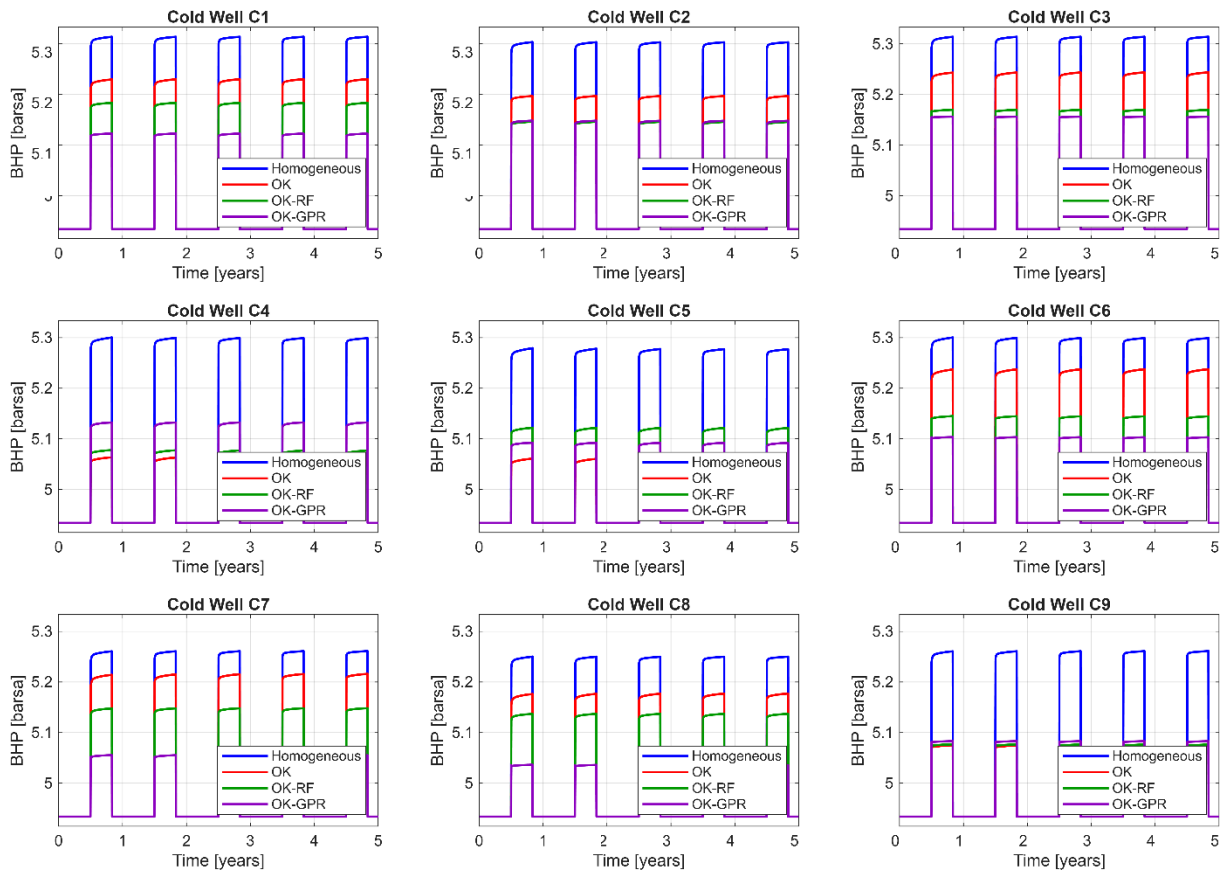


Figure 5.16 Bottom-hole pressure evolution of the nine cold wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields during the five-year simulation period.

5.11 Comparative Performance of the Hydraulic Conductivity Models

Previous sections demonstrated that the predicted thermo-hydraulic response of the numerical ATES model is sensitive to the adopted hydraulic conductivity representation. Since all simulations were performed using identical well configurations, operating schedules, boundary conditions, thermophysical properties, and numerical settings, the observed differences are solely attributable to the hydraulic conductivity field. Throughout this study, the homogeneous hydraulic conductivity field serves as the reference simulation, and the remaining results are interpreted relative to this reference case.

From a thermal perspective, all four simulations reproduced the same cyclic behaviour imposed by the prescribed seasonal operating schedule. Nevertheless, differences were observed in the magnitude of the predicted thermal response. Relative to the homogeneous reference simulation, the heterogeneous hydraulic conductivity fields produced different thermal plume geometries, different well-temperature evolutions, and different seasonal energy recovery. The heterogeneous simulations generally predicted higher heat recovery than the homogeneous simulation, whereas the homogeneous simulation produced the highest cold recovery under the investigated operating conditions.

The largest differences among the hydraulic conductivity representations were observed in the hydraulic response of the wells. Compared with the homogeneous reference simulation, the heterogeneous hydraulic conductivity fields generally predicted larger variations in well flow rates together with lower BHP. These differences were not spatially uniform and varied considerably from one well to another, indicating that the predicted hydraulic response depends strongly on the local hydraulic conductivity distribution surrounding each well.

Differences among the three heterogeneous hydraulic conductivity representations were generally smaller than the differences between the heterogeneous and homogeneous simulations. Although OK, OK-RF, and OK-GPR simulations produced different thermo-hydraulic responses, no single hydraulic conductivity representation consistently produced the largest or smallest values for every performance indicator. For example, OK-GPR simulation predicted the highest heat recovery but the lowest cold recovery, while the relative ranking of the heterogeneous representations varied among the individual wells with respect to well temperatures, flow rates, and bottom-hole pressures.

Overall, the results demonstrate that different hydraulic conductivity representations lead to different predicted thermo-hydraulic responses of the numerical ATES model under identical operating conditions. Relative to the homogeneous reference simulation, the heterogeneous hydraulic conductivity fields modify the predicted groundwater flow, thermal plume evolution, seasonal energy recovery, and well behaviour. These findings highlight the sensitivity of numerical ATES simulations to the adopted hydraulic conductivity representation and emphasize the importance of carefully evaluating alternative subsurface characterization approaches when interpreting simulation results.

6. CONCLUSIONS AND FUTURE WORK

6.1 Conclusion

This thesis investigated the influence of different hydraulic conductivity representations on the predicted thermo-hydraulic response of an ATEs system by integrating geostatistical interpolation, machine-learning-based residual modelling, and three-dimensional numerical simulation. A hybrid framework was developed in which OK was combined with RF and GPR to generate alternative hydraulic conductivity representations for numerical groundwater flow and heat transport simulations. Within the adopted Maximum Likelihood framework, OK predictions were treated as estimators of the unknown hydraulic conductivity, and cross-validation residuals were interpreted as estimation errors that were subsequently modelled using machine-learning techniques. The resulting hydraulic conductivity fields were incorporated into a three-dimensional numerical model developed using MRST.

Four hydraulic conductivity representations were investigated: a homogeneous reference field, a deterministic OK field, and two hybrid fields generated using OK–RF and OK–GPR. Throughout all numerical simulations, identical well configurations, operating schedules, pumping conditions, thermophysical properties, boundary conditions, and numerical settings were maintained. Consequently, differences observed among the simulation results are solely attributable to the adopted hydraulic conductivity representation.

Numerical simulations demonstrated that the predicted thermo-hydraulic response of ATEs system is sensitive to the adopted hydraulic conductivity representation. Relative to the homogeneous reference simulation, the heterogeneous hydraulic conductivity fields produced different thermal plume geometries, different well-temperature evolutions, different well flow rates, different BHP, and different

seasonal energy recovery. These results indicate that the spatial representation of hydraulic conductivity influences the predicted groundwater flow and heat transport behaviour within the numerical ATES model.

Among the investigated hydraulic conductivity representations, the OK, OK–RF, and OK–GPR simulations produced different thermo-hydraulic responses relative to the homogeneous reference simulation. OK–GPR simulation generally predicted the highest heat recovery together with the largest variations in well flow rates and the lowest BHP, whereas the homogeneous reference simulation predicted the highest cold recovery. These results demonstrate that different hydraulic conductivity representations modify different performance indicators in different ways under identical operating conditions. Therefore, the numerical simulations do not identify one hydraulic conductivity representation as universally superior, but instead quantify how alternative subsurface representations influence the predicted response of ATES model.

The simulations also demonstrated that the predicted hydraulic response is more sensitive to the adopted hydraulic conductivity representation than the predicted thermal response. Compared with well temperatures, substantially larger differences were observed in well flow rates and BHP, indicating that groundwater flow reacts more strongly to local variations in hydraulic conductivity than the corresponding thermal behaviour. In addition, the magnitude of the simulated response varied considerably among individual wells, demonstrating that local hydraulic conductivity distributions play an important role in controlling the predicted well behaviour.

Overall, this study demonstrates that combining geostatistical interpolation, machine-learning-based residual modelling, and numerical groundwater flow and heat transport simulation provides a systematic framework for investigating influence of hydraulic conductivity characterization on the predicted thermo-hydraulic response of ATES systems. The proposed methodology offers a practical approach for comparing alternative hydraulic conductivity representations within a consistent numerical modelling framework and contributes to a better understanding of how subsurface characterization influences simulated groundwater flow, heat transport, and seasonal thermal energy storage.

6.2 Study limitations

Although the proposed hybrid geostatistical–machine learning framework provides a systematic approach for generating alternative hydraulic conductivity

representations and evaluating their influence on the predicted thermo-hydraulic response of an ATES system, several limitations should be acknowledged.

First, the hydraulic conductivity fields investigated in this study are deterministic representations of subsurface heterogeneity. Although the hybrid OK–RF and OK–GPR approaches modify the hydraulic conductivity field through machine-learning-based residual modelling, they do not explicitly quantify the uncertainty associated with spatial interpolation. Consequently, the uncertainty arising from limited borehole observations is not propagated through the groundwater flow and heat transport simulations. The numerical results therefore represent deterministic simulation outcomes for the investigated hydraulic conductivity representations.

Second, the study was conducted for a single geological setting corresponding to Aquifer B beneath Parco Sempione, Milan. Although this case study provides an appropriate framework for comparing different hydraulic conductivity representations, the numerical results should not be generalized directly to aquifers with different geological structures, hydraulic properties, or operational conditions.

Third, all numerical simulations were performed using identical well locations, operating schedules, pumping conditions, injection temperatures, boundary conditions, and thermophysical properties. This modelling strategy was intentionally adopted to isolate the influence of the hydraulic conductivity representation on the predicted ATES response. Consequently, the conclusions of this study are limited to the operating conditions considered and do not account for alternative well configurations or operational strategies.

In addition, the machine-learning residual models were developed using spatial coordinates as predictor variables. Although this approach provided a consistent framework for generating alternative hydraulic conductivity representations, other sources of geological information, such as lithological classifications, sedimentary facies, geophysical surveys, or hydraulic testing data, were not incorporated into the modelling framework. The influence of these additional data sources therefore remains beyond the scope of the present study.

Finally, the numerical simulations were evaluated through comparisons among alternative hydraulic conductivity representations rather than through validation against long-term observations from an operating ATES system. Consequently, the study evaluates how different hydraulic conductivity representations influence the predicted thermo-hydraulic response of the numerical model, rather than determining which representation most accurately reproduces field behaviour.

Despite these limitations, the adopted framework provides a consistent basis for investigating the influence of hydraulic conductivity characterization on numerical ATES simulations and establishes a foundation for future studies that incorporate uncertainty analysis, additional geological information, and field-scale validation.

6.3 Future Work

Methodology developed in this study establishes a framework for investigating the influence of different hydraulic conductivity representations on the predicted thermo-hydraulic response of ATES systems. Several opportunities exist to further extend the proposed framework.

The most important extension is the incorporation of uncertainty analysis through conditional stochastic simulation. The present study employed deterministic hydraulic conductivity fields generated using OK and machine-learning-based residual modelling. Future research could generate multiple conditional realizations of hydraulic conductivity using Monte Carlo simulation while preserving the available borehole observations and spatial covariance structure. Such an approach would allow uncertainty in hydraulic conductivity to be propagated through the groundwater flow and heat transport simulations, providing probabilistic estimates of the predicted ATES response rather than a single deterministic solution.

Another promising direction is the application of the proposed framework to different geological settings and aquifer systems. Extending the methodology to sites with different geological characteristics, hydraulic properties, and hydrogeological conditions would provide further insight into the influence of hydraulic conductivity representation under a wider range of subsurface environments.

Future studies may also investigate alternative operational conditions for the numerical ATES model. The present work intentionally maintained identical well configurations, operating schedules, pumping rates, injection temperatures, and boundary conditions in order to isolate the influence of hydraulic conductivity representation. Additional investigations could evaluate different well layouts, seasonal operating schedules, storage capacities, pumping strategies, and injection temperatures to examine how these factors interact with different hydraulic conductivity representations.

The proposed framework may also be extended through the incorporation of additional geological and hydrogeological information into the hydraulic conductivity characterization process. Lithological classifications, sedimentary facies, geophysical surveys, hydraulic test data, and groundwater-level observations could

be integrated with the current methodology to generate alternative hydraulic conductivity representations based on a broader range of subsurface information.

Although RF and GPR were adopted in the present study, future work could evaluate additional machine-learning techniques for residual estimation, including Gradient Boosting, XGBoost, LightGBM, or deep-learning approaches. Ensemble-learning strategies combining multiple algorithms may also be investigated to evaluate their influence on hydraulic conductivity characterization.

Finally, validation of the proposed methodology using long-term observations from operating ATES systems would provide an important opportunity to compare numerical predictions with field measurements. Such studies would contribute to a better understanding of the applicability of different hydraulic conductivity representations under real operating conditions.

Overall, the methodology developed in this thesis provides a flexible framework for integrating geostatistics, machine learning, and numerical simulation. Extending this framework to include uncertainty quantification, additional geological information, alternative machine-learning algorithms, and field-scale validation would provide a more comprehensive assessment of how hydraulic conductivity characterization influences the predicted thermo-hydraulic response of ATES systems.

Bibliography

- [1] "Global Energy Review 2025," 2025. [Online]. Available: www.iea.org
- [2] IPCC, *Global Warming of 1.5°C*. Cambridge University Press, 2022. doi: 10.1017/9781009157940.
- [3] P. Fleuchaus, B. Godschalk, I. Stober, and P. Blum, "Worldwide application of aquifer thermal energy storage – A review," Oct. 01, 2018, *Elsevier Ltd*. doi: 10.1016/j.rser.2018.06.057.
- [4] P. Goovaerts, "Geostatistics for natural resources evaluation," 1997.
- [5] G. Matheron, "Principles of geostatistics. Economic Geology," 1963.
- [6] A. G. . Journel and C. . Huijbregts, *Mining geostatistics*. Academic Press, 1989.
- [7] T. Hastie, R. Tibshirani, and J. Friedman, "The elements of statistical learning: Data mining, inference, and prediction (2nd ed.)," 2009. [Online]. Available: <http://www.springer.com/series/692>
- [8] C. Edward. Rasmussen and C. K. I. . Williams, *Gaussian processes for machine learning*. MIT Press, 2006.
- [9] M. L. Buursink *et al.*, "fs20223082.pdf - Geologic Energy Storage," 2022.
- [10] E. Witter *et al.*, "A review of Geological Thermal Energy Storage for seasonal, grid-scale dispatching," Aug. 01, 2025, *Elsevier Ltd*. doi: 10.1016/j.rser.2025.115761.
- [11] F. Chen *et al.*, "Technical and Economic Feasibility Analysis of Underground Hydrogen Storage: A Case Study in Intermountain-West Region USA," 2022.
- [12] W. Ji, J. Wan, J. Li, S. Chen, H. Ma, and H. Yu, "Integration of large-scale underground energy storage technologies and renewable energy sources," Nov. 01, 2024, *Yandy Scientific Press*. doi: 10.46690/ager.2024.11.01.
- [13] de Marsily G., "Quantitative Hydrogeology: Groundwater Hydrology for Engineers," 1968.
- [14] C. V. Deutsch, *Geostatistical Reservoir Modeling*. 2002.
- [15] J. Bear, "dynamics of fluids in porous media," 1972.
- [16] S. M. Benson and D. R. Cole, "CO₂ sequestration in deep sedimentary formations," *Elements*, vol. 4, no. 5, pp. 325–331, 2008, doi: 10.2113/gselements.4.5.325.

- [17] David. Banks, *An introduction to thermogeology : ground source heating and cooling*. Blackwell Pub., 2008.
- [18] N. Heinemann *et al.*, “Enabling large-scale hydrogen storage in porous media-the scientific challenges,” Feb. 01, 2021, *Royal Society of Chemistry*. doi: 10.1039/d0ee03536j.
- [19] W. W. Woessner, A. C. Stringer, and E. P. Poeter, “An introduction to hydraulic testing in hydrogeology: Basic pumping, slug, and packer methods,” 2023.
- [20] Q. Liu *et al.*, “A Numerical Study of Slug Tests in a Three-Dimensional Heterogeneous Porous Aquifer Considering Well Inertial Effects,” *Water Resour. Res.*, vol. 56, no. 11, Nov. 2020, doi: 10.1029/2020WR027155.
- [21] B. Lora-Ariza, L. Silva Vargas, and L. D. Donado, “Estimation of Hydraulic Conductivity from Well Logs for the Parameterization of Heterogeneous Multilayer Aquifer Systems,” *Water (Switzerland)*, vol. 17, no. 16, Aug. 2025, doi: 10.3390/w17162439.
- [22] C. F. Panagiotou, P. Kyriakidis, and E. Tziritis, “Application of geostatistical methods to groundwater salinization problems: A review,” Dec. 01, 2022, *Elsevier B.V.* doi: 10.1016/j.jhydrol.2022.128566.
- [23] F. Heße, S. Müller, and S. Attinger, “Data-driven estimates for the geostatistical characterization of subsurface hydraulic properties,” *Hydrol. Earth Syst. Sci.*, vol. 28, no. 2, pp. 357–374, Jan. 2024, doi: 10.5194/hess-28-357-2024.
- [24] J. Lino Pereira, E. A. Varouchakis, G. P. Karatzas, and L. Azevedo, “Uncertainty Quantification in Geostatistical Modelling of Saltwater Intrusion at a Coastal Aquifer System,” *Math. Geosci.*, vol. 56, no. 5, pp. 867–885, Jul. 2024, doi: 10.1007/s11004-023-10120-7.
- [25] B. Sheil *et al.*, “Artificial intelligence transformations in geotechnics: progress, challenges and future enablers,” Jan. 01, 2026, *Elsevier Ltd.* doi: 10.1016/j.compgeo.2025.107604.
- [26] W. A. Agyare, S. J. Park, and P. L. G. Vlek, “Artificial Neural Network Estimation of Saturated Hydraulic Conductivity,” *Vadose Zone Journal*, vol. 6, no. 2, pp. 423–431, May 2007, doi: 10.2136/vzj2006.0131.
- [27] M. Farasati, M. Seyedian, and A. Fathaabadi, “Predicting soil hydraulic conductivity using random forest, SVM, and LSSVM models,” *Nat. Resour. Model.*, vol. 37, no. 4, Nov. 2024, doi: 10.1111/nrm.12407.
- [28] M. Kalumba, E. Nyirenda, I. Nyambe, S. Dondeyne, and J. Van Orshoven, “Machine Learning Techniques for Estimating Hydraulic Properties of the Topsoil across the Zambezi River Basin,” *Land (Basel)*, vol. 11, no. 4, Apr. 2022, doi: 10.3390/land11040591.
- [29] M. Rezaei *et al.*, “Incorporating machine learning models and remote sensing to assess the spatial distribution of saturated hydraulic conductivity in a light-textured

- soil,” *Comput. Electron. Agric.*, vol. 209, Jun. 2023, doi: 10.1016/j.compag.2023.107821.
- [30] A. Devi, A. Chandel, and V. Shankar, “Modelling hydraulic conductivity of porous media using machine learning techniques,” *Journal of Hydroinformatics*, vol. 27, no. 3, pp. 381–405, Mar. 2025, doi: 10.2166/hydro.2025.161.
- [31] T. Tilahun and J. Korus, “3D hydrostratigraphic and hydraulic conductivity modelling using supervised machine learning,” *Applied Computing and Geosciences*, vol. 19, Sep. 2023, doi: 10.1016/j.acags.2023.100122.
- [32] M. Elmorsy, W. El-Dakhakhni, and B. Zhao, “Enhancing Effective Thermal Conductivity Predictions in Digital Porous Media Using Transfer Learning,” *InterPore Journal*, vol. 2, no. 3, pp. IPJ250825-7, Aug. 2025, doi: 10.69631/ipj.v2i3nr75.
- [33] I. O. A. Odeh, A. B. Mcbratney, and D. J. Chittleborough, “Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging,” ELSEVIER, 1995.
- [34] M. Mohammadpour, H. Roshan, M. Arashpour, and H. Masoumi, “Machine learning assisted Kriging to capture spatial variability in petrophysical property modelling,” *Mar. Pet. Geol.*, vol. 167, Sep. 2024, doi: 10.1016/j.marpetgeo.2024.106967.
- [35] F. Fouedjio and H. Talebi, “Geostatistical semi-supervised learning for spatial prediction,” *Artificial Intelligence in Geosciences*, vol. 3, pp. 162–178, Dec. 2022, doi: 10.1016/j.aiig.2022.12.002.
- [36] J. Yao, W. Liu, Q. Liu, Y. Liu, X. Chen, and M. Pan, “Optimized algorithm for multipoint geostatistical facies modeling based on a deep feedforward neural network,” *PLoS One*, vol. 16, no. 6 June, Jun. 2021, doi: 10.1371/journal.pone.0253174.
- [37] S. Levy, L. Friedli, G. Mariéthoz, and N. Linde, “Conditioning of multiple-point statistics simulations to indirect geophysical data,” May 01, 2024, *Elsevier Ltd.* doi: 10.1016/j.cageo.2024.105581.
- [38] K.-A. Lie, “Knut-Andreas Lie An Introduction to Reservoir Simulation Using MATLAB User Guide for the Matlab Reservoir Simulation Toolbox (MRST),” 2016.
- [39] K. Vafai, “Handbook of porous media (3rd ed.),” 2015.
- [40] K.-A. Lie and O. Møyner, Eds., *Advanced Modeling with the MATLAB Reservoir Simulation Toolbox*. Cambridge University Press, 2021. doi: 10.1017/9781009019781.
- [41] D. W. Peaceman, “Interpretation of Well-Block Pressures in Numerical Reservoir Simulation With Nonsquare Grid Blocks and Anisotropic Permeability,” 1963.
- [42] R. Webster and M. A. Oliver, “Geostatistics for Environmental Scientists Second Edition,” 2007.
- [43] C. V Deutsch and A. G. Journel, “GSLIB: Geostatistical Software Library and User’s Guide Second Edition,” Oxford University Press, 1997.

- [44] J.-Paul. Chilès and Pierre. Delfiner, *Geostatistics : Modeling Spatial Uncertainty*. John Wiley & Sons, 2012.
- [45] F. Bachoc, "Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecification," *Comput. Stat. Data Anal.*, vol. 66, pp. 55–69, 2013, doi: 10.1016/j.csda.2013.03.016.
- [46] V. Yildirim, E. Todini, and Y. M. Kantar, "A robust maximum likelihood estimation approach for Ordinary Kriging with outlier-contaminated spatial data," *Environ. Ecol. Stat.*, Mar. 2026, doi: 10.1007/s10651-026-00706-9.
- [47] R. Peli, A. Menafoglio, M. Cervino, L. Dovera, and P. Secchi, "Physics-based Residual Kriging for dynamically evolving functional random fields," *Stochastic Environmental Research and Risk Assessment*, vol. 36, no. 10, pp. 3063–3080, Oct. 2022, doi: 10.1007/s00477-022-02180-8.
- [48] N. I. Loonat, A. I. J. M. Van Dijk, M. F. Hutchinson, and A. H. Weerts, "Anomaly Kriging Helps to Remove Bias in Spatial Model Runoff Estimates," *Water Resour. Res.*, vol. 56, no. 7, Jul. 2020, doi: 10.1029/2019WR026240.
- [49] T. Hengl, M. Nussbaum, M. N. Wright, G. B. M. Heuvelink, and B. Gräler, "Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables," *PeerJ*, vol. 2018, no. 8, 2018, doi: 10.7717/peerj.5518.
- [50] L. Breiman, "Random Forests," 2001.
- [51] H. V. Gupta, H. Kling, K. K. Yilmaz, and G. F. Martinez, "Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling," *J. Hydrol. (Amst.)*, vol. 377, no. 1–2, pp. 80–91, Oct. 2009, doi: 10.1016/j.jhydrol.2009.08.003.
- [52] G. James, D. Witten, T. Hastie, and R. Tibshirani, "An Introduction to Statistical Learning with Applications in R Second Edition," 2021.
- [53] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," 2005. [Online]. Available: www.int-res.com
- [54] A. Previati, V. Silvestri, and G. Crosta, "Deep learning text classification of borehole logs for regional scale modeling of hydrofacies (Po Plain, N Italy)," *J. Hydrol. Reg. Stud.*, vol. 58, Apr. 2025, doi: 10.1016/j.ejrh.2024.102157.
- [55] M. Bloemendal, T. Olsthoorn, and F. Boons, "How to achieve optimal and sustainable use of the subsurface for Aquifer Thermal Energy Storage," *Energy Policy*, vol. 66, pp. 104–114, Mar. 2014, doi: 10.1016/j.enpol.2013.11.034.
- [56] W. Sommer, J. Valstar, I. Leusbrock, T. Grotenhuis, and H. Rijnaarts, "Optimization and spatial pattern of large-scale aquifer thermal energy storage," *Appl. Energy*, vol. 137, pp. 322–337, Jan. 2015, doi: 10.1016/j.apenergy.2014.10.019.

List of Figures

Figure 2.1 Geological energy storage technologies and their associated geological settings (adapted from Buursink et al.) [9].	9
Figure 4.1 Study area and borehole locations.	62
Figure 4.2 Hydrogeological setting of Aquifer B.	64
Figure 4.3 Three-dimensional distribution of measured logK values.	65
Figure 4.4 Spatial distribution of training and testing boreholes.	66
Figure 4.5 Experimental and fitted exponential variograms in vertical and horizontal directions derived from training boreholes.	68
Figure 4.6 Comparison between observed and OK predicted logK values for independent testing boreholes.	70
Figure 4.7 Comparison between true and predicted Ordinary Kriging residual distributions for training dataset.	75
Figure 4.8 Comparison between true and predicted Ordinary Kriging residual distributions for testing dataset.	76
Figure 4.9 Comparison between observed hydraulic conductivity values and predictions obtained using Ordinary Kriging, OK-RF, and OK-GP for testing boreholes.	77
Figure 4.10 Comparison of RMSE, R^2 , and KGE obtained using Ordinary Kriging (OK), OK-RF, and OK-GP on independent testing dataset.	77
Figure 5.1 Annual operating schedule adopted for simulated ATEs system.	82
Figure 5.2 Spatial arrangement of warm and cold wells used in numerical ATEs simulations.	84

Figure 5.3 Homogeneous logarithmic hydraulic conductivity field adopted as the reference case for numerical ATEs simulations.....	90
Figure 5.4 Heterogeneous logarithmic hydraulic conductivity fields used in numerical ATEs simulations: (a) Ordinary Kriging (OK), (b) OK-RF, and (c) OK-GPR.....	92
Figure 5.5 Comparison of the simulated thermal plumes at the end of hot-water injection period (Month 4) of the first operating cycle for the four hydraulic conductivity models: (a) Homogeneous, (b) OK, (c) OK-RF, and (d) OK-GPR.....	94
Figure 5.6 Comparison of the simulated thermal plumes at the end of cold-water injection period (Month 10) of the first operating cycle for the four hydraulic conductivity models: (a) Homogeneous, (b) OK, (c) OK-RF, and (d) OK-GPR.....	95
Figure 5.7 Comparison of the simulated thermal plumes at the end of hot-water injection period (Month 4) of the fifth operating cycle for the four hydraulic conductivity models: (a) Homogeneous, (b) OK, (c) OK-RF, and (d) OK-GPR.....	96
Figure 5.8 Comparison of the simulated thermal plumes at the end of cold-water injection period (Month 10) of the fifth operating cycle for the four hydraulic conductivity models: (a) Homogeneous, (b) OK, (c) OK-RF, and (d) OK-GPR.....	97
Figure 5.9 Evolution of the heat recovery factor over the five-year simulation period for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity models. ...	98
Figure 5.10 Evolution of cold recovery factor over the five-year simulation period for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity models.	99
Figure 5.11 Temperature evolution of the four hot production wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields.....	101
Figure 5.12 Temperature evolution of the nine cold production wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields over the five-year simulation period.....	103
Figure 5.13 Flow-rate evolution of the four hot wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields during the five-year simulation period.	105
Figure 5.14 Flow-rate evolution of the nine cold wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields during the five-year simulation period.	106
Figure 5.15 Bottom-hole pressure evolution of the four hot wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields during the five-year simulation period.....	108

Figure 5.16 Bottom-hole pressure evolution of the nine cold wells for the homogeneous, OK, OK-RF, and OK-GPR hydraulic conductivity fields during the five-year simulation period.....	109
--	-----

List of Tables

Table 4.1 Main parameters of the modeled Aquifer B.	64
Table 4.2 Calibrated parameters of anisotropic exponential variogram model used for OK.....	69
Table 4.3 Performance metrics of OK evaluated using testing boreholes.....	71
Table 4.4 Quantitative comparison of Ordinary Kriging and hybrid models using independent testing dataset.....	78
Table 5.1 ATEs system design parameters accounting for seepage velocity of the groundwater.....	86
Table 5.2 Operating parameters used in numerical simulations.....	88
Table 5.3 Computational domain and structured grid characteristics used in MRST simulations.	89

2. List of Symbols

Variable	Description	SI unit
B	Formation Volume Factor	–
g	Gravitational Acceleration	m/s ²
c	Well Completion Factor	–
c_f	Fluid Specific Heat Capacity	J/(kg·K)
$c_{p,i}$	Specific Heat Capacity of Phase i	J/(kg·K)
D	Dataset	–
$e(x_0)$	Estimation Error	–
F_f	Fluid Mass Flux	kg/(m ² ·s)
h	Reservoir Thickness	m
h_f	Fluid Specific Enthalpy	J/kg
H_f	Fluid Conductive Heat Flux	W/m ²
H_r	Rock Conductive Heat Flux	W/m ²
I	Identity Matrix	–
k	Intrinsic Permeability	m ²
k_x	Permeability In X-Direction	m ²

Variable	Description	SI unit
k_y	Permeability In Y-Direction	m ²
k_e	Effective Permeability	m ²
k_f	Thermal Conductivity of Fluid	W/(m·K)
k_r	Thermal Conductivity of Rock	W/(m·K)
K	Hydraulic Conductivity	m/s
C	Covariance Matrix in GPR (Recommended)	–
c_*	Covariance Vector Between the Prediction Location and The Training Observations	–
$C_\delta(x, x')$	Covariance Function of Discrepancy Process	–
$C_\theta(x, x')$	Covariance Function of The Hydraulic Conductivity Field	–
h	Lag Distance Used in Variogram Analysis	m
l_H	Horizontal Variogram Length Scale	m
l_V	Vertical Variogram Length Scale	m
l_k	GP Kernel Length Scale (Recommended)	–
M_f	Fluid Mass	kg
M_r	Rock Mass	kg
$m(x)$	GP Mean Function	–
n	Outward Unit Normal Vector	–
N	Number Of Validation Observations	–
$N(h)$	Number Of Observation Pairs	–
p	Fluid Pressure	Pa
p_a	Prescribed Boundary Pressure	Pa
p_R	Reservoir Pressure	Pa

Variable	Description	SI unit
p_{bh}	Bottom-Hole Pressure	Pa
q	Source/Sink Term	m ³ /s
Q_f	Fluid Source/Sink Term	kg/(m ³ ·s)
Q_T	Thermal Source/Sink Term	W/m ³
N_T	Number Of Trees in The Random Forest Ensemble	–
r_e	Equivalent Drainage Radius	m
r_w	Well Radius	m
S	Skin Factor	–
T	Temperature	K or °C
u_f	Fluid Internal Energy	J/kg
u_r	Rock Internal Energy	J/kg
v	Darcy Velocity	m/s
$v(x)$	Feature Vector Associated with Spatial Location x	–
n	Number Of Observations	–
x	Spatial Coordinate	m
x_0	Estimation Location	m
x_i	Observation Location i	m
x_j	Observation Location j	–
x_*	Prediction Location	m
X	Predictor Matrix	–
$Z(x)$	Log-Transformed Hydraulic Conductivity Field (Regionalized Variable)	
$\hat{Z}_{OK}(x)$	Ordinary Kriging Estimate at Location x	

Variable	Description	SI unit
$\hat{\mathbf{Z}}_{corr}(\mathbf{x})$	The Corrected Prediction at Location $\mathbf{x}$	–
$\hat{\mathbf{Z}}_{OK}^{(-i)}(\mathbf{x}_i)$	LOOCV Ordinary Kriging Prediction at Location $\mathbf{x}_i$	–
$\mathbf{r}(\mathbf{x}_i)$	LOOCV Residual	–
z	Elevation Coordinate	m
$\mathbf{z}_{bh}$	Bottom-Hole Elevation	m
$\mathbf{z}_{scaled}$	Scaled Vertical Coordinate	m
N_0	Nugget Effect (Semivariance at Zero Separation Distance)	
$\mathbf{Range}_H$	Horizontal Practical Range of Spatial Correlation	m
$\mathbf{Range}_V$	Vertical Practical Range of Spatial Correlation	m
$T_t(\cdot)$	Prediction Of The T-Th Decision Tree	–
$f_t(\cdot)$	Decision-Tree Model	–
$\hat{\mathbf{Z}}_{corr}^{GP}(\mathbf{x})$	GP-Corrected Prediction	–
$\hat{\mathbf{Z}}_{corr}^{RF}(\mathbf{x})$	RF-Corrected Prediction	–
$\bar{\mathbf{Z}}_{obs}$	Mean Of the Observed Values	–
$\mathbf{Z}_{pred}(\mathbf{x}_i)$	Predicted Value at Validation Location $\mathbf{x}_i$	–
r	Pearson Correlation Coefficient	–
σ_{obs}	Standard Deviation of Observed Values	–
σ_{pred}	Standard Deviation of Predicted Values	–
μ_{obs}	Mean Of Observed Values	–
μ_{pred}	Mean Of Predicted Values	–
α	KGE Variability Ratio	–
β	KGE Bias Ratio	–

Variable	Description	SI unit
β_f	Fluid Compressibility	Pa ⁻¹
β_r	Rock Compressibility	Pa ⁻¹
β_t	Total Compressibility	Pa ⁻¹
$\delta(x)$	Discrepancy Field	–
$\widehat{\delta}_{GP}(x)$	GP Discrepancy Estimate	–
$\widehat{\delta}_{RF}(x)$	RF Discrepancy Estimate	–
t	Time	s
t^n	Time at Step n	s
Δt^n	Time Step Size	s
Δx	Grid Spacing In X-Direction	m
Δy	Grid Spacing In Y-Direction	m
$\varepsilon(x)$	Random Spatial Component	–
ε_i	Irreducible Random Error	–
$\gamma(h)$	Semivariogram	–
$\widehat{\gamma}(h)$	Experimental Semivariogram	–
Γ_a	Dirichlet Boundary Portion	m ²
λ_i	Kriging Weight	–
λ_e	Effective Thermal Conductivity	W/(m·K)
μ	Dynamic Viscosity	Pa·s
μ_K	Lagrange Multiplier (Recommended)	–
μ_δ	Discrepancy Mean Function	–
μ_{OK}	Unknown Locally Constant Mean in Ordinary Kriging	–

Variable	Description	SI unit
ρ	Fluid Density	kg/m ³
ρ_f	Fluid Density (Heat Transport)	kg/m ³
ϕ	Porosity	–
σ^2	Partial Sill	–
σ_f^2	GP Signal Variance	–
σ_n^2	Observation Noise Variance	–
σ_{OK}^2	Ordinary Kriging Variance	–
σ_{GP}^2	GP Predictive Variance	–
σ_δ^2	Discrepancy Variance	–
σ_{corr}^2	Corrected Prediction Variance	–
θ	Covariance Parameter Vector	–
$\hat{\theta}$	Estimated Covariance Parameters	–
Ω	Control Volume	m ³
$\partial\Omega$	Boundary Of Domain	m ²
$(\rho c)_e$	Effective Volumetric Heat Capacity of the Saturated Porous Medium	J/(m ³ ·K)
$Z_{obs}(x_i)$	Observed Value of The Log-Transformed Hydraulic Conductivity at Location x_i	

3. Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor Ehsan Ranaee, for his guidance, support, and valuable feedback throughout this research. His expertise, patience, and constructive suggestions were instrumental in shaping the direction of this study and improving the quality of the final work.

I would also like to express my sincere appreciation to Professor Alberto Guadagnini for his valuable advice, constructive feedback, and continuous support throughout this research.

I am also grateful to the faculty members and researchers of the master's programme in Energy Engineering for providing a stimulating academic environment and for sharing their knowledge throughout my studies.

My sincere appreciation goes to the research team at the University of Milano-Bicocca for providing the hydrogeological dataset used in this work. Their efforts in data collection, processing, and hydrogeological characterization made this research possible.

I would also like to thank my colleagues and friends for their support, encouragement, and insightful discussions throughout the development of this thesis. Their contributions were particularly valuable during the implementation of the geostatistical analyses, machine-learning models, and numerical simulations.

Finally, I owe my deepest gratitude to my family for their unconditional love, patience, and continuous encouragement. Their support has been a constant source of motivation throughout my academic journey and has made the completion of this work possible.

