Doctoral School of Geosciences

DATA-DRIVEN ASSESSMENT AND OPTIMIZATION OF HIGH-TEMPERATURE AQUIFER THERMAL ENERGY STORAGE IN DEPLETED CLASTIC HYDROCARBON RESERVOIRS

PhD Dissertation

Author

Hawkar Ali Abdulhaq

Supervisor

János Szanyi, PhD

Department of Geology, University of Szeged, Szeged, Hungary

Consultant

István Vass, PhD

MOL Plc

Department of Geology

University of Szeged

Faculty of Science and Informatics

Szeged, 2026

TABLE OF CONTENTS

CHAPTER I..... 1
INTRODUCTION..... 1
I.1. BACKGROUND..... 1
I.2. PROBLEM STATEMENT..... 2
I.3. OBJECTIVES..... 2
I.4. STRUCTURE OF THESIS..... 3
CHAPTER II..... 5
II.1. Introduction..... 6
II.2. Glossary of Terms..... 9
II.3. Materials and Methods..... 11
II.3.1. Methodological Framework..... 11
II.3.2. Data Collection..... 14
II.3.3. Data Modeling..... 15
II.a.Modeling of Stratigraphy..... 16
II.b. Modeling Porosity and Permeability..... 18
II.c. Lithology Model..... 20
II.d. Heat Capacity..... 25
II.e. Thermal Conductivity..... 26
II.f. Criteria Selection..... 27
II.g. Case Study Selection..... 28
II.3.4. MCDA Application Process..... 29
II.3.5. Sensitivity Analysis..... 30
II.4. Results..... 31
II.4.1. HT-ATES..... 31
II.a. Probability Distribution of the lowest risk P90..... 33
II.b.Probability Distribution of the Medium risk P50..... 33
II.c. Probability Distribution of High Values P90..... 33
II.4.2. High-Temperature Direct Application..... 34
II.5. Discussion..... 36
II.6. Conclusion..... 38
CHAPTER III..... 39
III.1. Introduction..... 40

III.2. Glossary of Terms.....44
III.3. Background and Regional.....46
III.3.1 Geological and Hydrogeological Setting of the Békés Basin.....46
III.3.2 Hydrodynamic Systems and Pressure Regimes.....49
III.3.3 Reservoir Properties and Geothermal Potential.....50
III.3.4 Hydrocarbon History and Well Infrastructure.....51
III.3.5 Relevance to Seasonal Heat Storage.....51
III.4. Materials and Methods.....52
III.4.1. Methodological Framework.....52
III.4.2. Data Collection and Data Preparation.....53
III.4.3. Hydrogeological Modelling.....55
III.4.4 Heat Transport Modelling.....56
III.4.6 Training Data for Machine Learning.....57
III.4.7 Model Calibration and Validation.....60
III.4.8 Sensitivity Analysis.....60
III.5. Results.....60
III.5.1 Heat Simulation result.....61
III.5.2. Machine Learning Result:.....62
III.6. Discussion.....64
III.6.1 Alignment with Previous Studies and Theoretical Outcomes.....64
III.6.2 Key Influencing Parameters.....65
III.6.3 Strengths and Limitations of Hydrogeological Model Calibration.....65
III.6.4 Enhancing Decision-Making for UTES Site Selection.....65
III.6.5 Implications for Scaling Geothermal Storage Projects.....66
III.6.6 Assumptions and Simplifications.....66
III.6.7 Uncertainties in Input Data.....66
III.6.8 Recommendations and Future Work.....66
III.7. Conclusion.....67
CHAPTER IV.....69
IV.1 . Introduction.....70
IV.2. Glossary of Terms.....74
IV.3. Background and Regional.....76
IV.3.1 Geological and Hydrogeological Setting.....76
IV.3.2 Hydrodynamic Systems and Pressure Regimes.....78

IV.3.3 Reservoir Properties and Geothermal Potential..... 79
IV.3.4 Hydrocarbon History and Well Infrastructure.....80
IV.3.5 Relevance to Seasonal Heat Storage.....80
IV.4. Materials and Methods.....81
IV.4.1. Methodological Framework.....81
IV.4.1. Data Collection and Data Preparation.....83
IV.4.3. Data Modelling.....85
IV.4.5 Machine Learning Setting.....87
IV.4.6 Machine Learning Process.....89
IV.4.7 Sensitivity Analysis.....91
IV.4.8 Model Calibration and Validation.....92
IV.5. Results.....93
IV.5.1 FZI Prediction result..... 93
IV.5.2. Residual Distribution and Model Robustness.....95
IV.6. Discussion..... 96
IV.6.1 Interpretation of Predicted 3D FZI Clusters.....97
IV.6.2 Subsurface Complexity of High-FZI Regions in 3D..... 98
IV.6.3 Enhancing Decision-Making for UTES Site Selection.....99
IV.6.5 Computational Efficiency in Sand Channel Prediction.....100
IV.6.6 Managing Model Uncertainty and Reproducibility..... 101
IV.6.7 Recommendations and Future Work.....102
IV.7. Conclusion.....103
CHAPTER V.....105
CONCLUSIONS..... 105
V.1. RECOMMENDATIONS.....107
SUMMARY.....109
INTRODUCTION AND AIMS.....109
APPLIED METHODS.....110
NEW SCIENTIFIC RESULTS..... 112
ACKNOWLEDGEMENTS.....119
References.....120

LIST OF THE FIGURES

II.Figure 1. Data Processing Methodological Framework.....12

II.Figure 2. Criteria Scoring Distribution for Each Parameter and Their Weights (Visualized by Python:(H. Abdulhaq, 2024)). Quantitative criteria and thresholds are shown in Abdulhaq, 2025f from Tab Thresholds. The threshold of each criterion is supported by different references, which are provided for each parameter (Bridger & Allen, 2010; Shi et al., 2023; Stricker et al., 2020) 14

II.Figure 3. Modeled Values Prepared for Spatial Tiling Process.....15

II.Figure 4. The top of the Bekes Formation obtained using Universal Kriging.....16

II.Figure 5. The results of the thickness simulation. A= E-type estimation; B= coefficient of variation (uncertainty of the simulation model); C= P10; D=P50; E=P90.....17

II.Figure 6. The base of the Bekes Formation is obtained by subtracting the thickness grid from the top grid.....18

II.Figure 7. The results of the porosity simulation. A= E-type estimation; B= coefficient of variation (uncertainty of the simulation model); C= P10; D=P50; E=P90.....19

II.Figure 8. The results of the permeability simulation (all scales are mD). A= E-type estimation; B= coefficient of variation (uncertainty of the simulation model); C= P90; D=P50; E=P10.....20

II.Figure 9. The electro-lithological types are defined using EM clustering. “A”: the final result of the EM clustering; “B”: the defined electro-lithologies with the corresponding smallest and largest gamma values.....21

II.Figure 10. The 3D Gamma model (A) and the corresponding 3D lithology model (B).....22

II.Figure 11. Map of the dominant electro-lithologies.....23

II.Figure 12. shows the applied linear regression describing the change in temperature with increasing depth. This regression was applied to both the top and bottom grid to obtain the temperature distribution at the top and bottom of the Bekes Formation (II.Figure 13, A and B).....24

II.Figure 13. Maps of the temperature distributions [°C] at the top (A) and base (B) of the Bekes Formation.....24

II.Figure 14. Heat Capacity in MJ/kg·K..... 25

II.Figure 15. Thermal Conductivity Distribution Maps for P10, P50, P90, and Mean Scenarios with Well Locations. These maps illustrate the distribution of thermal conductivity for different scenarios: P10 (10th percentile), P50 (50th percentile or median), P90 (90th percentile), and the mean. The color scale ranges from purple (low thermal conductivity) to yellow (high thermal conductivity). Red dots represent well locations used in the study.....27

II.Figure 16. Overview of the Study Area in Hungary: Bottom Left - Lithological Section (Tóth & Almási, 2001); Bottom Right - Position of Bekes Formations in the Study Grid with Vertical Exaggeration (13.33X); Above - MCDA Grids with Well Locations Indicated by Black Dots.....29

II.Figure 17. Results for HT-ATES Suitability. As illustrated in this figure, the areas marked in color indicate potential zones for energy storage as they meet the necessary criteria. The white areas represent regions that do not meet the required thresholds for suitability. The figure clearly delineates suitable zones for High Temperature-Aquifer Thermal Energy Storage (HT-ATES), aiding in the identification of optimal locations for energy storage initiatives.....32

II.Figure 18. Direct Heat Utilization Zones. This figure highlights zones suitable for direct heat utilization, with temperatures ranging between 100°C and 140°C. The marked areas indicate regions that meet the criteria for effective direct heat utilization, ensuring optimal use of geothermal energy for heating purposes..... 35

III.Figure 1. shows the location of the study area in Hungary, the modelled section within the Békés Formation lithological map, and a lithological cross-section; modified after Abdulhaq et al. (2024)..... 48

III.Figure 2. The workflow of the methodology employed in this study.....53

III.Figure 3. shows the location of the qualified wells over the temperature grid, highlighting the modelled area. White patches indicate areas where the reservoir thickness is less than 2 meters.....54

III.Figure 4. shows the distribution of the wells and the boundary of the urban town that can be considered the potential candidate for the district heating beneficiary. The Bekes F. Temperature near the town is around 70° C, which makes it an ideal location to be used as a UTES site. For this reason, we narrowed down the area of interest to be modelled..... 55

III.Figure 5 illustrates the application of the Random Forest algorithm using a radial relational diagram..... 59

III.Figure 6 presents simulation results for a specific time frame, showing the injection of hot

fluid through two wells..... 61

III.Figure 7 shows the thermal performance curve of the simulated hot well across repeated seasons..... 62

III.Figure 8 presents the thermal performance of wells located outside the modelled area.... 63

III.Figure 9 shows the percentage increase in thermal performance for each well based on the prediction results..... 64

IV.Figure 1 Thickness map of the Pannonian sediments in Hungary, highlighting the extent of the Szolnok Formation (outlined by purple lines). The green rectangle delineates the specific area investigated in this study (B 2204-B: Pannonian Basin Province, Central Europe (Province 4808)—Petroleum Geology, Total Petroleum Systems, and Petroleum Resource Assessment, n.d.; Geofizika, 2011; Nádor & Lenkey, 2023)..... 78

IV.Figure 2 Radial flow chart illustrating the workflow implemented in this research. The red arrows represent data outputs or processed information, while the blue arrows denote data inputs or feedback pathways..... 83

IV.Figure 3 Reservoir characterization plots illustrating key petrophysical properties from well-log data within the Szolnok Formation: (A) Permeability distribution (mD); (B) Effective porosity (%); (C) Volume of shale (VSHA; %, v/v), defined as the percentage of rock volume composed of shale; (D) Volume of sand (VSND; %, v/v); (E) Cross-sectional representation of permeability distribution (mD) along wells. The vertical axis shows depth (m, Total Vertical Depth), with wells sorted horizontally from thickest (left) to thinnest (right). Note: These plots represent petrophysical properties derived from well-log data, not physical core images..... 85

IV.Figure 4 Spatial and statistical representation of the Szolnok Formation thickness; the left panel displays a map illustrating reservoir thickness distribution across the study area, while the right panel shows the frequency distribution of thickness values..... 86

IV.Figure 5 Spatial distribution maps of modelled Flow Zone Index (FZI) values; the left map represents the predicted FZI values using Inverse Distance Weighted (IDW) interpolation, while the right map highlights regions with FZI greater than 4.5, indicative of potential sand-rich channel systems..... 87

IV.Figure 6 Feature importance based on the XGBoost regression model, ranked by gain, highlighting key petrophysical attributes and derived features that most significantly contributed to predicting sand-channel distributions within the Szolnok Formation..... 88

IV.Figure 7: Radial flow chart depicting the structured workflow of this research. Red nodes

represent input stages, such as loading and preprocessing data, while blue nodes indicate output stages, including model predictions, validation, and generation of final maps and quality control charts. Arrows illustrate the directional data flow and method dependencies throughout the analysis pipeline..... 91

IV.Table 1: Optimal hyperparameters and corresponding model performance metrics from the sensitivity analysis of the FZI prediction workflow..... 91

IV.Figure 8 A; Normalized maps illustrating spatial distributions of Flow Zone Index (FZI), B; sand content (VSND), and C; shale content (VSHA). D; Scatter plots showing correlations between normalized FZI and sand content (positive correlation, $r=0.333$, $p<0.001$) and E; shale content (negative correlation, $r=-0.384$, $p<0.001$), highlighting sandy channels..... 93

IV.Figure 9 Observed versus predicted Flow Zone Index (FZI) values from the 20% validation dataset, highlighting strong model performance . The dashed line indicates a perfect prediction (1:1 line), illustrating the model's accuracy in capturing subsurface heterogeneity..... 94

IV.Figure 10 Variance map showing the mean residuals (Predicted minus Actual FZI). The spatial distribution highlights minimal prediction errors, characterized by negligible bias.... 96

IV.Figure 11 Map of areas with predicted FZI values greater than 4.5, highlighting inferred sand channel systems. (A) Cumulative distribution of FZI values, (B) frequency histogram of FZI, and (C) spatial map displaying the mean distribution of FZI clusters..... 97

IV.Figure 12 (Left) Spatial distribution of mean FZI values greater than 4.5, with wells shown in red if they penetrate these predicted high-FZI zones. (Right) 3D visualization of predicted FZI values exceeding 4.5, highlighting the complex geometry of the inferred channel systems..... 99

LIST OF TABLES

II.Table 1 Assigned weights of evaluation criteria used in the sensitivity analysis.....30
III.Table 1. Key Parameters for the MODFLOW Model.....56
Table 2. Input parameters and model settings used in the MT3DMS heat transport simulation. 57
IV.Table 1: Optimal hyperparameters and corresponding model performance metrics from the sensitivity analysis of the FZI prediction workflow..... 91
IV.Table 2 clearly summarizes these computational differences:..... 101

CHAPTER I

INTRODUCTION

I.1. BACKGROUND

The transition toward low-carbon energy systems requires solutions that balance seasonal mismatches between heat supply and demand. In Central and Eastern Europe, underground thermal energy storage (UTES) can store surplus heat in the subsurface for later recovery, and high-temperature aquifer thermal energy storage (HT-ATES) is especially relevant for district-heating and industrial applications. Depleted hydrocarbon reservoirs are attractive HT-ATES targets because they offer proven porosity–permeability systems, dense well networks, and long production histories that reduce geological uncertainty. Repurposing such assets can complement or replace costly decommissioning and abandonment, allowing legacy infrastructure to contribute to decarbonization while limiting environmental and economic liabilities.

Hungary's Pannonian Basin provides an ideal testbed, where elevated geothermal gradients, well coverage, and substantial district-heating demand coincide with decades of subsurface data. Yet clastic-reservoir heterogeneity—channelized sands embedded in lower-permeability facies—complicates predictions of heat transport, thermal recovery efficiency, and long-term system performance.

This thesis synthesizes the geothermal and subsurface-energy contributions developed during my PhD alongside a broader publication record that includes contributions to 27 peer-reviewed papers spanning geothermal systems, reservoir characterization, geological mapping, and mineral-resource studies. Three HT-ATES-focused papers, developed with my supervisors and collaboration with Professor János Geiger, form the thesis backbone—covering MCDA–AHP screening, MODFLOW–MT3DMS heat-transport simulation with Random-Forest surrogates, and XGBoost-based 3D channel mapping for thermal-breakthrough prevention. Two complementary geothermal publications—a review on geothermal energy and critical metal extraction co-authored with Professor Ladislaus Rybach, and a national overview of harnessing geothermal energy in Hungary—provide wider context and actively strengthen stakeholder

engagement across academia, industry, and policy in Hungary and internationally. Together, these works establish a coherent progression from regional screening to reservoir-scale simulation and scalable, data-driven optimization for practical HT-ATES deployment.

I.2. PROBLEM STATEMENT

Although the potential of depleted hydrocarbon fields for UTES has been widely acknowledged, current assessment and design approaches remain fragmented and incomplete. Regional screening studies typically rely on simplified geological representations and deterministic parameter values, which are insufficient for formation-scale planning under uncertainty. Conversely, high-fidelity numerical simulations, while capable of capturing complex heat-transport processes, are computationally intensive and difficult to apply systematically across large numbers of candidate sites.

A critical limitation in many existing studies is the inadequate treatment of geological heterogeneity. In clastic reservoirs, sand-channel architecture exerts first-order control on fluid flow paths, thermal plume evolution, and breakthrough timing. Failure to explicitly resolve this architecture can lead to misleading performance predictions and suboptimal well placement. In addition, uncertainty associated with subsurface properties is often treated implicitly or ignored, reducing the robustness of feasibility assessments and decision-making.

Furthermore, there is a lack of integrated frameworks that connect site screening, numerical modeling, and data-driven prediction within a coherent decision-support system. As a result, the transition from regional potential assessment to optimized HT-ATES design remains poorly defined, limiting the deployment of subsurface thermal storage as a practical component of energy transition strategies.

I.3. OBJECTIVES

The overarching objective of this PhD research is to develop an integrated, uncertainty-aware, and data-driven framework for the assessment, design, and optimization of high-temperature aquifer thermal energy storage (HT-ATES) systems in depleted clastic hydrocarbon reservoirs.

This overarching aim is achieved through three interconnected objectives that collectively address site selection, performance evaluation, and scalable system optimization.

Objective 1: To develop a systematic and probabilistic framework for screening and ranking depleted hydrocarbon reservoirs according to their suitability for HT-ATES by integrating multi-criteria decision analysis with geostatistical simulation. This objective aims to enable efficient and transparent site selection by explicitly accounting for geological uncertainty and by supporting formation-scale prioritization and evaluation of repurposing potential.

Objective 2: To quantitatively assess the thermal performance of HT-ATES systems in heterogeneous clastic reservoirs using physically based numerical flow and heat-transport modeling. This objective focuses on evaluating the effects of reservoir architecture, anisotropy, well spacing, and operational strategies on heat transport, thermal breakthrough, and recovery efficiency, thereby supporting optimized system design.

Objective 3: To develop and apply scalable machine-learning and hybrid modeling approaches for rapid performance prediction and high-resolution geological characterization, and to integrate these tools into HT-ATES design workflows. This objective addresses the computational limitations of traditional numerical modeling by enabling efficient evaluation of large numbers of scenarios while retaining geological realism and physical interpretability.

I.4. STRUCTURE OF THESIS

This chapter summarises the thesis on the PhD research topic of data-driven assessment and optimization of high-temperature aquifer thermal energy storage (HT-ATES) in depleted clastic hydrocarbon reservoirs. After this brief introduction, each of the following three chapters (Chapters II–IV) represents a research paper that was accepted and published during the different stages of the research, followed by a conclusion chapter (Chapter V).

Chapter II: It presents the development and application of a probabilistic MCDA–AHP framework integrated with geostatistics to screen and rank depleted hydrocarbon fields for HT-ATES suitability. The chapter formalizes a criteria-based workflow that maps key reservoir,

thermal, and infrastructure attributes into spatial suitability indices to identify the most promising storage zones under uncertainty.

Chapter III: It presents a detailed, high-resolution numerical modelling study of groundwater flow and heat transport in heterogeneous clastic reservoirs, focusing on how geological variability and operational design choices control thermal recovery efficiency and breakthrough behaviour. The chapter also introduces a machine-learning surrogate approach to rapidly predict and generalize thermal performance across candidate wells based on simulation-generated training data.

Chapter IV: Based on the need for better representation of reservoir architecture, it presents a data-driven machine-learning and hybrid modelling workflow to delineate sand-channel systems and high-quality flow zones from well-log-derived attributes. The predicted 3D channel architecture is then linked to practical UTES/HT-ATES design implications, especially for improved well placement and thermal breakthrough prevention.

Chapter V: It presents a compiled and comprehensive summary of the main findings of the PhD research work, discusses implications for subsurface energy repurposing, and outlines future research directions. Finally, a summarized reference list is provided at the end of the thesis.

CHAPTER II

Transforming Abandoned Hydrocarbon Fields into Heat Storage Solutions: A Hungarian Case Study Using Enhanced MCDA-AHP and Geostatistical Methods

Hawkar Ali Abdulhaq
Department of Mineralogy, Geochemistry, and Petrology, University of Szeged
hawkar.ali.abdulhaq@szte.hu

János Geiger
Department of Mineralogy, Geochemistry, and Petrology, University of Szeged
matska@geo.u-szeged.hu

István Vass
MOL Hungary, MOL Plc
isvass@mol.hu

Tivadar M. Tóth
Department of Mineralogy, Geochemistry, and Petrology, University of Szeged
mtoth@geo.u-szeged.hu

Tamás Medgyes
Geothermal Energy Applied Research Department, University of Szeged
medgyes.tamas@szetav.hu

János Szanyi
Department of Geology and Paleontology, University of Szeged
szanyi@iif.u-szeged.hu

Energies
17, 2024
doi.org/10.3390/en17163954

Transforming Abandoned Hydrocarbon Fields into Heat Storage Solutions: A Hungarian Case Study Using Enhanced MCDA-AHP and Geostatistical Methods

Hawkar Ali Abdulhaq, János Geiger, István Vass, Tivadar M. Tóth, Tamás Medgyes and János Szanyi

Abstract

This study introduces a robust methodology utilizing Multi-Criteria Decision Analysis (MCDA) combined with the Analytic Hierarchy Process (AHP) to repurpose abandoned hydrocarbon fields for energy storage, supporting the transition to renewable energy sources. We use a geostatistical approach integrated with Python scripting to analyze reservoir parameters—including porosity, permeability, thickness, lithology, temperature, heat capacity, and thermal conductivity—from a decommissioned hydrocarbon field in Southeast Hungary. Our workflow leverages stochastic simulation data to identify potential zones for energy storage, categorizing them into high, moderate, and low suitability scenarios. This innovative approach provides rapid and precise analysis, enabling effective decision-making for energy storage implementation in depleted fields. The key finding is the development of a methodology that can quickly and accurately assess the feasibility of repurposing abandoned hydrocarbon reservoirs for underground thermal energy storage, offering a practical solution for sustainable energy transition.

II.1. Introduction

Revitalizing abandoned hydrocarbon fields for sustainable energy utilization presents a promising pathway to address the intermittent nature of renewable resources and the global challenge of carbon emissions reduction (Duggal et al., 2022; Gayayev, 2023). This initiative aligns with the urgent call for innovative energy solutions underscored in pivotal studies and reports (IEA, 2017; Paksoy et al., 2000; REN21, 2019; WEF, 2021), aiming to bridge the seasonal supply-demand mismatch and enhance energy security (Dincer & Rosen, 2011; Van Der Roest et al., 2021). It embodies the drive towards energy decentralization and independence, highlighting the critical need for adaptable and reliable energy systems in the face of fluctuating renewable outputs (Lee, 2013; G. Li, 2016). Hydrocarbon fields, upon reaching depletion,

typically require decommissioning and plugging to ensure environmental safety and compliance with regulatory standards. This process is often costly, with median decommissioning expenses estimated at around $76,000 per well, and can exceed $1 million in more complex scenarios due to factors like well depth, age, and location (Decommissioning Orphaned and Abandoned Oil and Gas Wells, n.d.). These costs encompass plugging the well to prevent leaks, removing infrastructure, and restoring the surface to its natural state.

Various initiatives have explored converting depleted hydrocarbon reservoirs for different pore utilization, encompassing CO2 storage (Qin et al., 2023), electricity generation (Duggal et al., 2022), hydrogen storage (Zhu et al., 2024), and Aquifer Thermal Energy Storage (Matos et al., 2019; Stricker et al., 2020), further underscoring the versatility and potential of these underground spaces. Utilizing depleted hydrocarbon reservoirs for Aquifer Thermal Energy Storage (ATES) signifies a pivotal move toward sustainable energy practices. This approach leverages existing geological insights and infrastructure for renewable energy storage and adeptly meets fluctuating energy demands across seasons. ATES, a proven technology, boasts over 2,800 sites globally, particularly in the Netherlands, evidencing its efficacy and adoption (Dickinson et al., 2009; Fleuchaus et al., 2020; Kastner et al., 2017; van Heekeren & Bakema, 2015).

ATES systems underscore a pivotal advancement in renewable energy utilization, offering a strategic solution to the energy mismatch challenge by enabling the seasonal storage and retrieval of heat. This enhances energy security by stabilizing the supply-demand continuum and fosters the integration of renewable sources into the energy grid. Critically, ATES's contribution extends to environmental conservation, showcasing a potential pathway for substantial reductions in CO2 emissions. The proven operational success in regions like the Netherlands is a testament to ATES's role in promoting sustainable energy solutions and advancing environmental preservation goals, marking it as a cornerstone technology in the quest for a greener future (Fleuchaus et al., 2020).

As the lifespan of many hydrocarbon fields approaches its end, stakeholders face a critical decision: undertake the expensive decommissioning process or pivot towards repurposing these sites for renewable energy projects (Mehmood et al., 2019; Soldo et al., 2020). With their inherent geological characteristics, depleted hydrocarbon reservoirs offer a promising avenue for

such energy storage solutions. Their permeability, porosity, and extensive historical data availability position them as ideal candidates for geothermal battery energy storage applications (Green et al., 2021). The operational efficiency of High Temperature-ATES (HT-ATES) systems, particularly in deeper reservoirs that do not affect near-surface groundwater horizons, coupled with the favorable geothermal conditions in regions like the Upper Rhine Graben, underscores the potential of depleted hydrocarbon reservoirs as ideal candidates for geothermal energy storage solutions (Holstenkamp et al., 2017). The study by (Jello et al., 2022) on Advanced Geothermal Energy Storage systems, which investigates field-scale implementation in the low-temperature Illinois basin by utilizing an oil well, complements the operational efficiency of HT-ATES systems in deeper reservoirs, making it a promising solution for sustainable geothermal energy storage (Holstenkamp et al., 2017). Numerous hydrocarbon fields, once bustling with oil and gas production, now stand depleted. We have selected one of the potential formations in the southeast, renowned for its extensive gas production history and substantial water volume, to emerge as a potential site for different pore utilizations, including ATES exploration. To effectively harness these reservoirs for ATES, either for low-temperature (LT) or high-temperature (HT) applications, a detailed methodology is essential for pinpointing the optimal locations within these reservoirs, ensuring their successful transformation into viable energy storage solutions.

The robustness of Multi-Criteria Decision Analysis (MCDA) as a methodological approach for this research is affirmed through its broad application across diverse geoscience and environmental management domains. Its efficacy is evidenced in the selection of sites for compressed air energy storage systems (Matos et al., 2019), where MCDA facilitated the integration of spatial and non-spatial criteria to discern optimal locations, and in water resource management (Manikkuwahandi et al., 2019), where it enabled the balancing of economic, social, and environmental objectives among multiple stakeholders. Furthermore, using MCDA in reservoir system operational planning through a two-stage procedure combining multi-objective optimization and MCDA techniques underscores its versatility in handling quantitative and qualitative objectives (Ko et al., 1994). These varied applications illustrate MCDA's comprehensive capacity to navigate complex decision-making landscapes, making it a fitting choice for evaluating the Depleted Hydrocarbon field's potential for HT-ATES and other pore utilization strategies.

This study proposes a comprehensive methodology for evaluating abandoned hydrocarbon fields for various pore utilization purposes. It utilizes a real hydrocarbon field in Hungary as a case study to assess its suitability for HT-ATES, among other applications. Leveraging MCDA and extensive reservoir data, we aim to pinpoint optimal zones for ATES deployment. The goal is to deliver a lucid, systematic analysis to inform future ATES projects or other energy pore utilizations, thereby enhancing energy sustainability and independence amid escalating environmental and demand-driven challenges.

II.2. Glossary of Terms

Abandoned Hydrocarbon Fields: Oil and gas fields that have ceased production and are no longer economically viable for extraction.

Analytic Hierarchy Process (AHP): A structured technique for organizing and analyzing complex decisions, using mathematics and psychology. It helps in setting priorities and making decisions by reducing complex decisions to a series of pairwise comparisons.

Aquifer Thermal Energy Storage (ATES): A technology that stores thermal energy in underground aquifers for later use, helping to balance the supply and demand of energy.

CO2 Storage: The process of capturing and storing carbon dioxide to reduce its presence in the atmosphere, usually in underground geological formations.

Consistency Check: In AHP, it refers to a process that ensures the consistency of judgments or comparisons made during the decision-making process.

Decommissioning: The process of safely closing and dismantling a facility, such as an oil well, that is no longer in use.

Electro-Lithologies: Rock types identified based on their electrical properties, often derived from gamma-ray logs.

Geostatistical Methods: Techniques used in geosciences to model spatial patterns and distributions.

High Temperature-Aquifer Thermal Energy Storage (HT-ATES): A type of ATES that operates at higher temperatures, typically used for applications requiring more heat energy.

Hydrocarbon Fields: Geological formations containing accumulations of oil and/or gas.

Lithology: The study and description of the petrological characteristics of rocks.

MCDA (Multi-Criteria Decision Analysis): A decision-making process that evaluates multiple conflicting criteria in decision-making, used to evaluate complex problems.

P10, P50, P90: They are parameters describing some characters of a cumulative distribution function (cdf). P10 is the value with a 0.1 probability of being exceeded, P50 with a 0.5 probability, and P90 with a 0.9 probability of being exceeded. These terms help quantify the uncertainty and variability in data, providing a range of possible outcomes.

Permeability: A measure of the ability of a porous material (like rock) to allow fluids to pass through it.

Porosity: The percentage of a rock's volume that consists of open spaces (voids), which can store fluids such as water or hydrocarbons.

Python Scripting: Writing scripts in Python programming language to automate processes and perform complex calculations.

Reservoir: A subsurface pool of hydrocarbons contained in porous or fractured rock formations.

Sensitivity Analysis: A method to predict the outcome of a decision given a certain range of variables, used to understand the impact of different variables on the outcome.

Sequential Gaussian Simulation (SGS): A geostatistical simulation technique used to model spatial distributions of geological variables.

Stochastic Simulation: A method of simulating random variables to predict the probability distribution of possible outcomes in a process.

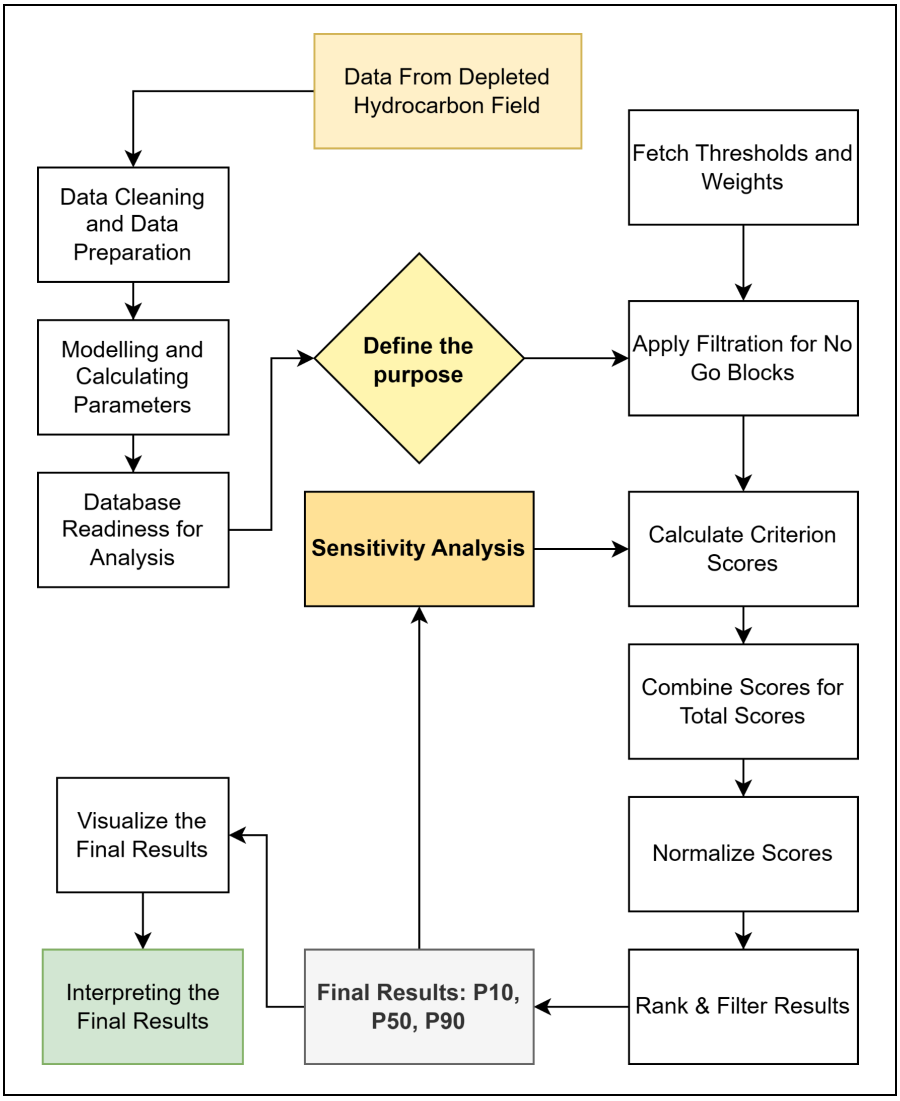
Universal Kriging: A geostatistical method used to interpolate the value of a random field at an unobserved location based on observations at nearby locations, taking into account both the trend and the spatial correlation of the data.

Variogram: A function describing the degree of spatial dependence of a spatial random field or stochastic process.

II.3. Materials and Methods

II.3.1. Methodological Framework

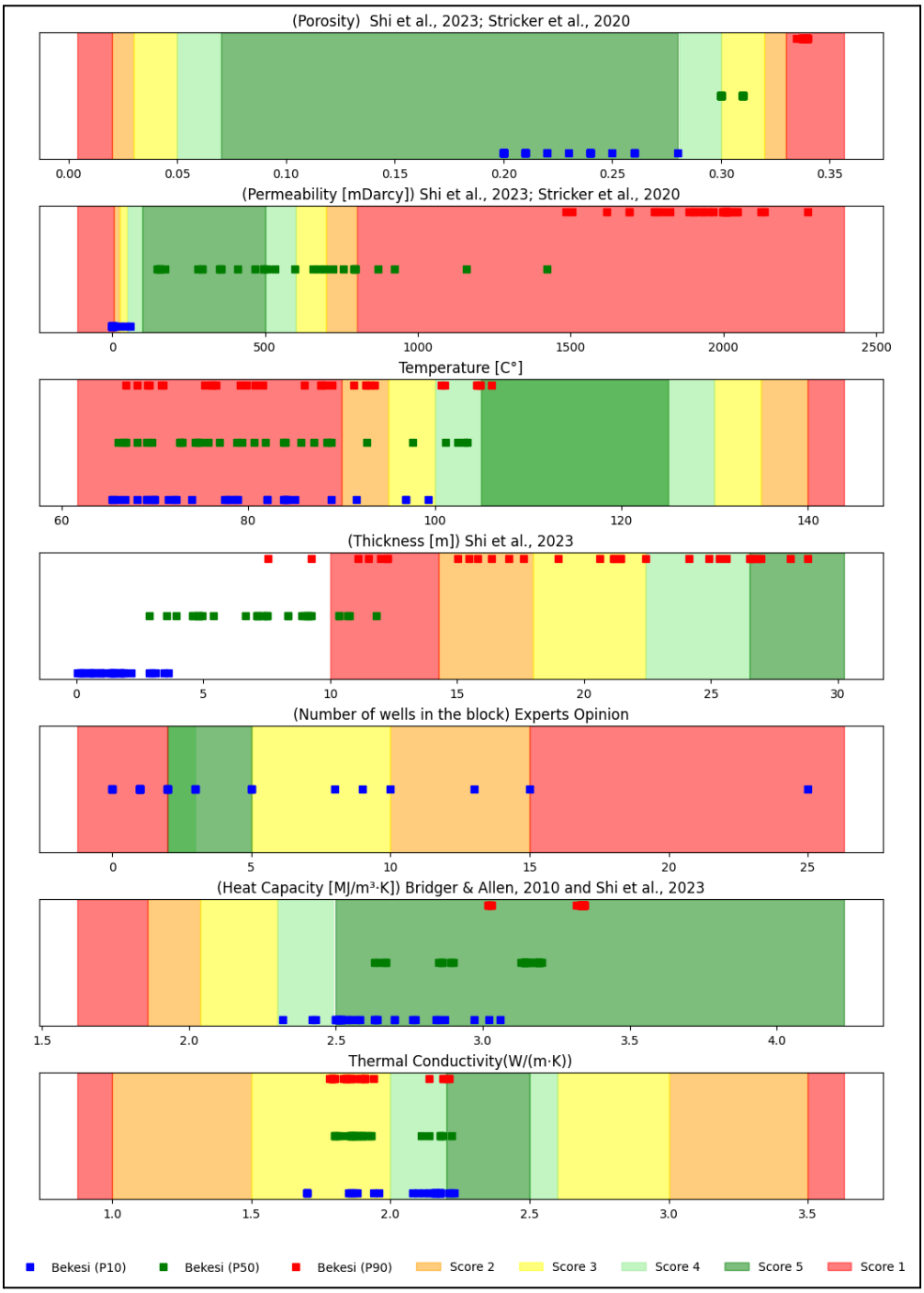
In our study, we adopted the MCDA with an emphasis on the Analytic Hierarchy Process (AHP) for our geostatistical modeled data due to its robust framework for handling complex decision-making scenarios (Yalcin & Kilic Gul, 2017). AHP facilitates the breakdown of the decision problem into a hierarchy of more easily comprehended sub-problems, each of which can be analyzed independently (II.Figure 1). This method is particularly suited to our study's goal of evaluating the suitability of different blocks within the study area for heat storage, as it allows for incorporating both quantitative data and qualitative expert judgments.



II.Figure 1. Data Processing Methodological Framework.

The criteria were weighted based on a thorough local sensitivity analysis reflecting the relative importance of each parameter in the context of ATES systems and high-temperature applications similar to electricity generation. This participatory approach ensures that the weighting scheme accurately captures the priorities and insights of professionals in geothermal energy and heat storage fields. By engaging experts in the weighting process, we aimed to ground our analysis in practical, field-informed perspectives. The criteria were further validated through a consistency check in the Analytical Hierarchy Process (AHP), ensuring the robustness of the weighting scheme. II.Figure 2 provides a detailed visual representation of each parameter, displaying their respective score ranges through a color-coded system: green, pale green, yellow, orange, and red.

The green color signifies the best scores, indicating optimal conditions, while red denotes the worst scores, highlighting areas of concern. Each parameter is illustrated with a series of colored square points—red, green, and blue—representing the data array for that parameter.



II.Figure 2. Criteria Scoring Distribution for Each Parameter and Their Weights (Visualized by

Python:(H. Abdulhaq, 2024)). Quantitative criteria and thresholds are shown in Abdulhaq, 2025f from Tab Thresholds. The threshold of each criterion is supported by different references, which are provided for each parameter (Bridger & Allen, 2010; Shi et al., 2023; Stricker et al., 2020) .

II.3.2. Data Collection

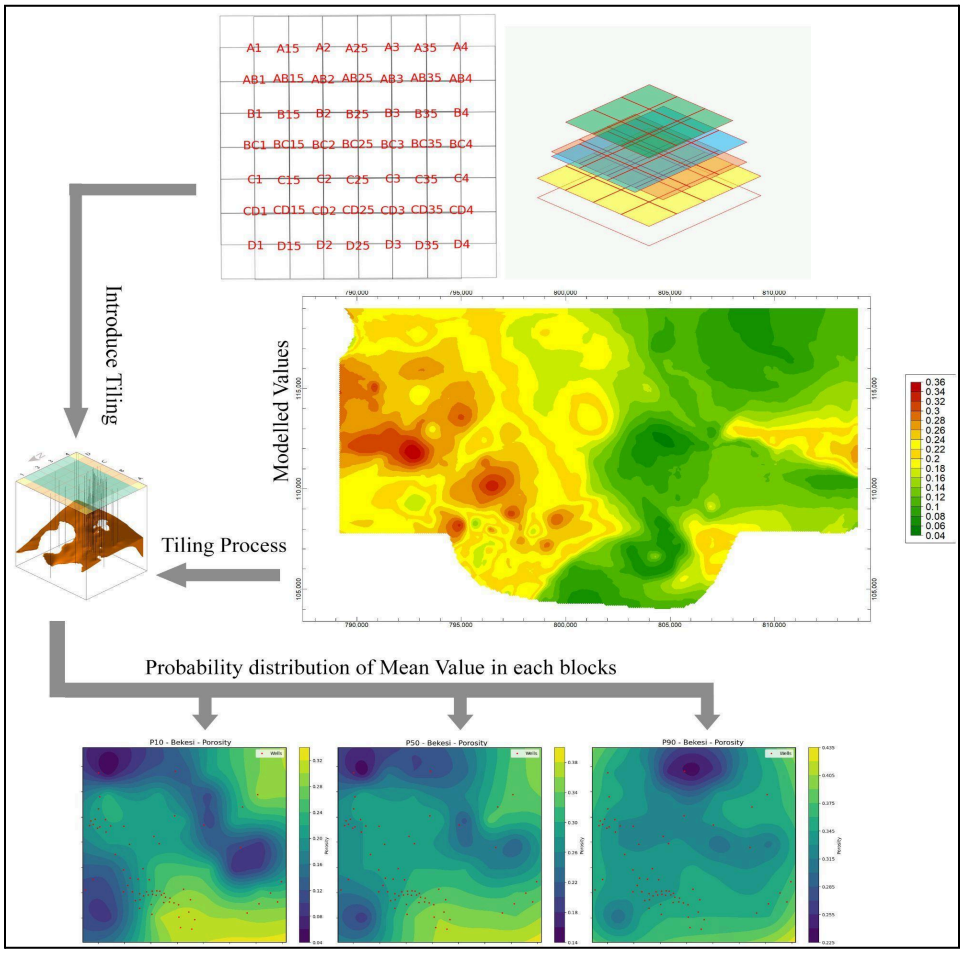
For the Data Collection segment of our study, we utilized data from 67 wells in the study area which is located in the Southern Part of the Bekes Basin, Hungary, a site previously explored for hydrocarbon but now considered for its geothermal potential (Kovács & Teleki, 1994). This field encompasses two potential formations, the shallow Szolnok Formation as an aquifer (Varga et al., 2019) and the deeper Bekes Formation (Varga et al., 2019), offering a unique lens to examine heat storage capabilities. The selection of 3.88 km as the side length for our analysis squares derives from the optimal area typically utilized for injection and production processes within such formations. Though the range for these dimensions could vary between 3 to 4 km, we chose 3.88 km to maximize coverage and efficiency across the study area. Implementing a quarter-overlapping array was used to mitigate potential spatial biases, thereby enhancing the probability of identifying the most suitable locations for our geothermal applications. This overlapping design ensures comprehensive coverage and uniformity in data analysis and is essential for accurately evaluating the site's potential. The integration of different types of data, including core data, density logs, resistivity logs, and gamma-ray measures, was conducted through a stochastic simulation process detailed further in the Data Modelling Chapter. This integration, crucial for converting well data into a comprehensive spatial model, facilitated the mapping of average reservoir properties. We applied SGeMS for geostatistical simulation, while RockWorks was used to integrate the different simulation results obtained for gross thickness, effective porosity, and permeability. Heat capacity, thermal conductivity, and temperature were modeled using Python scripts.

The Python script utilized for further data processing employed libraries such as NumPy, pandas, and matplotlib for advanced computational tasks. This script, executed on the Google Colab platform (Elabd et al., 2022), generated cumulative histograms that enhanced our understanding of reservoir properties with significant data variability. The analysis of uncertainty must come from the geostatistical simulations. The histograms by themself are not appropriate for this

purpose. All processed data were systematically organized into Google Sheets, preparing for the MCDA.

II.3.3. Data Modeling

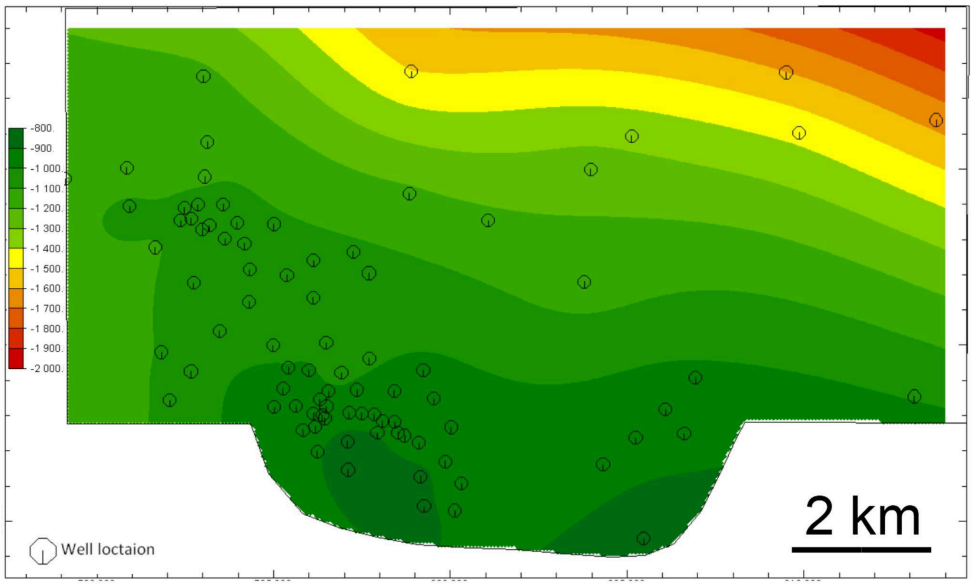
Data modeling was based on 67 wells in the study area. Porosity and permeability data predominantly came from quantitative log interpretations, but core measurements were also used to validate log data. The main purpose of data modeling was to produce 2D stochastic models of gross thickness and 3D stochastic models of porosity, permeability, and lithology distributions for geothermal modeling. The dimensions of each 3D model were 100 x 100 x 2m (II.Figure 3).



II.Figure 3. Modeled Values Prepared for Spatial Tiling Process.

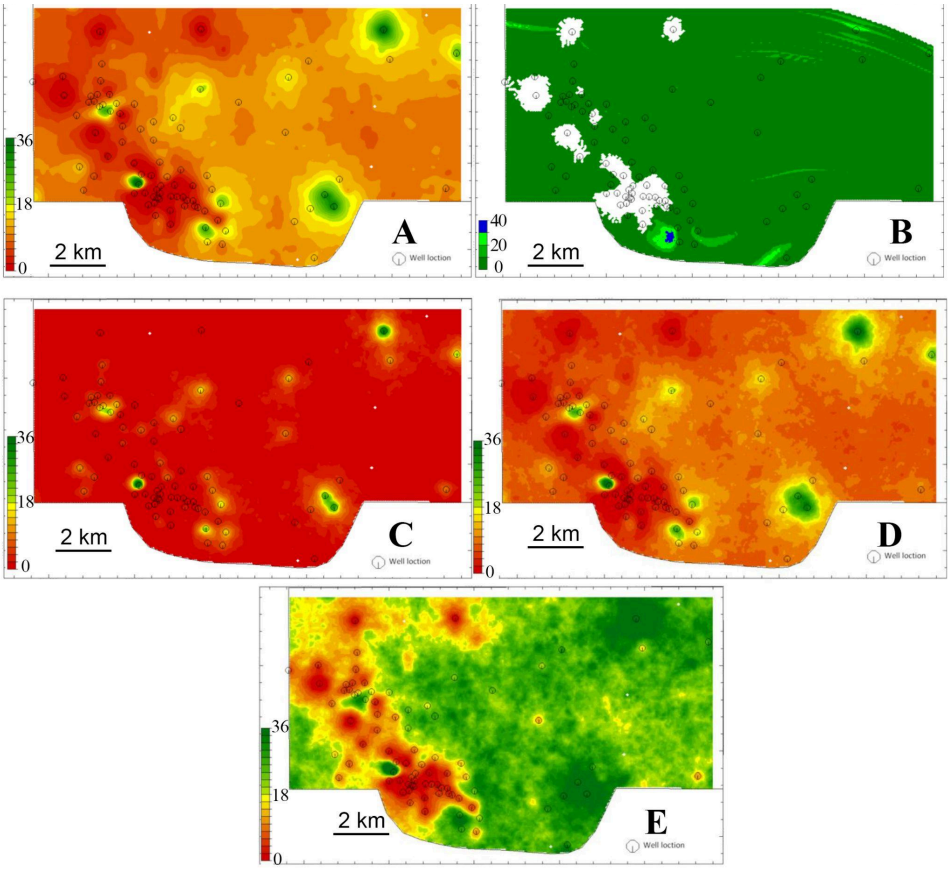
II.a.Modeling of Stratigraphy

The top of the Bekes Formation showed a significant north-south trend (II.Figure 4). Therefore, a non-stationary geostatistical model was the most appropriate choice. The applied Universal Kriging (Wackernagel, 1995) decomposed the variable into its trend and a stationary residual. A variogram was fitted to the residuals. In this approach, the trend residuals were kriged, and the estimated residuals were added to the trend model.



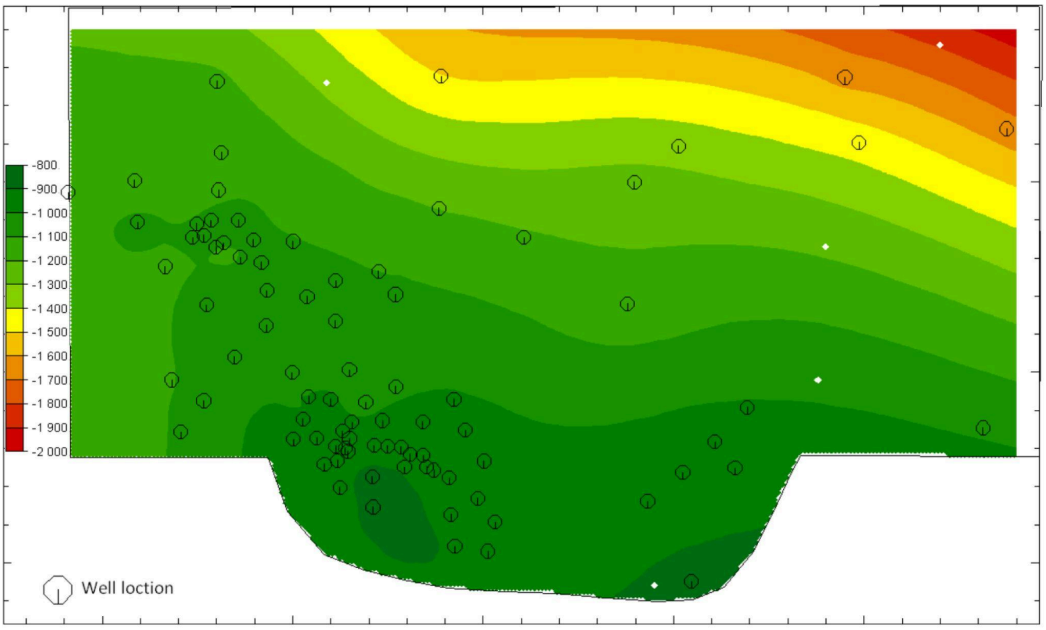
II.Figure 4. The top of the Bekes Formation obtained using Universal Kriging.

The thickness data were transformed into normal space; then, declustering weights were applied to obtain a representative probability distribution. Finally, the declustered normal values were extended into the plan using Sequential Gaussian Simulation (*Az SGS a világ vezető minőségellenőrző, ellenőrző, vizsgáló és tanúsító cége.*, n.d.; Swan, 1999). We generated 100 independent and equally probable stochastic realizations from which the following grids were calculated: (1) an E-type grid showing the most probable lateral thickness distribution; (2) a grid of the coefficient of variation calculated at each grid point from the simulated 100 values to show the uncertainty of the model; (3) three quantile grids showing the P10, P50, and P90 quantiles at each grid point (II.Figure 5).



II.Figure 5. The results of the thickness simulation. A= E-type estimation; B= coefficient of variation (uncertainty of the simulation model); C= P10; D=P50; E=P90.

The base of the formation was estimated by subtracting the E-type thickness grid from the previously estimated top grid ((II.Figure 6)).

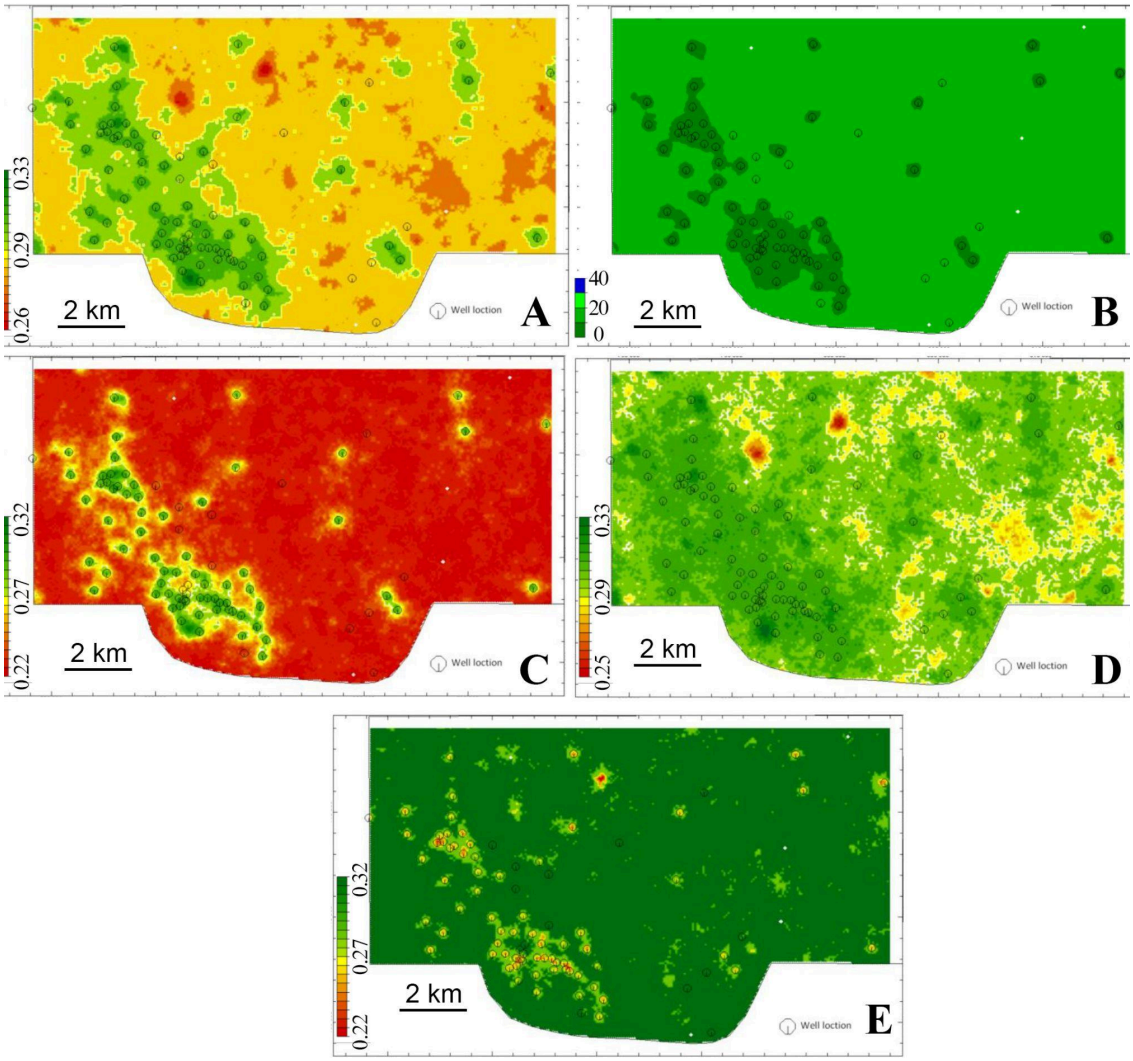


II.Figure 6. The base of the Bekes Formation is obtained by subtracting the thickness grid from the top grid.

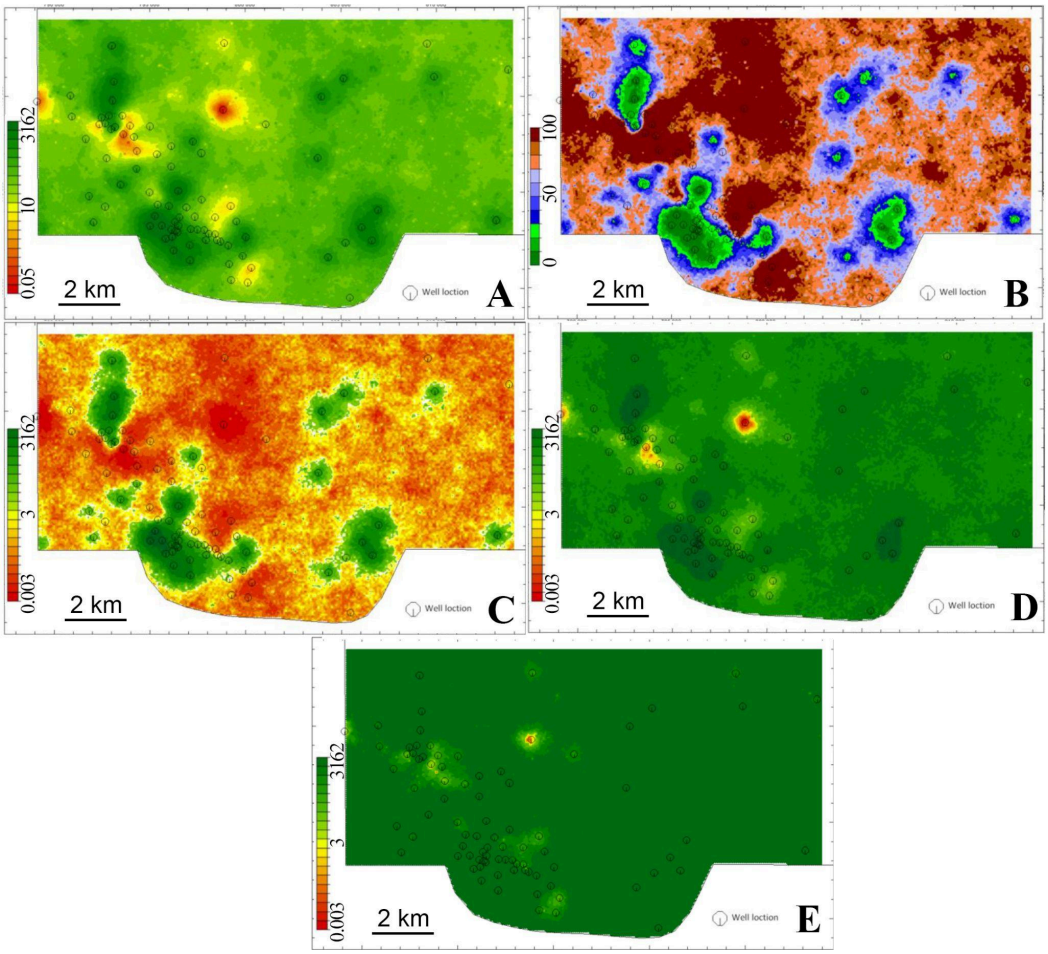
II.b. Modeling Porosity and Permeability

During the exploratory analysis phase, we decided to transform all log-derived porosity and permeability data to the scale of the core measurements. We applied a q-q transformation (Swan, 1999) of the log scale porosity and permeability distributions to the core scale distributions. To simulate the 3D spatial distributions of these two variables, the SGS approach was applied to the data sets transformed into normal space using 3D declustering weights.

Again, 100 independent equally likely 3D realizations were generated, from which the 3D distribution of the P10, P50, and P90 quantiles was calculated in addition to the 3D array of expectation and the coefficient of variation. The means of these 3D models per grid column were fitted to a 2D grid system (II.Figure 7 & 8).



II.Figure 7. The results of the porosity simulation. A= E-type estimation; B= coefficient of variation (uncertainty of the simulation model); C= P10; D=P50; E=P90.



II.Figure 8. The results of the permeability simulation (all scales are mD). A= E-type estimation; B= coefficient of variation (uncertainty of the simulation model); C= P90; D=P50; E=P10.

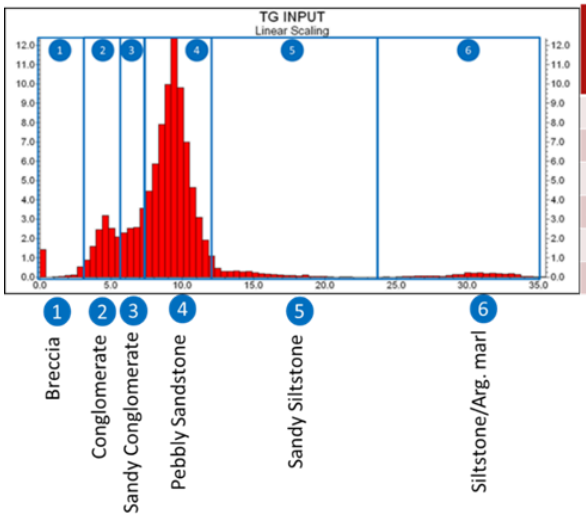
II.c. Lithology Model

The lithology model of the study area was created using the 3D gamma model. In this approach, we first defined the electro-lithologies using the gamma log in the available boreholes. The next step was to build a 3D gamma model. Finally, we transformed the voxel values of the 3D gamma model into electro-lithology types using the minimum and maximum gamma values of the previously defined electro-lithologies.

There were 28 boreholes from which gamma logs were available. At the EDA phase of the gamma data analysis, the main purpose was to identify some electro-lithological types based on

the gamma values. We used EM clustering (Dempster et al., 1977) to segment the original histogram of gamma values into different groups. We were then able to compare the EM clusters to the real (geological) rock types of the borehole columns by regressing the gamma values of each cluster to the real stratigraphic positions. These comparisons calculated for different numbers of clusters were used to identify the optimal number of groups in the EM clustering. As a result, we defined six electro-lithological types (II.Figure 9).

A

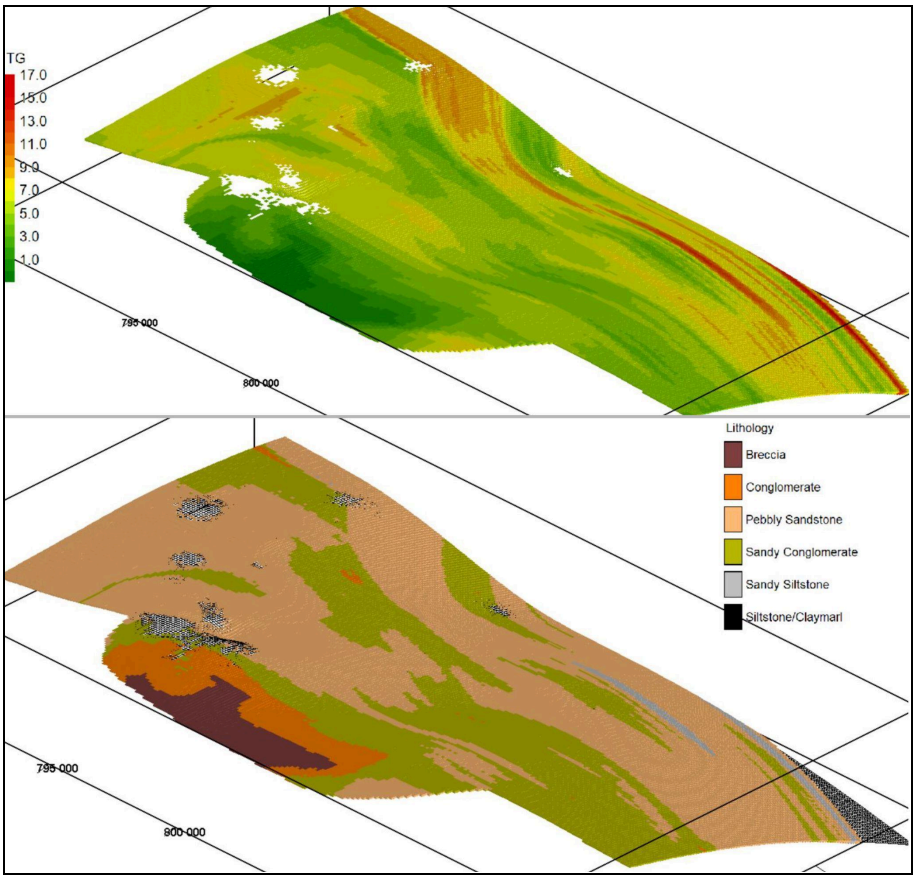


B

Rock types	TG	
	Lower Limit	Upper Limit
1. Breccia	0	3.12
2. Conglomerate	3.12	5.50
3. Sandy Conglomerate	5.50	7.32
4. Pebbly Sandstone	7.32	12.00
5. Sandy Siltstone	12.00	23.86
6. Siltstone/Argillaceous marl	23.86	36.00

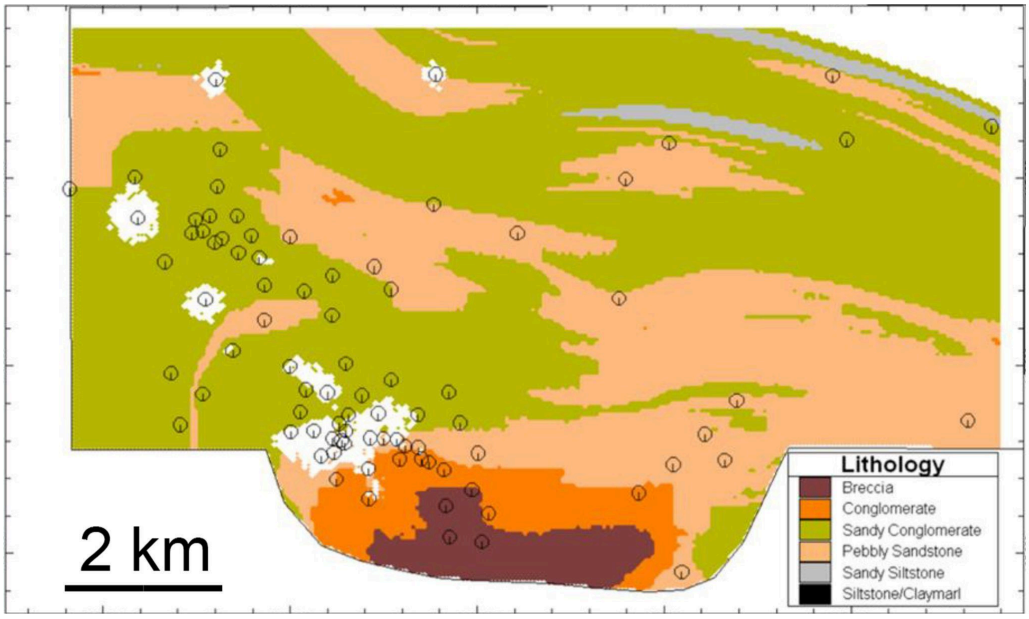
II.Figure 9. The electro-lithological types are defined using EM clustering. “A”: the final result of the EM clustering; “B”: the defined electro-lithologies with the corresponding smallest and largest gamma values.

To increase the stability of the 3D estimation of the gamma values, we added 12 virtual (auxiliary) locations to the 3D input. First, we generated an initial, very coarse 3D gamma model with a resolution of 500x500x2 m using a simple inverse distance estimation with a high-fidelity constraint. From this model, we back-estimated the gamma values to each vertical voxel of the 12 auxiliary wells and added these back-estimated values to the original 3D gamma input set. This extended input set generated the final high-resolution 3D gamma model (Fig. 10, A). The corresponding 3D lithology model (Fig. 10, B) was obtained by converting the voxel values of the 3D gamma model into electro-lithological categories using the minimum and maximum gamma values of the EM clusters (II.Figure. 9, B).



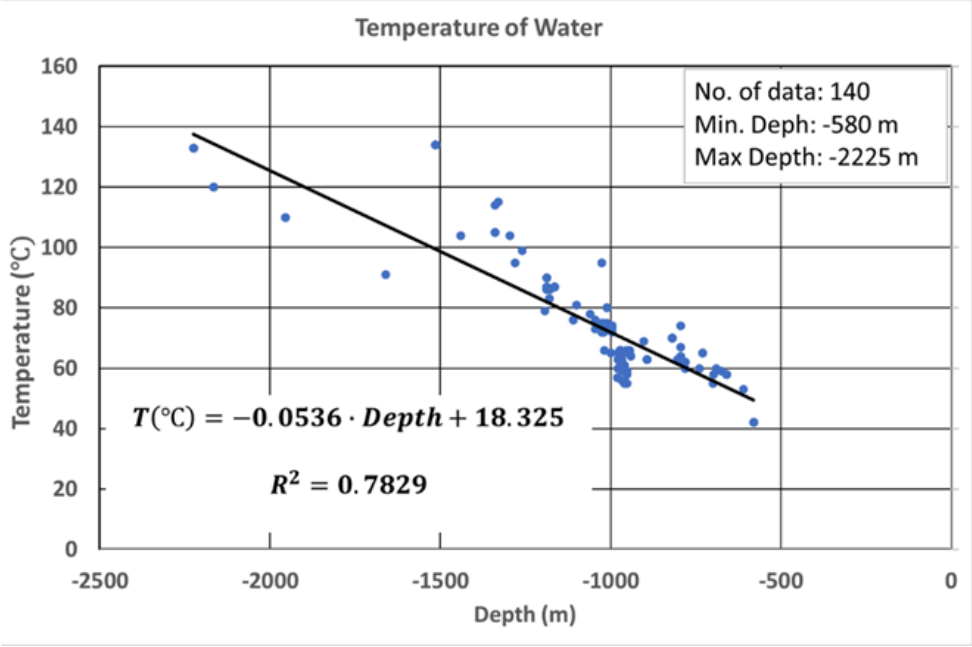
II.Figure 10. The 3D Gamma model (A) and the corresponding 3D lithology model (B).

Fig.11 shows the 2D map of the most frequent electro-lithological category in the voxel columns of the 3D electro-lithological model (Fig. 10, B).

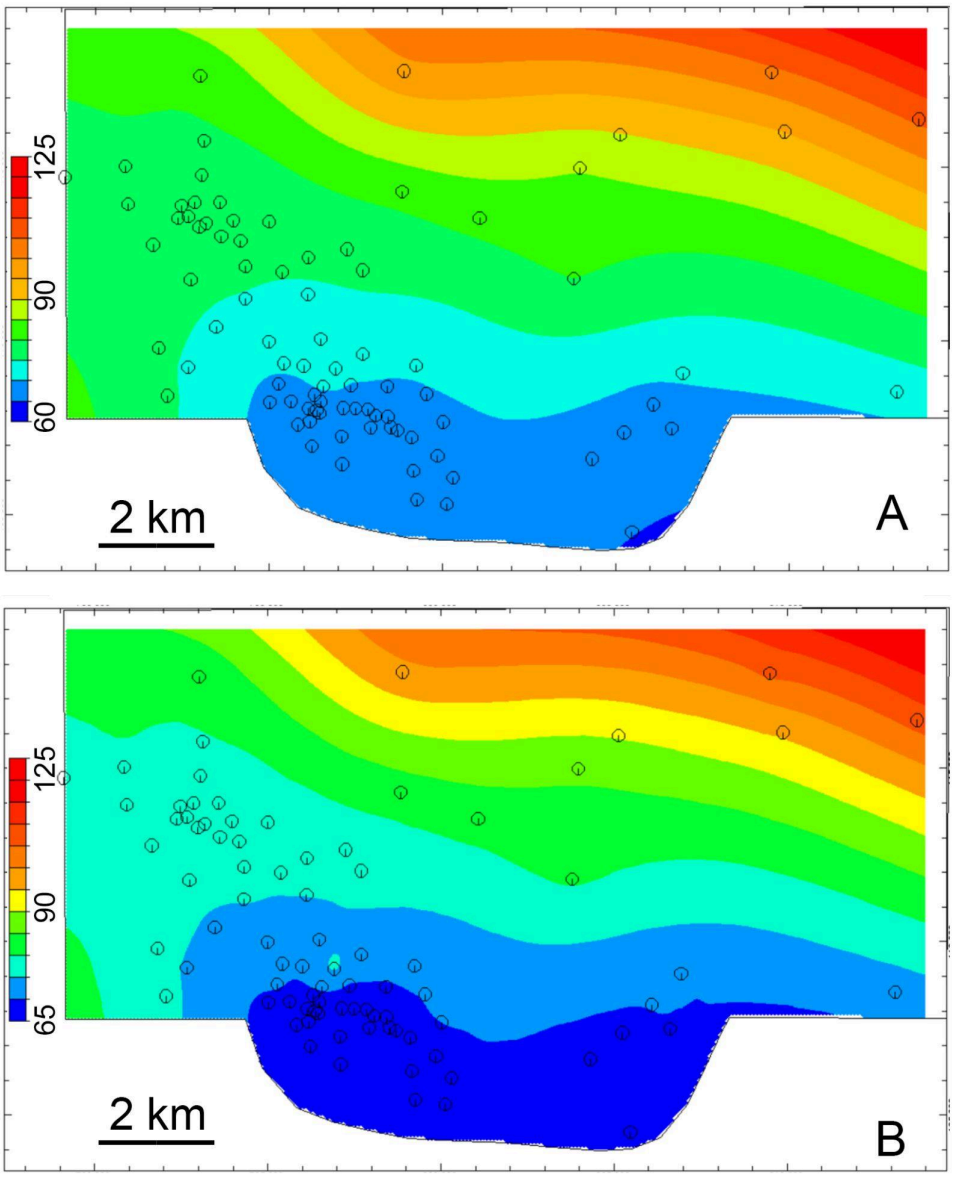


II.Figure 11. Map of the dominant electro-lithologies.

The temperature parameter was modeled using data from Drilling Steam Tests and bottom hole temperatures from the MOL database. There was an acceptable strong linear regression describing the temperature change with increasing depth (II.Figure 12).



II.Figure 12. shows the applied linear regression describing the change in temperature with increasing depth. This regression was applied to both the top and bottom grid to obtain the temperature distribution at the top and bottom of the Bekes Formation (II.Figure 13, A and B).



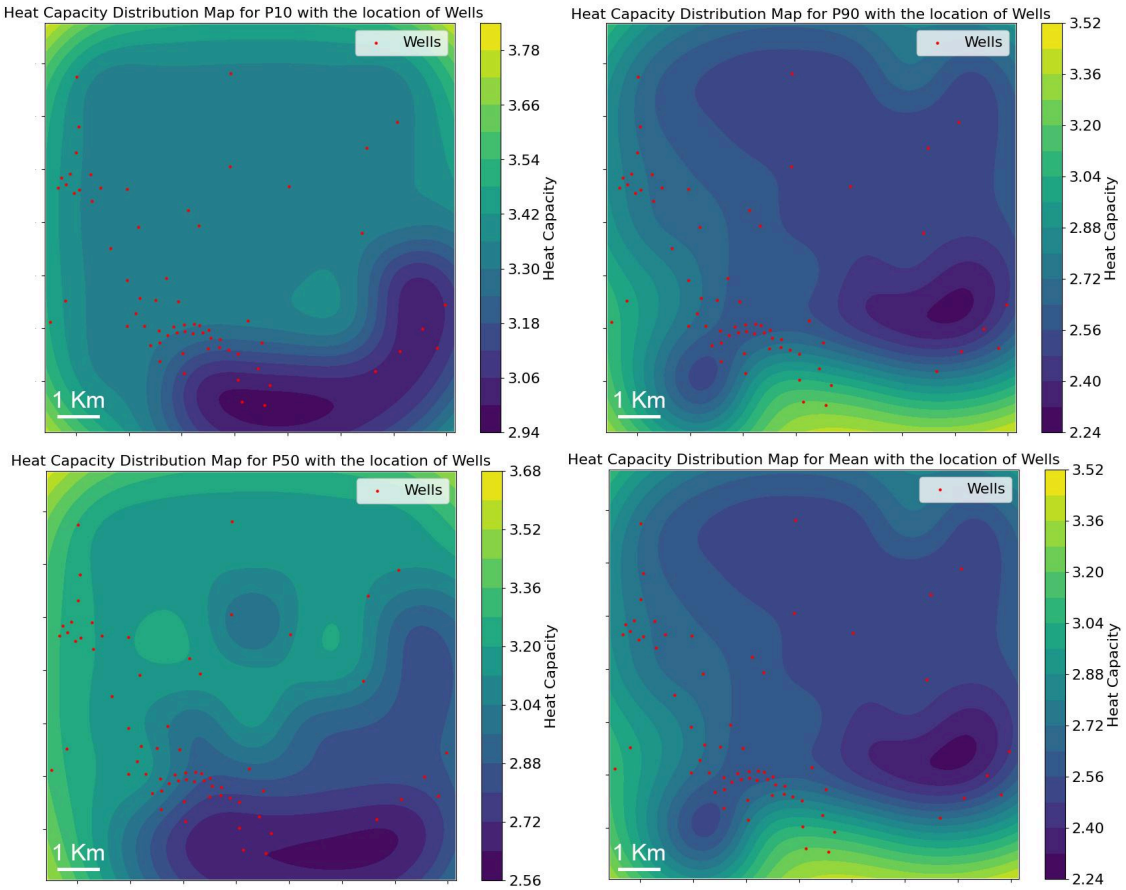
II.Figure 13. Maps of the temperature distributions [°C] at the top (A) and base (B) of the Bekes Formation.

II.d. Heat Capacity

Heat capacity for each block was calculated based on porosity, rock type, density, and estimated specific heat capacity. The formula used for this calculation is as follows:

$$Heat\ Capacity\ (MJ/m^3) = (\frac{Density\ (kg/m^3) \times Specific\ Heat\ Capacity\ (J/kg\backslash.K)}{1-Porosity}) \times 10 - 6$$

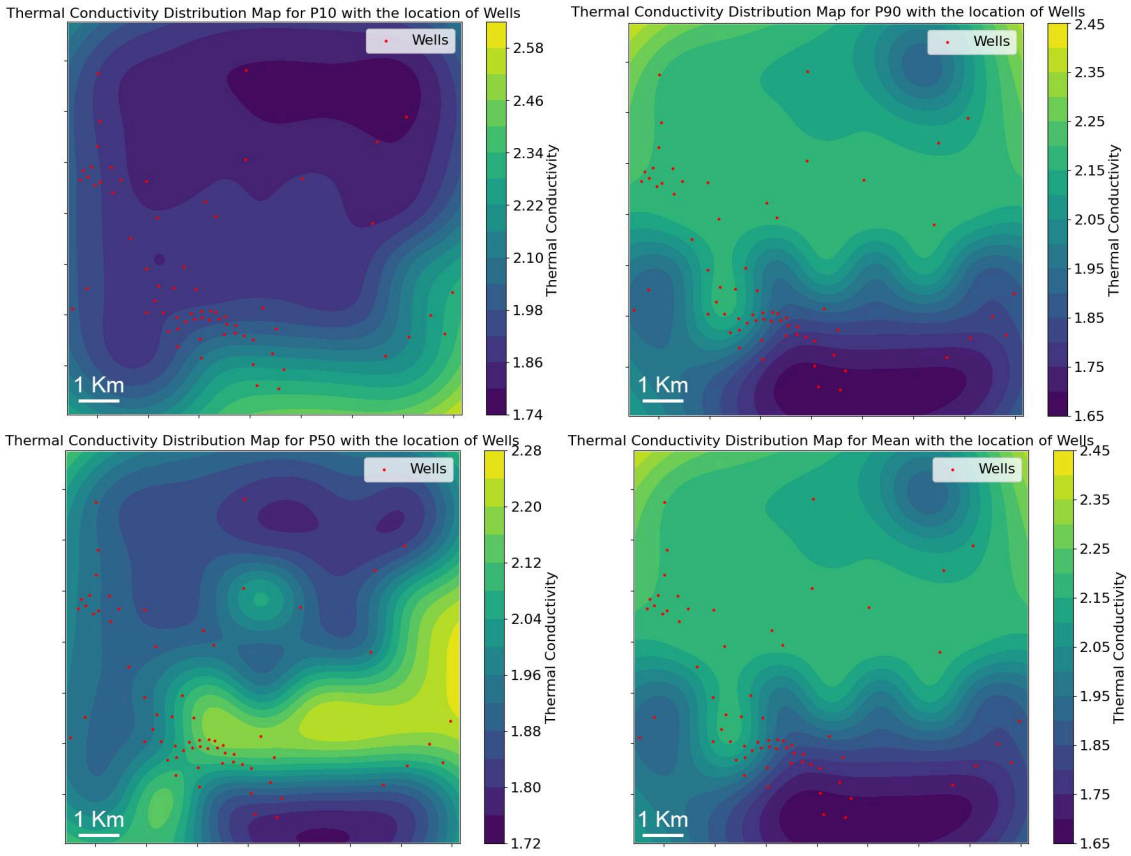
This formula derives the heat capacity in megajoules per cubic meter (MJ/m³). This method was applied across different scenarios—P10, P50, and P90—to assess the heat capacity variability in each block (II.Figure 14).



II.Figure 14. Heat Capacity in MJ/kg·K.

II.e. Thermal Conductivity

In our study, we referenced the findings of (Vass et al., 2018), who reported an average thermal conductivity of 2.03 W/m·K for the Bekes formation, corrected for temperature effects. This average served as a benchmark for comparing our calculated values. Given the variability in lithological characteristics discerned from total Gamma Ray modeling, it was imperative to evaluate how thermal conductivity changes in response to lithology, temperature, and porosity within the formation (Haffen et al., 2017). To this end, we established base thermal conductivity values specific to each identified rock type within the formation—Conglomerate, Breccia, Sandy Conglomerate, and Pebbly Sandstone. These values reflect the inherent thermal properties influenced by the geological composition. The thermal conductivity for each geological scenario—P10, P50, and P90—was then adjusted by a factor that accounts for the porosity, reducing conductivity by 10% per unit porosity to model the impact of increased pore space on heat transfer capabilities. Furthermore, we incorporated a temperature correction, decreasing conductivity by 0.2% for every degree Celsius increase above a baseline temperature of 25°C (II.Figure 15). With this methodology, we remained within the average thermal conductivity range of the Bekes Formation. However, we generated local changes in the function of the thermal conductivity equation to estimate the variation of thermal conductivity across the formation, providing a nuanced understanding of how specific geological conditions affect thermal properties.



II.Figure 15. Thermal Conductivity Distribution Maps for P10, P50, P90, and Mean Scenarios with Well Locations. These maps illustrate the distribution of thermal conductivity for different scenarios: P10 (10th percentile), P50 (50th percentile or median), P90 (90th percentile), and the mean. The color scale ranges from purple (low thermal conductivity) to yellow (high thermal conductivity). Red dots represent well locations used in the study.

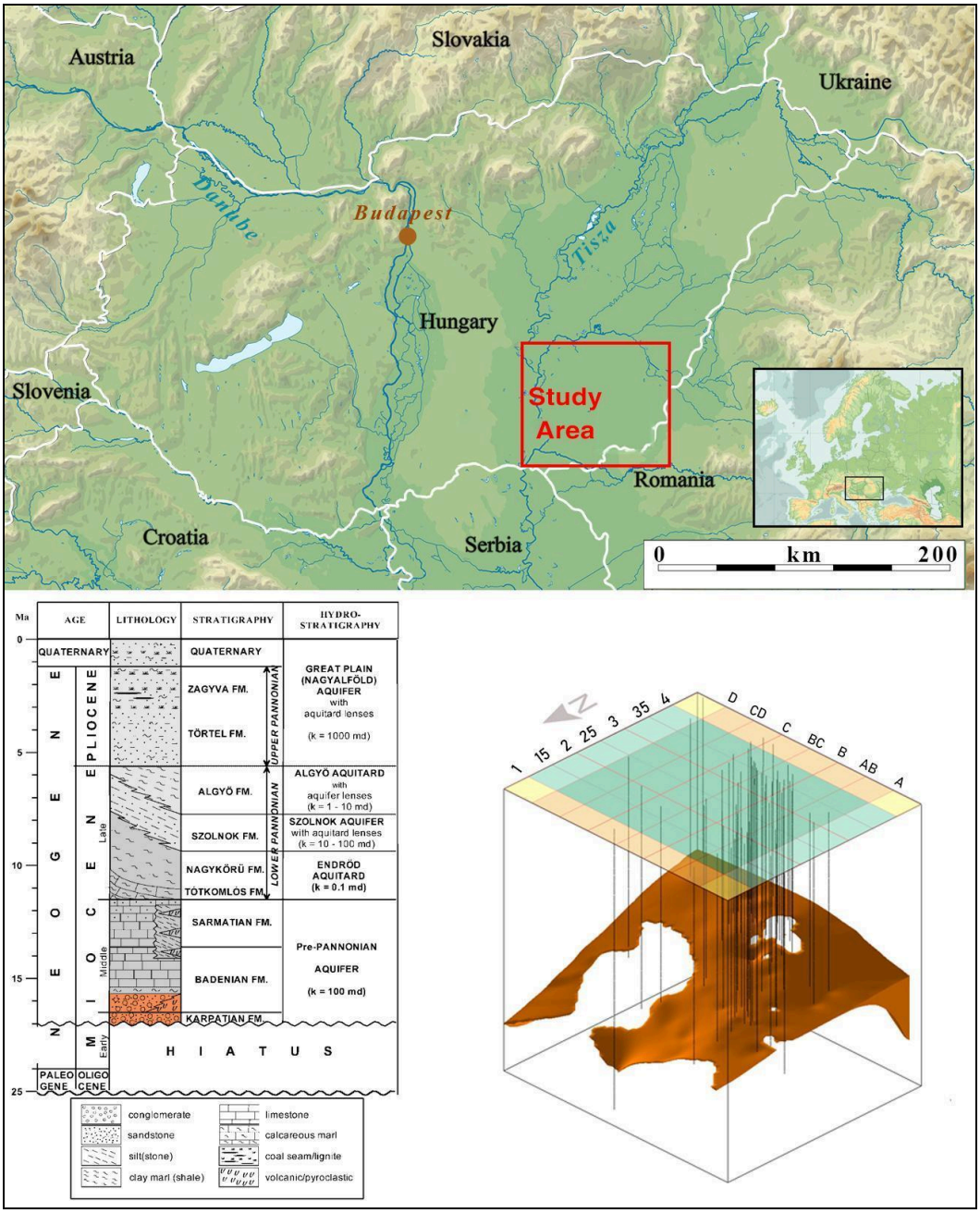
II.f. Criteria Selection

A comprehensive literature review was conducted to select criteria for the transition of hydrocarbon reservoirs to ATES systems. (Collignon et al., 2020; Dickinson et al., 2009; Fleuchaus et al., 2020; Holstenkamp et al., 2017; Paksoy et al., 2000; Réveillère et al., 2013; Todorov et al., 2020). This review identified key factors critical to the assessment of potential sites: Porosity, Permeability, Reservoir Temperature, and Thickness, along with Heat Capacity and Thermal Conductivity. Additionally, the presence and availability of wells within the reservoir were considered essential due to their direct impact on the feasibility of converting the

hydrocarbon reservoir for ATES use. Python script was used for the visualization of the modeled data (Abdulhaq, 2025f).

II.g. Case Study Selection

The selection of this study area for MCDA application is underscored by its distinct geological features, notably within the Bekes Formation, known for its conducive geothermal attributes (Czauner & Mádl-Szőnyi, 2013; Kun et al., 2022). The Lower Pannonian Békés Conglomerate, with an average thickness of 30-40 meters and a maximum of 50 meters (II.Figure 16). The reservoir thickness decreases towards the top of the basement high. The reservoir has an average porosity of 25% and a permeability of 400 milli Darcy, with a maximum of 2.5 D. The reservoir pressure is hydrostatic, and the geothermal gradient is 57°C/km. Oil production from this reservoir is limited to a few wells, with water cut more significantly than 95%. However, the reservoir's proximity to a potential heat market and the dense well network of over 70 wells make it a potential candidate for geothermal completion. However, no water injection exists, and significant oilfield water production occurs, estimated at 1000-1200 cubic meters per day. The Paleozoic granite-quartz porphyry contains porous rock inclusions resulting from the surface and near-surface corrosion, fracturing, and crack network caused by tectonic movements. Fluid movement is upward; maximum oil production occurred between 1961 and 63 at 25,000 m3/year.



II.Figure 16. Overview of the Study Area in Hungary: Bottom Left - Lithological Section (Tóth & Almási, 2001); Bottom Right - Position of Bekes Formations in the Study Grid with Vertical Exaggeration (13.33X); Above - MCDA Grids with Well Locations Indicated by Black Dots.

II.3.4. MCDA Application Process

The MCDA process for assessing the suitability of locations for heat storage involves a rigorous and methodical application, incorporating various tools and data sources. Temperature, permeability, and thickness data were critical in filtering potential sites based on predefined criteria tailored to the purpose of HT-ATES or Direct High-Temperature Utilization (Abdulhaq, 2025f). This selection process was further refined by applying weights and thresholds to evaluate each site's compatibility with the project goals. The MCDA analysis harnessed Python's analytical capabilities, particularly its pandas, numpy, and matplotlib libraries, to process and visualize the data, thus providing a comprehensive and clear depiction of the most viable locations for heat storage.

II.3.5. Sensitivity Analysis

The sensitivity analysis conducted on various criteria, such as porosity, permeability, temperature, thickness, well density, heat capacity, and thermal conductivity, reveals crucial insights into the stability and robustness of these parameters under weight adjustments. This local sensitivity analysis method involved varying one criterion's weight at a time while proportionally adjusting the others to ensure the total sum of weights remained 1. The impact of these adjustments on the overall scores was evaluated using variance minimization. The equations used in this analysis include the weighted sum calculation:

$Total\ Score\ =\ \Sigma i$ (Criterion i * Weight i)

and the variance calculation:

$$Variance\ =\ \frac{1}{N}\ \Sigma i\ (Score\ i\ -\ Mean\ Score)^{\ 2}$$

This analysis was essential in determining the resilience of the heat storage site selection process to changes in evaluation criteria weights. As illustrated by the color-coded maps, areas exhibiting more intense colors indicate regions where slight weight changes could significantly impact the overall scores of potential sites. The results highlight the need to carefully calibrate these weights to ensure a balanced and accurate site assessment.

II.Table 1 Assigned weights of evaluation criteria used in the sensitivity analysis

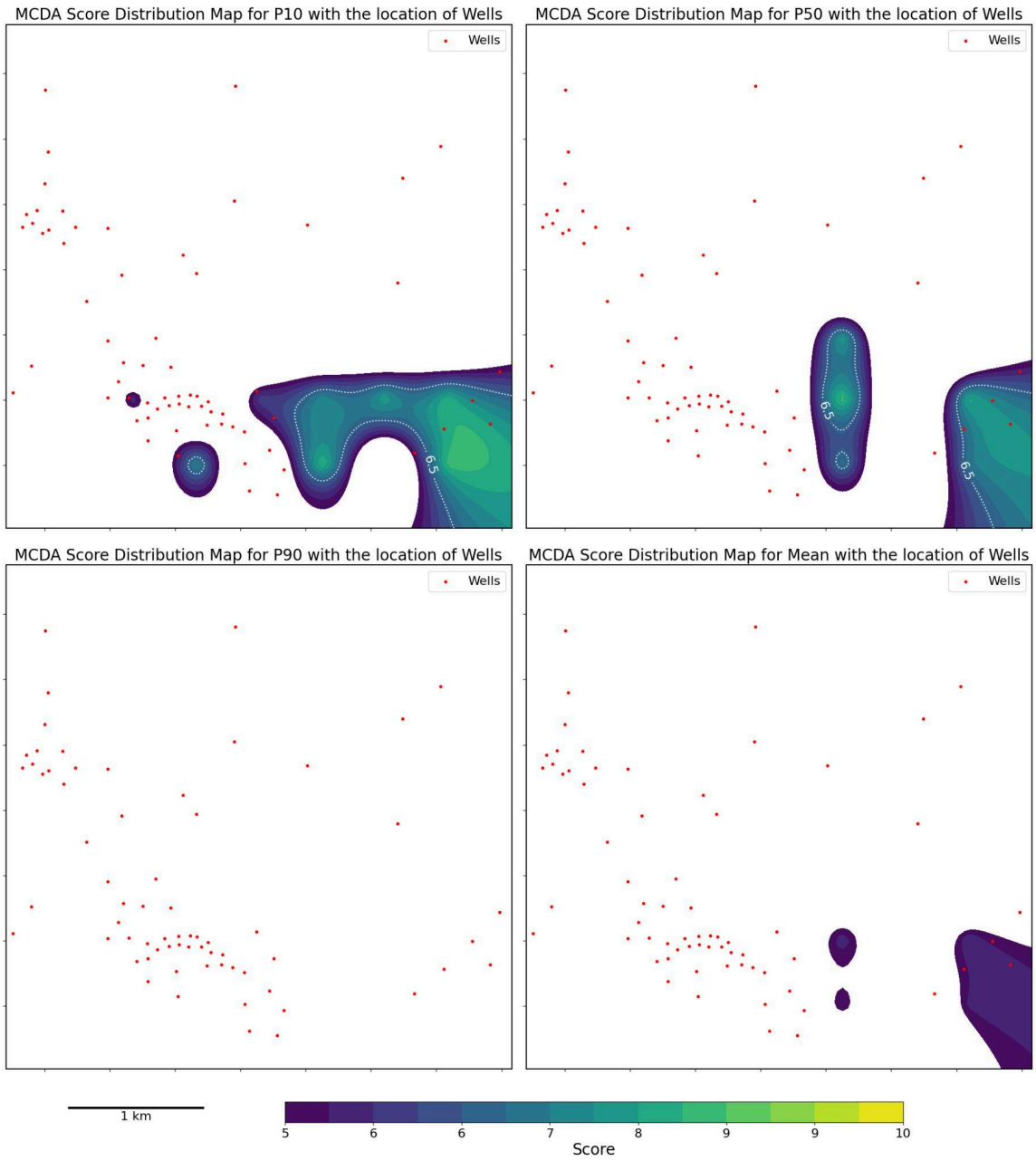
Criterion	Suggested Weight
Porosity	0.122
Permeability	0.079
Temperature	0.069
Thickness	0.141
Well Density	0.119
Heat Capacity	0.208
Thermal Conductivity	0.263

II.4. Results

In the Results section of our MCDA study on heat storage, we focused on evaluating potential locations based on a set of criteria critical for heat storage capabilities. Our analysis revealed that permeability and reservoir temperature are paramount, with acceptable permeability not nearing zero and temperatures set for specific purposes at the outset. The MCDA analysis assigned scores, setting a threshold at 6.5 to distinguish optimal sites. Scores above 6.5 indicated prime locations, while those below 6.4, if not filtered out, appeared as NaN, demonstrating a method to assess potentiality against established criteria rapidly (Abdulhaq, 2025f).

II.4.1. HT-ATES

Setting the focus on HT-ATES, we defined critical parameters, including a temperature range of 50 to 70°C, a minimum thickness of 10 m, and a minimum permeability threshold of 0.1 milli Darcy. Our MCDA framework identified the Bekes Formation as an exemplary candidate for HT-ATES in all scenarios, scoring an outstanding above threshold repeating in the same area, reflective of its favorable conditions for heat storage (II.Figure 17).



II.Figure 17. Results for HT-ATES Suitability. As illustrated in this figure, the areas marked in color indicate potential zones for energy storage as they meet the necessary criteria. The white areas represent regions that do not meet the required thresholds for suitability. The figure clearly delineates suitable zones for High Temperature-Aquifer Thermal Energy Storage (HT-ATES), aiding in the identification of optimal locations for energy storage initiatives.

II.a. Probability Distribution of the lowest risk P90

The assessment of potential heat storage sites, as illustrated in II.Figure 18, reveals that no sites were identified under the lowest risk category due to the stringent criteria established prior to the MCDA. The Python script employed in the analysis excluded any blocks that did not meet all purpose criteria, effectively functioning as a 'no-go' filter. The primary issues identified were temperatures falling outside the desired range of 50.0-75.0°C and thicknesses consistently below the required 10.0 meters threshold, with observed temperatures reaching up to 99.3°C and thicknesses as low as 0.04 meters. This rigorous filtering underscores the challenge of finding suitable heat storage sites and emphasizes the importance of precise criteria in ensuring the selection of viable locations.

II.b.Probability Distribution of the Medium risk P50

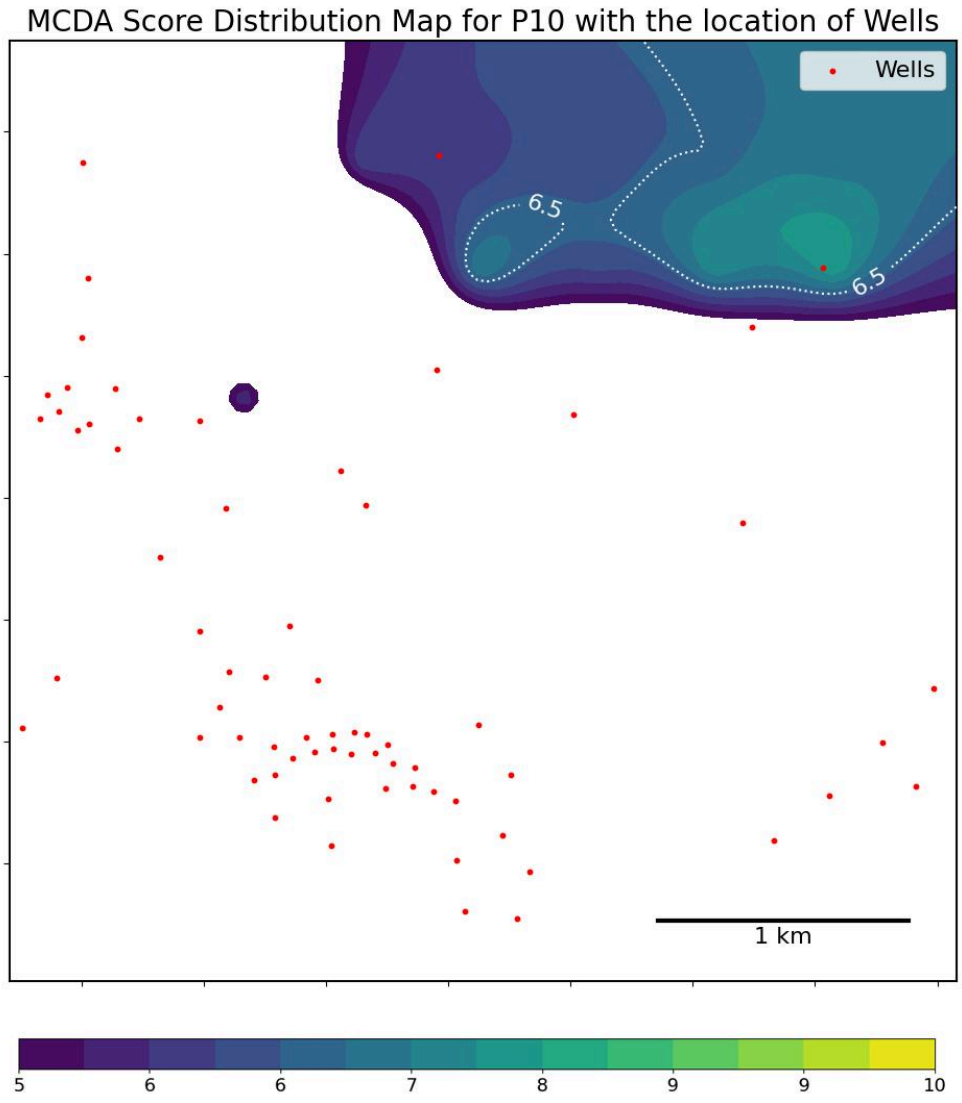
In the probability range of P50, three blocks were qualified based on their favorable characteristics, including well densities of 2 and 5, porosities of 0.3 and 0.31, and permeabilities of 411.63 and 795.97 milliDarcy. These blocks also met the criteria with densities of 2.5, temperatures between 70.1°C and 72.6°C, thicknesses ranging from 10.35 to 11.8 meters, heat capacities from 2.85 to 2.89 MJ/m³·K, and thermal conductivities of 2.19 to 2.2. These parameters highlight their suitability as potential heat storage sites within the P50 probability range.

II.c. Probability Distribution of High Values P90

In the high-risk probability range, many blocks were qualified due to their favorable characteristics, including well densities ranging from 2 to 21, porosities between 0.33 and 0.33, and permeabilities from 1728.78 to 2189.89 milliDarcy. These blocks also demonstrated suitable densities of 2.4 to 2.5, temperatures from 66.9°C to 74.9°C, and thicknesses between 10.5251 and 28.8116 meters. Additionally, their heat capacities ranged from 3.01 to 3.34 MJ/m³·K, and thermal conductivities varied from 1.92 to 2.21. These parameters underscore the potential of these blocks as viable heat storage sites within the high-risk probability range.

II.4.2. High-Temperature Direct Application

For direct high-temperature utilization, neither P90 (lowest risk) nor P50 (medium risk) were qualified due to not meeting the required criteria of a minimum temperature of 100°C, maximum temperature of 140°C, permeability of 0.01, and thickness of 10 meters, with observed temperatures ranging from 66.0°C to 99.3°C and thicknesses consistently below 10 meters. However, several blocks were qualified based on favorable attributes, including well densities ranging from 0 to 2, porosities between 0.34 and 0.34, and permeabilities from 1485.54 to 2004.90 milliDarcy (II.Figure 18). These qualified sites also demonstrated densities of 2.4, temperatures from 100.7°C to 106.1°C, and thicknesses between 15.80 and 26.51 meters. Additionally, the heat capacities ranged from 3.33 to 3.34 MJ/m³·K, with thermal conductivities varying from 1.78 to 1.80. These parameters underscore the suitability of these blocks for direct high-temperature applications, meeting the required criteria for temperature and thickness.



II.Figure 18. Direct Heat Utilization Zones. This figure highlights zones suitable for direct heat utilization, with temperatures ranging between 100°C and 140°C. The marked areas indicate regions that meet the criteria for effective direct heat utilization, ensuring optimal use of geothermal energy for heating purposes.

II.5. Discussion

Operators accumulate extensive datasets in hydrocarbon extraction, including well data, seismic records, reservoir characteristics, core analyses, and fluid testing results, alongside comprehensive stochastic simulations. These legacy data, crucial for repurposing initiatives, underpin our methodology to strategize the conversion of existing wells, which can be costly to decommission. Our approach leverages stochastic simulation data when available, which is a common asset among hydrocarbon companies. However, even without such data, well data can still facilitate the modeling process, albeit with increased uncertainty in the predictions.

Permeability and reservoir temperature critically influence the viability of HT-ATES systems. The strict threshold for permeability (>0.1 milli Darcy) serves as a fundamental criterion, immediately disqualifying blocks that do not meet this requirement. This is crucial for ensuring fluid can effectively circulate through the reservoir, which is non-negotiable for the operational success of HT-ATES (Tas et al., 2023). Similarly, the temperature criteria, tailored to high-temperature applications, ensure only those blocks capable of providing the necessary thermal conditions are considered (Zeghici et al., 2015). These parameters ensure a focused selection of potential sites, aligning closely with HT-ATES's technical demands.

However, the analysis's current scope primarily centers on subsurface characteristics, omitting surface factors that could significantly influence site feasibility. Integrating geographic and environmental considerations through GIS-based methodologies could enrich the MCDA framework. This addition would enable a more comprehensive assessment, accounting for potential land use conflicts, proximity to infrastructure, and environmental protection areas, strengthening the decision-making process for HT-ATES site selection.

Our study's exploration into heat storage capabilities through the MCDA aligns with and diverges from existing literature. Drawing parallels with Jello et al. (2022) and Stricker et al. (2020), we reaffirm the crucial role of subsurface characteristics like permeability and temperature in site viability for heat storage. However, unlike Stricker et al.’s regional analysis within the Upper Rhine Graben, our approach applies a broader, criteria-based MCDA framework. This difference showcases our method's adaptability, bridging the gap between overarching selection frameworks and specific geological assessments.

The differentiation between the Bekes Formation's wide range of results, as highlighted in our P10, P50, and P90 results, underscores the subtle understanding our MCDA framework brings to evaluating potential HT-ATES sites. Bekes Formation, with its consistently high scores and high contrast in the results, emerges as a prime candidate, notably in the P10 scenario, due to its favorable well characteristics. The contrast aligns with the broader criteria emphasized by Jello et al. and also diverges by providing a clear, quantifiable assessment of site suitability across a national landscape.

The MCDA-AHP approach significantly enhances the speed and accuracy of simultaneously evaluating multiple oil fields for various heat storage purposes, setting it apart from traditional methods. Traditionally, potential zone detection relies on time-consuming expert evaluations. However, while efficient, the grid-based analysis used in MCDA-AHP could be further refined. Currently utilizing equidimensional rectangles for grid creation, adopting a grid method that mirrors the reservoir's structural nuances would yield a more precise evaluation. Implementing such an overlapping grid requires accessible structural data, acknowledging the need for detailed geological insights to optimize the MCDA framework's efficacy.

The enhanced MCDA-AHP methodology significantly advances the evaluation of oil fields for various applications, including HT-ATES, LT-ATES, electricity generation, and CO2 storage. This approach outperforms traditional methods reliant on expert analysis by streamlining the process, enabling swift, precise identification of viable zones across various fields.

Moving forward, a detailed examination of these potential locations through advanced modeling and simulation, particularly transient modeling, is crucial. This step is vital for assessing the practical viability of these sites for diverse pore utilization purposes. Improving the gridding process to reflect the intricate structure of reservoirs better will further refine the accuracy of site evaluations. These advancements are poised to inform strategic decisions in energy storage development and policy, setting a foundation for future research and practical applications in the energy sector.

Future research should explore integrating machine learning and artificial intelligence to refine the site selection process for pore utilization. By harnessing these technologies, future studies can overcome current MCDA limitations, enabling predictive analytics and adaptive learning

from vast datasets. This approach promises to significantly enhance decision-making accuracy, tailor solutions to specific environmental and economic contexts, and revolutionize how we approach sustainable energy storage and utilization strategies.

II.6. Conclusion

This methodology paves the way to empower decision-making regarding the repurposing of abandoned hydrocarbon fields, facilitating the transition to renewable energy storage. The study showcases the feasibility and efficiency of employing Multi-Criteria Decision Analysis (MCDA) combined with the Analytic Hierarchy Process (AHP) and geostatistics to identify optimal underground heat storage locations. By leveraging comprehensive data analysis and expert evaluations on parameters like porosity, permeability, and reservoir temperature, this research outlines a strategic pathway for the sustainable utilization of former hydrocarbon sites. Future efforts will focus on refining the MCDA-AHP framework through advanced modeling and integrating artificial intelligence to enhance the precision of sustainable energy storage solutions, marking a significant advancement in renewable energy.

CHAPTER III

Predicting Thermal Performance of Aquifer Thermal Energy Storage Systems in Depleted Clastic Hydrocarbon Reservoir via Machine Learning: Case Study from Hungary

Hawkar Ali Abdulhaq
Department of Mineralogy, Geochemistry, and Petrology, University of Szeged
hawkar.ali.abdulhaq@szte.hu

János Geiger
Department of Mineralogy, Geochemistry, and Petrology, University of Szeged
matska@geo.u-szeged.hu

István Vass
MOL Hungary, MOL Plc
isvass@mol.hu

Tivadar M. Tóth
Department of Mineralogy, Geochemistry, and Petrology, University of Szeged
mtoth@geo.u-szeged.hu

Tamás Medgyes
Geothermal Energy Applied Research Department, University of Szeged
medgyes.tamas@szetav.hu

Gábor Bozsó
Department of Geology and Paleontology, University of Szeged
bozso.gabor@geo.u-szeged.hu

Balázs Kóbor
SZETAV District Heating Company of Szeged
kobor.balazs@yahoo.com

Éva Kun
Szabályozott Tevékenységek Felügyeleti Hatósága

eva.kun@sztfh.hu

János Szanyi
Department of Geology and Paleontology, University of Szeged
szanyi@iif.u-szeged.hu

Energies
18, 2025
doi.org/10.3390/en18102642

Predicting Thermal Performance of Aquifer Thermal Energy Storage Systems in Depleted Clastic Hydrocarbon Reservoirs via Machine Learning: Case Study from Hungary

Hawkar Ali Abdulhaq, János Geiger, István Vass, Tivadar M. Tóth, Tamás Medgyes, Gábor Bozsó, Balázs Kóbor, Éva Kun, and János Szanyi

Abstract

This study explores a novel strategy to repurpose depleted clastic sediment hydrocarbon reservoirs in Hungary as High-Temperature Aquifer Thermal Energy Storage (HT-ATES) systems, incorporating machine learning to enhance system optimization. Hungary's extensive inventory of depleted fields, predominantly featuring clastic formations, presents significant potential for geothermal energy storage applications. Initially, detailed reservoir models were constructed by analyzing existing well logs and core data. Subsequent advanced numerical simulations of heat transport and groundwater flow were performed within the Bekesi Formation, concentrating on a dual-well configuration—one dedicated to hot fluid injection and extraction and the other to managing cold fluids. State-of-the-art simulation tools, including SGeMS, RockWorks, Python, MODFLOW, and GMS MT3DMS, were utilized to pinpoint optimal brine injection sites by evaluating critical parameters such as thermal conductivity, porosity, and permeability; additional core analyses filled essential gaps in thermal conductivity data. The study's central innovation lies in deploying a Random Forest algorithm to optimize thermal recovery efficiency. Data generated from comprehensive simulations across multiple wells were used to train the model, which then predicted and refined thermal performance for the remaining wells in the field. The outcomes are expected to yield precise identification of optimal injection locations, rigorous heat transport analyses, accurate estimates of storage capacities, and improved predictions of thermal recovery efficiency, thereby establishing a sustainable and data-driven methodology for converting depleted hydrocarbon reservoirs into effective thermal energy storage systems.

III.1. Introduction

The end-of-life management of oil and gas wells—including plugging, abandonment, and site remediation—poses significant economic and environmental challenges. In the U.S., median decommissioning costs average USD 76,000 but can exceed USD 1 million for deep wells. In Hungary, abandonment costs typically range from 50 million to several hundred million HUF, depending on well depth and complexity. (Raimi et al., 2021; Vass, 2025). With over 3 million inactive or orphaned wells in the U.S. alone, many pose environmental risks, including methane emissions with a global-warming potential 25–34 times that of CO_2 (IPCC, 2021; Kang et al., 2019). In Pennsylvania, legacy wells contribute an estimated 5–8% of anthropogenic methane, prompting stricter plugging regulations (Kang et al., 2019; Osundare et al., 2018). Equally important is reclaiming and stabilizing well sites for land safety and ecosystem restoration. Recent advances in high-fidelity reservoir simulation paired with supervised machine-learning surrogates—such as artificial neural networks, Random Forests, and Gaussian-process emulators—now enable rapid, data-driven screening of thousands of well candidates to pinpoint those most suitable for conversion into geothermal energy assets, effectively transforming environmental liabilities into low-carbon, revenue-generating infrastructure (Jin et al., 2022; Duplyakin et al., 2022; Rohmer et al., 2023).

Amid these challenges, data-driven revitalization of depleted hydrocarbon fields has emerged as a powerful strategy to tackle renewable-energy intermittency and advance global decarbonization goals (Duggal et al., 2022; Gayayev, 2023). Recent policy roadmaps from the International Energy Agency, and the World Economic Forum (IEA, 2017; REN21, 2019; WEF, 2021) all highlight the need for adaptable, seasonal storage solutions that bolster energy security and grid flexibility (Dincer & Rosen, 2011; Van Der Roest et al., 2021). Repurposing depleted reservoirs directly supports this agenda: their well-characterized permeability, porosity, and extensive historical datasets can be leveraged to transform environmental liabilities into low-carbon, revenue-generating assets (Green et al., 2021; Lee, 2013; Li, 2016). Diverse pore-space applications—CO_2 sequestration (Qin et al., 2023), hydrogen storage (Zhu et al., 2024), subsurface electricity generation (Duggal et al., 2022), and Aquifer Thermal Energy Storage (ATES) (Matos et al., 2019; Stricker et al., 2020)—underscore their versatility in a decarbonizing energy landscape. Crucially, advances in high-resolution reservoir modelling and supervised

machine-learning algorithms now allow rapid screening of thousands of well trajectories, prediction of storage performance under uncertainty, and optimisation of injection–production schemes—dramatically reducing time-to-deployment and de-risking investment (Khosravi et al., 2024; Liu et al., 2024).

With over 2,800 systems worldwide—mainly in the Netherlands supplying ~2.5 TWh/yr—ATES has proven technically mature and scalable (Dickinson et al., 2009; Fleuchaus et al., 2020; Kastner et al., 2017; van Heekeren & Bakema, 2015). It uses existing wells and data to store surplus heat for later use, helping balance energy supply and support renewable integration (Fleuchaus et al., 2020; Paksoy et al., 2000). High-Temperature ATES (HT-ATES), targeting fluid temperatures ≥ 90 °C, further extends storage potential. Recent studies now combine coupled reservoir simulations with supervised-learning surrogates—such as neural-network emulators (Jin et al., 2022), Gaussian-process metamodelling (Rohmer et al., 2023), and Random Forest regressors (Duplyakin et al., 2022)—to forecast system performance across diverse geological scenarios in seconds rather than hours. In the Upper Rhine Graben, 90% of surveyed depleted oil fields are HT-ATES-ready, with projected storage of up to 12 GWh/yr and ~82% recovery after ten years (Holstenkamp et al., 2017; Liu et al., 2024; Stricker et al., 2020). Field-scale pilots confirm feasibility: the Middenmeer project stored 85–90 °C heat in a 400 m aquifer, using real-time monitoring and MODFLOW/MT3DMS to manage issues like sand production (HEATSTORE, 2025; Oerlemans et al., 2022). In the U.S., geothermal battery projects in California and Texas repurpose oilfields for long-duration thermal storage, leveraging machine-learning surrogates to optimise injection and reduce testing time by over 70% (Khosravi et al., 2024; Liu et al., 2024; Zhu et al., 2024). These systems offer superior discharge duration over batteries and cut CO_2 emissions (Fleuchaus et al., 2020).

Depleted reservoirs offer vast storage potential. The Carrizo-Wilcox aquifer could store 554 TWh of heat—63 TWh as electricity (Akindipe et al., 2024)—while the Upper Rhine Graben could supply 10 TWh/yr of heat, covering much of the region's demand (Stricker et al., 2020). These capacities far exceed those of pumped hydro and grid batteries (World Energy Outlook

2021 – Analysis, 2021). Repurposing wells cuts capital costs, and integration with renewables enables low levelized costs—USD 0.11/kWh for electricity, USD 0.02/kWh for heat (Anttila, 2021; Zhu et al., 2024). Though round-trip efficiency is 40–50%, geological storage fills a key seasonal balancing role (Van Der Roest et al., 2021). Although high-temperature ATES is advancing, substantial uncertainties still cloud the design and operation of dual-well systems in heterogeneous clastic reservoirs—uncertainties that brute-force deterministic simulations cannot efficiently close. Chief among these is selecting the hot-cold well spacing that suppresses premature thermal breakthrough yet preserves a high heat-recovery factor. Thermal breakthrough can erode efficiency and jeopardise long-term viability (Bloemendal & Hartog, 2018; Sommer et al., 2013). While ATES research is growing, few studies have formally linked spacing to thermal performance in repurposed hydrocarbon fields with complex stratigraphy (Kastner et al., 2017; Pellegrini et al., 2019). Cutting-edge work now couples high-fidelity thermo-hydraulic models with machine-learning surrogates—Random-Forest, artificial-neural-network, and Gaussian-process emulators—to scan thousands of spacing scenarios, propagate geological uncertainty, and pinpoint Pareto-optimal dual-well layouts in minutes rather than days (Jin et al., 2022; Rohmer et al., 2023; Duplyakin et al., 2022). Yet these data-driven methods have not been systematically deployed in depleted clastic basins (Liu et al., 2024; Khosravi et al., 2023).

This study addresses four tightly linked questions aimed at advancing seasonal thermal-energy storage in depleted clastic reservoirs. It first pinpoints the inter-well distance in a dual-well configuration that suppresses premature thermal breakthrough while maximizing heat-recovery efficiency. It then examines how key reservoir attributes in Hungarian clastic formation—porosity, permeability, and lithological anisotropy —shift the optimal spacing and shape of the overall system performance. Next, it evaluates which alternating operating schedule—summer hot-storage/winter hot-production versus summer cold-production/winter cold-storage—delivers the greatest annual energy return. Finally, it tests whether supervised machine-learning models, calibrated on suites of coupled MODFLOW-MT3DMS simulations, can predict and optimize these design variables across the full inventory of candidate wells, thereby slashing computational time and accelerating field deployment. A preliminary version of

this study was presented as an abstract at the 16th European Geothermal PhD Days, held in Szeged, Hungary (H. Abdulhaq, 2025b).

Hungary is an exceptional testbed for high-temperature ATES because it (i) hosts a dense network of depleted oil-and-gas wells situated close to district-heating loads, (ii) lies within a moderate-to-high geothermal-gradient province that delivers initial reservoir temperatures of ≈ 70 °C, and (iii) operates under strong national directives for decarbonisation and energy-security gains (Nádor et al., 2022; J. D. Szanyi et al., 2025). Decades of well-log, core, and production records have generated a richly labelled subsurface dataset that can be mined with geostatistics and supervised learning (Topór et al., 2023). Leveraging these assets, we integrate ML with high-resolution MODFLOW/MT3DMS heat-transport simulations and Random-Forest surrogates—an approach shown to cut optimisation runtimes by an order of magnitude while preserving predictive accuracy (Jin et al., 2022; Duplyakin et al., 2022). The resulting workflow delivers both mechanistic insight and actionable design rules for dual-well HT-ATES implementation in Hungarian clastic reservoirs.

By coupling high-fidelity MODFLOW-MT3DMS heat-transport simulations with Random-Forest surrogates trained on thousands of synthetic well trajectories, we cut optimisation runtimes by more than 90 % while maintaining high predictive skill (Jin et al., 2022; Rohmer et al., 2023). Feature-importance analysis of these meta-models ($R^2 \approx 0.87$ on a hold-out set) pinpoints inter-well distance, reservoir anisotropy, and cycle length as the dominant controls on Heat-Recovery Factor—insights that translate into practical spacing rules and seasonally phased operating schedules for Hungarian clastic reservoirs (Duplyakin et al., 2022; Jin et al., 2022). Validation against independent datasets from analogous Central European basins confirms the transferability of the workflow (Topór et al., 2023). Collectively, these results provide a scalable blueprint for converting depleted hydrocarbon assets into long-duration, low-carbon thermal batteries, simultaneously reducing decommissioning liabilities and advancing national energy-security and climate-mitigation goals (REN21, 2019; WEF, 2021).

III.2. Glossary of Terms

ATES (Aquifer Thermal Energy Storage): A technology for storing and retrieving thermal energy in aquifers, enabling seasonal energy management by injecting heat in summer and recovering it during winter.

HT-ATES (High-Temperature Aquifer Thermal Energy Storage): An advanced form of ATES designed for storage and recovery of thermal energy at fluid temperatures ≥90 °C, suitable for industrial and district heating applications.

MODFLOW: A modular three-dimensional finite-difference groundwater flow model developed by the U.S. Geological Survey, widely used for simulating groundwater conditions and flows.

MT3DMS: Modular Three-Dimensional Multi-Species Transport Model, used in conjunction with MODFLOW to simulate the transport of heat, solutes, or other contaminants in groundwater systems.

Random Forest: An ensemble machine learning method based on decision trees, used for regression and classification tasks, valued for its robustness and ability to model complex relationships.

Heat-Recovery Factor (HRF): The ratio of recovered thermal energy to the initially injected energy during an ATES cycle, often used as a performance metric for system efficiency.

Thermal Breakthrough: The phenomenon where injected hot or cold fluid reaches the production well too quickly, reducing system efficiency and potentially shortening operational lifetime.

UCN File (Unformatted Concentration File): A binary output file generated by MODFLOW/MT3DMS containing spatially and temporally resolved simulation results, in this study representing temperature distributions.

Residual Heat Accumulation: The progressive build-up of stored heat in the aquifer over multiple ATES cycles, typically leading to higher thermal recovery efficiencies over time.

Surrogate Model: A fast-running, data-driven model (e.g., Random Forest, neural network) trained on outputs from complex numerical simulations to predict system behavior efficiently.

Hydrogeological Model: A numerical model simulating groundwater flow based on hydraulic and geological parameters to understand subsurface water movement and storage characteristics.

Thermal Recovery Efficiency: The percentage of injected thermal energy that can be successfully recovered during production phases in a seasonal thermal energy storage system.

Overpressure: Subsurface pressure exceeding hydrostatic pressure, often due to geological compaction, tectonic forces, or fluid generation processes, which can influence reservoir behavior.

Pannonian s.l. (sensu lato):

A stratigraphic term referring broadly to the Upper Miocene sedimentary sequences in the Pannonian Basin, including formations such as Újfalu and Zagyva.

Hot Well: A well designated for the injection of heated water during storage periods and extraction during production periods in an ATES system.

Cold Well: A complementary well used to manage temperature balance in an ATES system, typically used for extracting cooler water during storage or injecting cooler water during production, depending on the operating scheme.

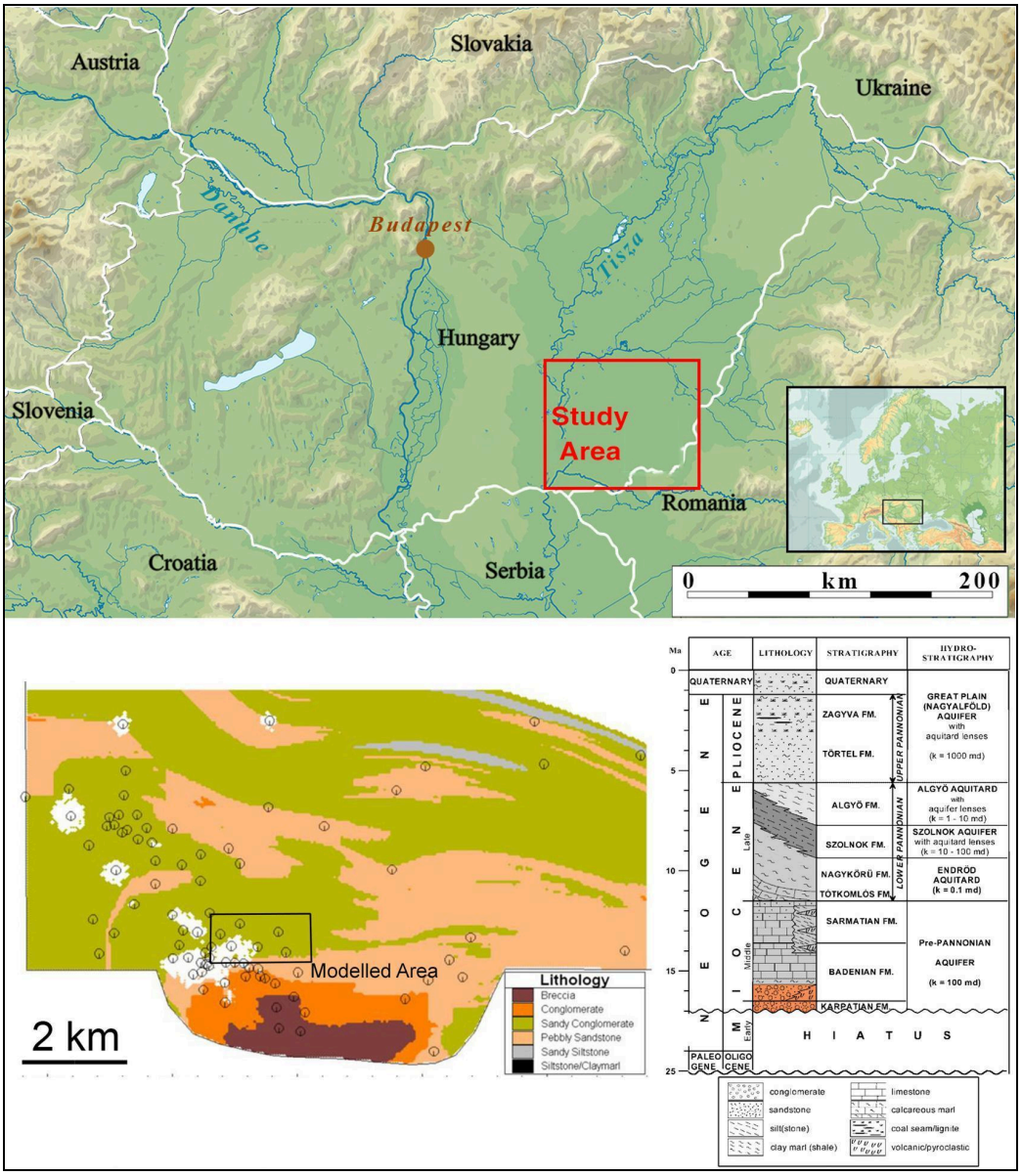
Stress Period: A defined time interval in a MODFLOW/MT3DMS simulation during which external stresses (e.g., injection, pumping) are assumed constant or follow a specified pattern.

Permutation Importance: A machine learning method used to assess the relative importance of input features by measuring the decrease in model performance when feature values are randomly shuffled.

III.3. Background and Regional

III.3.1 Geological and Hydrogeological Setting of the Békés Basin

The Pannonian Basin is a sedimentary basin located in East-Central Europe, characterized by a complex geological structure consisting of variously subsided basins and horst-like blocks. The basement primarily comprises metamorphic Paleozoic rocks, with Mesozoic carbonate formations present in some areas that can serve as good aquifers (Horváth et al., 2015). Within this larger geological context, the Békés Basin represents one of the two main depressions of the Southern Great Plain of Hungary, alongside the Makó Depression, with these two significant depressions divided by the Battonya Ridge (Juhász, 1991). The Békés Basin is particularly notable for its exceptional depth, reaching approximately 7,000 meters of post-Cretaceous sedimentary fill (USGS, 2023), making it one of the deepest sub-basins within the Pannonian Basin system. III.Figure 1 shows the location of the study area within Hungary, along with the lithological map highlighting the modelled section of the Békés Formation and a corresponding lithological cross-section.



III.Figure 1. shows the location of the study area in Hungary, the modelled section within the Békés Formation lithological map, and a lithological cross-section; modified after Abdulhaq et al. (2024).

The stratigraphic sequence of the Békés Basin follows the general pattern of the Pannonian Basin, with important variations in thickness and characteristics. At the beginning of the Lower Pannonian period, the Endrőd Marl Formation was deposited, consisting of calcareous marl and clay marl. This formation is overlain by the fine sand turbidite set of the Szolnok Formation, which reaches several hundred meters in thickness in some locations. Above the turbidites,

particularly in shallower basin areas, the hemipelagic marls are covered by the thick clayey-silty layers of the Algyő Formation with a prodelta facies (Haas, 2013). A key characteristic of this formation sequence is the extremely high overpressure below and throughout the set. The sand content of the Algyő Formation increases in areas with a shallower basement, allowing the upper part of the formation to function as a water-bearing unit in certain locations. Generally, however, the Lower Pannonian formations exhibit poor water-bearing characteristics.

The Pannonian s.l. sequence, which overlies the Lower Pannonian layers, consists primarily of the Újfalu Formation and the Zagyva Formation. The Újfalu Formation, characterized by delta front and delta plain facies, represents the most hydrogeologically significant Pannonian s.l. sediment. The Zagyva Formation features deltaic background and alluvial plain facies, with dominant sediments being bed-filling and bay-mouth bar deposits that demonstrate good water-bearing properties despite their limited horizontal dimensions. These formations are hydrodynamically connected through multiple linear erosions and overlapping (Juhász, 1991). In the Békés Basin region, the bottom of the Pannonian s.l. sequences typically lie at depths of 2,000-2,500 meters from the ground surface, with the total thickness of Pannonian s.l. sediments exceeding 2,000 meters in the Békés Basin—among the thickest in the entire Pannonian region.

III.3.2 Hydrodynamic Systems and Pressure Regimes

The Békés Basin, like the broader Carpathian basin, features two distinct flow regimes: an upper, gravity-driven flow system within the Pannonian s.l. sequences, and a deeper, overpressure-driven system within the Lower Pannonian formations, primarily affecting the finer deep-sea sediments and underlying formations (Mádl-Szőnyi & Tóth, 2009; Tóth & Almási, 2001). The overpressure in the deeper system is remarkably high, reaching up to 40 MPa above hydrostatic pressure. This extreme overpressure primarily results from tectonic compression of the formations, with additional contribution from gas formation during sediment maturation processes (Tóth and Almási, 2001).

In the Békés Basin region, pressure-depth profiles indicate that the dynamic pressure gradient exceeds the hydrostatic pressure by approximately 0.13 MPa (equivalent to about 13 m hydrostatic head) in Quaternary formations and by approximately 0.44 MPa (about 44 m hydrostatic head) in the Pannonian s.l. sequence. The Lower Pannonian sequence exhibits even

more dramatic super-hydrostatic pressure, with the dynamic pressure gradient exceeding hydrostatic pressure by more than 60 MPa. This significant pressure differential creates complex hydrodynamic conditions that must be carefully considered when designing and implementing any subsurface fluid management system, including seasonal thermal energy storage.

III.3.3 Reservoir Properties and Geothermal Potential

The Pannonian s.l. sandstone reservoirs in the Békés Basin region exhibit favorable characteristics for geothermal applications, with effective porosity values typically reaching 22-25%. The permeability of these Pannonian s.l. reservoirs, which consist of highly permeable sand layers, can reach up to 2000 mD (1.97 × 10^-12 m²), corresponding to a hydraulic conductivity of 5-10 m/day(Bálint & Szanyi, 2015; Korim, 1991; J. Szanyi et al., 2015). These values represent some of the most favorable reservoir conditions in the Hungarian geothermal context.

The consolidation state of the sandstone varies depending on depth and cementation processes. The sandstone can range from consolidated to unconsolidated, with cementation typically occurring through quartz overgrowth, calcite, or kaolin precipitation. The degree of cementation significantly influences both porosity and stability, particularly during production and injection operations. Generally, sandstone induration increases with depth as cementitious material precipitates into the pores from fluid extracted during compaction. The sand bodies are typically separated by thinner fine-grained sediments, creating a complex, heterogeneous reservoir structure (Korim, 1991; Bálint and Szanyi, 2015).

The Békés Basin is characterized by an exceptionally high geothermal gradient, approximately 50°C/km, significantly above global averages due to the relatively thin crust beneath the Pannonian Basin (Lenkey et al., 2021). This elevated thermal gradient results in reservoir temperatures around 70°C at depths of approximately 1,500-1,800 meters, making these formations particularly suitable for thermal energy storage and district heating applications. At greater depths of 2,000-2,500 meters, temperatures can reach 90-120°C, offering potential for higher-temperature applications.

The combination of favorable reservoir properties (high porosity and permeability) and outstanding geothermal conditions makes the Békés Basin exceptionally suitable for geothermal energy utilization, including seasonal heat storage applications (J. Szanyi et al., 2009). The region's depleted hydrocarbon wells, many of which penetrate these favorable Pannonian s.l. sandstone formations, present valuable opportunities for repurposing as components of thermal energy storage systems.

III.3.4 Hydrocarbon History and Well Infrastructure

The Békés Basin has a rich history of hydrocarbon exploration and production, with extensive drilling activities dating back to the mid-20th century. These activities have resulted in a substantial inventory of wells throughout the region, many of which have now reached the end of their productive lifespan as hydrocarbon producers. The basin contains significant natural gas resources, with gases produced from multiple reservoir intervals at depths ranging from 1,800 to 2,900 meters (Survey, 2023).

These Neogene sedimentary sequences overlying the basement highs have demonstrated the best hydrocarbon reservoir characteristics in the southeastern part of Hungary (Horváth et al., 2015). The extensive exploration and production history has generated valuable geological and reservoir data, including detailed information on formation properties, temperature profiles, pressure regimes, and fluid characteristics. This wealth of data provides a significant advantage for assessing the potential of these formations for thermal energy storage applications.

The existing well infrastructure, though aging, offers potential for repurposing rather than decommissioning. Many wells have been completed with telescopic designs, with casing diameters ranging from approximately 340 mm (13 3/8") at the surface to 140-178 mm (5 1/2" - 7") at reservoir depths. While some wells may require workover or partial recompletion to ensure mechanical integrity for long-term thermal storage operations, the basic infrastructure represents a valuable asset that could significantly reduce the capital costs associated with implementing seasonal heat storage systems.

III.3.5 Relevance to Seasonal Heat Storage

The hydrogeological characteristics of the Békés Basin make it particularly suitable for Aquifer Thermal Energy Storage (ATES) applications, especially High-Temperature ATES (HT-ATES) targeting temperatures up to 90°C or higher. The Pannonian s.l. sandstone formations, with their high porosity, good permeability, and favorable temperature conditions, provide an excellent medium for seasonal storage and retrieval of thermal energy.

The proximity of many depleted wells to population centers in the region creates opportunities for integrating seasonal heat storage with district heating systems, similar to successful implementations in other parts of Hungary, such as Szeged. The initial reservoir temperatures of approximately 70°C in the target formations are ideal for enhancement through additional heat input during summer months, with subsequent extraction during winter heating periods.

However, the complex pressure regimes, particularly the significant overpressure in deeper formations, present challenges that must be carefully managed. Additionally, the heterogeneous nature of the reservoir formations, with sand bodies separated by fine-grained sediments, creates potential for compartmentalization that could affect thermal breakthrough patterns between injection and production wells.

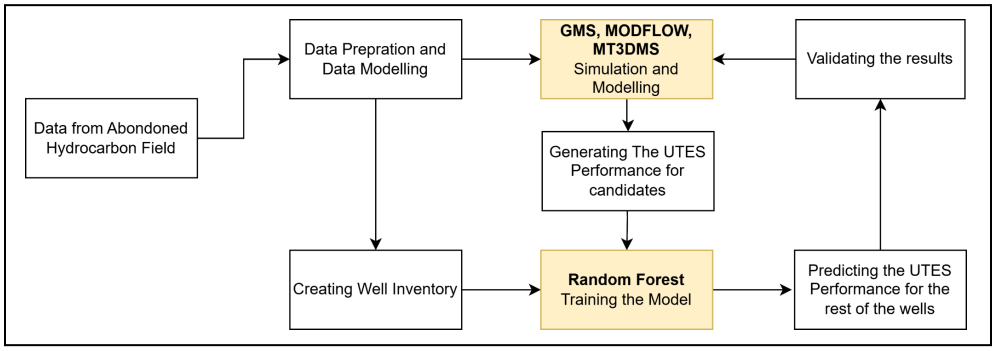
The dual-well system proposed for seasonal heat storage—with one well serving for summer hot storing/winter hot producing and another for summer cold producing/winter cold storing—must be carefully designed to account for these hydrogeological characteristics. The optimal spacing between wells must balance the need to prevent premature thermal breakthrough while maximizing energy recovery efficiency, taking into consideration the specific reservoir properties of the Békés Basin formations.

By leveraging the extensive geological knowledge, existing well infrastructure, and favorable reservoir conditions of the Békés Basin, seasonal heat storage systems can be optimized to provide sustainable, efficient thermal energy solutions while extending the productive life of otherwise abandoned hydrocarbon assets.

III.4. Materials and Methods

III.4.1. Methodological Framework

The methodological framework of this study began with the comprehensive curation and preparation of existing data from abandoned hydrocarbon fields, which involved extensive data modeling, cleaning, and refinement. A hydrogeological model was then constructed using MODFLOW (Harbaugh, 2005), providing the foundational flow and transport parameters needed for subsequent analyses. Building on this, a heat transport model was employed to simulate thermal performance in potential Underground Thermal Energy Storage (UTES) candidates. The resulting simulation outputs served as the training dataset for a Random Forest algorithm (Breiman, 2001; Pedregosa et al., 2011), which was designed to predict thermal performance in other areas. Finally, the predictions generated by the Random Forest were validated against actual simulation results. III.Figure 2 outlines the primary steps involved in this integrated approach.



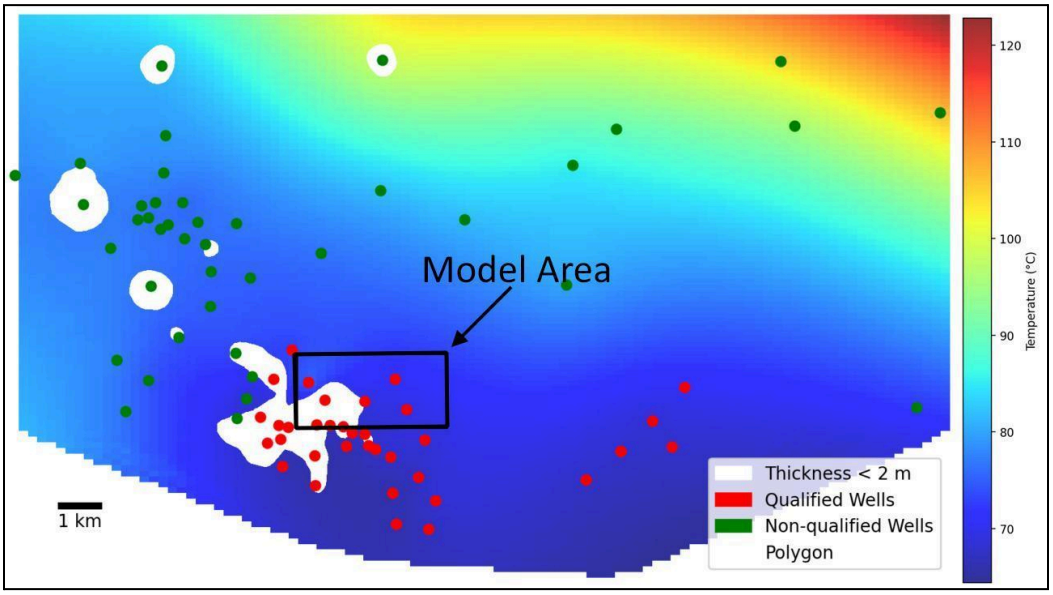
III.Figure 2. The workflow of the methodology employed in this study.

III.4.2. Data Collection and Data Preparation

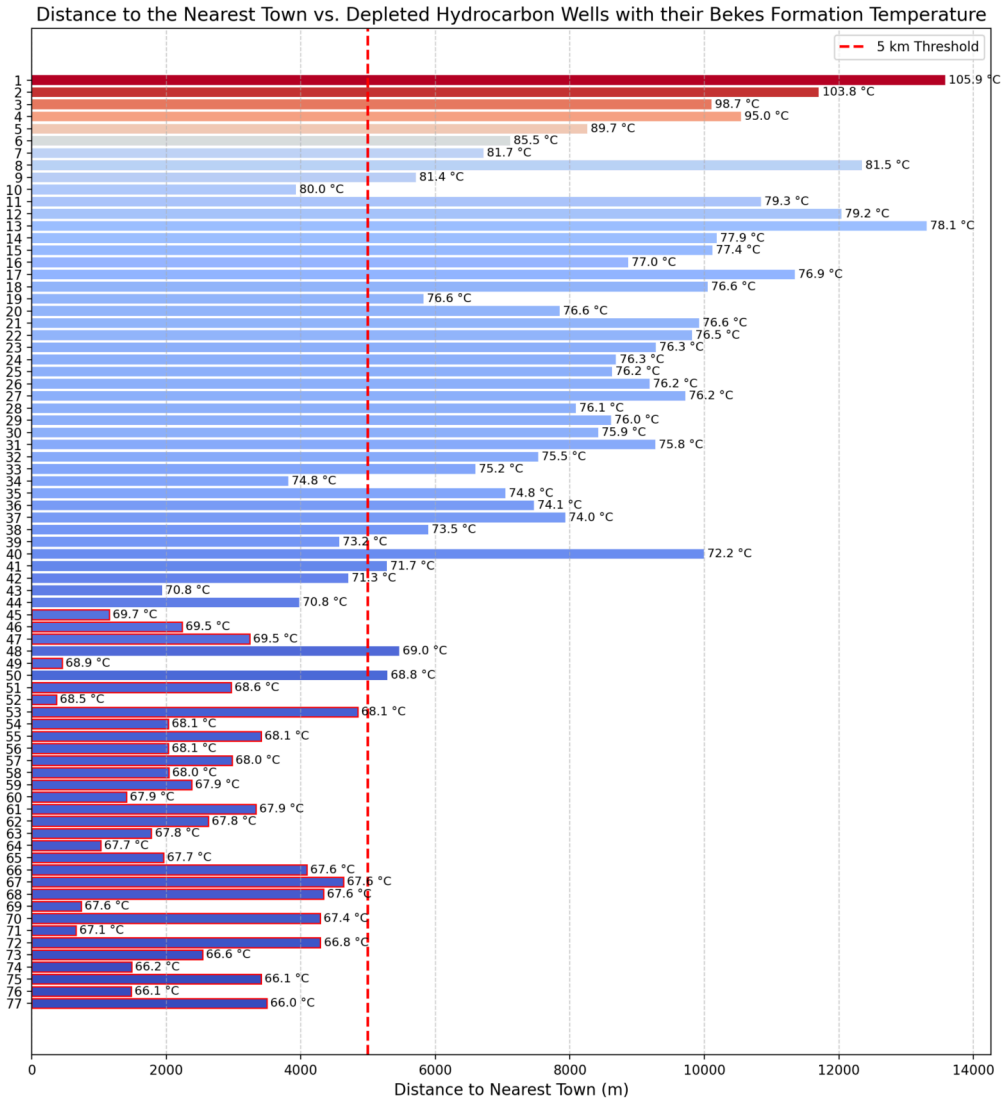
For the data collection segment of our study, we utilized well data from the southern part of the Bekes Basin in Hungary—a site formerly exploited for hydrocarbons and now recognized for its geothermal potential (Kovács & Teleki, 1994). This field encompasses two key formations: the shallow Szolnok Formation, which functions as an aquifer (Varga et al., 2019), and the deeper Bekes Formation, thereby providing a unique opportunity to evaluate heat storage capabilities. We integrated a diverse set of data—including core samples, density logs, resistivity logs, and gamma-ray measurements—using a stochastic simulation process. SGeMS was employed for the geostatistical simulation, while RockWorks facilitated the integration of simulation outcomes for gross thickness, effective porosity, and permeability. A total of 100 stochastic realizations were

generated for each grid or voxel point, with the median (Md-type estimation) calculated to represent the central tendency, effectively minimizing the impact of skewed or outlier values. For the Bekes Formation, this Md-type estimation was deemed most representative of the expected geological parameters.

In a prior study by Abdulhaq et al., (2024), the southeastern section of the study area was identified as a prime candidate for energy storage due to the Bekes Formation's average temperature of approximately 70°C, coupled with its favorable porosity and permeability characteristics. Based on these insights, a polygon delineating this promising area was selected for detailed hydrogeological and heat transport simulations. III.Figure 3 illustrates the distances from each well to the candidate town, with only those wells penetrating the Bekes Formation included in the analysis—wells that did not extend into the Bekes were excluded. The figure also displays the temperature distribution within the Bekes Formation. By applying a 5 km threshold for effective thermal transport and selecting wells with temperatures below 70°C as potential candidates for HT-ATES, only wells marked with red borders were retained for further simulation of their thermal performance (Fig 4).



III.Figure 3. shows the location of the qualified wells over the temperature grid, highlighting the modelled area. White patches indicate areas where the reservoir thickness is less than 2 meters.



III.Figure 4. shows the distribution of the wells and the boundary of the urban town that can be considered the potential candidate for the district heating beneficiary. The Bekes F. Temperature near the town is around 70° C, which makes it an ideal location to be used as a UTES site. For this reason, we narrowed down the area of interest to be modelled.

III.4.3. Hydrogeological Modelling

The groundwater flow within the Bekesi Conglomerate reservoir was modeled using MODFLOW-2000, a modular finite-difference groundwater flow modeling software developed

by the U.S. Geological Survey (Version 1.19.01, March 25, 2010). The groundwater flow model was developed using MODFLOW and discretized into three layers, 197 rows, and 400 columns with a cell size of 10 by 10 meters. In this configuration, the second layer represents the Bekes Formation. The model parameters were defined at either the cell-by-cell or grid scale, based on data-driven simulations and assumptions derived from field data and laboratory analyses. III.Table 1 summarizes the key parameters used in the MODFLOW processing model.

III.Table 1. Key Parameters for the MODFLOW Model

Parameters	**Value / Description**	**Source**
Initial Prescribed Hydraulic Head	Varies spatially	Derived from Kun et al., 2022
Horizontal Hydraulic Conductivity	Derived from permeability modelling	Estimated from well log data
Vertical Hydraulic Conductivity	Assumed as 50% of horizontal conductivity	Based on lithological assumptions
Specific Storage	0.001 m^{-1}	Literature-based estimate
Effective Porosity	Derived from porosity modelling	Estimated from well log data
Specific Yield	0.15	Literature-based estimate
Bulk Density	Calculated via gamma ray log surface simulation	Derived from natural gamma ray log simulation

III.4.4 Heat Transport Modelling

For heat transport modeling, we employed MT3DMS within the GMS framework. Following the successful initiation and simulation of the MODFLOW model, we activated the Basic Transport Package in MT3DMS by introducing the starting temperature for each cell. To simulate the heat

transport processes, we selected the advection, dispersion, source/sink mixing, and chemical reaction modules. The parameters utilized for this heat transport model are listed below.

Table 2. Input parameters and model settings used in the MT3DMS heat transport simulation

Parameter	**Value / Description**	**Justification**
Initial Temperature	Varies spatially	Derived from drill stem tests and bottom-hole temperature data
Advection Package	Third order TVD scheme Ultimate	Selected for numerical stability and accuracy
TRPT	0.1	Assumed based on typical sedimentary conditions (Gelhar et al., 1992)
TRVT	0.01	Assumed based on typical sedimentary conditions (USGS, 2022)
DMCOEF (Effective Molecular Diffusion Coefficient)	0.01 m²/day	Literature-based estimate (ModelMuse, 2024)
longitudinal Dispersivity	Varies with lithology	Based on the Rock Type Calculation and thermal conductivity of Bekes Fm. from (Vass et al., 2018)
Sorption	Linear isotherm	Common assumption for initial reactive transport modelling
Kinetic Rate Reaction	zero order reaction	Assumed for simplification of reactive processes
Preconditioner	Jacobi	Default iterative solver preconditioner

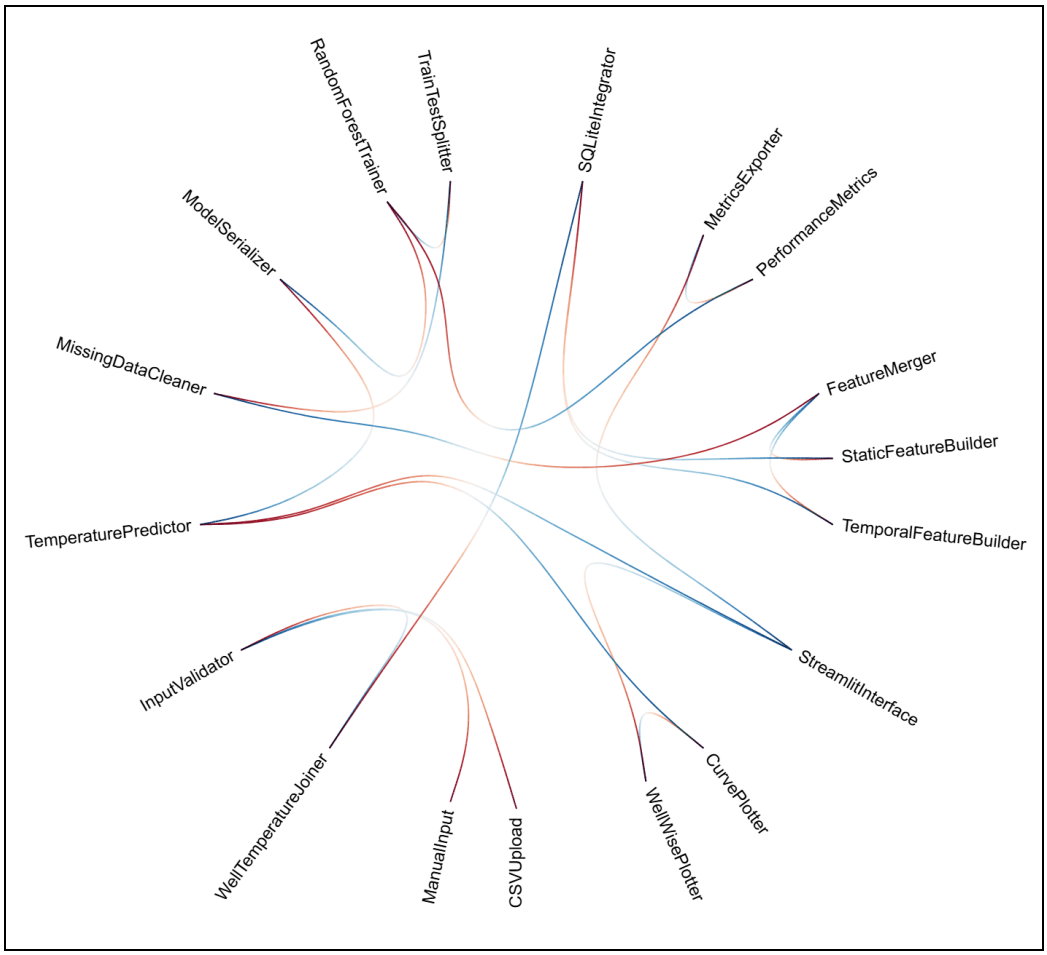
4.5 *Simulation Setting*

To replicate ATES operations, the stress periods were structured into one month of system downtime, followed by five months dedicated to heat storage, another one-month break, and five months allocated for heat production, repeating this cycle over a seven-year period. A pair of wells was selected where the distance between them is more than 500 meters: one well for hot injection in summer and subsequent hot production in winter, and the other well for cold production in summer with cold injection during winter.

III.4.6 Training Data for Machine Learning

To ensure the integrity and consistency of our training dataset, we developed an in-house data entry module using Streamlit that integrates both manual input and CSV-based uploads. This module facilitates the incorporation of critical well parameters—including well ID, name, spatial coordinates, porosity, permeability, gamma ray measurements, thickness, distance to the cold well, and initial temperature—while also allowing for the efficient addition of temporal temperature data. For the temperature data, the module supports file uploads in CSV format, validates the presence of required columns (TimeDays, WellType, Row, Col, Temperature), and links the data to the corresponding well ID via an interactive selection box (Fig. 5). The system provides real-time previews and feedback, ensuring that all data are consistently and accurately stored in the database.

The training process involves leveraging a SQLite database that consolidates data from two tables—one containing well properties and the other recording temperature measurements. First, the data from both sources are merged to form a comprehensive DataFrame, which includes attributes such as porosity, permeability, gamma ray, thickness, distance to cold wells, initial temperatures, and the time variable, while spatial reference columns (Row, Col) are retained only for reference. After cleaning the dataset by removing any rows with missing values, the data is split into training and testing subsets. A Random Forest regressor is then trained on the assembled features (excluding the spatial reference columns) to predict well temperatures, with performance evaluated through metrics including MAE, RMSE, and R^2. Finally, the trained model is serialized and saved to a predefined path, ensuring reproducibility and ease of deployment within our predictive framework.



III.Figure 5 illustrates the application of the Random Forest algorithm using a radial relational diagram.

The prediction phase leverages a pre-trained Random Forest model to forecast temperature evolution for all available wells over a specified time range. In this stage, users interact with a Streamlit-based interface where they define the starting time, ending time, and time step for the predictions. The system loads the trained model and, for each well in the database, constructs an input DataFrame populated with static parameters such as porosity, permeability, gamma ray, thickness, distance to the cold well, and initial temperature, while dynamically varying the time parameter. The model then predicts the temperature for each time step, and the results are combined into a unified dataset. To facilitate analysis, the predicted curves for individual wells are plotted on a single graph—each curve clearly labeled with its corresponding WellID—providing a comprehensive visualization of thermal performance over time. The

complete repository of the scripts is available online; however, the data content is not included (H. Abdulhaq, 2025a).

III.4.7 Model Calibration and Validation

For model calibration and validation, the hydrogeological simulation was first calibrated against historical field data—specifically, reported hydraulic head measurements—to ensure that the MODFLOW model accurately reflected the aquifer's behavior (Anderson et al., 2015). These historical observations provided key benchmarks that guided the tuning of parameters related to flow and transport processes. Concurrently, the machine learning component was iteratively calibrated using simulation outputs as training data. As the simulation yielded more refined thermal performance data, continuous retraining of the Random Forest model progressively improved its predictive accuracy (Khosravi et al., 2024). This integrated calibration strategy, leveraging both empirical historical data and dynamically generated simulation inputs, has enhanced the overall robustness and reliability of the modeling framework.

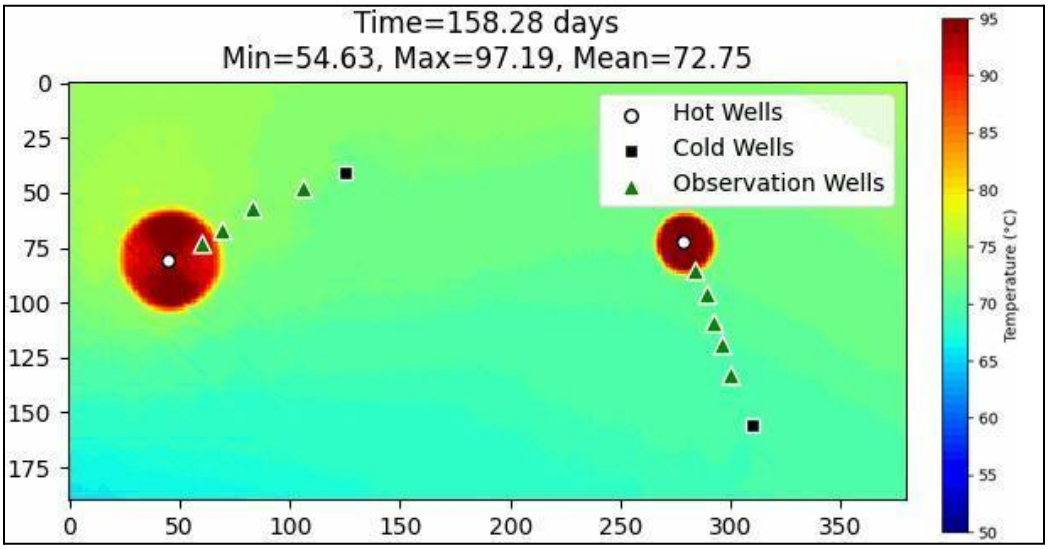
III.4.8 Sensitivity Analysis

The sensitivity analysis of the machine learning model revealed that TimeDays is by far the most influential parameter affecting temperature prediction outcomes, with a permutation importance score of 1.96, substantially higher than all other features. This result underscores the dominant role of temporal evolution in determining thermal behavior within the reservoir, reflecting the accumulation and dissipation of heat over successive cycles. Other features—such as Initial Temperature (0.034), Thickness (0.029), GammaRay (0.006), and Permeability (0.006)—exhibited comparatively minor effects, suggesting that while geological and petrophysical properties contribute to system performance, their impact is secondary to the time dimension (Fisher et al., 2019). Notably, Porosity had no measurable influence, and DistanceToColdWell showed a negligible effect (0.00007), indicating that these factors may be less critical under the modeled conditions or were inadequately represented in the available dataset. These findings provide insight into which parameters should be prioritized for accurate performance prediction and model refinement.

III.5. Results

III.5.1 Heat Simulation result

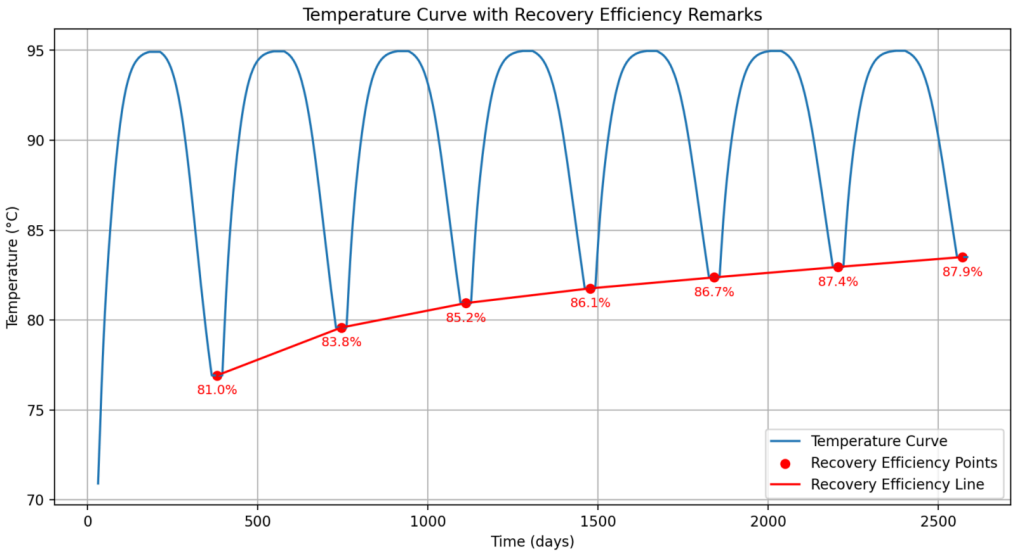
The output of the heat simulation is stored in a UCN file, which encapsulates the final temperature distribution computed by the model (Ishikawa et al., 2014). To efficiently leverage this output, we developed an in-house Python module designed to load, process, and analyze UCN files (Fig 6). This module automatically identifies hot and cold wells and allows users to designate observation wells within the simulation setting. Users can load the UCN file and select specific layers of interest, while the integrated visualization sub-module generates plots of temperature versus time for defined stress periods, providing clear insights into the thermal performance of individual wells. Additionally, an animation sub-module enables dynamic playback of temperature evolution by allowing adjustable speeds and frame skips, thereby enhancing interpretability. Furthermore, a dedicated recovery efficiency sub-module computes the thermal recovery efficiency for any simulated wells. The complete suite of module scripts is available online and can be accessed independently of the dataset (H. Abdulhaq, 2025c).



III.Figure 6 presents simulation results for a specific time frame, showing the injection of hot fluid through two wells.

In our thermal simulation studies, two sets of candidate results were obtained, both exhibiting robust injection performance with maximum temperatures consistently around 94.9°C across cycles (Fig. 7). In the first candidate, the break phases—representing reservoir conditions when

the system strikes—showed a gradual increase in maximum temperatures from 79.67°C up to 84.33°C, with thermal recovery efficiency improving from 83.92% to 88.82%. In the second candidate, although the initial break phase efficiency was lower at 81.05%, a notable enhancement was observed, with the efficiency rising to a maximum of 87.93% over repeated cycles. Key performance metrics for this candidate include a ratio of last-to-first efficiency of 1.08, a percent increase of 8.50% from the initial break phase, an average efficiency of 85.46%, and a slope of 1.04 per cycle, indicating a steady improvement with each cycle. Importantly, while these reservoir-level efficiency improvements are significant, the production efficiency during winter remains even higher than these baseline values, ensuring superior operational performance. Together, these observations demonstrate that repeated cycles of heat injection and reservoir cooling not only stabilize the thermal regime but also enhance both the inherent recovery and the actual production efficiency during winter operations.

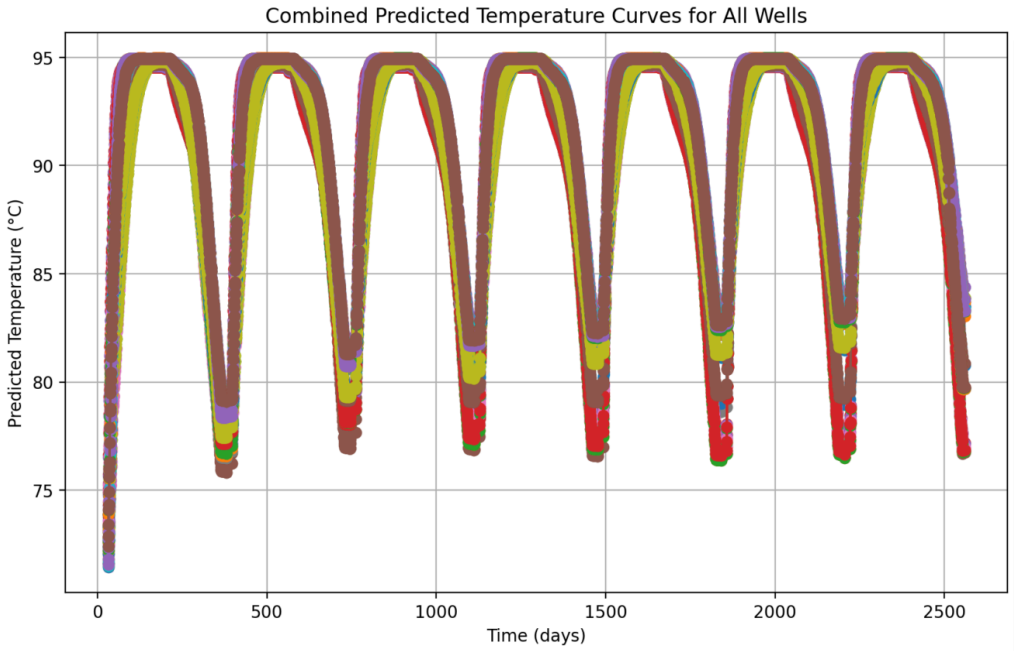


III.Figure 7 shows the thermal performance curve of the simulated hot well across repeated seasons.

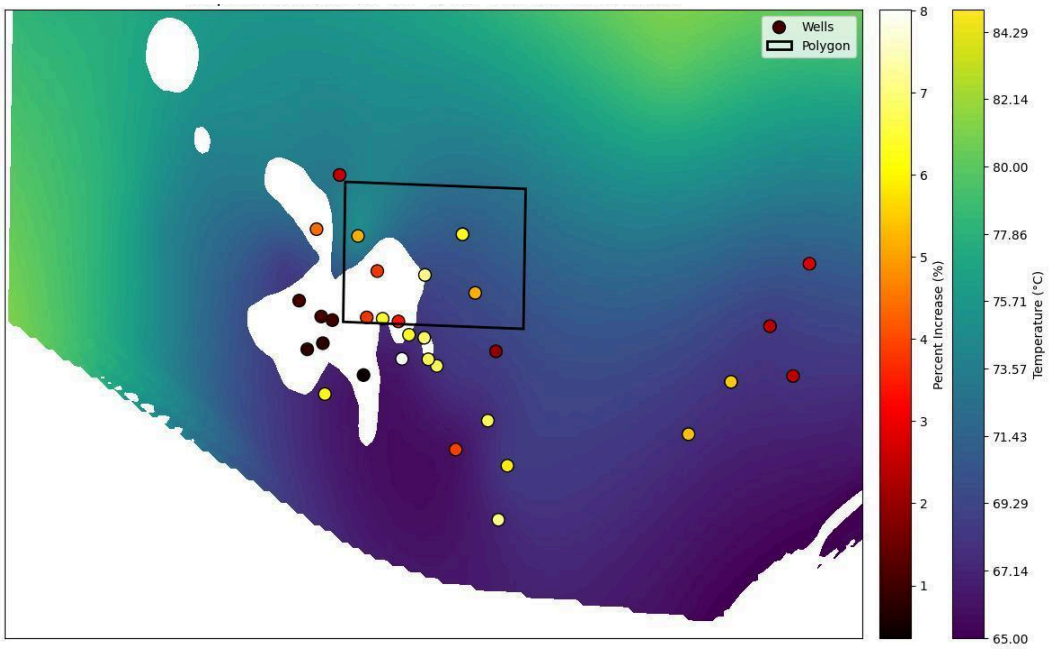
III.5.2. Machine Learning Result:

When initiating temperature performance predictions for candidate wells, we generated thermal efficiency results for multiple wells, predominantly located outside the originally modeled area. This indicates that our predictive approach is effective even beyond the immediate boundaries of

our hydrogeological model, highlighting the potential for broader applicability (Fig. 8). Analysis of these external candidates reveals initial efficiencies ranging between approximately 80% to 84%, with a consistent improvement observed through repeated injection and recovery cycles. Specifically, efficiencies showed an average of 85.46%, and maximum values reached up to 88.84%. The ratio of final-to-initial efficiency averaged around 1.05, corresponding to an overall efficiency improvement of up to 8.01%. Furthermore, the observed positive slopes (up to 1.0 per cycle) clearly illustrate that efficiency systematically increases over successive cycles. These findings underscore the reliability and robustness of our predictive methodology in forecasting thermal performance beyond the initially calibrated region, as visualized in III.Figure 9, which displays a distribution map of each well and the corresponding predicted improvement in thermal performance.



III.Figure 8 presents the thermal performance of wells located outside the modelled area.



III.Figure 9 shows the percentage increase in thermal performance for each well based on the prediction results.

III.6. Discussion

III.6.1 Alignment with Previous Studies and Theoretical Outcomes

Our simulation and machine learning outcomes align closely with prior theoretical and empirical findings related to High-Temperature Aquifer Thermal Energy Storage (HT-ATES). Previous studies indicate a typical improvement in thermal recovery efficiency with each successive injection-production cycle, attributed primarily to residual heat accumulation within the aquifer (Collignon et al., 2020; Drijver et al., 2012). Our simulation results reflect similar trends, with initial efficiencies around 80-84% gradually increasing to as high as approximately 88%. This progressive efficiency improvement closely mirrors published benchmarks from international case studies, which typically report HT-ATES efficiencies stabilizing in the range of 60–80% after several operational cycles (Collignon et al., 2020; Winterleitner et al., 2018).

The stability of our modeled temperature curves over multiple years also echoes the theoretical predictions, which suggest that a thermal equilibrium or steady state emerges after multiple cycles (Tang & Rijnaarts, 2023). However, subtle contrasts exist; for instance, our simulations may reflect idealized conditions, potentially omitting complexities such as density-driven convection or significant vertical heat migration reported in heterogeneous clastic reservoirs (Winterleitner et al., 2018). This highlights a limitation where idealized models may yield slightly optimistic efficiency and thermal stability predictions compared to more complex real-world scenarios.

III.6.2 Key Influencing Parameters

Our machine learning analysis identified time ('TimeDays') as the most influential parameter on temperature and efficiency predictions, surpassing all other geological and operational parameters by a substantial margin. Specifically, the permutation importance of time was approximately 1.96, whereas other parameters like initial temperature, thickness, gamma ray, and permeability demonstrated notably lower impacts (0.034, 0.029, 0.006, and 0.006, respectively). This finding emphasizes that temporal factors—specifically, cumulative cycle duration and residual heat buildup—predominantly govern thermal performance, consistent with the literature where residual thermal energy strongly influences long-term operational outcomes (Drijver et al., 2012).

III.6.3 Strengths and Limitations of Hydrogeological Model Calibration

Calibrating the hydrogeological model using historical head data provides several advantages, chiefly realistic and site-specific insights into subsurface dynamics. However, this method's main limitation lies in the necessity to modify boundary conditions to match historical measurements adequately, potentially introducing bias or oversimplification. Adjustments made to boundary conditions, while essential for aligning simulations with empirical data, might restrict the generalizability of the hydrogeological model to conditions significantly different from historical scenarios.

III.6.4 Enhancing Decision-Making for UTES Site Selection

Our predictive methodology substantially accelerates the decision-making process for selecting potential UTES sites by providing rapid, high-quality predictive outcomes derived from historical data and simulations. By transforming abandoned hydrocarbon reservoirs into a data-driven analytical framework, our approach reduces evaluation times and increases confidence in site assessments. This can significantly streamline site selection workflows, particularly valuable in regions with numerous abandoned wells and substantial historical datasets.

III.6.5 Implications for Scaling Geothermal Storage Projects

The successful application of this predictive approach using gamma ray logs to differentiate rock types indicates substantial potential for scaling geothermal storage projects, particularly within clastic sedimentary basins. Internationally, numerous clastic basins exist with similar sedimentological characteristics, making this methodology widely applicable. This capability facilitates rapid assessments across varied geographies, promoting efficient, scalable deployment of HT-ATES technologies globally.

III.6.6 Assumptions and Simplifications

Key assumptions and simplifications within our modeling framework include idealized boundary conditions, homogeneous or simplified geological heterogeneities, and consistent thermal properties across the model domain. Such assumptions likely influence the accuracy and predictive power of the model, potentially limiting its applicability to real-world scenarios with pronounced geological complexity or significantly variable hydrogeological conditions.

III.6.7 Uncertainties in Input Data

A major uncertainty arises from the inherent accuracy and reliability of the initial simulation data used for machine learning training. Because our predictive model relies heavily on accurate thermal simulations, any errors or oversimplifications in the initial simulations propagate through the predictions, potentially affecting reliability. This highlights the critical importance of accurate and comprehensive simulation input data.

III.6.8 Recommendations and Future Work

Future studies should focus on enhancing simulation fidelity, incorporating more detailed heterogeneity, and conducting diverse scenario sampling to bolster the robustness of the machine learning model. Additionally, exploring advanced physics-informed machine learning techniques could significantly improve predictive accuracy and generalizability. Extending simulation runs, including more diverse geological and operational parameters, and conducting further validation with independent field data would further strengthen confidence in our model predictions and their applicability to broader contexts.

III.7. Conclusion

This study set out to advance the design and deployment of high-temperature Aquifer Thermal Energy Storage (HT-ATES) in depleted clastic reservoirs through four interlinked research questions. First, by coupling high-resolution MODFLOW-MT3DMS heat-transport simulations with an in-house Python toolkit, we identified the optimal inter-well spacing that minimizes premature thermal breakthrough while maximizing cumulative heat-recovery efficiency. Our two candidate well pairs consistently achieved peak injection temperatures near 94.9 °C and demonstrated steady efficiency gains—up to 8.5 % over repeated cycles—when spaced to balance thermal front propagation and lateral heat recharge.

Second, we quantified how key reservoir properties in Hungarian clastic formations modulate this optimal spacing and overall system performance. Sensitivity analysis revealed that porosity and permeability variations shift thermal breakthrough timing and adjust peak recovery efficiencies by several percentage points, while anisotropy primarily affects the shape of the thermal plume. These insights enable tailoring well spacing to site-specific hydraulic and thermal heterogeneities.

Third, comparison of seasonal operating schedules confirmed that the conventional hot-storage/winter production cycle yields marginally higher annual energy returns than the inverse (cold-production/winter storage), owing to more effective residual-heat carryover. Specifically, winter production efficiencies exceeded baseline recovery values, underscoring the value of aligning storage and demand cycles to ambient temperature differentials.

Finally, our Random Forest surrogate models—trained on several coupled simulations—proved capable of predicting thermal recovery efficiency across a broad inventory of candidate wells outside the original calibration domain. Surrogate predictions achieved average accuracies within 2–3 % of full numerical simulations, with efficiency improvements up to 8 % over multiple cycles. This machine-learning workflow accelerates design optimization by an order of magnitude, enabling rapid screening of sites and operational schedules.

Collectively, these findings deliver a data-driven framework for HT-ATES implementation in depleted clastic reservoirs: mechanistic insights into spacing and scheduling trade-offs, parameter-specific performance adjustments, and a scalable surrogate modeling approach for design optimization. Future work should extend this framework by incorporating more complex heterogeneities, exploring physics-informed learning techniques, and validating predictions against field pilot data to further refine site-selection criteria and operational guidelines.

CHAPTER IV

A Data-Driven ML Model for Sand Channel Prediction from Well Logs for UTES Site Optimization and Thermal Breakthrough Prevention: Hungary Case Study

Hawkar Ali Abdulhaq
Department of Mineralogy, Geochemistry, and Petrology, University of Szeged
hawkar.ali.abdulhaq@szte.hu

János Geiger
Department of Mineralogy, Geochemistry, and Petrology, University of Szeged
matska@geo.u-szeged.hu

István Vass
MOL Hungary, MOL Plc
isvass@mol.hu

Tivadar M. Tóth
Department of Mineralogy, Geochemistry, and Petrology, University of Szeged
mtoth@geo.u-szeged.hu

Gábor Bozsó
Department of Geology and Paleontology, University of Szeged
bozso.gabor@geo.u-szeged.hu

János Szanyi
Department of Geology and Paleontology, University of Szeged
szanyi@iif.u-szeged.hu

Energies
18, 2025
doi.org/10.3390/en18164230

A Data-Driven ML Model for Sand Channel Prediction from Well Logs for UTES Site Optimization and Thermal Breakthrough Prevention: Hungary Case Study

Hawkar Ali Abdulhaq, János Geiger, István Vass, Tivadar M. Tóth, Gábor Bozsó, János Szanyi

Abstract

This study presents a data-driven approach to predict the three-dimensional distribution of sand-rich channels in hydrocarbon reservoirs using well log data, aiming to optimize site selection for Underground Thermal Energy Storage (UTES) and manage hot and cold well pairs effectively. Leveraging detailed petrophysical datasets from 128 hydrocarbon exploration wells within the Szolnok Formation in southern Hungary, the developed machine-learning workflow—combining XGBoost regression and spatial residual correction—accurately delineated permeable channel systems suitable for thermal energy injection and extraction. The model achieved robust predictive performance (R^2 = 0.92; RMSE = 0.24), and correlation analyses confirmed significant relationships between predicted channels and sand content and shale content. Clearly identified high-permeability channel zones facilitated strategic well placement, significantly reducing the risk of premature thermal breakthrough and enhancing the reliability and efficiency of UTES operations.

IV.1 . Introduction

The accelerating global shift toward renewable energy has significantly heightened the demand for effective and scalable energy storage solutions, which are essential for addressing renewable intermittency, ensuring grid flexibility, and providing reliable seasonal energy storage (Duggal et al., 2022; REN21, 2019; WEF, 2021) . While short-term battery storage technologies have matured significantly, their inherent limitations in scalability and cost-effectiveness for long-duration or seasonal storage underscore the need for alternative, large-scale solutions. Geological energy storage, particularly Underground Thermal Energy Storage (UTES), is increasingly recognized as a promising alternative due to its capacity for extensive seasonal storage and its minimal environmental footprint (Bloemendal & Hartog, 2018; Fleuchaus et al., 2020; Van Der Roest et al., 2021) .

UTES systems, notably Aquifer Thermal Energy Storage (ATES), are now well-established worldwide, with more than 2,800 operational installations—most prominently in the Netherlands, where they collectively provide approximately 2.5 TWh of thermal storage annually (Fleuchaus et al., 2020; Kastner et al., 2017; van Heekeren & Bakema, 2015). Recent technological advances, including High-Temperature ATES (HT-ATES) systems that operate at temperatures above 90 °C, further extend the potential of UTES by enabling efficient integration with renewable energy sources, industrial heat supply, and district heating networks. These developments substantially expand both the practical impact and the relevance of UTES technologies (Oerlemans, n.d.; Stricker et al., 2020). Recent studies, such as Stricker et al. (Stricker et al., 2024), examined mechanical risks like surface deformation during HT-ATES in depleted hydrocarbon reservoirs; however, explicit modeling of subsurface permeability channels to prevent thermal breakthrough remains comparatively underdeveloped.

A critical determinant of UTES system effectiveness is the strategic placement of injection and production wells, which directly influences the efficiency, reliability, and lifespan of thermal storage operations. One of the most significant operational challenges is thermal breakthrough, where injected thermal fluids prematurely reach production wells, dramatically reducing heat-storage efficiency and overall economic viability (Bloemendal & Hartog, 2018; Sommer et al., 2013). Preventing or mitigating thermal breakthrough fundamentally depends on precise geological characterization—specifically, understanding the spatial distribution and geometry of permeability, porosity, and, most importantly, channelized sand bodies and low-permeability barriers that dictate subsurface fluid-flow pathways (Dickinson et al., 2009; Fleuchaus et al., 2020).

To robustly characterize subsurface formations, the Flow Zone Indicator (FZI) has emerged as a valuable petrophysical metric, integrating permeability and porosity into a single spatially explicit indicator that describes fluid-flow capacity within geological formations (Amaefule et al., 1993; Astsauri et al., 2024; Jin et al., 2022; Man et al., 2021) (Amaefule et al., 1993; Jin et al., 2022). However, in complex fluvial and turbidite systems such as the Szolnok Sandstone, it is the explicit geometry, connectivity, and separation of sand-rich channels and intervening low-permeability units that most directly control fluid-flow behavior, thermal breakthrough, and

UTES performance (Cao et al., 2024; The Importance of Physiochemical Processes in Decarbonisation Technology Applications Utilizing the Subsurface: A Review | Earth Science, Systems and Society, n.d.). While FZI has been historically employed within hydrocarbon reservoir analysis for enhanced oil recovery, and has proven highly effective at delineating subsurface heterogeneity into distinct permeability zones, a more targeted modeling of channel systems offers even clearer identification of both optimal regions for fluid injection and production, as well as the natural low-permeability barriers that separate them (Duplyakin et al., 2022; A. (刘奥迪) Liu et al., 2024; Rohmer et al., 2023).

Despite its established utility in reservoir characterization, the application of explicit channel system modeling, alongside or beyond FZI analysis, for UTES system optimization—identifying injection-production well sites and assessing barrier quality for thermal breakthrough prevention—remains comparatively underdeveloped (Stricker et al., 2020). Focusing on the three-dimensional architecture of channelized sands and their separation by low-FZI barriers represents a critical advance in targeting, designing, and optimizing UTES operations. Addressing this knowledge gap presents a valuable opportunity for advancing UTES technology by precisely targeting and optimizing subsurface geological conditions (Jin et al., 2022; Pellegrini et al., 2019; Rangel-Jurado et al., 2023; Stricker et al., 2020).

This research builds upon FZI analysis by developing a structured, data-driven methodology that explicitly models and predicts the three-dimensional architecture of sand-rich channel systems in the Szolnok Sandstone, with direct relevance for optimizing UTES site selection and operational planning (H. Abdulhaq, 2025b). Leveraging comprehensive petrophysical datasets—including detailed records of permeability, porosity, lithological characteristics, and historical well logs—the study employs advanced geostatistical techniques such as Ordinary Kriging and Sequential Gaussian Simulation (SGS)(Varouchakis, 2021). Machine-learning models, widely utilized in petroleum exploration for various complex predictive tasks (Q. Li et al., 2025; Qingchao et al., 2025) [REF1, REF2], complement these traditional methods. These methods enable the generation of high-resolution, three-dimensional maps that reflect generalized heterogeneity; further, it clearly delineates the geometry and connectivity of channelized sands and their separation by low-permeability zones (Rangel-Jurado et al., 2023). While FZI-based mapping remains a valuable tool, this approach advances reservoir characterization by resolving

the spatial distribution of channel bodies that most directly govern UTES performance (H. Abdulhaq, 2025b; Duplyakin et al., 2022; Rohmer et al., 2023).

By clearly identifying the spatial extent and continuity of sand channels suitable for injection and production wells, and mapping the intervening low-FZI barriers, this approach directly addresses the critical challenge of thermal breakthrough (H. Abdulhaq, 2025b). The explicit assessment of channel geometry, barrier thickness, continuity, and spatial arrangement allows for targeted well placement that leverages natural geological compartmentalization, enhancing thermal storage efficiency and long-term system sustainability (Challenges and Opportunities for Aquifer Thermal Energy Storage (ATES) in EU Energy Transition Efforts—An Overview, n.d.). Quantitative criteria such as barrier effectiveness and spatial distribution are thus evaluated to support robust thermal storage operations (Baird et al., n.d.; Bloemendal & Hartog, 2018; Jin et al., 2022).

Additionally, this study applies a data-driven machine learning workflow, integrating XGBoost regression with advanced spatial feature engineering and thin-plate spline correction, to explicitly predict the three-dimensional presence, continuity, and separation of sand-channel systems within the reservoir (Khosravi et al., 2024). By interpolating well log-derived properties using Inverse Distance Weighting (IDW) and correcting residuals with Radial Basis Function (RBF) interpolation, the approach captures both local heterogeneity and large-scale channel connectivity (Duplyakin et al., 2022). This method enables the efficient evaluation of numerous well-placement scenarios, providing rigorous uncertainty quantification and detailed sensitivity analysis—key for robust strategic planning and rapid deployment of UTES systems in channelized settings (Rohmer et al., 2023). Such a hybrid modeling strategy substantially reduces computational complexity while maximizing spatial accuracy and interpretability, consistent with best practices in the field (Khosravi et al., 2024; A. (刘奥迪) Liu et al., 2024; Rohmer et al., 2023).

Hungary represents an ideal context for this investigation due to its dense network of depleted hydrocarbon wells, moderate-to-high geothermal gradients (around 70°C initial reservoir

temperatures), and close proximity to urban centers with significant district heating demands. The presence of extensive Szolnok Sandstone Formation with its channel systems and comprehensive historical well data further enhances Hungary's suitability for channel-focused UTES modeling (IV.Figure 1). Moreover, Hungary's commitment to decarbonization and energy security, combined with decades of subsurface data from hydrocarbon exploration, makes it an exceptional testbed for detailed geostatistical analysis and machine-learning-driven reservoir optimization (Nádor, 2022; J. D. Szanyi et al., 2025; Topór et al., 2023).

This research systematically characterizes the 3D distribution and connectivity of sand channels—supported by FZI—to identify optimal injection and production well sites for UTES. For the first time, we apply a hybrid machine-learning approach integrating XGBoost regression with spatial residual correction to explicitly predict channel architecture and barrier distribution specifically within the Szolnok Formation. Emphasizing the mapping of high-permeability channels and low-FZI barriers, the approach mitigates thermal breakthrough risks, enhancing thermal isolation. Integrating numerical simulations with supervised machine-learning techniques optimizes system design under geological uncertainty, providing a reproducible framework for scalable, efficient UTES systems aligned with global decarbonization and energy resilience goals (REN21, 2019; WEF, 2021).

IV.2. Glossary of Terms

ATES (Aquifer Thermal Energy Storage): A technology for storing and retrieving thermal energy in aquifers, enabling seasonal energy management by injecting heat in summer and recovering it during winter.

HT-ATES (High-Temperature Aquifer Thermal Energy Storage): An advanced form of ATES designed for storage and recovery of thermal energy at fluid temperatures ≥90 °C, suitable for industrial and district heating applications.

Channelization: The development of laterally and vertically connected, sand-rich channel bodies within sedimentary reservoirs, which strongly control subsurface fluid-flow pathways and are critical for optimizing UTES performance and preventing thermal breakthrough.

Flow Zone Indicator (FZI): A petrophysical metric combining permeability and porosity into a single value that characterizes the fluid-flow capacity of a rock interval, widely used for identifying reservoir heterogeneity and mapping high-permeability channels and low-permeability barriers.

XGBoost: An advanced, scalable machine learning algorithm based on gradient-boosted decision trees, used for regression and classification tasks in this study to predict the three-dimensional distribution of sand channel systems from well log data.

IDW (Inverse Distance Weighting): A spatial interpolation technique that estimates unknown values at unsampled points based on the values of nearby observations, with closer points weighted more heavily; used here for spatial feature engineering of well log data.

RBF (Radial Basis Function) Interpolation: A mathematical method for smoothly interpolating values across multidimensional space; in this study, thin-plate spline RBF correction was used to improve the spatial accuracy of model residuals.

Random Forest: An ensemble machine learning method based on decision trees, used for regression and classification; valued for its robustness and ability to model complex relationships.

Gaussian Process Emulator: A statistical machine learning model used to create fast, probabilistic surrogate models of complex systems, enabling uncertainty quantification and efficient scenario evaluation.

Residual Heat Accumulation: The progressive build-up of stored heat in an aquifer over multiple ATES cycles, typically leading to higher thermal recovery efficiencies over time.

Thermal Breakthrough: The phenomenon where injected hot or cold fluid reaches the production well too quickly, reducing system efficiency and potentially shortening operational lifetime; a key risk in UTES that can be mitigated by natural or engineered barriers.

Hydrogeological Model: A numerical or conceptual model simulating groundwater flow and heat transport based on geological and hydraulic properties, used to inform site selection and UTES system design.

Hot Well: A well used to inject and store heated water during UTES charging phases and extract thermal energy during production phases.

Cold Well: A complementary well used for injecting cooler water during UTES extraction or extracting cold water during charging, essential for managing thermal gradients and system balance.

Permutation Importance: A machine learning method for quantifying the importance of individual input features by measuring the decrease in model performance when feature values are randomly shuffled.

Stress Period: A defined time interval in a hydrogeological or thermal simulation during which external stresses (e.g., injection, extraction) are assumed constant or follow a specified pattern.

MODFLOW: A modular three-dimensional finite-difference groundwater flow model developed by the U.S. Geological Survey, widely used for simulating groundwater and heat transport in aquifers.

MT3DMS: Modular Three-Dimensional Multi-Species Transport Model, used with MODFLOW to simulate the movement of heat, solutes, or contaminants in groundwater systems.

Overpressure: Subsurface pressure exceeding hydrostatic pressure, often due to geological compaction, tectonic forces, or fluid generation, affecting reservoir management.

Pannonian s.l. (sensu lato): A stratigraphic term referring broadly to the Upper Miocene sedimentary sequences in the Pannonian Basin, including formations such as Újfalu and Zagyva.

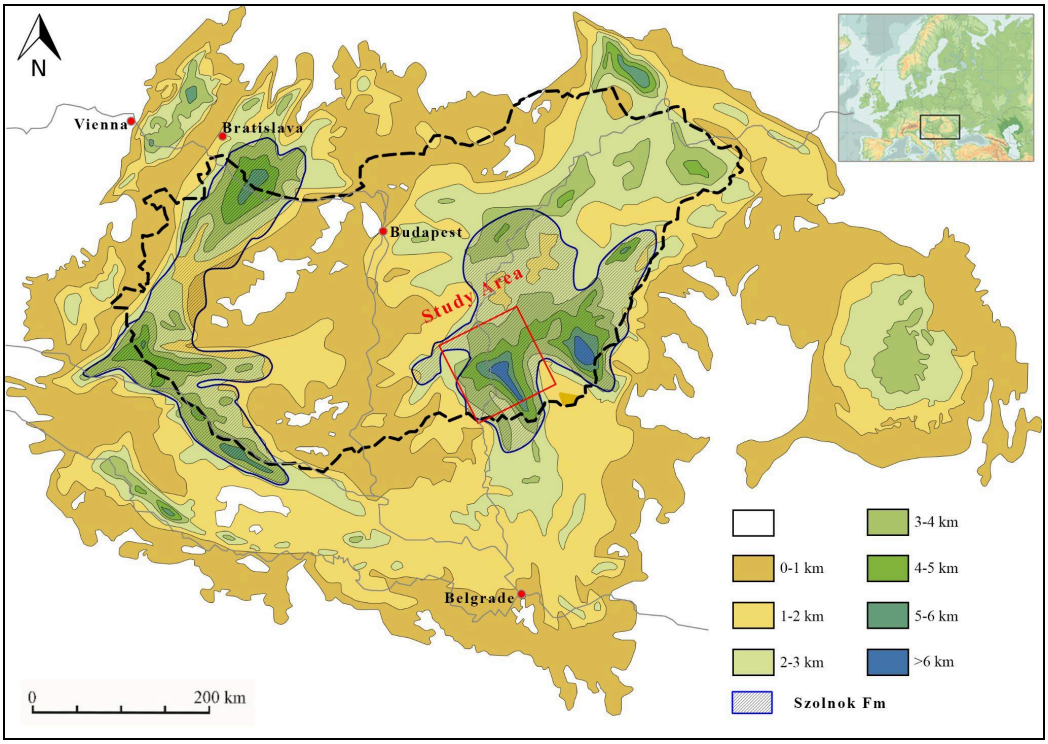
IV.3. Background and Regional

IV.3.1 Geological and Hydrogeological Setting

The Pannonian Basin, located in East-Central Europe, features complex geological structures, including subsided sedimentary basins separated by horst-like uplifted blocks (H. Abdulhaq, 2025b; H. A. Abdulhaq et al., 2024). The basement primarily comprises Paleozoic metamorphic rocks, occasionally overlain by Mesozoic carbonate formations, which act as significant aquifers in various locations, providing favorable conditions for geothermal exploration and thermal

storage projects (Njeru et al., 2024; J. D. Szanyi et al., 2025; Topór et al., 2023). Within this intricate geological setting, the Szolnok Formation in southern Hungary represents an essential Lower Pannonian stratigraphic unit, widely recognized for its considerable reservoir potential and suitability for geothermal applications (Njeru et al., 2024; Vass et al., 2018).

Compositionally, the Szolnok Sandstone is predominantly fine-grained and well-sorted, composed primarily of quartz, feldspar, and lithic fragments derived from mixed siliciclastic and carbonate source rocks, with variable carbonate and clay cementation influencing reservoir quality (Sendula et al., 2025). The Szolnok Formation predominantly consists of fine-grained sand turbidites interbedded with layers of siltstone and claystone, deposited within turbidite fan environments that developed partly in prodelta regions and partly in deeper basin settings of the extensive Lake Pannon basin during the Late Miocene (Lower Pannonian) (H. Abdulhaq, 2025b; Nagymarosy & Báldi-Beke, 1993; Varga et al., 2019). This formation exhibits substantial variations in thickness, typically ranging from a few meters up to about 50 meters, reflecting complex depositional processes involving channelized turbidity flows and lobe systems sourced from adjacent paleo-deltas (Sztanó et al., 2015). These depositional patterns significantly influence reservoir heterogeneity, directly impacting the effectiveness and optimization of geothermal energy storage operations, particularly for Aquifer Thermal Energy Storage (ATES) and Underground Thermal Energy Storage (UTES) systems (Fleuchaus et al., 2018; Kim et al., 2010; Project Thermo-Hydro-Mechanical Description – Geothermal Energy and Geofluids, n.d.; Stricker et al., 2020).



IV.Figure 1 Thickness map of the Pannonian sediments in Hungary, highlighting the extent of the Szolnok Formation (outlined by purple lines). The green rectangle delineates the specific area investigated in this study (B 2204-B: Pannonian Basin Province, Central Europe (Province 4808)—Petroleum Geology, Total Petroleum Systems, and Petroleum Resource Assessment, n.d.; Geofizika, 2011; Nádor & Lenkey, 2023).

Typically, the Szolnok Formation directly overlies thick marls and claystones of the Endrőd Formation, serving as regional seals, and is subsequently capped by prodeltaic silty clay sediments of the Algyő Formation (Haas, 2013). This stratigraphic stacking results in substantial overpressure conditions, significantly influencing fluid dynamics and reservoir management strategies (Mádl-Szőnyi & Tóth, 2009; J. Szanyi et al., 2015; Tóth & Almási, 2001; Vass et al., 2018).

IV.3.2 Hydrodynamic Systems and Pressure Regimes

The hydrodynamic environment of the Szolnok Formation is dominated by distinct flow regimes: shallow intervals influenced by gravity-driven groundwater systems and deeper intervals characterized by significant overpressure (H. Abdulhaq, 2025b; J. D. Szanyi et al., 2025). This

overpressure is primarily generated by tectonic compaction, sediment loading, and gas formation from organic maturation processes, creating challenging conditions for subsurface fluid management. Pressure-depth relationships in the Szolnok Formation reveal significant super-hydrostatic pressures, often exceeding hydrostatic gradients by several tens of megapascals (Mádl-Szőnyi & Tóth, 2009; Tóth & Almási, 2001).

IV.3.3 Reservoir Properties and Geothermal Potential

Reservoir characterization of the Szolnok Formation indicates favorable geological properties suitable for thermal energy storage (H. A. Abdulhaq et al., 2024). Effective porosity within the formation typically ranges between approximately 15% to 23%, with permeabilities generally varying from 10 to 270 millidarcies (mD) (Elabd et al., 2022; Permeability and Mineralogy of the Újfalu Formation, Hungary, from Production Tests and Experimental Rock Characterization: Implications for Geothermal Heat Projects, n.d.; Vass et al., 2018). Hydraulic conductivity, derived from permeability estimates, extends from approximately 11 to 2500 mmD, supporting considerable fluid flow capabilities and reservoir connectivity, essential for successful geothermal operations (H. A. Abdulhaq et al., 2024; Jello et al., 2022; Shi et al., 2023).

Detailed reservoir quality analysis highlights a moderate sand content typically between 50% and 70%, representing channelized sand bodies interbedded with finer-grained sediments such as siltstones and claystones, which act as natural intra-reservoir barriers or baffles (Bridger & Allen, 2010; Varga et al., 2019). The formation's Flow Zone Indicator (FZI) typically exceeds values of 5.0, further emphasizing the excellent fluid flow potential for geothermal injection and production operations (Jin et al., 2022; Rohmer et al., 2023).

Geothermal gradients in the region of the Szolnok Formation are notably high, averaging approximately 50°C/km, reflecting relatively thin crustal conditions beneath the Pannonian Basin (Elbalawy et al., 2025; Tóth & Almási, 2001). Consequently, reservoir temperatures are commonly around 70°C at economically viable depths of approximately 1,500–1,800 meters, with deeper intervals at around 2,000–2,500 meters potentially reaching temperatures as high as 90–120°C (Pellegrini et al., 2019; J. D. Szanyi et al., 2025; Vass et al., 2018). Such conditions strongly favor geothermal energy storage and district heating applications (Fleuchaus et al., 2020; Pellegrini et al., 2019).

IV.3.4 Hydrocarbon History and Well Infrastructure

The Szolnok Formation has experienced extensive hydrocarbon exploration and production since its discovery in the mid-20th century (H. A. Abdulhaq et al., 2024; Czauner & Mádl-Szőnyi, 2013; Tari et al., 2023). Numerous wells penetrate this formation, providing extensive geological, reservoir, and operational data critical for subsequent geothermal assessments and thermal energy storage implementation (Fleuchaus et al., 2020; Réveillère et al., 2013). Typically, wells penetrating the Szolnok Formation in the studied region have employed telescopic casing designs ranging from approximately 340 mm diameter at the surface to 140–178 mm at reservoir depths (Vass et al., 2018; Zeghici et al., 2015).

The existing extensive well infrastructure, predominantly from past hydrocarbon production, offers significant economic advantages for repurposing wells into geothermal and seasonal thermal storage wells (Fleuchaus et al., 2018; Pellegrini et al., 2019; J. Szanyi et al., 2023). Although these wells vary in mechanical integrity and may require selective workover or recompletion, their presence substantially reduces initial investment requirements, making the implementation of geothermal projects more financially viable (H. A. Abdulhaq et al., 2024; Wesselink et al., 2018).

IV.3.5 Relevance to Seasonal Heat Storage

The geological, hydrodynamic, and geothermal conditions of the Szolnok Formation in southern Hungary make it particularly suitable for Aquifer Thermal Energy Storage (ATES), including advanced High-Temperature ATES (HT-ATES) applications (H. Abdulhaq, 2025b). Its combination of favorable porosity (15–17%), significant permeability (up to 270 mD), and excellent hydraulic conductivity (up to 2500 mmD) supports efficient fluid movement and heat exchange essential for effective thermal storage and recovery (H. Abdulhaq, 2025b; Topór et al., 2023).

The geological, hydrodynamic, and geothermal conditions of the Szolnok Formation in southern Hungary make it particularly suitable for Aquifer Thermal Energy Storage (ATES), including advanced High-Temperature ATES (HT-ATES) applications (Fleuchaus et al., 2018; Wesselink et al., 2018). Its combination of favorable porosity (15–17%), significant permeability (up to 270

mD), and excellent hydraulic conductivity (up to 2500 mmD) supports efficient fluid movement and heat exchange essential for effective thermal storage and recovery (Elabd et al., 2022; Shi et al., 2023; Vass et al., 2018).

The formation's inherent stratigraphic heterogeneity, with sand-rich channel systems surrounded by finer-grained sediments, provides naturally compartmentalized conditions (H. Abdulhaq, 2025b; Bridger & Allen, 2010; Varga et al., 2019). These heterogeneous reservoir characteristics create ideal scenarios for strategically positioning injection and extraction wells, leveraging high-permeability channelized sands for thermal storage, while low-permeability regions serve as natural barriers preventing premature thermal breakthrough (Dickinson et al., 2009; Sommer et al., 2013). Optimal well-spacing and operational strategies, therefore, rely on detailed geological characterization and modeling of reservoir heterogeneity.

Furthermore, the proximity of depleted hydrocarbon wells to district heating infrastructures in southern Hungary presents advantageous conditions for efficiently integrating thermal storage solutions into existing heating networks (Holstenkamp et al., 2017; Pellegrini et al., 2019). However, careful management of the formation's substantial overpressure conditions and stratigraphic complexities remains critical to ensure sustainable and efficient thermal energy storage operations (Czauner & Mádl-Szőnyi, 2013; Zeghici et al., 2015).

Collectively, leveraging the Szolnok Formation's geological insights, favorable reservoir conditions, existing well infrastructure, and high geothermal potential positions it as an outstanding candidate for developing sustainable, scalable seasonal thermal energy storage systems (Réveillère et al., 2013; Topór et al., 2023).

IV.4. Materials and Methods

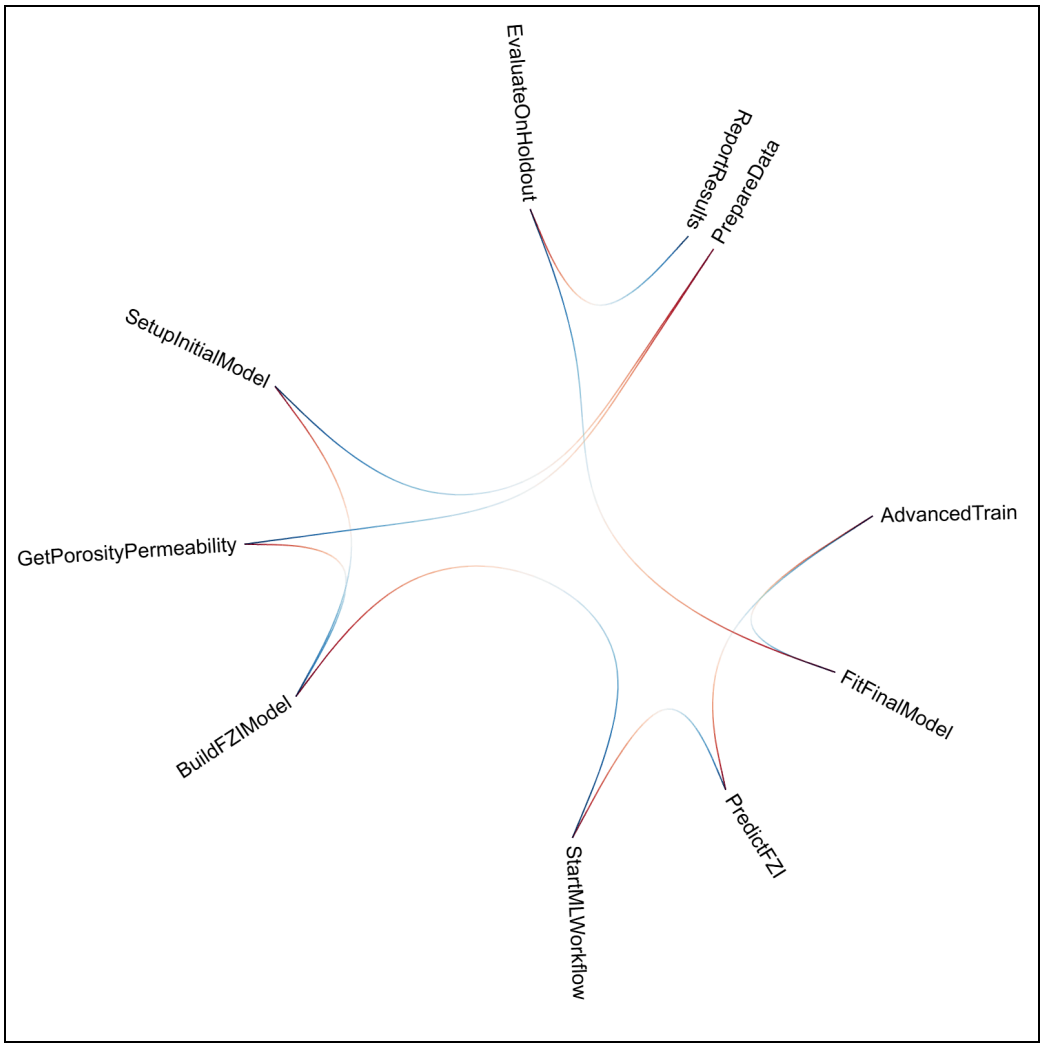
IV.4.1. Methodological Framework

The methodological framework of this study began with comprehensive data curation and preparation, involving extensive modeling, cleaning, and refinement of existing hydrocarbon field datasets. Using RockWorks software, an initial Flow Zone Index (FZI) model was constructed based on reliable porosity and permeability estimates derived from detailed reservoir characterizations (Egbele et al., 2005; Zecchin & Catuneanu, 2015). Subsequently, a

machine-learning workflow was developed, leveraging well-log data to predict spatial distributions of the FZI (IV.Figure 2). XGBoost regression was chosen for its robustness and efficiency in capturing complex geological patterns, while spatial continuity and residual smoothing were effectively addressed through Inverse Distance Weighting (IDW) interpolation and Radial Basis Function (RBF) interpolation using a thin-plate spline kernel, defined as:

$\mathrm{TPS}(r)=r^2\ln(r)$, where $r=\sqrt{(x-xi)^2+(y-yi)^2}$

The predicted FZI outputs from the initial machine-learning model served as input training data for a HistGradientBoostingRegressor (Astsauri et al., 2024; Kassab et al., 2024; Man et al., 2021), which was evaluated on a withheld dataset representing approximately 30% of the study area. The final predictive model demonstrated strong performance, achieving an R^2 of approximately 0.92, a mean absolute error (MAE) of 0.17, and a root mean squared error (RMSE) of 0.24 (Astsauri et al., 2024).

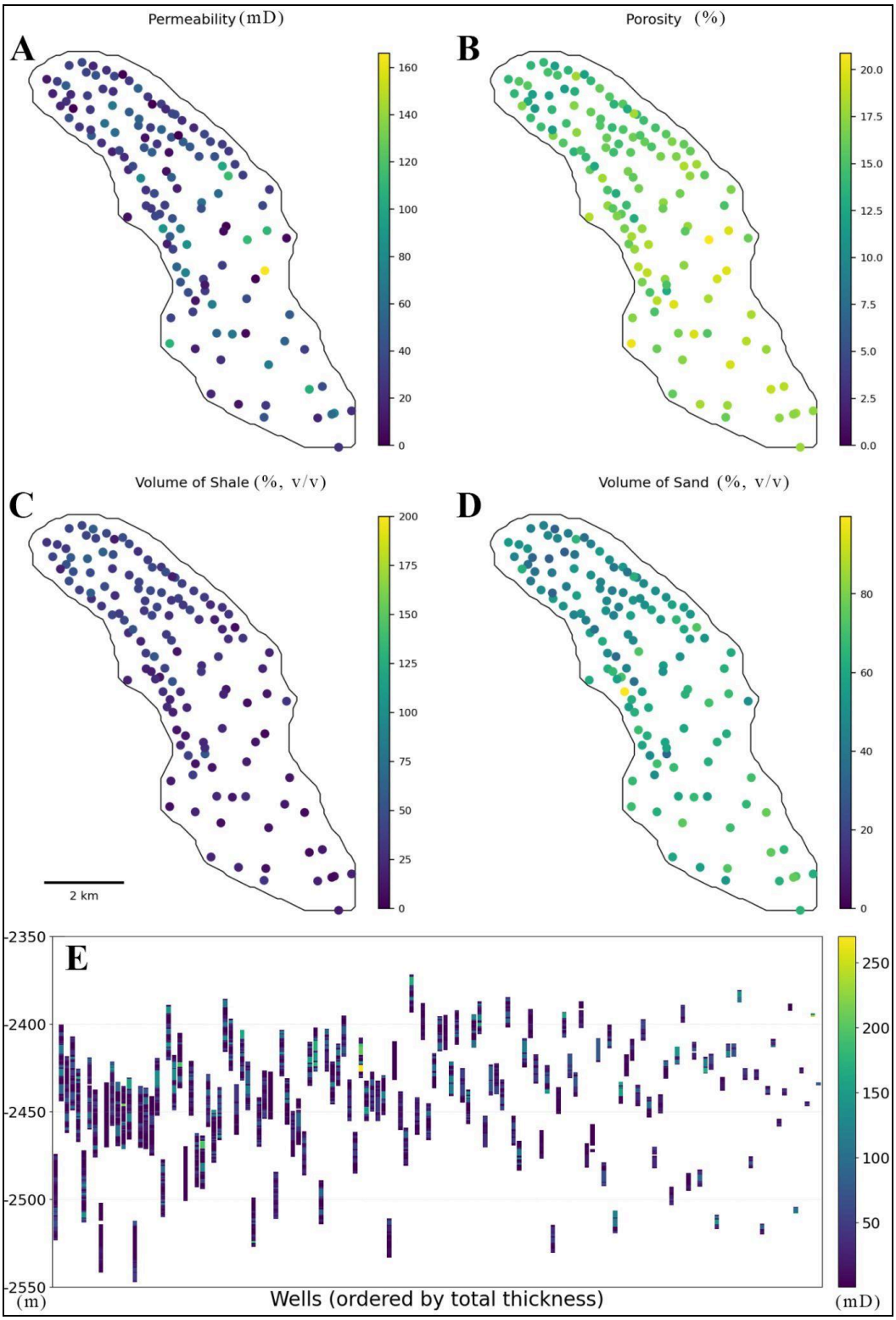


IV.Figure 2 Radial flow chart illustrating the workflow implemented in this research. The red arrows represent data outputs or processed information, while the blue arrows denote data inputs or feedback pathways.

IV.4.1. Data Collection and Data Preparation

The data collection phase of this study involved systematically compiling detailed subsurface geological and petrophysical data from the Szolnok Formation in southern Hungary, utilizing comprehensive historical records from 128 hydrocarbon exploration and production wells. The primary datasets included well-log measurements, core analyses, and stratigraphic interpretations, providing key information such as formation tops and bases, thickness variations,

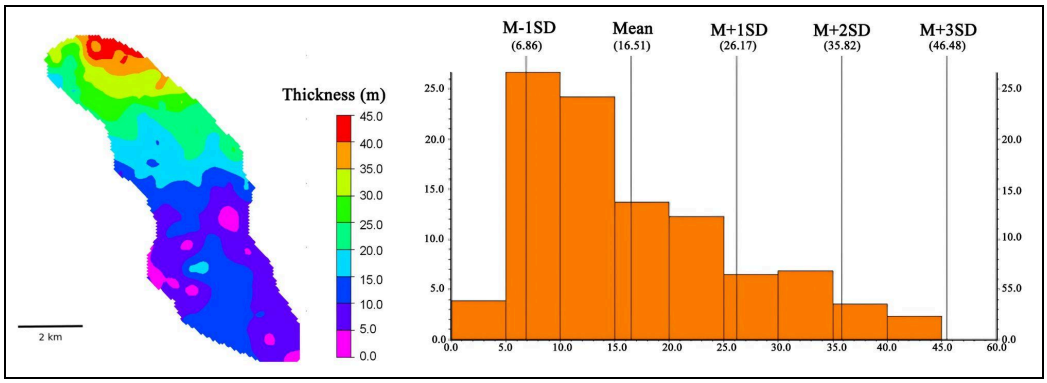
sand content distributions, effective porosity measurements, permeability data, and hydraulic conductivity values (IV.Figure 3) (Abdallah et al., 2016).



IV.Figure 3 Reservoir characterization plots illustrating key petrophysical properties from well-log data within the Szolnok Formation: (A) Permeability distribution (mD); (B) Effective porosity (%); (C) Volume of shale (VSHA; %, v/v), defined as the percentage of rock volume composed of shale; (D) Volume of sand (VSND; %, v/v); (E) Cross-sectional representation of permeability distribution (mD) along wells. The vertical axis shows depth (m, Total Vertical Depth), with wells sorted horizontally from thickest (left) to thinnest (right). Note: These plots represent petrophysical properties derived from well-log data, not physical core images.

IV.4.3. Data Modelling

The data collection process for this study involved acquiring detailed subsurface data from the Szolnok Formation in southern Hungary, leveraging information derived from 128 hydrocarbon exploration and production wells. The dataset primarily comprised stratigraphic and reservoir property data, including top and base surfaces, thickness, sand content, effective porosity, effective pore volume, permeability, and reservoir quality indices (IV.Figure 4) (H. A. Abdulhaq et al., 2024). Stratigraphic surfaces (top and base) were estimated using kriging with a trend model (Czauner et al., 2023), while sand content, porosity, and permeability were modeled using both Inverse Distance Weighted (IDW) interpolation—specifically optimized to emphasize lateral reservoir continuity—and Ordinary Kriging techniques to enhance spatial prediction reliability (Malvić et al., 2020). The effective porosity within the formation ranged between 15% and 22%, while permeability values ranged from 10 to 270 mD, corresponding to hydraulic conductivity values of approximately 11 to 2500 mmD (Sandstone Heterogeneity - 3-D Reservoir Characterization, n.d.). Reservoir Quality Index (RQI) and Flow Zone Index (FZI) models were constructed by combining these spatially interpolated porosity and permeability distributions, yielding average FZI values exceeding 5.0, indicative of excellent fluid-flow capabilities within the channelized sand bodies (Cheng et al., 2022; Xing et al., 2022). Additionally, a dedicated uncertainty analysis for sand content in thin beds (<2 m) using Monte Carlo simulations indicated moderate prediction uncertainties, with an average 95% confidence interval width of 14.79, highlighting areas where predictions should be interpreted cautiously due to limited data or geological complexity.



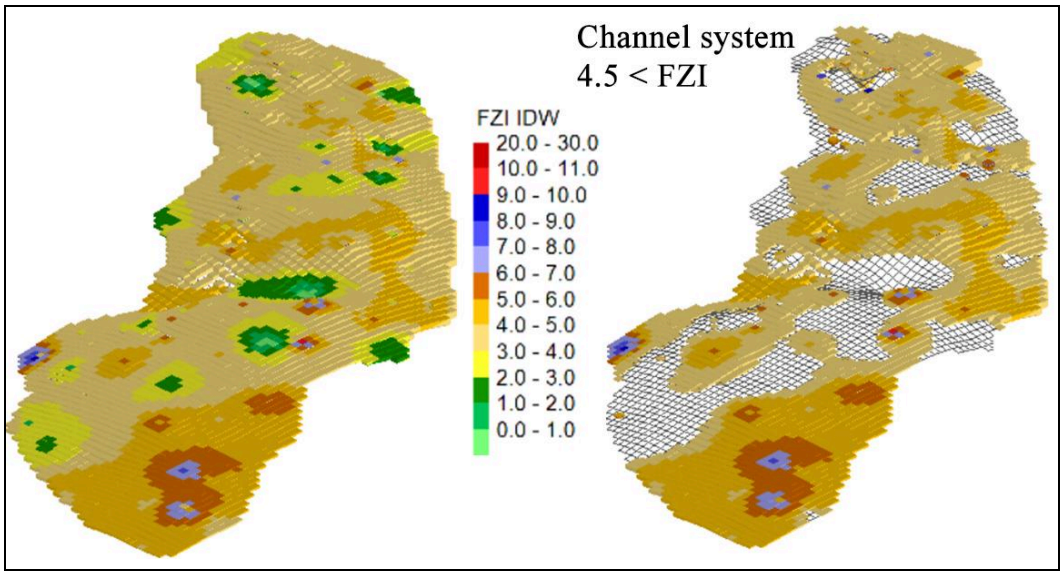
IV.Figure 4 Spatial and statistical representation of the Szolnok Formation thickness; the left panel displays a map illustrating reservoir thickness distribution across the study area, while the right panel shows the frequency distribution of thickness values.

4.4 *Flow Zone Index Modelling*

The Flow Zone Index (FZI) modelling for the Szolnok Formation involved integrating spatially interpolated porosity (ϕ) and permeability (k) datasets to quantitatively characterize and visualize reservoir fluid-flow capacities (Man et al., 2021). FZI was calculated using the standard reservoir engineering equation (Abbaszadeh et al., 1996; Amaefule et al., 1993; Corbett et al., 2003; Svirsky et al., 2004):

$$FZI = \frac{RQI}{NPI} = \frac{0.0314\sqrt{\frac{k}{\phi}}}{\frac{\phi}{1-\phi}}$$

where RQI is the Reservoir Quality Index, k is permeability (mD), and ϕ is effective porosity (expressed as a percentage). Spatial interpolation was conducted using both Inverse Distance Weighted (IDW) methods, optimized specifically to emphasize lateral continuity, and Ordinary Kriging supported by directional variogram analyses (IV.Figure 5) (Baisad et al., 2023; Otchere et al., 2021). These geostatistical approaches were selected to capture regional trends and localized heterogeneities effectively, ensuring accurate and reliable delineation of subsurface flow zones.



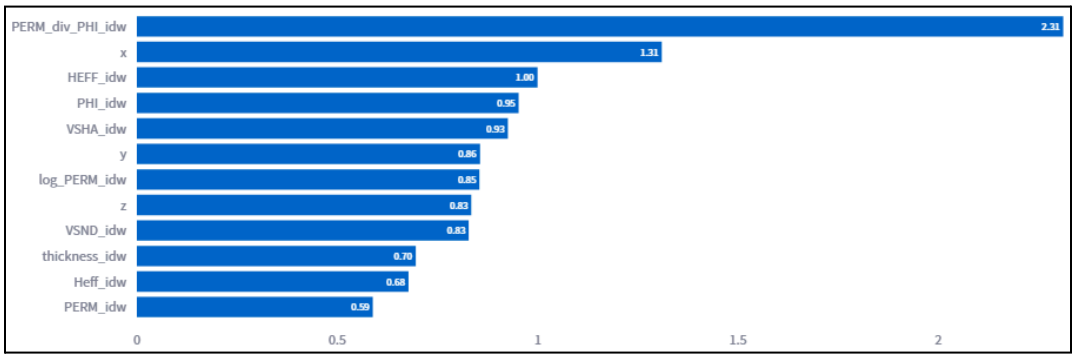
IV.Figure 5 Spatial distribution maps of modelled Flow Zone Index (FZI) values; the left map represents the predicted FZI values using Inverse Distance Weighted (IDW) interpolation, while the right map highlights regions with FZI greater than 4.5, indicative of potential sand-rich channel systems.

The resulting FZI models, generated independently using both IDW and Kriging methods, consistently indicated highly favorable fluid-flow conditions, particularly within channelized sand bodies. Specifically, the IDW-derived FZI values exceeded 4.5, while Kriging-derived estimates indicated slightly higher values, generally exceeding 5.0. The consistently high FZI values reflect excellent reservoir connectivity and permeability within these turbiditic channel systems, highlighting their suitability for fluid injection and extraction operations (Al-Dousari et al., 2022). By clearly delineating areas of optimal flow capacity, the FZI modelling results provided critical guidance for strategic site selection and well placement decisions aimed at maximizing thermal storage efficiency and preventing premature thermal breakthrough (Fadel et al., 2022).

IV.4.5 Machine Learning Setting

The training dataset was meticulously prepared for modeling the Flow Zone Index (FZI) in three-dimensional space, integrating rich petrophysical features and rigorous spatial preprocessing. Initially, interval-level data extracted from multiple well logs were processed to

obtain geospatially accurate samples characterized by depth (z) and crucial petrophysical attributes, including permeability (PERM), porosity (PHI), shale volume (VSHA), sandstone volume (VSND), effective thickness (HEFF), and total thickness (thickness) (IV.Figure 6).



IV.Figure 6 Feature importance based on the XGBoost regression model, ranked by gain, highlighting key petrophysical attributes and derived features that most significantly contributed to predicting sand-channel distributions within the Szolnok Formation.

An essential preprocessing step was the targeted removal of outliers from both well intervals and grid node observations, employing robust statistical techniques: Z-score (±3.0) and Interquartile Range (IQR; 1.5 × IQR) filtering. This careful cleaning significantly improved data quality, stability, and model accuracy (Iglesias et al., 2025).

Geometrically, the spatial domain was defined and optionally refined through the application of a buffer around core wells, effectively isolating a central region for focused analysis. Within this spatially constrained area, a systematic 80/20 train-test split was performed on the available 3D grid nodes, ensuring a reliable validation framework for predictive performance (Ahmadi, 2025; Shen et al., 2018).

Subsequent feature engineering leveraged Inverse Distance Weighting (IDW) interpolation to generate robust, spatially contextualized attributes for each grid node. Specifically, petrophysical features from the nearest well intervals were interpolated, forming IDW-derived features (PERM_idw, PHI_idw, VSHA_idw, VSND_idw, HEFF_idw, and thickness_idw). Further feature augmentation included logarithmic transformations (e.g., log_PERM_idw) to normalize skewed

data distributions and derived ratio features (e.g., permeability-to-porosity ratio PERM_div_PHI_idw) to capture meaningful petrophysical relationships (Nieminen, 2022).

A powerful hybrid modeling approach was employed, combining XGBoost regression (trained via 5-fold cross-validation) with a spatially explicit residual correction using Thin-Plate Spline interpolation (Zhang et al., 2024). This integrated methodology significantly enhanced prediction accuracy and spatial consistency.

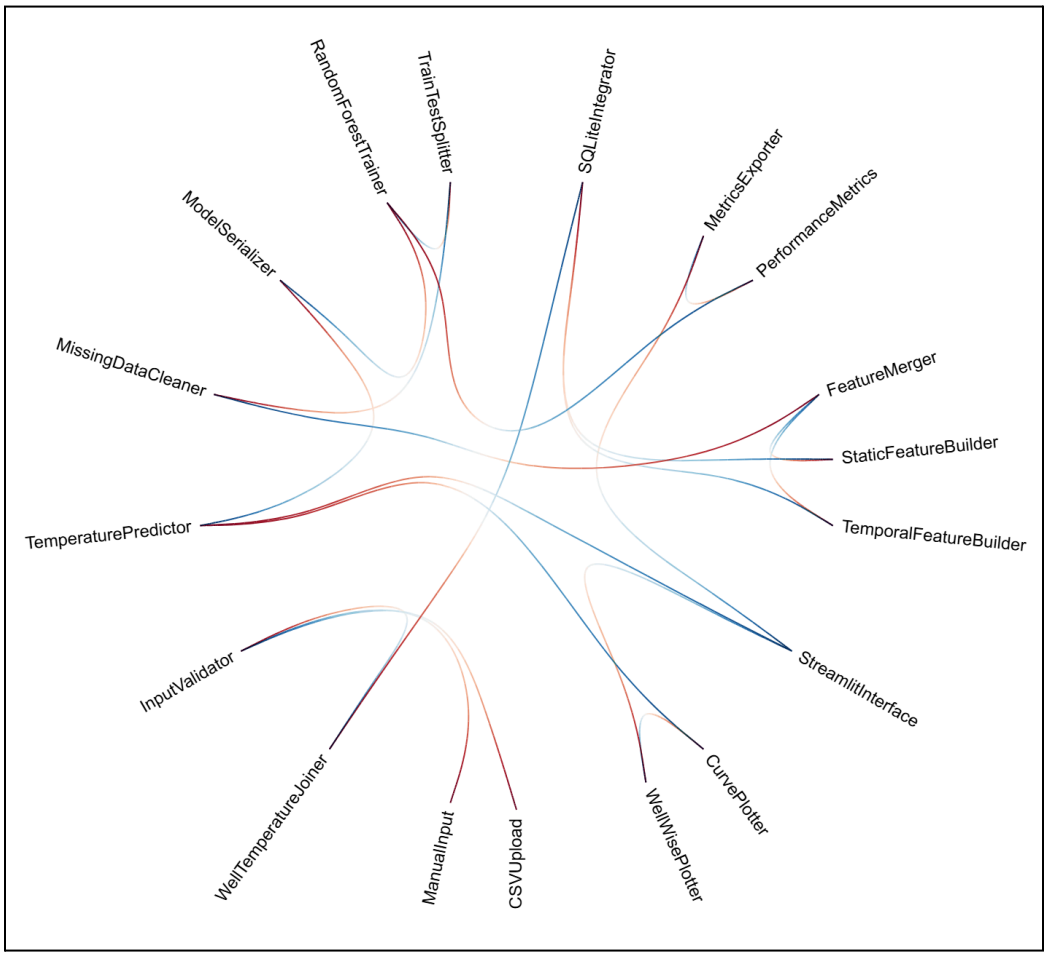
Comprehensive quality control and diagnostics were systematically generated, including scatter plots, residual analyses, depth slices, feature importance rankings, and geospatial prediction maps, facilitating robust model validation and interpretability (Ahmadi, 2025). Collectively, these engineered and spatially informed features provided a robust foundation for accurate, reliable, and interpretable three-dimensional FZI predictions (Ge et al., 2011).

IV.4.6 Machine Learning Process

The machine learning process implemented in this study integrated rigorous data preprocessing, spatial feature engineering, and robust predictive modeling to accurately forecast three-dimensional sand-channel systems indicated by the Flow Zone Index (FZI) (Cheng et al., 2022). Initially, interval-level petrophysical well data, including permeability and porosity (input features), were systematically loaded and filtered for outliers using robust Z-score statistical methods to ensure the reliability of model inputs (Badem et al., 2024). Permeability values in the training set ranged from approximately 0.01 mD to 980 mD, while porosity values spanned from approximately 5% to 28%. Subsequently, the spatial modeling domain was defined using a polygonal boundary shapefile, optionally refined by buffering to retain core, most representative reservoir areas (H. Abdulhaq, 2025b). Advanced spatial interpolation techniques, specifically Inverse Distance Weighting (IDW), generated spatially informed predictive features at each 3-D grid node, further augmented by composite attributes and log-transformed petrophysical features designed to capture geological relationships (Das, 2025). The complete workflow, including each critical step from data loading to final predictions and evaluation, is illustrated in IV.Figure 7.

In the second stage, an optimized XGBoost regression model was rigorously trained using an 80/20 train-test split of the grid nodes, and validated via a comprehensive 5-fold cross-validation

to assess predictive accuracy (Ahmadi, 2025). The output predictions of the model were subsequently improved by applying spatially explicit residual corrections using thin-plate spline interpolation (RBF), significantly enhancing spatial consistency (Zhao, 2021). Comprehensive evaluation metrics (R^2, MAE, RMSE) were calculated from the independent test set, confirming the model's high accuracy (test $R^2 = 0.92$) (Mellal et al., 2023). Ultimately, this integrated modeling workflow yielded robust, accurate, and interpretable three-dimensional predictions of sand-channel distribution, directly informing reservoir characterization and operational decision-making for optimized UTES well management (H. Abdulhaq, 2025b).



IV.Figure 7: Radial flow chart depicting the structured workflow of this research. Red nodes represent input stages, such as loading and preprocessing data, while blue nodes indicate output stages, including model predictions, validation, and generation of final maps and quality control charts. Arrows illustrate the directional data flow and method dependencies throughout the analysis pipeline.

IV.4.7 Sensitivity Analysis

A sensitivity analysis was conducted by systematically evaluating combinations of key modeling parameters, specifically the number of neighbors used in inverse distance weighting (idw_k_neigh), radial basis function interpolation (rbf_neigh), the XGBoost learning rate, and the Z-score threshold (z_thresh) applied for outlier filtering (Hosamo & Mazzetto, 2025). A total of 50 parameter combinations were randomly sampled from a structured parameter grid to assess their impact on predictive performance, measured by cross-validation R^2, test R^2, mean absolute error (MAE), and root mean squared error (RMSE) (Gao et al., 2025). The analysis revealed that the model was particularly sensitive to the learning rate and the spatial complexity controlled by idw_k_neigh and rbf_neigh. The optimal predictive performance (test $R^2 = 0.971$, MAE = 0.093, RMSE = 0.178) was achieved using moderate spatial complexity (idw_k_neigh=7, rbf_neigh=40), a relatively higher learning rate (0.05), and a stringent Z-score threshold (3.0) (Bărbulescu & Saliba, 2024). The detailed results of this optimal parameter combination and corresponding performance metrics are presented in Table 1. These findings emphasize the critical balance required in parameter selection, highlighting that intermediate values for spatial parameters combined with higher learning rates effectively improve predictive robustness while mitigating potential overfitting or underfitting scenarios.

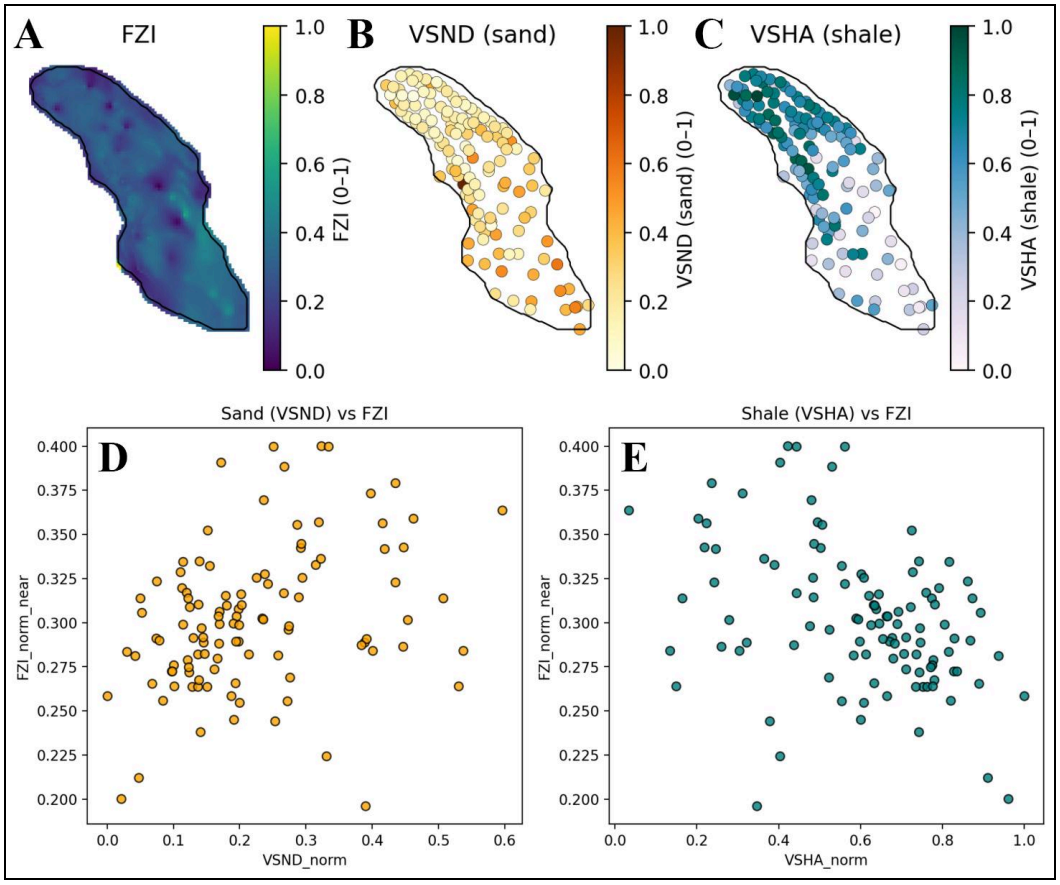
IV.Table 1: Optimal hyperparameters and corresponding model performance metrics from the sensitivity analysis of the FZI prediction workflow.

IDW Neighbors (idw_k_neigh)	RBF Neighbors (rbf_neigh)	Learning Rate	Z-Threshold	CV R^2	Test R^2	Test MAE	Test RMSE
7	40	0.05	3	0.961	0.971	0.093	0.178

9	40	0.05	3	0.961	0.971	0.094	0.18
9	60	0.05	3	0.961	0.971	0.094	0.18
5	60	0.05	3	0.961	0.972	0.093	0.177
5	20	0.05	4	0.955	0.962	0.098	0.211
3	40	0.05	4	0.954	0.963	0.098	0.209
5	40	0.05	3.5	0.958	0.965	0.098	0.2
3	60	0.05	3.5	0.957	0.965	0.098	0.2
9	20	0.05	3.5	0.958	0.964	0.099	0.202
9	40	0.05	4	0.954	0.962	0.099	0.212

IV.4.8 Model Calibration and Validation

The model calibration and validation involved analyzing the relationships between the normalized predicted FZI values (FZI_norm) and key petrophysical indicators, specifically normalized sand content (VSND_norm) and normalized shale content (VSHA_norm). Correlation analysis showed a statistically significant positive correlation between FZI_norm and VSND_norm, with Pearson's correlation coefficient (r) of 0.333 (p=0.000461) and Spearman's rank correlation coefficient (ρ) of 0.360 (p=0.00014). Conversely, a statistically significant negative correlation was observed between FZI_norm and VSHA_norm, yielding a Pearson's r of -0.384 (p=4.45e-05) and Spearman's ρ of -0.389 (p=3.41e-05). These findings confirm that the predicted normalized FZI is consistent with geological expectations, positively correlating with sandy, permeable zones and negatively with shaly, lower-quality intervals (IV.Figure 8) (Abraham-A et al., 2022).



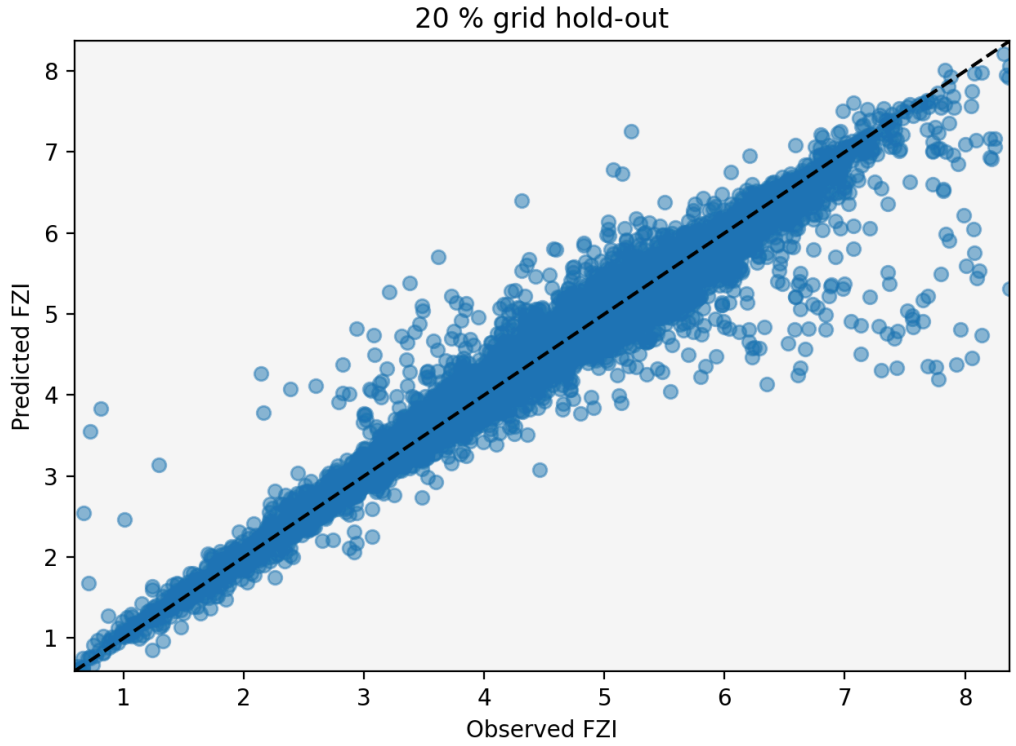
IV.Figure 8 A; Normalized maps illustrating spatial distributions of Flow Zone Index (FZI), B; sand content (VSND), and C; shale content (VSHA). D; Scatter plots showing correlations between normalized FZI and sand content (positive correlation, r=0.333, p<0.001) and E; shale content (negative correlation, r=-0.384, p<0.001), highlighting sandy channels.

IV.5. Results

IV.5.1 FZI Prediction result

The three-dimensional FZI prediction model demonstrated high accuracy and stability across all depth slices in the reservoir. Aggregated performance metrics reveal that the mean absolute error (MAE) averaged just 0.056, while the mean root mean squared error (RMSE) was approximately 0.17, with the majority of predictions exhibiting errors well below these values. The distribution of MAE and RMSE across individual slices is tightly clustered, with 75% of grid predictions falling under 0.055 MAE and 0.19 RMSE, and only a handful of locations reaching a maximum

MAE of 0.43 and RMSE of 0.87. These statistics underscore the overall precision and consistency of the model, ensuring that both local and global FZI variations are faithfully captured (IV.Figure 9) (Djebbas et al., 2023).

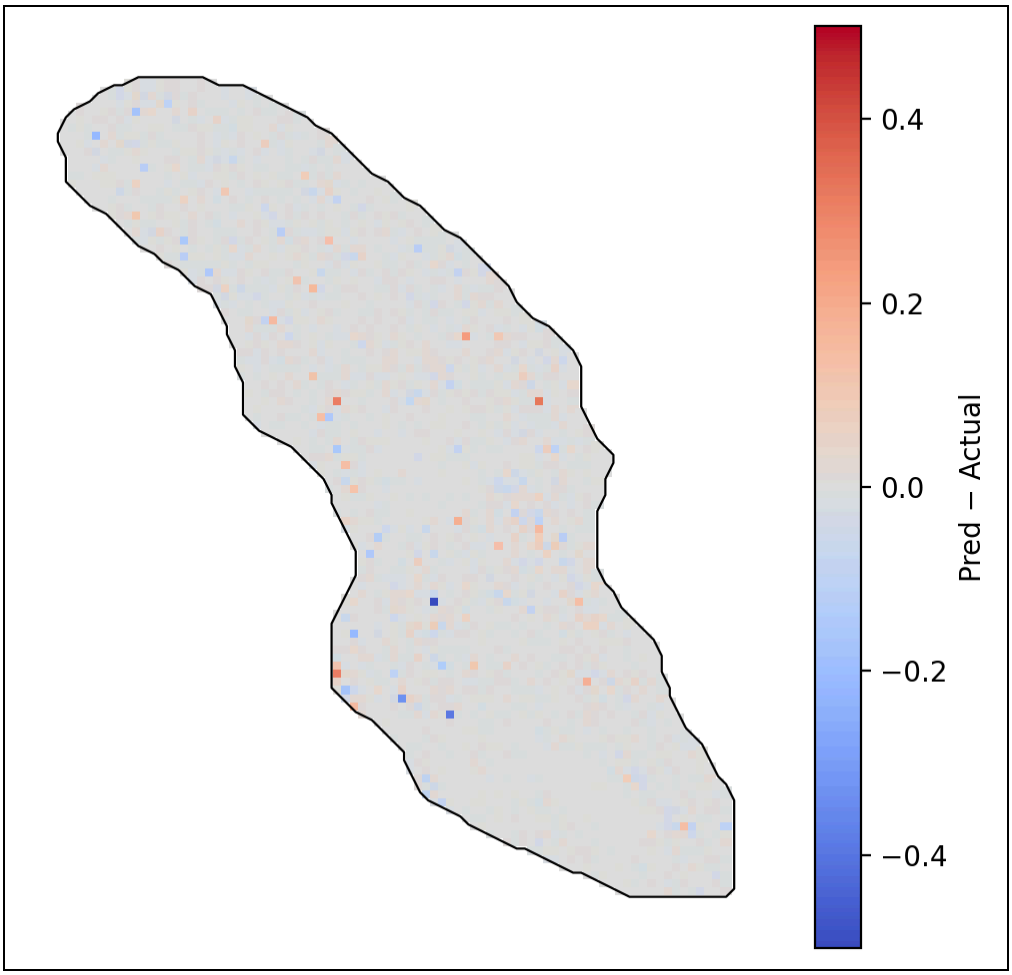


IV.Figure 9 Observed versus predicted Flow Zone Index (FZI) values from the 20% validation dataset, highlighting strong model performance . The dashed line indicates a perfect prediction (1:1 line), illustrating the model's accuracy in capturing subsurface heterogeneity.

The predicted Flow Zone Index (FZI) values across the three-dimensional grid display a high degree of concordance with the observed data, as demonstrated by the distribution of residuals between predictions and measurements. The majority of grid points exhibit very small residuals, with the largest concentration falling in the lowest error bins—thousands of points have absolute residuals less than 0.01, indicating that the model's predictions are extremely close to actual FZI values for most of the reservoir (Djebbas et al., 2023). The gradual increase in bin counts with increasing error, followed by a sharp decline at higher error levels, further confirms that large deviations are rare and the prediction error is dominated by minor discrepancies.

IV.5.2. Residual Distribution and Model Robustness

A closer examination of the residual histogram reveals that over 90% of the prediction errors are less than 0.2 in absolute value, with the most populated bins spanning error ranges from 10^{-6} to approximately 0.2 (Mohammadian et al., 2022). Only a small fraction of points have larger residuals, with very few exceeding 1.0. This skewed, tightly centered error distribution reflects the robustness of the machine learning workflow, which includes extensive outlier filtering, careful spatial feature engineering, and a two-step modeling process (XGBoost regression followed by RBF residual correction)(Mohammadpour et al., 2024). The results suggest the model successfully captures both the spatial heterogeneity and the petrophysical complexity of the reservoir, yielding predictions with minimal systematic bias (IV.Figure 10).



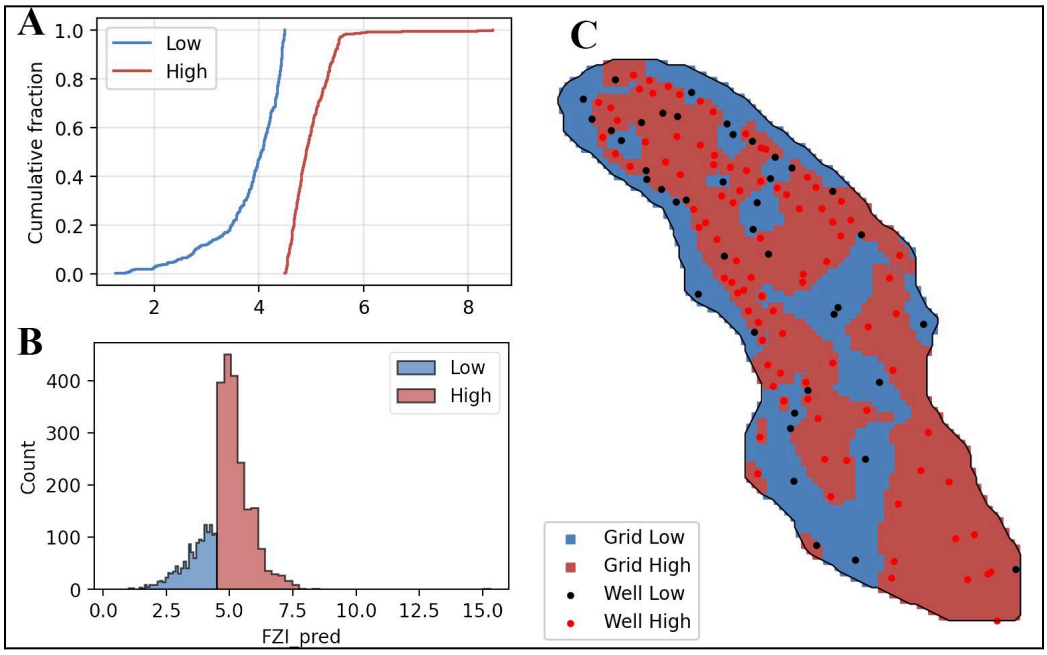
IV.Figure 10 Variance map showing the mean residuals (Predicted minus Actual FZI). The spatial distribution highlights minimal prediction errors, characterized by negligible bias.

Overall, the prediction outcome demonstrates both accuracy and stability across the study area, with negligible evidence of overfitting or model drift. The extremely low and sparsely distributed large residuals imply that prediction failures, when they do occur, are isolated and do not compromise the reliability of the overall FZI map (Al-Dhafeeri & Nasr-El-Din, 2007; Man et al., 2021). Such a result is crucial for reservoir characterization workflows, as it provides a high-resolution, spatially continuous prediction surface with quantified uncertainty, supporting confident operational decisions and further geological analysis.

IV.6. Discussion

IV.6.1 Interpretation of Predicted 3D FZI Clusters

The predicted three-dimensional FZI clusters, classified into two distinct categories—Low-FZI (<4.5) and High-FZI (≥4.5)—clearly delineate reservoir units into sand-rich channel systems and surrounding low-quality lithologies (IV.Figure 10) (Koroncz et al., 2022; Man et al., 2021; Nabawy et al., 2018). The Low-FZI cluster includes 363 grid nodes and 72 wells, characterized by an average FZI value of 3.85 (std=0.67), typically indicating fine-grained or cemented lithologies with reduced permeability and limited fluid transmissivity. In contrast, the High-FZI cluster comprises 366 grid nodes and 60 wells, exhibiting a higher average FZI of 4.99 (std=0.41), with maximum values up to 8.47. These high-FZI zones strongly suggest well-connected, sandy channel systems favorable for fluid flow (Cheng et al., 2022).



IV.Figure 11 Map of areas with predicted FZI values greater than 4.5, highlighting inferred sand channel systems. (A) Cumulative distribution of FZI values, (B) frequency histogram of FZI, and (C) spatial map displaying the mean distribution of FZI clusters.

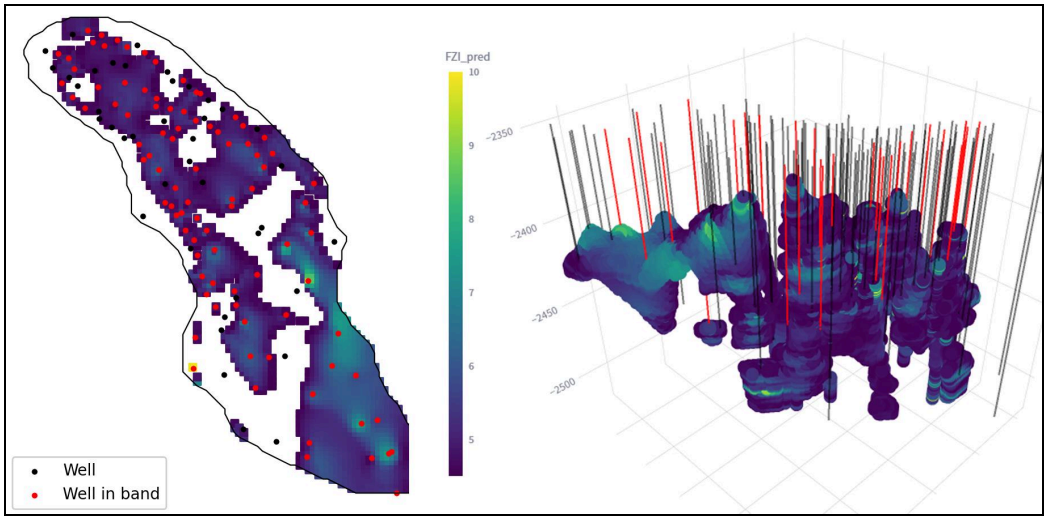
Wells that directly penetrate the High-FZI channels are shown in red; however, some wells located within high-FZI zones remain marked in black, indicating that although they appear

spatially associated, they did not penetrate the channelized system vertically (IV.Figure 11C). This highlights the importance of precise 3D delineation for accurately identifying reservoir connectivity and optimizing well placement (Fu et al., 2025; A. (刘奥迪) Liu et al., 2024).

The clear separation between the Low and High-FZI clusters emphasizes the model's ability to accurately distinguish between sand-rich channels and surrounding lower-quality lithologies, reflecting significant petrophysical contrasts (Steventon et al., 2021). The minimal overlap and tightly constrained standard deviations within each cluster confirm the robustness of the classification, indicating internally consistent reservoir properties. The balanced spatial distribution of nodes and wells across both clusters highlights the geological complexity and heterogeneity of the reservoir (Shirinkin et al., 2024). Crucially, the distinct delineation of High-FZI clusters provides essential insight for reservoir management, enabling targeted placement of wells within high-permeability channel systems, thus optimizing fluid flow, enhancing energy recovery, and reducing the risk of premature thermal breakthrough (Shirinkin et al., 2024).

IV.6.2 Subsurface Complexity of High-FZI Regions in 3D

The horizontal (depth-mean) analysis of the predicted FZI clusters identifies distinct high-FZI regions (FZI $\geq$ 4.5) distributed across the reservoir, but the three-dimensional visualization highlights considerably greater complexity and detail (Singh et al., 2024). In the 3D view, the high-FZI zones form vertically stacked, laterally extensive, channel-like features, consistent with known channel deposition characteristics of the Szolnok Formation (IV.Figure 12 Right) (Fathy et al., 2025). Specifically, out of 60,954 analyzed grid nodes, extensive regions exhibit FZI values exceeding 4.5, closely aligning with interpreted sandy channel systems (El-Gendy et al., 2023). Notably, 29 out of 136 wells directly penetrate these high-FZI channels (highlighted in red), while the remaining 107 wells fall outside these zones (displayed in black). This clear delineation provides valuable spatial context for reservoir characterization, reflecting geological controls on fluid-flow pathways (Singh et al., 2024).



IV.Figure 12 (Left) Spatial distribution of mean FZI values greater than 4.5, with wells shown in red if they penetrate these predicted high-FZI zones. (Right) 3D visualization of predicted FZI values exceeding 4.5, highlighting the complex geometry of the inferred channel systems.

This detailed three-dimensional understanding of the high-FZI zones offers substantial advantages for the design and optimization of UTES (Underground Thermal Energy Storage) cold and hot well duplets (Mahbaz et al., 2021). The naturally channelized high-FZI zones, indicative of enhanced permeability and reservoir quality, can be strategically utilized to maximize heat transfer and storage efficiency while simultaneously mitigating the risk of thermal breakthrough. By placing injection (cold) and production (hot) wells in separate, isolated channel segments, operators can effectively use the natural geologic architecture as barriers, minimizing unintended thermal communication (Wachowicz-Pyzik et al., 2020). Consequently, leveraging the inherent channel geometry of the Szolnok Formation supports optimal placement of duplet wells and also enhances overall UTES system performance and sustainability (Mahbaz et al., 2021).

IV.6.3 Enhancing Decision-Making for UTES Site Selection

Direct well data-driven prediction of channel systems within the Szolnok Formation significantly enhances decision-making for UTES (Underground Thermal Energy Storage) site selection and the management of cold and hot wells in depleted hydrocarbon reservoirs (H. Abdulhaq, 2025b). By utilizing high-resolution, spatially continuous models that clearly identify channelized sand

bodies based on actual well log data, operators can precisely target highly permeable zones ideal for thermal energy injection and extraction. Accurate delineation of channel geometries enables strategic placement of cold and hot well duplets, effectively utilizing natural geological boundaries to minimize the risk of premature thermal breakthrough (Rangel-Jurado et al., 2023). The model's ability to reliably capture the spatial distribution and connectivity of channelized zones reduces subsurface uncertainty, ensuring that well planning aligns closely with geological reality. Ultimately, this channel-focused predictive approach promotes more efficient, sustainable, and cost-effective thermal energy storage operations within repurposed hydrocarbon reservoirs (H. Abdulhaq, 2025b).

6.4 Generalizing Channelization Prediction for the Szolnok Formation

This model offers substantial potential for predicting channel systems in other basins containing the Szolnok Formation, particularly if comparable well log datasets are available (H. Abdulhaq, 2025b). Given the widespread occurrence and consistent depositional characteristics of the Szolnok Formation across the region, the established relationships between well log responses and channel presence are likely transferable to areas lacking detailed channelization data ("(PDF) CO2 Mineralization in a Saline Aquifer-like Environment, Pannonian Basin, Hungary," 2025). However, the methodology may require recalibration or adaptation in regions with significantly different depositional environments, sediment composition, or structural complexities. By applying the same preprocessing and feature-engineering workflows to well logs from the new basin, operators can rapidly generate high-resolution predictions of channel geometries and connectivity (Duplyakin et al., 2022). This approach enables efficient identification of sandy channel bodies ideal for targeted UTES, geothermal, or hydrocarbon development, even in locations where direct geological observations are limited (Farkas et al., 2025). Ultimately, the ability to generalize channel predictions significantly reduces uncertainty in reservoir characterization, guiding the optimal placement of wells and enhancing data-driven decisions across the broader Pannonian Basin (H. Abdulhaq, 2025b).

IV.6.5 Computational Efficiency in Sand Channel Prediction

Computational efficiency has a major impact on the practical applicability of predictive models in reservoir characterization, as conventional reservoir simulations often demand significant

computational resources and time – especially for complex geology or large datasets (Otmane et al., 2025). Conventional geostatistical methods, such as Sequential Gaussian Simulation (SGS), are widely used due to their robust spatial modeling capabilities (*Geostatistical Modeling of Porosity and Evaluating the Local and Global Distribution | Journal of Petroleum Exploration and Production Technology*, n.d.). However, SGS typically involves substantial computational time (approximately 8–12 hours per scenario) and intensive expert-driven manual calibration.

The hybrid machine-learning workflow developed in this study—combining XGBoost regression with Radial Basis Function (RBF) spatial corrections—offers a significant improvement, delivering comparable or better accuracy (RMSE = 0.24) with drastically reduced computational time (~15 minutes per run) (Otmane et al., 2025; Zarei et al., 2021). The direct, data-driven nature of the approach reduces implementation complexity and enhances operational efficiency (Table 2).

IV.Table 2 clearly summarizes these computational differences:

Method	**RMSE**	**Computational Time**
XGBoost + RBF (this study)	0.24	~15 minutes
Sequential Gaussian Simulation	~0.35–0.45*	~8–12 hours

IV.6.6 Managing Model Uncertainty and Reproducibility

Although our model for predicting subsurface channel systems demonstrates high accuracy, some uncertainty inevitably remains due to potential differences in well log calibration, local geological variations, and the representativeness of training data—particularly when transferring the model to other basins or datasets (Brcković et al., 2024; H. Liu et al., 2020). While this study specifically focused on one oil field in southern Hungary, several other oil fields have historically produced from the Szolnok Formation, potentially benefiting significantly from the

methodological framework developed here (Clayton et al., 1994). Empirical substantiation of the model's generalizability through pilot testing in geologically analogous settings (e.g., the Vienna Basin within the broader Pannonian system) would significantly strengthen its scalability and operational transferability (H. Abdulhaq, 2025c). Applying this structured, data-driven approach to these additional sites could enhance reservoir characterization, optimize thermal energy storage, and facilitate effective site selection and management across the broader region (H. Abdulhaq, 2025b). To maximize transparency and enhance its applicability to new geological settings, I have shared the fully documented code and methodology, including the trained XGBoost model, RBF residual correction model, and feature engineering schema.(H. Abdulhaq, 2025a) These files ensure consistent and reproducible predictions of channel systems by providing the complete workflow, including feature preparation, model application, and residual correction (*Digitalization of Legacy Datasets and Machine Learning Regression Yields Insights for Reservoir Property Prediction and Submarine-Fan Evolution: A Subsurface Example From the Lewis Shale, Wyoming | Published in The Sedimentary Record*, n.d.). Users can leverage these saved artifacts to quantify uncertainty, verify each modeling step, and effectively adapt the channel prediction workflow to other regions with similar log data, ensuring robust and geologically meaningful interpretations (*Digitalization of Legacy Datasets and Machine Learning Regression Yields Insights for Reservoir Property Prediction and Submarine-Fan Evolution: A Subsurface Example From the Lewis Shale, Wyoming | Published in The Sedimentary Record*, n.d.; Mahjour et al., 2025).

IV.6.7 Recommendations and Future Work

Future work should focus on improving the model's predictive accuracy and transferability for identifying sand-rich channel systems across different geological settings by incorporating diverse training datasets from multiple basins. Developing a standardized well-log preprocessing pipeline—including consistent calibration checks and unified feature engineering—will reduce biases when applying the model to new areas. Furthermore, integrating seismic data could provide independent structural validation, while core analyses and previous hydrocarbon production histories would offer additional calibration points, substantially enhancing the resolution and confidence of channel delineations. Employing advanced uncertainty quantification methods, such as probabilistic modeling or ensemble predictions, will also provide

valuable assessments of prediction reliability. Lastly, collaborative data sharing and integration of these validation sources will significantly enhance model validation and practical utility.

IV.7. Conclusion

This research successfully demonstrated a structured, data-driven methodology for accurately predicting the three-dimensional distribution of sand-rich channel systems, indicated by high Flow Zone Index (FZI ≥ 4.5), within the Szolnok Formation in southern Hungary. Leveraging extensive petrophysical data from 128 wells, the study clearly delineated high-permeability sandy channels from surrounding lower-quality intervals characterized by low FZI. The resulting spatial models distinctly identified interconnected channelized sands separated by continuous, low-permeability zones, effectively forming natural barriers crucial for mitigating premature thermal breakthrough in Underground Thermal Energy Storage (UTES) systems.

Although this research focused on one representative oil field in southern Hungary, numerous other fields historically producing from the Szolnok Formation could similarly benefit from adopting this predictive framework. The application of this methodology across these additional sites would enhance reservoir management practices, reduce geological uncertainty, and significantly improve the operational effectiveness of UTES installations regionally.

The integrated modeling workflow, which included rigorous outlier filtering, spatial feature engineering using Inverse Distance Weighted (IDW) interpolation, and predictive modeling via XGBoost regression enhanced with residual corrections, achieved high predictive accuracy. Validation results revealed a robust model performance with an R^2 of 0.92, a mean absolute error (MAE) of approximately 0.056, and a root mean squared error (RMSE) of 0.17. The precision and reliability of these predictions ensure confident strategic planning and optimal well placement, significantly enhancing the operational sustainability and efficiency of thermal storage.

By explicitly identifying regions of high FZI suitable for hot and cold well placement and delineating effective low-FZI barriers, the methodology directly addresses the critical challenge of thermal breakthrough, enhancing UTES system longevity and thermal recovery efficiency. Future research should focus on further refining model accuracy and generalizability through

standardized data preprocessing, integration with seismic surveys, and advanced uncertainty quantification. Ultimately, this robust, reproducible methodology substantially advances reservoir characterization, supporting the global expansion of UTES technology for sustainable, scalable, and economically viable thermal energy storage solutions.

CHAPTER V

CONCLUSIONS

This PhD research is centred on the assessment, design, and optimization of high-temperature aquifer thermal energy storage (HT-ATES) systems in depleted clastic hydrocarbon reservoirs, with a particular focus on the Hungarian subsurface within the Pannonian Basin. The basin represents an exemplary geological and socio-economic setting where mature hydrocarbon fields, elevated geothermal gradients, and extensive district heating demand coexist. By integrating geological characterization, uncertainty analysis, numerical simulation, and machine-learning approaches, this research aims to advance the practical feasibility of repurposing depleted hydrocarbon reservoirs as long-term subsurface heat storage assets.

The research adopts a multi-scale and data-driven methodology, combining probabilistic screening, physically based numerical modelling, and scalable surrogate modelling. Geological heterogeneity, particularly in channelized clastic reservoirs, is treated as a first-order control on heat transport and storage performance rather than as a secondary uncertainty. The results presented in this thesis demonstrate that integrating uncertainty-aware decision frameworks with high-resolution reservoir characterization is essential for reliable HT-ATES feasibility assessment and system optimization.

In the first part of this research, a probabilistic multi-criteria decision analysis framework integrated with geostatistical simulation was developed to screen and rank depleted hydrocarbon reservoirs for HT-ATES suitability. Key geological, thermal, and operational parameters were spatially represented and populated using stochastic realizations to explicitly capture subsurface uncertainty. The resulting suitability maps reveal strong spatial heterogeneity within individual formations and demonstrate that high-potential HT-ATES zones can be identified even in data-limited settings. The application of this framework to the Békés/Békési Formation shows that robust suitability persists across multiple uncertainty realizations, highlighting its potential as a reference formation for HT-ATES deployment in clastic basins. Importantly, the framework provides a transparent and reproducible bridge between regional-scale screening studies and formation-scale design decisions.

The second part of the research focuses on the quantitative evaluation of HT-ATES performance using physically based groundwater flow and heat-transport modelling. High-resolution numerical models were constructed to investigate thermal plume evolution, heat recovery efficiency, and breakthrough timing under realistic operating conditions. The simulations demonstrate that reservoir architecture, anisotropy, and well configuration exert a strong control on storage performance. In particular, an optimal range of inter-well spacing exists in which thermal recovery efficiency is maximized, whereas both overly tight and overly wide well configurations lead to suboptimal performance. The results further show that operational scheduling significantly influences long-term efficiency, and that storage performance improves over successive cycles as thermal conditions stabilize. These findings confirm that HT-ATES performance cannot be reliably predicted without explicitly resolving geological heterogeneity and system dynamics.

Building on the computational demands of numerical modelling, the third part of the research introduces machine-learning and hybrid modelling approaches to enable scalable performance prediction. Surrogate models trained on ensembles of numerical simulations successfully capture the relationships between reservoir properties, well configurations, and thermal performance metrics. These models achieve high predictive accuracy while reducing computational cost by more than an order of magnitude, enabling rapid evaluation of large parameter spaces. The application of these surrogate models beyond the original numerical model domain demonstrates their potential for formation-wide screening and optimization, provided that input parameters remain within physically meaningful bounds.

A further key contribution of this research is the development of hybrid geostatistical–machine-learning workflows for high-resolution prediction of sand-channel architecture in clastic reservoirs. By explicitly delineating high-permeability flow zones and adjacent low-permeability barriers, the predicted channel networks provide geologically interpretable inputs for HT-ATES design. The results show that assigning injection and production wells to hydraulically separated channel bodies can significantly delay thermal breakthrough and enhance storage efficiency. These findings demonstrate that detailed characterization of sedimentary architecture is a critical prerequisite for long-lived and efficient HT-ATES systems in channelized reservoirs.

Overall, this research demonstrates that depleted hydrocarbon reservoirs should not be treated solely as decommissioning liabilities, but rather as strategic subsurface assets capable of supporting the energy transition. The integrated workflow developed in this thesis links geological uncertainty, physical modelling, and data-driven optimization into a coherent decision-support framework. While the Hungarian subsurface serves as a primary case study, the methodologies and insights presented here are transferable to other clastic basins with similar legacy datasets and heating demands. By bridging the gap between geological characterization and energy-system design, this PhD research contributes to the advancement of subsurface thermal energy storage as a key component of sustainable energy infrastructure.

V.1. RECOMMENDATIONS

1. Application of the integrated screening and modelling framework to additional depleted hydrocarbon fields within the Pannonian Basin and other sedimentary basins would improve understanding of the regional variability in HT-ATES potential and strengthen confidence in large-scale deployment strategies.

2. Further coupling of numerical HT-ATES models with geomechanical and geochemical simulations is recommended to assess long-term reservoir integrity, thermal stress effects, and potential fluid–rock interactions under repeated thermal cycling.

3. Expansion of machine-learning surrogate models using larger and more diverse training datasets would improve robustness and allow for broader extrapolation across different reservoir types and operational scenarios.

4. Detailed sedimentological and petrophysical studies integrating core, image logs, and high-resolution seismic data are recommended to further constrain sand-channel connectivity and improve geological realism in HT-ATES design.

5. Long-term pilot HT-ATES operations in repurposed hydrocarbon fields should be monitored to validate model predictions and to refine operational strategies under real-world conditions.

6. Future research should investigate the integration of HT-ATES with other subsurface energy technologies, such as geothermal production and CO_2 storage, to maximize the multifunctional use of mature reservoirs and enhance overall system sustainability.

SUMMARY

INTRODUCTION AND AIMS

The decarbonization of energy systems requires reliable and large-scale solutions for balancing the temporal mismatch between heat supply and demand. In regions with strong seasonal heating requirements, underground thermal energy storage (UTES) has emerged as a key enabling technology for integrating renewable and excess heat into district heating and industrial energy systems. Among UTES concepts, high-temperature aquifer thermal energy storage (HT-ATES) is particularly promising due to its ability to store heat at temperatures suitable for direct use without extensive upgrading.

Depleted clastic hydrocarbon reservoirs represent attractive candidates for HT-ATES deployment because they possess proven porosity–permeability systems, extensive well infrastructure, and long production histories that significantly reduce geological uncertainty and development risk. However, despite their potential, the assessment and design of HT-ATES systems in such reservoirs remain challenging due to strong geological heterogeneity, uncertainty in subsurface properties, and the high computational cost of detailed numerical simulations. In particular, channelized clastic reservoirs exhibit complex flow architectures that strongly control heat transport, thermal breakthrough, and storage efficiency.

The Hungarian subsurface, located within the Pannonian Basin, provides an ideal setting for investigating these challenges. The basin combines elevated geothermal gradients, dense well coverage, extensive legacy datasets, and significant district heating demand. These characteristics make it a representative testbed for developing transferable, data-driven workflows for HT-ATES assessment and design.

The primary aim of this PhD research is to develop an integrated, uncertainty-aware, and data-driven framework for assessing, designing, and optimizing HT-ATES systems in depleted clastic hydrocarbon reservoirs. The specific aims of this research are:

1. to establish a probabilistic framework for screening and ranking HT-ATES suitability using multi-criteria decision analysis integrated with geostatistical simulation;

2. to evaluate HT-ATES performance in heterogeneous clastic reservoirs using physically based numerical flow and heat-transport modelling;
3. to develop scalable machine-learning and hybrid modelling approaches for rapid performance prediction and high-resolution geological characterization; and
4. to demonstrate the applicability of the proposed framework through case studies from the Hungarian subsurface.

APPLIED METHODS

To address the objectives of this research, an integrated methodological approach was adopted, combining geological characterization, probabilistic analysis, numerical simulation, and machine learning. Geological and petrophysical input data were derived from legacy hydrocarbon datasets, including well logs, core measurements, and production records from depleted fields in the Pannonian Basin.

A probabilistic multi-criteria decision analysis (MCDA) framework based on the analytical hierarchy process (AHP) was developed to screen HT-ATES suitability. Key geological, thermal, and operational parameters—including porosity, permeability, reservoir thickness, temperature, heat capacity, and well density—were spatially represented on a grid and populated using geostatistical simulation. Multiple stochastic realizations (P10, P50, and P90) were generated to explicitly capture subsurface uncertainty and to produce formation-scale suitability maps.

Physically based numerical modelling of groundwater flow and heat transport was carried out using coupled flow and thermal transport simulators. High-resolution models were constructed to investigate thermal plume evolution, heat recovery efficiency, and breakthrough timing under realistic operational scenarios. Sensitivity analyses were performed to assess the effects of reservoir heterogeneity, anisotropy, well spacing, and operational scheduling on HT-ATES performance.

To overcome the computational limitations of large ensembles of numerical simulations, machine-learning surrogate models were developed. Supervised learning algorithms were trained

on numerical simulation outputs to predict key performance metrics, such as thermal recovery efficiency and breakthrough time. In addition, hybrid geostatistical–machine-learning workflows were implemented to predict high-permeability sand-channel architectures using petrophysical and spatial features derived from well data. These predicted channel networks were directly integrated into HT-ATES design and optimization workflows.

NEW SCIENTIFIC RESULTS

The research presented in this thesis has yielded the following new scientific results concerning the assessment and optimization of HT-ATES systems in depleted clastic hydrocarbon reservoirs:

T1: The multi-criteria decision analysis–analytical hierarchy process (MCDA–AHP) framework, integrated with geostatistics, is a time-efficient and systematic way to assess the suitability of abandoned hydrocarbon fields in Hungary for underground heat storage. In this framework, key reservoir and operational parameters—such as porosity, permeability, thickness, temperature, heat capacity, thermal conductivity, and well density—are represented on a grid and populated using stochastic geostatistical simulations to capture subsurface uncertainty. Expert-derived weights are assigned to these criteria and combined into composite suitability indices, which classify each grid block as having high, medium, or low potential for thermal energy storage. The resulting suitability maps show that the method can rapidly highlight the most promising zones within each field, even in data-limited settings. These findings indicate that MCDA–AHP, when coupled with geostatistics, provides a practical decision-support tool for repurposing mature fields as underground thermal energy storage sites rather than treating them solely as decommissioning liabilities.

T2: The Békés Formation in Hungary is consistently identified as a prime candidate for high-temperature aquifer thermal energy storage (HT-ATES) by the probabilistic MCDA–AHP screening framework. Analyses of P10, P50, and P90 realizations generated from geostatistical simulations show that this formation repeatedly attains high suitability scores over a large fraction of its area, even when accounting for uncertainty in porosity, permeability, and temperature. The contrast in suitability between favorable and marginal zones within the formation demonstrates the framework's ability to resolve spatial heterogeneity and to distinguish robust targets from more uncertain ones. When compared with earlier, coarser national screening studies, the results show that the method effectively bridges regional-scale ranking and formation-scale design insight. These observations imply that the Békés/Békési Formation can serve as a reference case for HT-ATES deployment in clastic basins and that the MCDA–AHP approach provides a rigorous basis for prioritizing formations and fields under geological uncertainty.

T3: Repurposing depleted hydrocarbon fields for underground thermal energy storage is shown to be economically competitive with, and in many cases preferable to, conventional well decommissioning when evaluated within the MCDA–AHP framework. Typical decommissioning costs, which can range from tens of thousands to over one million USD per well, contrast with the option of retaining and adapting existing boreholes and reservoir infrastructure for long-term energy storage services. In the decision framework, criteria related to economic performance, environmental risk reduction, and regulatory constraints can be combined with geological parameters to rank fields according to their repurposing potential. The results demonstrate that many fields currently categorized as end-of-life liabilities could instead be treated as strategic subsurface assets for heat storage, CO_2 storage, or other low-carbon applications. This indicates that integrating repurposing options into decommissioning planning can materially change the cost–benefit balance for operators and regulators, with significant implications for national energy transition strategies.

T4: A high-resolution hydrogeological and thermal transport model of the Békési Formation demonstrates that a depleted clastic reservoir can sustain high-temperature aquifer thermal energy storage (HT-ATES) under realistic operating conditions. The model, implemented with MODFLOW and MT3DMS, uses well logs, core analyses, and production histories to reconstruct the three-dimensional distribution of porosity, permeability, and anisotropy and to translate these properties into hydraulic and thermal parameters. Simulations of dual-well (duplet) configurations show that injection temperatures on the order of 95 °C are feasible and that the resulting thermal plumes remain largely contained within the formation over multiple seasonal cycles. The spatial evolution of the heat plume reveals that reservoir architecture and anisotropy strongly control heat migration pathways, storage efficiency, and the timing of thermal breakthrough at the production well. These outcomes indicate that physically based numerical modeling can directly link clastic reservoir heterogeneity to HT-ATES performance and provide quantitative design guidance.

T5: Systematic numerical experiments show that optimal inter-well spacing and operational scheduling can increase HT-ATES thermal recovery efficiencies in the Békési Formation from about 81 % initially to nearly 89 % after several cycles, corresponding to an improvement of roughly 8.5 %. Suites of MODFLOW/MT3DMS simulations performed

under varying porosity, permeability, and anisotropy conditions quantify how thermal breakthrough timing and recovery factors depend on both geology and well layout. The results indicate that there exists an intermediate range of spacing in which heat recovery is maximized: wells placed too close together experience early breakthrough, whereas wells placed too far apart underutilize the available reservoir volume. Scenario comparisons of different seasonal injection–production schedules further show that certain combinations of storage and extraction periods yield systematically higher annual energy returns than others. These findings demonstrate that jointly optimizing well spacing and operational strategy, informed by geological sensitivity analysis, is essential for robust and efficient HT-ATES design in clastic reservoirs.

T6: A Random Forest–based surrogate model trained on ensembles of HT-ATES simulations provides accurate and computationally efficient predictions of thermal performance across large numbers of candidate wells and configurations. The model learns relationships between input descriptors—such as reservoir properties, well spacing, injection temperature, and operational schedule—and outputs including thermal recovery efficiency and breakthrough time, achieving a coefficient of determination of approximately 0.87 on validation data. Once calibrated, the surrogate can evaluate thousands of scenarios at a cost that is more than 90 % lower than running full MODFLOW/MT3DMS simulations for each case, enabling extensive parameter sweeps and uncertainty analyses. Application of the model to wells beyond the original numerical model domain shows that it can successfully extrapolate within a broader geological context, provided the input parameters remain within physically reasonable bounds. These results indicate that hybrid numerical–machine-learning workflows can scale HT-ATES planning from small, expert-curated scenario sets to exhaustive, data-driven optimization.

T7: The Hungarian subsurface is demonstrated to be a particularly favorable and instructive testbed for data-driven repurposing of depleted clastic hydrocarbon reservoirs into HT-ATES systems. The country's dense network of mature wells, relatively high geothermal gradients, and extensive district heating demand create a setting in which stored heat can be practically delivered to end users. Historical datasets—including well logs, core measurements, and production records—are shown to be sufficient to construct geostatistical models, numerical flow and heat-transport simulations, and machine-learning surrogates that collectively support HT-ATES design. Case studies in the Békés/Békési Formation illustrate how

this integrated workflow can identify optimal well configurations, predict long-term performance, and manage uncertainty. These facts imply that the Hungarian example offers a transferable blueprint for other clastic basins with similar legacy data and heating needs, transforming depleted fields into key components of national decarbonization and energy security strategies.

T8: A hybrid machine-learning workflow combining XGBoost regression with radial basis function (RBF)–based spatial residual correction accurately predicts the three-dimensional distribution of high Flow Zone Index ($FZI \geq 4.5$) sand-rich channels in the Szolnok Formation. Petrophysical data from 128 wells, including porosity, permeability, and lithological indicators, are processed into features that capture both local rock properties and their spatial context before being used for training. The resulting model achieves an R^2 of about 0.92 and a root-mean-square error of roughly 0.24, clearly distinguishing high-permeability channel bodies from low-FZI background intervals that behave as barriers to flow. The predicted channel network exhibits geologically plausible geometries and connectivity patterns, matching independent interpretations of the depositional system. These results show that hybrid geostatistical–machine-learning approaches can deliver high-resolution, geologically interpretable channel architectures that exceed the efficiency and, in many cases, the predictive quality of traditional stochastic simulations.

T9: Explicit prediction of sand-channel architecture in the Szolnok Formation is shown to be directly usable for optimizing underground thermal energy storage (UTES) site selection and well placement to reduce thermal breakthrough risks. The model outputs delineate continuous high-FZI channel zones suitable for injecting and producing fluids and identify adjacent low-FZI intervals that act as natural baffles or barriers between channel segments. By assigning hot and cold wells to different, geologically separated channel bodies that are shielded by these low-FZI zones, the design can minimize unwanted hydraulic and thermal communication between well pairs. Simulated flow and heat-transport behavior based on the predicted architecture confirm that such configurations delay thermal breakthrough and improve storage efficiency relative to designs that ignore channel geometry. These findings indicate that detailed channel mapping is a critical prerequisite for long-lived, high-performance UTES systems in channelized clastic reservoirs.

T10: The hybrid XGBoost–RBF workflow for sand-channel prediction is demonstrated to be significantly more computationally efficient than traditional Sequential Gaussian Simulation (SGS) while maintaining comparable or superior accuracy, and it is designed for reuse in other basins. Benchmarking shows that a typical SGS realization requiring approximately 8–12 hours of computation can be replaced by a machine-learning–based prediction generated in around 15 minutes once the model is trained. The workflow, which includes standardized feature engineering from well logs and spatial residual correction, is shown to be portable to other Szolnok-bearing fields and analogous channelized clastic systems, subject to recalibration with local data. The codebase and model configuration are documented and shareable, allowing other practitioners to reproduce, audit, and extend the approach. These facts suggest that the workflow can serve as a practical standard for channel-focused characterization in UTES, geothermal, and hydrocarbon projects across the Pannonian Basin and similar sedimentary provinces.

ACKNOWLEDGEMENTS

This PhD journey has been both challenging and deeply rewarding, and it would not have been possible without the support, guidance, and encouragement of many people.

First and foremost, I would like to express my deepest gratitude to Dr. János Szanyi, who was far more than a supervisor to me. He has been a true mentor and a scientific compass throughout my PhD. His guidance, insight, and constant encouragement shaped not only this research but also my development as a researcher. I am sincerely thankful for his trust, patience, and unwavering support.

I am profoundly grateful to Dr. Tivadar M. Tóth, whose dedication and determination made this PhD possible despite the many administrative and institutional challenges along the way. His support, persistence, and willingness to stand by this project were instrumental, and I deeply appreciate everything he has done to help bring this work to completion.

I would also like to thank Dr. István Vass, my co-supervisor, for his support, constructive discussions, and valuable perspectives that strengthened this research. My sincere thanks go as well to Dr. János Geiger, whose help was absolutely vital. His expertise, availability, and readiness to assist at critical moments had a significant impact on the success of this work.

I extend my appreciation to the Department of Geology, with special thanks to Szilvia Csató, the department coordinator, for her continuous support, kindness, and help with countless organizational matters that made everyday academic life much easier.

I would like to gratefully acknowledge SZETÁV, which represented the closest industry connection to my research. Their openness, professional support, and practical insight into district heating systems were invaluable and greatly enriched the applied aspects of this work.

Finally, my deepest and most heartfelt thanks go to my family. Above all, I thank my wife, Srwa Fadhil. A significant part of this PhD was lived and completed at home, and her belief in me and in my work was a constant source of strength—like a candle that never went out. Her patience, encouragement, and unwavering support carried me through the most demanding moments of this journey. I also thank my wonderful children, Evar, Ela, and Ervin, whose presence, love, and smiles gave meaning and balance to everything I did.

To everyone who supported me directly or indirectly along this journey, I am truly grateful.

REFERENCES

Abbaszadeh, M., Fujii, H., & Fujimoto, F. (1996). Permeability Prediction by Hydraulic Flow Units—Theory and Applications. SPE Formation Evaluation, 11(04), 263–271. https://doi.org/10.2118/30158-PA

Abdallah, S., Sid Ali, O., & Benmalek, S. (2016). Rock type and permeability prediction using flow-zone indicator with an application to Berkine Basin (Algerian Sahara). In SEG Technical Program Expanded Abstracts 2016 (pp. 3068–3072). Society of Exploration Geophysicists. https://doi.org/10.1190/segam2016-13943527.1

Abdulhaq, H. (2024). Supplementary data for the manuscript "Guided Transformation of Abandoned Hydrocarbon Fields into Heat Storage Solutions Using Enhanced MCDA-AHP Geostatistical Based Method: Hungarian Case Study. [Dataset]. Zenodo. https://doi.org/10.5281/zenodo.11507335

Abdulhaq, H. (2025a). Machine Learning Model for Predicting Thermal Performance of High-Temperature Aquifer Thermal Energy Storage (HT-ATES) in Depleted Clastic Reservoirs. Zenodo. https://doi.org/10.5281/ZENODO.15294846

Abdulhaq, H. (2025b). Predicting Thermal Performance of Aquifer Thermal Energy Storage Systems in Depleted Clastic Hydrocarbon Reservoir via Machine Learning: Case Study from Hungary. The 16th European Geothermal PhD Days Book of Abstracts. https://www.egpd2025.com

Abdulhaq, H. (2025c). Thermal Performance Analysis and Visualization App for UCN-Based Heat Simulations. Zenodo. https://doi.org/10.5281/ZENODO.15294958

Abdulhaq, H. (2025d). 3D FZI Prediction Pipeline and Interactive Streamlit App for Well Log Analysis. https://zenodo.org/records/15814254

Abdulhaq, H. (2025e). XGBoost Machine Learning Model for 3D FZI Prediction in Szolnok Formation. https://zenodo.org/records/16217944

Abdulhaq, H. (2025f). Supplementary Materials for the Hungarian Heat Storage Case Study. DatasetDatasetDataset. Zenodo. https://doi.org/10.5281/zenodo.18070837

Abdulhaq, H. A., Geiger, J., Vass, I., Tóth, T. M., Medgyes, T., & Szanyi, J. (2024). Transforming Abandoned Hydrocarbon Fields into Heat Storage Solutions: A Hungarian Case Study Using Enhanced Multi-Criteria Decision Analysis–Analytic Hierarchy Process and Geostatistical Methods. Energies, 17(16), Article 16. https://doi.org/10.3390/en17163954

Abraham-A, R. M., Taioli, F., & Nzekwu, A. I. (2022). Physical properties of sandstone reservoirs: Implication for fluid mobility. Energy Geoscience, 3(4), 349–359. https://doi.org/10.1016/j.engeos.2022.06.001

Ahmadi, M. (2025). Interpretable Machine Learning for High-Accuracy Reservoir Temperature Prediction in Geothermal Energy Systems. Energies, 18(13), Article 13. https://doi.org/10.3390/en18133366

Akindipe, D., McTigue, J., Dobson, P., Atkinson, T., Witter, E., Kumar, R., Sonnenthal, E., Umbro, M., Lederhos, J., Adams, D., & Zhu, G. (2024). Techno-Economic Analysis and Market Potential of Geological Thermal Energy Storage (GeoTES) Charged With Solar Thermal and Heat Pumps (NREL/TP--5700-91225, 2474842, MainId:93003; p. NREL/TP--5700-91225, 2474842, MainId:93003). https://doi.org/10.2172/2474842

Al-Dhafeeri, A. M., & Nasr-El-Din, H. A. (2007). Characteristics of high-permeability zones using core analysis, and production logging data. Journal of Petroleum Science and Engineering, 55(1), 18–36. https://doi.org/10.1016/j.petrol.2006.04.019

Al-Dousari, D. M., Almudhhi, S., & Garrouch, A. (2022). Predicting the flow zone indicator of carbonate reservoirs using NMR echo transforms, and routine open-hole log measurements : Journal of Engineering Research, 10(1A), Article 1A. https://doi.org/10.36909/jer.13185

Amaefule, J. O., Altunbay, M., Tiab, D., Kersey, D. G., & Keelan, D. K. (1993, October 3). Enhanced Reservoir Description: Using Core and Log Data to Identify Hydraulic (Flow) Units and Predict Permeability in Uncored Intervals/Wells. SPE Annual Technical Conference and Exhibition. https://doi.org/10.2118/26436-MS

Anderson, M. P., Woessner, W. W., & Hunt, R. J. (Eds.). (2015). Front Matter. In Applied Groundwater Modeling (Second Edition) (p. iii). Academic Press. https://doi.org/10.1016/B978-0-08-091638-5.00018-3

Anttila, A. (2021). Techno-economic comparison of thermal energy storage solutions for decarbonizing heat in Espoo by 2025. https://aaltodoc.aalto.fi/server/api/core/bitstreams/a65b9c93-33d1-4975-8a97-af3fdf3d39b1/content

Astsauri, T., Habiburrahman, M., Ibrahim, A. F., & Wang, Y. (2024). Utilizing machine learning for flow zone indicators prediction and hydraulic flow unit classification. Scientific Reports, 14(1), 4223. https://doi.org/10.1038/s41598-024-54893-1

Az SGS a világ vezető minőségellenőrző, ellenőrző, vizsgáló és tanúsító cége. (n.d.). SGSCorp. Retrieved June 5, 2024, from https://www.sgs.com/hu-hu/

B 2204-B: Pannonian Basin Province, Central Europe (Province 4808)—Petroleum Geology, Total Petroleum Systems, and Petroleum Resource Assessment. (n.d.). Retrieved July 8, 2025, from https://pubs.usgs.gov/bul/2204/b/

Badem, A. C., Yılmaz, R., Cesur, M. R., & Cesur, E. (2024). Advanced Predictive Modeling for Dam Occupancy Using Historical and Meteorological Data. Sustainability, 16(17), Article 17. https://doi.org/10.3390/su16177696

Baird, K., Arnold, D., Doster, F., Geiger, S., Petrovskyy, D., Hampson, G. J., Jacquemyn, C., Jackson, M. D., Silva, J. M., Rahman, F., Judice, S., & Sousa, M. C. (n.d.). Assessment of the Impacts of Multi-Scale Sedimentological Heterogeneity on Low-Enthalpy Geothermal Energy Production.

Baisad, K., Chutsagulprom, N., & Moonchai, S. (2023). A Non-Linear Trend Function for Kriging with External Drift Using Least Squares Support Vector Regression. Mathematics, 11(23), Article 23. https://doi.org/10.3390/math11234799

Bálint, A., & Szanyi, J. (2015). A half century of reservoir property changes in the Szentes geothermal field, Hungary. https://doi.org/10.1556/24.58.2015.1-2.2

Bărbulescu, A., & Saliba, Y. (2024). Sensitivity Analysis of the Inverse Distance Weighting and Bicubic Spline Smoothing Models for MERRA-2 Reanalysis PM2.5 Series in the Persian Gulf Region. Atmosphere, 15(7), Article 7. https://doi.org/10.3390/atmos15070748

Bloemendal, M., & Hartog, N. (2018). Analysis of the impact of storage conditions on the thermal recovery efficiency of low-temperature ATES systems. Geothermics, 71, 306–319. https://doi.org/10.1016/j.geothermics.2017.10.009

Breiman, L. (2001). Random Forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brcković, A., Orešković, J., Cvetković, M., & Marić-Đureković, Ž. (2024). Enhancing the Understanding of Subsurface Relations: Machine Learning Approaches for Well Data Analysis in the Drava Basin, Pannonian Super Basin. Applied Sciences, 14(14), Article 14. https://doi.org/10.3390/app14146039

Bridger, D. W., & Allen, D. M. (2010). Heat transport simulations in a heterogeneous aquifer used for aquifer thermal energy storage (ATES). Canadian Geotechnical Journal, 47(1), 96–115. https://doi.org/10.1139/T09-078

Cao, T., Cui, J., He, Y., Ma, L., Qiao, W., & Liu, Y. (2024). The Influence of Reservoir Architecture on the Connectivity of the Shahejie Formation in the Liuzhong Oilfield. Energies, 17(1), Article 1. https://doi.org/10.3390/en17010115

Challenges and Opportunities for Aquifer Thermal Energy Storage (ATES) in EU Energy Transition Efforts—An Overview. (n.d.). Retrieved July 5, 2025, from https://www.mdpi.com/1996-1073/18/4/1001

Cheng, X., Zhao, B., Gao, C., & Gao, Y. (2022). Reservoir Permeability Calculation under Flow Unit Control. Energies, 15(20), Article 20. https://doi.org/10.3390/en15207637

Clayton, J. L., Spencer, C. W., & Koncz, I. (1994). Tótkomlós–Szolnok(.) Petroleum System of Southeastern Hungary. In L. B. Magoon & W. G. Dow (Eds.), The Petroleum System—From Source to Trap (Vol. 60, p. 0). American Association of Petroleum Geologists. https://doi.org/10.1306/M60585C35

Collignon, M., Klemetsdal, Ø. S., Møyner, O., Alcanié, M., Rinaldi, A. P., Nilsen, H., & Lupi, M. (2020). Evaluating thermal losses and storage capacity in high-temperature aquifer thermal energy storage (HT-ATES) systems with well operating limits: Insights from a study-case in the Greater Geneva Basin, Switzerland. Geothermics, 85, 101773. https://doi.org/10.1016/j.geothermics.2019.101773

Corbett, P., Ellabad, Y., Mohammed, K., & Pososyaev, A. (2003). Global Hydraulic Elements – Elementary Petrophysics for Reduced Reservoir Modelling. cp. https://doi.org/10.3997/2214-4609-pdb.6.F26

Czauner, B., & Mádl-Szőnyi, J. (2013). Regional hydraulic behavior of structural zones and sedimentological heterogeneities in an overpressured sedimentary basin. Marine and Petroleum Geology, 48, 260–274. https://doi.org/10.1016/j.marpetgeo.2013.08.016

Czauner, B., Szabó, Z., Márton, B., & Mádl-Szőnyi, J. (2023). Basin-Scale Hydraulic Evaluation of Groundwater Flow Controlled Biogenic Gas Migration and Accumulation in the Central Pannonian Basin. Water, 15(18), Article 18. https://doi.org/10.3390/w15183272

Das, A. (2025). A data-driven approach utilizing machine learning (ML) and geographical information system (GIS)-based time series analysis with data augmentation for water quality assessment in Mahanadi River Basin, Odisha, India. Discover Sustainability, 6(1), 555. https://doi.org/10.1007/s43621-025-01464-7

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data Via the EM Algorithm. Journal of the Royal Statistical Society: Series B (Methodological), 39(1), 1–22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

Dickinson, J. S., Buik, N., Matthews, M. C., & Snijders, A. (2009). Aquifer thermal energy storage: Theoretical and operational analysis. Géotechnique, 59(3), 249–260. https://doi.org/10.1680/geot.2009.59.3.249

Digitalization of Legacy Datasets and Machine Learning Regression Yields Insights for Reservoir Property Prediction and Submarine-Fan Evolution: A Subsurface Example From the Lewis Shale, Wyoming | Published in The Sedimentary Record. (n.d.). Retrieved July 5, 2025, from https://thesedimentaryrecord.scholasticahq.com/article/36638

Dincer, I., & Rosen, M. A. (2011). Thermal energy storage: Systems and applications (2. ed). Wiley.

Djebbas, F., Ameur-Zaimeche, O., Kechiched, R., Heddam, S., Wood, D. A., & Movahed, Z. (2023). Integrating hydraulic flow unit concept and adaptive neuro-fuzzy inference system to accurately estimate permeability in heterogeneous reservoirs: Case study Sif Fatima oilfield, southern Algeria. Journal of African Earth Sciences, 206, 105027. https://doi.org/10.1016/j.jafrearsci.2023.105027

Drijver, B., Aarssen, M. van, & Zwart, B. de. (2012). High-temperature aquifer thermal energy storage (HT-ATES): Sustainable and multi-usable. The 12th International Conference on Energy Storage.

Duggal, R., Rayudu, R., Hinkley, J., Burnell, J., Wieland, C., & Keim, M. (2022). A comprehensive review of energy extraction from low-temperature geothermal resources in hydrocarbon fields. Renewable and Sustainable Energy Reviews, 154, 111865. https://doi.org/10.1016/j.rser.2021.111865

Duplyakin, D., Beckers, K. F., Siler, D. L., Martin, M. J., & Johnston, H. E. (2022). Modeling Subsurface Performance of a Geothermal Reservoir Using Machine Learning. Energies, 15(3), Article 3. https://doi.org/10.3390/en15030967

Egbele, E., Ezuka, I., & Onyekonwu, M. (2005, August 1). Net-To-Gross Ratios: Implications in Integrated Reservoir Management Studies. Nigeria Annual International Conference and Exhibition. https://doi.org/10.2118/98808-MS

Elabd, M. A., González-Guzmán, R., Weber, B., Solís, C., Bernard-Romero, R., Velasco-Tapia, F., & Marín-Camacho, P. (2022). Determination of permeability data and 3-D modelling of the host rock and sinters from a geothermal field: Los Geysers, northern Trans-Mexican Volcanic Field. Data in Brief, 45, 108637. https://doi.org/10.1016/j.dib.2022.108637

Elbalawy, M. A., Balash, M., Eid, M. H., Takács, E., & Velledits, F. (2025). Innovative method integrates play fairway analysis supported with GIS and seismic modeling for geothermal potential evaluation in a basement reservoir. Scientific Reports, 15(1), 1325. https://doi.org/10.1038/s41598-024-79943-6

El-Gendy, N. H., Mabrouk, W. M., Waziry, M. A., Dodd, T. J., Abdalla, F. A., Alexakis, D. E., & Barakat, M. K. (2023). An Integrated Approach for Saturation Modeling Using Hydraulic Flow Units: Examples from the Upper Messinian Reservoir. Water, 15(24), Article 24. https://doi.org/10.3390/w15244204

Fadel, M., Reinecker, J., Bruss, D., & Moeck, I. (2022). Causes of a premature thermal breakthrough of a hydrothermal project in Germany. Geothermics, 105, 102523. https://doi.org/10.1016/j.geothermics.2022.102523

Farkas, M. P., Magyar, G., Hofmann, H., Zimmermann, G., & Fedor, F. (2025). Comparison of stimulation techniques in a geothermal injection well in a sedimentary aquifer in Szentes, Hungary. Geothermal Energy, 13(1), 20. https://doi.org/10.1186/s40517-025-00344-7

Fathy, M., Reda, M., Abdel-Fattah, M. I., Raef, A., & Al-Hashim, M. H. (2025). Seismic and petrophysical characterization of the Sequoia Pliocene channel reservoirs: Insights into hydrocarbon potential in the West Delta Deep Marine, Egypt. Petroleum Geoscience, 31(3), petgeo2024-070. https://doi.org/10.1144/petgeo2024-070

Fisher, A., Rudin, C., & Dominici, F. (2019). All Models are Wrong, but Many are Useful: Learning a Variable's Importance by Studying an Entire Class of Prediction Models Simultaneously (arXiv:1801.01489). arXiv. https://doi.org/10.48550/arXiv.1801.01489

Fleuchaus, P., Godschalk, B., Stober, I., & Blum, P. (2018). Worldwide application of aquifer thermal energy storage – A review. Renewable and Sustainable Energy Reviews, 94, 861–876. https://doi.org/10.1016/j.rser.2018.06.057

Fleuchaus, P., Schüppler, S., Bloemendal, M., Guglielmetti, L., Opel, O., & Blum, P. (2020). Risk analysis of High-Temperature Aquifer Thermal Energy Storage (HT-ATES). Renewable and Sustainable Energy Reviews, 133, 110153. https://doi.org/10.1016/j.rser.2020.110153

Fu, L., Wang, X., Zhong, X., & Cao, X. (2025). Technology and Applications for the Interpretation and Modeling of Advanced Sand Body Architectures. Energies, 18(3), Article 3. https://doi.org/10.3390/en18030504

Gao, H., Wang, Y., Ye, J., Chen, A., Wen, M., Wang, D., Du, B., Li, S., & Cui, Q. (2025). Machine learning-based prediction of laser ablation impulse from plasma spectra. Physics of Plasmas, 32(6), 063301. https://doi.org/10.1063/5.0257411

Gayayev, I. (2023). Conversion of Abandoned Hydrocarbon Structures into Geothermal Wells for Sustainable Energy Production in Sedimentary Basins [Laurea, Politecnico di Torino]. https://webthesis.biblio.polito.it/29039/

Ge, Y., Thomasson, J. A., Sui, R., & Wooten, J. (2011). Regression-kriging for characterizing soils with remotesensing data. Frontiers of Earth Science. https://doi.org/10.1007/s11707-011-0174-1

Gelhar, L. W., Welty, C., & Rehfeldt, K. R. (1992). A Critical Review of Data on Field-Scale Dispersion in Aquifers. Water Resources Research, 28(7), 1955–1974. https://doi.org/10.1029/92WR00607

Geofizika M. (2011). Magyar geofizika 52. Évf. 2. Sz. (2011.).

Geostatistical modeling of porosity and evaluating the local and global distribution | Journal of Petroleum Exploration and Production Technology. (n.d.). Retrieved July 20, 2025, from https://link.springer.com/article/10.1007/s13202-021-01308-w

Green, S., McLennan, J., Panja, P., Kitz, K., Allis, R., & Moore, J. (2021). Geothermal battery energy storage. Renewable Energy, 164, 777–790. https://doi.org/10.1016/j.renene.2020.09.083

Haas, J. (Ed.). (2013). Geology of Hungary. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-21910-8

Haffen, S., Géraud, Y., Rosener, M., & Diraison, M. (2017). Thermal conductivity and porosity maps for different materials: A combined case study of granite and sandstone. Geothermics, 66, 143–150. https://doi.org/10.1016/j.geothermics.2016.12.005

Harbaugh, A. W. (2005). MODFLOW-2005: The U.S. Geological Survey modular ground-water model--the ground-water flow process. Techniques and Methods, Article 6-A16. https://doi.org/10.3133/tm6A16

HEATSTORE. (2025). https://www.heatstore.eu/national-project-netherlands.html

Holstenkamp, L., Meisel, M., Neidig, P., Opel, O., Steffahn, J., Strodel, N., Lauer, J. J., Vogel, M., Degenhart, H., Michalzik, D., Schomerus, T., Schönebeck, J., & Növig, T. (2017). Interdisciplinary Review of Medium-deep Aquifer Thermal Energy Storage in North Germany. Energy Procedia, 135, 327–336. https://doi.org/10.1016/j.egypro.2017.09.524

Horváth, F., Musitz, B., Balázs, A., Végh, A., Uhrin, A., Nádor, A., Koroknai, B., Pap, N., Tóth, T., & Wórum, G. (2015). Evolution of the Pannonian basin and its geothermal resources. Geothermics, 53, 328–352. https://doi.org/10.1016/j.geothermics.2014.07.009

Hosamo, H., & Mazzetto, S. (2025). Performance Evaluation of Machine Learning Models for Predicting Energy Consumption and Occupant Dissatisfaction in Buildings. Buildings, 15(1), Article 1. https://doi.org/10.3390/buildings15010039

IEA. (2017). Energy Technology Perspectives 2017 – Analysis. IEA. https://www.iea.org/reports/energy-technology-perspectives-2017

Iglesias, F., Massana, J., Burgas, L., Planellas, N., & Colomer, J. (2025). Methodological Advances in Temperature Dynamics Modeling for Energy-Efficient Indoor Air Management Systems. Applied Sciences, 15(8), Article 8. https://doi.org/10.3390/app15084291

IPCC. (2021). Climate change widespread, rapid, and intensifying – IPCC — IPCC. https://www.ipcc.ch/2021/08/09/ar6-wg1-20210809-pr/

Ishikawa, T., Morita, A., Fukushima, T., & Ono, H. (2014). Three-Dimensional Cerebral Aneurysm Models for Surgical Simulation and Education—Development of Aneurysm Models with Perforating Arteries and for Application of Fenestrated Clips. Open Journal of Modern Neurosurgery, 04(02), 59–63. https://doi.org/10.4236/ojmn.2014.42013

Jello, J., Khan, M., Malkewicz, N., Whittaker, S., & Baser, T. (2022). Advanced geothermal energy storage systems by repurposing existing oil and gas wells: A full-scale experimental and numerical investigation. Renewable Energy, 199, 852–865. https://doi.org/10.1016/j.renene.2022.07.145

Jin, W., Atkinson, T. A., Doughty, C., Neupane, G., Spycher, N., McLing, T. L., Dobson, P. F., Smith, R., & Podgorney, R. (2022). Machine-learning-assisted high-temperature reservoir thermal energy storage optimization. Renewable Energy, 197, 384–397. https://doi.org/10.1016/j.renene.2022.07.118

Juhász, G. (1991). Lithostratigraphical and sedimentological framework of the Pannonian (sl) sedimentary sequence in the Hungarian Plain (Alföld), Eastern Hungary. Acta Geologica Hungarica, 34(1–2), 53–72.

Kang, M., Mauzerall, D. L., Ma, D. Z., & Celia, M. A. (2019). Reducing methane emissions from abandoned oil and gas wells: Strategies and costs. Energy Policy, 132, 594–601. https://doi.org/10.1016/j.enpol.2019.05.045

Kassab, M. A., Abbas, A. E., Osman, I. A., & Eid, A. A. (2024). Reservoir rock typing for optimum permeability prediction of Nubia formation in October Field, Gulf of Suez, Egypt. Journal of Petroleum Exploration and Production Technology, 14(6), 1395–1416. https://doi.org/10.1007/s13202-024-01774-y

Kastner, O., Norden, B., Klapperer, S., Park, S., Urpi, L., Cacace, M., & Blöcher, G. (2017). Thermal solar energy storage in Jurassic aquifers in Northeastern Germany: A simulation study. Renewable Energy, 104, 290–306. https://doi.org/10.1016/j.renene.2016.12.003

Khosravi, R., Simjoo, M., & Chahardowli, M. (2024). A new insight into pilot-scale development of low-salinity polymer flood using an intelligent-based proxy model coupled with particle swarm optimization. Scientific Reports, 14(1), 29000. https://doi.org/10.1038/s41598-024-78210-y

Kim, J., Lee, Y., Yoon, W. S., Jeon, J. S., Koo, M.-H., & Keehm, Y. (2010). Numerical modeling of aquifer thermal energy storage system. Energy, 35(12), 4955–4965. https://doi.org/10.1016/j.energy.2010.08.029

Ko, S.-K., Fontane, D. G., & Margeta, J. (1994). Multiple reservoir system operational planning using multi-criterion decision analysis. European Journal of Operational Research, 76(3), 428–439. https://doi.org/10.1016/0377-2217(94)90278-X

Korim K. (1991). A szentesi hévízmező feltárásának és termelésének három évtizedes története (Three decades of research and utilization in the Szentes Geothermal Field). Bányászati és Kohászati Lapok, 124(6), 179–184

Koroncz, P., Vizhányó, Z., Farkas, M. P., Kuncz, M., Ács, P., Kocsis, G., Mucsi, P., Fedorné Szász, A., Fedor, F., & Kovács, J. (2022). Experimental Rock Characterisation of Upper Pannonian Sandstones from Szentes Geothermal Field, Hungary. Energies, 15(23), Article 23. https://doi.org/10.3390/en15239136

Kovács, A., & Teleki, P. G. (1994). History of Oil and Natural Gas Production in the Békés Basin. In P. G. Teleki, R. E. Mattick, & J. Kókai (Eds.), Basin Analysis in Petroleum Exploration (pp. 237–256). Springer Netherlands. https://doi.org/10.1007/978-94-011-0954-3_12

Kun É., Zilahi-Sebess L., & Szanyi J. (2022). Battonya–Pusztaföldvári-hát térségének nagy entalpiájú geotermikusenergia-vagyona (I. rész): Hidrodinamikai és hőtranszportmodell. Földtani Közlöny, 152(1), 53–75. https://doi.org/10.23928/foldt.kozl.2022.152.1.53

Lee, K. S. (2013). Underground thermal energy storage. Springer.

Lenkey, L., Mihályka, J., & Paróczi, P. (2021). Review of geothermal conditions of Hungary. Földtani Közlöny, 151(1), 65. https://doi.org/10.23928/foldt.kozl.2021.151.1.65

Li, G. (2016). Sensible heat thermal storage energy and exergy performance evaluations. Renewable and Sustainable Energy Reviews, 53, 897–923. https://doi.org/10.1016/j.rser.2015.09.006

Li, Q., Li, Q., Wang, F., Xu, N., Wang, Y., & Bai, B. (2025). Settling behavior and mechanism analysis of kaolinite as a fracture proppant of hydrocarbon reservoirs in CO2 fracturing fluid. Colloids and Surfaces A: Physicochemical and Engineering Aspects, 724, 137463. https://doi.org/10.1016/j.colsurfa.2025.137463

Liu, A., Li, J., Bi, J., Chen, Z., Wang, Y., Lu, C., Jin, Y., & Lin, B. (2024). A novel reservoir simulation model based on physics-informed neural networks. Physics of Fluids, 36(11), 116617. https://doi.org/10.1063/5.0239376

Liu, H., Wu, Y., Cao, Y., Lv, W., Han, H., Li, Z., & Chang, J. (2020). Well Logging Based Lithology Identification Model Establishment Under Data Drift: A Transfer Learning Method. Sensors, 20(13), Article 13. https://doi.org/10.3390/s20133643

Mádl-Szőnyi, J., & Tóth, J. (2009). A hydrogeological type section for the Duna-Tisza Interfluve, Hungary. Hydrogeology Journal, 17, 961–980. https://doi.org/10.1007/s10040-008-0421-z

Mahbaz, S. B., Yaghoubi, A., Dehghani-Sanij, A., Sarvaramini, E., Leonenko, Y., & Dusseault, M. B. (2021). Well-Doublets: A First-Order Assessment of Geothermal SedHeat Systems. Applied Sciences, 11(2), Article 2. https://doi.org/10.3390/app11020697

Mahjour, S. K., Saleh, A., & Mahjour, S. S. (2025). Dimension-Adaptive Machine Learning for Efficient Uncertainty Quantification in Geological Carbon Storage Models. Processes, 13(6), Article 6. https://doi.org/10.3390/pr13061834

Malvić, T., Ivšinović, J., Velić, J., Sremac, J., & Barudžija, U. (2020). Application of the Modified Shepard's Method (MSM): A Case Study with the Interpolation of Neogene Reservoir Variables in Northern Croatia. Stats, 3(1), Article 1. https://doi.org/10.3390/stats3010007

Man, H. Q., Hien, D. H., Thong, K. D., Dung, B. V., Hoa, N. M., Hoa, T. K., Kieu, N. V., & Ngoc, P. Q. (2021). Hydraulic Flow Unit Classification and Prediction Using Machine Learning Techniques: A Case Study from the Nam Con Son Basin, Offshore Vietnam. Energies, 14(22), Article 22. https://doi.org/10.3390/en14227714

Manikkuwahandi, T. D. S., Hornberger, G. M., & Baroud, H. (2019). Decision Analysis for Expansion of Mahaweli Multipurpose Reservoir System in Sri Lanka. Journal of Water Resources Planning and Management, 145(9), 05019013. https://doi.org/10.1061/(ASCE)WR.1943-5452.0001094

Matos, C. R., Carneiro, J. F., & Silva, P. P. (2019). Overview of Large-Scale Underground Energy Storage Technologies for Integration of Renewable Energies and Criteria for Reservoir Identification. Journal of Energy Storage, 21, 241–258. https://doi.org/10.1016/j.est.2018.11.023

Mehmood, A., Yao, J., Fan, D., Bongole, K., Liu, J., & Zhang, X. (2019). Potential for heat production by retrofitting abandoned gas wells into geothermal wells. PLOS ONE, 14(8), e0220128. https://doi.org/10.1371/journal.pone.0220128

Mellal, I., Latrach, A., Rasouli, V., Bakelli, O., Dehdouh, A., & Ouadi, H. (2023). Water Saturation Prediction in the Middle Bakken Formation Using Machine Learning. Eng, 4(3), Article 3. https://doi.org/10.3390/eng4030110

Menear, K., Duplyakin, D., Oliver, M. C., Shah, M., Martin, M. J., Martinek, J., Nithyanandam, K., & Ma, Z. (2023). One System, Many Models: Designing a Surrogate Model for Sulfur Thermal Energy Storage: Preprint. Renewable Energy.

Mohammadian, E., Kheirollahi, M., Liu, B., Ostadhassan, M., & Sabet, M. (2022). A case study of petrophysical rock typing and permeability prediction using machine learning in a heterogenous carbonate reservoir in Iran. Scientific Reports, 12(1), 4505. https://doi.org/10.1038/s41598-022-08575-5

Mohammadpour, M., Roshan, H., Arashpour, M., & Masoumi, H. (2024). Machine learning assisted Kriging to capture spatial variability in petrophysical property modelling. Marine and Petroleum Geology, 167, 106967. https://doi.org/10.1016/j.marpetgeo.2024.106967

ModelMuse: A Graphical User Interface for Groundwater Models | U.S. Geological Survey. (2024). https://www.usgs.gov/software/modelmuse-a-graphical-user-interface-groundwater-mode

Nabawy, B. S., Basal, A. M. K., Sarhan, M. A., & Safa, M. G. (2018). Reservoir zonation, rock typing and compartmentalization of the Tortonian-Serravallian sequence, Temsah Gas Field, offshore Nile Delta, Egypt. Marine and Petroleum Geology, 92, 609–631. https://doi.org/10.1016/j.marpetgeo.2018.03.030

Nádor, A. (2022). De-risking geothermal exploration by public data portals and financial incentives – good examples from Hungary.

Nádor, A., & Lenkey, L. (2023). Geothermal Resources of Hungary: A Play-Based Review.

Nádor, A., Kujbus, A., & Tóth, A. (2022). Geothermal Energy Use, Country Update for Hungary. European Geothermal Congress 2022, 1–13.

Nagymarosy, A., & Båldi-Beke, M. (1993). The Szolnok unit and its probable paleogeographic position. Tectonophysics, 226(1), 457–470. https://doi.org/10.1016/0040-1951(93)90132-4

Nieminen, P. (2022). Application of Standardized Regression Coefficient in Meta-Analysis. BioMedInformatics, 2(3), Article 3. https://doi.org/10.3390/biomedinformatics2030028

Njeru, R. M., Sofyan, A., Halisch, M., Kóbor, B., & Szanyi, J. (2024). Optimizing Micro-CT Resolution for Geothermal Reservoir Characterization in the Pannonian Basin. Energies, 17(13), Article 13. https://doi.org/10.3390/en17133081

Oerlemans, P. (n.d.). Performance of HT-ATES in practice.

Oerlemans, P., Drijver, B., Koenen, M., Koornneef, J., Dinkelman, D., & Godschalk, B. (2022). First field results on the technical risks and effectiveness of mitigation measures for the full scale HT-ATES demonstration project in Middenmeer.

Osundare, O., Teodoriu, C., Falcone, G., & Ichim, A. (2018). Estimation of Plugging and Abandonment Costs Based on Different EU Regulations with Application to Geothermal Wells.

Otchere, D. A., Hodgetts, D., Ganat, T. A. O., Ullah, N., & Rashid, A. (2021, August 9). Static Reservoir Modeling Comparing Inverse Distance Weighting to Kriging Interpolation Algorithm in Volumetric Estimation. Case Study: Gullfaks Field. Offshore Technology Conference. https://doi.org/10.4043/30919-MS

Otmane, M., Imtiaz, S., Jaluta, A. M., & Aborig, A. (2025). Boosting Reservoir Prediction Accuracy: A Hybrid Methodology Combining Traditional Reservoir Simulation and Modern Machine Learning Approaches. Energies, 18(3), Article 3. https://doi.org/10.3390/en18030657

Paksoy, H. O., Andersson, O., Abaci, S., Evliya, H., & Turgut, B. (2000). Heating and cooling of a hospital using solar energy coupled with seasonal thermal energy storage in an aquifer. Renewable Energy, 19(1–2), 117–122. https://doi.org/10.1016/S0960-1481(99)00060-9

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (2011). Scikit-learn: Machine Learning in Python. MACHINE LEARNING IN PYTHON.

Pellegrini, M., Bloemendal, M., Hoekstra, N., Spaak, G., Andreu Gallego, A., Rodriguez Comins, J., Grotenhuis, T., Picone, S., Murrell, A. J., & Steeman, H. J. (2019). Low carbon heating and cooling by combining various technologies with Aquifer Thermal Energy Storage. Science of The Total Environment, 665, 1–10. https://doi.org/10.1016/j.scitotenv.2019.01.135

Permeability and Mineralogy of the Újfalu Formation, Hungary, from Production Tests and Experimental Rock Characterization: Implications for Geothermal Heat Projects. (n.d.). Retrieved July 5, 2025, from https://www.mdpi.com/1996-1073/14/14/4332

Project Thermo-Hydro-Mechanical Description – Geothermal Energy and Geofluids. (n.d.). Retrieved July 5, 2025, from https://geg.ethz.ch/project-thermo_hydro-description/

Qin, J., Song, J., Tang, Y., Rui, Z., Wang, Y., & He, Y. (2023). Well applicability assessment based on fuzzy theory for CO2 sequestration in depleted gas reservoirs. Renewable Energy, 206, 239–250. https://doi.org/10.1016/j.renene.2023.01.090

Qingchao, L., Jingjuan, W., Qiang, L., Fuling, W., & Yuanfang, C. (2025). Sediment Instability Caused by Gas Production from Hydrate-Bearing Sediment in Northern South China Sea by Horizontal Wellbore: Sensitivity Analysis. Natural Resources Research, 34(3), 1667–1699. https://doi.org/10.1007/s11053-025-10478-x

Raimi, D., Krupnick, A. J., Shah, J.-S., & Thompson, A. (2021). Decommissioning Orphaned and Abandoned Oil and Gas Wells: New Estimates and Cost Drivers. Environmental Science & Technology, 55(15), 10224–10230. https://doi.org/10.1021/acs.est.1c02234

Rangel-Jurado, N., Hawkins, A. J., & Fulton, P. M. (2023). Influence of extreme fracture flow channels on the thermal performance of open-loop geothermal systems at commercial scale. Geothermal Energy, 11(1), 19. https://doi.org/10.1186/s40517-023-00261-7

REN21. (2019). RENEWABLES 2019 GLOBAL STATUS REPORT. https://www.ren21.net/gsr-2019

Réveillère, A., Hamm, V., Lesueur, H., Cordier, E., & Goblet, P. (2013). Geothermal contribution to the energy mix of a heating network when using Aquifer Thermal Energy Storage: Modeling and application to the Paris basin. Geothermics, 47, 69–79. https://doi.org/10.1016/j.geothermics.2013.02.005

Rohmer, J., Armandine Les Landes, A., Loschetter, A., & Maragna, C. (2023). Fast prediction of aquifer thermal energy storage: A multicyclic metamodelling procedure. Computational Geosciences, 27(2), 223–243. https://doi.org/10.1007/s10596-023-10192-8

Sandstone Heterogeneity—3-D Reservoir Characterization. (n.d.). Retrieved July 4, 2025, from https://pubs.usgs.gov/dds/dds-033/USGS_3D/ssx_txt/heterog.htm

Sendula, E., Albrecht, R., Conceição de Castro, C., Keresztény-Borbás, E., Szabó-Krausz, Z., & Kovács, J. (2025). Characteristics of Heat-Depleted Thermal Water Re-Injection-Induced Water–Rock Interactions in a Sandstone Reservoir Containing Carbonate and Silicate Minerals (Szentes, Hungary). Minerals, 15(5), Article 5. https://doi.org/10.3390/min15050513

Shen, J., Chen, L., Wu, Y., & Jing, N. (2018). Approach to Accelerating Dissolved Vector Buffer Generation in Distributed In-Memory Cluster Architecture. ISPRS International Journal of Geo-Information, 7(1), Article 1. https://doi.org/10.3390/ijgi7010026

Shi, Y., Cui, Q., Song, X., Liu, S., Yang, Z., Peng, J., Wang, L., & Guo, Y. (2023). Thermal performance of the aquifer thermal energy storage system considering vertical heat losses through aquitards. Renewable Energy, 207, 447–460. https://doi.org/10.1016/j.renene.2023.03.044

Shirinkin, D., Kochnev, A., Krivoshchekov, S., Putilov, I., Botalov, A., Kozyrev, N., & Ozhgibesov, E. (2024). High Permeability Streak Identification and Modelling Approach for Carbonate Reef Reservoir. Energies, 17(1), Article 1. https://doi.org/10.3390/en17010236

Singh, M., Mahmoodpour, S., Schmidt-Hattenberger, C., Sass, I., & Drews, M. (2024). Influence of Reservoir Heterogeneity on Simultaneous Geothermal Energy Extraction and CO2 Storage. Sustainability, 16(1), Article 1. https://doi.org/10.3390/su16010387

Soldo, E., Alimonti, C., & Scrocca, D. (2020). Geothermal Repurposing of Depleted Oil and Gas Wells in Italy. Proceedings, 58(1), Article 1. https://doi.org/10.3390/WEF-06907

Sommer, W., Valstar, J., van Gaans, P., Grotenhuis, T., & Rijnaarts, H. (2013). The impact of aquifer heterogeneity on the performance of aquifer thermal energy storage. Water Resources Research, 49(12), 8128–8138. https://doi.org/10.1002/2013WR013677

Steventon, M. J., Jackson, C. A.-L., Johnson, H. D., Hodgson, D. M., Kelly, S., Omma, J., Gopon, C., Stevenson, C., & Fitch, P. (2021). Evolution of a sand-rich submarine channel–lobe system, and the impact of mass-transport and transitional-flow deposits on reservoir heterogeneity: Magnus Field, Northern North Sea. Petroleum Geoscience, 27(3), petgeo2020-095. https://doi.org/10.1144/petgeo2020-095

Stricker, K., Egert, R., Schill, E., & Kohl, T. (2024). Risk of surface movements and reservoir deformation for high-temperature aquifer thermal energy storage (HT-ATES). Geothermal Energy, 12(1), 4. https://doi.org/10.1186/s40517-024-00283-9

Stricker, K., Grimmer, J. C., Egert, R., Bremer, J., Korzani, M. G., Schill, E., & Kohl, T. (2020). The Potential of Depleted Oil Reservoirs for High-Temperature Storage Systems. Energies, 13(24), Article 24. https://doi.org/10.3390/en13246510

Survey, U. S. G. (2023). Mineral commodity summaries 2023. https://doi.org/10.3133/mcs2023

Svirsky, D., Ryazanov, A., Pankov, M., Corbett, P. W., & Posysoev, A. (2004, March 29). Hydraulic Flow Units Resolve Reservoir Description Challenges in a Siberian Oil Field. SPE Asia Pacific Conference on Integrated Modelling for Asset Management. https://doi.org/10.2118/87056-MS

Swan, A. (1999). DEUTSCH, C. V. & JOURNEL, A. G. 1998. GSLIB. Geostatistical Software Library and User's Guide, 2nd ed. x + 369 pp. + CD-ROM. Oxford, New York: Oxford University Press. Price £45.00 (hard covers). ISBN 0 19 510015 8. Geological Magazine, 136(1), 83–108. https://doi.org/10.1017/S0016756899531774

Szanyi, J. D., Kovács, B., & Abdulhaq, H. A. (2025). Harnessing geothermal energy in Hungary. Geological Society, London, Special Publications, 555(1), SP555-2024–22. https://doi.org/10.1144/SP555-2024-22

Szanyi, J., Kovacs, B., & Scharek, P. (2009). Geothermal energy in Hungary: Potentials and barriers. Eur. Geol., 27, 15–19.

Szanyi, J., Medgyes, T., Kóbor, B., & Pál-Molnár, E. (2015). Technologies of injection into sandstone reservoirs: Best practices, case studies. GeoLitera; Institute of Geosciences, University of Szeged. https://publicatio.bibl.u-szeged.hu/13471/

Szanyi, J., Rybach, L., & Abdulhaq, H. A. (2023). Geothermal Energy and Its Potential for Critical Metal Extraction—A Review. Energies, 16(20), 7168. https://doi.org/10.3390/en16207168

Tang, D. W. S., & Rijnaarts, H. H. M. (2023). Dimensionless Thermal Efficiency Analysis for Aquifer Thermal Energy Storage. Water Resources Research, 59(11), e2023WR035797. https://doi.org/10.1029/2023WR035797

Tari, G., Bada, G., Boote, D. R. D., Krézsek, C., Koroknai, B., Kovács, G., Lemberkovics, V., Sachsenhofer, R. F., & Tóth, T. (2023). The Pannonian Super Basin: A brief overview. AAPG Bulletin, 107(8), 1391–1417. https://doi.org/10.1306/02172322098

Tas, L., Simpson, D., & Hermans, T. (2023). Assessing the potential of low transmissivity aquifers for ATES systems: A case study in Flanders (Belgium) [Preprint]. Geology. https://doi.org/10.31223/X53371

The Importance of Physiochemical Processes in Decarbonisation Technology Applications Utilizing the Subsurface: A Review | Earth Science, Systems and Society. (n.d.). Retrieved July 5, 2025, from https://www.lyellcollection.org/doi/full/10.3389/esss.2022.10043

Todorov, O., Alanne, K., Virtanen, M., & Kosonen, R. (2020). A method and analysis of aquifer thermal energy storage (ATES) system for district heating and cooling: A case study in Finland. Sustainable Cities and Society, 53, 101977. https://doi.org/10.1016/j.scs.2019.101977

Topór, T., Słota-Valim, M., & Kudrewicz, R. (2023). Assessing the Geothermal Potential of Selected Depleted Oil and Gas Reservoirs Based on Geological Modeling and Machine Learning Tools. Energies, 16(13), Article 13. https://doi.org/10.3390/en16135211

Tóth, J., & Almási, I. (2001). Interpretation of observed fluid potential patterns in a deep sedimentary basin under tectonic compression: Hungarian Great Plain, Pannonian Basin. Geofluids, 1(1), 11–36. https://doi.org/10.1046/j.1468-8123.2001.11004.x

USGS. (2022). Techniques and Methods (Techniques and Methods) [Techniques and Methods].

Van Der Roest, E., Beernink, S., Hartog, N., Van Der Hoek, J. P., & Bloemendal, M. (2021). Towards Sustainable Heat Supply with Decentralized Multi-Energy Systems by Integration of Subsurface Seasonal Heat Storage. Energies, 14(23), 7958. https://doi.org/10.3390/en14237958

Varga, A., Bozsó, G., Garaguly, I., Raucsik, B., Bencsik, A., & Kóbor, B. (2019). Cements, Waters, and Scales: An Integrated Study of the Szeged Geothermal Systems (SE Hungary) to Characterize Natural Environmental Conditions of the Thermal Aquifer. Geofluids, 2019, 1–21. https://doi.org/10.1155/2019/4863814

Varouchakis, E. A. (2021). Median Polish Kriging and Sequential Gaussian Simulation for the Spatial Analysis of Source Rock Data. Journal of Marine Science and Engineering, 9(7), Article 7. https://doi.org/10.3390/jmse9070717

Vass, I. (2025). Well abandonment cost estimates for Hungary (personal communication).

Vass, I., Tóth, T. M., Szanyi, J., & Kovács, B. (2018). Hybrid numerical modelling of fluid and heat transport between the overpressured and gravitational flow systems of the Pannonian Basin. Geothermics, 72, 268–276. https://doi.org/10.1016/j.geothermics.2017.11.013

Wackernagel, H. (1995). Multivariate Geostatistics. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-03098-1

WEF. (2021). 5 green energy milestones from around the world. World Economic Forum. https://www.weforum.org/agenda/2021/04/renewables-record-capacity-solar-wind-nuclear/

Wesselink, M., Liu, W., Koornneef, J., & van den Broek, M. (2018). Conceptual market potential framework of high temperature aquifer thermal energy storage—A case study in the Netherlands. Energy, 147, 477–489. https://doi.org/10.1016/j.energy.2018.01.072

Winterleitner, G., Schütz, F., Wenzlaff, C., & Huenges, E. (2018). The Impact of Reservoir Heterogeneities on High-Temperature Aquifer Thermal Energy Storage Systems. A Case Study from Northern Oman. Geothermics, 74, 150–162. https://doi.org/10.1016/j.geothermics.2018.02.005

World Energy Outlook 2021 – Analysis. (2021). IEA. https://www.iea.org/reports/world-energy-outlook-2021

Xing, Y., Wang, H., Zhang, L., Cheng, M., Shi, H., Guo, C., Chen, P., & Yu, W. (2022). Depositional and Diagenetic Controls on Reservoir Quality of Callovian-Oxfordian Stage on the Right Bank of Amu Darya. Energies, 15(19), Article 19. https://doi.org/10.3390/en15196923

Yalcin, M., & Kilic Gul, F. (2017). A GIS-based multi criteria decision analysis approach for exploring geothermal resources: Akarcay basin (Afyonkarahisar). Geothermics, 67, 18–28. https://doi.org/10.1016/j.geothermics.2017.01.002

Zarei, M., Bozorg-Haddad, O., Baghban, S., Delpasand, M., Goharian, E., & Loáiciga, H. A. (2021). Machine-learning algorithms for forecast-informed reservoir operation (FIRO) to reduce flood damages. Scientific Reports, 11(1), 24295. https://doi.org/10.1038/s41598-021-03699-6

Zecchin, M., & Catuneanu, O. (2015). High-resolution sequence stratigraphy of clastic shelves III: Applications to reservoir geology. Marine and Petroleum Geology, 62, 161–175. https://doi.org/10.1016/j.marpetgeo.2014.08.025

Zeghici, R. M., Oude Essink, G. H. P., Hartog, N., & Sommer, W. (2015). Integrated assessment of variable density–viscosity groundwater flow for a high temperature mono-well aquifer thermal energy storage (HT-ATES) system in a geothermal reservoir. Geothermics, 55, 58–68. https://doi.org/10.1016/j.geothermics.2014.12.006

Zhang, J., Wang, R., Jia, A., & Feng, N. (2024). Optimization and Application of XGBoost Logging Prediction Model for Porosity and Permeability Based on K-means Method. Applied Sciences, 14(10), Article 10. https://doi.org/10.3390/app14103956

Zhao, N. (2021). An Efficient Downscaling Scheme for High-Resolution Precipitation Estimates over a High Mountainous Watershed. Remote Sensing, 13(2), Article 2. https://doi.org/10.3390/rs13020234

Zhu, S., Shi, X., Yang, C., Bai, W., Wei, X., Yang, K., Li, P., Li, H., Li, Y., & Wang, G. (2024). Site selection evaluation for salt cavern hydrogen storage in China. Renewable Energy, 120143. https://doi.org/10.1016/j.renene.2024.120143

DECLARATION OF THE SUPERVISOR

I hereby confirm that the content of the dissertation is based on the independent work of the doctoral candidate and that he has contributed decisively to the results through his independent creative activity. I consider the entire dissertation to be eligible for support from a professional and academic point of view and recommend its acceptance.

Szeged, 16 Feb 2026

János Szanyi, Dr.

Supervisor