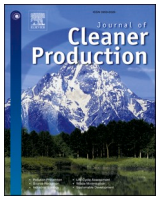
Journal of Cleaner Production 460 (2024) 142452

Contents lists available at ScienceDirect

Journal of Cleaner Production

journal homepage: www.elsevier.com/locate/jclepro



Artificial intelligence applications for accurate geothermal temperature prediction in the lower Friulian Plain (north-eastern Italy)

Danial Sheini Dashtgoli [a,b,*], Michela Giustiniani [b], Martina Busetti [b], Claudia Cherubini [a,b]

[a] Department of Mathematics, Informatics, and Geosciences, University of Trieste, Italy
[b] National Institute of Oceanography and Applied Geophysics - OGS, Sgonico, Trieste, Italy

ARTICLE INFO

Handling Editor: Zhen Leng

Keywords:
Geothermal resources
Aquifer temperature
Machine learning
XGBoost
Carbonate reservoir

ABSTRACT

Geothermal energy as a sustainable and clean energy source depends on the accurate estimation of reservoir temperatures. Understanding aquifer temperatures is crucial for optimizing low-enthalpy geothermal system exploitation. Advances in predictive algorithms can improve geothermal efficiency, while conventional methods of indirect temperature measurement and assumptions in geochemical analysis lead to uncertainties. As a solution, this study presents a comprehensive evaluation of six machine learning algorithms including eXtreme gradient boosting (XGBoost), decision tree, generalized regression neural network, extreme randomized trees, radial basis function, and elastic net. We employed essential performance metrics including coefficient of determination (R^2) score, root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and variance accounted for (VAF) to elucidate their predictive accuracy and generalization potential in the lower Friulian Plain (north-eastern Italy) where a geothermal reservoir is present. Among the algorithms scrutinized, XGBoost emerges as a predictive exemplar, achieving a remarkable R^2 score of 0.9930 on the test dataset, with consistently low RMSE of 0.788, MAE of 0.587, MAPE of 1.909, and high VAF of 99.30, reaffirming its exceptional precision and robustness. It is worth noting that the other four models show slightly weaker performance than XGBoost, while elastic net shows moderate predictive power, which illustrates the complexity of the database. The Wilcoxon signed-rank test confirmed the superior performance of XGBoost in estimating geothermal temperatures compared to other algorithms, with statistical evidence supporting its precision and reliability. A Monte Carlo simulation for uncertainty analysis underlined the importance of model selection, accuracy and uncertainty management in the planning of geothermal projects in the lower Friulian Plain. A sensitivity analysis was performed to identify the main factors influencing the temperature prediction. Among the parameters considered, bicarbonate the highest significance at 0.51, which is essential for accurate temperature prediction because of its buffering capacity which directly influences water's thermal properties. Magnesium and electrical conductivity each contribute with 0.11, also play significant roles due to their impact on the water's heat retention and distribution capabilities. Water depth, with a value of 0.08, also has a significant influence on the temperature profiles in prediction models. In summary, the accurate prediction of XGBoost for the temperature of aquifer in carbonate reservoirs in the lower Friulian Plain, underline its value for optimizing geothermal resources and highlight most important influences on temperature.

Nomenclature

ANN	Artificial Neural Network	**ML**	Machine Learning
AUC	Area Under the Curve	**NGB**	Natural Gradient Boosting
Ca	Calcium	**NH_4**	Ammonium
Cl	Chloride	**NO_3**	Nitrate
DNN	Deep Neural Network	**Na**	Sodium

(*continued on next column*)

(*continued*)

DT	Decision Tree	**PNRR**	Piano Nazionale di Ripresa e Resilienza
EC	Electrical Conductivity	**R^2**	Coefficient of Determination
Fe	Iron	**RBF**	Radial Basis Function Neural Network
FVG	Friuli Venezia Giulia	**REC**	Regression Error Curve

(*continued on next page*)

* Corresponding author. Department of Mathematics, Informatics, and Geosciences, University of Trieste, Italy.
E-mail address: dsheinidashtgoli@ogs.it (D. Sheini Dashtgoli).

https://doi.org/10.1016/j.jclepro.2024.142452
Received 6 November 2023; Received in revised form 5 April 2024; Accepted 2 May 2024
Available online 9 May 2024

(*continued*)

GRNN	Generalized Regression Neural Network	**RF**	Random Forest
HCO_3	Bicarbonate	**RMSE**	Root Mean Square Error
iNEST	Interconnected North-East Innovation Ecosystem	**SNN**	Shallow Neural Networks
K	Potassium	**SO_4**	Sulfate
LR	Linear Regression	**TDS**	Total Dissolved Solids
MAE	Mean Absolute Error	**VAF**	Variance Accounted For
MAPE	Mean Absolute Percentage Error	**XGBoost**	eXtreme Gradient Boosting
Mg	Magnesium	**XRT**	Extreme Randomized Trees

1. Introduction

Many countries are currently dependent on imported fossil fuels to meet their energy and development needs. Estimated global oil reserves have raised concerns about future energy security, particularly for developing countries. In response to this challenge, it is proposed for both developed and developing countries to diversify their energy sources and reduce dependence on fossil fuels. Among these alternatives, non-conventional energy sources, especially geothermal resources, are proving to be a promising solution. The utilization of geothermal resources not only promotes energy independence, but also contributes to the reduction of CO_2 emissions, thus ensuring the environmental preservation for future generations. Geothermal systems are characterized by their enthalpy level including low, medium, and high enthalpy, each offering unique opportunities for exploration and utilization (Stober and Bucher, 2021). In particular, low-enthalpy geothermal system utilization has experienced significant growth in recent years due to its greater accessibility compared to conventional high-temperature applications. In addition, the increasing demand for cooling in the context of global warming has contributed to this trend (Tomaszewska et al., 2018).

Geothermal energy depends on the production rate and the borehole temperature, which in turn depends on the aquifer temperature. Knowledge of the underground temperature is therefore crucial for the efficient planning of geothermal plants, as higher temperatures increase performance and cost efficiency. In addition to its scientific importance, the accurate measurement of subsurface temperatures also has practical implications. It serves as a guide for the planning and optimization of geothermal energy extraction methods, ultimately increasing energy efficiency and minimizing environmental impacts (Agemar et al., 2012).

In the domain of geothermal exploration, diverse methods have emerged for predicting subsurface temperatures. Seismic tomography unveils structural complexities, while electrical resistivity surveys decode temperature-related conductivity variations (Jones, 1998; Maercklin, 2005; Mechie et al., 2004; Muñoz et al., 2010). Numerical simulations and groundwater flow models provide insights into heat conduction and fluid movement (Pastore et al., 2021). Notably, geothermometer utilizes isotopic compositions to deduce thermal history (Arnórsson et al., 1983). Together, these methods offer intricate insights, enhancing resource assessment and sustainable energy extraction possibilities.

Despite the advantages of conventional geothermal exploration techniques, they have considerable limitations. Seismic tomography and electrical resistivity surveys, as illustrated by Guan et al. (2023), have difficulties in directly quantifying subsurface temperatures due to the indirect nature of their measurements. Moreover, numerical modeling and groundwater simulations highlight the uncertainties connected to input parameters, which can lead to significant inaccuracies if not carefully calibrated (Jia et al., 2024). Furthermore, as discussed by Kadri et al. (2023), the reliability of geochemical methods and geothermometers depends on the assumption that elemental concentrations remain unchanged post-formation, a premise that can lead to errors in temperature estimation if proven incorrect. These studies emphasized the necessity for accurate petrophysical relationships and rigorous model calibration when assessing geothermal resources.

In the context of geothermal resources, the extensive application of machine learning (ML) enables the discovery of complex relationships among different features, providing valuable insights. In addition, ML offers benefits such as improved resource utilization and improved predictive modeling for geothermal systems (i.e. Ahmmed and Vesselinov, 2022; Mudunuru et al., 2023; Rau et al., 2023). In response to the limitations of conventional methods, ML has proven to be an effective tool for predicting subsurface temperatures in geothermal reservoirs, optimizing multifaceted aspects.

Okoroafor et al. (2022) conducted a comprehensive review spanning two decades, investigating the increasing integration of ML algorithms in various areas of geothermal research, covering the areas of exploration, drilling, reservoir characterization, seismicity and production engineering. Moraga et al. (2022) proposed an innovative approach that integrates remote sensing, ML and Artificial Intelligence (AI) for an initial assessment of geothermal potential. By analyzing mineral markers, surface temperatures, faults and deformations, their AI model was able to accurately predict the geothermal potential based on surface manifestations. Subsequently, the potential of ML in predicting various geothermal parameters came to the fore. Kolawole and Assaad (2023) demonstrated the capability of ML in predicting changes in rock mechanical properties at different scales – micro, meso, and mega. Similarly, Qiao et al. (2023) addressed geothermal resource assessment and utilized ML to assess the geothermal potential of the Dongpu depression based on heat flow, geothermal gradient, and thermal properties.

In addition, researchers have used ML beyond the prediction of subsurface temperatures. For example, Porkhial et al. (2015) modeled the temperature behavior in geothermal reservoirs using experimental data and the group method of neural network type of data processing. Sharifi et al. (2016) conducted a thorough study of the Takab geothermal field (Iran) using chemical analyses of water samples from hot and cold springs. The hot springs showed elevated element concentrations, suggesting ongoing circulation and rock interactions. The heat source was attributed to mixing of deep geothermal fluid with cold groundwater or low conductivity heat flow. Estimates of reservoir temperatures varied, with the silica-enthalpy mixing model predicting a range of 62–90 °C.

Neural networks were used effectively for field-scale temperature estimation using resistivity data by Ishitsuka et al. (2018). Despite their success with limited temperature data, the accuracy of the method decreased with increasing distance from the measurements. To address this, a resistivity-based neural kriging approach was developed that incorporates a temperature data variogram to improve accuracy. The proposed method harmonized the estimated temperatures with the geology and offers potential for different geophysical data to estimate various physical parameters. Later, Pérez-Zárate et al. (2019) predicted geothermal reservoir temperatures using gas composition inputs through three-layer neural networks. Six optimal architectures, including artificial neural network (ANN) were identified through comparisons and externally validated. Their effectiveness in gas geothermometry was demonstrated by high accuracy (2–11% error). Shahdi et al. (2021) utilized temperature data from oil and gas wells and applied eXtreme gradient boosting (XGBoost) and random forest (RF) models to predict temperature-at-depth and geothermal gradients. Varol Altay et al. (2022) extended the applications of ML in geothermal energy utilization by using a hybrid metaheuristic ANN model to predict geothermal fluid properties. Furthermore, Puppala et al. (2023) addressed the complexity of reservoir temperature prediction and investigated the feasibility of using convolutional neural networks, recurrent neural networks and deep neural networks. Their study demonstrated the effectiveness of neural network models in predicting production temperatures while significantly reducing computational

effort. Yang et al. (2022) investigated the practicality of ML using data from the Lindian geothermal field. By using various methods, including ANN, they estimated reservoir temperatures and evaluated the prediction accuracy based on temperature logging data. Their results emphasized the effectiveness of ANN as an accurate method for estimating reservoir temperature.

Tut Haklidir and Haklidir (2020) used various ML algorithms, including deep learning model (DNN), to predict the temperature of geothermal reservoirs based on hydrogeochemical data from different geothermal systems. The results showed that DNN algorithm outperformed the other methods and provided very accurate reservoir temperature predictions that closely matched the geothermometer calculations. Later, Ibrahim et al. (2023) conducted another study on the same database to improve reservoir temperature prediction in geothermal exploration using ML techniques. Among the developed models, the Natural Gradient Boosting (NGB) model exhibited superior performance by achieving an impressive R^2 value of 0.9959 and showing low RMSE and mean absolute error (MAE) values. They used the Shapley additive explanation technique to gain insights into the decision-making process of the NGB model. In particular, they found that the concentration of SiO_2, a marker for volcanic geothermal resources, plays a central role in influencing temperature. The study's robust NGB model provided a reliable approach for estimating reservoir temperature – a critical factor in the exploration and utilization of geothermal energy resources. Abrasaldo et al. (2024) highlighted the central role of data-driven algorithms in optimizing the operation of geothermal energy systems and underlined the transformative potential of ML in this domain. Concurrently, Gudala et al. (2024) demonstrated the utility of a thermo-hydro-mechanical model for understanding the complex interactions in fractured geothermal reservoirs, and pointed out the effectiveness of neural network models in forecasting temperature changes in production wells. Qin et al. (2024) further advanced this integration and developed a physics-guided ML approach to overcome the inherent limitations of data-driven models for geothermal reservoir management by proposing a more reliable and generalizable method for optimizing field operations. Similarly, Yan et al. (2024) presented a physics-informed ML framework tailored to enhanced geothermal systems, demonstrating how computational efficiency and accurate predictions can work together to optimize thermal recovery in a sustainable manner.

In summary, the integration of ML in geothermal resource research has led to transformative advances in several areas with a particular focus on the prediction of subsurface temperature. These applications highlight the central role of ML in shaping the future of geothermal exploration, development and sustainability.

2. Research significance

Knowledge of aquifer temperatures, especially in relation to the low-enthalpy geothermal systems prevailing in the lower Friulian Plain, located in Friuli Venezia Giulia (FVG) region (north-eastern Italy), is essential for optimizing the use of geothermal resources. This need is emphasized by the geothermal system in Grado (Della Vedova et al., 2015). Accurate temperature prediction not only enables the efficient use of geothermal energy, but also helps to raise funds and provide data for numerical modeling, which is crucial for environmental sustainability and energy planning. However, despite few recent advances in temperature prediction algorithms, research in this specific area remains limited, both in the lower Friulian Plain and globally. This study aims to fill this gap by presenting a comprehensive evaluation of six ML algorithms. Taking into account the complex interplay of geochemical elements and unique geological settings that pose a challenge to model accuracy, the potential of each algorithm for more accurate predictions is explored.

3. Materials and methods

3.1. Study area

The study area is located in the lower Friulian Plain in north-eastern Italy (Fig. 1). This region is characterized by a complex geological and tectonic history, with the sedimentation of several kilometers of Mesozoic carbonates (mainly limestones) during the extensional regime, and subsequent Cenozoic Dinaric and Alpine compressional phases with the sedimentation of terrigenous sequences. The Cenozoic terrigenous sequences consist of Eocene turbitidic sediments (Trieste Flysch) that filled the Dinaric foreland and of the Miocene-Quaternary alternation of consolidated and unconsolidated terrigenous sediments that filled the Alpine foreland (e.g. Busetti et al., 2010; Dal Cin et al., 2022; Fantoni, Catellani, Merlini, Rogledi, & Venturini, 2002; Zecchin et al., 2022 and references therein). The lithological characteristics have led to the formation of carbonate and clastic aquifers (e.g. Cimolino et al., 2010; Della Vedova et al., 2015; Zini et al., 2011), in which fluid circulation could be locally influenced by the presence of faults and fracture systems that provide pathways for fluid migration (Busetti et al., 2013; Giustiniani et al., 2022; Petrini et al., 2013; Vesnaver et al., 2021).

3.1.1. Hydrogeological setting

The hydrogeological setting of the lower Friulian Plain is complex, as it is characterised by a strong heterogeneity of permeability depending on the different lithologies that form a deep Meso-Cenozoic carbonate aquifer and a clastic multilayered aquifer in the Cenozoic terrigenous sediments (e.g. Cimolino et al., 2010; Della Vedova et al., 2015; Zini et al., 2011). There is little data on the hydrogeological properties of aquifers and aquitards. However, laboratory tests on samples from permeable layers give values of 10^{-5}-10^{-6} m/s, while permeability measurements on some clayey-silty samples from depth give values of 10^{-10}-10^{-11} m/s (Zini et al., 2008).

The carbonates host thermal waters whose isotopic composition suggests that the deep saline reservoir may be the remnants of seawater trapped in the Meso-Cenozoic carbonates during the marine conditions of the late Oligocene - Miocene (Petrini et al., 2013). The carbonate

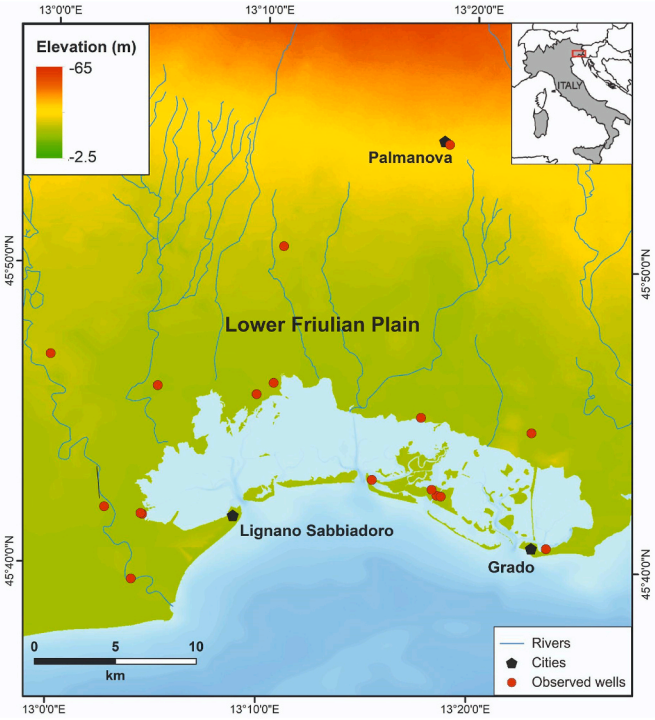


Fig. 1. Study area and location of the wells analyzed in this study.

aquifer has a multifaceted geometry and is confined, with its dynamics significantly influenced by prominent fault systems that strongly influence flow directions. Moreover, these fault systems contribute to the formation of hydrothermal cells, and may also allow some local mixing with the shallower freshwater aquifers (Petrini et al., 2013). Such interactions between different aquifers contribute to the complexity of the overall system. The shallower multi-layered aquifer system, located in Miocene and Plio-Quaternary sediments, consists of eleven aquifers characterised by high variability in depth and lateral continuity (Zini et al., 2011). These aquifers are characterized by the presence of Ca/HCO_3-type waters because there is a strong contamination from deep fractured and carbonates karstified as dimostrated by different studies (e.g. Cimolino et al., 2010; Petrini et al., 2013). It is important to emphasize that the water in the deeper layers is characterised by a high concentration of dissolved substances and temperatures of more than 35 °C.

3.2. Workflow

The workflow shown in Fig. 2 outlines the entire process, from using the database to implementation and evaluating the data. It includes the exploration of six different ML models, all focused on the prediction of reservoir temperature based on hydrochemical characteristics. Three of these models utilize various ensemble techniques, such as boosting and bagging, to improve model performance. In addition, two models are structured as shallow neural networks (SNN), specifically selected for the limited number of data points in the database, with the goal of effectively capturing and analyzing the nonlinearity of the data. Moreover, a linear regression model (LR) is integrated to provide a comparative analysis of different modeling approaches. These strategies are explained in more details in section 3.5 where the specifics of the algorithms used are discussed. Each model undergoes a comprehensive evaluation using metric measurements and hyperparameter tuning to optimize performance to ensure the most accurate predictions possible. The performance of the model is rigorously evaluated using five metrics, AUC and statistical tests. In addition, the uncertainty of each model's predictions is carefully analyzed. The process culminates in a sensitivity analysis for the best performing model to estimate the impact of variable inputs, significantly improving the reliability of the model and its applicability in real-world scenarios.

3.3. Database

3.3.1. Statistical overview of the data

This study is based on a database of key hydrogeochemical parameters and temperature data obtained from various boreholes in the FVG region (Fig. 1). It includes 74 data points, containing information on depth, pH, electrical conductivity (EC), total dissolved solids (TDS), major cations (Na, K, Mg, Ca), bicarbonate (HCO_3), ammonium (NH_4), nitrate (NO_3), chloride (Cl), sulfate (SO_4) and iron (Fe). The temperature data is particularly important for understanding the geothermal activity and conditions of the reservoir in this area. This database forms the basis for training and testing ML models to predict deep water temperatures in the FVG region, contributing to a comprehensive understanding of the geothermal potential of the region. Information about the database is summarized in Table 1.

3.4. Data visualization

Visualization plays a crucial role in data science by translating complex patterns and trends into accessible visual representations, enabling better insights and decision-making (Correa et al., 2009). In the scatter plot in Fig. 3 showing the relationship between temperature and various input features, such as depth, HCO, and NH_4, an interesting pattern emerges. In particular, there is a clear trend towards an increase in temperature as the values for depth, HCO_3, and NH_4 increase. This observation suggests a possible correlation between these factors and the temperature variations.

The correlation heatmap is a powerful visualization tool that provides a concise overview of the relationships between the variables in a dataset. By mapping correlation coefficients (CC) to a color spectrum, these visualizations provide an efficient way to identify patterns and connections between different features. In this context, we use a correlation heatmap to analyze the intricate relationships between aquifer temperature and various hydrochemical parameters. This analysis aims to elucidate the underlying dynamics of temperature variations in the aquifer with respect to the measured parameters.

The correlation heatmap in Fig. 4 underscores insights into the relationships between the aquifer temperature and various hydrochemical parameters. The aquifer temperature exhibits a strong robust positive correlation with depth (CC = 0.744), suggesting that greater depths are associated with increased temperatures. This correlation can be partially attributed to the geothermal gradient (Stefanini, 1980). Conversely,

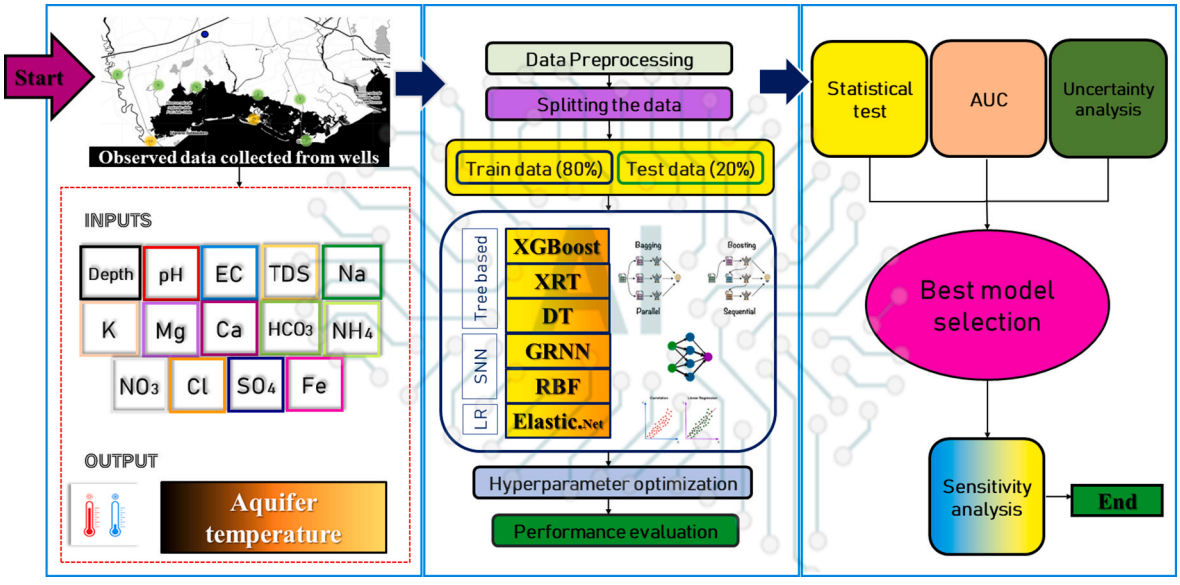


Fig. 2. Workflow of the procedure adopted.

Table 1
Overview of the statistical values of the parameters in the database.

Symbol	Units	Average	STD	Min	1st quartile	2nd quartile	3rd quartile	Max
Depth	m	353.04	112.08	120.00	282.00	340.00	423.25	594.00
pH	–	7.75	0.38	6.78	7.48	7.84	7.92	8.67
EC	µS-cm	3874.56	7366.15	235.00	352.25	407.50	1735.25	27400.00
TDS	–	676.23	735.94	136.00	201.00	228.90	939.25	2000.00
Na	mg-l	708.78	1399.33	6.05	32.93	58.53	309.91	4997.00
K	mg-l	35.07	62.18	1.16	2.36	3.87	25.35	248.60
Mg	mg-l	34.16	55.31	1.23	9.00	10.46	19.01	215.75
Ca	mg-l	79.21	164.96	2.66	12.57	18.25	32.13	698.60
HCO_3	mg-l	273.28	83.24	159.82	219.73	257.43	290.45	497.31
NH_4	mg-l	5.49	6.05	0.18	1.53	2.59	6.34	24.76
NO_3	mg-l	3.85	5.49	0.01	0.73	1.48	4.27	27.57
Cl	mg-l	1368.28	2946.95	0.42	3.51	14.35	388.47	11028.82
SO_4	mg-l	72.40	201.49	0.40	0.88	2.08	15.91	800.30
Fe	µg-l	100.80	105.05	4.00	40.00	74.50	120.50	676.00
Temp	°C	28.81	7.49	16.30	24.10	26.90	32.85	47.40

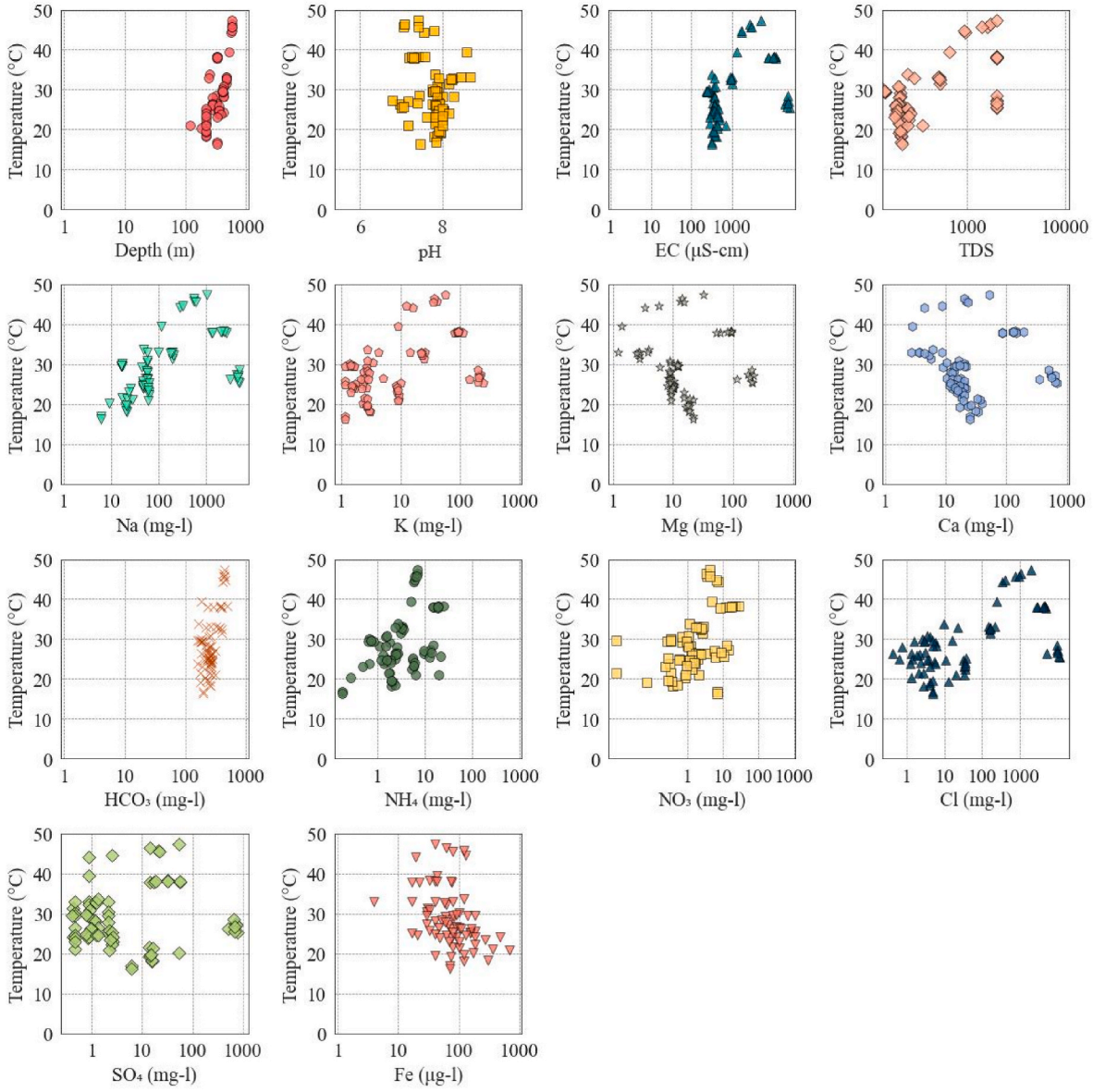


Fig. 3. Scatter plots with the temperature values compared to other input parameters.

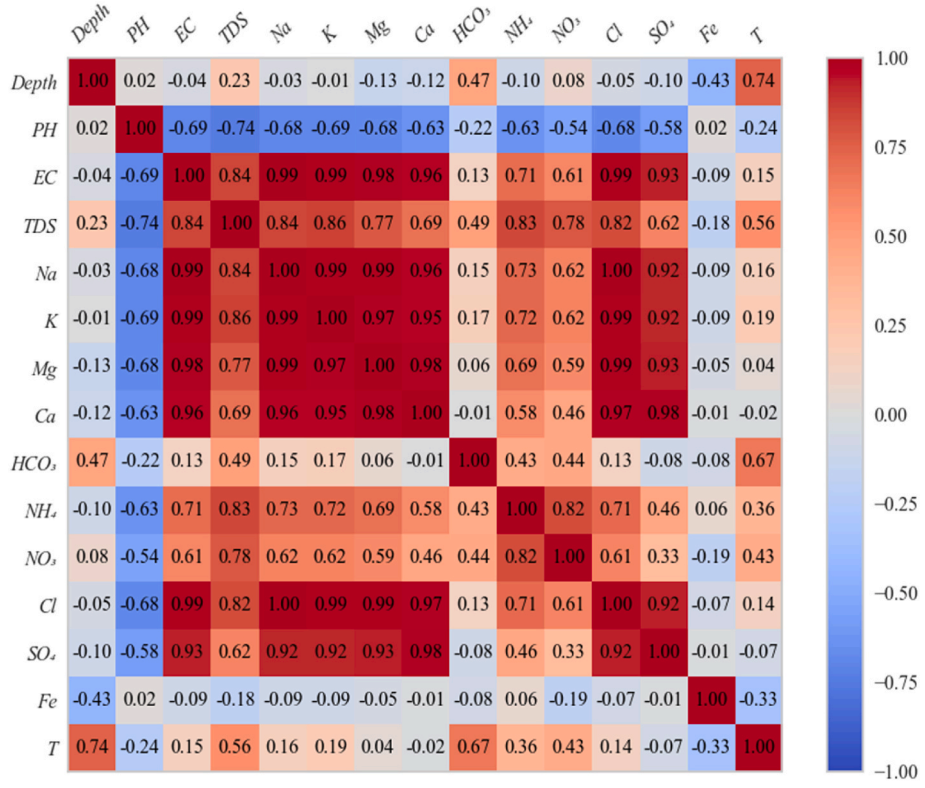


Fig. 4. Correlation heatmap of geochemical features and aquifer temperature (T).

parameters such as HCO_3 (CC = 0.668), TDS (CC = 0.563), NO_3 (CC = 0.427), and NH_4 (CC = 0.363) show moderate to strong correlations with temperature increases. These correlations indicate that the contributions of mineral dissolution, chemical reactions involving bicarbonate, and nitrogen transformations contribute to temperature dynamics. Conversely, parameters such as pH (CC = −0.239) and SO_4 (CC = −0.070) exhibit negative correlations with temperature, suggesting that higher acidity and higher SO_4 concentrations are associated with slightly lower temperatures. Na, K, and Cl exhibit weaker positive correlations, while Mg, Ca, and Fe show weaker negative correlations with temperature, suggesting subtle influences on temperature dynamics. While these observed correlations provide insights into potential relationships, further multidimensional analyses and expertise are essential to unravel the intricate interplay of hydrogeological and geochemical factors responsible for aquifer temperature fluctuations in the aquifer.

3.5. Model development

3.5.1. XGBoost

The XGBoost technique constructs an ensemble of decision tree (DT), where each individual tree is based on the residuals of the previous tree. This means that each tree is carefully designed to reduce the residual error of the previous tree and thus improve the overall accuracy of the model. An important advantage of XGBoost is its fast execution and scalability. It is able to manage large datasets with a variety of features, making it an extremely popular option in numerous ML scenarios. In addition, XGBoost contains an array of built-in functionalities designed to improve the performance of the model, such as early stopping and cross-validation. Another notable feature of XGBoost is its capacity for comprehensibility. Within XGBoost, one gains insight into the meaning of each feature within the model, allowing for a deeper understanding of the decision-making process (Chen and Guestrin, 2016). The general form of the objective function per iteration is given by:

$$Obj = \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t)}\right) + \sum_{k=1}^{K} \Omega(f_k) \tag{1}$$

$$where\ \ \Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{i=1}^{T} w_j^2 \tag{2}$$

where:

Obj represents the objective function to be minimized.

n is the number of instances in the training data.

y_i is the actual value of the ith instance.

$\hat{y}_i^{(t)}$ is the predicted value for the ith instance at the tth iteration.

l is a differentiable convex loss function that measures the difference between the predicted value and the actual value (squared loss).

K is the number of trees in the model.

f_k represents the k^{th} tree.

$\Omega(f_k)$ is the regularization term for the k^{th} tree.

T is the number of leaves in the tree.

w_j is the score on the j^{th} leaf.

γ and λ are parameters that control the complexity of the model. γ penalizes the number of leaves, and λ penalizes the size of the leaf scores.

3.5.2. XRT

Extremely randomized trees (XRT), a new method for supervised classification and regression tasks, was proposed by Geurts et al. (2006) and is based on a tree-based ensemble approach. The core idea is that considerable randomness is introduced both in the selection of attributes and in the selection of cut-point during node splitting of the trees. In some cases, the trees are constructed completely at random, without any dependence on the output values of the training dataset. The degree of randomization is adjustable by a specific parameter, tailored to the specific requirements of the problem.

3.5.3. Elastic net

The elastic net algorithm is a linear regression technique that combines the properties of the Lasso (L1 regularization) and Ridge (L2 regularization) regression methods. It was developed to deal with situations where many variables are present and some of them may be correlated or redundant. Elastic net aims to find a balance between feature selection and regularization to improve the performance and interpretability of linear regression models (Zou and Hastie, 2005). The objective function for elastic net regression can be formulated as follows:

$$\text{Minimize}: \frac{1}{2n}\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ij}\right)^2+\lambda\left(\alpha\sum_{j=1}^{p}|\beta_j|+\frac{1-\alpha}{2}\sum_{j=1}^{p}\beta_j^2\right) \quad (3)$$

where:

n is the number of observations.

y_i is the response variable for the i^{th} observation.

x_{ij} is the value of the j^{th} predictor (feature) for the i^{th} observation.

β_0 is the intercept term.

β_j is the coefficient for the j^{th} predictor.

p is the total number of predictors.

λ is the regularization parameter that controls the strength of the penalty. The higher the value of λ, the stronger the penalty.

α is the mixing parameter that controls the balance between L1 and L2 penalties. $\alpha = 1$ corresponds to Lasso regression (pure L1 penalty), and $\alpha = 0$ corresponds to Ridge regression (pure L2 penalty).

The first term in the objective function is the Residual Sum of Squares (RSS), which measures the fit of the model to the data. The second term is the elastic net penalty, which is a linear combination of the L1 and L2 penalties.

3.5.4. DT

DT is a core ML algorithm known for its intuitive decision-making process and its versatility in classification and regression tasks. Like a flowchart, a DT algorithm breaks down the data into a hierarchy of decisions based on features, and finally arrives at a prediction outcome. At each internal node, the algorithm selects a feature and applies a split based on a threshold, segmenting the data into increasingly homogeneous subsets. This recursive splitting continues until specified conditions are met, such as a certain tree depth. DT is valued for its transparency, as it offer a visual representation of the decision path that allows for easy interpretation of predictions. Moreover, DT can handle non-linear relationships and accommodate different data types, so they can adopted to different datasets. Nevertheless, they can lead to overfitting if not properly regularized and their sensitivity to data variability can lead to model instability. Despite these limitations, DT remains a fundamental tool in the ML toolkit, valued for the balance between comprehensibility and predictive accuracy (Quinlan, 1986).

3.5.5. RBF neural network

The Radial Basis Function (RBF) model, influenced by neural networks, uses basis functions with centers and spreads to transform data to mimic neural connections structured as shallow neural networks. These functions act as synapses that determine the influence of the feature on the output. This neural-inspired approach allows RBF to capture complex patterns and relationships, especially in non-linear data. It is ideal for regression, classification, and clustering tasks. In regression, weighted sums of the basis function outputs are calculated for predictions, and in classification, decision boundaries are created in the transformed space. However, RBF requires careful parameter tuning to prevent overfitting. Despite its power, it requires careful implementation and regularization (Safaei-Farouji et al., 2022b; Schwenker et al., 2001).

3.5.6. GRNN

General regression neural network (GRNN) is one of the most powerful ANN methods developed to continuously predict output values (Specht, 1991a). The GRNN is essentially a general quantitative approach known as kernel regression, which can be defined as a normalized radial basis neural network (Mahdaviara et al., 2022; Vo Thanh et al., 2023). This type of ANN topology has two main advantages: low cost and fast learning (Afrasiabi et al., 2022). The repeated computation approach is not used in GRNN. This method can accurately predict any relationship between the input and output matrices by using only the training samples. The input layers, the pattern layers, the summation layers and the output layers are the four layers of GRNN. As their terms suggest, the input and output layers are responsible for accepting data and giving model outputs. In addition, the scatter variable (σ) is the only variable that can be used to fine-tune the efficiency of the GRNN algorithm during the learning phase (Cigizoglu and Alp, 2006).

3.6. Limitations

3.6.1. Database limitation

Determining the minimum dataset size for accurate modeling depends on the domain and data characteristics, although even small datasets can potentially provide important insights if properly analyzed (Ranstam and Cook, 2018). The challenge of limited datasets is particularly pronounced in fields such as geochemical analysis, where data collection is difficult and costly (Shorten and Khoshgoftaar, 2019). To address issues such as overfitting in such data-poor environments, the adoption of appropriate algorithms and the use of techniques such as regularization and cross-validation is crucial for improving model performance and generalizability (Babyak, 2004). Advanced regularization methods (Chakrabarti, 2022) and statistical model selection techniques (Brunton et al., 2016) further help in developing of robust models that can be generalized to unseen data despite the limitations of the dataset. The key is to focus on methodological rigor and the right analytical tools to create accurate and generalizable models in different contexts.

3.6.2. ML limitation

ML models are widely recognized for their high prediction accuracy. These models are excellent due to their flexibility in processing different types of data and their ability to model complex nonlinear relationships, making them useful in fields such as geoscience (i.e. Ghaffari-Razin et al., 2023; Kianoush et al., 2023; Ma et al., 2023; Male and Duncan, 2020; Safaei-Farouji et al., 2022a; Sarailidis et al., 2023; Sheini Dashtgoli et al., 2024; Wang et al., 2023; Wei et al., 2020; Xu et al., 2024; Yao et al., 2023; Zhang et al., 2017). Despite these strengths, challenges such

Table 2
Limitation of the ML models used.

Model	Limitations	References
XGBoost	Overfitting Computationally intensive for large datasets	Chen and Guestrin (2016), Friedman (2001)
DT	Instability Overfitting Difficulty with linear relationships	Breiman (2001)
XRT	Complexity Overfitting Similar biases as RF	Geurts et al. (2006), Louppe (2014)
Elastic net	Parameter sensitivity Underperformance in non-linear	Tibshirani (1996), Zou and Hastie (2005)
GRNN	Computationally intensive for large datasets Incapability of extrapolation	Kurup and Griffin (2006), Specht (1991b), Wasserman (1993)
RBF	Function choice dependence Overfitting risk	Meng et al. (2002), Wu et al. (2013)

as overfitting, computational complexity, and sensitivity to outliers require careful model selection, parameter tuning, and validation (Hawkins, 2004). As shown in Table 2, the challenges addressed therein ensure robust and effective ML solutions and emphasize the importance of understanding model limitations and the trade-offs involved in both model development and application (i.e. Tavares et al., 2022; Westphal and Brannath, 2020). This balanced approach to model selection and evaluation allows the benefits of ML models to be effectively utilized while reducing their limitations.

3.7. Hyperparameter tuning and optimization

Tuning the hyperparameter is of paramount importance in the development of the ML model as it has a direct impact on the performance and generalization ability of the model. A proper configuration of hyperparameters can significantly improve the accuracy and effectiveness of the model. For this crucial aspect, the Optuna® library provides an invaluable solution (Akiba et al., 2019). Optuna® is a widely used Python library tailored to automate the complicated process of hyperparameter optimization. By using different optimization algorithms, such as the covariance matrix adaptation evolution strategy (CMA-ES), Optuna® systematically explores the hyperparameter space. This systematic exploration aims to identify the configuration that yields the best optimal model performance. Here, we outline the step-by-step process employed for each model.

1. **Initialization**: We started by defining a high-dimensional search space for the hyperparameters of each model. Initial hyperparameter values were selected based on preliminary experiments.
2. **Objective Function**: For each model, an objective function was defined that calculates the cross-validation value (R^2 value) of the model with a given set of hyperparameters. The goal of the optimization was to maximize this score.
3. **CMA-ES Optimization**:
   - **Population Size**: at each iteration, a population of candidate solutions (hyperparameter sets) was generated, with the size determined based on the dimensionality of the search space.
   - **Adaptation**: At each step, CMA-ES updated the mean and covariance of the hyperparameter distribution based on the performance of the current population, guiding the search towards regions of the search space with promising hyperparameters.
   - **Selection**: The best-performing candidates from the population were selected to influence the distribution in the next generation, ensuring a focus on promising areas of the search space.
4. **Convergence Criteria**: The optimization process ends after a certain number of iterations, fixed at 2000 trials. This predetermined stopping condition ensures a comprehensive exploration of the hyperparameter space.
5. **Model-Specific Tuning**: For all models, the tuning process focused on optimizing key parameters such as tree depth, regularization strengths, spread coefficients, and network architecture to significantly improve the predictive performance of each model.
6. **Final Model Selection**: Upon completion of the CMA-ES optimization process, the hyperparameter set yielding the best cross-validation score was selected for each model, ensuring optimal performance.

By employing CMA-ES, we leveraged a sophisticated and adaptive approach to navigate the complex hyperparameter spaces of our models. This methodology allowed for efficient and effective tuning, significantly improving the predictive performance of our machine learning algorithms for accurate geothermal temperature estimation.

3.8. ML evaluation metrics

Model validation and analysis as part of the modelling process is a critical phase in which it is determined whether or not the intelligent model has delivered sufficiently accurate results for the intended objective. Different statistical indicators are used to recognize the relationships between the predicted reservoir temperature and the experimental reservoir temperature that are R^2, RMSE, MAE, MAPE, and VAF. The equations for the three performance factors is described as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(Temp_i - Temp_i^*\right)^2}{\sum_{i=1}^{n} \left(Temp_i^* - \overline{Temp}\right)^2} \tag{4}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \left(Temp_i - Temp_i^*\right)^2} \tag{5}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \left|Temp_i - Temp_i^*\right| \tag{6a}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n} \left|\frac{Temp_i^* - Temp_i}{Temp_i^*}\right| \tag{7a}$$

$$VAF = \left(1 - \frac{Var\left(Temp_i^* - Temp_i\right)}{Var\left(Temp_i^*\right)}\right) \times 100\% \tag{8a}$$

where n is the number of experimental data points, $Temp_i$ is the predicted temperature values, $Temp_i^*$ is the measured temperature values, Var ($Temp_i^* - Temp_i$) is the variance of the residuals (the differences between the measured and predicted temperatures), and $\overline{Temp}$ is the average temperature.

The effectiveness of statistical metrics in the evaluation of regression models depends on factors such as data quality, the complexity of the model, the selection of characteristics, the data distribution and the evaluation methodology. Data noise and outliers significantly affect model predictions, while both overfitting and underfitting can affect model performance (Choi, 2009; Hastie et al., 2009). Proper feature selection increases model accuracy by focusing on relevant data (Chandrashekar and Sahin, 2014). Furthermore, dealing with heteroscedasticity of data and scaling issues is crucial for accurate model evaluation (Kotsiantis et al., 2007).

4. Results and discussion

4.1. Tuned hyperparameters

In this study, we harnessed six robust ML models to predict temperature. During the optimization phase, we explored different ratios between training and testing, including 70:30, 75:25, and 80:20, to fine-tune our models' performance. The 80:20 train-test ratio ultimately provided the best results. Advanced hyperparameter tuning techniques were applied to optimize prediction performance, and the corresponding tuned parameters are listed in Table 3.

4.2. Statistical analysis of results

In this study, a rigorous quantitative assessment of different ML algorithms is performed using performance metrics - R^2 value, RMSE, MAE, MAPE, and VAF. The evaluation aims to elucidate the predictive accuracy and generalization potential of each algorithm on a given dataset. Among the analyzed algorithms, XGBoost proves to be a prime example of predictive prowess. As shown in Table 4, it attains excellent R^2 values of 0.9999 and 0.9930 in the training and test datasets, respectively. Its consistently low RMSE of 0.788, MAE of 0.587, MAPE of 1.909, and high VAF of 99.30% underscore its exceptional precision and robustness in different scenarios. The DT model shows remarkable

Table 3
The hyper-parameter tuning is defined for this paper.

Model	Hyperparameter	Value
XGBoost	n_estimators	100
	learning rate	0.099
	max depth	380
	min child weight	1
	gamma	0.0049
	reg alpha	0.0001
	reg lambda	0.0000129
	tree method	exact
DT	criterion	friedman_mse
	max_depth	7
	min_samples_split	2
	min_samples_leaf	1
GRNN	spread coeficient	0.0792
XRT	criterion	poisson
	max_depth	32
	min_samples_split	4
	min_samples_leaf	2
	Max feature	None
RBF	nNeuron	20
	spreed	17
Elastic net	λ	0.1210
	α	0.9595

predictive prowess, achieving R^2 values of 0.9955 and 0.9921 for the training and test sets, respectively. These values emphasize the model's ability to capture complicated relationships within the data. Despite a slightly higher RMSE value of 0.834, MAE value of 0.602, MAPE value of 2.061, and VAF value of 99.24%, the performance of the DT model remains good. The GRNN model has a high predictive capacity and achivie R^2 values of 0.9956 (train) and 0.9848 (test). Although the RMSE value of 0.818, MAE value of 0.635, MAPE value of 2.439 and VAF value of 98.56% are slightly higher than those of XGBoost and DT, they demonstrate the ability of GRNN to effectively capture the underlying data patterns. In the case of the XRT model, commendable R^2 values of 0.9842 (train) and 0.9602 (test) are observed. However, relatively high RMSE value of 1.879, MAE value of 1.371, MAPE value of 4.618, and VAF value of 96.23% indicate potential for improvement, so that the use of this model should be considered in specific contexts. The RBF model maintains consistent R^2 values of 0.9830 in all datasets. While its RMSE value of 1.756, MAE value of 1.435, MAPE value of 6.441, and VAF value of 90.25% are comparatively higher, but the stability and steady performance recommend its use in certain applications. Elastic net has moderate predictive capabilities, which is reflected in R^2 values of 0.8582 (train) and 0.8294 (test). Although its RMSE value of 3.895, MAE value of 2.673, MAPE value of 11.248, and VAF value of 83.74% are relatively high, elastic net shows potential in scenarios where alternative models are less suitable. These results contribute to a deeper understanding of algorithm performance and serve as a basis for informed model selection for temperature prediction in the FVG region.

4.3. Visual representation of results

Fig. 5 illustrates the relationship between the predicted and the actual temperature for each ML model used. The correlation coefficients for the predicted and the observed temperature in both the training and the test datasets are in substantial agreement with the fitted line, which is characterized by a slope of 1. It is noteworthy that the elastic net model shows a large scatter distribution. Nevertheless, the remaining five models show commendable performance in predicting reservoir temperature.

Fig. 6 shows a detailed overview through comparative bar charts of R^2, RMSE, MAE, MAPE and VAF, highlighting the performance of the six ML models in the training and test datasets.

4.3.1. AUC

In this study, we evaluated the accuracy of the regression models using the area under the curve (AUC) of the regression error curve (REC). This innovative approach, adopted from classification analysis, facilitates the visualization of a model's predictive accuracy over a range of error tolerances. The REC curve shows the accuracy of model predictions within defined limits of deviation from actual results, with an AUC value of 1 indicating perfect prediction accuracy and 0.5 indicating performance equivalent to chance. This metric is particularly useful to illustrate a model's ability to keep prediction errors within acceptable limits. The AUC is a consistent, comparative measure that captures the model's precision across different thresholds and provides an overview of the model's overall performance in regression tasks (Tahmassebi et al., 2018). As shown in Fig. 7, the REC curves for the training and test datasets show a clear pattern in the performance accuracy and consistency of the evaluated ML algorithms. For the training dataset, XGB achieves an AUC of 0.986, indicating near-perfect prediction accuracy. This is reflected in a steep REC curve, indicating that XGB maintains a low prediction error across most of the dataset. For the test dataset, XGB shows a slight decrease in AUC to 0.975, which is lower but still indicates high accuracy in predicting unseen data. DT shows comparable robustness with an AUC of 0.982 in the training dataset, indicating that it has high accuracy. The REC curve of DT is almost indistinguishable from that of XGB, highlighting its ability to keep prediction errors within a narrow range. In the test set, DT's AUC remains high at 0.975, suggesting that its performance is consistent even when applied to new data. GRNN and XRT both show strong predictive capabilities, with AUCs of 0.980 and 0.982, respectively, in the training phase. Their REC curves closely follow the curves of XGB and DT, indicating similar accuracy at different error thresholds. In the test phase, both GRNN and XRT achieve high AUCs of 0.973 and 0.975, confirming their efficiency in the test set. While RBF shows a solid AUC of 0.964 in the training set, it exhibits a REC curve that flattens out with increasing deviation, suggesting that its accuracy decreases at higher error margins. This trend continues in the testing phase, with the AUC of the RBF dropping to 0.954, highlighting a decrease in performance on unseen data. elastic net has the lowest AUCs in both the training (0.942) and testing (0.916) phases. The REC curve for elastic net flattens more than that of the other

Table 4
The comparison of statistical indicators is highlighted for six ML models.

Model	Training dataset					Testing dataset					Reference
	RMSE	MAE	R^2	MAPE	VAF	RMSE	MAE	R^2	MAPE	VAF	
XGBoost	0.026	0.020	0.9999	0.072	99.99	0.788	0.587	0.9930	1.909	99.30	This study
DT	0.448	0.220	0.9955	0.811	99.55	0.834	0.602	0.9921	2.061	99.24	This study
GRNN	0.505	0.270	0.9956	1.193	99.44	0.818	0.635	0.9848	2.439	98.56	This study
XRT	0.847	0.570	0.9842	2.231	98.42	1.879	1.371	0.9602	4.618	96.23	This study
RBF	1.333	0.995	0.983	4.775	94.36	1.756	1.435	0.9830	6.441	90.25	This study
Elastic net	2.542	1.833	0.8582	7.061	85.82	3.895	2.673	0.8294	11.248	83.74	This study
NGB	3.800	2.880	0.9988	–	–	4.593	3.967	0.9959	–	–	Ibrahim et al. (2023)
DNN	10.240	7.960	–	–	–	8.292	6.451	0.9791	–	–	Tut Haklidir and Haklidir (2020)

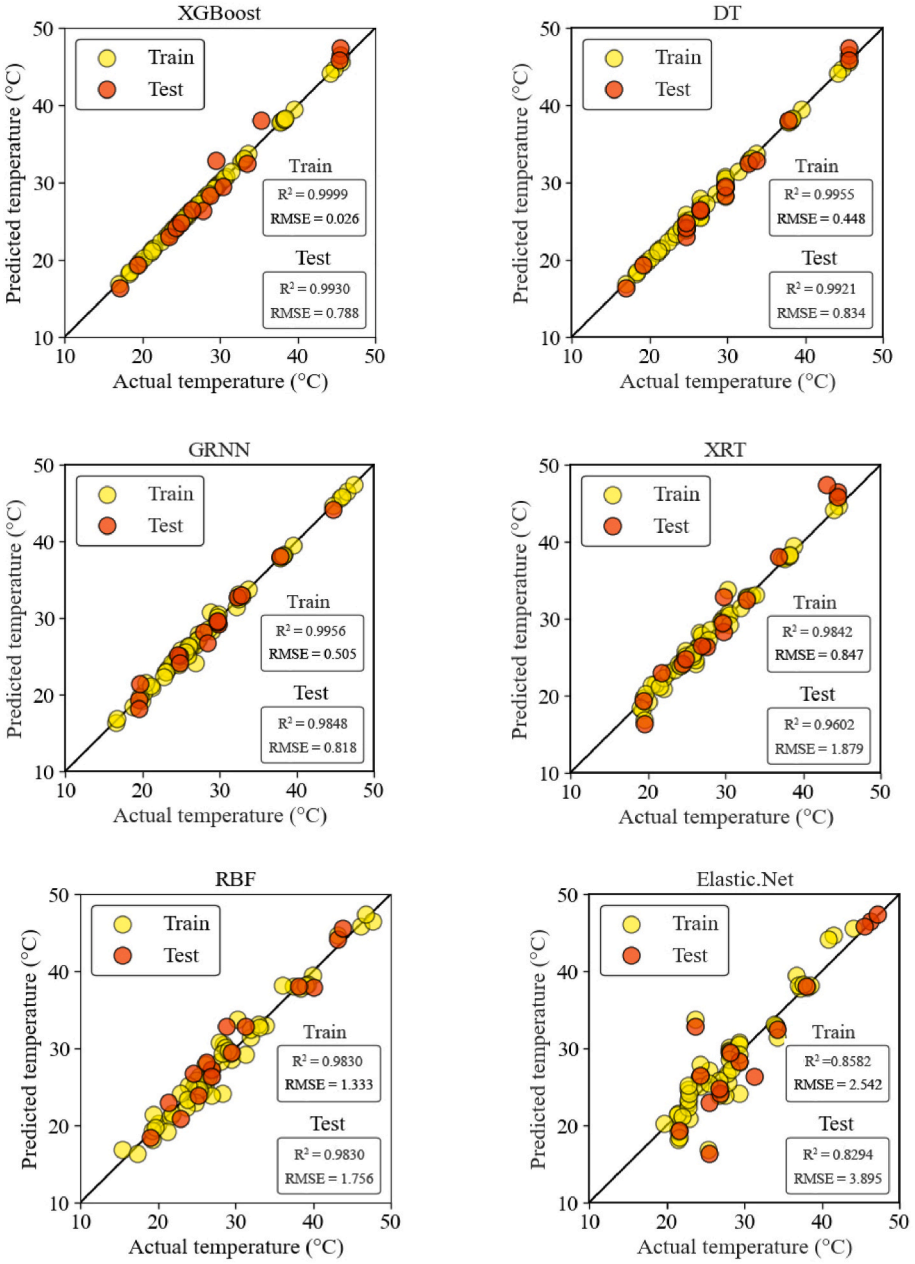


Fig. 5. Cross-plots showing the relationship between the experiment and the predicted temperatures for each ML model used.

models, meaning that its predictions are less consistent and deviate more from the actual values. The steeper drop in AUC from training to test suggests that elastic net may have difficulty generalizing its predictions to new datasets. In summary, analyzing the REC curves provides a comprehensive assessment of the models' performance. The steepness of the REC curves for XGB, DT, GRNN and XRT indicates that these models have consistently high accuracy at both low and high error tolerances, demonstrating their robustness and reliability in different scenarios. In contrast, the flatter REC curves of RBF and elastic net in particular indicate areas where these models may need further refinement to achieve comparable performance.

4.3.2. *Wilcoxon signed-rank test*

Statistical tests play an important role in ML evaluation, as they provide a basis for evaluating, model performance and rigorously comparing different algorithms. The Wilcoxon signed-rank test is a non-parametric statistical test that is particularly useful when comparing two related samples or repeated measurements on a single sample to determine whether their mean ranks differ in the population (Rosner et al., 2006). It is an alternative to the paired Student's t-test when it cannot be assumed that the data are normally distributed. The mathematical basis of the Wilcoxon test is to rank the absolute differences between paired observations, ignoring the signs, and then add the ranks for the positive and negative differences to calculate the test statistic. The test statistic is used to derive the p-value, which indicates the probability that the test results are observed under the null hypothesis. The z-value, a measure of the statistical effect size, is calculated by normalizing the test statistic based on the expected mean and standard deviation of the ranks under the null hypothesis. These statistical measures are crucial in ML evaluation, as they help to make informed decisions about model improvements and algorithm selection based on empirical evidence rather than assumptions or heuristic rules (Chandra and Verma, 2020; Uçar et al., 2020).

The statistical analysis performed with the Wilcoxon signed-rank test

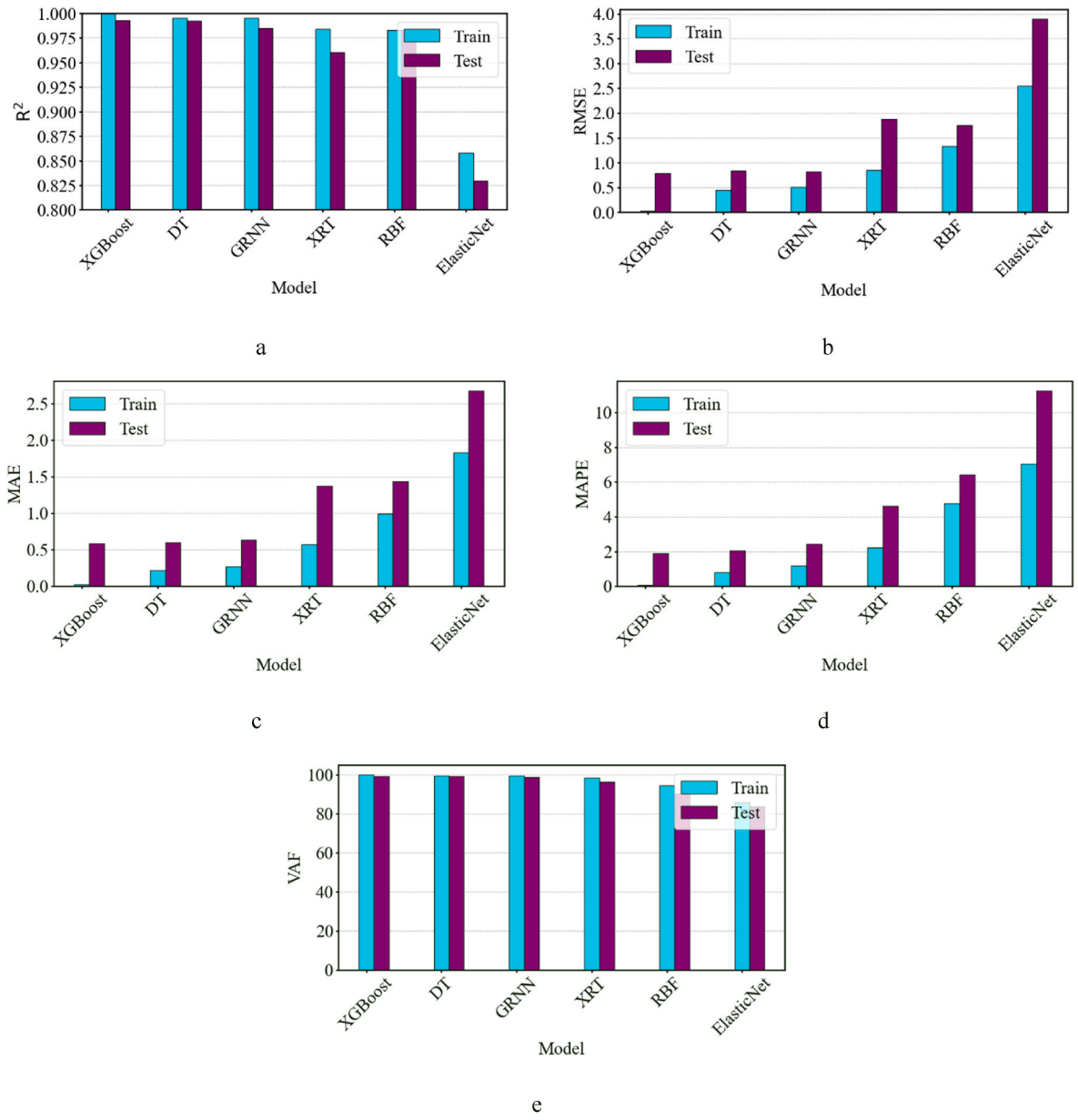


Fig. 6. The statistical efficacy of the six machine learning models a) R^2, b) RMSE, and c) MAE, d) MAPE, and e) VAF.

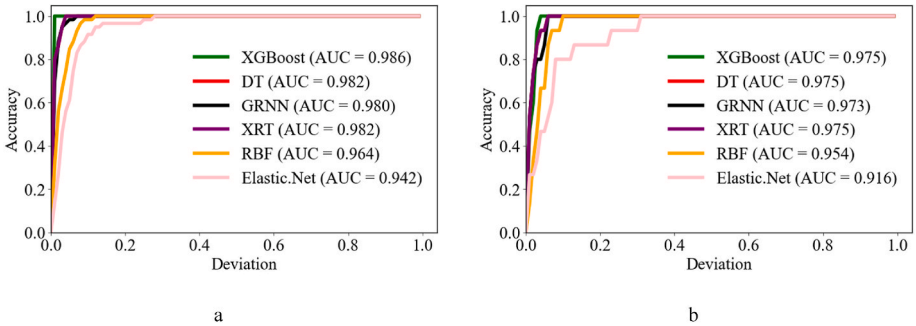


Fig. 7. The reported AUC values correspond to the a) train phase b) test phase.

Table 5
Wilcoxon signed rank test.

Pair comparison	P-Value	Z-Value
XGBoost vs DT	0.488708	0.738350
XGBoost vs XRT	0.005371	2.669421
XGBoost vs Elastic net	0.010254	2.499032
XGBoost vs GRNN	0.000122	3.350975
XGBoost vs RBF	0.000061	3.407771

provides compelling evidence for the differential performance of the different prediction models with respect to the RMSE compared to the XGBoost algorithm (Table 5). In particular, the comparison between XGBoost and the DT model yields a p-value of 0.488708, indicating no statistically significant difference in prediction error, suggesting that both models are comparable in terms of accuracy. Conversely, significant differences are observed when XGBoost is compared to XRT (p-value: 0.005371), elastic net (p-value: 0.010254), GRNN (p-value: 0.000122), and RBF (p-value: 0.000061) with the low p-values strongly indicating superior performance of the XGBoost model in minimizing prediction error. The negative Z-values in these comparisons confirm the lower RMSE values for XGBoost to the other models. This analysis not only highlights the robustness and efficiency of XGBoost in providing accurate predictions, but also emphasizes its significant advantages over traditional tree-based methods (such as XRT and DT), linear elastic net, GRNN, and RBF within the dataset studied. These results demonstrate the statistical rigor behind the evaluation of model performance, supporting the argument for the use of XGBoost in prediction tasks requiring high accuracy and reliability.

4.4. Uncertainty assessment

4.4.1. Method 1

As the first method, an analysis of uncertainty was performed for the developed model to quantitatively assess the uncertainty involved in the estimation of the temperatures. This process includes the computation of the mean error ($\bar{e}$) and the standard deviation of the error (S_e) by applying the following formulas:

$$\bar{e} = \frac{\sum_{i=1}^{n} e_i}{n} \tag{6b}$$

$$S_e = \sqrt{\frac{\sum_{i=1}^{n} (e_i - \bar{e})^2}{n-1}} \tag{7b}$$

$$uncertainty\ band\ width = \pm 1.96 S_e \tag{8b}$$

$$e_i = Temp_{(predicted)} - Temp_{(actual)} \tag{9}$$

By using $\bar{e}$ and S_e, it is possible to construct a confidence interval around the predicted error values by utilizing the Wilson score approach without applying a continuity correction, as described by Newcombe (1998). The application of $1.96S_e$ allows the approximation of a 95% confidence interval for these predictions.

The performance analysis of different predictive models shows a range of accuracy and uncertainty (Table 6). XGBoost slightly underestimates with a very narrow uncertainty bandwidth (1.5437), indicating precise predictions. DT and GRNN both overestimate, with GRNN being slightly more accurate in its predictions. The XRT model significantly underestimates the values and has considerable uncertainty (uncertainty bandwidth = 3.5891), while elastic net has the largest overestimation and the highest uncertainty of all models (uncertainty bandwidth = 7.4540). Fig. 8 shows the relationship between the relative error and the predicted temperature for the six ML models: XGBoost, DT, GRNN, XRT, RBF, and elastic net. It is evident from the plot and mean error calculated in Table 6 that moving from XGBoost to elastic net, there is a clear trend towards an increase in relative error. This trend means that the accuracy of the predictions tends to decrease with the transition from XGBoost, a high-performing model, to elastic net, which exhibits comparatively higher relative errors in temperature prediction.

This summary highlights the critical balance between model selection, accuracy and uncertainty management in predictive analysis.

Table 6
Uncertainty analysis result for method 1.

Model	$\bar{e}$	S_e	Uncertainty bandwidth
XGBoost	−0.0184	0.7876	1.5437
DT	0.1569	0.8194	1.6059
GRNN	0.1195	0.7959	1.5599
RBF	0.3342	2.2126	4.3366
XRT	−0.4226	1.8312	3.5891
Elastic net	0.8433	3.8031	7.4540

4.4.2. Monte Carlo simulation - method 2

The second method of our uncertainty analysis was to perform a Monte Carlo simulation to critically evaluate the variability and confidence in the temperature predictions of our XGBoost model. By running 500 random scenarios, this strategy allowed us to thoroughly investigate the inherent uncertainties contained in the hydrogeochemical data from drilled wells. The feature ranges are listed in Table 7. The simulation outcomes provided a comprehensive temperature distribution, highlighting critical percentiles at P90 (38.4 °C), P50 (median, 33.6 °C), and P10 (25.4 °C) (Fig. 9). These metrics provide an insight into the expected temperature range under different conditions and set both conservative and optimistic benchmarks that are crucial for the design of geothermal systems. In particular, the P90 percentile indicates that temperatures are expected to be below 38.4 °C in 90% of the simulated scenarios, while the P10 percentile suggests that temperatures will be above 25.4 °C in 90% of cases, setting key efficiency and operational benchmarks. The median value, P50 at 33.6 °C, provides a realistic mean value for assessing project feasibility. The significant difference of 13.0 °C between the P10 and P90 percentiles underlines the predictive uncertainty and highlights the importance of developing versatile geothermal systems. This comprehensive Monte Carlo simulation approach supports the planning and design phases of geothermal projects in north-eastern Italy and their economic feasibility and adaptability to geological variability, which is crucial for the promotion of sustainable geothermal energy initiatives.

Fig. 10 shows the density distribution of the XGBoost predictions and reveals a good fit between the density distribution of the XGBoost predictions and the observed values.

Fig. 11 shows a comprehensive comparison of the predicted and actual measurements for each temperature value in both the training and testing datasets. This analysis highlights the ability of the XGBoost model to accurately approximate the target values in the testing dataset and emphasizes its remarkable performance in estimating reservoir temperature at different levels.

Fig. 12 illustrates the Williams plot, which shows points exceeding the leverage limit. If a point exceeds this limit but maintains a standardized residual within the acceptable range of −3 to 3, this indicates a high influence on the model without a significant deviation from the predicted value. The choice of the range from −3 to 3 for the standardized residuals is justified by the properties of the normal distribution. In the normal distribution, approximately 99.7% of the values are within three standard deviations of the mean (Lee et al., 2015). Therefore, standardized residuals within this range are assumed to be consistent with the expected variability of the data, assuming the model is correctly specified, and the error terms are normally distributed. Data points with extreme predictor values can significantly affect the fit of the model, even if they are not outliers in the dependent variable. These influential points can strongly affect the coefficients of the models and

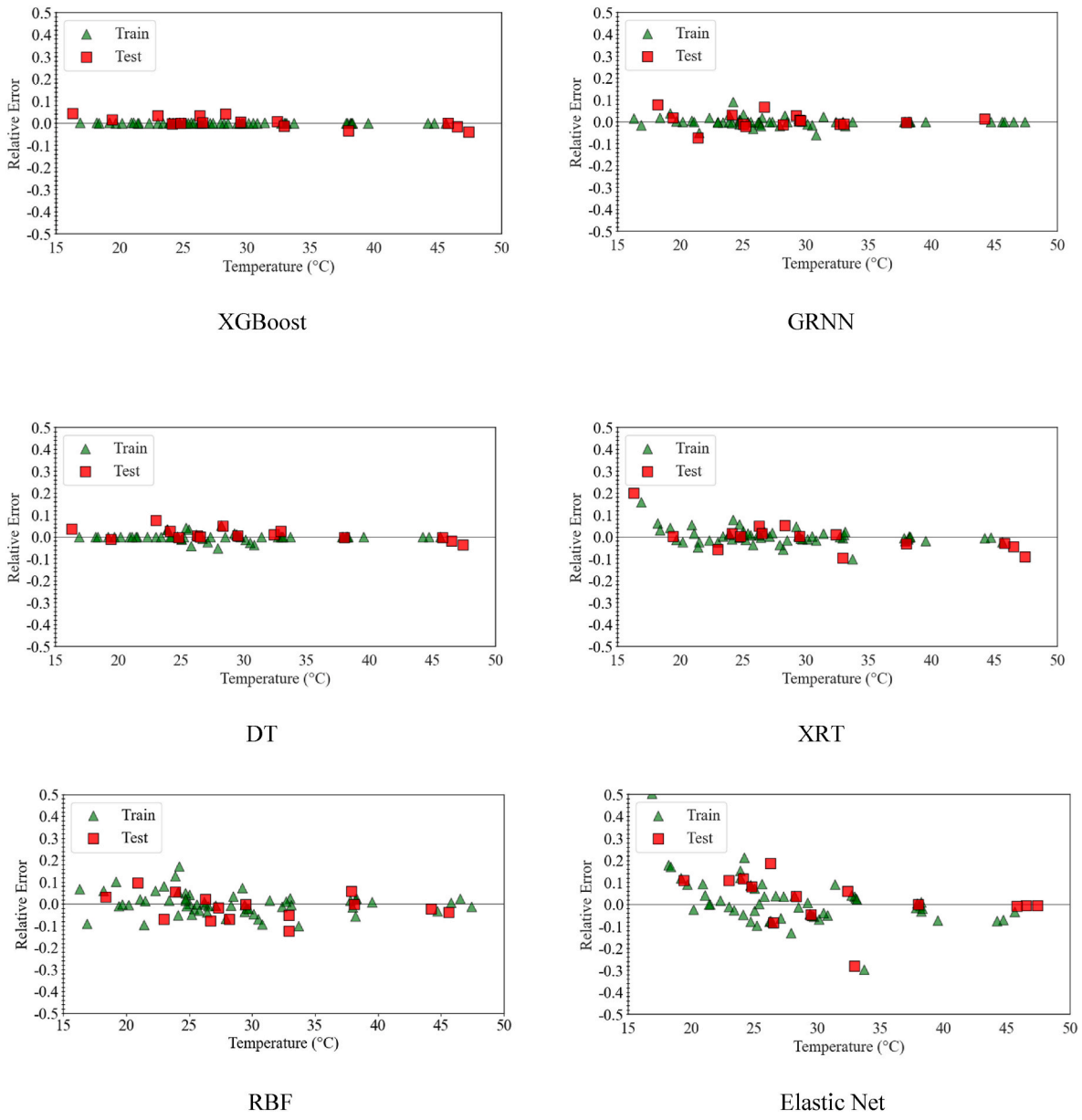


Fig. 8. The Plots show the relative error of the predicted temperature obtained for each ML model.

Table 7
Features range for Monte Carlo simulation.

Parameter	Minimum value	Maximum value
Depth (m)	100	600
pH	6	9
EC (μS-cm)	200	30000
TDS	100	2000
Na (mg-l)	1	5000
K (mg-l)	1	250
Mg (mg-l)	1	250
Ca (mg-l)	1	700
HCO_3 (mg-l)	100	500
NH_4 (mg-l)	0.01	30
NO_3 (mg-l)	0.01	30
Cl (mg-l)	0.1	12000
SO_4 (mg-l)	0.1	850
Fe (μg-l)	1	700

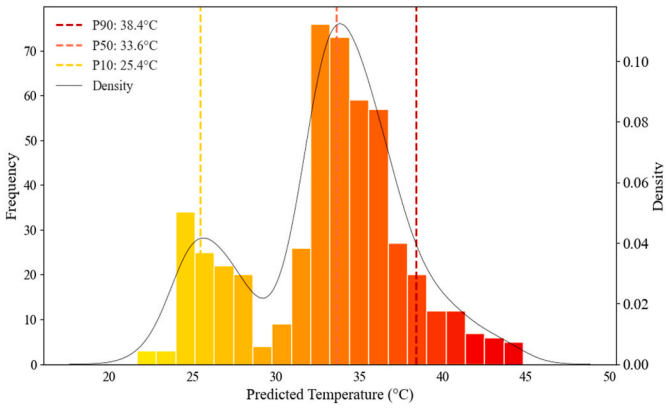


Fig. 9. Monte Carlo simulation result for temperature predictions.

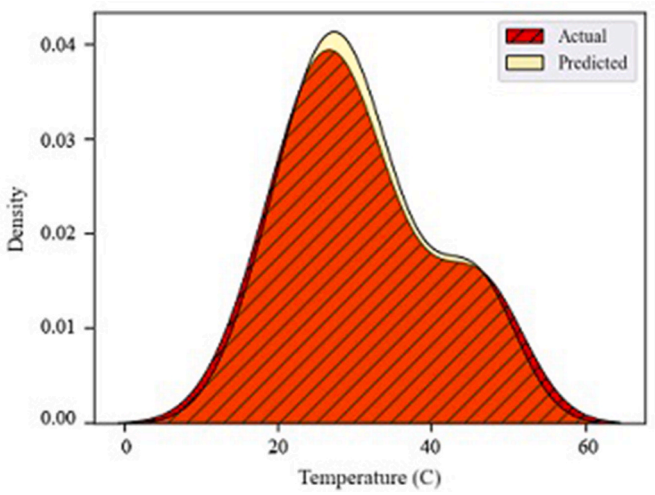


Fig. 10. XGBoost density distribution.

change the overall fit. The Williams Plot identifies these high-leverage or influential points. Points beyond the leverage limit (h*) should be investigated, even if their standardized residuals are within the acceptable range, as they can influence the model. To summarize, exceeding the leverage limit (h* = 0.608) indicates extreme predictor values that could influence the coefficients of the model. An accuracy of 91.5% for training and 93.3% for test samples suggest that the model effectively predicts the majority of data points, and minimizes problems with influential outliers or poorly predicted instances.

4.5. Sensitivity analysis

From a statistical point of view, the sensitivity analysis using the XGBoost algorithm clearly shows the features that most strongly influence the model's ability to predict water temperature fluctuations. (Fig. 13). Of these, HCO_3 proves to be the most important factor with a value of 0.51. Bicarbonate has a significant influence on the buffering capacity of water and helps to maintain a stable environment for geochemical reactions in the aquifer, preserving the thermal properties of the geothermal fluids, which has a direct impact on the efficiency of geothermal heating systems (Wanders et al., 2019). These effects are particularly significant in lower Friulian Plain where the geothermal resources are located in the carbonates. Mg and EC are identified as the next most influential parameters, each with an importance value of 0.11. The influence of magnesium is primarily based on its interaction with water molecules, because the magnesium changes the molecular structure of water and influeces the thermal conductivity. (Li et al., 2020). EC value serves as a measure of the ion content of the water and is linked to the thermal conductivity of the water. Higher EC values indicate a higher ion content, which allows for more efficient heat distribution within the water body, as found by Zhang et al. (2020). The analysis also shows the importance of water body depth, which has a value of 0.08.

The increase of water temperature with depth is partially relate to geothermal gradient, that in the study area ranges from 23 °C/km to 35 °C/km (Stefanini, 1980). This observation is consistent with the data in Fig. 3, showing an increasing trend of temperature with depth. Other parameters such as SO_4, Na, NO_3, NH_4, and Ca have moderately important values around 0.02. These constituents can potentially affect water temperature through various mechanisms, including interactions with other water components and impacts on the heat capacity and thermal properties of the water. TDS and Fe, with importance values of 0.01 each, complete the list of analyzed parameters. Although their direct influence on water temperature is relatively small, their presence and interactions within the aquifer can still subtly influence thermal dynamics. In conclusion, this sensitivity analysis offers a comprehensive overview of the interplay between water quality parameters and temperature regulation. The dominant influence of HCO_3 concentration, alongside the significant contributions of Mg, EC, and depth, highlights

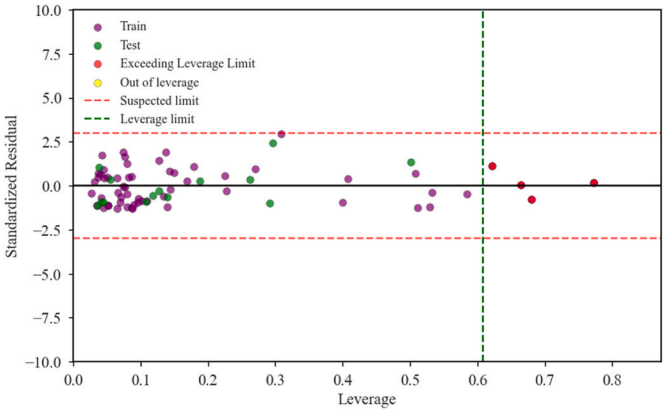


Fig. 12. Williams plot highlighting leverage and residuals in the XGBoost model's temperature predictions.

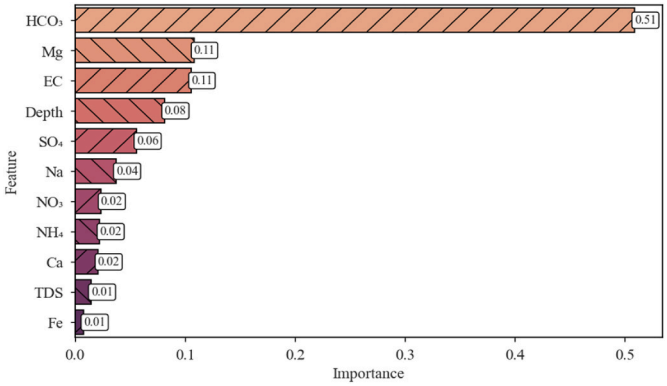


Fig. 13. Sensitivity analysis of each feature used for XGBoost algorithm.

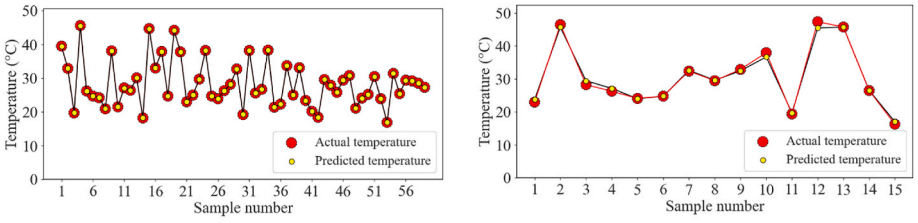


Fig. 11. Comprehensive comparison of predicted and actual temperature measurements for each value of temperature used for training (panel A) and testing (panel B).

the complex mechanisms underlying water temperature dynamics. This analysis provides valuable insights for the development of more accurate aquifer temperature prediction models and thus contributes to a deeper understanding of low enthalpy geothermal resources in the area.

4.6. Comparison with ML models from recent research

The results of the studies proposed by Ibrahim et al. (2023) and Tut Haklidir and Haklidir (2020) were compared with the results of the present study to emphasize the potential of the latter. As shown in Table 4, the XGBoost model in this study had an RMSE of 0.788 and an MAE of 0.587 on the test dataset, demonstrating a high level of accuracy. In comparison, Ibrahim et al. (2023) reported an RMSE of 4.594 and an MAE of 3.968 for the natural gradient boosting (NGB) model, while Tut Haklidir and Haklidir (2020) reported an RMSE of 8.293 and an MAE of 6.451 for their deep neural networks (DNN) model. These comparisons underscore the superior performance of the XGBoost model in this study, even when considering the different databases used in the research. Additionally, Fig. 14 provides a visual comparison of the predictive accuracy and error margins of the XGBoost model compared to the other models, illustrating the nuances of performance on different datasets. In addition, the aforementioned studies were conducted in regions characterized by active and intense tectonic and volcanic activity that exhibit a variety of geochemical features. In contrast, our research focused on low-enthalpy geothermal systems that differ significantly in their geologic settings. This context may provide further insight into the observed discrepancies in model performance. In particular, the models may respond differently to the unique geochemical signatures present in each environment. The differences in performance can generally be attributed to several factors, including the inherent strengths of the XGBoost algorithm in handling different data types and its robustness against overfitting compared to NGB and DNN models. Moreover, the preprocessing techniques, feature selection and hyperparameter tuning specific to each study can significantly affect the ability of the model to generalize to unseen data. The efficiency of the XGBoost model, particularly in minimizing prediction errors in the test dataset, highlights its effectiveness in practical applications, and suggests a more sophisticated approach to model development and validation compared to the referenced studies.

5. Conclusion and future directions

To overcome the limitations of conventional geothermal exploration methods, in this comprehensive study we have harnessed the power of six ML algorithms to predict temperatures in geothermal reservoirs. These data obtained from different wells in the lower Friulian Plain in FVG region, contain not only temperature values, but also key hydrogeochemical parameters and form a solid basis for our models. Of the six ML models analyzed, the XGBoost model performed best. With an R^2 value of 0.9930 and a VAF value of 99.30% on the test data set, as well as the lowest RMSE, MAE, and MAPE values, it demonstrated excellent precision and robustness. This achievement not only illustrates the effectiveness of the XGBoost model in navigating the complex, nonlinear relationships in geothermal temperature data, but also emphasizes its potential to improve prediction accuracy for geothermal reservoir temperatures. Other models, such as DT, showed remarkable predictive capabilities and effectively captured complex data relationships. GRNN showed a strong predictive capacity, while XRT, RBF and elastic net showed varying degrees of predictive success, enriching the comparative analysis of the study.

Our results emphasize the important role of HCO_3, Mg, EC and water depth as critical predictors and provide valuable insights into the geochemical influences on geothermal systems. The careful optimization of model parameters and the application of a robust 80:20 train test split confirm the reliability of our results and pave the way for real applications in geothermal exploration.

This study significantly advances geothermal research through methodological innovation and the strategic use of ML. The dominance of the XGBoost model is particularly noteworthy as it outperforms models such as NGB and DNN, which were employed by other researchers in recent studies. This comparison demonstrates the effectiveness of advanced ML techniques in capturing the complex dynamics of temperatures in geothermal reservoirs. Such advances highlight the significant potential of ML to revolutionize geothermal studies by providing a more accurate and sophisticated understanding of geothermal temperature dynamics.

However, the study also acknowledges limitations, such as the need for larger data sets to generalize the results and the challenge of transferring the methodology to different geological settings.

Future activities of this research include the development of a numerical model that will better characterize the geothermal reservoir. Of course, temperature is an essential component for these future research activities.

In the field of geothermal exploration, where predictive modeling is of great importance, there is still room for further research. In particular, attention can be drawn to the following areas.

- Optimized model transferability: to achieve this, efforts need to focus on developing techniques that improve the transferability of ML models to different geological regions. This includes the refinement of algorithms and the integration of domain-specific knowledge to ensure effective predictions in different geological contexts.
- Integration of real-time data: In addition, it is crucial to explore methods to seamlessly integrate real-time data streams, such as

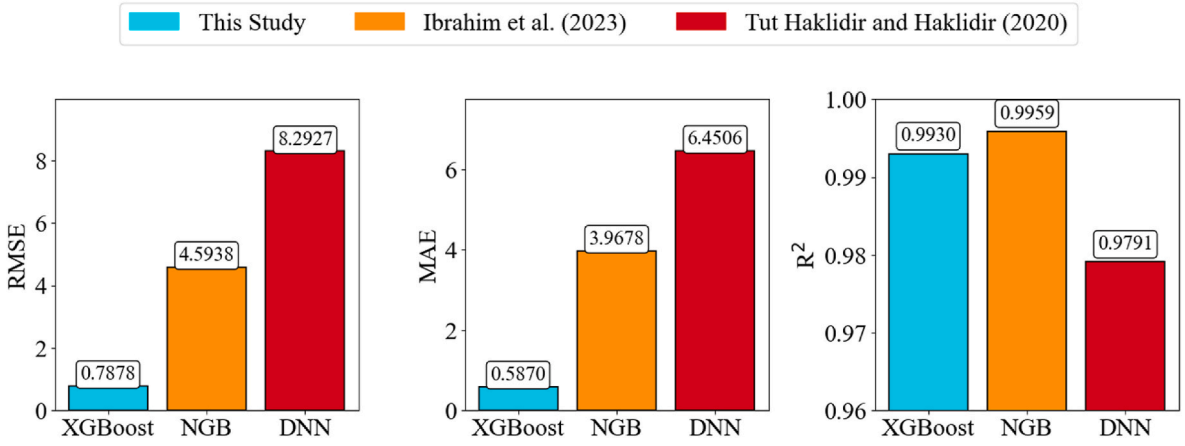


Fig. 14. Plots showing the comparison between the errors of the best ML model of this study and those of Ibrahim et al. (2023) and Tut Haklidir and Haklidir (2020).

continuous monitoring of temperature and geochemical parameters, into ML models. This integration facilitates the dynamic updating of the model and thus improves prediction accuracy by capturing temporal variations in geothermal systems.

By pursuing these future directions, the field can move toward more accurate, reliable, and understandable predictive models that will significantly advance our understanding and utilization of geothermal resources. This progress will not only benefit the scientific community, but also support the global transition to sustainable and renewable energy sources.

CRediT authorship contribution statement

Danial Sheini Dashtgoli: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Data curation, Conceptualization. **Michela Giustiniani:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Project administration, Methodology, Funding acquisition. **Martina Busetti:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation. **Claudia Cherubini:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Danial Sheini Dashtgoli reports financial support was provided by the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR)). If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This study was carried out within the PNRR research activities of the consortium iNEST (Interconnected North-East Innovation Ecosystem) funded by the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) – Missione 4 Componente 2, Investimento 1.5 – D.D. 1058 23/06/2022, ECS_00000043). This manuscript reflects only the Authors' views and opinions.

References

Abrasaldo, P.M.B., Zarrouk, S.J., Kempa-Liehr, A.W., 2024. A systematic review of data analytics applications in above-ground geothermal energy operations. Renew. Sustain. Energy Rev. 189, 113998 https://doi.org/10.1016/j.rser.2023.113998.

Afrasiabi, S., Afrasiabi, M., Parang, B., Mohammadi, M., Samet, H., Dragicevic, T., 2022. Fast GRNN-based method for distinguishing inrush currents in power transformers. IEEE Trans. Ind. Electron. 69 (8), 8501–8512. https://doi.org/10.1109/TIE.2021.3109535.

Agemar, T., Schellschmidt, R., Schulz, R., 2012. Subsurface temperature distribution in Germany. Geothermics 44, 65–77. https://doi.org/10.1016/j.geothermics.2012.07.002.

Ahmmed, B., Vesselinov, V.V., 2022. Machine learning and shallow groundwater chemistry to identify geothermal prospects in the Great Basin, USA. Renew. Energy 197, 1034–1048. https://doi.org/10.1016/j.renene.2022.08.024.

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna: a next-generation hyperparameter optimization framework. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631. https://doi.org/10.1145/3292500.3330701.

Arnórsson, S., Gunnlaugsson, E., Svavarsson, H., 1983. The chemistry of geothermal waters in Iceland. III. Chemical geothermometry in geothermal investigations. Geochem. Cosmochim. Acta 47 (3), 567–577. https://doi.org/10.1016/0016-7037(83)90278-8.

Babyak, M.A., 2004. What you see may not Be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom. Med. 66 (3), 411–421. https://doi.org/10.1097/01.psy.0000127692.23278.a9.

Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32. https://doi.org/10.1023/A:1010933404324.

Brunton, S.L., Proctor, J.L., Kutz, J.N., 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proc. Natl. Acad. Sci. USA 113 (15), 3932–3937. https://doi.org/10.1073/pnas.1517384113.

Busetti, M., Volpi, V., Nicolich, R., Barison, E., Romeo, R., Baradello, L., Giustiniani, M., Marchi, M., Zonlla, C., Wardell, N., 2010. Dinaric tectonic features in the gulf of trieste (northern Adriatic). Bollettino Di Geofisica Teorica e Applicata 51 (2–3), 117–128.

Busetti, M., Zgur, F., Vrabec, M., Facchin, L., Pelos, C., Romeo, R., Sormani, L., Slavec, P., Tomini, I., Visnovich, G., 2013. Neotectonic reactivation of Meso-Cenozoic structures in the Gulf of Trieste and its relationship with fluid seepings. In: Atti del 32° Convegno del Gruppo Nazionale di Geofisica della Terra Solida (GNGTS), pp. 29–34. *Trieste, 19-21 novembre 2013*.

Chakrabarti, D., 2022. Robust linear classification from limited training data. Mach. Learn. 111 (5), 1621–1649. https://doi.org/10.1007/s10994-021-06093-5.

Chandra, T.B., Verma, K., 2020. Analysis of quantum noise-reducing filters on chest X-ray images: a review. Measurement 153, 107426. https://doi.org/10.1016/j.measurement.2019.107426.

Chandrashekar, G., Sahin, F., 2014. A survey on feature selection methods. Comput. Electr. Eng. 40 (1), 16–28. https://doi.org/10.1016/j.compeleceng.2013.11.024.

Chen, T., Guestrin, C., 2016. Xgboost: a scalable tree boosting system. Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining 785–794.

Choi, S.-W., 2009. The effect of outliers on regression analysis: regime type and foreign direct investment. Quarterly Journal of Political Science 4 (2), 153–165. https://doi.org/10.1561/100.00008021.

Cigizoglu, H.K., Alp, M., 2006. Generalized regression neural network in modelling river sediment yield. Adv. Eng. Software 37 (2), 63–68. https://doi.org/10.1016/j.advengsoft.2005.05.002.

Cimolino, A., Della Vedova, B., Nicolich, R., Barison, E., Brancatelli, G., 2010. New evidence of the outer Dinaric deformation front in the Grado area (NE-Italy). Rendiconti Lincei 21 (S1), 167–179. https://doi.org/10.1007/s12210-010-0096-y.

Correa, C.D., Chan, Y.-H., Ma, K.-L., 2009. A framework for uncertainty-aware visual analytics. In: 2009 IEEE Symposium on Visual Analytics Science and Technology, pp. 51–58. https://doi.org/10.1109/VAST.2009.5332611.

Dal Cin, M., Böhm, G., Busetti, M., Picotti, S., Zgur, F., Camerlenghi, A., 2022. 3D velocity-depth model from multichannel seismic in the Dinaric foredeep of the Gulf of Trieste (Adriatic Sea), at the NE edge of Adria plate. Tectonophysics 838, 229470. https://doi.org/10.1016/j.tecto.2022.229470.

Della Vedova, B., Cimolino, A., Castelli, E., Brancatelli, G., 2015. Geothermal heating and cooling in the FVG region: the Grado district heating and the Pontebba ice rink plants. Status and future in the Peri–Adriatic Area–Veli Lošinj (Croatia). Proceedings of the Workshop on Geothermal Energy 65–76.

Fantoni, R., Catellani, D., Merlini, S., Rogledi, S., Venturini, S., 2002. La registrazione degli eventi deformativi cenozoici nell'avampaese Veneto-Friulano. Mem. Soc. Geol. It 57, 301–313.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn. 63 (1), 3–42. https://doi.org/10.1007/s10994-006-6226-1.

Ghaffari-Razin, S.R., Rastbood, A., Hooshangi, N., 2023. Regional application of generalized regression neural network in ionosphere spatio-temporal modeling and forecasting. GPS Solut. 27 (1), 51. https://doi.org/10.1007/s10291-022-01389-y.

Giustiniani, M., Busetti, M., Dal Cin, M., Barison, E., Cimolino, A., Brancatelli, G., Baradello, L., 2022. Geophysical and geological views of potential water resources in the north-eastern adriatic sea. Geosciences 12 (3), 139. https://doi.org/10.3390/geosciences12030139.

Guan, B., Xia, J., Liu, Y., Zhang, H., Zhou, C., 2023. Near-surface radial anisotropy tomography of geothermal reservoir using dense seismic nodal array. J. Phys. Conf. 2651 (1), 012023 https://doi.org/10.1088/1742-6596/2651/1/012023.

Gudala, M., Tariq, Z., Govindarajan, S.K., Yan, B., Sun, S., 2024. Fractured geothermal reservoir using CO_2 as geofluid: numerical analysis and machine learning modeling. ACS Omega. https://doi.org/10.1021/acsomega.3c07215 acsomega.3c07215.

Hastie, T., Tibshirani, R., Friedman, J.H., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed. Springer.

Hawkins, D.M., 2004. The problem of overfitting. J. Chem. Inf. Comput. Sci. 44 (1), 1–12. https://doi.org/10.1021/ci0342472.

Ibrahim, B., Konduah, J.O., Ahenkorah, I., 2023. Predicting reservoir temperature of geothermal systems in Western Anatolia, Turkey: a focus on predictive performance and explainability of machine learning models. Geothermics 112, 102727. https://doi.org/10.1016/j.geothermics.2023.102727.

Ishitsuka, K., Mogi, T., Sugano, K., Yamaya, Y., Uchida, T., Kajiwara, T., 2018. Resistivity-based temperature estimation of the kakkonda geothermal field, Japan, using a neural network and neural kriging. Geosci. Rem. Sens. Lett. IEEE 15 (8), 1154–1158. https://doi.org/10.1109/LGRS.2018.2832647.

Jia, X., Lin, Y., Ouyang, M., Wang, X., He, H., 2024. Numerical simulation of hydrothermal flow in the North China Plain: a case study of Henan Province. Geothermics 118, 102910. https://doi.org/10.1016/j.geothermics.2023.102910.

Jones, A.G., 1998. Waves of the future: superior inferences from collocated seismic and electromagnetic experiments. Tectonophysics 286 (1–4), 273–298. https://doi.org/10.1016/S0040-1951(97)00270-9.

Kadri, M., Mohd Muztaza, N., Mohd Nordin, M.N., Zakaria, M.T., Rosli, F.N., Mohammed, M.A., Zulaika, S., 2023. Integrated geophysical methods used to explore geothermal potential areas in Siogung-Ogung, North Sumatra, Indonesia. Bulletin of the Geological Society of Malaysia 76, 47–53. https://doi.org/10.7186/bgsm76202304.

Kianoush, P., Mohammadi, G., Hosseini, S.A., Keshavarz Faraj Khah, N., Afzal, P., 2023. ANN-based estimation of pore pressure of hydrocarbon reservoirs—a case study. Arabian J. Geosci. 16 (5), 302. https://doi.org/10.1007/s12517-023-11373-6.

Kolawole, O., Assaad, R.H., 2023. Modeling and prediction of temporal biogeomechanical properties using novel machine learning approach. Rock Mech. Rock Eng. https://doi.org/10.1007/s00603-023-03353-9.

Kotsiantis, S.B., Kanellopoulos, D., Pintelas, P.E., 2007. Data Preprocessing for Supervised Leaning. https://doi.org/10.5281/ZENODO.1082415.

Kurup, P.U., Griffin, E.P., 2006. Prediction of soil composition from CPT data using general regression neural network. J. Comput. Civ. Eng. 20 (4), 281–289.

Lee, D.K., In, J., Lee, S., 2015. Standard deviation and standard error of the mean. Korean Journal of Anesthesiology 68 (3), 220.

Li, S., Yang, X., Hou, J., Du, W., 2020. A review on thermal conductivity of magnesium and its alloys. J. Magnesium Alloys 8 (1), 78–90. https://doi.org/10.1016/j.jma.2019.08.002.

Louppe, G., 2014. Understanding Random Forests: from Theory to Practice arXiv Preprint arXiv:1407.7502.

Ma, J., Lei, D., Ren, Z., Tan, C., Xia, D., Guo, H., 2023. Automated machine learning-based landslide susceptibility mapping for the three gorges reservoir area, China. Math. Geosci. https://doi.org/10.1007/s11004-023-10116-3.

Maercklin, N., 2005. Characterizing a large shear-zone with seismic and magnetotelluric methods: the case of the Dead Sea Transform. Geophys. Res. Lett. 32 (15), L15303 https://doi.org/10.1029/2005GL022724.

Mahdaviara, M., Larestani, A., Nait Amar, M., Hemmati-Sarapardeh, A., 2022. On the evaluation of permeability of heterogeneous carbonate reservoirs using rigorous data-driven techniques. J. Petrol. Sci. Eng. 208, 109685 https://doi.org/10.1016/j.petrol.2021.109685.

Male, F., Duncan, I.J., 2020. Lessons for machine learning from the analysis of porosity-permeability transforms for carbonate reservoirs. J. Petrol. Sci. Eng. 187, 106825 https://doi.org/10.1016/j.petrol.2019.106825.

Mechie, J., Sobolev, S.V., Ratschbacher, L., Babeyko, A.Y., Bock, G., Jones, A.G., Nelson, K.D., Solon, K.D., Brown, L.D., Zhao, W., 2004. Precise temperature estimation in the Tibetan crust from seismic detection of the α-β quartz transition. Geology 32 (7), 601. https://doi.org/10.1130/G20367.1.

Meng, Joo Er, Wu, Shiqian, Lu, Juwei, Toh, Hock Lye, 2002. Face recognition with radial basis function (RBF) neural networks. IEEE Trans. Neural Network. 13 (3), 697–710. https://doi.org/10.1109/TNN.2002.1000134.

Moraga, J., Duzgun, H.S., Cavur, M., Soydan, H., 2022. The geothermal artificial intelligence for geothermal exploration. Renew. Energy 192, 134–149. https://doi.org/10.1016/j.renene.2022.04.113.

Mudunuru, M.K., Ahmmed, B., Rau, E., Vesselinov, V.V., Karra, S., 2023. Machine learning for geothermal resource exploration in the tularosa basin, New Mexico. Energies 16 (7), 3098. https://doi.org/10.3390/en16073098.

Muñoz, G., Bauer, K., Moeck, I., Schulze, A., Ritter, O., 2010. Exploring the Groß Schönebeck (Germany) geothermal site using a statistical joint interpretation of magnetotelluric and seismic tomography models. Geothermics 39 (1), 35–45. https://doi.org/10.1016/j.geothermics.2009.12.004.

Okoroafor, E.R., Smith, C.M., Ochie, K.I., Nwosu, C.J., Gudmundsdottir, H., Jabs, Aljubran, M., 2022. Machine learning in subsurface geothermal energy: two decades in review. Geothermics 102, 102401. https://doi.org/10.1016/j.geothermics.2022.102401.

Pastore, N., Cherubini, C., Giasi, C.I., 2021. Analysis of gravel back-filled borehole heat exchanger in karst fractured limestone aquifer at local scale. Geothermics 89, 101971. https://doi.org/10.1016/j.geothermics.2020.101971.

Pérez-Zárate, D., Santoyo, E., Acevedo-Anicasio, A., Díaz-González, L., García-López, C., 2019. Evaluation of artificial neural networks for the prediction of deep reservoir temperatures using the gas-phase composition of geothermal fluids. Comput. Geosci. 129, 49–68. https://doi.org/10.1016/j.cageo.2019.05.004.

Petrini, R., Italiano, F., Ponton, M., Slejko, F.F., Aviani, U., Zini, L., 2013. Geochemistry and isotope geochemistry of the Monfalcone thermal waters (northern Italy): inference on the deep geothermal reservoir. Hydrogeol. J. 21 (6), 1275–1287. https://doi.org/10.1007/s10040-013-1007-y.

Porkhial, S., Salehpour, M., Ashraf, H., Jamali, A., 2015. Modeling and prediction of geothermal reservoir temperature behavior using evolutionary design of neural networks. Geothermics 53, 320–327. https://doi.org/10.1016/j.geothermics.2014.07.003.

Puppala, H., Saikia, P., Kocherlakota, P., Suriapparao, D.V., 2023. Evaluating the applicability of neural network to determine the extractable temperature from a shallow reservoir of Puga geothermal field. International Journal of Thermofluids 17, 100259. https://doi.org/10.1016/j.ijft.2022.100259.

Qiao, Y., Zuo, Y., Tu, S., Zhang, J., Yang, M., Zhang, T., 2023. Strata temperatures and geothermal resource evaluation in the Dongpu depression, bohai bay basin, north China. Sci. Rep. 13 (1), 3630. https://doi.org/10.1038/s41598-023-30760-3.

Qin, Z., Jiang, A., Faulder, D., Cladouhos, T.T., Jafarpour, B., 2024. Physics-guided deep learning for prediction of energy production from geothermal reservoirs. Geothermics 116, 102824. https://doi.org/10.1016/j.geothermics.2023.102824.

Quinlan, J.R., 1986. Induction of decision trees. Mach. Learn. 1 (1), 81–106. https://doi.org/10.1007/BF00116251.

Ranstam, J., Cook, J.A., 2018. LASSO regression. Br. J. Surg. 105 (10) https://doi.org/10.1002/bjs.10895, 1348–1348.

Rau, E.G., Ahmmed, B., Yeates, D., Atchley, S., Mudunuru, M., Frash, L., 2023. *Unsupervised machine Learning and Magnetotellurics for geothermal play Development in the tularosa basin, New Mexico* [preprint]. SSRN. https://doi.org/10.2139/ssrn.4459216.

Rosner, B., Glynn, R.J., Lee, M.T., 2006. The Wilcoxon signed rank test for paired comparisons of clustered data. Biometrics 62 (1), 185–192. https://doi.org/10.1111/j.1541-0420.2005.00389.x.

Safaei-Farouji, M., Vo Thanh, H., Sheini Dashtgoli, D., Yasin, Q., Radwan, A.E., Ashraf, U., Lee, K.-K., 2022a. Application of robust intelligent schemes for accurate modelling interfacial tension of CO2 brine systems: implications for structural CO2 trapping. Fuel 319, 123821. https://doi.org/10.1016/j.fuel.2022.123821.

Safaei-Farouji, M., Vo Thanh, H., Sheini Dashtgoli, D., Yasin, Q., Radwan, A.E., Ashraf, U., Lee, K.-K., 2022b. Application of robust intelligent schemes for accurate modelling interfacial tension of CO2 brine systems: implications for structural CO2 trapping. Fuel 319, 123821. https://doi.org/10.1016/j.fuel.2022.123821.

Sarailidis, G., Wagener, T., Pianosi, F., 2023. Integrating scientific knowledge into machine learning using interactive decision trees. Comput. Geosci. 170, 105248 https://doi.org/10.1016/j.cageo.2022.105248.

Schwenker, F., Kestler, H.A., Palm, G., 2001. Three learning phases for radial-basis-function networks. Neural Network. 14 (4–5), 439–458. https://doi.org/10.1016/S0893-6080(01)00027-2.

Shahdi, A., Lee, S., Karpatne, A., Nojabaei, B., 2021. Exploratory analysis of machine learning methods in predicting subsurface temperature and geothermal gradient of Northeastern United States. Geoth. Energy 9 (1), 18. https://doi.org/10.1186/s40517-021-00200-4.

Sharifi, R., Moore, F., Mohammadi, Z., Keshavarzi, B., 2016. Estimation of deepwater temperature and hydrogeochemistry of springs in the Takab geothermal field, West Azerbaijan, Iran. Environ. Monit. Assess. 188 (1), 75. https://doi.org/10.1007/s10661-015-5037-x.

Sheini Dashtgoli, D., Dehnad, M.H., Mobinipour, S.A., Giustiniani, M., 2024. Performance comparison of machine learning algorithms for maximum displacement prediction in soldier pile wall excavation. Undergr. Space 16, 301–313. https://doi.org/10.1016/j.undsp.2023.09.013.

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. Journal of Big Data 6 (1), 60. https://doi.org/10.1186/s40537-019-0197-0.

Specht, D.F., 1991a. A general regression neural network. IEEE Trans. Neural Network. 2 (6), 568–576. https://doi.org/10.1109/72.97934. Scopus.

Specht, D.F., 1991b. A general regression neural network. IEEE Trans. Neural Network. 2 (6), 568–576.

Stefanini, S., 1980. Il termalismo delle acque artesiane nelle lagune di Marano di Grado e nelle aree adiacenti. Rassegna Tecnica del Friuli Venezia Giulia 5, 19–24.

Stober, I., Bucher, K., 2021. Geothermal Energy: from Theoretical Models to Exploration and Development. Springer International Publishing. https://doi.org/10.1007/978-3-030-71685-1.

Tahmassebi, A., Gandomi, A.H., Meyer-Baese, A., 2018. A pareto front based evolutionary model for airfoil self-noise prediction. In: 2018 IEEE Congress on Evolutionary Computation (CEC), pp. 1–8. https://doi.org/10.1109/CEC.2018.8477987.

Tavares, C., Nascimento, N., Alencar, P., Cowan, D., 2022. Adaptive method for machine learning model selection in data science projects. In: 2022 IEEE International Conference on Big Data (Big Data), pp. 2682–2688. https://doi.org/10.1109/BigData55660.2022.10020386.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. Roy. Stat. Soc. B Stat. Methodol. 58 (1), 267–288.

Tomaszewska, B., Pająk, L., Bundschuh, J., Bujakowski, W., 2018. Low-enthalpy geothermal energy as a source of energy and integrated freshwater production in inland areas: technological and economic feasibility. Desalination 435, 35–44. https://doi.org/10.1016/j.desal.2017.12.032.

Tut Haklidir, F.S., Haklidir, M., 2020. Prediction of reservoir temperatures using hydrogeochemical data, western anatolia geothermal systems (Turkey): a machine learning approach. Nat. Resour. Res. 29 (4), 2333–2346. https://doi.org/10.1007/s11053-019-09596-0.

Uçar, M.K., Nour, M., Sindi, H., Polat, K., 2020. The effect of training and testing process on machine learning in biomedical datasets. Math. Probl Eng. 2020, 1–17. https://doi.org/10.1155/2020/2836236.

Varol Altay, E., Gurgenc, E., Altay, O., Dikici, A., 2022. Hybrid artificial neural network based on a metaheuristic optimization algorithm for the prediction of reservoir temperature using hydrogeochemical data of different geothermal areas in Anatolia (Turkey). Geothermics 104, 102476. https://doi.org/10.1016/j.geothermics.2022.102476.

Vesnaver, A., Böhm, G., Busetti, M., Dal Cin, M., Zgur, F., 2021. Broadband Q-factor imaging for geofluid detection in the gulf of trieste (northern adriatic sea). Front. Earth Sci. 9, 640194 https://doi.org/10.3389/feart.2021.640194.

Vo Thanh, H., Sheini Dashtgoli, D., Zhang, H., Min, B., 2023. Machine-learning-based prediction of oil recovery factor for experimental CO2-Foam chemical EOR: implications for carbon utilization projects. Energy 278, 127860. https://doi.org/10.1016/j.energy.2023.127860.

Wanders, N., Van Vliet, M.T.H., Wada, Y., Bierkens, M.F.P., Van Beek, L.P.R., 2019. High-Resolution global water temperature modeling. Water Resour. Res. 55 (4), 2760–2778. https://doi.org/10.1029/2018WR023250.

Wang, Z., Guo, L., Gong, H., Li, X., Zhu, L., Sun, Y., Chen, B., Zhu, X., 2023. Land subsidence simulation based on Extremely Randomized Trees combined with Monte Carlo algorithm. Comput. Geosci. 178, 105415 https://doi.org/10.1016/j.cageo.2023.105415.

Wasserman, P.D., 1993. Advanced Methods in Neural Computing. John Wiley & Sons, Inc.

Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., Liu, L., Wu, H., Song, Y., 2020. Improved 1 km resolution $PM_{2.5}$ estimates across China using enhanced space–time extremely randomized trees. Atmos. Chem. Phys. 20 (6), 3273–3289. https://doi.org/10.5194/acp-20-3273-2020.

Westphal, M., Brannath, W., 2020. Evaluation of multiple prediction models: a novel view on model selection and performance assessment. Stat. Methods Med. Res. 29 (6), 1728–1745. https://doi.org/10.1177/0962280219854487.

Wu, X.-J., Jiang, G.-C., Wang, X.-J., Fang, N., Zhao, L., Ma, Y.-M., Luo, S.-J., 2013. Prediction of reservoir sensitivity using RBF neural network with trainable radial basis function. Neural Comput. Appl. 22 (5), 947–953. https://doi.org/10.1007/s00521-011-0787-z.

Xu, J., Liu, Z., Hong, G., Cao, Y., 2024. A new machine-learning-based calibration scheme for MODIS thermal infrared water vapor product using BPNN, GBDT, GRNN, KNN, MLPNN, RF, and XGBoost. IEEE Trans. Geosci. Rem. Sens. 62, 1–12. https://doi.org/10.1109/TGRS.2024.3356578.

Yan, B., Xu, Z., Gudala, M., Tariq, Z., Sun, S., Finkbeiner, T., 2024. Physics-informed machine learning for reservoir management of enhanced geothermal systems. Geoenergy Science and Engineering 234, 212663. https://doi.org/10.1016/j.geoen.2024.212663.

Yang, F., Zhu, R., Zhou, X., Zhan, T., Wang, X., Dong, J., Liu, L., Ma, Y., Su, Y., 2022. Artificial neural network based prediction of reservoir temperature: a case study of Lindian geothermal field, Songliao Basin, NE China. Geothermics 106, 102547. https://doi.org/10.1016/j.geothermics.2022.102547.

Yao, P., Yu, Z., Zhang, Y., Xu, T., 2023. Application of machine learning in carbon capture and storage: an in-depth insight from the perspective of geoscience. Fuel 333, 126296. https://doi.org/10.1016/j.fuel.2022.126296.

Zecchin, M., Busetti, M., Donda, F., Dal Cin, M., Zgur, F., Brancatelli, G., 2022. Plio-Quaternary sequences and tectonic events in the northern Adriatic Sea (northern Italy). Mar. Petrol. Geol. 142, 105745 https://doi.org/10.1016/j.marpetgeo.2022.105745.

Zhang, K., Wu, X., Niu, R., Yang, K., Zhao, L., 2017. The assessment of landslide susceptibility mapping using random forest and decision tree methods in the Three Gorges Reservoir area, China. Environ. Earth Sci. 76 (11), 405. https://doi.org/10.1007/s12665-017-6731-5.

Zhang, Z., Jia, Y., Zhao, J., 2020. Effect of magnesium ion concentration on the scale inhibition of heat exchanger in circulating cooling water under alternating electric field. Appl. Sci. 10 (16), 5491. https://doi.org/10.3390/app10165491.

Zini, l., Calligaris, C., Treu, F, Iervolino, D., Lippi, F., Cimolino, A., Mereu, A., Barison, E., Zavagno, E., Cucchi, F., Schak, R., Deana, A., 2011. Risorse idriche sotterranee del Friuli Venezia Giulia: sostenibilità dell'attuale utilizzo. Università degli Studi di Trieste, 89 pp.

Zini, L., Cucchi, F., Franceschini, G., Treu, F., 2008. Caratteristiche idrologiche e geochimiche delle riserve acquifere sotterranee della pianura del Friuli Venezia Giulia.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic Net. J. Roy. Stat. Soc. B 67 (2), 301–320. JSTOR. http://www.jstor.org/stable/3647580.