AiION: novel deep learning chemical geothermometer for temperature prediction of deep geothermal reservoirs.

ALGAIAR, M., BANO, S., LASHIN, A., HOSSAIN, M., FAISAL, N.H. and ABU SALEM, H.S.

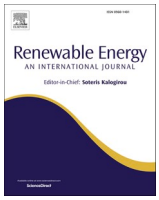
2025

Renewable Energy 248 (2025) 123154

Contents lists available at ScienceDirect

Renewable Energy

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



AiION - Novel deep learning chemical geothermometer for temperature prediction of deep geothermal reservoirs

Mahmoud AlGaiar [a,*], Shahana Bano [a], Aref Lashin [b], Mamdud Hossain [a], Nadimul Haque Faisal [a,**], Hend S. Abu Salem [c]

[a] School of Computing, Engineering and Technology, Robert Gordon University, Garthdee Road, Aberdeen, AB10 7GJ, UK
[b] Petroleum and Natural Gas Engineering Department, College of Engineering, King Saud University, Riyadh, 11421, Saudi Arabia
[c] Geology Department, Faculty of Science, Cairo University, Giza, 12613, Egypt

ARTICLE INFO

Keywords:
Geothermometry
Artificial intelligence
Geothermal energy
Geothermal exploration
Machine learning
Random forest
Gradient boosting
Artificial neural network
Deep neural network

ABSTRACT

This study introduces AiION, a novel deep learning chemical geothermometer designed to predict deep geothermal reservoir temperatures and address the limitations of traditional geothermometry methods. By integrating classical geothermometry, multi-component geothermometry, and existing machine learning insights, AiION was trained on a comprehensive dataset of 674 water samples from Nevada. Among four evaluated machine learning algorithms, AiION, a deep neural network model, demonstrated superior performance, explaining over 97 % of the variance in both training and test data. The global applicability of AiION was validated through successful evaluation on 42 new well samples from diverse geothermal fields worldwide. This research significantly advances solute geothermometry by providing a reliable, data-driven tool for geothermal exploration and development, contributing to sustainable energy efforts. The novelty of AiION lies in its large training dataset, high prediction accuracy, and global applicability, which overcome the limitations of traditional and existing machine learning methods for reliable subsurface temperature prediction in diverse geothermal systems.

1. Introduction

Geothermal exploration has traditionally relied on classical geothermometers, including chemical, isotopic and gas geothermometers [1]. These methods assume hydrogeochemical equilibrium between reservoir fluids and host rocks and use concentrations of dissolved minerals such as silica to estimate subsurface temperatures. However, their effectiveness is often compromised by factors such as fluid mixing and degassing during ascent, which can alter the original fluid composition [2]. In addition, site-specific variability and the complexity of subsurface processes further challenge accurate temperature estimates, highlighting the need for more robust geothermometric approaches [3].

Multicomponent geothermometry has emerged as a more advanced method for estimating reservoir temperatures by evaluating saturation indices for multiple mineral-fluid reactions simultaneously [4]. This technique uses numerical modelling and thermodynamic databases to interpret complex geochemical signatures and has demonstrated good performance in various geothermal fields [5]. Despite its advantages, the practical application of multicomponent geothermometry is limited by the need for detailed a priori assumptions regarding the thermodynamic conditions and specific chemical reactions occurring in the subsurface process. Furthermore, its computational intensity often limits its applicability to specific sites rather than broader regions [4].

The emergence of ML techniques represents a transformative shift in addressing the limitations of traditional geothermometric methods. By leveraging extensive geochemical datasets and sophisticated algorithms, ML can identify patterns in the data that may be missed by traditional analyses [6]. This ability is particularly beneficial in complex geological settings and diverse geochemical characteristics. The integration of ML not only improves the accuracy of subsurface temperature predictions but also facilitates the discovery of blind geothermal prospects. Furthermore, the combination of ML with traditional geoscience knowledge allows for a deeper understanding of the interactions between geochemical signatures and subsurface thermal regimes, thereby advancing geothermal resource assessment and management.

Research into the application of ML in geothermal reservoir temperature prediction has developed significantly over the last two

* Corresponding author.
** Corresponding author.
E-mail addresses: m.algaiar@rgu.ac.uk (M. AlGaiar), n.h.faisal@rgu.ac.uk (N.H. Faisal).

https://doi.org/10.1016/j.renene.2025.123154
Received 9 January 2025; Received in revised form 31 March 2025; Accepted 11 April 2025
Available online 12 April 2025

Abbreviations

°C	Degrees Celsius
$\delta^{18}O$	Isotopic Composition of Oxygen-18
δD	Deuterium - Isotopic Composition of Hydrogen
AI	Artificial Intelligence
AiION	Artificial Intelligence Geothermometer Using Ionic Concentrations
ANN	Artificial Neural Network
BHT	Bottom-hole Temperature
BPNN	Back Propagation Neural Network
Ca	Calcium
CBE	Charge Balance Equilibrium
CH_4	Methane
Cl	Chloride
CNN	Convolutional Neural Network
CO_2	Carbon Dioxide
DNN	Deep Neural Network
DT	Decision Tree
F	Fluorine
GB	Gradient Boosting
GBGG	Great Basin Groundwater Geochemical Database
GMM	Gaussian Mixture Models
GWO-MLP	Grey Wolf Optimizer-Multi-Layer Perceptron
HCO_3	Bicarbonate
H_2	Hydrogen Gas
H_2S	Hydrogen Sulfide
IQR	Interquartile Range
K	Potassium
K_T	Thermodynamic Equilibrium Constant
LLNL	Lawrence Livermore National Laboratory
LR	Linear Regression
MAE	Mean Absolute Error
Mg	Magnesium
Mg/L	Milligrams per Liter
ML	Machine Learning
MLP	Multilayer Perceptron
MSE	Mean Square Error
Na	Sodium
NBMG	Nevada Bureau of Mines and Geology
NGB	Natural Gradient Boosting
pH	Power of Hydrogen
ppm	Parts Per Million
Q	Ion Activity Product
R^2	Coefficient of Determination
RF	Random Forest
RMSE	Root Mean Square Error
SHAP	SHapley Additive exPlanations
SI	Saturation Index
SiO_2	Silica
SMU	Southern Methodist University
SO_4	Sulfate
TDS	Total Dissolved Solids
VIF	Variance Inflation Factor
XGB	Xtreme Gradient Boosting

decades. Early studies, such as those by Bayram [7] and Can [8], used artificial neural networks (ANNs) to improve upon conventional sodium-potassium geothermometer methods. Later advances included the work of Díaz-Gonzalez et al. [9], who proposed three improved sodium-potassium geothermometers using ML techniques. In 2019, Perez-Zarate et al. [10] presented a novel gaseous ANN geothermometer based on the concentrations of CO_2, H_2S, CH_4 and H_2. Similarly, Acevedo-Anicasio et al. [11] also trained ANN models to obtain eight novel gaseous geothermometers. The aforementioned studies were mostly limited to the improvement of classical geothermometers using ML algorithms. More recently, Haklidir and Haklidir [12] developed a deep neural network (DNN) solute geothermometer that integrated several hydrochemical parameters but relied on a limited dataset of 83 samples to predict reservoir temperatures in Western Anatolia, Turkey. Ibrahim et al. [13] further tested various ML algorithms on Haklidir's data, highlighting the importance of SHapley Additive exPlanations (SHAP) for understanding model contributions. In addition, Altay et al. [14] explored several ML methods for temperature prediction in Anatolia, Turkey, using an extended training dataset of 112 samples, finding success with a Grey Wolf Optimizer Multi-Layer Perceptron (GWO-MLP).

To address the challenges of data availability for training ML models, Yang et al. [15] used numerical simulations of water-rock interactions to generate a geochemical and thermal dataset for the Lindian geothermal field in China. They developed five ANN architectures and used silica and major cations as input variables, achieving lower prediction errors than conventional geothermometers. However, they noted limitations related to simulated data and specific subsurface conditions. To improve generalization and minimize error propagation, Ystroem et al. [16] used global geochemical and in-situ temperature data of 155 water samples from nine geothermal sites to train their supervised multi-layer perceptron (MLP) ANN geothermometer, AnnRG, which demonstrated high accuracy with an RMSE of 9.405 and R^2 of 0.978. A recent study by Sheini Dashtgoli et al. [17] also applied ML techniques to predict geothermal temperatures in the lower Friulian Plain of northeastern Italy using a database of 74 hydrogeochemical data points. The study evaluated six ML models, with the XGBoost model being the most effective, with an R^2 of 0.993 and low error metrics. However, limitations in data set size were noted, suggesting that future research should focus on improving model transferability and integrating real-time data.

This study represents a significant advancement in geothermometry by presenting, for the first time, a large and comprehensive thermogeochemical dataset of 647 water samples from the complex geological environments of Nevada. Recognizing the limitations of classical and multi-component geothermometry and the challenges in data availability and predictive capabilities of existing ML models, this research aimed to develop an improved, data-driven geothermometer. The resulting novelty lies in AiION, a highly accurate and globally applicable deep learning geothermometer trained on this unique dataset, which effectively addresses the scientific problem of unreliable subsurface temperature prediction in diverse geothermal systems. The value of this work stems from its potential to provide accurate temperature predictions for a wide range of geothermal environments, thereby significantly advancing geothermal exploration and development efforts worldwide.

2. Background

The Great Basin region of western North America is a significant contributor to geothermal energy production, with an installed capacity of approximately 1200 MWe and approximately 28 operational power plants [18,19]. This study focuses on the state of Nevada, which is almost entirely within the Great Basin region. Nevada has an installed capacity of approximately 720 MWe from 14 active geothermal power plants, contributing approximately 8 % of the state's electricity generation [19]. These figures highlight Nevada's leadership in the geothermal sector in the United States and its potential for future development. The state has a diverse range of geothermal resources

characterized by unique geological features and significant thermal activity, including areas of known geothermal resources such as hot springs and fumaroles. In addition, approximately 75 % of Nevada's geothermal potential is believed to be in "blind" systems - those without surface manifestations - indicating significant untapped resources [20]. Recent geological assessments have successfully identified new geothermal systems, further highlighting Nevada's potential for future development [21]. The availability of comprehensive geochemical data, such as the Great Basin Groundwater Geochemical Database (GBGG), allows researchers to analyze geochemical signatures that are critical to identifying viable geothermal resources within the region, making Nevada an ideal candidate for in-depth geothermal analysis and research initiatives aimed at optimizing resource development and management.

2.1. Data source

The data used in this study are from the GBGG database, originally compiled by the Nevada Bureau of Mines and Geology (NBMG) and updated by the University of Nevada at Reno, which is a comprehensive repository of hydrogeochemical data containing 51,454 samples from a variety of sites with different characteristics designed to investigate the regional geothermal potential within the Great Basin [18]. This comprehensive dataset includes sample temperature measurements, aqueous species, isotopes, geothermometry, system parameters and spatial information. This diverse dataset facilitates the application of different analytical techniques such as classical geothermometry, multicomponent geothermometry and hydrogeochemical modelling for geothermal resource assessment and exploration in the Great Basin region.

2.2. Geology of the study area

The Great Basin region is one of the largest geothermal provinces in the world, characterized by predominantly amagmatic geothermal systems. These systems are associated with anomalously high geothermal gradients due to crustal extension and thinning, as described by Lachenbruch and Sass [22], and Blackwell et al. [23]. The landscape of the region is characterized by deep valleys separating large mountain ranges, with faults defining the boundaries between these valleys and mountains. The repeating pattern of mountain ranges flanked by parallel faults and valleys defines the geological structure of the region. Most geothermal systems in the Great Basin are controlled by Quaternary normal faults and are typically located near the margins of actively subsiding basins. Consequently, fluids upwelling along these faults often flow into permeable subsurface sediments within the basins, rather than reaching the surface directly along the fault lines. These upwells can manifest many kilometers from their deeper sources or remain hidden with no surface expression, as noted by Siler et al. [24].

Located within the Great Basin, Nevada has a complex and heterogeneous structural framework shaped by diverse tectonic and depositional environments. Its geological settings are primarily influenced by extensional and transtensional tectonics within the Basin and Range Province, with a wide variety of geological formations including volcanic, sedimentary and metamorphic rocks that have undergone significant tectonic activity since the late Miocene [25]. The presence of numerous fault systems, reflecting their rich tectonic history, has shaped the characteristic ridge and graben structures of the region. These faults often act as conduits for geothermal fluids and heat transfer, leading to the formation of geothermal systems, particularly in areas with high strain rates and a greater density of Quaternary faults, such as the Walker Lane [26]. Table 1 provides a detailed summary of the structural settings for known geothermal systems in Nevada, showing that step-overs or relay ramps in normal fault zones are the most favorable setting, hosting approximately 39 % of the systems [21]. The systems are grouped into eight types of structural settings, with few systems classified as hot sedimentary aquifers. The systems were also grouped according to the maximum temperature obtained either by direct measurement of wells or springs or by geothermometry. The percentage of electricity produced from each setting is also included in the table, excluding direct-use applications.

The complex interplay of igneous, sedimentary and metamorphic processes over billions of years has resulted in a diverse range of rock types and structures in Nevada. These processes are primarily associated with major geological events such as seafloor spreading, ancient plate collisions, and volcanic activity, all of which are critical to understanding the geothermal potential of the region [27]. Major geographical features such as the Sierra Nevada and the East Humboldt Range further define the geological landscape of Nevada. The Sierra Nevada creates a rain shadow effect that imparts desert-like conditions to much of the Great Basin. Fig. 1 provides detailed insights into the tectonic map of the western United States highlighting the major plate boundaries, including the Central Nevada Seismic Belt (CNSB), Eastern California Shear Zone (ECSZ), Garlock Fault (GF), Hurricane Sevier Fault System (HSFS), Juan de Fuca Plate (JDF), San Andreas Fault (SN), Sierra Nevada (SN), Walker Lane Belt (WLB), and Wasatch Fault System (WFS). Yellow arrows indicate geodetic velocities relative to a fixed "Colorado Plateau" reference frame [28].

2.3. Hydrogeology of the study area

The hydrogeology of Nevada is characterized by basin and range physiography, involving alternating mountain ranges and valleys with predominantly internal drainage systems that influence groundwater flow, recharge and discharge mechanisms. Groundwater flow is highly variable, influenced by lithology and structural features [29]. The water table generally falls from the mountains to the basin floor, with gradients varying according to factors such as recharge and permeability

Table 1
Summary of structural settings for known geothermal systems in Nevada based on percentage of total number of systems [21].

Structural Setting Type	Average Abundance (%) of Each Setting for All Geothermal Systems	Relative Abundance (%) of each Structural Setting for each Category of Geothermal Systems				
		Grouped by Maximum Temperature				Power Producing (%)
		36–90 °C	90–130 °C	130–160 °C	>160 °C	
Accommodation Zones	5.2	5.8	3.3	0.0	10.7	15.8
Displ. Transfer Zones	4.1	1.9	3.3	12.1	3.6	15.8
Pull-Aparts	4.1	2.9	0.0	6.1	10.7	5.3
Step-Overs	39.2	40.8	33.3	36.4	42.9	47.4
Fault Intersections	20.6	18.4	30.0	27.3	10.7	0.0
Fault Terminations	23.2	25.2	30.0	15.2	17.9	15.8
Fault Bends	1.0	1.9	0.0	0.0	0.0	0.0
Major Normal Faults	1.0	1.9	0.0	0.0	0.0	0.0
Hot Sed. Aquifers	1.5	1.0	0.0	3.0	3.6	0.0
% Sum	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**	**100.0**

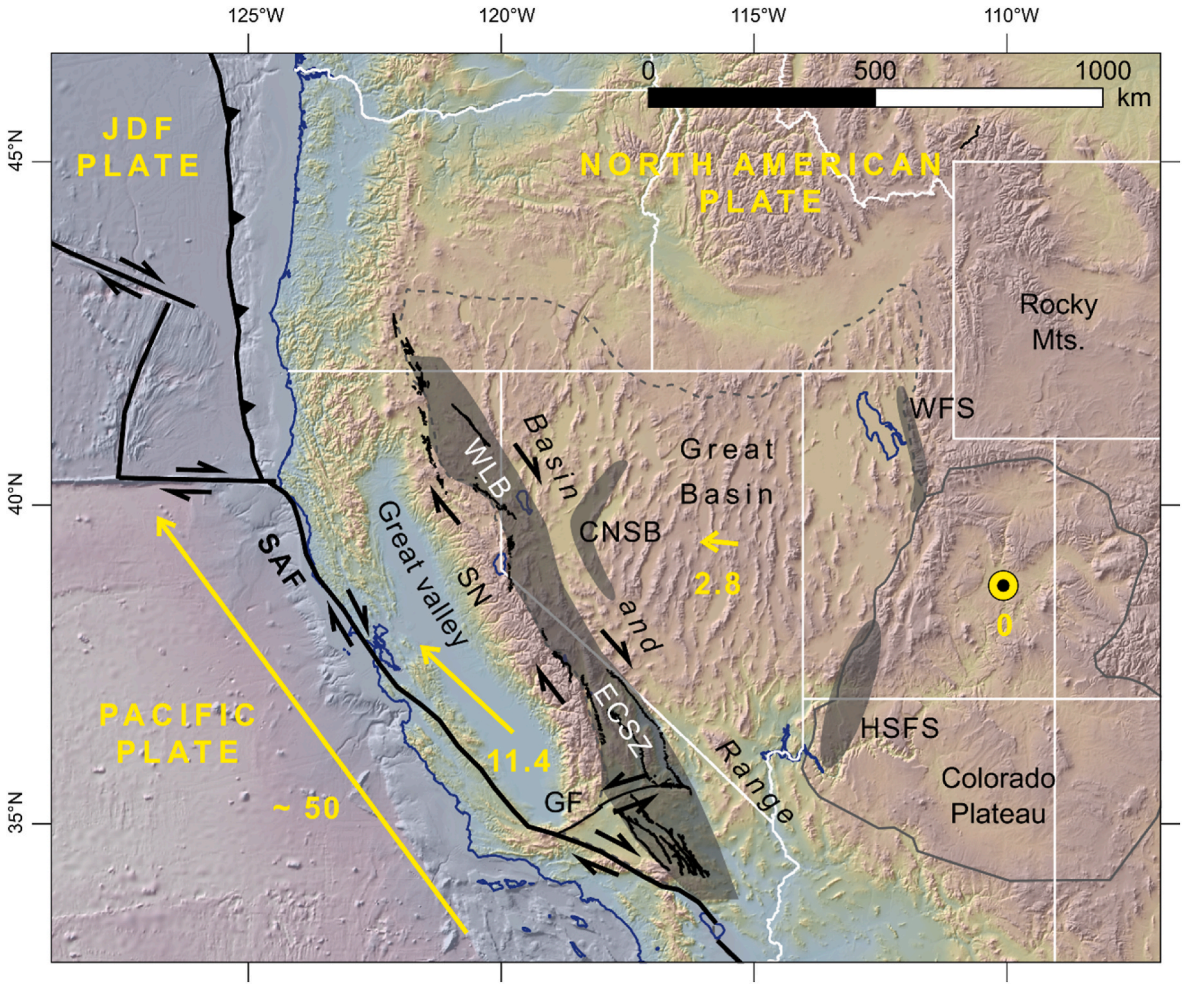


Fig. 1. Tectonic map of the western United States and the major plate boundaries [28] (Reproduced under the terms of "Fair Use" Permission).

[30].

The subsurface consists of two primary hydrogeological units: (1) consolidated rocks (including carbonate rocks, volcanic flows, and intrusive and metamorphic rocks), with carbonate rocks exhibiting the highest horizontal hydraulic conductivity and playing a critical role in groundwater flow and storage, and potentially hosting both conduction- and convection-dominated geothermal systems; (2) unconsolidated sediments (including alluvial slopes, valley floors, fluvial deposits, and playas), which are categorized based on flow regime, topography, and the presence of stream channels [31]. Fig. 2 illustrates the groundwater

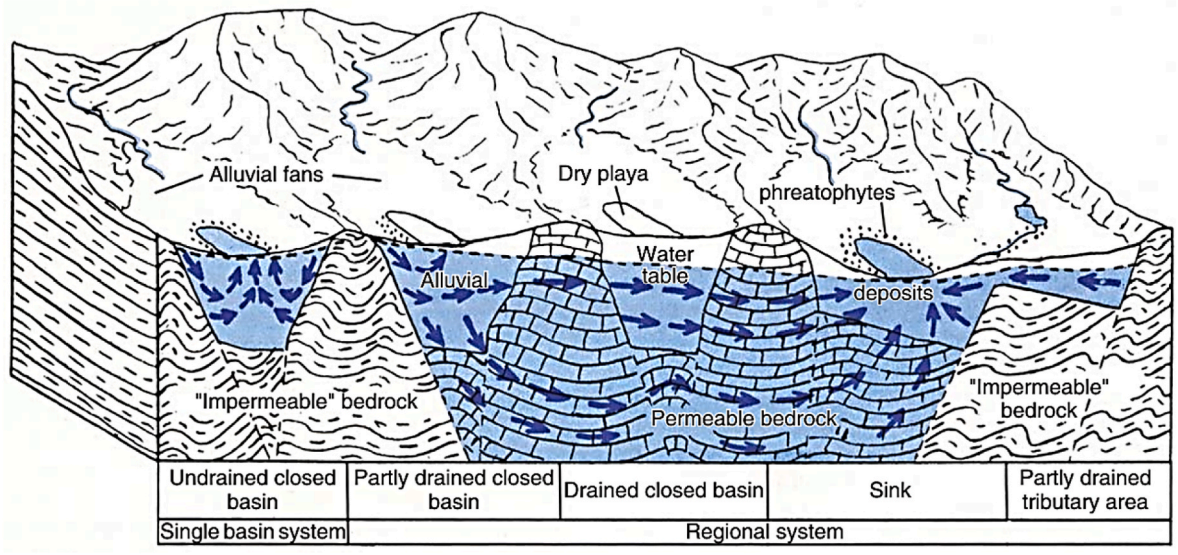


Fig. 2. Groundwater flow characteristics for different types of hydrographic areas in Nevada [29] (Figure courtesy of the U.S. Geological Survey).

flow characteristics for different types of hydrographic areas in Nevada and shows that the amount of groundwater discharge depends on the rock types underlying and bounding the hydrographic areas [29]. Major aquifer systems include carbonate-rock aquifers and basin-fill aquifers. Carbonate-rock aquifers are extensive in the eastern Great Basin, while basin-fill aquifers are composed of older and younger basin-fill deposits.

Over time, mineral precipitation and tectonic activity create alternating periods of enhanced and restricted flow in consolidated rocks. The complex interplay between high hydraulic gradients in low permeability rocks and the availability of recharge creates distinct hydrological features, including perennial streams and springs [32].

Recharge is primarily from winter precipitation stored as snowpack in adjacent mountains, with considerable variation in the timing and frequency of recharge, particularly in the higher mountain ranges where precipitation is greater. Most groundwater discharge is by evapotranspiration, with significant areas of discharge in topographically low parts of valleys where the water table is close to the land surface [33]. Spring discharge is also an important component in the carbonate rock province [34]. The interplay between high hydraulic gradients in low permeability rocks and the availability of precipitation recharge creates distinctive hydrological features, including perennial streams and springs [35]. High groundwater temperatures, often exceeding the mean annual air temperature, with three distinct ranges (4 °C–15 °C, mid-13 °C–18 °C, and mid-18 °C–32 °C), are indicative of deep circulation systems, often associated with faults and thermal springs. These thermal springs, found throughout the state, often have higher temperatures than nearby cold springs, suggesting localized upwelling of geothermal fluids [34].

These hydrogeological characteristics are critical for assessing geothermal resources and understanding groundwater flow patterns in Nevada's diverse geological landscape. The state's unique hydrogeological setting requires site-specific investigations to accurately assess groundwater dynamics and geothermal potential.

2.4. Water chemistry & minerology of the study area

Geothermal waters in Nevada exhibit diverse chemical compositions and mineralogical characteristics, reflecting the complex geological and hydrological setting. The geochemistry of these waters is primarily influenced by deep circulation of meteoric fluids along fault systems, interaction with host rocks, and mixing with magmatic volatiles [36]. The major constituents of Nevada geothermal waters include potassium (K^+), sodium (Na^+), magnesium (Mg^{2+}), calcium (Ca^{2+}), chloride (Cl^-), fluorine (F^-), silica (SiO_2), sulfate (SO_4^{2-}), and bicarbonate (HCO_3^-). Their relative concentrations provide valuable information about reservoir conditions and fluid origin. For example, high Na/K ratios often indicate higher reservoir temperatures, while elevated Ca and Mg concentrations may indicate lower temperatures or shallow circulation [37]. Fig. 3 shows a Schoeller diagram [38] of geothermal wells in Nevada to further illustrate relative ionic concentrations. The diagram highlights the dominance of Na and Cl in many geothermal systems in Nevada, particularly those associated with high temperature reservoirs. The variable HCO_3 concentrations reflect the diversity of water types and the influence of both deep and shallow circulation patterns [39].

The total dissolved solids (TDS) content of Nevada's geothermal waters varies considerably among geothermal systems, reflecting the diverse geological settings and water-rock interactions throughout the state. In general, Nevada's geothermal waters have moderate to high TDS concentrations, ranging from approximately 1000 mg/L to over 10,000 mg/L in some cases [40]. This relatively high TDS content is typical of many geothermal systems in Nevada, particularly those associated with deep circulation in fault-controlled basins. Water-rock interactions and residence time contribute to the "maturity" of geothermal fluids. Mature waters typically have higher TDS and approach chemical equilibrium with the reservoir rocks, as reflected in their saturation indices for key minerals.

Nevada's geothermal waters vary in composition across the state and can be generally classified into several types based on their chemical composition. The most common types include sodium-bicarbonate, calcium-bicarbonate, sodium-chloride, and mixed types [41]. Fig. 4 shows a Piper diagram [42] analysis of geothermal wells in Nevada. The clustering of samples in the Na-K-Cl corner of the diagram indicates the presence of sodium-chloride type waters. These waters are typically associated with deep, high-temperature geothermal systems and extensive water-rock interaction. The diagram also shows a spread towards the HCO_3 region, representing sodium-bicarbonate waters often found

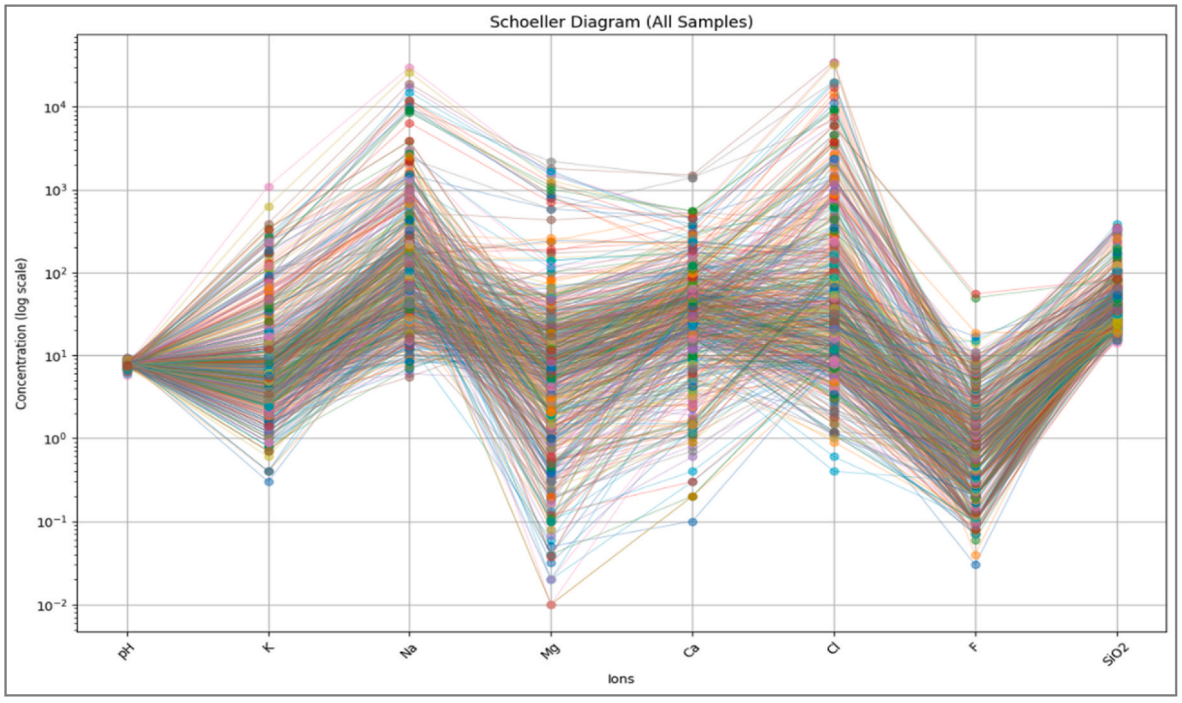


Fig. 3. Schoeller diagram for selective Nevada water samples.

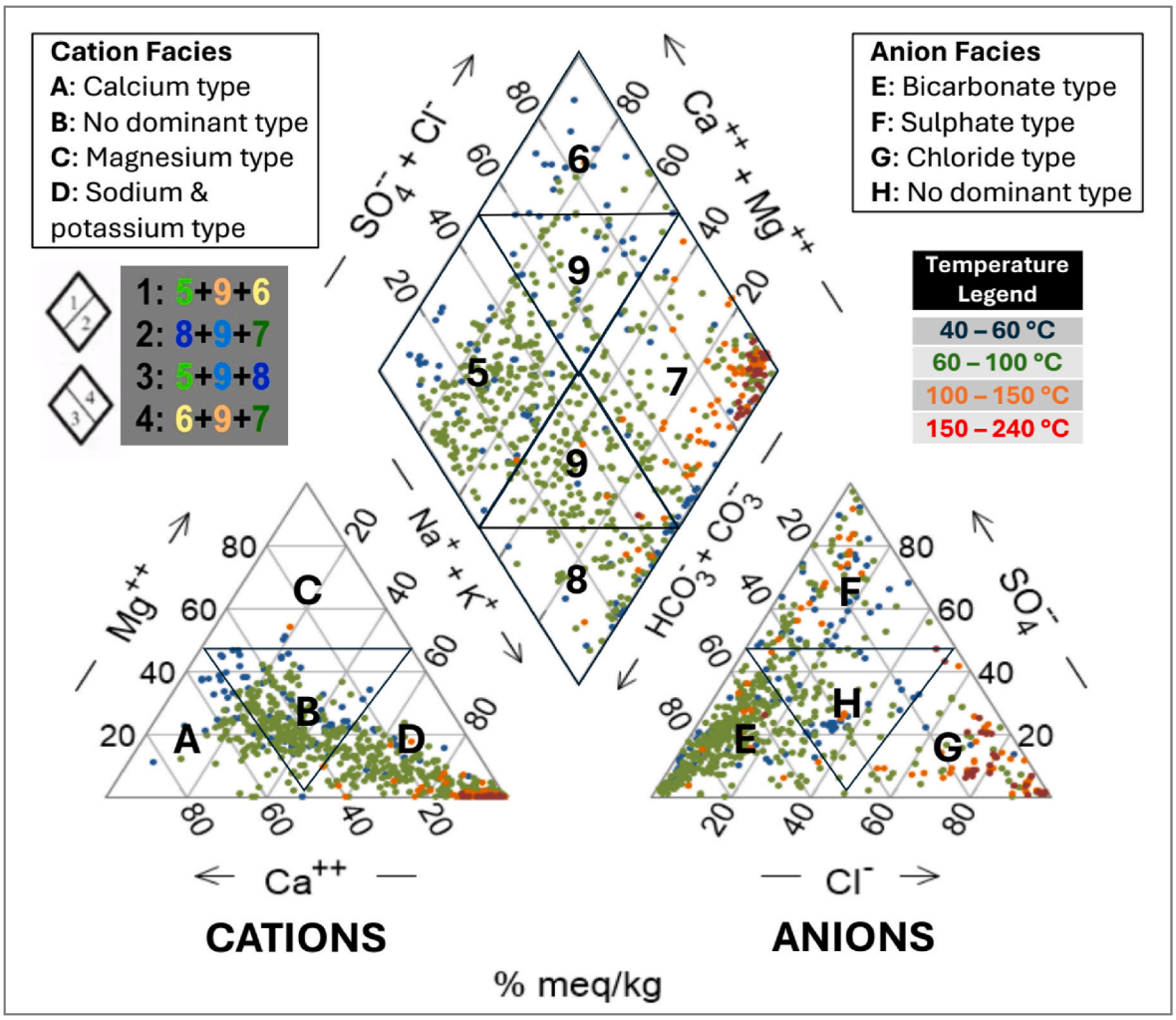


Fig. 4. Piper trilinear diagram showing the hydrochemical characteristics and hydro-chemical facies of Nevada groundwater samples. The Piper diagram can be separated in hydrochemical facies. Legend: 1: Alkaline earths exceed alkalies; 2: Alkalies exceed alkaline earths; 3: Weak acids exceed strong acids; 4: Strong acids exceed weak acids; 5: Magnesium bicarbonate type; 6: Calcium chloride type; 7: Sodium chloride type; 8: Sodium bicarbonate type; 9: Mixed type.

in lower temperature systems or those with significant meteoric input. This classification provides a general overview of the most common water types found in the state's geothermal resources, but water types may vary depending on the geological setting and depth of individual geothermal systems. Isotopic analysis, particularly of Deuterium (δD) and Oxygen-18 ($\delta^{18}O$), provides insight into the fluid's origin and mixing processes. Most Nevada geothermal waters are of meteoric origin, with some systems showing oxygen isotope shifts due to high-temperature water-rock interactions [36].

Temperature plays a critical role in determining the mineralogical assemblages associated with geothermal waters. At higher temperatures (>200 °C), typical of many geothermal systems in Nevada, quartz is the dominant silica phase and controls silica solubility. In contrast, lower temperature systems may be saturated with chalcedony or amorphous silica [37]. The precipitation of secondary minerals such as calcite, silica polymorphs and clay minerals are temperature dependent and influences the permeability and fluid chemistry of geothermal reservoirs. For example, the formation of travertine and siliceous sinter at the surface is indicative of specific temperature ranges and fluid compositions [43].

The geochemical characteristics of Nevada's geothermal waters provide valuable tools for exploration and resource evaluation. In addition, the presence of critical elements such as lithium in geothermal fluids offers the potential for mineral extraction as a by-product of geothermal energy production.

3. Target variable inference

The primary objective of the ML Geothermometer is to predict subsurface reservoir temperature. However, the geochemical database used to train the model only provided sample temperature measurements, not recorded subsurface temperatures, as confirmed by the data owners (University of Nevada, Reno). To overcome this limitation, additional methods including classical geothermometers, multicomponent geothermometry, regional temperature records, and existing ML geothermometers were implemented to infer reservoir temperatures while minimizing the uncertainty of the predicted subsurface temperature to enable effective model training.

3.1. *Classical geothermometers*

Classical chemical geothermometers have evolved since their introduction in the 1960s [44]. The basic hypothesis of geothermometry is the chemical equilibrium of the reservoir fluid and the host rock [45]. Secondary processes may alter the fluid composition and hence the equilibrium as it migrates to the earth's surface. The choice of geothermometer is also critical, with different types being more suitable for specific temperature ranges and mineral assemblages.

To better understand the mineral assemblages and to aid in the selection of the correct geothermometer, the PHREEQC software [46] was used to determine the saturation indices of the minerals in all water samples using the Lawrence Livermore National Laboratory (LLNL) thermodynamic database. The LLNL contains reliable data for a large number of minerals and aqueous species in the temperature range 0–300 °C, which corresponds to the temperature ranges of the Nevada geothermal fields [47]. The saturation index (SI) of a mineral indicates the saturation condition of a solution with respect to the particular mineral phase. It is defined as the logarithm of the ratio between the activity product, Q, and the dissolution constant at temperature T, K_T. SI = log [Q/K_T]. SI = 0 for a given mineral indicates full equilibrium in solution. A positive SI indicates oversaturation of the mineral in the solution and possible precipitation. A negative SI indicates undersaturation of the mineral in the solution and possible dissolution.

In general, the most commonly dissolved minerals in Nevada water samples are silica minerals. In order of solubility, they are chalcedony, cristobalite(α), tridymite, and quartz. Fig. 5 shows the saturation indices for these four mineral species in the water samples. The groundwater was almost saturated with tridymite (average SI = - 0.038), slightly unsaturated with chalcedony (average SI = - 0.094), unsaturated with cristobalite(α) (average SI = - 0.309) and slightly oversaturated with quartz (average SI = 0.136). The silica content of the groundwaters investigated varied widely from 14.0 to 389.0 ppm. Groundwater samples with higher SiO_2 concentrations indicate higher circulation temperatures within an aquifer.

A variety of silica geothermometer equations have been developed to estimate subsurface temperatures based on the equilibrium of quartz or chalcedony in reservoir water [48]. Silica minerals with different crystal structures have different solubilities and construction temperatures, which affects their use in geothermometry. Chalcedony, a less ordered form of silica, is more soluble than quartz and crystallizes at temperatures below 180 °C. This property difference makes chalcedony-based geothermometers ideal for lower temperature settings, while quartz geothermometers are better suited to high temperature environments above 150 °C. Below this temperature, chalcedony rather than quartz becomes the dominant control of dissolved silica concentrations [49]. The accuracy of these geothermometers is influenced by the silica concentration in the water, which reflects the thermal history of the reservoir.

Considering the equilibrium with these minerals, two geothermometer equations were used in this analysis:

(Eq. (1)) Chalcedony geothermometer which can be expressed as [50].

$$t = 1112 / (4.91 - \log SiO_2) - 273.15 \quad (1)$$

(Eq. (2)) Quartz geothermometer which can be expressed as [51].

$$t = 1309 / (5.19 - \log SiO_2) - 273.15 \quad (2)$$

where t (°C) is the equilibration temperature at depth, and SiO_2− is the silica concentration, mg/L.

The output temperature values from both geothermometers were compared with the additional methods to derive the required target variable and it was found that the mean of both geothermometers (T_{mean}) accurately represents the reservoir temperature. T_{mean} is then used to compare the different temperature estimation methods.

3.2. Multicomponent geothermometry

Classical geothermometers, while effective for high temperature geothermal systems, often prove unreliable in medium to low enthalpy systems due to the inability to achieve full fluid-mineral equilibration [2]. Multicomponent geothermometry provides a more comprehensive approach to estimating subsurface reservoir temperatures by evaluating the saturation indices of multiple reservoir minerals over a temperature range using thermodynamic databases. This method can consider various subsurface processes such as dilution and degassing. It determines the reservoir temperature by identifying where the saturation indices (log [Q/KT]) for multiple minerals simultaneously reach zero, indicating equilibrium conditions [4]. Unlike classical geothermometers, this approach doesn't rely on specific mineral solubilities, making it applicable to different geological settings, provided the reservoir mineral types and their detailed chemical compositions are known. The main challenge in this method is the selection of the mineral assemblage to be used for the calculations.

To cross-validate the 'classical' geothermometers and to better constrain the target variable, multicomponent geothermometry was performed on all water samples using the iGeoT program and the thermodynamic database (tk-slt.h06_jun16) to calculate both the ion activity product (Q) and the thermodynamic equilibrium constant (K_T) of various minerals over a range of temperatures (e.g. 40–300 °C). iGeoT has the ability to optimize some of the unknown and/or erroneous parameters, including the effect of mixing/dilution. Optionally, iGeoT can reconstruct the subsurface reservoir temperature by varying critical parameters such as aluminum concentration and steam loss to minimize the uncertainty in the temperature estimate. The detailed procedure of this method is described by Spycher et al. [5]. For each water sample, the saturation indices were plotted as a function of temperature, and the clustering of log(Q/K) curves near zero at a given temperature (for specific reservoir minerals assemblage) was derived to obtain the reservoir temperature (T_{iGeoT}). The temperature values were plotted

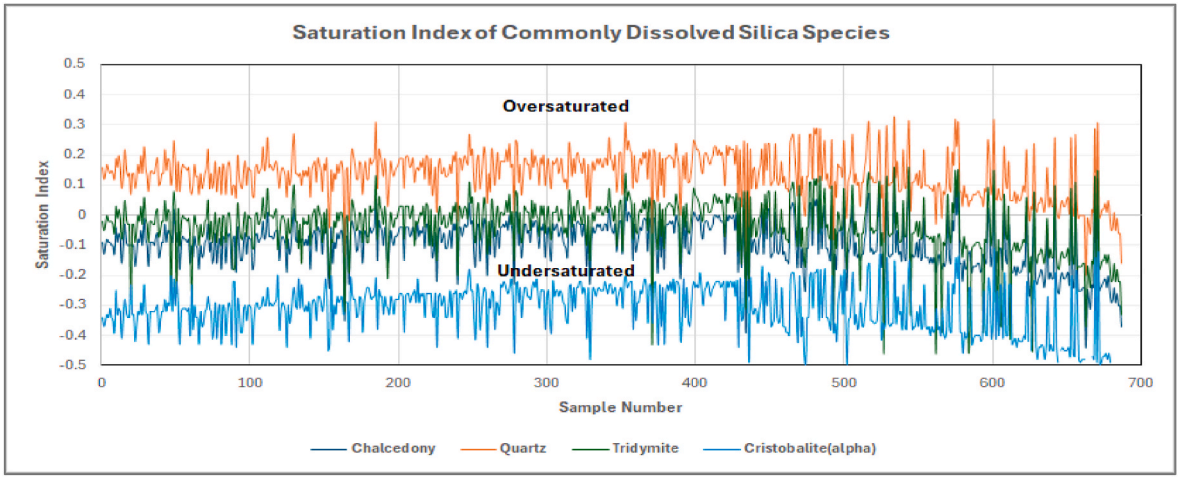


Fig. 5. The saturation indices for the four mineral species in the water samples.

against the mean SiO_2 geothermometer temperature (T_{mean}) as shown in Fig. 7.

3.3. Regional thermal database

The Southern Methodist University (SMU) Geothermal Lab developed one of the 'nodes' of the National Geothermal Data System, collecting, cataloguing and making available to end users a wide range of geothermal and oil and gas information [52]. The SMU database included bottom-hole temperature (BHT) measurements, heat flow data, thermal conductivity, log depth, geothermal gradient and other spatial information. The SMU database had 2365 data entries covering the state of Nevada. Most BHTs are measured from sedimentary rocks drilled by the oil and gas industry to depths of 1.2–3 km, depending on the depth to the reservoir. In areas of geothermal energy production, wells are typically drilled to depths of 1.0–3.0 km in the western United States.

An ensemble weighted average of four models, Linear Regression, Random Forest (RF), Xtreme Gradient Boosting (XGB) and Kriging, was developed to impute the missing temperature values in the Nevada dataset (T_{SMU}) based on the available BHT data at nearby locations. The model included multiple feature engineering, including depth-coordinate interaction and gradient-depth interaction, as shown in Fig. 6, to improve its imputation accuracy. It was then evaluated using the R^2 and RMSE evaluation matrices and the predicted temperature values (T_{SMU}) were plotted against (T_{mean}) as shown in Fig. 7. The low (T_{SMU}) values are consistent with the (T_{mean}) values, whereas the high (T_{SMU}) values are inconsistent with the calculated SiO_2 geothermometer and the multicomponent geothermometry approach.

3.4. Existing data-driven model

The ANN solute geothermometer (AnnRG) developed by Ystroem et al. [16] used geochemical data (mainly K^+, Na^+, Mg^{2+}, Ca^{2+}, Cl^-, SiO_2 and the power of hydrogen 'pH') to predict subsurface temperatures in geothermal systems. The AnnRG model was trained on a total of 155 data pairs and validated on 45 water samples from four different geothermal fields worldwide. Given the availability of the seven features of the AnnRG model, the model was run on all the Nevada geochemical datasets and the output temperature was plotted against T_{mean}, T_{iGeoT} and T_{SMU} values, as shown in Fig. 7. The output temperature values (T_{AnnRG}) agreed with the calculated classical and multicomponent geothermometers, except for some samples with high TDS values (>20,000 ppm), possibly due to the model training range.

All four temperature inference methods were combined to provide a more accurate representation of subsurface temperatures for training the ML model. The temperature difference between all four inference methods was examined and 23 samples were rejected from the training dataset due to large variance in temperature values.

4. Exploratory data analysis

The GBGG database contained several features that could be used to train the data-driven geothermometer, including multiple ion concentrations, trace element concentrations, fluid system parameters, and spatial data of the water samples. The selection of the model training features followed a systematic approach including domain knowledge, availability of the selected feature throughout the dataset to avoid data imputation, and inclusion of the selected feature in standard water geochemical analyses to extend the applicability of the model to different regions. This resulted in the extraction of eight key features including the major cations (K^+, Na^+, Mg^{2+}, Ca^{2+}), major anions (Cl^-, F^-) and SiO_2, and pH.

4.1. Data pre-processing

The GBGG database included 14,369 water samples from Nevada, covering springs (3544 samples) and drilled wells data (10,825 samples) in the 17 counties of the state. Only wells data were considered in this study, and all missing chemical constituents and duplicate records were eliminated from the drilled wells data, resulting in a refined total of 937 samples with complete feature sets.

In order to validate the chemical equilibrium of the 937 water samples and to ensure that the aqueous solutions were electrically neutral, the charge balance equilibrium (CBE) of the water samples was calculated using the PHREEQC software according to "Eq. (3)" below [53], prior to statistical processing of the data:

$$E = \frac{\sum Zm_C - \sum Zm_A}{\sum Zm_C + \sum Zm_A} \cdot 100\% \quad (3)$$

where E is the percent error in the change balance, z is the charge on the ionic species and m the molality of cationic (C) and anionic (A) species.

Ideally, the sum of all positive charges (cations) should be equal to the sum of all negative charges (anions). CBE is used to assess the validity and quality of water analyses. Its value can be positive or negative, positive indicating a higher concentration of cations than anions and vice versa. Acceptable water analyses have a CBE between ±5 %. Possible causes of electrical imbalance include laboratory error during analysis, unmeasured dissolved species and the use of unfiltered samples containing particulate matter that dissolves when acid is added. The CBE calculations showed that 263 samples were outside the acceptable

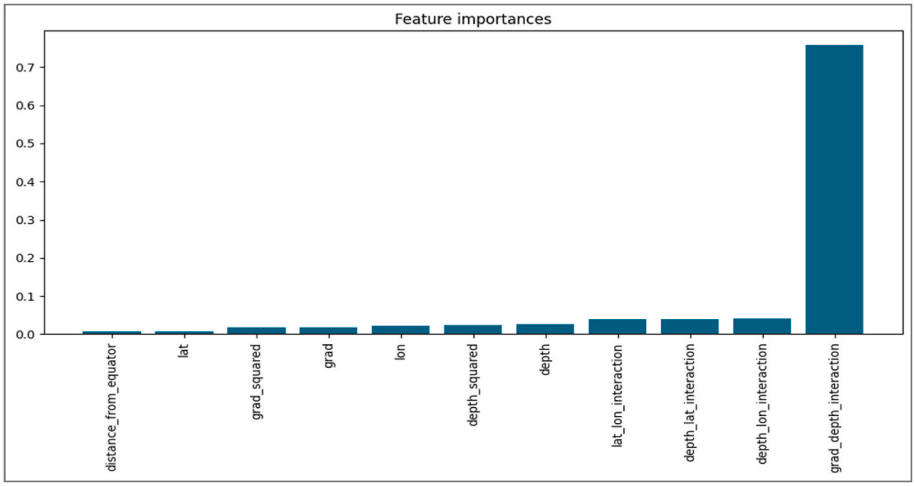


Fig. 6. Feature engineering variables for the temperature imputation algorithm.

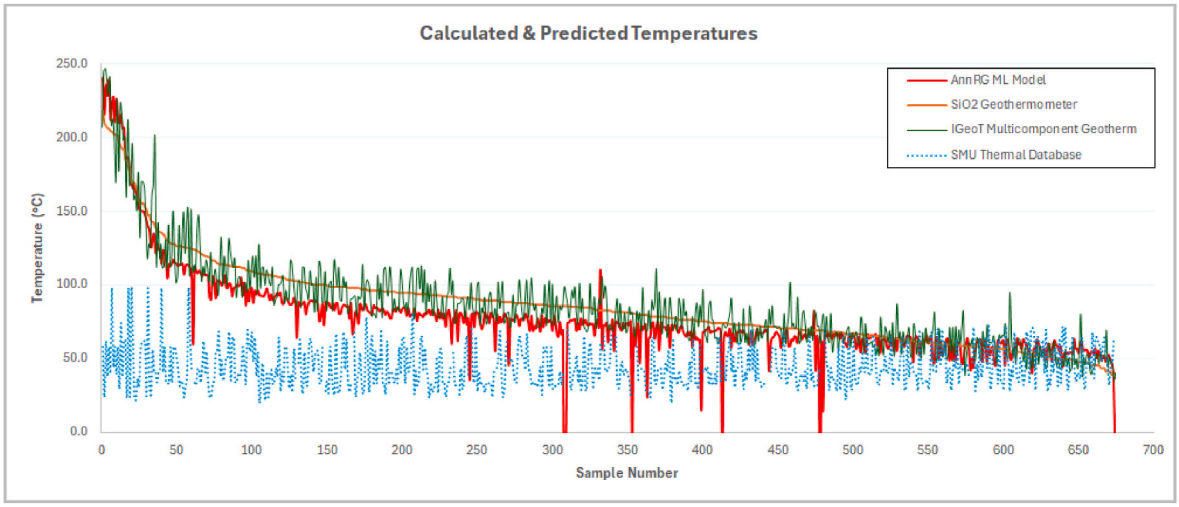


Fig. 7. Calculated and predicted temperature values for all samples.

charge balance error range, so these samples were excluded, leaving a final dataset of 674 samples with acceptable chemical compositions.

4.2. Data distribution

Statistical analysis of the target variable, temperature, shows a range from 37 °C to 218 °C, with a mean of approximately 86.1 °C and a median of 82.6 °C, indicating a slightly right-skewed distribution. Table 2 summarizes the minimum, maximum, mean, standard deviation, median and interquartile range of the chemical constituents. Notably, potassium has a wide range from 0.30 to 1100.00 ppm, while sodium has an even wider range from 5.50 to 30,000.00 ppm. The distribution of these characteristics suggests significant variability in the chemical composition of the groundwater across the study area, which may influence the geothermal potential and hydrological dynamics. Understanding these statistical parameters is essential for interpreting relationships between temperature and chemical constituents, as well as for modelling geothermal systems within the Great Basin region.

The distribution of the target variable, along with all selected features, was visualized through graphical representations, including histogram and box plots. The distribution plots (Fig. 8) illustrate the frequency and distribution of each feature and the target variable, making it easier to identify the shape of the data, deviations from normality, and potential outliers. With the exception of pH, the data shows a high degree of non-linearity and skewness due to the large geographic area of Nevada and the diverse geologic settings of the water samples. The box plots (Fig. 9) provide a more structured summary of the distribution for each feature and the target variable. The box represents the interquartile range (IQR), the median is indicated by a line within the box, and the whiskers extend to 1.5 times the IQR [54]. Points beyond the whiskers are identified as potential outliers, providing insight into the variability and extremes within the dataset.

Table 2
Chemical constituents' data statistics.

	Mean	Std	Min	Max	25 %	75 %
pH	7.80	0.50	5.70	9.30	7.50	8.00
K	22.98	66.61	0.30	1100.00	3.43	12.00
Na	544.54	2269.21	5.50	30000.00	31.00	210.00
Mg	45.47	199.42	0.01	2200.00	2.30	19.00
Ca	63.24	119.82	0.10	1500.00	17.00	64.00
Cl	642.03	2953.29	0.40	35000.00	10.00	118.50
F	1.58	3.57	0.03	55.30	0.20	1.58
SiO_2	57.42	52.26	14.00	389.00	31.00	62.00
T_{mean}	86.14	30.39	36.40	218.20	66.40	97.80

Std: Standard Deviation

While these graphical methods highlight data points that are statistically outliers from the majority, they do not inherently distinguish between true errors or anomalies and natural variations due to different underlying conditions. In the context of geothermal systems, geochemical concentrations and subsurface temperatures can vary significantly based on geographic locations and specific geological settings, including host rock types, fault systems, and hydrothermal processes. Therefore, potential outliers identified in the overall dataset may simply represent data points from geothermal systems with different characteristics due to their location or geological setting. In this case, clustering analysis is helpful in confirming data anomalies because it is used to group data points based on their similarity across multiple features.

4.3. Data clustering

The geochemical samples analyzed in this study reflect the evolution of water during its subsurface transit, which can vary in duration from short-term surface flows to long-term deep fluid circulation. The geological setting of each source is influenced by several factors, including the mineralogical composition of the surrounding rocks, the specific dissolution rates of the minerals, and the physicochemical parameters of the water as it interacts with these minerals. In addition, the mixing of water masses that have followed different underground paths is a crucial factor affecting the composition of spring water. External influences such as climatic conditions also play an important role. To uncover meaningful trends within this multivariate system, we explored various clustering techniques - K-means, hierarchical clustering, DBSCAN and Gaussian Mixture Models (GMM) - to identify underlying structures in the dataset (Fig. 10). Of these methods, hierarchical clustering using the Ward criterion proved most effective in organizing the data into meaningful clusters. This approach was applied to all water samples across a range of parameters including concentrations of major anions and cations, pH, ionic strength, latitude and T_{mean} temperature.

Hierarchical clustering revealed eight distinct clusters characterized by similar water types, maturity levels and temperature ranges, indicating that the previously identified anomalies in the data were not outliers, but rather the result of a different environment. Each cluster was carefully examined to identify potential outliers and to gain insight into the geological and hydrogeochemical parameters influencing its composition. Although a detailed analysis of these clusters is beyond the scope of this article and will be addressed in future publications, it is noteworthy that Ward's method - an objective function-based criterion for hierarchical cluster analysis - was instrumental in this study. This method minimizes variance within clusters by merging pairs based on

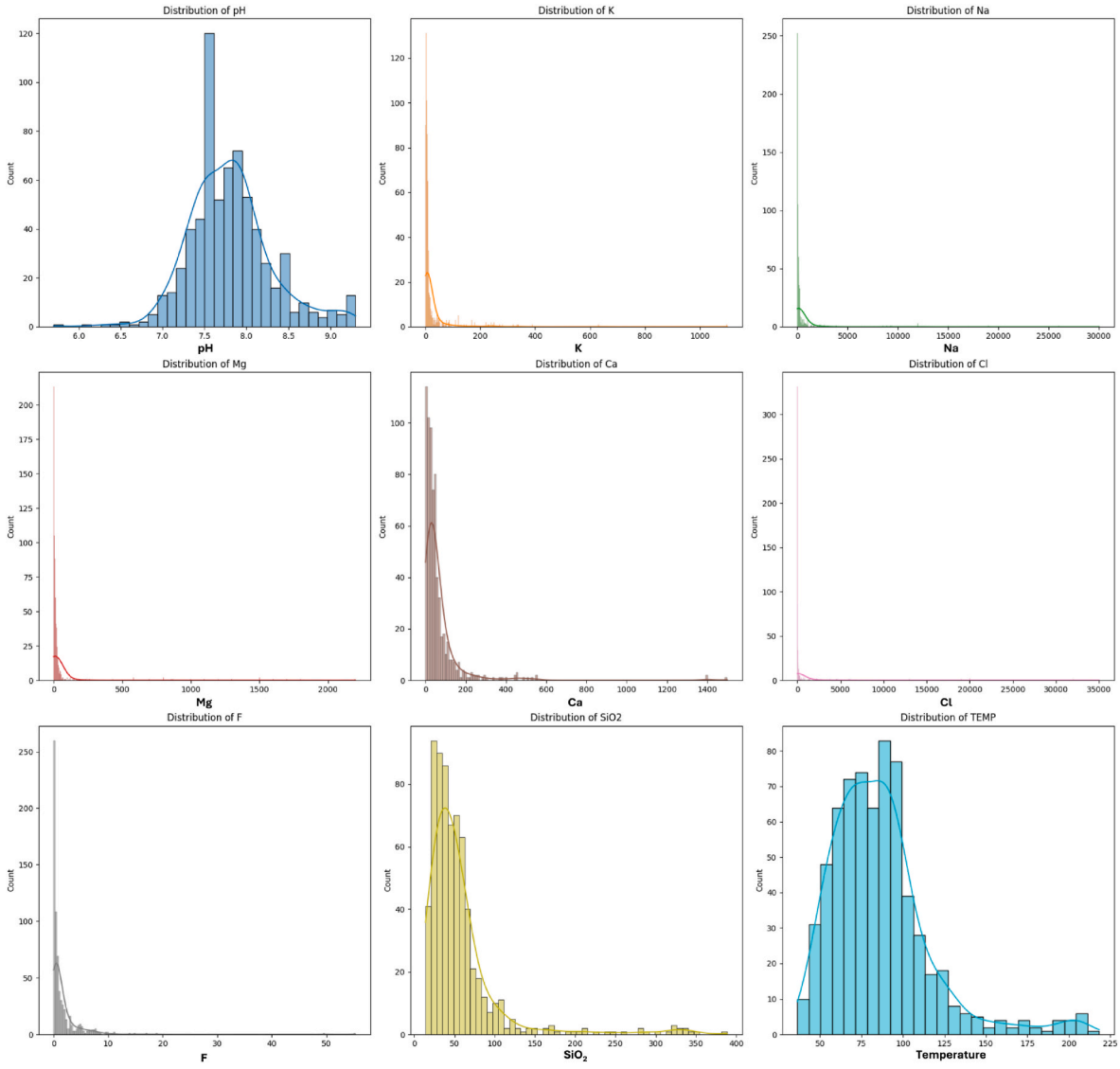


Fig. 8. Distribution plots of each feature and the target variable.

optimal values of an objective function, which can be tailored to reflect specific investigative goals. The results of different clustering techniques were evaluated using Silhouette [55] and Calinski-Harabasz [56] scores to assess their effectiveness, with higher scores indicating better defined clusters. These metrics provide critical insights into the significance of the identified clusters and guide further exploration of the complex interactions that govern water chemistry in this region.

4.4. *Features correlation*

A correlation heatmap was generated to illustrate the relationship between the selected features and temperature, providing a comprehensive overview of their interactions prior to data normalization. The Pearson heatmap (Fig. 11) shows that silica has a strong positive correlation (0.94) with the target variable, indicating its importance as a predictor in the dataset. In addition, fluorine and potassium show moderate positive correlations with temperature. In contrast, pH, magnesium and calcium show weak negative correlations. Scatter plots (Fig. 12) further illustrate the relationships between the top three most correlated features and temperature, providing a visual representation of the strength and direction of these relations.

A multicollinearity check was also performed between the features and revealed important insights. A high correlation of 0.96 between Cl and Na indicates a strong relationship, suggesting redundancy in the information they provide. Variance Inflation Factor (VIF) results show that Na, Cl, Ca and Mg all have VIF values greater than 10, indicating very high multicollinearity, while K has a VIF between 5 and 10, also indicating high multicollinearity. In addition, pH and SiO_2 show moderate multicollinearity with VIF values between 1 and 5. The implications of these findings suggest that the highly intercorrelated features, particularly Na and Cl, may warrant consideration for exclusion from modelling or the need to create composite features to mitigate redundancy, or the use of tree-based models which are less affected by multicollinearity. Furthermore, while pH does not show strong correlations with any single feature, its moderate negative correlations with several features suggest that it could provide unique information in the model. The hierarchical clustering dendrogram (Fig. 13) illustrates how features cluster based on similarity, further highlighting the relationships between variables and guiding future feature selection strategies. Variables that cluster early (lower in the tree) are more similar, while the height of the branches indicates the distance between clusters.

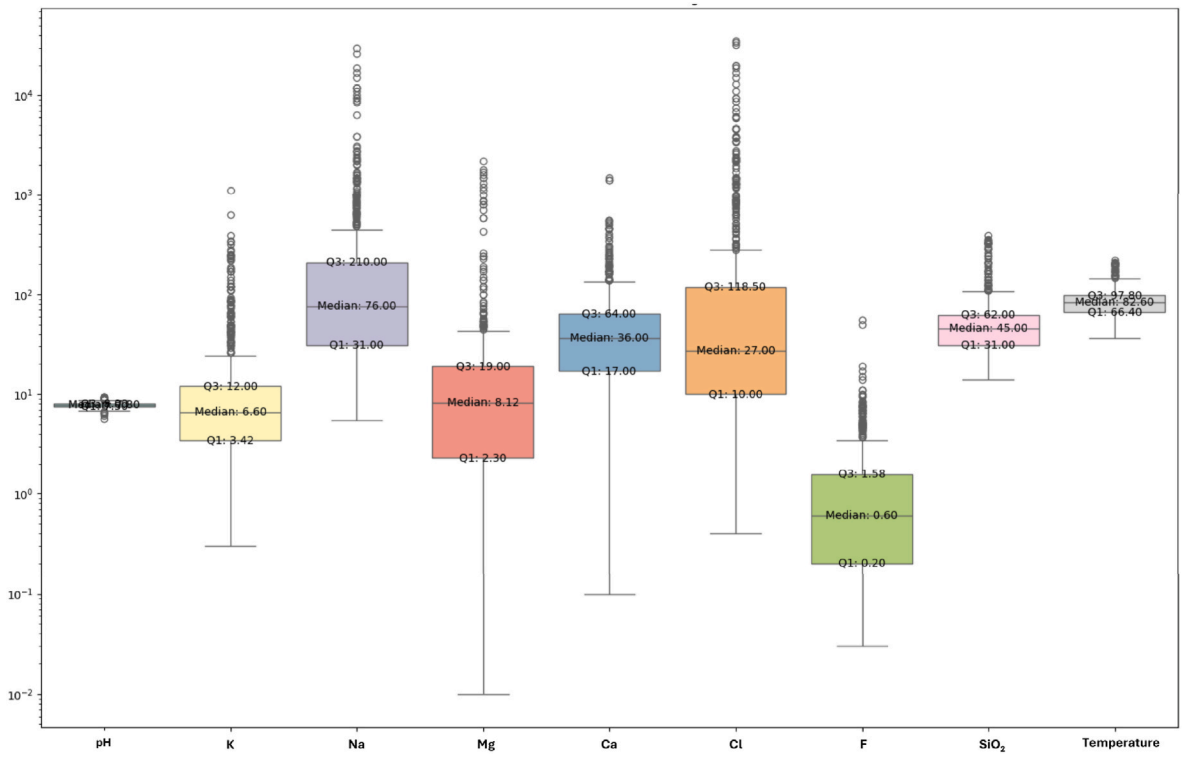


Fig. 9. Box plots in log scale for each feature and the target variable.

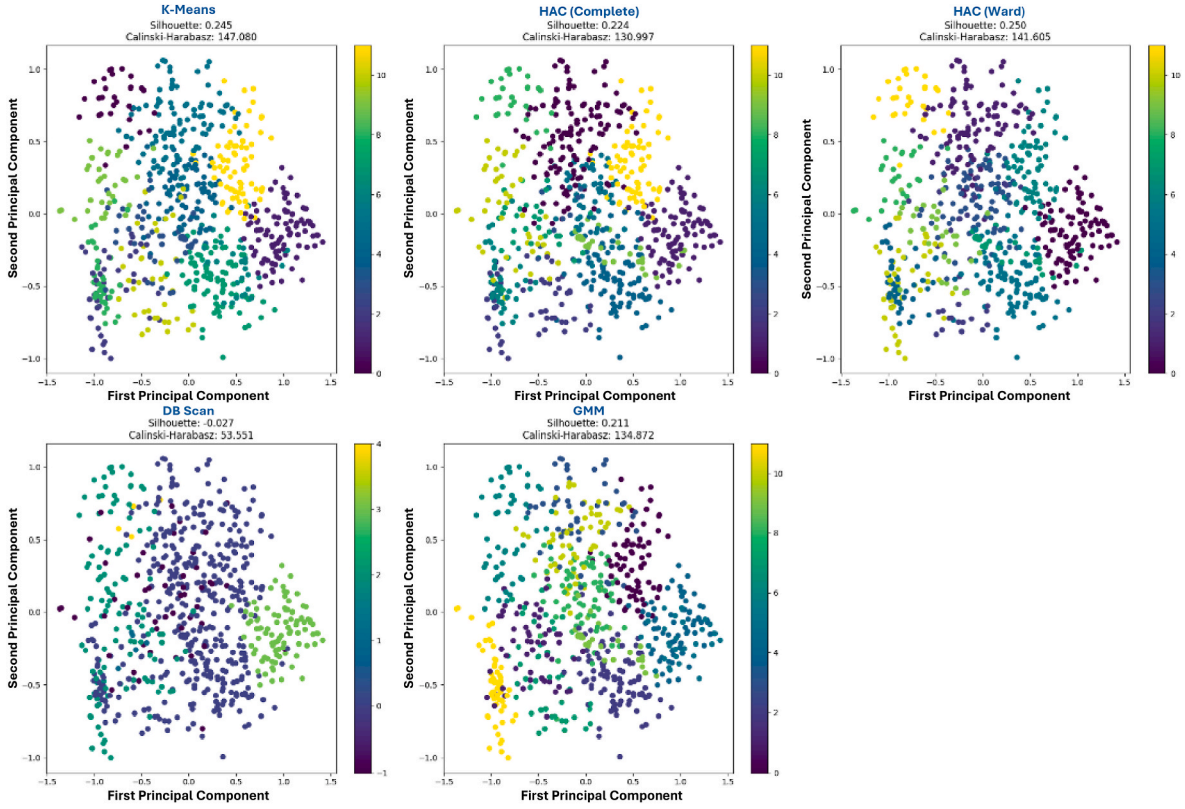


Fig. 10. Various data clustering techniques and their corresponding Silhouette and Calinski-Harabasz Scores.

5. Model development

In this study, we define a supervised learning problem aimed at predicting subsurface temperature using geochemical data. A comprehensive summary of ML models used to estimate subsurface temperature is provided, including various transformation methods such as Z-score

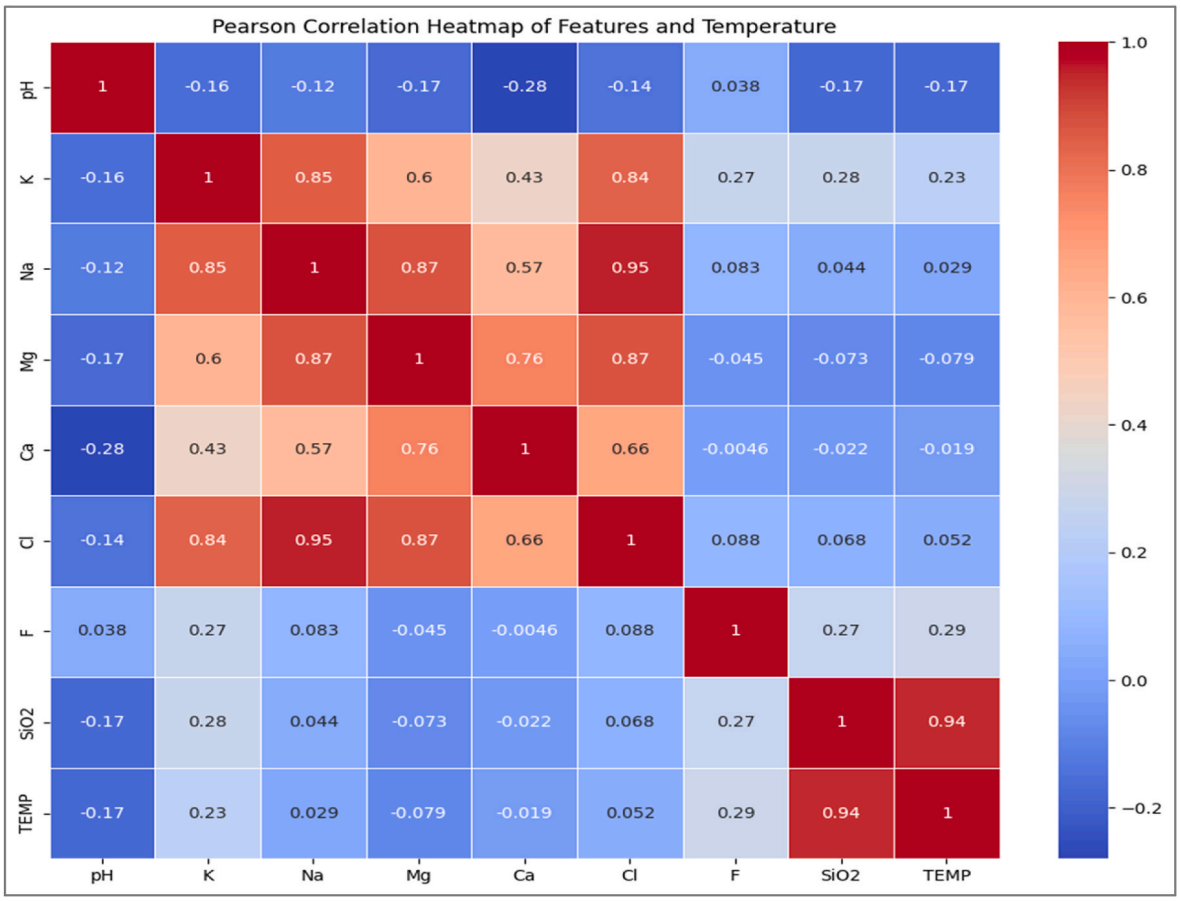


Fig. 11. Pearson correlation heatmap of features and temperature.

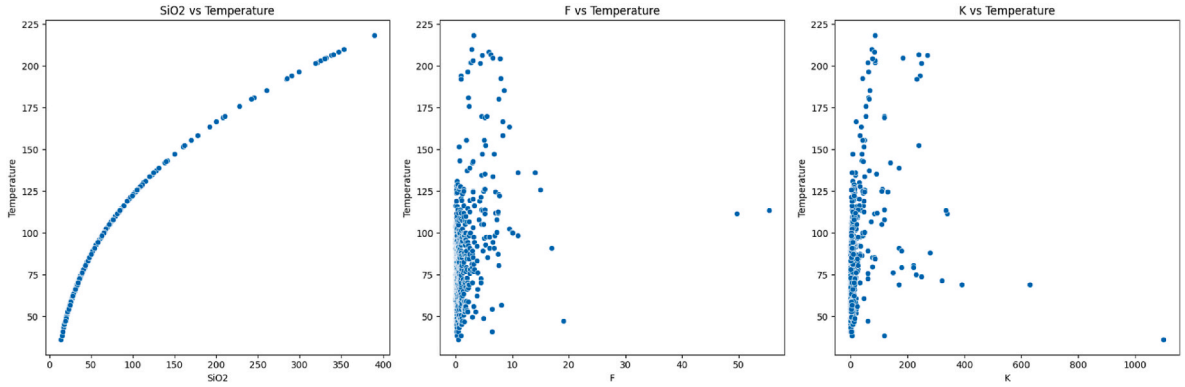


Fig. 12. Scatter plots of relationship between the top 3 most correlated features (SiO_2, F and K) and temperature.

normalization, logarithmic transformation, quantile normalization, Box-Cox transformation and quantile-uniform transformation. Each normalization technique was evaluated for its effect on model performance, with particular attention paid to skewness and kurtosis values to assess the distributional characteristics of the transformed data.

5.1. Baseline models

The dataset was divided into 80 % training and 20 % test sets and different regression models - including random forest, gradient boosting, back-propagation neural networks (BPNN) and deep neural networks - were tested with the above transformations. The results showed that normalization techniques had an impact on prediction accuracy, as reflected in the scores of the applied models, as shown in Fig. 14. In this study, the Z-score transformation was used to improve the input data by spreading frequent values, thereby reducing the influence of marginal outliers and improving model performance [57].

5.1.1. Random forest

RF is an ensemble ML technique that combines multiple decision trees to improve prediction accuracy and mitigate overfitting, making it particularly effective for regression tasks [58]. In this study, we used the RF regressor because of its ability to handle complex datasets with inherent variability, as demonstrated by its successful application in the geosciences [59]. The model uses an ensemble of 100 decision trees, each with a maximum depth of 10. It uses bootstrap sampling and

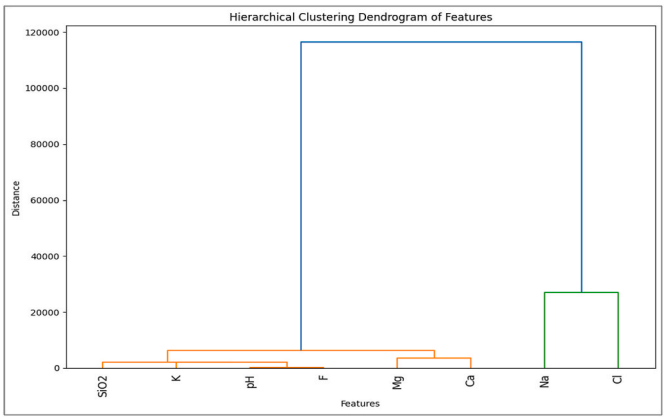


Fig. 13. Hierarchical clustering dendrogram of all features.

considers the square root of the total number of features when making splits. The minimum number of samples required to split an internal node is 5, and the minimum number of samples required to be at a leaf node is 2. The strength of RF lies in its ability to aggregate predictions from multiple trees, thereby improving accuracy and reducing variance [60]. Furthermore, we chose an ensemble-based algorithm because it is robust to imbalances in the target and is good at capturing nonlinear relationships in the data [61].

5.1.2. Gradient boosting

Gradient boosting is a powerful supervised learning technique that constructs predictive models through an iterative process of combining multiple weak learners, typically decision trees, to form a robust ensemble model [62]. Gradient boosting is similar to RFs, but instead of building all the trees simultaneously and averaging them, gradient boosting builds new trees sequentially to reduce the residuals from all the previous trees. This allows the prediction model to be gradually 'boosted' as more trees are built [63].

In this study, we used the XGB library, a scalable and efficient implementation of gradient boosting developed by Chen and Guestrin [64]. XGB builds decision trees sequentially, where each new tree is trained to correct the errors made by the previous trees, thereby improving the overall accuracy of the model. This approach allows the predictions to be fine-tuned as more trees are added, making it particularly effective for complex datasets with non-linear relationships [65]. This gradient boosting model uses 100 trees with a maximum depth of 3. It has a learning rate of 0.01, which helps to reduce overfitting by making the model more conservative. The subsample and colsample_bytree parameters are both set to 0.8, which introduces randomness and further prevents overfitting. The min_child_weight of 3 helps to control the complexity of the model.

Initially, the XGB model showed clear signs of overfitting, as evidenced by a narrow distribution of training errors but wider test errors, overwhelming importance of the SiO_2 feature, and the highly variable cross-validation mean square error (MSE) across different fold sizes. To address this and improve generalization, we made several improvements, including increasing regularization by adjusting parameters such as max_depth, min_child_weight, gamma, and L1/L2 regularization; reducing model complexity through fewer estimators and early stopping; and exploring and possibly transforming the dominant SiO_2 feature. The code developed aimed to constrain the model and achieve a balance between bias and variance for improved generalization across both training and test datasets.

5.1.3. Artificial neural network

ANNs are a leading ML approach that mimics the functionality of the human brain. They are structured with input, output and intermediate hidden layers. These layers are connected by nodes, each of which performs a specific task through an activation function that introduces non-linear properties to the ANN. The strength of the flow of information between nodes is determined by the weight assigned to their connections [66]. The flow of information in ANNs occurs in two main phases: forward and backward propagation. In forward propagation, data flows between neurons from the input layer through hidden layers to the output layer, generating predictions. This is followed by backward propagation, where the network adjusts its parameters based on the prediction errors in an attempt to minimize them [67]. The training process of an ANN is an iterative cycle of these two phases. It continuously refines the model's parameters through repeated forward and backward passes, gradually improving the network's predictive accuracy. This cycle continues until the network achieves optimal predictive performance, balancing accuracy with the risk of overfitting [68].

In this study we used a simple BPNN with four layers: an input layer, two hidden layers of 64 and 32 neurons respectively, and an output layer. It uses rectified linear unit (ReLU) activation functions and is trained using the Adam optimizer with a learning rate of 0.001. The model is constructed using MSE as the loss function. This architecture is chosen for its simplicity and ability to capture non-linear relationships in the data. While it doesn't perform as well as the tree-based models, it still provides reasonable predictions and serves as a baseline for neural network approaches.

5.1.4. Deep neural network

DNNs are an advanced subset of ANNs that use multiple layers of interconnected neurons to model complex relationships between input and output variables [69]. The large feature dimension in our study

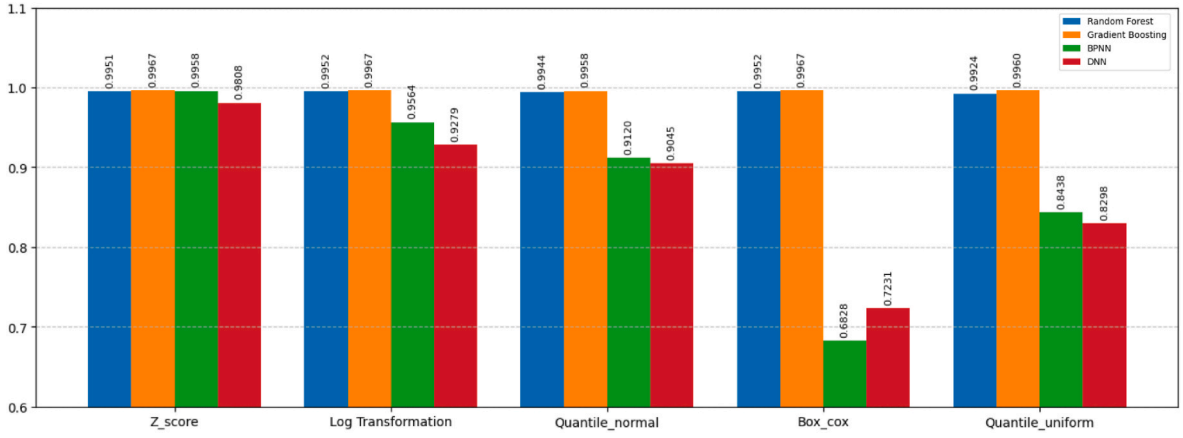


Fig. 14. Various data normalization techniques and their corresponding R^2 Scores (test set).

necessitated the exploration of DNNs to effectively capture the nonlinear relationships inherent in geochemical properties and subsurface temperatures.

We explored different DNN architectures to find a balance between complexity and computational efficiency. In this study, the DNN has a more complex architecture than the BPNN and is designed with a focus on performance and generalization. It consists of three hidden layers (64, 32 and 16 neurons) using ReLU activation functions, followed by a single output neuron. The model uses Batch Normalization after each hidden layer to stabilize learning and improve generalization. The Adam optimizer with a learning rate of 0.001 is chosen for adaptive learning. The model uses MSE as the loss function, which is suitable for regression tasks. Training includes early stopping (patience of 20 epochs) to avoid overfitting and optimize convergence. The combination of these techniques - regularization, normalization, adaptive learning rates and early stopping - work together to create a model that can learn complex patterns in the data while maintaining good generalization to unseen examples. The shape of the loss function is visualized as learning curves in Fig. 15.

In Fig. 15, the learning curves of the DNN model are shown over 88 epochs of training. The blue line represents the training loss, while the orange line represents the validation loss. Initially, the validation loss starts higher than the training loss, indicating the poor generalization of the model in the beginning. As training progresses, both losses decrease rapidly, with the validation loss decreasing more steeply in the early epochs. Around epoch 20, the curves intersect and continue to decrease together, indicating that the model is learning effectively without overfitting. The curves flattened out after about 40 epochs, indicating that the model has reached a stable level of performance. The close alignment of training and validation losses in later epochs indicates good generalization, confirming the ability of the DNN to perform well on unseen data.

5.2. Model evaluation & results

Three key metrics were used to assess the model's performance: R^2, RMSE, and MAE. The R^2 value quantifies how well the predicted values match the actual observations. A higher R^2 value, closer to 1, indicates a better fit of the model to the data and better prediction performance. The RMSE provides an estimate of how much the predictions differ from the actual values. This metric is particularly useful for assessing model accuracy during cross-validation, although it is not suitable for direct comparison between different datasets due to its dependence on response size [70].

The MAE is a widely used metric in regression analysis that quantifies the average size of errors between predicted and actual values, regardless of their direction. The importance of the MAE lies in its ability to provide a straightforward interpretation of prediction accuracy, as it expresses the average error in the same units as the target variable.

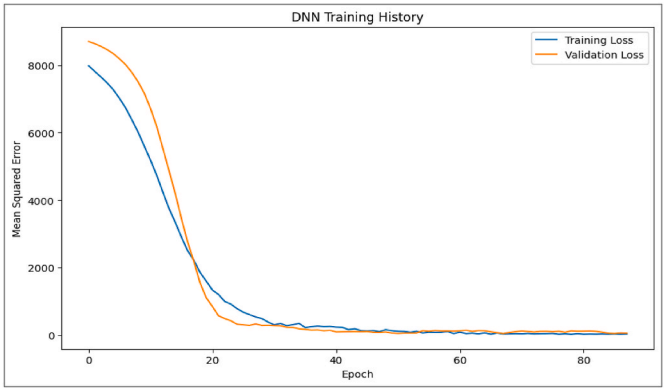


Fig. 15. Learning curves of the training and validation of the DNN.

Unlike RMSE, which squares the errors and can disproportionately affect the results due to larger errors, MAE gives equal weight to all errors, making it particularly useful for understanding model performance in a more intuitive way [71]. The evaluation metrics can be expressed in "Eq. (4)":

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|T_{actual} - T_{predicted}\right|$$

$$RMSE = \sqrt{\sum_{i=1}^{N}\frac{T_{actual} - T_{predicted}}{N}}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(\frac{T_{actual} - T_{predicted}}{N}\right)^2}{\sum_{i=1}^{N}\left(T_{actual} - \widehat{T}\right)^2} \quad (4)$$

where; T_{actual} = actual (measured) temperature values, $T_{predicted}$ = predicted temperature values, $\widehat{T}$ = mean temperature values, and N = number of observations.

The results of the developed ML models for subsurface temperature prediction provide critical insights into the performance of different algorithms in geothermal exploration. Of the four models tested, DNN emerged as the most accurate with R^2 values of 0.9784 for training and 0.9783 for testing. This model not only closely matches the training performance of the RF model, but actually outperforms it on the test set, indicating superior generalization capabilities.

Table 3 summarizes the statistical indicator values obtained from the developed models for both training and test data using the three-evaluation metrics adopted. The results show a high performance on the training data, indicating the success of the training of the different models. In addition, a close performance score between the training and test data sets indicates the absence of either overfitting or underfitting.

Fig. 16 provides a visual representation of the model performance using cross plots. The predicted reservoir temperatures using the training data set are plotted against the actual reservoir temperatures, the true response, in the form of a regression line. The performance plots of the four algorithms are shown. The vertical distance from the regression line to any point is the prediction error. Maximum points on the regression line indicate the most accurate algorithm. The cross plots illustrate actual versus predicted temperatures, with DNN showing a tight clustering around the regression line, indicating high prediction accuracy. In contrast, while XGB and BPNN also showed reasonable performance, their points were more scattered from the line, indicating greater variability in the predictions.

The error distributions of the developed models were also plotted in bar charts as shown in Fig. 17. The hallmark of an effective model is its ability to generalize properly, as measured by its performance on unseen data in an independent test dataset [72]. In this comparison, the DNN outperformed all other models, with the lowest RMSE and MAE. This superior performance suggests that the DNN has a greater capacity for generalization, which can be attributed to its adaptive architecture and its enhanced ability to capture and model complex, non-linear relationships between input features and target variables.

The scatter plot in Fig. 18 compares the R^2 values of the four ML models on the training and test datasets. The DNN model shows superior performance and generalization, with high and close R^2 values (0.9784 train, 0.9783 test), which places it closest to the line of perfect fit. The RF model shows high performance, but with a larger gap between train (0.9779) and test (0.9591) scores, indicating some overfitting. BPNN shows moderate, consistent performance (around 0.89), while XGB shows the lowest performance (around 0.81). This visualization effectively illustrates the balance between performance and generalization of each model, with DNN emerging as the most effective approach to capture the underlying patterns of complex geothermal systems

Table 3
Evaluation scores for developed models.

	Random Forest		XGBoost		BPNN		DNN	
	Train	Test	Train	Test	Train	Test	Train	Test
R^2	0.9779	0.9591	0.8111	0.8267	0.9144	0.8836	0.9784	0.9783
RMSE	4.5012	5.5055	13.1643	11.3308	8.8634	9.2857	4.4529	4.0097
MAE	2.8024	3.6209	9.387	8.1612	6.5551	6.711	2.767	2.6363

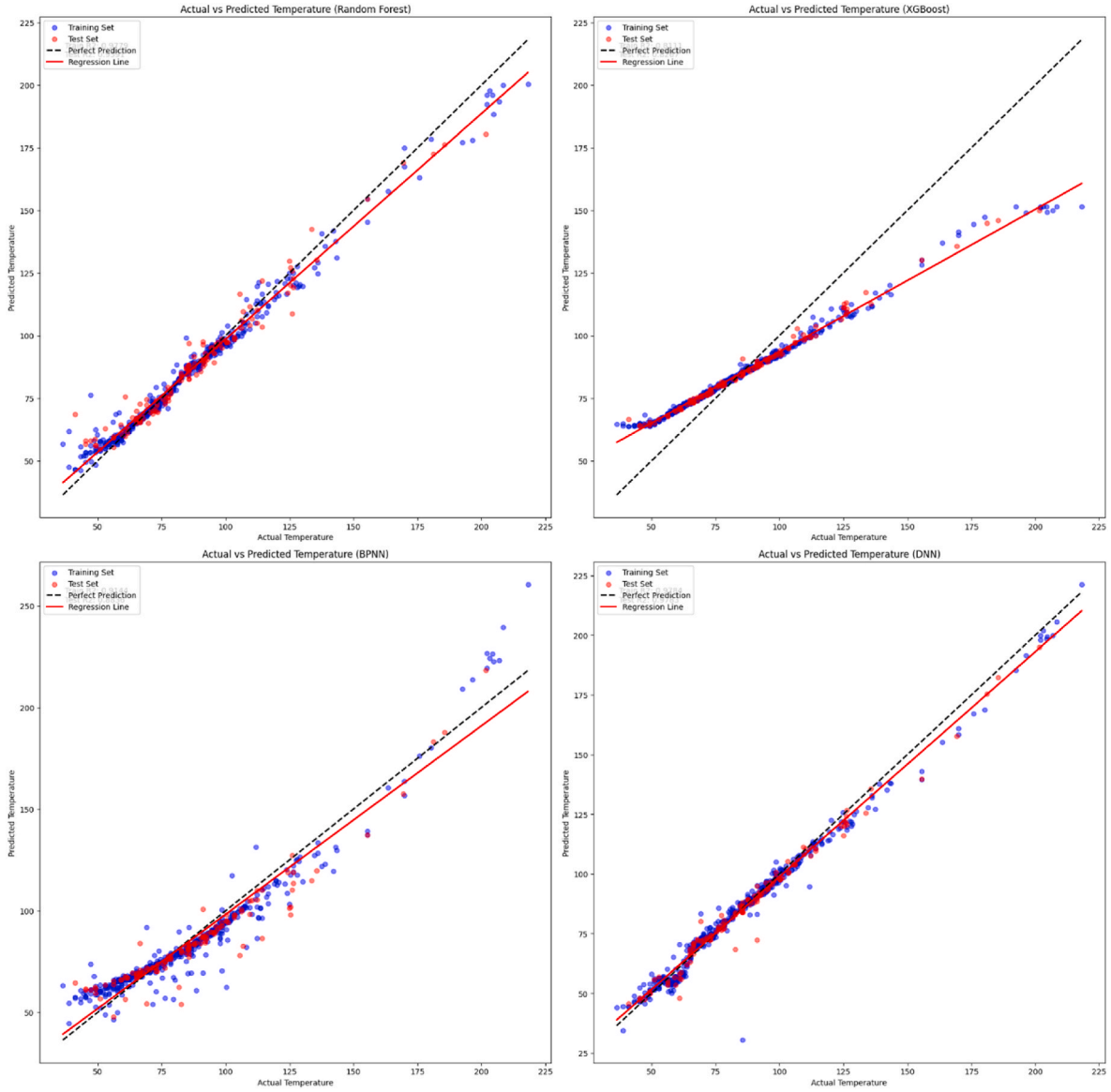


Fig. 16. Cross plots of the RF, XGB, BPNN and DNN models.

characterized by heterogeneous geological settings.

While the performance of the DNN model, named AiION, is comparable to existing data-driven geothermometry approaches, as shown in Table 4, this study primarily examines how the characteristics of the training dataset - specifically, variability in data volume and temperature range - affect the model's generalization and prediction accuracy. Direct comparisons of absolute evaluation metrics (e.g., R^2, RMSE) between models trained on different data distributions become particularly irrelevant when models operate under fundamentally different constraints regarding available training data and target temperature ranges. Instead, this analysis focuses on understanding how data-driven factors specific to the training process affect a model's ability to reliably extrapolate beyond its original training domain.

6. Model validation

This DNN model proved to be very effective in solving our regression problem. Its performance metrics are excellent, with very high R^2 values indicating that it explains over 97 % of the variance in the training and test data. The model's ability to generalize well to unseen data is particularly noteworthy, suggesting that it has captured the underlying patterns in the data without overfitting.

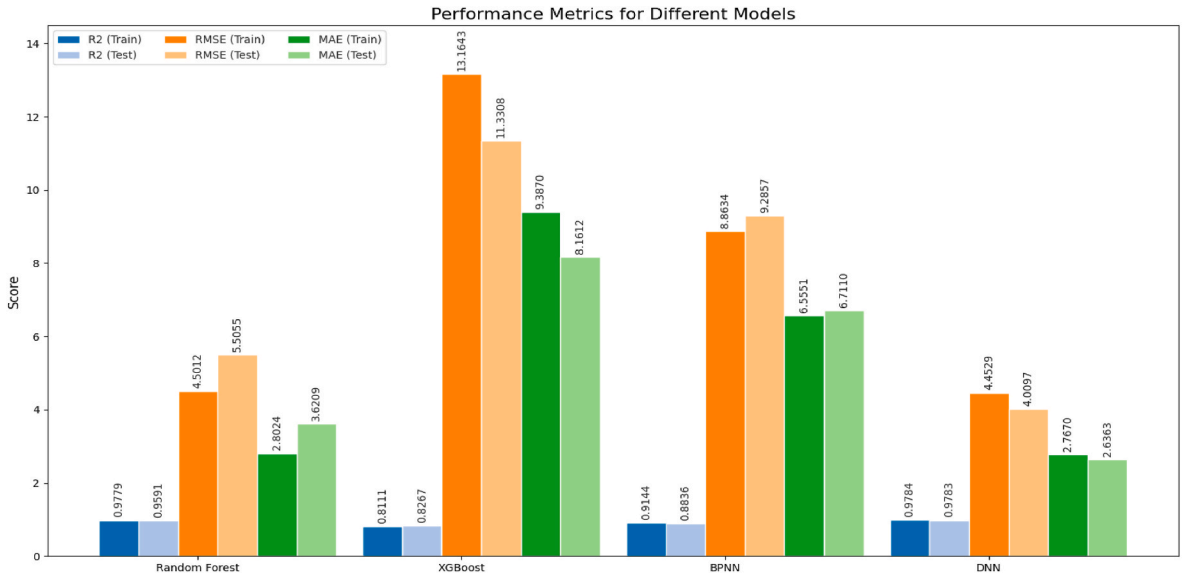


Fig. 17. Comparative plot of performance metrics of all models.

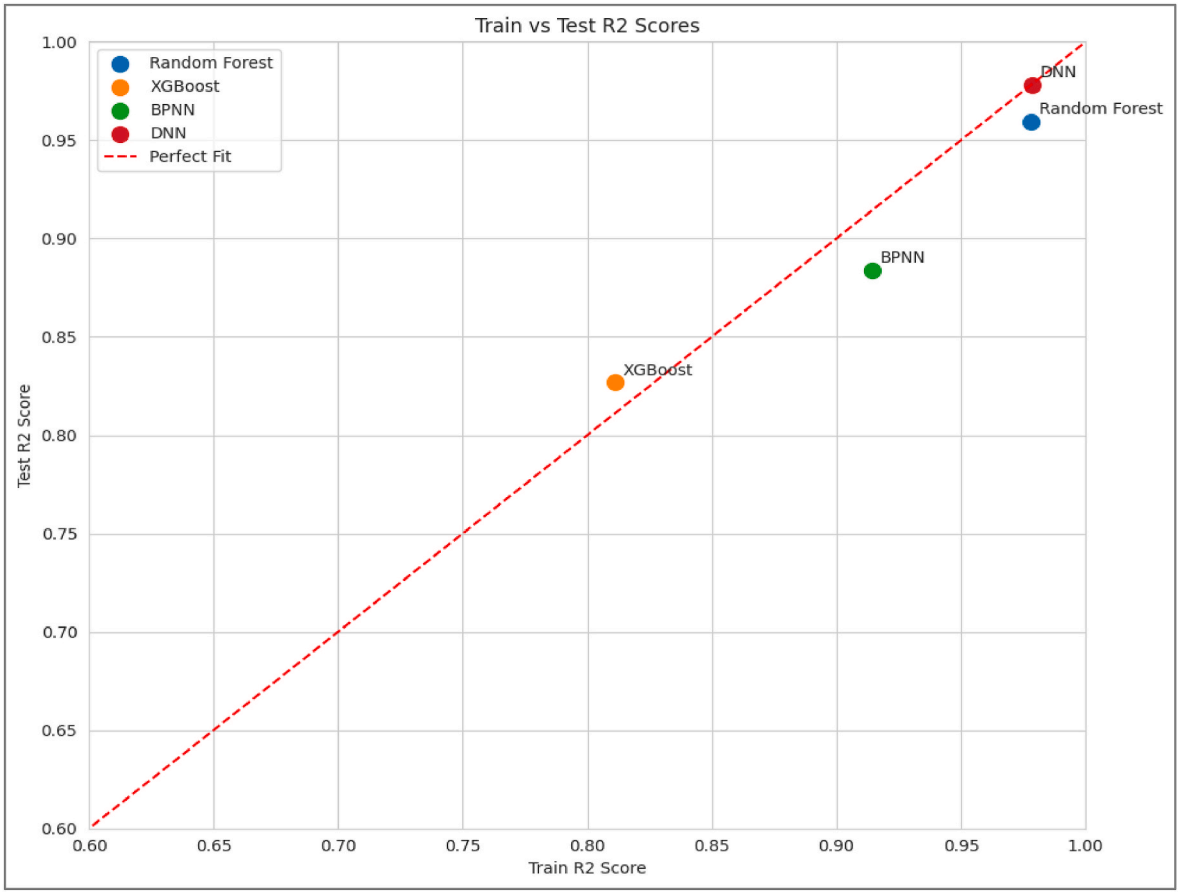


Fig. 18. Scatter plot for Train vs Test R^2 scores.

For further validation, a new geochemical dataset of 42 new wells, including actual borehole temperatures obtained from literature for different geothermal regions around the world [73–79], was compiled to test the model on external data. This dataset was rigorously tested against the best performing model, DNN, to assess its effectiveness as a global model for predicting subsurface temperatures. This comparison was crucial because, unlike the main dataset, the temperature profile dataset contains temperature measurements from different geographical regions, some of which are outside the training boundaries of the DNN model. Table 5 summarizes the statistics of the chemical constituents of the validation data.

The evaluation included several metrics, including R^2, RMSE and MAE. The results showed that the DNN model has good overall predictive capabilities for temperature, as evidenced by its R^2 value of 0.8299

Table 4
Evaluation results for AiION in comparison to previously developed models.

Article	Algorithm	Train/Test Data Size	Temperature Range	Evaluation Matrices (Test Set)		
Haklidir & Haklidir (2020) [12]	DNN	66/17	50–245 °C	RMSE of 8.29		
Ibrahim et al. (2023) [13]	NGB	67/17	50–245 °C	R^2 of 0.9959	RMSE of 4.5938	MAE of 3.9678
Ystroem et al. (2023) [16]	ANN	155/61	36–295 °C	R^2 of 0.9780	RMSE of 10.091	MAPE of 0.092
Dashtgoli et al. (2024) [17]	XGB	59/15	16.3–47.4 °C	R^2 of 0.9930	RMSE of 0.788	MAE of 0.587
AiION	DNN	539/135	36.4–218.2 °C	R^2 of 0.9783	RMSE of 4.0097	MAE of 2.6363

Table 5
Validation dataset constituents statistics.

	Mean	Std	Min	Max	25 %	75 %
pH	8.75	1.02	5.57	9.75	8.56	9.46
K	80.79	238.96	0.74	1553.91	3.10	101.88
Na	714.71	1659.82	34.00	10705.29	73.15	874.20
Mg	5.33	17.26	0.00	80.38	0.03	0.44
Ca	91.87	304.79	0.24	1893.78	1.85	34.40
Cl	677.17	3285.10	10.50	21351.57	29.16	143.00
F	6.12	8.33	0.23	24.80	0.70	9.64
SiO_2	246.91	191.86	38.00	829.84	97.50	382.00
T_{actual}	151.61	70.49	49.20	310.00	92.25	202.00

and the strong correlation between actual and predicted values, indicating that it explains about 83 % of the variance in the target variable. The model does a good job of capturing the general trend of temperature variations over different areas. However, the RMSE of 28.7 and MAE of 22.8 indicate that individual predictions can deviate from actual values by an average of 20–30°. It is also worth noting that the charge balance error of many of the validation water samples was outside the ±5 % acceptable range, *making these samples questionable*. Although the model's performance is solid, there's room for improvement, particularly in reducing the prediction error for individual cases. Strategies to improve performance include collecting more data, feature engineering, ensemble methods and hyperparameter tuning. Fig. 19 shows representation of actual versus predicted temperatures, which further validates the accuracy of the model by showing a tight clustering around the regression line. The results showed that the predictions of the DNN model were in close agreement with the measured data, reinforcing its reliability and applicability in different geological environments.

Despite the encouraging results, some limitations must be acknowledged. Although the training dataset used is larger than that used in previous studies (n = 539 samples), it may still not fully capture the heterogeneity of subsurface conditions in diverse global geothermal systems, leading to uncertainties in temperature predictions for new samples outside the training range of the developed model. The upper and lower bounds of the data points may also affect the reliability of the models, especially for extreme conditions with specific geological anomalies that may not be well represented in the training data.

7. Opportunities

Despite its overwhelming potential, ML has limitations. Most importantly, it is largely unable to extrapolate outside the domains in which it has learned with desirable accuracy [80]. Another major challenge in geothermal exploration is the diverse and complex nature of the subsurface environment. This variability means that no single ML approach can be universally effective across all geothermal systems worldwide. The unique characteristics of each system may require tailored or adaptive methods to achieve optimal results.

Looking ahead, several future directions can further enhance this research, including analyzing model performance across different water types and geological contexts to identify potential biases and limitations, and exploring targeted improvements through specialized models for specific water types or geological contexts; expanding the dataset to complement training intervals by incorporating additional samples from

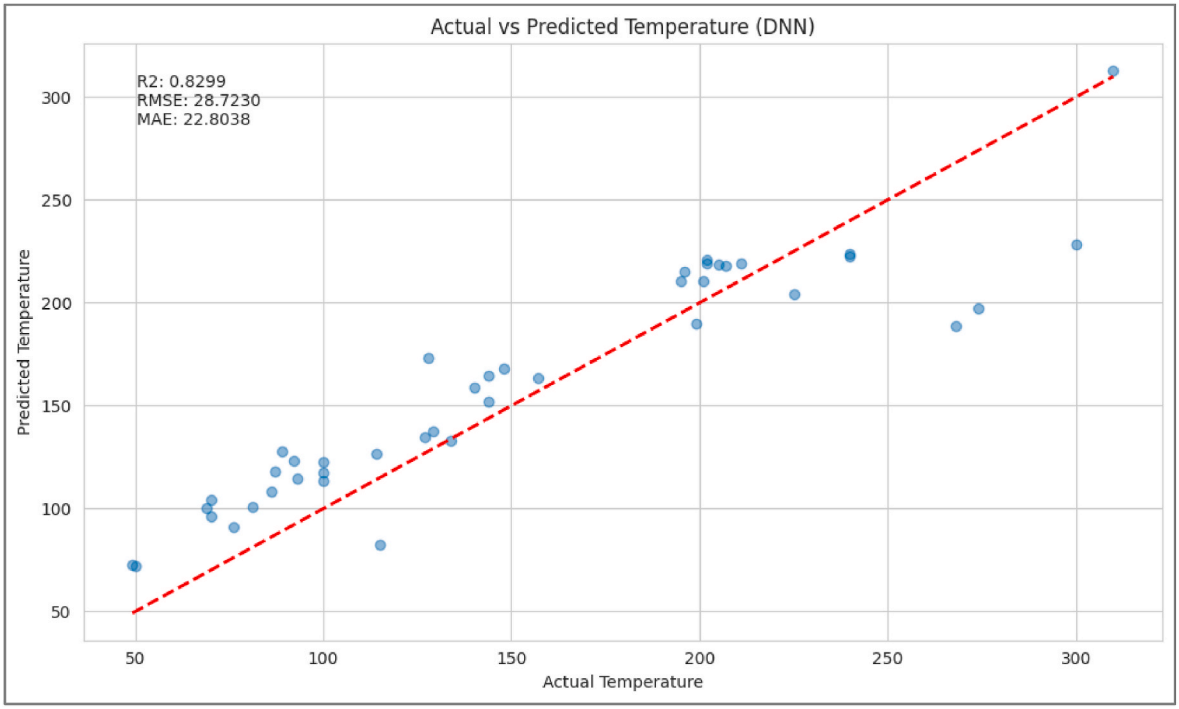


Fig. 19. Scatter plot for Train vs Test R^2 scores of the validation dataset.

different geothermal fields around the world will improve model robustness; incorporating more feature engineering into the model could refine predictions and broaden applicability to different geological contexts. In addition, the development of a user-friendly website (aiION.ai) to disseminate results and facilitate access to the model would support ongoing geothermal research and investment opportunities.

Overall, this comprehensive evaluation highlights the potential of ML models to improve geothermal resource assessments by providing accurate predictions of subsurface temperatures based on extensive hydrogeochemical data. The integration of these advanced modeling techniques with traditional geothermometry provides a robust framework for understanding geothermal potential in complex geological environments worldwide. Accurate and reliable assessment of geothermal resources is essential for informed governance and policy decisions related to energy planning and land use, and we believe this study will provide complementary analysis for geothermal exploration for future investment.

8. Conclusions

This research addresses the challenge of limited subsurface temperature data by employing a novel and integrated approach to infer target temperatures, combining classical and multi-component geothermometry, a regional thermal database, and insights from previous machine learning models to create a unique dataset of 674 water samples from Nevada, a well-known geothermal region with a wide variety of geological settings and high fluid chemistry complexity. This comprehensive dataset served as the basis for evaluating four ML algorithms - RF, XGB, BPNN, and DNN - for their ability to predict subsurface temperatures in geothermal reservoirs, ultimately demonstrating the superior suitability of the DNN model, named AiION, for subsurface temperature prediction. The key value of AiION lies in its streamlined and efficient approach that relies solely on readily available standard geochemical fluid analysis (pH and major ion concentrations), as opposed to traditional multi-component methods that require complex optimization and mineralogical expertise.

The key findings of this study underscore the exceptional performance and global potential of the AiION model. AiION achieved high prediction accuracy on the Nevada dataset with an R^2 of 0.978 and a low Mean Absolute Error (MAE) of approximately 2.7 % for both training and test data. This indicates a strong ability to explain the variance in the data with minimal prediction error. Furthermore, the study demonstrated the global applicability of AiION through successful testing on 42 new well samples from different geothermal fields worldwide, highlighting its reliability in different geological environments.

This research makes significant contributions by presenting a new, large and comprehensive thermo-geochemical dataset from Nevada and by introducing AiION, a novel and highly accurate deep learning chemical geothermometer that outperforms traditional methods and other tested ML algorithms. AiION's reliance on standard geochemical data simplifies the temperature prediction process and provides a more practical tool for geothermal exploration, in line with UN Sustainable Development Goals seven and thirteen for affordable and clean energy and climate action.

CRediT authorship contribution statement

Mahmoud AlGaiar: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Shahana Bano:** Writing – review & editing, Supervision, Methodology. **Aref Lashin:** Writing – review & editing, Supervision, Resources. **Mamdud Hossain:** Writing – review & editing, Supervision. **Nadimul Haque Faisal:** Writing – review & editing, Supervision, Resources, Project administration, Conceptualization. **Hend S. Abu Salem:** Writing – review & editing, Methodology.

Data statement

Available upon reasonable request.

Funding

This study did not receive any specific grants from funding agencies in the public, commercial, or non-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to acknowledge Robert Gordon University for providing resources and support for this paper and for their invaluable contributions to the completion of this research.

References

[1] A.H. Truesdell, Summary of section III-Geochemical techniques in exploration, in: Proc. 2nd UN Symp. on the Development and Use of Geothermal Resources, vol. 1, 1976, pp. 1iii-1xxix.
[2] F. Nitschke, S. Held, T. Neumann, T. Kohl, Geochemical characterization of the villarrica geothermal system, southern Chile, part II: site-specific re-evaluation of SiO2 and Na-K solute geothermometers, Geothermics 74 (2018) 217–225, https://doi.org/10.1016/j.geothermics.2018.03.006.
[3] W.E. Glassley, Geothermal Energy: Renewable Energy and the Environment, first ed., CRC press, 2010 https://doi.org/10.1201/EBK1420075700.
[4] M. Reed, N. Spycher, Calculation of pH and mineral equilibria in hydrothermal waters with application to geothermometry and studies of boiling and dilution, Geochem. Cosmochim. Acta 48 (7) (1984) 1479–1492, https://doi.org/10.1016/0016-7037(84)90404-6.
[5] N. Spycher, L. Peiffer, E.L. Sonnenthal, G. Saldi, M.H. Reed, B.M. Kennedy, Integrated multicomponent solute geothermometry, Geothermics 51 (2014) 113–123, https://doi.org/10.1016/j.geothermics.2013.10.012.
[6] B. Ahmmed, V.V. Vesselinov, Machine learning and shallow groundwater chemistry to identify geothermal prospects in the great basin, USA, Renew. Energy 197 (2022) 1034–1048, https://doi.org/10.1016/j.renene.2022.08.024.
[7] A.F. Bayram, Application of an artificial neural network model to a Na–K geothermometer, J. Volcanol. Geoth. Res. 112 (1–4) (2001) 75–81, https://doi.org/10.1016/S0377-0273(01)00235-9.
[8] I. Can, A new improved Na/K geothermometer by artificial neural networks, Geothermics 31 (6) (2002) 751–760, https://doi.org/10.1016/S0375-6505(02)00044-5.
[9] L. Diaz-Gonzalez, E. Santoyo, J. Reyes-Reyes, Three new improved Na/K geothermometers using computational and geochemiometrical tools: application to the temperature prediction of geothermal systems, Rev. Mex. Ciencias Geol. 25 (3) (2008) 465–482.
[10] D. Pérez-Zárate, E. Santoyo, A. Acevedo-Anicasio, L. Díaz-González, C. García-López, Evaluation of artificial neural networks for the prediction of deep reservoir temperatures using the gas-phase composition of geothermal fluids, Comput. Geosci. 129 (2019) 49–68, https://doi.org/10.1016/j.cageo.2019.05.004.
[11] A. Acevedo-Anicasio, E. Santoyo, D. Pérez-Zárate, K. Pandarinath, M. Guevara, L. Díaz-González, GaS_GeoT: a computer program for an effective use of newly improved gas geothermometers in predicting reliable geothermal reservoir temperatures, Geoth. Energy 9 (2021) 1–41, https://doi.org/10.1186/s40517-020-00182-9.
[12] F.S. Tut Haklidir, M. Haklidir, Prediction of reservoir temperatures using hydrogeochemical data, Western Anatolia geothermal systems (turkey): a machine learning approach, Nat. Resour. Res. 29 (4) (2020) 2333–2346, https://doi.org/10.1007/s11053-019-09596-0.
[13] B. Ibrahim, J.O. Konduah, I. Ahenkorah, Predicting reservoir temperature of geothermal systems in west Anatolia, Turkey: a focus on predictive performance and explainability of machine learning models. Turkey: a focus on predictive performance and explainability of machine learning models. https://doi.org/10.2139/ssrn.4239319, 2022.
[14] E.V. Altay, E. Gurgenc, O. Altay, A. Dikici, 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 (2022) 102476, https://doi.org/10.1016/j.geothermics.2022.102476.
[15] F. Yang, R. Zhu, X. Zhou, T. Zhan, X. Wang, J. Dong, L. Liu, Y. Ma, Y. Su, Artificial neural network based prediction of reservoir temperature: a case study of Lindian geothermal field, songliao basin, NE China, Geothermics 106 (2022) 102547, https://doi.org/10.1016/j.geothermics.2022.102547.

[16] L.H. Ystroem, M. Vollmer, T. Kohl, F. Nitschke, AnnRG-An artificial neural network solute geothermometer, Appl. Comput. Geosci. 20 (2023) 100144, https://doi.org/10.1016/j.acags.2023.100144.

[17] D.S. Dashtgoli, M. Giustiniani, M. Busetti, C. Cherubini, Artificial intelligence applications for accurate geothermal temperature prediction in the lower Friulian plain (north-eastern Italy), J. Clean. Prod. 460 (2024) 142452, https://doi.org/10.1016/j.jclepro.2024.142452.

[18] University of Nevada, Geoth. Energy (2024). https://nbmg.unr.edu/Research/GeothermalEnergy.html. (Accessed 25 September 2024).

[19] B.F. Ayling, 35 years of geothermal power generation in Nevada, USA: a review of field development, generation, and production histories, in: 45th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California, 2020, February.

[20] M.F. Coolbaugh, G.L. Raines, R.E. Zehner, L. Shevenell, C.F. Williams, Prediction and discovery of new geothermal resources in the great basin: multiple evidence of a large undiscovered resource base, in: GRC 2006 Annual Meeting: Geothermal Resources-Securing our Energy Future, vol. 30, 2006, pp. 867–873.

[21] J.E. Faulds, M.F. Coolbaugh, N.H. Hinz, Inventory of structural setting for active geothermal systems and late Miocene (~ 8 MA) to Quaternary epithermal mineral deposits in the basin and range province of Nevada, Rep. - Nev. Bur. Mines Geol. 58 (2021) 28, 3 plates, scale 1:2,500,000.

[22] A.H. Lachenbruch, J.H. Sass, Heat flow in the United States and the thermal regime of the crust, Earth's crust. 20 (1977) 626–675, https://doi.org/10.1029/GM020p0626.

[23] D.D. Blackwell, J.L. Steele, L.C. Carter, Heat flow patterns of the North American continent: a discussion of the DNAG geothermal map of North America, DOEEEGTP (USDOE Office Energ. Effic. Renew. Energ. Geother. Tech. Pgm) (1990), https://doi.org/10.2172/896322.

[24] D.L. Siler, J.E. Faulds, N.H. Hinz, G.M. Dering, J.H. Edwards, B. Mayhew, Three-dimensional geologic mapping to assess geothermal potential: examples from Nevada and Oregon, Geothermal Energy 7 (2019) 1–32, https://doi.org/10.1186/s40517-018-0117-0.

[25] J. Faulds, N. Hinz, Favorable tectonic and structural settings of geothermal systems in the great basin region, Western USA: proxies for discovering blind geothermal systems, in: Proceedings World Geothermal Congress, Melbourne, Australia, 19–25, Nevada Bureau of Mines and Geology, University of Nevada, Reno, 2015, April. DOE-UNR-06731-02).

[26] J.E. Faulds, M. Coolbaugh, G. Blewitt, C.D. Henry, Why Is Nevada in Hot Water? Structural Controls and Tectonic Model of Geothermal Systems in the Northwestern Great Basin, vol. 28, Geothermal Resources Council Transactions, 2004, pp. 649–654.

[27] J.H. Stewart, Geology of Nevada: a Discussion to Accompany the Geologic Map of Nevada, vol. 4, NV Bureau of Mines & Geology. Special Publication, 1980, p. 136.

[28] E. Pérouse, B.P. Wernicke, Spatiotemporal evolution of fault slip rates in deforming continents: the case of the great basin region, northern basin and range province, Geosphere 13 (1) (2017) 112–135.

[29] D.K. Maurer, T.J. Lopes, R.L. Medina, J.L. Smith, Hydrogeology and Hydrologic Landscape Regions of Nevada, vol. 5131, USGS Scientific Investigations Report, 2004.

[30] F.H. Olmsted, A.H. Welch, M.L. Sorey, D.H. Schaefer, The geothermal hydrology of southern grass valley, pershing county, Nevada (no. 96-4139), US Geol. Surv. Water Resour. Invest. Report (1997) 96–4139.

[31] S.W. Lohman, Definitions of Selected ground-water terms–revisions and Conceptual Refinements, US Government Printing Office, 1972.

[32] J.R. Harrill, D.E. Prudic, Aquifer Systems in the Great Basin Region of Nevada, Utah, and Adjacent States: Summary Report (No. 1409), US Government Printing Office, 1998.

[33] D.F. Schulke (Ed.), Great Basin Recharge Studies. Water Resources Center, Desert Research Institute, University of Nevada, 1987.

[34] M.D. Mifflin, Delineation of ground-water Flow Systems in Nevada, Doctoral dissertation, University of Nevada, Reno, 1968.

[35] M.R. Heynekamp, L.B. Goodwin, P.S. Mozley, W.C. Haneberg, Controls on fault-zone architecture in poorly lithified sediments, rio grande rift, New Mexico: implications for fault-zone permeability and fluid flow, Geophys. Monogr. 113 (1999) 27–49, https://doi.org/10.1029/GM113p0027.

[36] R.K. Hose, B.E. Taylor, Geothermal Systems of Northern Nevada (No. 74-271), US Geological Survey, 1974, https://doi.org/10.3133/ofr74271. Open File Report 74-271.

[37] S.K. Sanyal, Classification of geothermal systems–a possible scheme, in: Thirtieth Workshop on Geothermal Reservoir Engineering, vol. 2, Stanford University, Stanford, California, 2005. No. 02).

[38] H. Schoeller, Utilite de la notion des exchanges de bases pour le comparison des eaux souterraines. Société Géologie Comptes Rendus Sommaire et Bulletin, série 5 (1935) 651–657.

[39] J. Bundschuh, J.P. Maity, B. Nath, A. Baba, O. Gunduz, T.R. Kulp, J.S. Jean, S. Kar, H.J. Yang, Y.J. Tseng, P. Bhattacharya, Naturally occurring arsenic in terrestrial geothermal systems of Western Anatolia, Turkey: potential role in contamination of freshwater resources, J. Hazard Mater. 262 (2013) 951–959, https://doi.org/10.1016/j.jhazmat.2013.01.039.

[40] S. Simmons, S. Kirby, R. Allis, J. Moore, T. Fischer, Update on Production Chemistry of the Roosevelt Hot Springs Reservoir. University of Utah. Stanford Geothermal Workshop, Stanford University, Palo Alto, CA, 2018.

[41] L. Shevenell, M.F. Coolbaugh, A new method of evaluation of chemical geothermometers for calculating reservoir temperatures from thermal springs in Nevada, Trans. Geoth. Resour. Counc. 35 (2011) 657–661.

[42] A.M. Piper, A graphic procedure in the geochemical interpretation of water-analyses, Eos, Trans. American Geophys. Un. 25 (6) (1944) 914–928, https://doi.org/10.1029/TR025i006p00914.

[43] Nevada Bureau of Mines and Geology, Geothermal minerals of Nevada. https://storymaps.arcgis.com/stories/624b919af8cb4382824a47fd14fef13a, 2022. Available from: [Accessed Sept 2024].

[44] W.F. Giggenbach, Geothermal solute equilibria. derivation of Na-K-Mg-Ca geoindicators, Geochem. Cosmochim. Acta 52 (12) (1988) 2749–2765, https://doi.org/10.1016/0016-7037(88)90143-3.

[45] W.F. Giggenbach, Geothermal mineral equilibria, Geochem. Cosmochim. Acta 45 (3) (1981) 393–410, https://doi.org/10.1016/0016-7037(81)90248-9.

[46] D.L. Parkhurst, C.A.J. Appelo, PHREEQC (version 3)-A computer program for speciation. batch-reaction, one-dimensional transport, and inverse geochemical calculations, US Geol. Surv. Water Resour. Div. 6 (A43) (2013) 497, https://doi.org/10.3133/tm6A43.

[47] J.C.L. Meeussen, H.A. van der Sloot, J.J. Dijkstra, D.S. Kosson, Review of Thermodynamic and Adsorption Databases, Energy Research Centre of the Netherlands and Vanderbilt University/CRESP, 2009.

[48] X. Wang, G.L. Wang, H.N. Gan, Z. Liu, D.W. Nan, Hydrochemical characteristics and evolution of geothermal fluids in the chabu high-temperature geothermal system, southern Tibet, Geofluids 2018 (1) (2018) 8532840, https://doi.org/10.1155/2018/8532840.

[49] F.K. Zaidi, A. Lashin, E. Aboud, N. Al Arifi, A. Al-Bassam, E. Al-Homadhi, M. A. Anbar, Silica geothermometry and multi indices approach to characterize groundwater from Midyan region in northwestern Saudi Arabia, J. Afr. Earth Sci. 192 (2022) 104557, https://doi.org/10.1016/j.jafrearsci.2022.104557.

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

[51] R.O. Fournier, Chemical geothermometers and mixing models for geothermal systems, Geothermics 5 (1–4) (1977) 41–50, https://doi.org/10.1016/0375-6505(77)90007-4.

[52] Southern Methodist University, Geothermal laboratory heat flow database. http://smu.edu/geothermal/, 2024 [Accessed Sept 2023].

[53] R.A. Freeze, J.A. Cherry, Groundwater prentice-hall. Englewood Cliffs, NJ, vol. 176, 1979, pp. 161–177.

[54] D.F. Williamson, R.A. Parker, J.S. Kendrick, The box plot: a simple visual method to interpret data, Ann. Intern. Med. 110 (11) (1989) 916–921, https://doi.org/10.7326/0003-4819-110-11-916.

[55] P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J. Comput. Appl. Math. 20 (1987) 53–65, https://doi.org/10.1016/0377-0427(87)90125-7.

[56] T. Caliński, J. Harabasz, A dendrite method for cluster analysis, Commun. Stat. Theor. Methods 3 (1) (1974) 1–27, https://doi.org/10.1080/03610927408827101.

[57] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (2011) 2825–2830.

[58] Y. Liu, Y. Wang, J. Zhang, New machine learning algorithm: random forest, in: Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14-16, 2012. Proceedings, vol. 3, Springer Berlin Heidelberg, 2012, pp. 246–252, https://doi.org/10.1007/978-3-642-34062-8_32.

[59] B. Ghimire, J. Rogan, V.R. Galiano, P. Panday, N. Neeti, An evaluation of bagging, boosting, and random forests for land-cover classification in cape cod, Massachusetts, USA, GIScience Remote Sens. 49 (5) (2012) 623–643, https://doi.org/10.2747/1548-1603.49.5.623.

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

[61] N. Moniz, P. Branco, L. Torgo, Resampling strategies for imbalanced time series forecasting, Int. J. Data Sci. Anal. 3 (2017) 161–181, https://doi.org/10.1007/s41060-017-0044-3.

[62] D.A. Otchere (Ed.), Data Science and Machine Learning Applications in Subsurface Engineering, CRC Press, Taylor & Francis Group, 2024, https://doi.org/10.1201/9781003366980.

[63] J. Friedman, T. Hastie, R. Tibshirani, Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors), Ann. Stat. 28 (2) (2000) 337–407, https://doi.org/10.1214/aos/1016218223.

[64] T. Chen, C. Guestrin, Xgboost: a scalable tree boosting system, in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2016, August, pp. 785–794, https://doi.org/10.1145/2939672.293978.

[65] J. Montiel, R. Mitchell, E. Frank, B. Pfahringer, T. Abdessalem, A. Bifet, Adaptive xgboost for evolving data streams, in: 2020 International Joint Conference on Neural Networks (IJCNN), IEEE, 2020, July, pp. 1–8, https://doi.org/10.1109/IJCNN48605.2020.9207555.

[66] A.J. Maren, C.T. Harston, R.M. Pap, Handbook of Neural Computing Applications, Academic Press, 2014.

[67] A.J. Maren, Statistical Mechanics, Neural Networks, and Machine Learning: Using Powerful Brain Strategies to Improve Machine Learning DRAFT Chapter 3, 2017.

[68] J.J. Hopfield, Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. 79 (8) (1982) 2554–2558, https://doi.org/10.1073/pnas.79.8.2554.

[69] W. Samek, G. Montavon, S. Lapuschkin, C.J. Anders, K.R. Müller, Explaining deep neural networks and beyond: a review of methods and applications, Proc. IEEE 109 (3) (2021) 247–278, https://doi.org/10.1109/JPROC.2021.3060483.

[70] D. Chicco, M.J. Warrens, G. Jurman, The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, Peerj comput. sci. 7 (2021) e623.
[71] T.O. Hodson, Root mean square error (RMSE) or mean absolute error (MAE): when to use them or not, Geosci. Model Dev. Discuss. (GMDD) 2022 (2022) 1–10, https://doi.org/10.5194/gmd-15-5481-2022.
[72] T. Hastie, R. Tibshirani, J.H. Friedman, J.H. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, vol. 2, springer, New York, 2009, pp. 1–758, https://doi.org/10.1007/978-0-387-21606-5.
[73] L.H. Ystroem, F. Nitschke, S. Held, T. Kohl, A multicomponent geothermometer for high-temperature basalt settings, Geoth. Energy 8 (2020) 1–21, https://doi.org/10.1186/s40517-020-0158-z.
[74] B.F. Ayling, N.H. Hinz, Developing a conceptual model and power capacity estimates for a low-temperature geothermal prospect with two chemically and thermally distinct reservoir compartments, hawthorne, Nevada, USA, Geothermics 87 (2020) 101870, https://doi.org/10.1016/j.geothermics.2020.101870.
[75] A. Gokgoz, Geochemistry of the Kizildere-Tekkehamam-Buldan-Pamukkale geothermal fields, Turkey, Geoth. Train. Icel. 1998 (1998) 115–156.
[76] Ö. Avşar, G. Altuntaş, Hydrogeochemical evaluation of umut geothermal field (SW Turkey), Environ. Earth Sci. 76 (2017) 1–22, https://doi.org/10.1007/s12665-017-6929-6.
[77] S. Arnórsson, E. Gunnlaugsson, H. Svavarsson, The chemistry of geothermal waters in Iceland. II. Mineral equilibria and independent variables controlling water compositions, Geochem. Cosmochim. Acta 47 (3) (1983) 547–566, https://doi.org/10.1016/0016-7037(83)90277-6.
[78] S. Arnórsson, G. Axelsson, K. Sæmundsson, Geothermal systems in Iceland, Jökull 58 (2008) 269–302.
[79] R. Liu, H. Li, Z. Zhao, Z. Zhang, Analysis of geothermal fluid chemical characteristics and genetic model—A case study from the urban area of Jingmen China, Front. Earth Sci. 10 (2023) 1081781, https://doi.org/10.3389/feart.2022.1081781.
[80] B. Lantz, Machine Learning with R: Expert Techniques for Predictive Modelling, Packt publishing ltd, 2019.