

Evaluation of Machine Learning Models for Geothermal Downhole Temperature Prediction Using Hydrogeochemical Datasets

Mohammed Ali Belay

`mohammed.ali@aastu.edu.et`

Addis Ababa Science and Technology University

Elias Mamushet

Addis Ababa Science and Technology University

Assnake Bekele

Addis Ababa Science and Technology University

Tarekegn Tadesse

Addis Ababa Science and Technology University

Research Article

Keywords: Downhole temperature prediction, Geothermal exploration, Machine learning, Hydrogeochemical variables

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11427714/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Evaluation of Machine Learning Models for Geothermal Downhole Temperature Prediction Using Hydrogeochemical Datasets

Mohammed Ali Belay^{1,2,*}, Elias Mamushet^{1,2}, Assnake Bekele^{2,3}, Tarekegn Tadesse^{2,3}

¹ *Mining Engineering Department, Addis Ababa Science and Technology University, P.O. Box 16417, Addis Ababa, Ethiopia*

² *Mineral Exploration, Extraction, and Processing Center of Excellence, Addis Ababa Science and Technology University, P.O. Box 16417, Addis Ababa, Ethiopia*

³ *Geology Department, Addis Ababa Science and Technology University, P.O. Box 16417, Addis Ababa, Ethiopia*

* *Corresponding author, email: mohammed.ali@aastu.edu.et*

Abstract

Accurate estimation of geothermal downhole temperatures is predominantly useful in areas where direct subsurface temperature measurements are limited. This method supports resource assessment and helps to reduce uncertainty during exploration. In this study, we used hydrogeochemical and physicochemical variables to develop and assess machine learning (ML) models to estimate geothermal downhole temperatures in different regions and geological settings. Ridge regression (RR), support vector regression (SVR), random forest (RF), extreme gradient boosting (XGBoost), and CatBoost are the ML approaches used in this paper. We assessed model performance through an independent test dataset and leave-one-out cross-validation (LOOCV). The predictive performance of the models was evaluated using parameters such as the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE). CatBoost achieved the highest overall predictive performance, with a test R^2 of approximately 0.91 and an RMSE of 23.31 °C, whereas Random Forest produced the lowest MAE of 11.21 °C and showed close agreement with the ideal 1:1 relationship. Moreover, XGBoost demonstrated strong predictive capability, whereas RR and SVR presented substantially larger prediction errors and greater compression of the temperature range. In addition, LOOCV and residual diagnostics were used to assess model robustness, prediction bias, and temperature-dependent error patterns. The results clearly revealed that ML, especially tree-based models, is an excellent tool for geothermal development where direct subsurface temperature measurements are limited.

Keywords: Downhole temperature prediction, Geothermal exploration, Machine learning, Hydrogeochemical variables

1. Introduction

Ethiopia is endowed with different energy resources, including both renewable and nonrenewable sources such as hydropower, sunlight, geothermal energy, coal, and other fossil fuel resources. Among these energy sources, geothermal energy is recognized as a sustainable option that significantly contributes to reducing carbon dioxide (CO₂) emissions and mitigating the carbon footprint associated with fossil-fuel-based production.

At present, 90% of the total electric power generation in Ethiopia comes from hydropower, and 10% comes from wind, geothermal, solar, coal, and other sources. Major geothermal projects like the Aluto-Langano and Tulu-Moye geothermal fields are in development (Benti et al., 2023)).

The Ethiopian rift valley is one of the most promising areas for geothermal energy development in Ethiopia due to its favorable geological and tectonic settings (Samrock et al., 2021). In the central Main Ethiopian Rift, the most promising geothermal prospects are all located along a narrow network of en echelon-disposed NNE-SSW-trending faults, known as the Wonji fault belt (WFB). These faults play crucial roles in controlling the rise and circulation of geothermal fluids by providing permeable routes within the fractured crust (Samrock et al., 2021).

1.1. Overview of geothermal development in Ethiopia

Exploration of geothermal energy sources in Ethiopia started in 1969, with some prospecting mechanisms using surface geology, geochemistry, aerial photos, and infrared airborne survey investigations (Burnside et al., 2021). However, until 2013, Ethiopia had a single installed geothermal power plant at Aluto-Langano, which produced less than 1 MW per year (Burnside et al., 2021, after Tadesse, 2018). Among the several prospective areas depicted in Fig. 1, twenty-five geothermal prospects have been identified as good for electricity generation (Kebede et al., 2020). Many of these prospect regions are associated with Quaternary volcanic processes, which are crucial geological phenomena that contribute to their geothermal potential (Chandrasekharam et al., 2015)

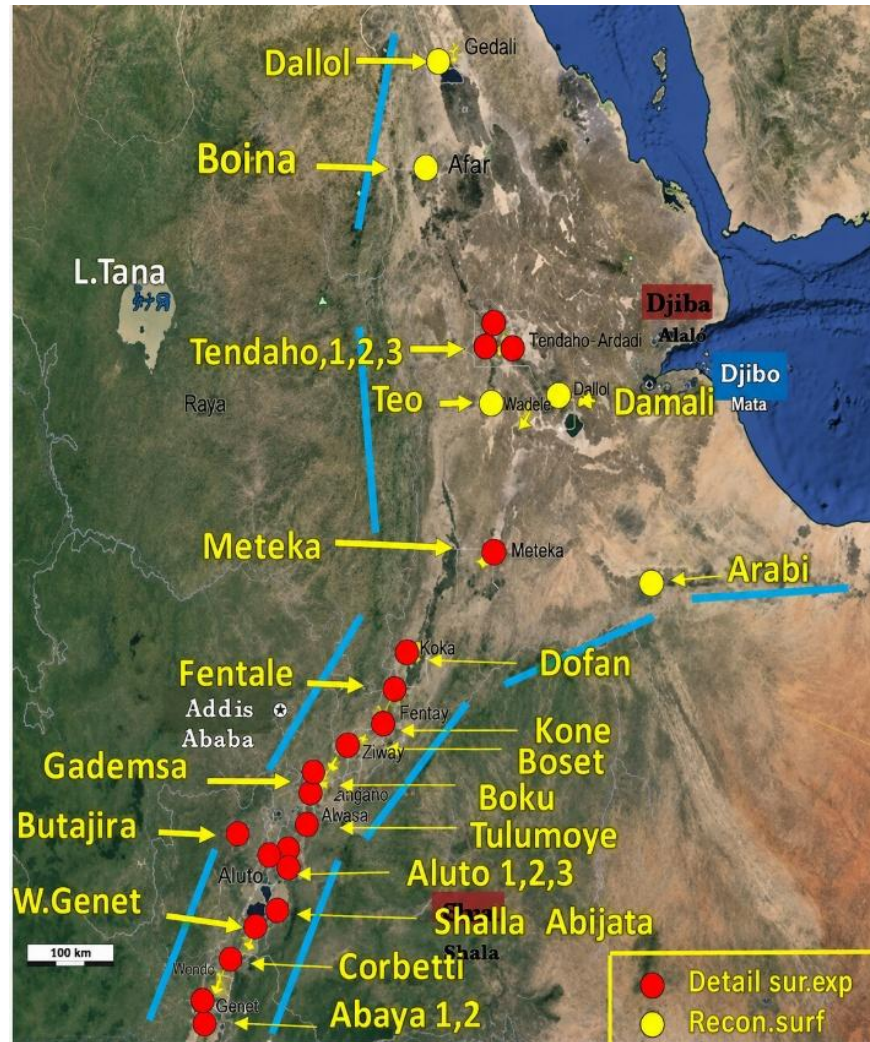


Fig. 1: Geothermal prospect regions of Ethiopia ((Genetu, 2025), after Solomon, K., W. Fikru, and K. Tesfaye, 2021)

The development of geothermal energy takes place in several consecutive stages, starting from a surface survey and followed by exploration drilling. As illustrated in Fig. 2, the geothermal development project stages, risks, and costs relative to each stage are clearly demonstrated. Traditional mechanisms of subsurface feature analysis depend on professional knowledge for evaluation and interpretation, as well as reserve estimation of the resource, leading to uncertainty in the discovery of blind geothermal resources (AlGaier et al., 2024). According to these scholars, recent advances in data-driven modeling have extended the application of artificial intelligence (AI) in geothermal exploration. By identifying complex patterns within large, multivariate datasets and developing predictive models, AI-based approaches can replace traditional professional-based

and statistical methods to reduce exploration uncertainty and improve prediction accuracy. In (AlGaiar et al., 2024), different ML algorithms like linear regression (LR), decision tree (DT), extreme gradient boosting (XGB), and artificial neural network (ANN) were used to facilitate geothermal exploration.

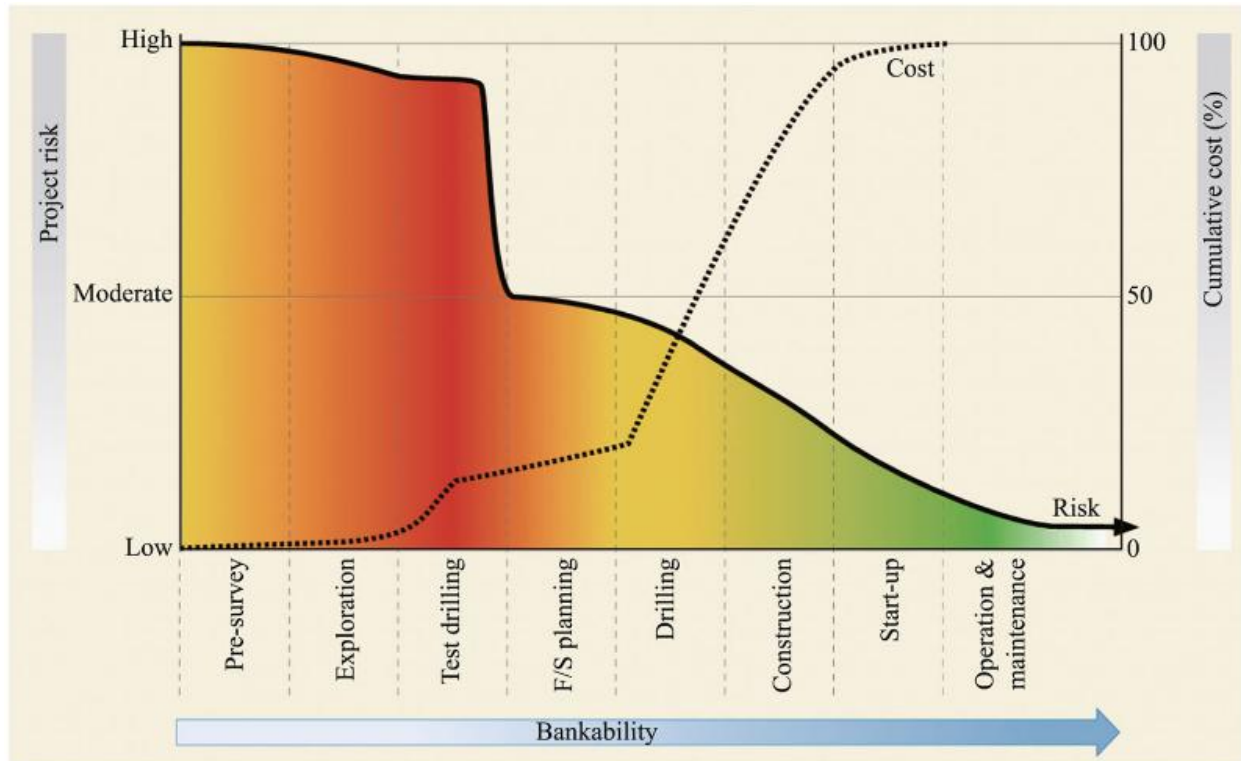


Fig. 2: Geothermal development project costs and risk profiles across the various stages of geothermal project development (AlGaiar et al., 2024)

On the basis of (Zhang et al., 2026), the application of these ML methods is different at different stages of geothermal development. During the resource exploration stage, random forest (RF), XGBoost, support vector machines (SVM), and clustering may be used. In the reservoir characterization category, the application of ML involves the prediction of temperature and thermophysical properties using these ML algorithms, as illustrated in Fig. 3. In addition, (Shoeibi Omrani et al., 2025) emphasize the application of artificial intelligence in geothermal energy systems, where k-means, k-nearest neighbors, random forest, and artificial neural network algorithms are also used in this field. According to Zhu (2024), several machine learning techniques, including random forest (RF), artificial neural networks (ANN), deep neural networks (DNN), k-means clustering, decision trees, and other data-driven methods, have been used to

characterize geothermal reservoirs. These methods are used to predict temperature, optimize well placement, track pressure changes, and estimate production enthalpy in geothermal wells.

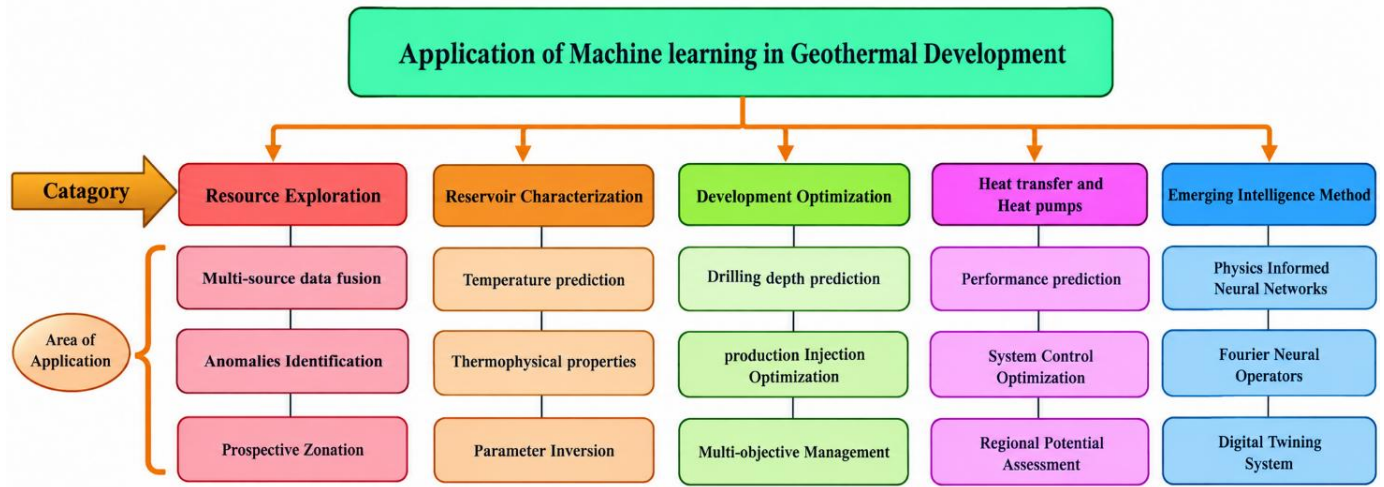


Fig. 3: Schematic framework of machine learning applications in geothermal development.

Modified after (Zhang et al., 2026)

Ethiopia has substantial geothermal potential, with several prospective areas identified across the country Fig. 1. However, geothermal development is constrained by the high cost, time requirements, and uncertainty associated with traditional exploration, particularly the risk of unsuccessful drilling. By identifying patterns in current geological, geochemical, geophysical, and geographical datasets, machine learning (ML) offers a viable complementary strategy to enhance prediction and lower exploration uncertainty. Advanced machine learning techniques are still not widely used in Ethiopia's geothermal industry, despite their increasing global utilization. Therefore, this research fills an important gap by developing and validating ML models for geothermal downhole temperature prediction, with the potential to support more cost-effective and informed geothermal development in Ethiopia.

Numerous studies have used hydrogeochemical indicators, well depth, ion concentration, and structural background data as inputs for developing regression models to predict geothermal reservoir temperatures (Zhang et al., 2026). Those scholars emphasize that machine learning models can integrate more variables and learn more complex nonlinear relationships than conventional geochemical thermometers can. Based on (Zhang et al., 2025), some studies use Shapley additive explanation (SHAP) or sensitivity analysis to identify key controlling variables

that enhance model interpretability. According to (Lundberg and Lee, 2017), SHAP is defined as a game theory-based approach for explaining machine learning models that determines the contribution of each covariate via estimation of the exact value for individual instances.

The predictive performance of the developed model is evaluated by three evaluation metrics; the coefficient of determination (R^2), mean absolute error (MAE), and root mean square error (RMSE). Based on (Ibrahim et al., 2022) the corresponding mathematical equations for those metrics are:

$$R^2 = \frac{[\sum_{i=1}^n (T_a - \bar{T}_a) (T_p - \bar{T}_p)]^2}{\sum_{i=1}^n (T_a - \bar{T}_a)^2 \sum_{i=1}^n (T_p - \bar{T}_p)^2}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (T_a - T_p)^2}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |T_p - T_a|$$

where n represents the total number of test samples, T_a represents the actual/observed temperature, T_p represents the predicted temperature, $\bar{T}_a$ represents the mean of the actual/observed temperature and $\bar{T}_p$ represents the mean of the predicted temperature.

2. Methodology

In this research, the necessary hydrogeochemical data points are extracted from the Database_Ethiopian_Geothermal_final_WIRES_Water dataset developed by Montcoudiol et al. (2021). This dataset is hosted by the University of Strathclyde. The database contains several pieces of information, such as surface and subsurface fluid quality information, including geochemical, physicochemical, and hydrogeochemical parameters. These parameters are crucial for geothermal studies in Ethiopia. From the datasets containing more than 2400 data points, we extracted 197 data points, as shown in Table 1, from 14 geothermal fields with corresponding downhole temperatures and the necessary parameters, with a few missing points. We used the selected major ionic species and some physicochemical parameters as predictor variables and the downhole temperature as a target variable.

The dataset was imported into the analysis environment in Google Colab and subjected to preprocessing workflows following the data collection process. The preprocessing workflow included data cleaning, exploratory data analysis, missing-value treatment, outlier detection, feature engineering, and other necessary preprocessing steps before model development. The general machine learning workflow and all processing procedures are illustrated in Fig. 4.

Finally, five supervised machine learning regression algorithms were selected for geothermal downhole temperature prediction. Those ML models include one regularized linear model (ridge regression), three tree-based ensemble models (random forest (RF), XGBoost, and CatBoost), and one kernel-based model (support vector regression (SVR)). The coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE) values were used to evaluate and compare the predictive performance of the developed models and determine the model that gives the best downhole temperature estimations.

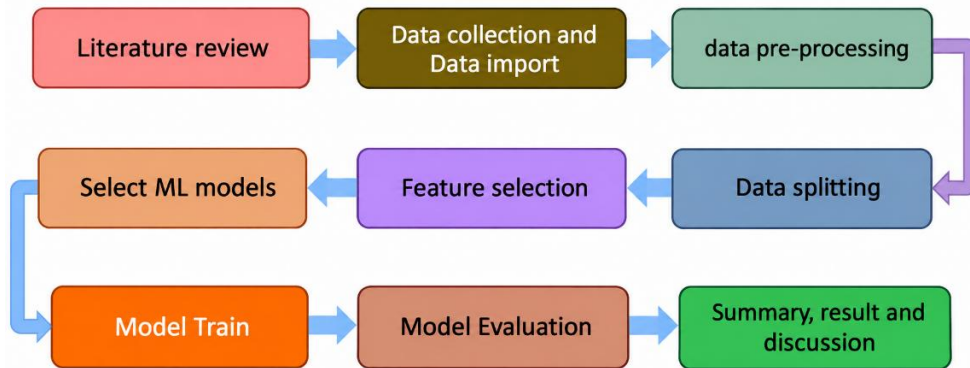


Fig. 4: The overall workflow for downhole temperature prediction using the proposed ML models

Table 1: The number of samples with the corresponding geothermal fields

Geothermal Fields	Aluto-Langano	Abaya	Tendaho	Corbetti	Dofan-	Rift Valley	Abhe	Sodere	Meteka	Tulu-Moye	Teo	Danab	Tendaho	Highlands 1	Total
No. of samples	86	26	18	12	12	12	7	7	4	4	4	2	2	1	197

3. Results and discussion

3.1. The correlation between the predictor and target variables

Correlation analysis shows the relationships between hydrogeochemical variables and geothermal downhole temperature (Fig. 5). Among the variables considered, SiO_2 has the strongest positive correlation with the downhole temperature ($r = 0.93$), followed by K ($r = 0.84$) and Na ($r = 0.67$). The noticeable relationship of SiO_2 with temperature is consistent with the sensitivity of dissolved silica concentrations to water-rock interactions and temperature-dependent dissolution processes in geothermal systems. In addition, the positive correlations of K and Na with temperature may reflect the influence of fluid-rock interactions and the evolution of geothermal fluids at elevated temperatures. Therefore, these strong relationships may indicate that these predictor variables may provide useful information for predicting geothermal downhole temperatures. In contrast, Mg ($r = -0.39$) and Ca ($r = -0.32$) exhibit relatively weak negative correlations with geothermal downhole temperature. Another interpretation is that the analysis highlights strong correlations among some predictor variables, particularly between K and SiO_2 ($r = 0.91$), indicating the presence of potential multicollinearity.

These characteristics were therefore considered in the selection of the machine-learning framework. Ridge regression provides a suitable regularized approach to reduce multicollinearity. In parallel, nonlinear algorithms, such as RF, XGBoost, CatBoost, and SVR, were considered since they can represent nonlinear and potentially interactive relationships between hydrogeochemical characteristics and geothermal downhole temperature. They also give the flexibility to capture more complex relationships between hydrogeochemical characteristics and downhole temperature.

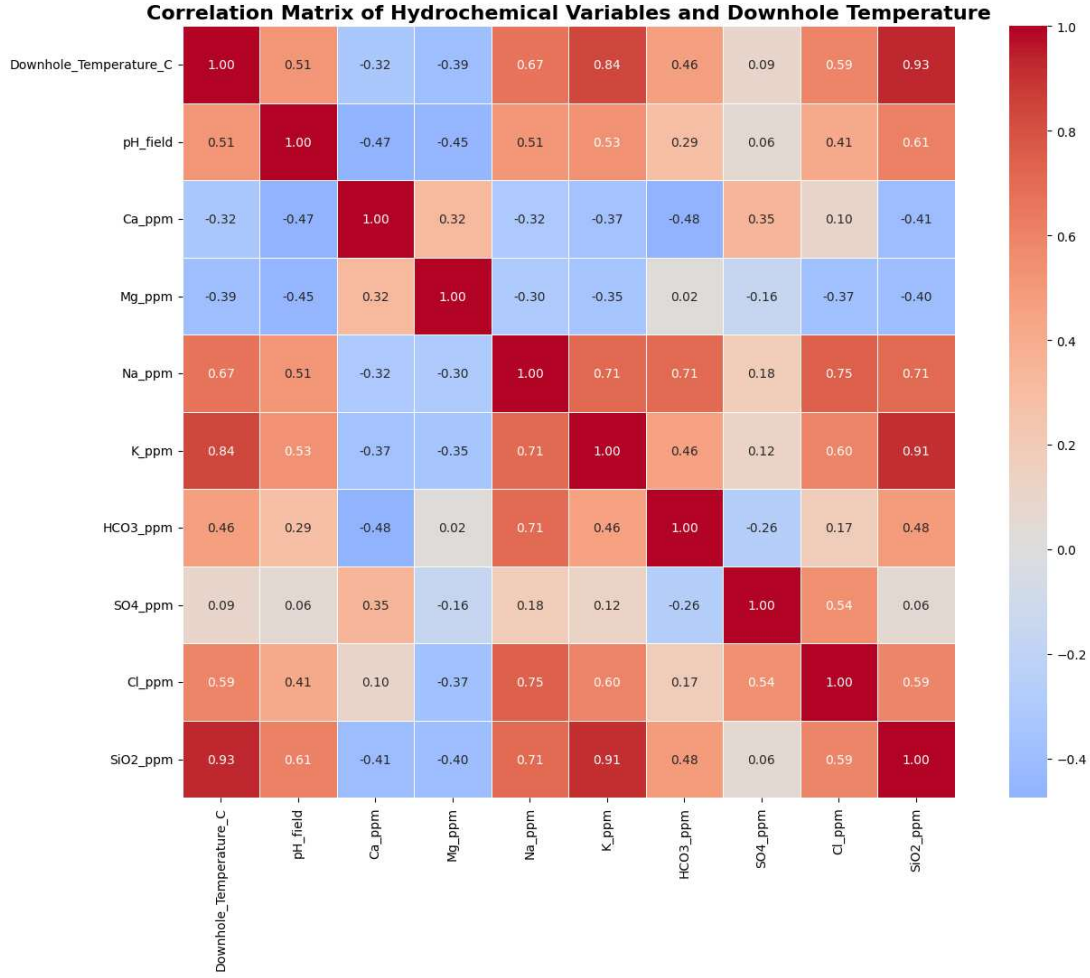


Fig. 5: Pearson's correlation matrix between predictor variables and the target variable.

3.2. Comparative Performance of the Developed ML Models

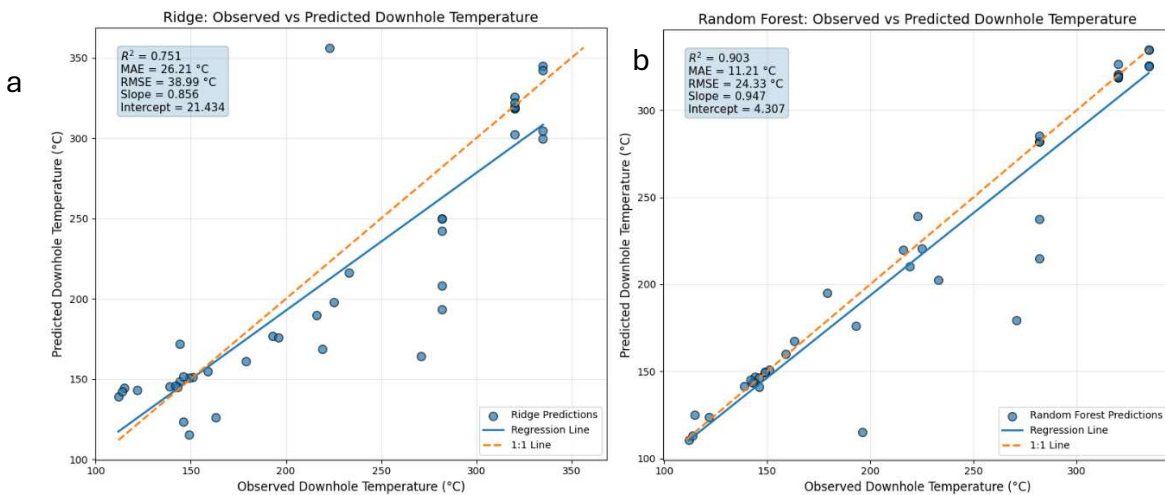
The nonlinear ensemble tree-based models showed a reliably stronger predictive capability than linear ridge regression for downhole temperature estimation (Fig. 6). RF and CatBoost achieved values of $R^2 = 0.903$ and $R^2 = 0.909$, respectively; see Fig. 6b and 6d, compared with $R^2 = 0.751$ for RR (Fig. 6a). This result shows that nonlinear relationships and interactions among the predictor variables make important contributions to the prediction of the downhole temperature. As indicated in Table 2, CatBoost produced the lowest RMSE (23.51 °C). In comparison, RF yielded the lowest MAE (11.21 °C), suggesting that CatBoost provides superior control of larger prediction errors, whereas RF provides slightly more accurate typical absolute predictions. XGBoost showed intermediate performance (RMSE = 29.78 °C) as shown in Fig. 6c, whereas the ridge regression model exhibited larger prediction errors.

However, the parity plots indicate some deviation from the 1:1 relationship, especially toward the upper-temperature range, suggesting possible systematic underprediction of extreme temperatures. This behavior confirms that further investigations using LOOCV residual diagnostics, bias analysis, and uncertainty quantification are warranted before concluding the overall behavior of the models.

Overall, the superior performance of CatBoost, RF, and XGBoost demonstrates that nonlinear relationships and interactions among hydrogeochemical variables are important controls on geothermal downhole temperature.

Table 2: Performance evaluation of the developed model on the test dataset

Model	Train-R ²	Test-R ²	MAE	RMSE
CatBoost	0.999	0.909	13.140	23.509
Random Forest	0.981	0.903	11.206	24.328
XGBoost	1.000	0.855	14.021	29.779
SVR	0.796	0.840	20.710	31.208
Ridge Regression	0.809	0.751	26.212	38.986



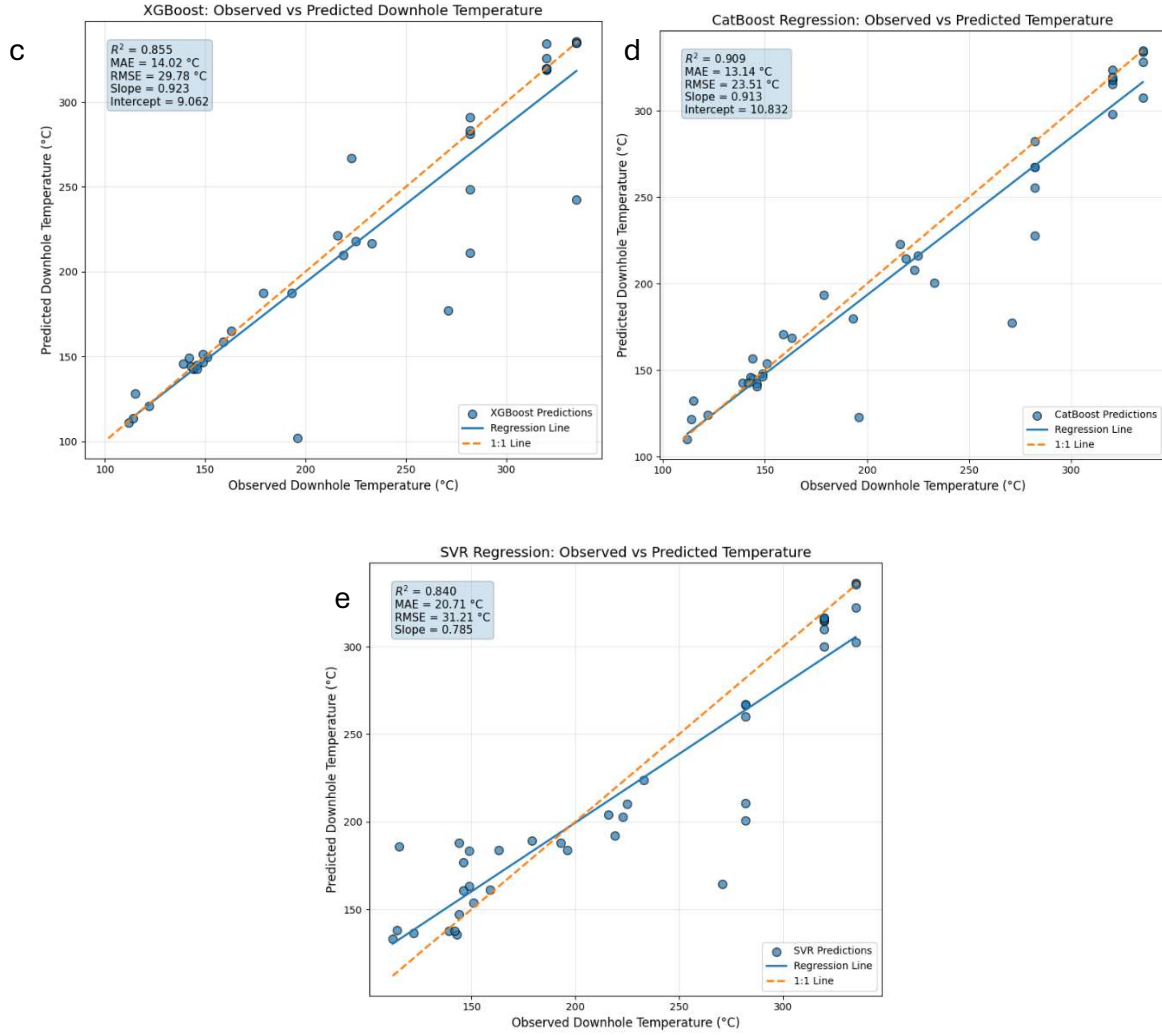


Fig. 6: Scatter plots of the (a) ridge regression, (b) random forest, (c) XGBoost, (d) CatBoost, and (e) SVR models for the actual and predicted downhole temperatures

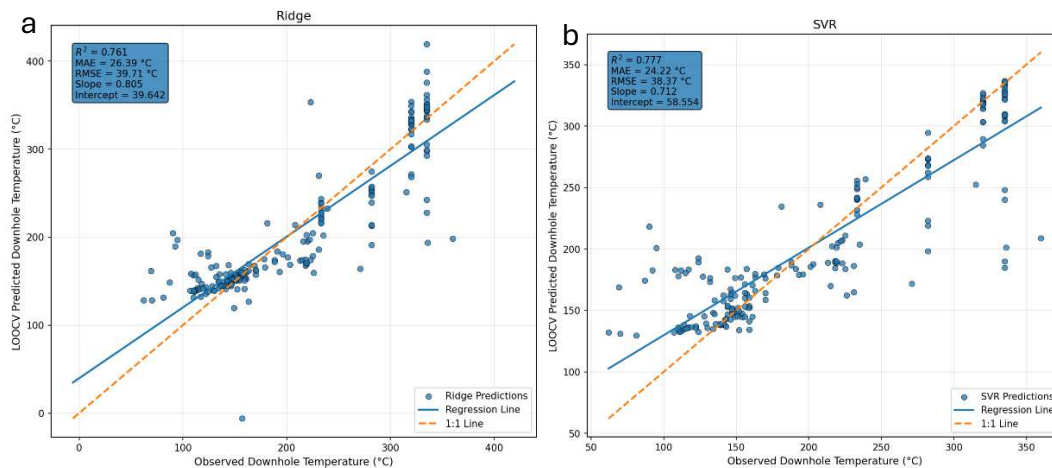
3.3. Leave-one-out cross-validation (LOOCV) results and model interpretation

The LOOCV results clearly demonstrate differences in the generalizability of the tested models, as illustrated in Fig.7. CatBoost, as shown in Fig.7e, achieved the highest predictive performance (RMSE = 26.52 °C, $R^2 = 0.894$), closely followed by random forest as shown in Fig. 7c (RMSE = 27.49 °C, $R^2 = 0.871$) and XGBoost ($R^2 = 0.842$) as shown in Fig. 7d whereas, ridge regression ($R^2 = 0.761$) and SVR ($R^2 = 0.777$) showed comparatively lower predictive accuracies, as illustrated in Fig. 7a, and 7b, respectively. The substantially improved performance of the ensemble models indicates that the downhole temperature exhibits nonlinear relationships and interactions among the controlling variables that are not adequately represented by a linear regression.

The regression slopes were less than 1 for all the models, particularly SVR (0.712), indicating prediction compression toward the mean and a tendency to underestimate higher and lower-temperature observations. The ensemble models reduced this effect (slopes ≈ 0.88), but some deviation may remain above 300 °C, suggesting reduced predictive reliability in the extreme-temperature range.

From a geothermal perspective, the superior performance of the ensemble models suggests that the downhole temperature is governed by nonlinear relationships and interactions among the predictor variables, which are not adequately represented via a linear regression

Overall, the LOOCV results support CatBoost, RF, and XGBoost as the most robust predictive approaches, and indicate that further improvement may require incorporating explicit geological controls such as lithology, fault or fracture density, alteration, depth, and hydrogeological characteristics.



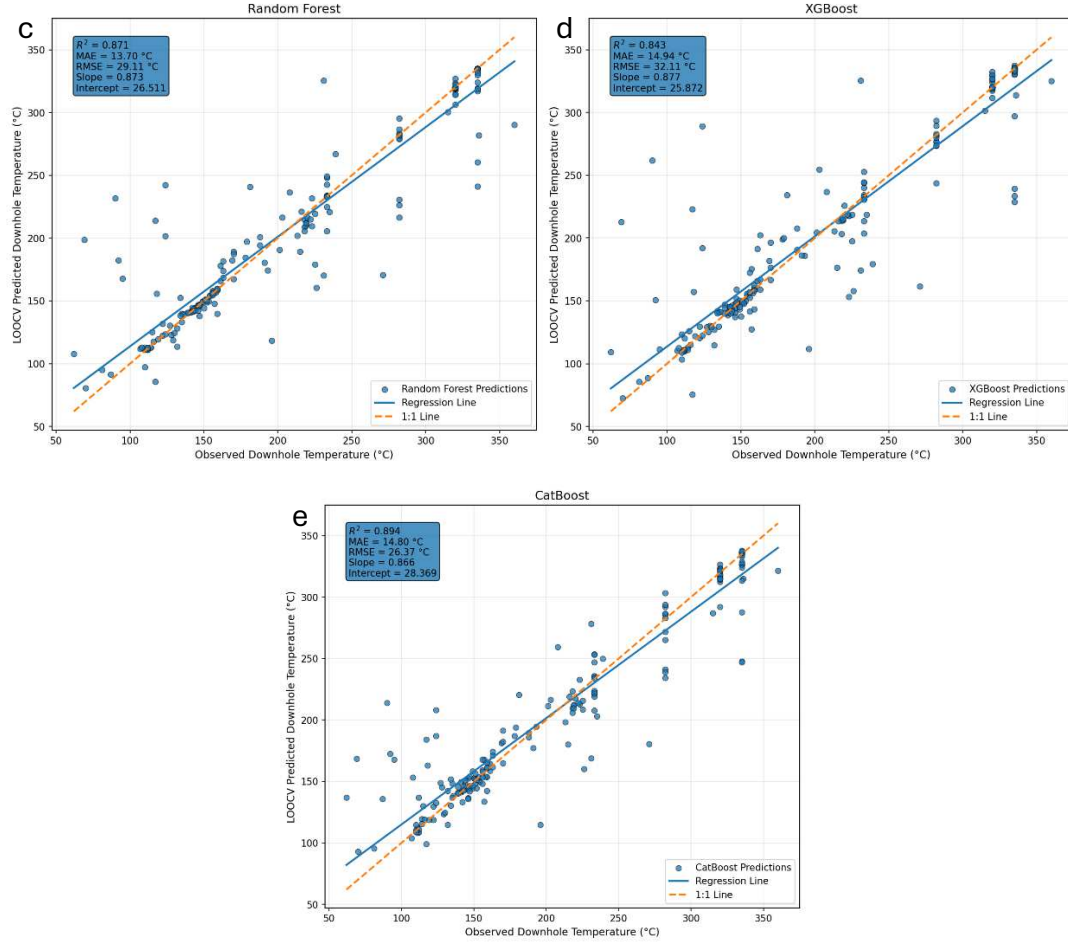


Fig. 7: LOOCV performance of the ML models for Observed vs. predicted downhole temperature prediction: (a) Ridge regression, (b) SVR, (c) RF, (d) XGBoost, (e) CatBoost.

LOOCV residual diagnostics indicate that the optimized CatBoost model generally provides well-centered predictions with limited systematic bias ($MBE = 1.68$ °C) and a good coefficient of determination ($R^2_{LOOCV} = 0.89$), as depicted in Fig. 8a. As shown in Fig. 8b, the locally weighted scatterplot smoothing (LOWESS) trend remains close to zero across most of the temperature range, demonstrating satisfactory calibration; however, as illustrated in Fig. 8c, the presence of several large positive and negative residuals and the asymmetric residual distribution indicate nonnegligible prediction uncertainty. The dispersion of residuals becomes more pronounced in the upper temperature range, suggesting potential heteroscedasticity and reduced predictive reliability under extreme geothermal conditions. Therefore, while the model demonstrates robust overall generalizability, uncertainty is not uniform across the temperature domain and should be considered when applying the model to high-temperature geothermal prediction

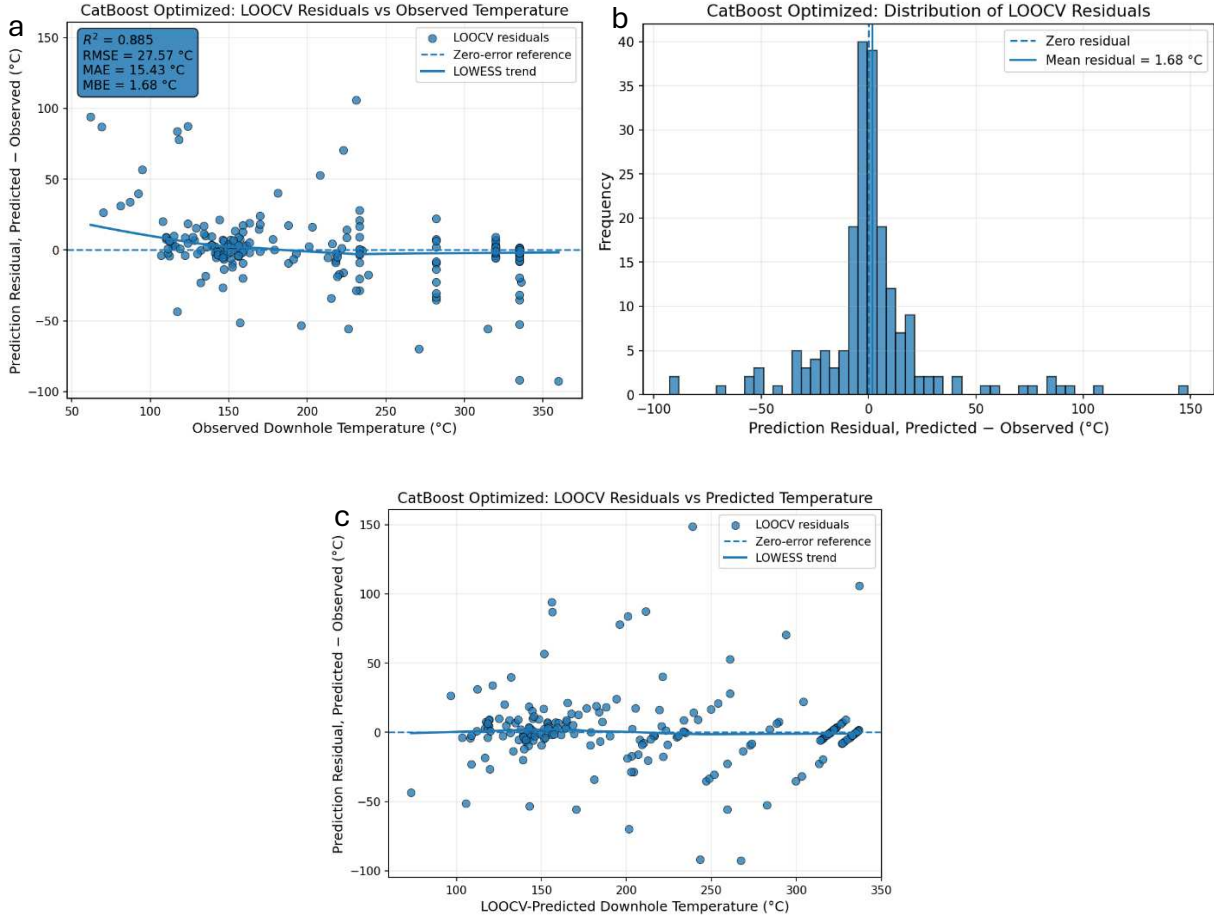


Fig. 8: LOOCV diagnostic: (a) LOOCV residual vs observed temperature (b) distribution of LOOCV residuals (c) LOOCV residuals vs predicted temperature

SHAP (Shapley additive explanations) analysis, as illustrated in Fig. 9, indicates that downhole temperature prediction is primarily controlled by K and SiO₂, followed by Mg, Ca, Cl, HCO₃, Na, SO₄, and PH. A similar study by (Ibrahim et al., 2022) identified SiO₂ as the most crucial predictor variable, followed by K. On the one hand, the strong positive contributions of elevated K and SiO₂ are consistent with enhanced water-rock interaction, silicate mineral dissolution, and progressive geothermal-fluid equilibration at elevated temperatures. On the other hand, the negative contribution associated with higher Mg is consistent with Mg depletion during hydrothermal alteration and the evolution of mature geothermal fluids. Ca and Na exhibit weaker, dataset-dependent negative associations; this may reflect coupled mineral dissolution, ion exchange, lithological effects, and multivariate interactions. Elevated Cl and HCO₃ contribute positively to temperature prediction, which may suggest that fluid maturity, deep circulation, and CO₂-related

hydrogeochemical processes are also captured by the model. In general, the SHAP results demonstrate that the CatBoost model doesn't rely solely on individual chemical concentrations but rather captures an integrated hydrogeochemical signature of geothermal fluid evolution and reservoir conditions. However, SHAP values represent predictive associations rather than causal relationships, and their geological interpretation should therefore be constrained by the regional lithology, alteration assemblages, and known hydrothermal-fluid pathways.

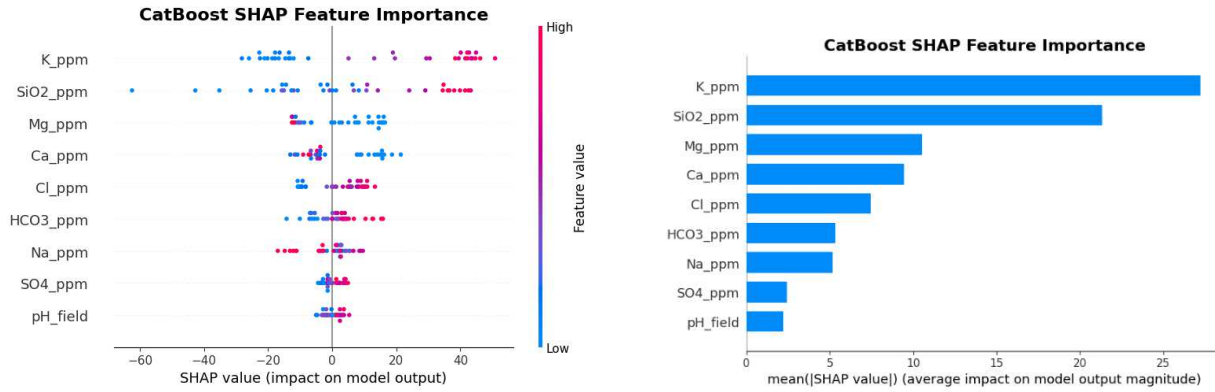


Fig. 9: SHAP value analysis of the best model (blue color=low feature value, pink & red =high feature value)
(CatBoost)

4. Conclusion

The combined test-set and LOOCV results demonstrate that nonlinear ensemble learning provides a more reliable representation of downhole temperature variability than linear ridge regression does. For the independent test data, CatBoost achieved the highest predictive performance ($R^2=0.909$), with an RMSE = 23.51 °C, closely followed by random forest ($R^2=0.903$), with an RMSE = 24.33 °C. The same overall behavior was retained under LOOCV, where CatBoost achieved $R^2_{LOOCV}=0.893$ and RMSE = 26.52 °C, whereas random forest achieved $R^2_{LOOCV}=0.885$ and RMSE = 27.49 °C. This consistency between independent testing and cross-validation provides evidence that the predictive performance is not solely attributable to overfitting.

The substantially lower performance of ridge regression ($R^2 = 0.751$ on the test data and 0.777 under LOOCV) indicates that the downhole temperature is governed by nonlinear and interacting controls that cannot be adequately represented by a linear relationship. The relatively similar performances of CatBoost and RF additionally suggest that the dominant predictive relationships are robust to different ensemble-learning architectures.

Despite the strong overall model performance, both the test results and LOOCV predictions demonstrate some deviation from the 1:1 relationship, predominantly at the higher-temperature range of the downhole. From a geological perspective, such behavior may indicate unresolved subsurface heterogeneity, such as variations in lithology, permeability, fracture-controlled fluid circulation, and localized convective heat transfer. These processes can produce thermal anomalies that are difficult to infer solely from the available predictor variables.

In general, the CatBoost machine learning method can be considered the most suitable final predictive model on the basis of having the highest R^2 and lowest RMSE across both evaluation strategies, while random forest provides the lowest MAE and therefore remains a highly competitive alternative for typical absolute prediction accuracy. The agreement between the test-set and LOOCV performance supports the generalizability and robustness of the ensemble models. However, the remaining errors should be investigated through SHAP analysis and integration with geological information to determine whether they represent random model uncertainty or systematic signatures of unresolved geothermal structure. This integration is essential for moving from statistical temperature prediction toward physically interpretable geothermal characterization.

CRedit authorship contribution statement:

Mohammed Ali Belay: Writing-original draft, Conceptualization, Visualization, Data curation, Methodology, Software, Writing-review & editing. **Elias Mamushet:** Writing-review & editing, Supervision, Methodology, Conceptualization. **Assnake Bekele:** Writing- review & editing, Visualization, Supervision. **Tarekegn Tadesse:** Writing-review & editing, Validation

Declaration of competing interests:

The authors declare that there are no known competing financial interests or personal relationships that could influence the work reported in this paper.

Data availability:

Some or all the data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

Data source:

Montcoudiol, N. (Creator), Burnside, N. (Supervisor), Becker, K. (Contributor), Lewi, E. (Contributor) (28 Jul 2021). Ethiopia ground and surface water quality review database. University of Strathclyde. Database_Ethiopia_Geothermal_final_WIRES_Water(.xlsx). 10.15129/68b5e36e-fc2f-47b8-b0a0-88ab6f9f0a21

References

1. AlGaïar, M., Hossain, M., Petrovski, A., Lashin, A., Faisal, N., 2024. Applications of artificial intelligence in geothermal resource exploration: A review. *Deep Undergr. Sci. Eng.* 3, 269–285. <https://doi.org/10.1002/dug2.12122>
2. Benti, N.E., Woldegiyorgis, T.A., Geffe, C.A., Gurmesa, G.S., Chaka, M.D., Mekonnen, Y.S., 2023. Overview of geothermal resources utilization in Ethiopia: Potentials, opportunities, and challenges. *Sci. Afr.* 19, e01562. <https://doi.org/10.1016/j.sciaf.2023.e01562>
3. Burnside, N., Montcoudiol, N., Becker, K., Lewi, E., 2021. Geothermal energy resources in Ethiopia: Status review and insights from hydrochemistry of surface and groundwaters. *WIREs Water* 8, e1554. <https://doi.org/10.1002/wat2.1554>
4. Genetu, M., 2025. Assessment of Geothermal Energy in Aluto-Langano and Tulu-Moye, Central Ethiopia. *Geol. J.* 60, 64–72. <https://doi.org/10.1002/gj.5073>
5. Ibrahim, B., Konduah, J.O., Ahenkorah, I., 2022. Predicting Reservoir Temperature of Geothermal Systems in West Anatolia, Turkey: A Focus on Predictive Performance and Explainability of Machine Learning Models. *SSRN Electron. J.* <https://doi.org/10.2139/ssrn.4239319>
6. Samrock, F., Grayver, A., Cherkose, B., Kuvshinov, A., Saar, M.O., n.d. Aluto-Langano Geothermal Field, Ethiopia: Complete Image of Underlying magmatic-hydrothermal System Revealed by Revised Interpretation of Magnetotelluric Data.
7. Shoeibi Omrani, P., Poort, J., Shahmohammadi, S., 2025. Artificial intelligence in the geothermal energy systems, in: *Geothermal Energy Engineering*. Elsevier, pp. 349–377. <https://doi.org/10.1016/B978-0-443-21662-6.00004-9>
8. Zhang, Q., Gou, L., Xu, L., 2026. Machine Learning for Geothermal Energy Systems: Prediction, Optimization, and Physics-Informed Hybrid Methods. *Energies* 19, 3193. <https://doi.org/10.3390/en19133193>

9. Tadesse, F. (2018). Ethiopia: Aluto-Langano geothermal expansion comes to life. Addis Fortune. Retrieved from <https://addisfortune.net/articles/aluto-langano-geothermal-expansion-comes-to-life/>
10. Kebede, S., Woldemariam, F., & Kassa, T. (2020). Status of geothermal exploration and development in Ethiopia. Paper Presented at the World Geothermal Congress 2020+1, Reykjavík, Iceland.
11. Zhu, Y., 2024. Leveraging Machine Learning for Subsurface Geothermal Energy Development. Highlights Sci. Eng. Technol. 121, 440–449.
<https://doi.org/10.54097/j8tjym72>
12. Chandrasekharam, D., Lashin, A., Al Arifi, N., Al Bassam, A. A., & Varun, C. (2015). Evolution of geothermal systems around the Red Sea. Environmental Earth Sciences, 73(8), 4215–4236. <https://doi.org/10.1007/s12665-014-3710-y>
13. Zhang, L.; Liang, X.; Yang, W.; Jia, Z.; Xiao, C.; Zhang, J.; Dai, R.; Feng, B.; Fang, Z. Identification of the formation temperature field by explainable artificial intelligence: A case study of Songyuan City, China. Energy 2025, 319, 135172
<https://doi.org/10.1016/j.energy.2025.135172>
14. Lundberg, S. M., and Lee, S.-I. (2017). "A unified approach to interpreting model predictions." 662 Advances in neural information processing systems, 30.