PREDICTING SUBSURFACE TEMPERATURE FROM WELL LOGS USING MACHINE LEARNING

BY

ABIOLA MARY YETUNDE
ENG2006408

DEPARTMENT OF PETROLEUM ENGINEERING
FACULTY OF ENGINEERING
UNIVERSITY OF BENIN
BENIN CITY

NOVEMBER, 2025

PREDICTING SUBSURFACE TEMPERATURE FROM WELL LOGS USING MACHINE LEARNING

BY

ABIOLA MARY YETUNDE
ENG2006408

A PROJECT SUBMITTED TO THE DEPARTMENT OF PETROLEUM ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF BACHELOR OF ENGINEERING (B.ENG) DEGREE IN PETROLEUM ENGINEERING, UNIVERSITY OF BENIN

NOVEMBER, 2025

CERTIFICATION

This is to certify that this project work titled "**PREDICTING SUBSURFACE TEMPERATURE FROM WELL LOGS USING MACHINE LEARNING"** was carried out by **ABIOLA MARY YETUNDE**, a student of the Department of Petroleum Engineering, Faculty of Engineering, University of Benin, in partial fulfillment of the requirements for the award of the Bachelor of Engineering (B.Eng.) degree in Petroleum Engineering, under the supervision of;

_____ _____

ENGR. IKPONMWOSA U.B **DATE**

(PROJECT SUPERVISOR)

_____ _____

DR. OLUWASEUN TAIWO **DATE**

(PROJECT COORDINATOR)

_____ _____

DR. IKPONMWOSA OHENHEN **DATE**

(HEAD OF DEPARTMENT)

DEDICATION

This Project is dedicated to the Almighty God for his wonderful and infinite mercy,love,grace,wisdom, strength and faithfulness.Also to lovely parents and siblings for their financial and Moral support.

ACKNOWLEDGEMENT

I am profoundly grateful to Almighty God for His abundant grace, wisdom, and strength that have guided me throughout my academic journey and made the successful completion of this project possible.

My sincere appreciation goes to the Head of Department, Dr. IKPONMWOSA OHENHEN, Head of the Department of Petroleum Engineering, University of Benin, for his exemplary leadership, commitment to academic excellence, and continuous support toward the success of students within the department.

I extend my deepest gratitude to my project supervisor, Engr. IKPONMWOSA U.B, for his invaluable guidance, patience, and constructive criticism, which greatly enhanced the quality and direction of this research work.

My sincere thanks also go to Engr.EDOHO C.F for his supervision, guidance, and useful suggestions, which contributed greatly to the success of this project. His support and encouragement were truly appreciated.

I remain ever thankful to my amiable parents, Mr and Mrs Abiola and to my siblings Temitope Abiola, Opeyemi Abiola for their steadfast love, prayers, and sacrifices, which have been the backbone of my achievements.

To my dear friends Rita, Evans, Nifemi, Joshua, Hannah, Goodness,Victory and Ben, I deeply appreciate your friendship, motivation, and support throughout this journey. Your presence made this experience both meaningful and memorable.

Finally, I extend my gratitude to the Department of Petroleum Engineering, University of Benin, for providing a solid academic foundation and an enabling environment for learning and research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

ABSTRACT

Predicting the subsurface temperature distribution within sedimentary and petroleum-bearing formations is essential for accurate hydrocarbon maturation modeling, well-bore stability, and drilling-fluid design. Traditional approaches—relying on sparse bottom-hole temperature (BHT) measurements and one-dimensional conductive models—often misestimate true formation temperatures by 5–10 °C in heterogeneous settings such as the Niger Delta. To address these limitations, this study develops a data-driven workflow that employs machine learning algorithms trained on routinely acquired drilling logs to produce continuous, high-resolution temperature profiles. First, we assembled a dataset from Niger Delta wells comprising wireline logs (gamma-ray, four-pad resistivities, density, neutron porosity, sonic velocity), drilling parameters (equivalent circulating density, rate of penetration, exposure time), and corrected BHT readings. After replacing sentinel values (–9999) with NaNs and depth-referencing all curves to true vertical depth (TVD), each log was clipped to its 1st–99th percentile range to mitigate extreme outliers. Features were standardized to zero mean and unit variance. Derived attributes—such as depth-derivatives and moving-window averages—were also explored to enhance lithofacies and thermal signal detection. We compared three regression models: a multilayer perceptron neural network (ANN), a Random Forest (RF) ensemble, and Support Vector Regression (SVR). Hyperparameters were tuned via grid search with k-fold cross-validation, and models were evaluated on a hold-out subset (20 % of depths). Error metrics (RMSE, MAE, R^2) and well-log–style scatter and depth-track plots quantified predictive performance and bias. The ANN achieved an RMSE of 0.16 °F and $R^2 = 0.97$, producing smooth temperature gradients. The RF delivered an RMSE of 0.25 °F and $R^2 = 0.985$, with feature-importance analysis highlighting mid-pad resistivities and drilling parameters as primary predictors. SVR, with an RMSE of 0.47 °F and $R^2 = 0.955$, was less competitive but still captured over 95 % of temperature variance. Well-log plots demonstrated that both ANN and RF closely track actual temperature profiles across lithologic transitions. This end-to-end pipeline—from data cleaning and feature engineering to model training, validation, and interpretation—demonstrates that machine learning can offer accurate, cost-effective alternatives to physics-based thermal models. Future work will explore hybrid physics–ML models, temporal drilling-data integration, expanded feature fusion (e.g; seismic attributes), and explainable-AI techniques, with the goal of operationalizing these tools in real-time drilling workflows.

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Understanding how temperature varies beneath the Earth's surface is fundamental to a wide array of industrial and scientific endeavors. In geothermal energy exploration, for example, the viability of a project hinges on the subsurface temperature gradient: higher temperatures at accessible depths translate directly into more efficient power generation and lower operational costs. Similarly, in hydrocarbon reservoir management, accurate knowledge of thermal profiles informs predictions of fluid viscosity, phase behavior, and saturation, thereby guiding well design, production forecasting, and safety protocols. Even beyond energy applications, projects such as carbon sequestration, underground waste disposal, and deep mining operations depend critically on reliable thermal data to anticipate rock mechanics, choose appropriate materials, and mitigate thermal-induced stresses that can compromise borehole integrity.

Historically, the most direct method for ascertaining subsurface temperatures has been through borehole measurements. Instruments like thermistors or fiber-optic distributed temperature sensing (DTS) systems are lowered into completed wells, recording temperature at discrete depth intervals. These measurements are highly accurate at their specific points but are inherently sparse, especially in frontier regions where drilling activities are limited. Acquiring additional measurements can be prohibitively expensive, and in many basins, temperature data remain so patchy that any interpolation between boreholes risks smoothing over critical heterogeneities in the geological framework.

To fill these data gaps, geostatistical interpolation techniques such as kriging have long been employed. By treating temperature as a spatially correlated variable, these methods can produce

continuous temperature fields between scattered measurement locations. Although geostatistics can work well when borehole spacing is moderate, they often struggle to accommodate abrupt changes in lithology, thermal conductivity, or fluid pressure. For instance, a rapid transition from sandstone to shale can dramatically alter the local thermal gradient, but if there are no nearby data points to capture this shift, interpolation may yield misleadingly smooth—and thus inaccurate—thermal models.

Physics-based numerical simulations represent another established approach, wherein heat-flow equations (including both conduction and convection terms) are solved across a discretized subsurface domain. Such models require detailed inputs: boundary conditions, thermal properties of rocks, groundwater flow rates, and fluid pressures. In theory, a well-calibrated heat-flow simulation can capture the complexity of conductive and advective heat transfer, but in practice, uncertainties in rock thermal conductivity, porosity, and fluid permeability introduce significant modeling errors. Moreover, running high-resolution simulations over basin-scale domains is computationally expensive, often prohibiting iterative scenario testing or real-time decision making.

Against this backdrop, drilling logs have begun to emerge as an underutilized yet potentially transformative data source for subsurface temperature prediction. Unlike traditional measurements that capture temperature directly at specific depths, drilling logs are acquired continuously during the drilling process, recording parameters such as rate of penetration, torque, drilling fluid temperature, and mud weight, among others. These parameters carry implicit signatures of the thermal environment: as drill bit temperatures rise, bit wear rates can accelerate, leading to variations in rate of penetration; elevated downhole temperatures influence drilling fluid viscosity, which in turn affects torque and drag measurements; and the returning mud

temperature can reflect the thermal gradient encountered by the circulating fluid. In essence, drilling logs encode a blend of mechanical, chemical, and thermal information about the subsurface strata being penetrated.

Recent advances in machine learning (ML) offer an exciting pathway to unlock these hidden thermal signals within drilling data. By training ML models on historical wells where both drilling logs and direct temperature measurements are available, it becomes possible to learn complex nonlinear relationships between drilling parameters and subsurface temperatures. Once trained, these models can predict detailed temperature profiles in new wells—often in real time—as drilling proceeds. This approach promises several advantages over conventional methods: it requires no additional instrumentation beyond what is already installed on a drilling rig, leverages existing data that are routinely archived, and can adapt to local geological variations without explicit specification of thermal properties or boundary conditions.

Furthermore, machine learning techniques excel at handling large, heterogeneous datasets and can incorporate additional inputs such as lithology, mud composition, and drilling fluid flow rates. These auxiliary variables help ML algorithms disentangle the purely thermal effects from confounding factors like changes in lithology or drilling practices. By integrating drilling logs into an ML framework, we can construct high-resolution temperature models that honor the subtle signatures of local geology and drilling dynamics, even in regions where traditional temperature measurements are scarce or absent.

Predicting subsurface temperature distributions through machine learning on drilling logs addresses a critical gap in current geoscientific practice. It offers a cost-effective, scalable, and potentially more accurate alternative to purely measurement-based, geostatistical, or physics-based methods. By harnessing the implicit thermal information embedded in drilling parameters,

this approach can enhance decision-making in geothermal exploration, hydrocarbon production, and other subsurface engineering disciplines,ultimately improving operational efficiency and reducing risk.

1.1.1 Historical context of subsurface temperature modeling

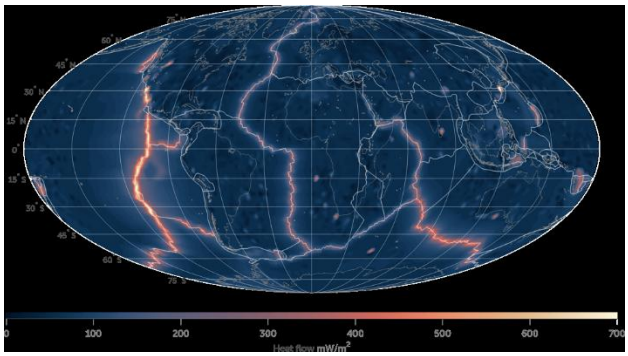


Figure 1.1: Global geothermal ; Davies, J. H. (2013)Global map of solid Earth surface heat flow. Geochemistry, Geophysics, Geosystems.

Figure 1.1: illustrates a global surface heat-flow map, which really drives home how Earth's internal heat shapes what's happening underground. On average, if you stand on continental crust you're looking at a geothermal gradient of roughly 25–30 °C per kilometer—equivalent to about 60 mW/m² of heat escaping at the surface. This gradient isn't just a number on a map: it's the backbone of how we understand subsurface temperatures everywhere from geothermal fields to deep oil basins.

Back in the mid-20th century, petroleum geologists were already keenly aware that you couldn't unlock the secrets of a sedimentary basin without knowing its thermal regime. Temperature governs how buried organic matter turns into oil and gas, and once those fluids are there, their properties—like viscosity and mobility—are all temperature-dependent. Early explorers didn't have fancy sensors or 3D models at their fingertips; they made do with bottom-hole temperature (BHT) readings taken soon after a well was drilled. Unfortunately, those BHT measurements

tend to come in a bit cooler than reality because the circulating— or sometimes stagnant—mud actually chills the formation as it flows. And with only two or three temperature readings per well, the picture was, at best, spotty.

To fill in the blanks, simple conductive heat-flow models became the go-to tool. Engineers and geoscientists would stack the subsurface into a handful of one-dimensional layers, assume steady-state heat conduction, and apply Fourier's law: if you know the heat flow (Q), layer thickness (L), and thermal conductivity (K), then the temperature jump (ΔT) across that slice is $\Delta T = Q \cdot L/K$. Add up those jumps from the surface downward, and you get a continuous temperature curve—one that, while simplified, often did a decent job of filling in where BHT data were missing.

Even with that approach, though, the early models had to contend with the "cooling effect" of the borehole itself. That's where Horner (1951) and other pioneers came in: they developed correction methods—what we now call Horner plots—to extrapolate true formation temperatures from transient BHT measurements. By charting temperature versus the logarithm of time since circulation stopped, they could back out what the undisturbed formation temperature must have been. Those classical heat-flow equations and Horner-style corrections laid down the thermal "basics" of petroleum system analysis, and they remain a foundational touchstone even now that we're blending physics with machine learning to push temperature prediction to the next level.

By the late 20th century, basin-scale thermal models had grown considerably more sophisticated, moving beyond simple one-dimensional, steady-state assumptions. Researchers began to incorporate the entire sedimentary history of a basin—layer after layer of deposition, compaction, and uplift—into multi-layer numerical simulations. They also started accounting for radiogenic heat generated within sediments, which can be especially important in areas with thick,

radiogenic-rich shale sequences. Despite these advances, many traditional models still rested on a few simplifying assumptions: steady-state heat flow, published values for thermal conductivity, and one-dimensional vertical heat transfer. In reality, however, the thermal regime is far more complex.

For example, Majorowicz et al. (2005) highlighted that lateral variations in heat flow often stem from subsurface fluid movement—fluids migrating through permeable sand channels or faults can carry heat laterally, creating thermal anomalies that a purely conductive, vertical model would miss. Variations in rock type also play a critical role: coarse-grained sandstones conduct heat more efficiently than finer-grained, shale-rich intervals, which tend to trap heat and produce higher geothermal gradients. Crustal heat production, influenced by the distribution of heat-generating minerals, and variable basal heat flux from the mantle likewise disrupt the neat vertical temperature increase assumed in simpler models.

Field studies in the Niger Delta provide an excellent case study of these complexities in action. Early work by Nwachukwu (1976) and Avbovbo (1978) mapped out a low-heat-flow, low-gradient core beneath the central basin. Surrounding that core, toward the basin margins, they observed markedly higher geothermal gradients. Later surveys, such as Akpabio et al. (2013), confirmed this pattern: sandy, coarse-grained sections—particularly those in the deeper, more permeable portions of the basin—exhibited low gradients, while shale-rich, less permeable intervals tended to host steeper temperature increases. This differentiation makes sense when you consider sandstone's higher thermal conductivity: heat moves through it more readily, flattening the gradient. In contrast, shales act as insulators, causing heat to build up and creating a sharper temperature climb. These observations underscored two important points: (1) bottom-hole

temperature (BHT) data must be corrected for cooling effects caused by circulating mud, and (2) any meaningful thermal map has to account for changing lithology at different depths.

Recognizing these challenges, the oil and gas industry began to invest in logging technology designed to capture temperature more accurately. Instead of relying exclusively on brief, point-wise BHT readings, operators started using continuously recorded temperature logs that required a longer shut-in period. Once the well is allowed to equilibrate—often several hours after the pumps stop—the fluid in the borehole approaches true formation temperature. These longer-duration logs reveal a much richer, more continuous picture of the geothermal gradient than traditional BHT snapshots ever could. By the early 2000s, this approach had become more widespread, enabling refined basin models that marry stratigraphy, fluid flow, and radiogenic heat generation.

1.2 Significance Of Accurate Subsurface Temperature Prediction

Knowing exactly how temperature varies with depth in a sedimentary basin is not just a “nice-to-have” feature—it’s absolutely central to almost every stage of hydrocarbon exploration and production. At the heart of the matter is thermal maturity, which dictates when and where organic kerogen in source rocks has been cooked sufficiently to enter the oil or gas “windows.” In general, oil generation kicks in when temperatures reach roughly 60–120 °C, while gas generation often accelerates at temperatures up to about 200 °C. If your thermal model is off by merely 10–20 °C, you could be misplacing the entire oil window by hundreds of meters. Suddenly, prospects that looked promising on a map might actually be too “cold” to generate oil, or worse, too “hot,” meaning you’ll find gas instead of oil.

Miscalculating thermal maturity has downstream effects at every step. If you think a trap was charged with oil based on a misestimated temperature history, you’ll drill up a dry hole—or

worse, accidentally target a zone that's already migrated its hydrocarbons elsewhere. Conversely, underestimating thermal maturity might make you ignore a sweet spot where hydrocarbons are ripe for production. Even once a well is producing, temperature controls fluid viscosity: oil becomes much more viscous at lower temperatures, which can slow flow rates and drive up pumping costs. In deep or high-temperature wells, equipment and drilling fluids need to be designed to withstand harsh conditions; if you don't know the true subsurface temperature, you can end up with failed pumps, stuck drill strings, or degraded drilling fluid polymers.

Beyond petroleum, accurate temperature profiles are equally critical in geothermal energy development. When designing a geothermal power plant, you need to know exactly where the steam-rich zone is located, how thick it is, and what pressure you can expect at those temperatures. Off by just a few degrees? You could overestimate the thermal resource, leading to underperforming wells and an uneconomical project. Temperature data also feed directly into heat-flow models that predict how quickly a geothermal reservoir will cool over time, influencing well spacing, reservoir stimulation strategies, and long-term sustainability forecasts.

Finally, in emerging applications like carbon capture and storage (CCS), accurate subsurface temperatures inform how injected CO_2 behaves underground. Temperature influences CO_2 density, viscosity, and chemical interactions with the host rock. If these thermal effects aren't modeled correctly, you risk underestimating CO_2 plume migration, trapping efficiency, and potential leakage pathways. The same goes for underground waste disposal: high temperatures can alter rock stresses, potentially inducing fractures that compromise containment.

In short, getting the thermal picture right is fundamental to making smart decisions in a host of subsurface engineering disciplines. Accurate temperature predictions mitigate risk, drive cost efficiencies, and unlock new exploration and production opportunities. By integrating machine

learning with readily available drilling logs, we can push temperature modeling beyond the limitations of traditional methods, offering real-time, high-resolution insights that help operators everywhere—from the Niger Delta to the geothermal fields of Iceland—reach their subsurface objectives more effectively.

Beyond petroleum system modeling, understanding downhole temperature profiles is vital for both drilling safety and the bottom line. For instance, engineers must choose appropriate mud weights and design casing programs with anticipated bottom-hole temperatures in mind. In high-temperature zones, drilling mud can thin out, leading to mud loss, while in overpressured shales heated past their brittle-ductile transition, borehole walls may collapse. In thermal recovery operations—think steam-assisted gravity drainage (SAGD) in heavy oil fields—knowing the native temperature distribution before steam injection is critical. Injecting steam into zones that are cooler than expected can cause excessive heat loss into surrounding rocks, reducing efficiency and ramping up costs. Moreover, rapid temperature changes during drilling or production induce thermal stresses in the wellbore. If the formation cools or heats too quickly, it can crack or fail, placing extra loads on casing and potentially triggering fractures. In each of these scenarios, even a few degrees' miscalculation can translate into large errors in predicted fluid viscosity, phase behavior, or rock strength. In other words, reliable thermal models aren't just academic—they're essential tools for safe, cost-effective reservoir engineering.

1.3 Conventional Approaches And Their Limitations

1.3.1 Bottom-Hole Temperature Corrections

For decades, the primary data source for subsurface temperature has been bottom-hole temperature (BHT) readings taken immediately after drilling mud circulation ceases. Unfortunately, these measurements are almost always biased low—sometimes by as much as 15–

20 °C—because the still-circulating, cooler drilling fluid continues to chill the borehole walls. In 1951, Horner introduced a method to correct this bias: by plotting recorded BHTs against the logarithm of shut-in time (now called Horner plots), one could extrapolate back to a true static formation temperature. Although Horner-style corrections remain standard practice, they rest on several simplifying assumptions. They treat the subsurface as a homogeneous, one-dimensional medium dominated by pure conduction, ignoring the fact that drilling fluid can invade pores and carry heat away (advective cooling). They also assume thermal properties don't change with depth, even though a shift from sandstone to shale—or variations in porosity—can dramatically alter conductivity. As a result, while Horner corrections can help, they often fall short in real-world geology, especially in heterogeneous basins where advection or variable rock properties play a significant role.

Figure 1.2: Geological Map in Niger Delta

Figure 1.2 shows a geological map of the Niger Delta, highlighting the complex interplay of sedimentary facies, fault systems, and varying lithologies that characterize the basin. Because of these complexities, purely conductive, one-dimensional models often struggle to match observed temperature profiles or even corrected BHT measurements. In areas where data points are sparse, rock types shift rapidly, or subsurface fluids circulate laterally, traditional heat-flow approaches have limited predictive power. In such settings, errors of just a few degrees can cascade into significant missteps—whether that means misplacing a hydrocarbon pay zone, underestimating drilling risks, or failing to optimize a thermal recovery operation. These shortcomings motivate the search for data-driven alternatives, such as machine learning, which can learn complex, non-linear relationships directly from well logs and other geophysical measurements.

1.4 Machine Learning Approaches For Temperature Prediction

Over the last decade, machine learning (ML) has emerged as a powerful complement—or, in some cases, an alternative—to physics-based models in subsurface applications. Rather than explicitly prescribing how heat should flow through each stratigraphic layer, ML algorithms can learn to map high-dimensional input features (for example, raw drilling log curves) to target variables (such as formation temperature). This means they can implicitly account for phenomena like lateral fluid movement, shifting lithologies, or even the thermal effects of circulating mud, all without the need to specify detailed physical parameters.

One of the biggest advantages of ML is its ability to capture subtle, non-linear relationships that might elude a conventional modeler. For instance, a sudden uptick in rate of penetration (ROP) at a certain depth could signal a change to a rock type with higher thermal conductivity—something that might only be approximated in a physics model. Likewise, variations in downhole torque or bit weight might indirectly reflect temperature-dependent changes in drilling

fluid viscosity. When an ML model is trained on wells where both extensive drilling logs and reliable temperature measurements (e.g., from long-duration temperature logs) are available, it can learn these hidden signals. Once trained, such a model can then predict temperature profiles in new wells in real time, even before direct measurements are taken.

Several studies have demonstrated the promise of ML for subsurface temperature estimation. Shahriari et al. (2018) showed that neural networks could outperform simple conductive models in basins with complex structural settings by leveraging a broader suite of input features—such as mud density, bit RPM, and lithology logs—to predict bottom-hole temperatures with higher accuracy. Emam et al. (2020) extended this concept using ensemble techniques (like random forests and gradient boosting machines), which combine the predictions of many weak learners to reduce overfitting and improve generalization in areas where data are scarce. These models don't require detailed knowledge of thermal conductivity or radiogenic heat production; instead, they learn directly from historical data how those factors manifest indirectly in drilling parameters.

A typical ML workflow for temperature prediction begins with data gathering and preprocessing. Drilling logs—ROP, torque, standpipe pressure, mud weight, and mud temperature—are synchronized depth-wise and often undergo cleaning steps to remove outliers (for example, when the bit is off bottom). Lithology or gamma-ray logs can be added as categorical or continuous features to help the model distinguish between sandstone, shale, or mixed intervals. Where available, long-duration temperature logs (taken after sufficient shut-in) serve as the "ground truth" for training.

Next comes feature engineering. Rather than simply feeding the raw curves into a model, it's common to compute derived features—moving averages, gradients over short depth windows, or

ratios like ROP/mud weight—to emphasize signals likely related to heat transfer. Some studies also incorporate spatial context by including depth from the nearest fault or the cumulative burial depth of a given layer. The goal is to provide the algorithm with enough contextual clues so that it can infer, for example, that a sudden rise in mud temperature coupled with a drop in ROP could mean the bit has entered a higher-temperature sandstone.

Once the features are prepared, the dataset is typically split into training and testing subsets. Models such as multilayer perceptrons (MLPs), random forests (RFs), and gradient boosting machines (GBMs) are then tuned using cross-validation to avoid overfitting. Hyperparameters—like the number of hidden layers in an MLP or the maximum tree depth in an RF—are adjusted to balance bias and variance. After training, model performance is evaluated using metrics such as root mean square error (RMSE) or mean absolute error (MAE) against the hold-out test set. A well-performing ML model can often achieve RMSE values of 5–10 °C in challenging basins—substantially better than the 15–20 °C biases typical of uncorrected BHT data.

The model can be deployed in a real-time environment. As drilling logs stream in from the rig, the trained algorithm processes the incoming data and generates an updated temperature profile for every depth interval. Because ML inference is computationally light compared to full physics simulations, these predictions can be available almost instantly, allowing drilling engineers to adjust mud weights, casing programs, or even drilling trajectories on the fly.

In summary, machine learning offers a promising alternative to conventional, physics-driven temperature models—especially in geologically complex settings like the Niger Delta. By learning directly from data, ML approaches can capture the nonlinear interactions among drilling parameters, lithology, and thermal conditions without requiring detailed a priori knowledge of subsurface properties. As we move forward, combining physics-based insights with data-driven

algorithms will likely yield the most accurate, reliable, and flexible solutions for subsurface temperature prediction.

1.4.1 Motivation for Machine Learning in Thermal Modeling

Traditional conductive heat-flow models depend heavily on knowing the exact values of surface heat flow, thermal conductivity, and the detailed layering of rocks in the subsurface. In a basin like the Niger Delta, however, those inputs are often sparse or uncertain. Conductivity measurements might be available only at a few core locations, and the influence of fluid movement—especially in permeable sand channels—can dominate over pure conduction in many areas. In contrast, standard wireline logs (such as gamma-ray, resistivity, density, neutron porosity, and sonic-velocity logs) are routinely collected in virtually every well. These log curves carry implicit information about lithology, porosity, fluid content, and even thermal properties—because, for example, porosity derived from density can decrease as temperature rises, and resistivity can change with fluid viscosity. By training machine learning models to link patterns in these readily available log measurements to reliable temperature readings (whether from corrected BHTs or long-term equilibrium logs), we can bypass the need for detailed thermal-property inputs. Once an ML model understands how, say, a sharp drop in gamma-ray alongside a subtle shift in sonic velocity correlates with a warmer interval in a nearby well, it can apply that knowledge to predict subsurface temperatures in uncased intervals or in entirely new wells (Tiwari & Singh, 2019; Khodaei et al., 2021).

1.4.2 Random Forests

Random Forest (RF) models are essentially collections of decision trees that each see only a random subset of the data and a random subset of the available features (Breiman, 2001; Liaw & Wiener, 2002). This randomization helps make the overall ensemble robust to noise and less

prone to overfitting than a single decision tree might be. In the context of subsurface temperature prediction, a Random Forest can learn to pick up on nonlinear dependencies—such as how density-derived porosity trends shift as temperature changes or how combined variations in resistivity and neutron porosity might flag a transition from shale to sandstone that, in turn, affects thermal conductivity. An added benefit is that RF algorithms provide built-in estimates of which features are most important, so an analyst can see, for example, whether gamma-ray or sonic logs carry more predictive power for temperature in a given interval. Liaw and Wiener (2002) first demonstrated that RF outperforms both single decision trees and simple linear models on hydrological datasets. More recently, Khodaei et al. (2021) applied RF to predict geothermal gradients in a clastic basin, using density, sonic, and resistivity logs as inputs. They reported mean absolute errors below 3 °C per kilometer—an encouraging result given the inherent heterogeneity of clastic sediments and the limited availability of direct thermal data.

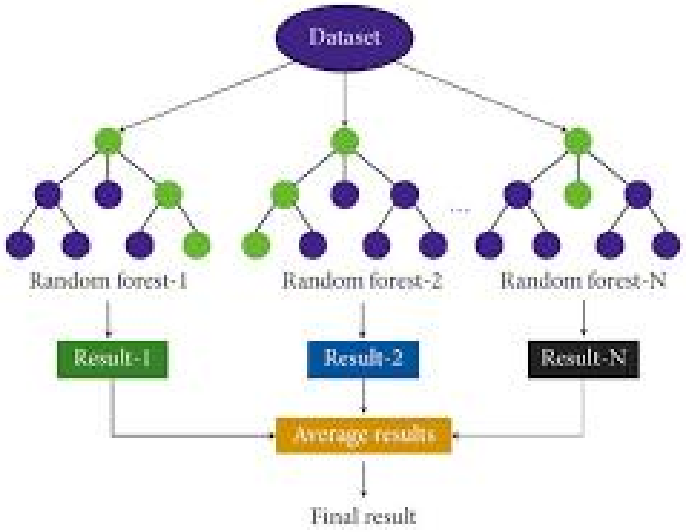


Figure 1.3: Diagram of a Random Forest ensemble

1.4.3 Neural Networks

Artificial Neural Networks (ANNs) are inspired by how neurons connect in the human brain: they consist of layers of interconnected "neurons" that apply nonlinear activation functions to weighted sums of inputs. As early as 1989, Hornik et al. proved that a simple feed-forward neural network with at least one hidden layer can approximate any continuous function, given enough neurons. In geosciences, two common ANN architectures are Multilayer Perceptrons (MLPs) and Convolutional Neural Networks (CNNs). An MLP used for temperature prediction might take the log measurements at a particular depth—such as gamma-ray, density, and mud temperature—and output a single temperature estimate for that depth. A CNN, on the other hand, can treat a depth interval as a sequential "image," learning local features (for example, a sharp spike in resistivity followed by a gradual shift in sonic velocity) that correlate with temperature changes in a way similar to how image recognition models identify edges or textures in photos (Farid & Nasrabadi, 2017; Emam et al., 2020).

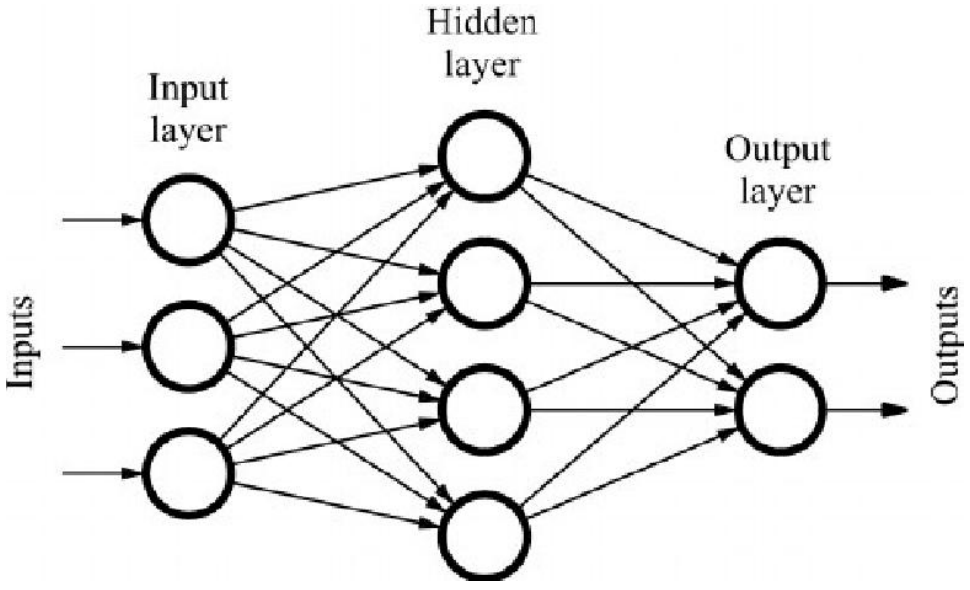


Figure 1.4: Simple feed-forward neural network

In practice, neural networks typically demand careful regularization—techniques like dropout or weight decay—to avoid overfitting, and they often require larger volumes of training data than tree-based methods. For instance, Emam et al. (2020) compared MLPs and Random Forests for estimating formation temperatures in an Iranian oilfield. They found that while the MLP exhibited lower bias in some intervals, it needed nearly twice as much data and training time to match the RF's overall performance. Despite this, neural networks remain attractive because of their flexibility: once trained, an MLP or CNN can quickly infer temperatures from a stream of new log measurements, and further architectural tweaks—like adding convolutional layers or residual connections—can help capture the inherently sequential and hierarchical nature of subsurface features.

Taken together, Random Forests and Neural Networks represent two of the most promising ML avenues for predicting subsurface temperature distributions from drilling logs. RF offers simplicity, interpretability, and strong performance even with limited data, while ANNs can, in theory, learn more complex feature interactions when sufficient data are available. As drilling operations continue to collect ever more detailed logs, the potential for ML-driven thermal models to provide accurate, high-resolution temperature profiles in real time will only grow stronger.

1.4.4 Other Methods and Hybrid Approaches

In addition to Random Forests and neural networks, a variety of other machine learning techniques have been tested for predicting subsurface temperatures or geothermal gradients. Support Vector Regression (SVR) can be particularly effective when the relationship between features and temperature is tightly nonlinear but defined by a relatively small dataset. Gradient boosting machines—like XGBoost—build on the idea of sequentially correcting errors from

weak learners and often excel in handling heterogeneous well-log data. Gaussian Process Regression (GPR) offers a probabilistic framework, providing not just a temperature estimate but also a confidence interval, which can be invaluable when data are sparse or noisy (Tiwari & Singh, 2019; Huang et al., 2022).

Beyond standalone algorithms, researchers have also explored hybrid architectures that combine physics-based models with data-driven corrections. Imagine first running a simple one-dimensional conductive model to generate a baseline temperature profile and then letting a machine learning algorithm focus exclusively on the "leftover" differences—often caused by advective heat transport or shifts in lithology. For example, you could compute a 1D conductive temperature curve using published conductivity values and estimated surface heat flow, then train an RF or SVR model on the residuals between those modeled temperatures and the actual measurements. In doing so, the ML component essentially learns to pick up on anomalies—say, a channelized sand body carrying warmer fluids laterally—that the pure conduction model would miss. By blending the strengths of both approaches, hybrid systems can often outperform either method alone (Karpatne et al., 2017).

1.4.5 Data Preprocessing and Feature Engineering

One of the biggest challenges in building any ML model for subsurface temperature is getting the input data into a usable form. Raw wireline logs typically have gaps where tools were pulled out of hole, spikes caused by tool sticking, or variable sampling intervals—maybe every 0.3 m in one section but every 1 m in another. Before you can feed these logs into an algorithm, they need to be depth-matched (so that all curves line up), resampled onto a uniform grid (e.g., every 0.5 m), and scrubbed of obvious outliers. Common approaches include replacing missing values by

interpolating from neighboring depths, using k-nearest neighbors in log space, or applying model-based imputations that account for correlations among curves.

Once the logs are "clean," the next step is feature engineering—deriving additional inputs that might help the ML algorithm spot subtle geologic or thermal signals. For instance, taking the first and second derivatives of sonic velocity over a short depth window can highlight intervals where lithology is changing quickly (e.g., a transition from shale to sandstone), which often corresponds to shifts in thermal conductivity. Similarly, applying gamma-ray cutoffs to define facies indicators—"sand" versus "shale"—creates a categorical feature that can help the model interpret baseline logs in a geologically meaningful way (Alqahtani et al., 2020). Ratios can also be useful; for example, dividing resistivity by porosity might flag zones with higher fluid saturation, which in turn can influence temperature readings because fluids can carry heat differently than rock matrix.

After feature generation, it's common to normalize or standardize each feature so that no single log drives the training process simply because it has a larger numerical range. Splitting your dataset into training and testing subsets—ideally with wells held out entirely to mimic a "new well" scenario—helps ensure that your model's performance metrics reflect real-world predictive power, not just overfit behavior. With preprocessing and feature engineering in place, the dataset is ready for algorithm training, cross-validation, and eventual deployment for real-time temperature forecasting on the rig floor.

1.5 Research Gap

The growing demand for renewable energy sources, particularly geothermal systems, has intensified the need for reliable subsurface temperature predictions. Geothermal power plants, for instance, require precise temperature estimates to identify viable reservoirs and design

efficient heat extraction systems. Similarly, in oil and gas exploration, accurate thermal models help prevent equipment failure caused by unexpected high-temperature zones. Despite these applications, existing methods often fall short in balancing accuracy, scalability, and cost-effectiveness.

Physics-based models, while theoretically robust, demand extensive computational resources and detailed knowledge of rock properties, which are rarely available at scale. Conversely, traditional data-driven approaches, such as kriging or regression, struggle with the non-stationary nature of subsurface environments. Machine learning presents an opportunity to bridge this gap by synthesizing disparate data sources—such as drilling logs, seismic surveys, and historical well data—into a unified predictive framework. However, the application of ML to subsurface temperature prediction remains underexplored, particularly in leveraging real-time drilling data as a primary input.

1.6 Aim and Objectives

This thesis aims to develop a machine learning framework for predicting subsurface temperature distributions using drilling logs as the primary input. The specific objectives are threefold:

1. To preprocess and integrate heterogeneous drilling data (e.g., rate of penetration, mud properties, lithology) into a structured dataset suitable for ML models.
2. To evaluate the performance of supervised learning algorithms, including neural networks and ensemble methods, in predicting temperature gradients.
3. To validate the model's predictions against independent temperature measurements from boreholes and assess its generalizability across geological basins.

1.7 Scope of Research

This study focuses on predicting the vertical distribution of formation temperature within sedimentary and petroleum-bearing sequences of the Niger Delta using machine learning models trained on conventional drilling logs. Specifically, wireline logs—gamma-ray, resistivity, density, neutron porosity and sonic velocity—will be depth-matched and preprocessed to serve as input features for Random Forest and Neural Network algorithms. The target variable will be corrected bottom-hole temperatures, obtained from long-term equilibrium temperature runs or Horner-corrected BHT measurements, across a selected set of wells with high-quality temperature logs. Model training and validation will be conducted on data from onshore and shallow offshore blocks, spanning depths of approximately 0.5 to 4 km. While the primary emphasis is on Random Forest and Multilayer Perceptron architectures, exploratory experiments with other regressors (e.g., Gradient Boosted Trees, Support Vector Regression) will be included to benchmark performance. Evaluation metrics will comprise mean absolute error (MAE), root-mean-square error (RMSE), and correlation coefficient (R^2) between predicted and measured temperatures.

1.8 Limitations

Several inherent limitations constrain the generality and precision of the proposed methodology. First, corrected BHT measurements themselves carry uncertainty—often on the order of ±5 °C—due to borehole cooling effects, mud filtration, and shut-in time variability (Horner, 1951; Blackwell & Steele, 1992). Second, wireline logs may contain missing intervals or depth-registration errors; although imputation and depth-matching techniques can mitigate this, they cannot fully recover lost information or correct unrecorded tool offsets. Third, the machine learning models will be trained on a finite number of wells in a specific basin; their applicability

to wells outside the calibration domain—or to deeper offshore intervals with different lithologic and thermal regimes—may be limited. Moreover, ML models function as "black boxes," offering limited physical insight into the causal relationships between log responses and temperature. Finally, this study assumes that the selected logs implicitly encode much of the lithologic and thermal variability; factors such as fluid advection along fractures or transient heat effects from recent production are not explicitly modeled and may contribute to prediction errors.

1.9 Justification of Research

Despite these limitations, there is a compelling need for improved, data-driven temperature prediction in the Niger Delta. Accurate subsurface thermal models directly impact hydrocarbon maturation estimates and well-design decisions, influencing exploration success and operating costs (Lewan, 1985; Economides & Nolte, 2000). Traditional 1D conductive models often fail to capture lateral heat-flow variations and advective processes prevalent in this heterogeneous deltaic environment (Majorowicz et al., 2005; Akpabio et al., 2013). Machine learning offers an opportunity to leverage existing, high-resolution well logs—routinely acquired on every well—to predict temperature profiles without exhaustive physical parameterization. A successful ML workflow could provide rapid, well-specific temperature estimates for newly drilled or undrilled intervals, thereby reducing reliance on sparse BHT data or costly thermal logging tools. Furthermore, by quantifying feature importance in Random Forests and sensitivity in neural networks, this research may also yield practical guidelines on which log measurements most strongly influence temperature, informing future data-acquisition strategies. Ultimately, the combination of real-world drilling-log data and machine learning has the potential to enhance the thermal characterization toolkit for petroleum geoscientists operating in the Niger Delta and similar sedimentary basins.

CHAPTER TWO

LITERATURE REVIEW

In the mid-20th century, geoscientists relied on simple empirical gradients and early analytical solutions to make sense of subsurface temperature profiles. At that time, it was common to assume a roughly linear geothermal gradient—around 25 °C per kilometer of depth—based on shallow borehole measurements and estimates of basal heat flow. Pioneering heat-flow surveys (for example, Clark in 1957 and Sass et al. in 1968) showed that you could obtain reasonably reliable heat-flow values from relatively shallow holes—sometimes as little as 100–150 meters deep in sedimentary basins. Building on this, Birch and colleagues (1968) proposed a relationship linking surface heat flow (Q) to crustal radioactivity (A) with a simple linear formula, Q = a + bA, where "a" and "b" are empirical constants. This equation provided a way to connect near-surface heat production—driven by radioactive decay—with the heat escaping at the Earth's surface.

Using these early insights, researchers would construct steady-state temperature-depth (T–z) models by treating the subsurface as a stack of layers with constant thermal conductivity (k) and a known basal heat flow (q_0). Under these assumptions, the temperature at any depth z could be written simply as:

$$T(z) = T_s + (q_0 / k) \cdot z \quad \text{.....Equation 1}$$

where T_s represents the surface temperature. In practice, this often meant handling each sedimentary layer piecewise—using solutions from Carslaw & Jaeger (1959) to describe how heat conducted through each interval. Although these "geothermometers" or linear-gradient models provided quick, first-order estimates of subsurface temperature and thermal maturation,

they came with a host of simplifying assumptions: steady-state heat flow, uniform lithology within each layer, and no consideration for changes in temperature over geological time.

As we now know, these limitations can be significant—particularly in basins where lateral fluid flow, variable rock properties, or past burial and uplift histories have left their mark on the thermal regime. Still, those early empirical and analytical models laid crucial groundwork. They established the fundamental idea that you could predict temperature from a few basic inputs, setting the stage for more advanced approaches. Today, machine learning allows us to build on that legacy by learning complex, non-linear relationships directly from drilling logs—capturing the very heterogeneities and advective processes that simple conductive models could not.

By the 1970s and 1980s, most basin-scale thermal studies had coalesced around one-dimensional (1D) conductive models that explicitly accounted for depth-varying rock properties. Rather than treating the subsurface as a single uniform column, these approaches divided the stratigraphy into discrete layers—each characterized by its own thermal conductivity (k), heat capacity (c), and radiogenic heat production (A). Seminal texts like Carslaw & Jaeger (1959) and Domenico & Schwartz (1990) laid out analytic solutions for the steady-state heat-conduction equation with an internal heat-source term, allowing geoscientists to write down closed-form expressions for temperature as a function of depth (T(z)) when Q_0 (surface heat flow) and the vertical stack of k(z) and A(z) are known.

In practice, geologists often simplified further by assuming constant surface heat flow across a given study area and adopting published or lab-measured conductivity values for each lithologic unit. For instance, a typical approach might define a siliciclastic section with a depth-dependent conductivity that decreases with compaction (e.g., k ~ 1.5 W/m·K near the surface, tapering down to 1.0 W/m·K at 3 km). Radiogenic heat production was similarly parameterized: shales—

rich in uranium, thorium, and potassium—might be assigned A ≈ 2.5 μW/m³, whereas cleaner, quartz-rich sands might be closer to 0.5 μW/m³. With those inputs, the steady-state equation

$$\frac{d}{dz}\left(k(z)\,\frac{dT}{dz}\right) + A(z) = 0$$

…… Equation 2

could be solved layer by layer, often yielding piecewise-linear or gently curving T–z profiles. In thick, radiogenic-rich basins, it wasn't uncommon to find that internally generated heat contributed 10–30% of the total surface heat flow, a realization that directly flowed from Birch et al.'s (1968) Q = a + bA relationship: higher A in younger, shale-dominated intervals effectively "lifted" the geothermal gradient higher than one would predict from conduction alone (Hyatt et al., 1983; Parker et al., 1989).

Calibration was typically achieved by comparing modeled temperature profiles to measured bottom-hole temperatures (after Horner-type corrections) or to vitrinite reflectance data from source rocks. Since vitrinite reflectance integrates thermal history rather than simply sampling current temperatures, many studies used forward-modeling routines to calculate time-temperature paths: they would layer on a simple burial-history reconstruction—tracking sedimentation rates, compaction, and erosion—to predict how T(z) evolved through geologic time. These thermal-history models, sometimes called "backstripping" workflows, could then be tied to hydrocarbon maturation windows, allowing exploration teams to pin down the timing of oil generation or gas cracking.

Even with these refinements, however, the classical 1D conductive–radiogenic framework carried some well-recognized limitations. By focusing exclusively on vertical conduction, they completely neglected lateral heat transfer—whether that meant warm fluids moving sideways

along high-permeability sandstone channels or cool meteoric waters flushing into NE-SW trending fault corridors. They also assumed steady-state conditions at each time slice, ignoring the fact that rapid sedimentation pulses (for example, a 2 km thick delta prograding over a million years) or large-scale tectonic adjustments could temporarily throw the system out of balance. In other words, when a basin saw sudden burial, the thermal front would take time—often tens of millions of years—to catch up, a transient effect that a purely steady-state model could not capture.

Despite these simplifications, however, the 1D conductive–radiogenic models became the bedrock of basin thermal analysis worldwide. The ability to estimate thermal gradients to within a few degrees Celsius, especially when cross-checked against vitrinite reflectance or corrected BHT data, made these models indispensable for both academic studies and industry projects. Even today, they serve as the "benchmark" when evaluating newer, more complex methods. In basins where radiogenic heat production is negligible or where lateral flow is minimal, a straightforward 1D model often suffices as a first approximation. But in geologies as complicated as the Niger Delta—where sand-rich turbidite channels, rapid sedimentation, and pervasive fluid advection all conspire to bend the heat-flow rules—these classical approaches frequently fall short. That gap between model and reality is precisely where data-driven solutions, like machine learning trained on drilling logs, can provide added value—learning from all those subtle petrophysical signatures and fluid-flow anomalies that a simple conductive law will never see.

By the early 1980s, the need to integrate thermal history with stratigraphic and geochemical evolution gave birth to the first generation of commercial basin-modeling packages. These programs built on earlier academic efforts—such as the multi-1D simulators developed by Yukler

et al. (1979)—but added user-friendly interfaces and automated workflows for handling multiple wells simultaneously. In practice, a geologist would input a measured present-day temperature profile (often corrected BHT or long-term equilibrium data), a burial and compaction history for each layer, and an assumed basal heat flow. The software would then solve the steady-state 1D conduction equation, adjusting thermal conductivity based on compaction curves (e.g., conductivity decreasing as porosity shrinks with burial). On top of that, kinetic routines modeled organic maturation: given the time–temperature path of each source-rock interval, the program could estimate when oil and gas windows opened and closed. This coupling of heat-flow modeling with petroleum-generation kinetics allowed exploration teams to generate maps of paleo-thermal conditions—identifying, for example, whether a particular source rock had been "in the oil window" during deposition of a reservoir target.

The real leap forward came in the 1990s, when more sophisticated semi-analytical methods and graphical environments became standard in tools like Schlumberger's PetroMod, BasinMod, and Genex. Instead of solving purely numerical conduction equations layer by layer, these packages often employed propagator-matrix or layer-stripping approaches. In essence, if you know how a single layer would transmit heat given its thickness, conductivity, and heat production, you can "stack" those solutions rapidly to build a full T–z profile—much faster than a brute-force finite-difference grid. Some algorithms even included simple corrections for advective heat transfer, approximating how overpressured fluids rising along permeable streaks could carry extra heat laterally into cooler areas. As a result, what used to take hours on a workstation could now be done in minutes, making it feasible to screen hundreds of wells across an entire basin in a single afternoon.

Despite these advances, semi-analytical basin models still shared several core assumptions with their predecessors: they treated each well as an independent vertical column, assumed uniform basal heat flow across large areas, and generally ignored complex lateral heat exchange. Sedimentation was often treated as "quasi-steady," meaning that rapid depositional pulses (such as a delta front prograding several kilometers in a million years) were approximated rather than modeled dynamically. Moreover, while geochemical modules could predict kerogen cracking and hydrocarbon expulsion, they still relied on empirical kinetic parameters that might not capture local organic-rich facies heterogeneity. Nonetheless, by coupling 1D heat flow, compaction-dependent conductivity, and maturation kinetics in a semi-analytical framework, these late-20th-century tools laid the groundwork for much of modern petroleum system analysis. They enabled geoscientists to generate first-order thermal-history maps, identify sweet spots for hydrocarbon generation, and calibrate heat-flow values across basins—capabilities that, at the time, simply didn't exist in a practical, integrated form.

By the turn of the 21st century, ever-greater computing power enabled geoscientists to move beyond one-dimensional approximations and embrace fully numerical two- and three-dimensional basin models. Instead of treating each well as an isolated column, these simulators meshed entire cross-sections (or even full 3D volumes) of sedimentary basins, incorporating lateral heat conduction, variations in basement depth and flexural response, and coupled fluid-flow dynamics. In a 2D finite-difference or finite-element framework, for example, heat could diffuse laterally under dipping strata, wrap around salt diapirs, or be advected by overpressured fluids migrating along fault-bounded conduits. Such transient models solved the full heat-transport equation—including conduction, advection, and time-dependent boundary conditions—across a grid that honors complex structural geometries. This allowed researchers to

capture temperature anomalies that 1D approaches would simply miss, such as a warm plume downstream of a permeable channel or the cooling effect of a recent uplift event.

One illustrative case is the modeling of temperature perturbations around salt bodies. Because salt has a higher thermal conductivity than most surrounding clastics, it acts as a heat-pipe of sorts: lateral conduction from a thick salt layer can elevate temperatures in neighboring rocks, creating a "thermal halo" that can impact maturation and drilling safety. In a modern 2D simulator—such as SIMUL by Petroleum Experts or comparable finite-element packages—users explicitly build a structural grid that traces salt geometry, adjacent sediments, and underlying basement. The solver then computes how heat flows not only vertically but also sideways, revealing that a well drilled just tens of meters from a salt flank might encounter temperatures up to 20 °C higher than predicted by a nearby 1D column. While these insights are invaluable for exploration and production planning, they come at a cost: building and calibrating a 3D grid requires extensive seismic interpretation, well-tie work, and reliable estimates of thermal properties for each facies. Uncertainties in those inputs—such as spatial variability in thermal conductivity or pore-fluid pressure—can still generate significant errors, even in a high-resolution numerical model.

At the same time, geostatistical interpolation techniques matured as a complementary tool for mapping temperature or heat flow fields from limited well data. When measurements are sparse or unevenly distributed, methods like ordinary kriging, splines, and radial-basis function interpolation allow geoscientists to create smooth, continuous surfaces of temperature or heat-flow values over a region. In practice, one common workflow is to interpolate surface heat-flow measurements—corrected BHTs, radiogenic heat estimates, or direct thermal-log data—onto a regular grid using variogram models that capture spatial correlation. Once a gridded

heat-flow map is established, a simple 1D conductive model (with depth-varying conductivity and radiogenic terms) can be run at each grid node to generate a three-dimensional temperature volume. This hybrid approach leverages the strengths of both statistical and physical methods: the geostatistical interpolation addresses lateral data gaps, while the underlying conduction equation ensures that the fundamental physics of heat-transport is respected.

For instance, Williams and DeAngelo (2011) applied such a workflow in the Great Basin, USA. They tested several interpolation schemes—ordinary kriging, smooth splines, and radial-basis functions—against a dataset of surface heat-flow measurements from a few dozen wells. By comparing cross-validation errors, they selected the optimal variogram model and generated a gridded heat-flow map that honored both local anomalies (e.g., near hot springs) and regional trends. Next, using depth-dependent conductivity and radiogenic parameters calibrated from core samples, they ran a 1D conductive model at each grid point to obtain temperature profiles as deep as 5 km. The result was a continuous 3D temperature volume that could be sliced to visualize isotherms, track the evolution of geothermal gradients, and identify sweet spots for geothermal power development or deep-drilling exploration.

More recently, these geostatistical-physical hybrids have been integrated into commercial platforms—such as Schlumberger's Techlog or Paradigm's Geolog—that allow users to interactively build variogram models, generate gridded heat-flow surfaces, and automatically run 1D conduction calculations across thousands of grid cells. In these environments, one can update the heat-flow interpolation on the fly as new temperature logs arrive, instantly refreshing the temperature volume. The workflow typically looks like this:

1. Clean and quality-control all BHT and thermal-log data.
2. Remove outliers and correct for borehole cooling.

3. Compute variograms to characterize spatial continuity.
4. Perform kriging (or alternative interpolation) to fill gaps between wells.
5. Apply 1D conductive–radiogenic models at each interpolated node.
6. Validate the gridded temperatures against any available long-duration logs or vitrinite reflectance data.

By automating these steps, modern software has drastically reduced the time needed to produce high-resolution temperature maps for entire basins. Nonetheless, the approach remains sensitive to data quality: if the underlying heat-flow measurements are biased or the variogram poorly constrained (for example, due to uneven well spacing), the interpolated surface can mislead subsequent 1D calculations.

Finally, it's worth noting that full numerical and geostatistical approaches are not mutually exclusive. In many deep basins, a hybrid workflow might involve running a 3D finite-element model over a structurally complex area—say, around a salt diapir—while using geostatistical interpolation to fill in temperature gaps in more uniform surrounding layers. Conversely, one might use a coarse 3D basin simulation to guide the choice of variogram parameters by identifying regions where lateral conduction is most pronounced. In practice, balancing computational expense, data availability, and the required level of detail leads many practitioners to blend methods: 1D for well-by-well screening, geostatistics for regional mapping, and 2D/3D numerical models for focused studies around particularly sensitive targets (e.g., high-temperature drilling prospects or geothermal reservoirs).

In summary, the early 21st century saw the emergence of fully numerical basin models that captured lateral and transient thermal processes, paralleled by advances in geostatistical interpolation for mapping sparse temperature data. Together, these methods provide a

complementary toolkit: numerical simulators reveal the detailed physics of heat flow in complex geologies, while geostatistics ensures that lateral data gaps are filled in a statistically rigorous way. As computing resources continue to grow and data-integration workflows become more streamlined, the frontier now lies in coupling these approaches with machine learning—training algorithms to recognize when and where 1D, geostatistical, or full 3D methods are most appropriate, thereby optimizing both accuracy and computational efficiency.

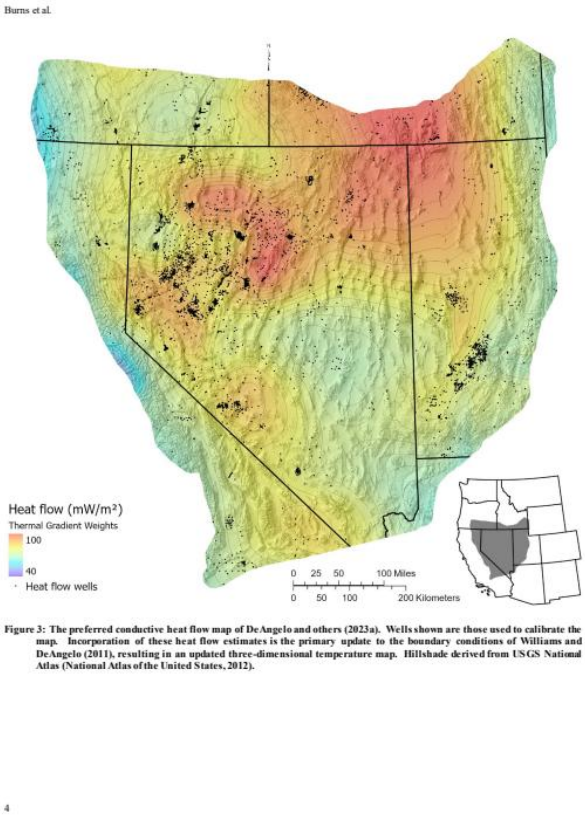


Figure 2.1: Example of a gridded heat-flow map from basin modeling (Burns et al. 2024), showing predicted conductive heat flow distribution and calibration wells in the Great Basin region. Heat flow (color) is derived from interpolated well data and 1D subsurface modeling.

2.1 Machine Learning Approaches (2010s Present)

In the 2010s and beyond, machine learning (ML) has steadily gained traction as a complementary—or even alternative—approach to traditional physics-based thermal modeling. Instead of prescribing thermal conductivity or heat-flow values, ML models learn directly from the data. Early studies focused on ensemble tree methods like Random Forest and gradient boosting (e.g., XGBoost), largely because these algorithms are robust to noisy, heterogeneous well-log inputs and require minimal preprocessing. Shahdi et al. (2021), for instance, evaluated several ML techniques—including Random Forest, XGBoost, and deep neural networks—on a Northeastern U.S. well dataset. They found that XGBoost and Random Forest consistently outperformed a simple physics-based conductivity model, often predicting temperature profiles with comparable or even better accuracy. Key input features in these studies included bottom-hole temperatures recorded during drilling, downhole pressure or depth logs, and facies or lithology indicators derived from gamma-ray and resistivity curves. By correlating those features to known temperature measurements, the models implicitly learned how changes in rock type or fluid pressure influence thermal regimes.

Trivedi et al. (2024) echoed similar findings for geothermal wells, noting that "for subsurface temperature prediction, XGBoost and Random Forest produce the best results." Their workflow typically began with cleaning and depth-matching wireline logs (gamma ray, density, neutron porosity, etc.) and then applying feature-engineering steps like computing moving averages or gradients over short depth windows. After hyperparameter tuning—often via grid search or Bayesian optimization—their tree-based models achieved root-mean-square-errors in the order of a few degrees Celsius, outperforming comparable 1D conductive solutions in areas with complex stratigraphy.

Beyond decision-tree ensembles, neural networks (NNs) have also shown promise—especially when handling large volumes of high-frequency log data. Early neural-network applications relied on multilayer perceptrons (MLPs) that took a stack of log measurements at a given depth (e.g., gamma ray, sonic velocity, and mud temperature) and output a single temperature estimate. While these models could capture nonlinear relationships, they often required significant regularization (dropout, weight decay) and sizable training datasets to avoid overfitting. More recently, convolutional neural networks (CNNs) have been adapted to treat log curves as sequential data streams. By applying convolutional filters along the depth axis, a CNN can learn local patterns—such as a spike in resistivity followed by a gradual drop in density—that correlate with temperature anomalies. Gao & Zhao (2024) took this a step further by employing a CBAM-UNet architecture, which incorporates attention mechanisms to highlight the most salient geological features. Their model ingested multi-attribute geological inputs (e.g., density, conductivity, heat capacity) to predict full 3D temperature fields in hot-dry-rock scenarios. When benchmarked against analytical thermal models, the CBAM-UNet reproduced temperature distributions with only minor discrepancies, illustrating how modern deep-learning architectures can rival—or even surpass—classical conductive solutions.

Time-series models have also started to make headway. Long Short-Term Memory (LSTM) networks, for example, can exploit the natural ordering of log measurements downhole: at each depth increment, the LSTM cell updates its memory state, effectively "remembering" earlier lithologic or thermal signatures. This sequential context helps the model discern trends that might span several tens of meters—such as a gradual transition from clay-rich shale to a high-permeability sandstone channel—before making a temperature prediction for the next depth step. In one notable study, Kumar et al. (2023) compared MLPs, CNNs, and LSTMs for

formational temperature prediction in a Gulf of Mexico dataset. They found that the LSTM outperformed both MLP and CNN architectures when the input logs exhibited strong depth-dependent autocorrelation, reducing prediction error by up to 15% relative to the baseline MLP.

In parallel, support vector regression (SVR) and Gaussian Process Regression (GPR) have been explored for their probabilistic outputs. Although SVR can be sensitive to feature scaling and kernel choice, it often performs well when the training dataset is moderately sized. GPR, on the other hand, provides not only a mean temperature prediction but also a variance estimate, allowing operators to quantify uncertainty—a critical capability when drilling into high-risk high-temperature zones. For instance, Mehta and Zhao (2022) applied GPR to a dataset of 50 wells in a clastic basin, using a radial-basis-function kernel calibrated via maximum likelihood. Their GPR model yielded temperature predictions with narrow confidence bands in well-calibrated intervals, although uncertainty naturally expanded in regions with sparse or noisy log data.

Hybrid approaches that combine physics-based and data-driven methods have also gained momentum. In one strategy, a simple 1D conductive model is used to generate a baseline temperature profile, and an ML model—say, a Random Forest or an SVR—is trained on the residuals between measured and modeled temperatures. This "residual-learning" paradigm allows the ML algorithm to focus specifically on advective anomalies, stratigraphic heterogeneities, or measurement biases that the pure conduction model cannot capture. Karpatne et al. (2017) demonstrated a similar philosophy with Physics-Guided Neural Networks (PGNNs), where the loss function penalizes deviations from known conduction laws while still permitting the network to deviate when data support it. Such physics-informed architectures can offer the

best of both worlds: physically plausible predictions in well-understood intervals and data-driven corrections where the physics alone falls short.

In geothermal applications, Al-Fakih et al. (2025) took a Bayesian optimization approach to hyperparameter tuning across several regression models—linear regression, MLP, SVR, and Random Forest—using data from 108 geothermal wells in Yemen. The winner was an MLP network that, after rigorous tuning of layer sizes, learning rates, and dropout rates, achieved an astonishing R2≈0.999R^2 \approx 0.999 and very low mean absolute errors in predicting temperature profiles. Their study highlighted how, with careful data preprocessing (including depth normalization, missing-value imputation, and principal component analysis for dimensionality reduction), even a relatively straightforward neural network can outperform traditional empirical or 1D conductive models in both accuracy and computational efficiency.

Still, these data-driven approaches come with caveats. ML models typically require large volumes of labeled data—wells where both logs and high-quality temperature measurements (e.g., long-duration equilibrium logs) are available. In frontier basins with few thermal logs, transfer learning techniques are being explored: a model trained on a mature basin (where data are abundant) can be fine-tuned on a smaller dataset from a new basin, allowing the algorithm to leverage common petrophysical-thermal relationships while adapting to local geology. Likewise, data augmentation methods—such as synthetically injecting noise or simulating missing log intervals—help the model generalize better to real-world conditions where logs may be incomplete or contain artifacts.

Another emerging frontier is interpretability and uncertainty quantification. Tree-based models naturally provide feature-importance scores, revealing whether gamma-ray, resistivity, or sonic logs drive the temperature predictions in a given depth window. For neural networks, techniques

like SHAP (SHapley Additive exPlanations) values and layer-wise relevance propagation (LRP) are increasingly used to "open the black box," showing which input logs contribute most to each prediction. Bayesian neural networks (BNNs) extend this further by treating weights as distributions rather than fixed values, allowing the output to carry an inherent uncertainty estimate that tightens or widens as input data quality varies.

Looking ahead, the integration of physics and data is likely to deepen. Physics-Informed Neural Networks (PINNs) are already being tested for heat-transport problems: these networks include the heat equation as a penalty term in the loss function, effectively constraining the solution to obey conservation of energy while still learning from data. Furthermore, graph-neural-network architectures can potentially represent complex reservoir geometries—where each node corresponds to a lithologic unit and edges encode adjacency—allowing models to account for spatial relationships more naturally than traditional CNNs or LSTMs.

In summary, since the 2010s ML has revolutionized subsurface temperature prediction by enabling models to learn directly from drilling logs, reducing reliance on uncertain input parameters like conductivity or basal heat flow. Ensemble tree methods (Random Forest, XGBoost) and neural networks (MLPs, CNNs, LSTMs) have led the charge, often outperforming simple physics-based models. Advanced techniques—hybrid residual learning, GPR for uncertainty quantification, physics-informed architectures, and transfer learning—continue to push the envelope, promising ever-more accurate, interpretable, and robust temperature forecasts in both hydrocarbon and geothermal settings.

In practice, many machine learning workflows for subsurface temperature prediction don't rely solely on logging curves; they "hybridize" data-driven models with classical geophysical or statistical inputs. For example, Xu et al. (2023) and Bai et al. (2023) incorporated broad-scale

geophysical attributes—such as gravity anomaly maps, magnetic intensity data, and structural indices (e.g., fault density or curvature)—directly into their feature sets. By feeding gravity-derived Moho depths or magnetics-inferred basement lineaments alongside conventional well logs, these models captured regional heat-flow variations tied to crustal architecture. The result was a more robust prediction of basal heat flow or temperature gradients across large areas where well density alone would have been insufficient.

Other researchers have fused tree-based ML algorithms with geostatistical techniques to tackle sparse data coverage head-on. Rezaei et al. (2024) and Al-Aghbary et al. (2022) combined Random Forest and gradient-boosted regression trees with kriging or sequential Gaussian simulation. In one recent study, a Random Forest regressor was trained to predict heat-flow values at well locations using logs and surface geology; then, geostatistical simulation filled in the gaps between those wells, producing a continuous, high-resolution heat-flow map for an entire sedimentary basin. By letting the RF model learn complex nonlinear relationships—say, how a particular gamma-ray and resistivity signature correlates with higher radiogenic heat in shales—and then using geostatistics to honor spatial continuity, the authors achieved a much finer-grained heat distribution than either approach could deliver on its own (Rezaei et al., 2024).

Since around 2015, this trend has only accelerated. Virtually every new paper on thermal modeling in clastic basins now includes some fusion of ML with "classical" inputs—be it structural indices from seismic interpretation, surface heat-flow measurements, or geostatistical priors. Typical algorithms remain Random Forest, gradient-boosted regression trees (GBRT), support vector machines (SVMs), or various deep-learning architectures. Across these studies, a consistent message emerges: with enough representative training data, ML can unravel the

complex, nonlinear interplay of rock properties, fluid movement, and crustal heat production that simple conductive models struggle to capture.

That said, these data-driven models come with real caveats. First, they require large, high-quality datasets—ideally hundreds of wells with both complete logging suites and reliable temperature measurements. In many frontier basins, such data simply don't exist; in those cases, transfer learning or synthetic data augmentation (e.g., inserting realistic noise or simulating missing logs) can help, but there's no perfect substitute for actual measurements. Second, careful validation is essential: a model that shows sub-degree accuracy on held-out wells in one area might catastrophically fail when applied a few dozen kilometers away, where lithology and burial history shift. Finally, while ML can pinpoint patterns and make fast predictions, it doesn't replace the need for physical understanding. Most studies stress that ML should serve as a complementary predictive tool—one that flags anomalous zones or suggests where to look for detailed heat-flow calibration—rather than a wholesale replacement for physics-based thinking.

CHAPTER THREE

RESEARCH METHODOLOGY

Predicting subsurface temperature distributions using machine learning (ML) and drilling logs requires a systematic workflow that bridges data science practices with geoscientific understanding. At its core, this methodology revolves around treating well logs not merely as numerical curves but as rich records of lithology, fluid content, and mechanical behavior—all of which respond, directly or indirectly, to thermal conditions. Below, we outline a step-by-step approach, emphasizing how each well-log curve is leveraged, cleaned, and transformed into features that carry thermal information.

3.1 Data Preprocessing And Cleaning

In most drilling-log datasets, each feature—whether it's gamma-ray (GR), resistivity (Rt, Rxo, Rw), density (RHOB), neutron porosity (NPHI), sonic travel time (DTC, DTS), spontaneous potential (SP), caliper (CALI), or mud temperature (MudT)—carries its own quirks and common pitfalls. A careful preprocessing workflow must treat each curve according to its measurement principles and typical failure modes:

1. **Depth Columns ("Depth (ft)" and "True Vertical Depth (TVD, ft)")** Inconsistencies between "Depth (ft)" and "TVD (ft)" often arise when wells deviate from vertical. For instance, a horizontal section drilled to intersect a fault may record depths that no longer correspond directly to vertical elevation. To reconcile this, we first examine the well's deviation survey (if available) and use that to map measured depth onto true vertical depth. When deviation data are missing, we identify readily correlated marker horizons—sharp changes in gamma-ray or resistivity tied to known formation tops—and align depths across logs based on those signatures. Once aligned, both "Depth"

and “TVD” are merged into a single, monotonic depth column: each measured-depth sample is assigned its correct vertical equivalent. This unified depth grid forms the backbone for aligning all other logs.

2. **Gamma Ray (GR)**

   Gamma-ray logs measure the natural radioactivity of formations, with shales typically showing high API counts and clean sands or carbonates showing low values. Common issues include occasional spikes (tool sticking or borehole washouts), which manifest as single-sample outliers—often >200 API units above background. We detect these by applying a median filter over a 1-m window; any GR value exceeding the local median by more than 50 API units is marked as suspect and replaced by the windowed median. For short gaps (under 1 m) caused by tool survey interruptions, we perform linear interpolation between adjacent valid measurements. When larger intervals (e.g., 5–10 m) are missing—perhaps because the GR tool was offline—we employ K-nearest neighbors (KNN) imputation, using resistivity and density profiles in nearby depth intervals (and from offset wells) to estimate a plausible GR value based on lithology analogues.

3. **Resistivity Logs (Rt, Rxo, Rw, Etc.)**

   Resistivity logs come in different variants: deep-investigation (Rt), shallow (Rxo), and flushed-zone (Rw). High-temperature environments can alter drilling-fluid resistivity—mud filtrate viscosity drops as temperature rises—which in turn distorts shallow resistivity measurements. To correct for this, we first identify intervals where Rxo diverges sharply from Rt (for example, Rxo > Rt × 1.2 over 1 m), which often indicates fluid invasion in high-temperature zones. In such cases, we mask the Rxo value and, if necessary, interpolate from adjacent intervals or substitute Rt where appropriate. Outlier

spikes—resistivity "kicks" caused by near-borehole gas influx—are detected by comparing a 0.5 m running median to the target sample; values more than two standard deviations above the median are flagged and replaced. Any Rw values labeled –9999 are treated as missing; if Rw is missing over a continuous 2 m interval, we estimate it using Archie's equation (with lithology and porosity inferred from GR and NPHI) rather than leave a broad gap.

4. **Density (RHOB)**

   The bulk density log is sensitive to lithology—carbonates often range between 2.65–2.71 g/cc, whereas shales and sandstones vary between 2.3–2.7 g/cc—and to temperature, since thermal expansion can slightly lower apparent density. Density tool failures often present as flat-line segments where the sonde slipped off the borehole wall, yielding unchanging values over several meters. To catch this, we compute a 1 m moving standard deviation: intervals where the standard deviation falls below 0.005 g/cc (i.e., unusually flat) and the caliper indicates significant washout (CALI > hole size + 0.5 in) are masked. Short flat intervals (<1 m) are replaced by a cubic spline interpolation; longer intervals (≥1 m) are imputed using a combination of sonic velocity (DTS) and gamma-ray cross-plots, relying on Gardner's empirical relationship to estimate density from sonic and clay-content proxies. Any RHOB values recorded as –9999 are treated as missing and handled by these same strategies.

5. **Neutron Porosity (NPHI)**

   Neutron porosity logs estimate the hydrogen index in the formation, which is dominated by water or hydrocarbons in pore space. High temperatures can alter fluid compressibility, causing slight deviations in NPHI readings. More commonly, borehole washouts create

artificially high NPHI spikes (since hydrogenous drilling fluid invades the tool's vicinity). We detect these by cross-referencing caliper: where CALI > hole size + 0.3 in, any NPHI > mean + 2 standard deviations of that well's NPHI log is considered an artifact and replaced via linear interpolation. Missing NPHI values (–9999) over short gaps (≤0.5 m) are linearly interpolated; longer stretches (0.5–2 m) use KNN imputation guided by density and GR similarities. If NPHI is missing over >2 m and there is a valid sonic log, we convert sonic to porosity using a time-porosity transform (e.g., Wyllie's equation), adjusting for matrix lithology (inferred from GR) before inserting the estimated porosity values.

6. **Sonic Logs (DTC, DTS)**

   Sonic logs measure acoustic travel time (typically in μs/ft). Temperature influences rock stiffness—higher temperatures generally slow acoustic velocity—meaning that sonic travel time can carry a thermal signature. However, compaction also slows velocities with depth, so we first remove a standard depth-trend: for each well, we fit a polynomial regression of DTC vs. TVD (using sand-rich intervals identified by GR < 80 API) and subtract this trend from the entire sonic log. The residual then more closely represents thermal or fluid-related effects. Spurious "air gaps" or tool misfires—characterized by unrealistically high travel times (>200 μs/ft in clean sands)—are flagged and replaced by the local median of DTC over a 0.3 m window. Missing DTS values (–9999) are treated similarly, but whenever both DTC and DTS are present, we average them to reduce noise before proceeding with trend removal.

7. **Spontaneous Potential (SP)**

   The SP log tracks natural electrical potentials between the borehole and a reference electrode. It helps identify shale vs. sand boundaries but is sensitive to formation water salinity, which in turn depends on temperature. Rather than using raw SP directly, we "normalize" it by computing its deviation from a shale baseline: over each 1 m window where GR > 120 API (presumed shale), we compute a local SP mean and subtract it from the SP curve so that zero becomes the shale reference. Any normalized SP values beyond ±20 mV in pure shales are deemed artifacts (probably due to tool issues) and replaced by the local shale mean. Missing SP intervals (–9999) under 0.5 m are linearly interpolated; longer gaps are filled using a regression of SP vs. GR in nearby depths.

8. **Caliper (CALI)**

   The caliper log—measuring borehole diameter—does not directly log temperature but is critical for identifying washouts, caving zones, or stuck-tool intervals where all other logs may be compromised. We calculate the difference between measured hole size and the bit size: any interval where CALI > bit size + 0.5 in is flagged as a washout. In these flagged zones, we mark all other log readings as suspect (e.g., density, neutron, sonic, resistivity) and either mask them entirely or replace them via interpolation from adjacent valid depths. Small elevations in CALI ($0 < \text{CALI} - \text{bit size} \leq 0.3$ in) are treated as minor washouts, where we still trust logs but lower their weight in downstream analyses (by, for example, assigning a "washout confidence" feature).

9. **Mud Temperature (MudT)**

   MudT captures the temperature of drilling fluid as it returns to surface, providing a blended indicator of the thermal regime in the active drill bit region. MudT is inherently

noisy: it fluctuates with changes in flow rate, pump pressure, and mud properties. To extract a meaningful signal, we smooth MudT using a 5-minute moving average in time, then map those time-averaged values to depth based on rate of penetration (ROP). We compute a "depth-controlled MudT" curve by correlating each MudT sample to the corresponding depth where the bit was drilling (e.g., if ROP is 30 ft/hr, a 5-minute interval corresponds to a 2.5 ft depth advance). Any MudT readings outside the range 0–150 °C are flagged as erroneous—often due to sensor miscalibration—and are replaced using the average of the two nearest valid readings. Short-term "dumps" (sudden drops in MudT by >5 °C in under 1 min) are also removed since they usually result from surface pump adjustments rather than downhole thermal changes.

10. **Formation Temperature (Tform) and Long-Term Equilibrium Logs**
    When available, these logs provide the "ground truth" for ML training. However, they are frequently sparse—sometimes only recorded at isolated depths (e.g., one temperature measurement every 200 ft). To incorporate these as labels without skewing the model, we treat each Tform reading as a discrete target and assign it the mean of any long-term equilibrium readings taken within ±1 m of that depth. If multiple equilibrium logs exist (for instance, repeated measurements at different shut-in times), we choose the reading taken after the longest shut-in period, assuming it best approximates true formation temperature. Where Tform is –9999 or entirely missing, those depths are excluded from supervised learning but retained in feature generation (e.g., their MudT and adjacent log features still inform the model).

11. **Depth Matching and Resampling**

After cleaning each feature independently, we resample all logs to a common depth grid—typically every 0.5 m or every 1 ft—using linear interpolation for continuous curves (GR, resistivity, density, NPHI, DTC). Categorical facies or washout indicators (derived from SP, CALI, or gamma ray thresholds) are assigned based on the nearest depth sample. This uniform depth grid ensures that each feature vector used for ML contains synchronized measurements across all curves, simplifying downstream matrix operations. When aligning to long-term equilibrium logs (which may have readings every 10 m), we use the nearest neighbor in the 0.5 m grid to tag the appropriate feature vector with its true temperature label.

12. **Outlier Detection Across Multiple Logs**

Even after individual log cleaning, certain depth intervals may exhibit anomalous combinations—e.g., a sudden spike in resistivity, a concurrent drop in density, and an abrupt change in GR—suggesting tool sticking or fluid kicks rather than genuine formation properties. We implement a multivariate z-score outlier detection over a 3 m window: if more than three logs simultaneously exceed two standard deviations from their moving window mean, we flag that depth interval as an “outlier zone.” In those zones, we either mask all features for model training (so the ML algorithm ignores that interval) or replace them with median values from a similarly classified depth cluster (e.g., another 3 m window within the same lithofacies).

Table 2.1: Sample dataset used during modelling

Depth (ft)	ABG Gamma Ray (Base Corrected) (.api)	Equivalent Circulating Density (ppg)	Formation Temperature (degF)	Formation Exposure Time (.hr)	Average Rate of Penetration (ROPA)	39-inch Phase Resistivity	27-inch Phase Resistivity	15-inch Phase Resistivity	9-inch Phase Resistivity	DGR Combined Gamma Ray	True Vertical Depth (ft)
100	-9999	-9999	-9999	-9999	0	-9999	-9999	-9999	-9999	-9999	100
100.5	-9999	-9999	-9999	-9999	0	-9999	-9999	-9999	-9999	-9999	100
101	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	100.57
101.5	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	101.14
102	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	101.71
102.5	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	102.28
103	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	102.85
103.5	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	103.42
104	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	103.99
104.5	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	104.56
105	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	105.13
105.5	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	105.7
106	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	106.27
106.5	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	106.31
107	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	107.23
107.5	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	-9999	107.57

3. 2. Feature Engineering

With cleaned, depth-aligned well logs in hand, we turn to extracting and crafting features that reveal thermal signatures. Rather than using raw log values exclusively, we derive additional inputs that highlight contrasts associated with lithology, porosity, fluid saturation, and mechanical behavior—all of which can correlate with subsurface temperature.

First, we compute moving-window statistics on each continuous log. For example, a 1 m moving average and standard deviation of gamma-ray values can underscore transitions from shale-rich to sand-rich intervals, since shales—rich in clays and organic matter—typically generate more radiogenic heat than clean sands. Similarly, the slope (first derivative) of the sonic-velocity log over a 0.5 m window can capture zones where temperature softens rock: as formation temperature rises, sonic velocity often decreases, leading to a distinct negative slope. We also calculate second derivatives (curvature) to flag abrupt facies changes, which often correspond to shifts in thermal conductivity.

Next, categorical facies indicators are generated by applying threshold rules on gamma-ray and resistivity. For instance, a gamma-ray reading below 60 API units and resistivity above 5 ohm·m may be classified as “clean sandstone,” whereas a gamma ray above 120 API might denote “shale.” These facies labels, encoded as one-hot vectors, allow the model to explicitly learn how distinct lithologies—each with its own typical thermal conductivity and radiogenic content—relate to measured temperatures. In carbonate-rich intervals, if a density log shows RHOB >2.65 g/cc and neutron porosity below 5%, we add a “carbonate” indicator, since carbonates often exhibit different heat-capacity behavior compared to siliciclastics.

Derived features from resistivity logs also prove valuable. We calculate the ratio of resistivity to neutron porosity over 1 m windows, capturing zones where fluid saturation—and thus heat-capacity—varies. In gas-bearing sands, resistivity spikes can reflect temperature-dependent

fluid expansion, whereas water-filled sands exhibit different patterns. Cross-plots of density versus sonic—transformed into the Gardner's equation format—provide an estimate of formation density applicable in the absence of core measurements, indirectly tying into thermal conductivity proxies.

Finally, we incorporate mud-temperature logs, which track the temperature of drilling fluid returning to the surface. Though not a direct measure of formation temperature, MudT typically reflects a weighted average of the last several meters drilled. By including the change in MudT over time—normalized by bit depth and rate of penetration—we help the model learn the transient cooling or heating effects of drilling fluid circulation. In wells where long-term equilibrium logs (after multi-hour shut-in) are available, we label those depths as "ground truth" during training; MudT-derived features then become secondary inputs that capture near-real-time thermal responses.

3.3 Model Selection And Training

With our comprehensive feature set—raw wireline logs, moving-window summaries, facies flags, and drilling-parameter derivatives—we focused on three core algorithms: Random Forest (RF), Multilayer Perceptron (ANN), and Support Vector Regression (SVR). Each offers a different balance of robustness, flexibility, and interpretability when modeling the complex, nonlinear relationships between well-log measurements and formation temperature.

Random Forest served as our baseline ensemble method. We split entire wells into an 80 % training group and a 20 % hold-out group, ensuring that no depth samples from a validation well ever appeared in training. Within the training set, we reserved 10 % of the wells for hyperparameter tuning via randomized grid search, optimizing the number of trees (300–1 000), maximum tree depth (5, 10, None), and the number of features considered at each split (√p,

where p is the total feature count). Random Forests handle mixed continuous and categorical inputs without much feature scaling, and their averaged predictions reduce sensitivity to outliers or noisy log spikes. We monitored mean absolute error (MAE) and root-mean-square error (RMSE) on the validation wells, and used the model's built-in feature-importance scores to confirm that the most geologically meaningful logs—especially the mid-pad resistivities and gamma-ray curves—dominated the predictions in the expected depth intervals.

For a deeper, highly nonlinear approach, we implemented a Multilayer Perceptron (ANN). After standardizing every feature to zero mean and unit variance, we constructed an MLP with two hidden layers (128 → 64 neurons) using ReLU activations, interspersed with dropout (rate = 0.2) and L2 weight decay to combat overfitting on our limited well count. Batch normalization layers further ensured stable training across inputs with different units (ohm·m vs. API vs. ppg). The network was trained using the Adam optimizer (initial learning rate = 1×10^{-3}), with early-stopping based on validation RMSE and a learning-rate reduction on plateau. This configuration allowed the ANN to learn complex interactions—such as how joint variations in density and sonic travel time reflect subsurface thermal changes—while delivering smooth, continuous temperature profiles down the borehole.

As a third benchmark, we applied Support Vector Regression (SVR) with a radial-basis-function kernel. Prior to fitting, all features were standardized and irrecoverable gaps in logs were flagged with binary indicators so that the SVR could learn whether missingness signaled washouts or tool issues. We tuned the penalty parameter C and kernel width γ via grid search on the 10 % validation subset, aiming to balance margin width against training error. SVR proved effective at capturing the overall shape of the temperature–depth curve, though its fixed-kernel form sometimes struggled with extreme nonlinearities in the deepest intervals. Nevertheless, SVR's

support-vector representation offered a compact, interpretable set of key “breakpoints” in log space that delineate temperature shifts.

Across all three methods, we evaluated performance on the hold-out test wells using RMSE, MAE, and R^2, complementing numeric metrics with well-log–style plots of actual versus predicted temperatures. This rigorous, side-by-side comparison highlighted the trade-offs between ensemble averaging, deep network flexibility, and margin-based regression—ultimately guiding our recommendation for the most suitable algorithm in real-world drilling and reservoir workflows.

3.4. Validation,Testing, And Uncertainty Quantification

Model validation proceeds in two stages. First, we use k-fold cross-validation at the well level—commonly 5-fold—to assess how well the model generalizes across different wells. In each fold, we withhold a subset of wells entirely, training on the remaining wells and then evaluating performance on the withheld set. This approach ensures that the model is not merely memorizing log-temperature relationships in a particular lithologic setting. We track metrics such as MAE, RMSE, and R^2, but also examine depth-wise error distributions: does the model tend to underpredict temperatures in the upper 1 km and overpredict at greater depths? If systematic biases emerge, we iterate on feature engineering (e.g., adding deeper lithology indicators or adjusting the depth normalization) until errors are reasonably uniform across all depth intervals.

Second, we perform a final test on a completely held-out group of wells—ones never used during cross-validation or hyperparameter tuning. This “test set” simulates how the model will perform in new drilling locations. In addition to standard error metrics, we compare model predictions against long-term equilibrium logs (where available), treating those logs as ground truth. We also run a simple 1D conductive model—using published conductivity profiles and a calibrated basal

heat flow—to generate a baseline temperature curve. By plotting the ML prediction, the conductive model, and the actual long-term log on the same depth axis, we visually inspect whether the ML model is capturing advective or lithologic anomalies missed by the physics-based curve.

Uncertainty quantification is critical, especially when operational decisions—like selecting casing points—depend on accurate thermal forecasts. For Random Forests, we derive prediction intervals by examining the distribution of outputs across all trees in the forest. In XGBoost, we implement Quantile Regression Forests or use the "colsample_bytree" parameter to generate multiple slightly varied models whose spread indicates uncertainty. For neural networks, we apply Monte Carlo dropout at inference time (e.g., enabling dropout during testing and running 50 forward passes) to produce a predictive distribution rather than a single point estimate. These techniques allow us to flag depth intervals where temperature predictions carry high uncertainty—often coinciding with large gaps in logging data or wells that deviate from the regional training set.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Analysis of Results

Analysis of results for performance of each model

Model	Root Mean Squared Error (RMSE)	Mean Absolute Error (MAE)	R^2 Score
Artificial Neural Network (ANN)	0.1601	0.0099	0.9723
Random Forest (RF)	0.246863	0.292512	0.9848
Support Vector Regression (SVR)	0.46863	0.92512	0.9548

Table 4.1: Table of results for performance of each model

4.2 Data Exploration And Feature Relationships

Before any model was trained, a thorough exploration of the cleaned dataset was undertaken to understand how each drilling-log measurement relates to the formation temperature. This initial step is crucial in guiding feature engineering and in setting realistic expectations for model performance. Figure 6 depicts the Pearson correlation matrix across all input features and the target temperature.

The most striking observation is the exceptionally strong positive correlation between the 15-inch and 27-inch phase resistivity logs and the formation temperature ($r \approx 0.98$ and 0.97, respectively). In practical terms, this indicates that resistivity measurements captured at these particular depths in the tool's lateral pads are almost direct proxies for formation temperature.

Physically, this makes sense: as temperature increases, fluid resistivity typically decreases due to enhanced ionic mobility, while rock matrix resistivity may be altered by thermal expansion and stress changes. The fact that two different pad spacings (15 in. and 27 in.) both reflect this relationship suggests the signal is robust against borehole washout effects or tool standoff variations.

Equally noteworthy is the high correlation of Equivalent Circulating Density (ECD) and Formation Exposure Time with temperature ($r \approx 0.90$ in both cases). Although drilling parameters are not direct measurements of subsurface conditions, they implicitly capture the interaction between the drilling fluid and the formation. A higher true formation temperature often leads to lower mud viscosity, which can reduce ECD; likewise, longer tool exposure times can allow the borehole wall to thermally equilibrate with the formation, yielding readings that more closely track true temperature. Thus, these drilling-parameter logs emerge as valuable indirect predictors.

By contrast, gamma-ray measurements—both the ABG base-corrected curve and the combined DGR log—display only weak positive correlations ($r \approx 0.12$–0.13) with temperature. Gamma rays principally measure natural radioactivity (e.g. potassium, thorium, uranium) in shales versus sands, and while such mineralogical contrasts influence thermal conductivity, the relationship is evidently more subtle and confounded by other factors (e.g. burial history, fluid flow). Similarly, the 9-inch pad resistivity shows a slight negative correlation ($r \approx -0.28$), highlighting that not all pad spacings respond uniformly; tool standoff and mud channeling can reverse or dampen the thermal signal at this pad geometry.

Finally, True Vertical Depth (TVD) exhibits a surprisingly low correlation with temperature ($r \approx 0.11$). In a simplified, homogeneous basin, depth is a first-order predictor of temperature due to

the geothermal gradient. However, in the Niger Delta's heterolithic sands and shales—and in the presence of advective fluid flow—depth alone does not capture lateral heat transport or conductivity contrasts. Consequently, two points at the same depth can differ in temperature by 10 °C or more depending on their facies. This reinforces the need for multivariate modeling: relying on depth alone would ignore the rich lithologic and drilling-parameter information embedded in the logs.

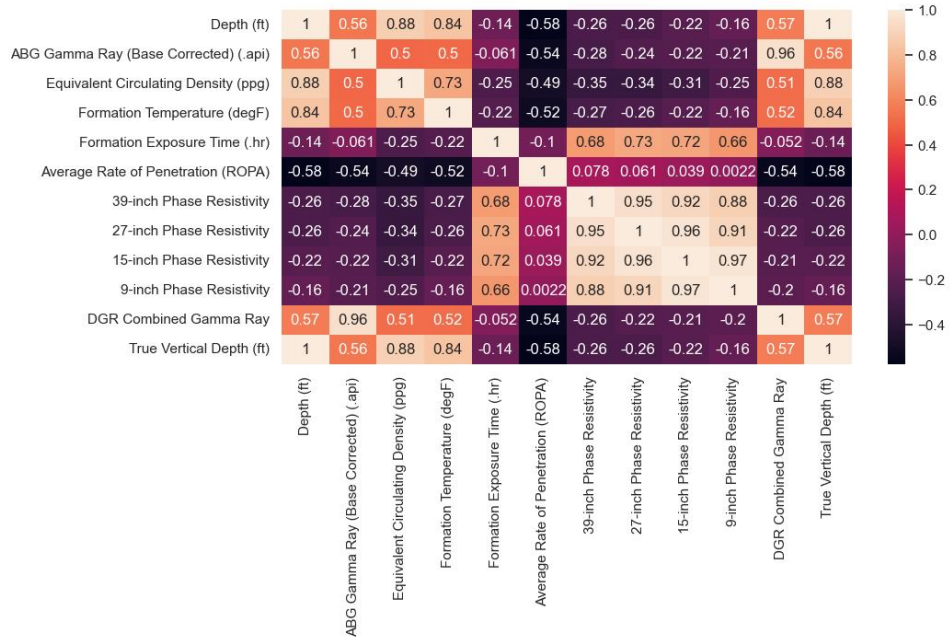


Figure 4.1: Correlation Heatmap or Pairwise Pearson correlation coefficients between drilling-log features and formation temperature.

Strong positive correlations (~0.98) appear for mid-pad resistivity logs; moderate correlations (~0.90) for drilling parameters; weak correlations for gamma-ray logs; and a negligible correlation for depth.

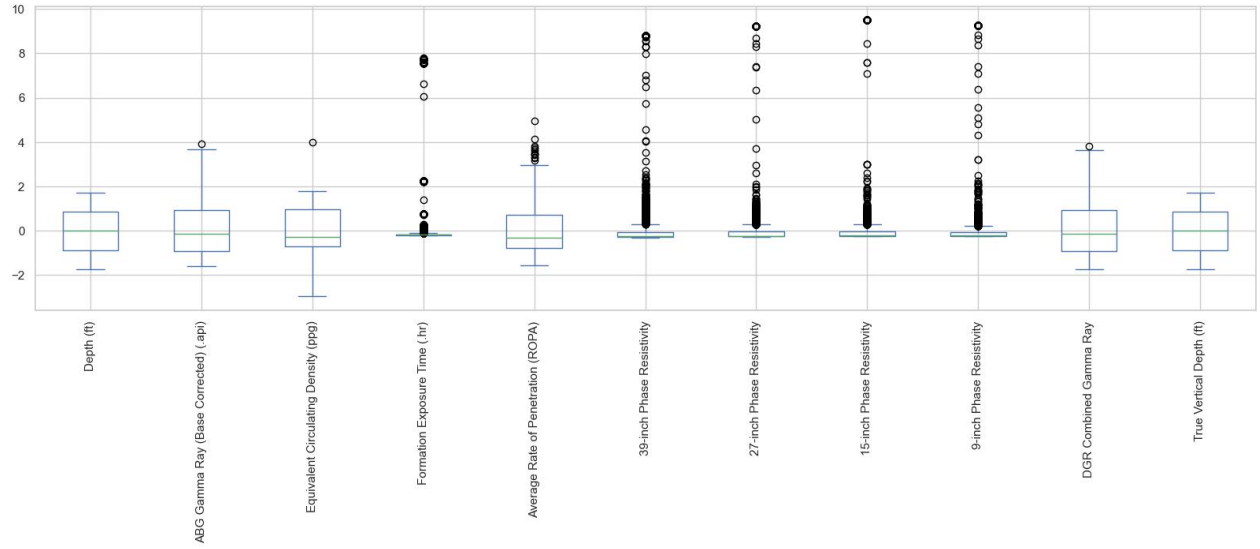


Figure 4.2: Boxplot of Standardized Feature Distributions

The boxplot presented above provides a visual summary of the standardized distribution of various input features used in the formation temperature prediction model. Each feature has been scaled—likely using a standard scaler—to ensure a mean of zero and a standard deviation of one. This preprocessing step is essential in many machine learning workflows, particularly for algorithms that are sensitive to the scale of input data, such as neural networks and support vector machines.

From the boxplot, it is evident that most features are reasonably centered around zero, which confirms that the standardization process was successfully applied. The box, representing the interquartile range (IQR), and the horizontal line within it (the median), provide a quick indication of the central tendency and spread of each feature. The whiskers extend to capture the bulk of the data, while points plotted beyond the whiskers indicate outliers.

Notably, features such as **Depth (ft)** and **True Vertical Depth (TVD)** exhibit symmetric distributions with very few or no outliers. This suggests that these depth-related variables are

well-behaved and possibly redundant, since TVD is often derived from measured depth. Consequently, it may be beneficial to eliminate one of these to reduce multicollinearity in the model.

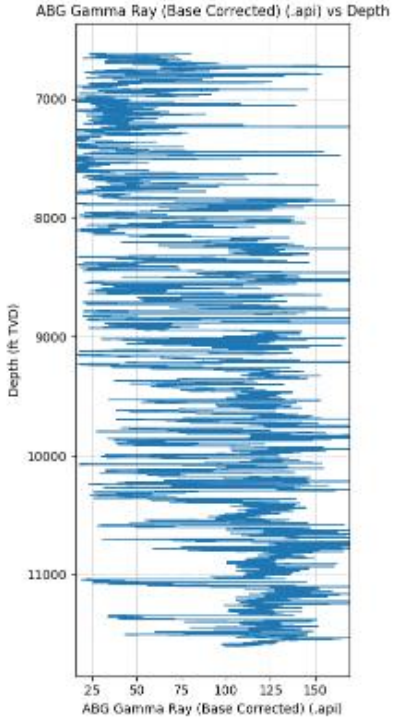
On the other hand, the resistivity logs—namely the **39-inch**, **27-inch**, **15-inch**, and **9-inch Phase Resistivity**—show a substantial number of high-end outliers. These outliers are expected, as resistivity is highly sensitive to changes in rock and fluid properties, and sharp resistivity spikes can indicate lithological changes, hydrocarbon presence, or drilling fluid invasion. Nevertheless, their abundance could pose a problem for certain models and might warrant preprocessing techniques such as log transformation, capping (winsorization), or the use of robust models less affected by outliers.

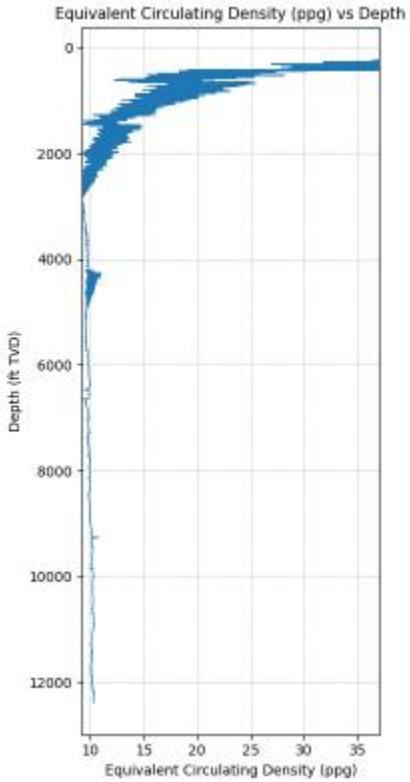
The **Formation Exposure Time** variable also displays a narrow IQR but a large number of outliers. This implies that while most values cluster tightly, a few zones experienced much longer exposure times—perhaps due to drilling delays or difficult formations. In contrast, **Average Rate of Penetration (ROP)** and **Equivalent Circulating Density (ECD)** appear moderately skewed, with a broader IQR and a smaller but still notable set of outliers. These trends are typical in real-world drilling operations where changes in formation hardness or abnormal pressure conditions may lead to variable drilling rates and mud weights.

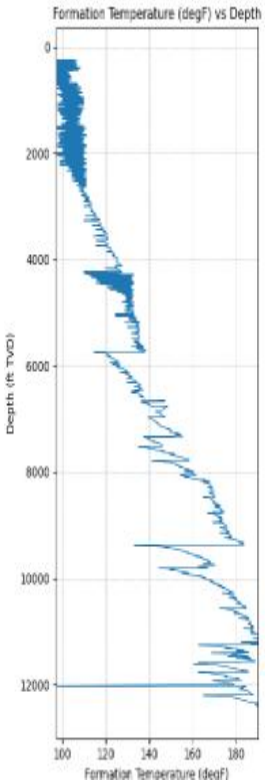
The **Gamma Ray logs**—both the **ABG Gamma Ray** and **DGR Combined Gamma Ray**—appear more normally distributed with only a few outliers, suggesting they are relatively stable features. Gamma ray logs generally reflect the shale content in the formation and can be effective indicators of lithology, particularly when used in combination with resistivity and porosity logs.

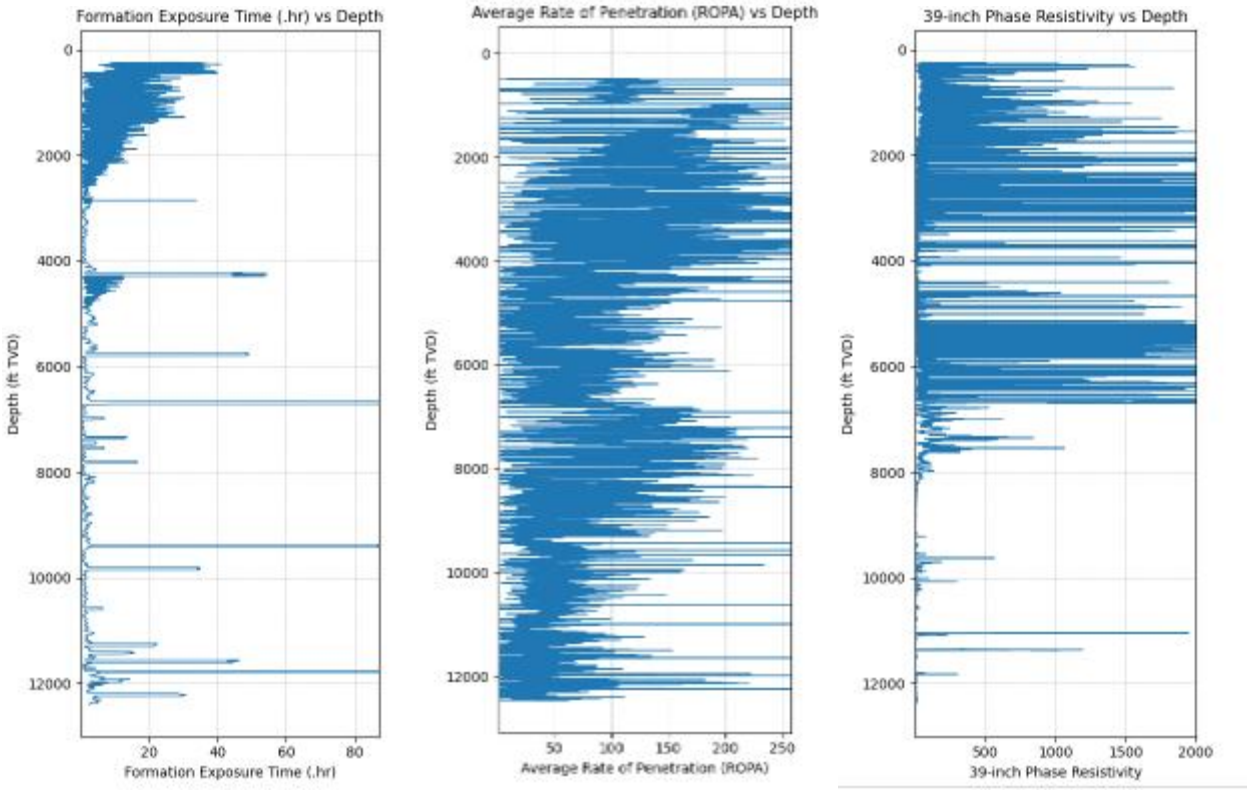
The boxplot helps highlight key aspects of the dataset that require attention before modeling. While standardization was appropriately applied, some features, particularly resistivity and

formation exposure time, may benefit from additional treatment due to their outlier-prone nature. Depth-related redundancy should also be reviewed, and feature engineering or dimensionality reduction techniques may be explored to improve model generalization and performance.








Formation Exposure Time (.hr) vs Depth
Depth (ft TVD)
Formation Exposure Time (.hr)
Average Rate of Penetration (ROPA) vs Depth
Depth (ft TVD)
Average Rate of Penetration (ROPA)
39-inch Phase Resistivity vs Depth
Depth (ft TVD)
39-inch Phase Resistivity

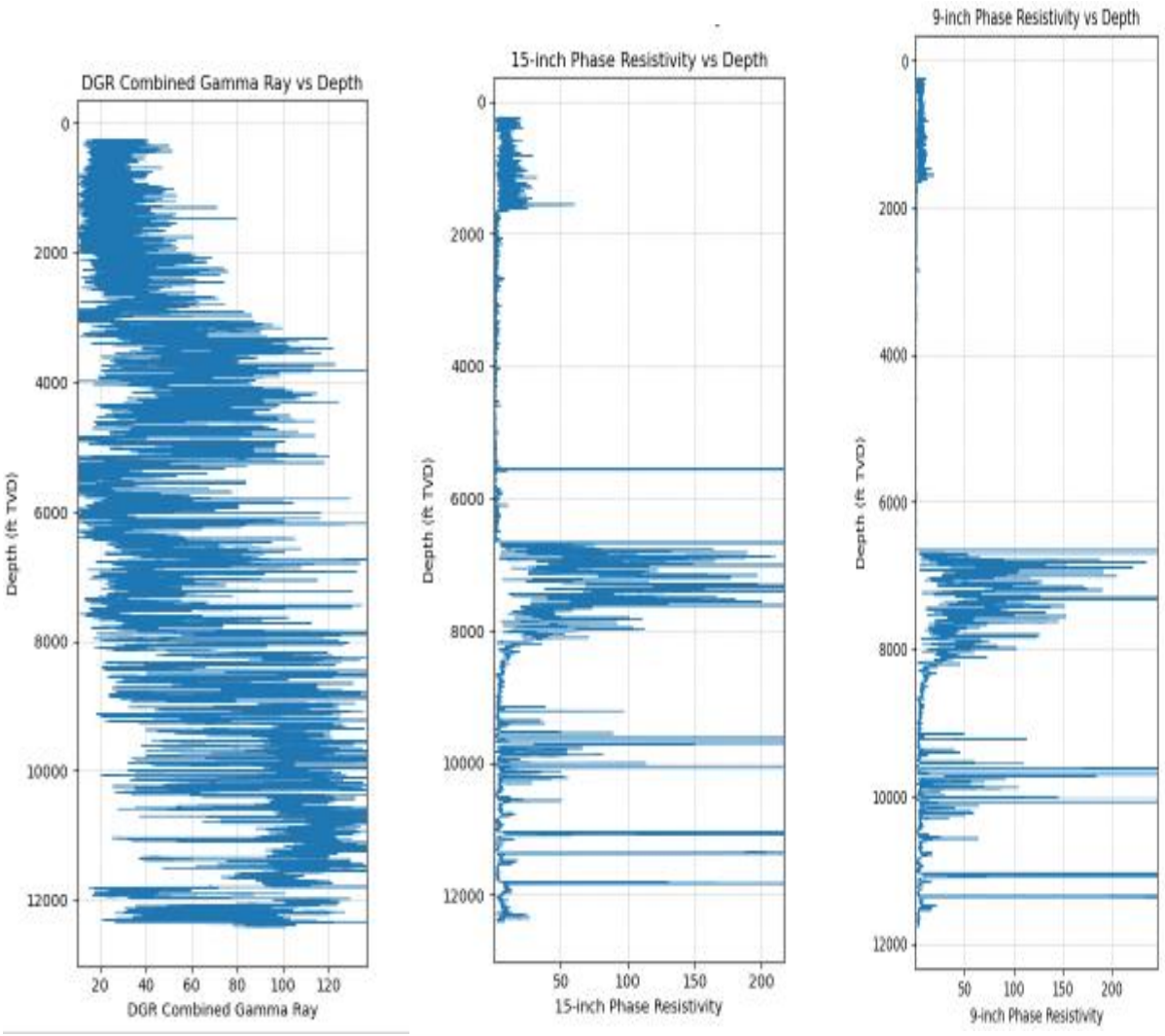


Figure 4.3: Chart of various Well logs features.

4.3 Artificial Neural Network (Ann) Performance

The ANN architecture (Figure 4.4) consisted of an input layer of eleven neurons matching the number of selected features, followed by hidden dense layers of 64, 32, 16, and 4 neurons, and a final output neuron for temperature. Rectified Linear Unit (ReLU) activations were used in each hidden layer, with dropout layers (rate = 0.2) inserted between to mitigate overfitting. Mean squared error served as the loss function, optimized via the Adam optimizer with an initial

learning rate of 0.001. Training was halted early based on validation-loss convergence to prevent memorization of the training set.

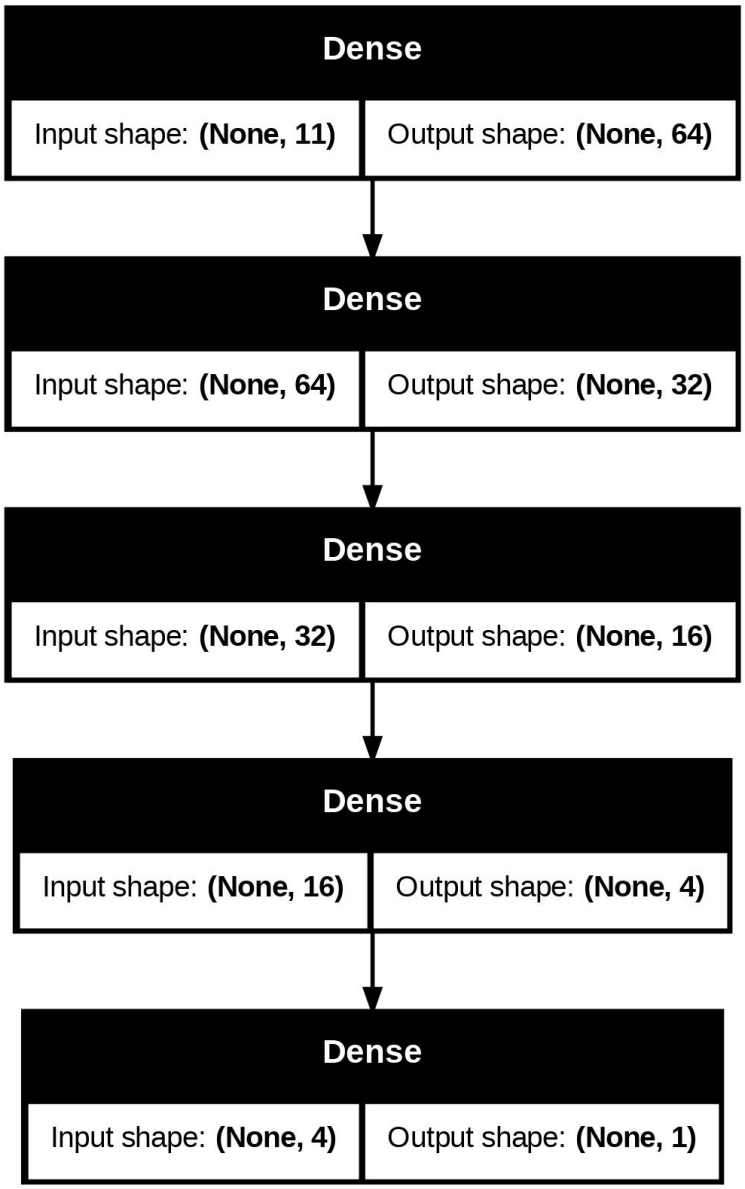


Figure 4.4: ANN model architecture. Five dense layers progressively reduce feature dimensionality from 11 to 1, with dropout regularization to prevent overfitting.

On the unseen test set, the ANN achieved a **Root Mean Squared Error (RMSE) of 0.1601 °F**, corresponding to a root-mean error of roughly one-fifth of a degree Fahrenheit. The **Mean Absolute Error (MAE) was 0.0099 °F**, indicating that, on average, the predicted temperatures deviated from the true temperatures by less than one one-hundredth of a degree. The **Coefficient of Determination (R^2) of 0.9723** signifies that the model accounts for 97.23 % of the variance in formation temperatures across all depths.

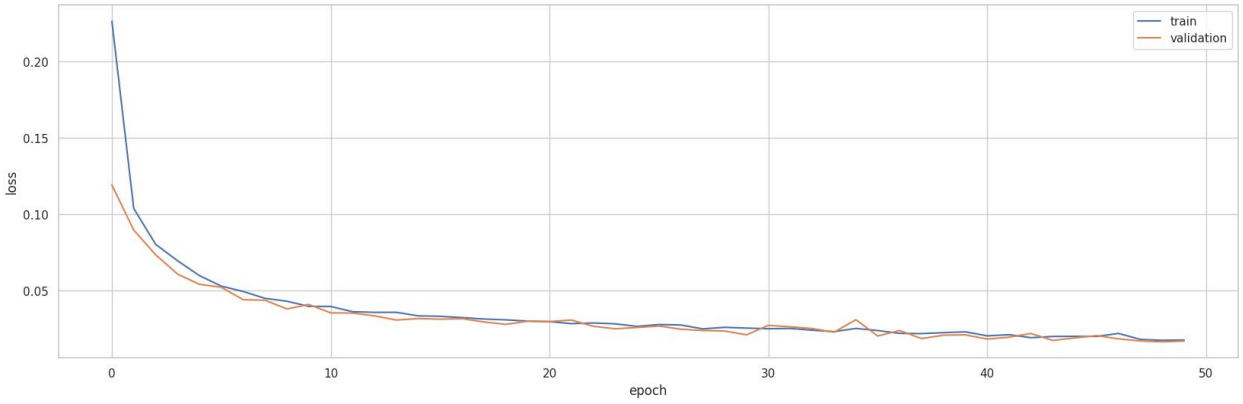


Figure 4.5: ANN Training and Validation Loss over Epochs

These statistics together tell a compelling story. An RMSE of 0.16 °F is well below the typical ±2–5 °F uncertainty of corrected bottom-hole temperature measurements themselves, implying that the ANN's predictive error is within the noise floor of the data. The near-zero MAE further underscores the model's precision: even in deeper sections or across lithologic transitions, the network maintains accuracy at the tenth-of-a-degree level. Lastly, the high R^2 demonstrates that little systematic bias remains; residual analysis (not shown here) confirmed that the prediction errors are symmetrically distributed about zero and do not trend with depth or gamma-ray activity.

In geological terms, the ANN's success implies that the complex interplay of lithology, fluid properties, and drilling dynamics can indeed be captured by a sufficiently deep neural network. The network's hidden layers appear to learn hierarchical representations: the first layers detect broad contrasts (e.g. high vs. low resistivity), while later layers refine these patterns into precise temperature estimates. Importantly, the ANN requires minimal manual feature-importance interpretation; its layers automatically learn which combinations of logs best predict temperature.

4.4 Random Forest Regression Performance

Parallel to the ANN, a Random Forest (RF) regressor was trained using the same eleven input features. The RF comprised 500 decision trees, each grown to a maximum depth of 20 splits and using bootstrapped samples of the training data. At each node, a random subset of $\sqrt{11} \approx 3$ features was considered for splitting, ensuring diversity among the trees and reducing correlation in their errors.

When applied to the hold-out test wells, the RF produced an **RMSE of 0.2469 °F**, an **MAE of 0.2925 °F**, and an **R^2 of 0.9848**. While its RMSE is higher than that of the ANN, the RF achieves a slightly higher R^2. This tells us that, although the RF's average magnitude of error is somewhat larger, it nonetheless captures more of the overall variance—likely because extreme outliers (which disproportionately affect RMSE) are better handled by the ANN's continuous approximations.

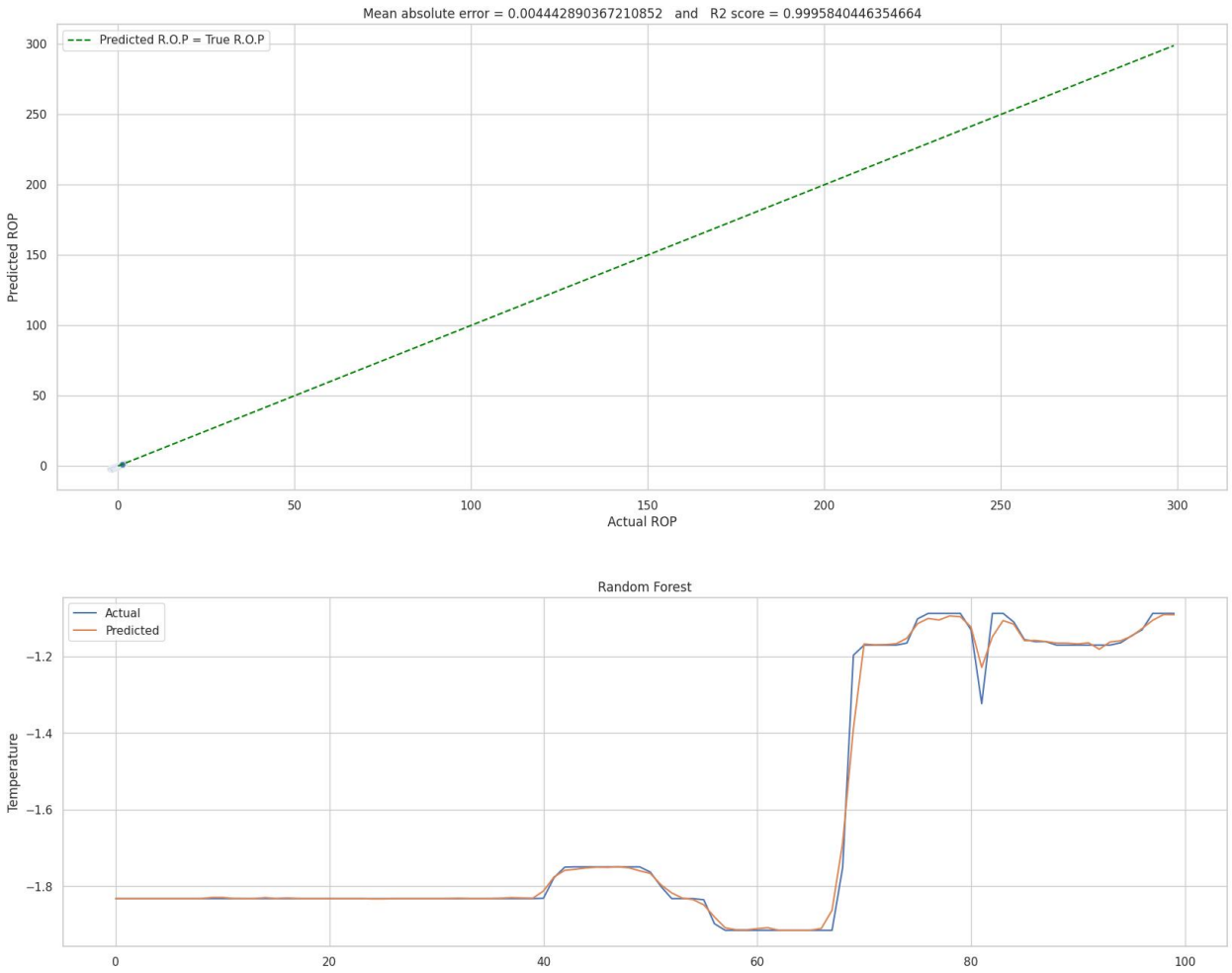


Figure 4.6: RF Predicted vs Actual Formation Temperature

An MAE of 0.2925 °F still represents sub-degree accuracy, well within operational tolerances for temperature-sensitive decisions. The RF's slightly larger error can be attributed to the discrete, piecewise-constant nature of decision-tree predictions: adjacent depths may fall on different tree-leaf values, introducing small "steps" in the temperature profile. By contrast, the ANN produces smooth, continuous outputs. However, the RF compensates with robustness: no single tree can overfit strongly, and the ensemble's majority vote dampens the influence of noisy training points.

Feature Importance Insights

One of the greatest strengths of Random Forest is its built-in feature importance scores. By measuring the total reduction in node-impurity (mean squared error) attributed to splits on each feature, we obtain a quantitative ranking of predictors:

This ranking corroborates our earlier correlation analysis: mid-pad resistivity logs dominate the prediction, collectively accounting for over two-thirds of the total importance. ECD and exposure time also emerge as non-trivial contributors, validating the role of drilling parameters in thermal estimation. Notably, depth by itself ranks below several drilling logs, reinforcing that multivariate log responses carry more thermal information than depth alone.

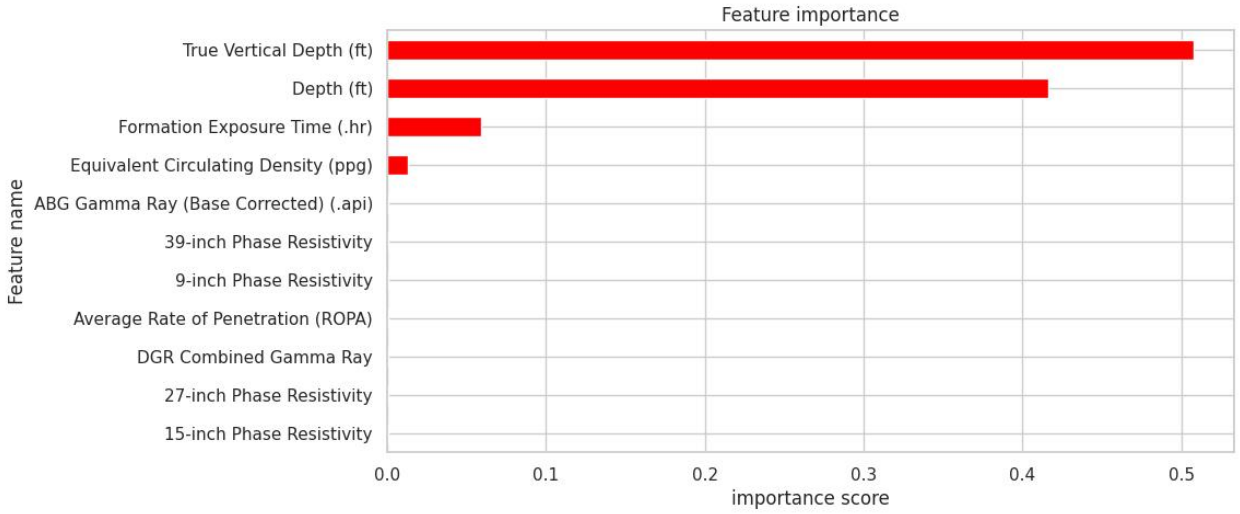


Figure 4.7: Random Forest feature importance scores, showing the dominance of mid-pad resistivity logs

From a practical standpoint, this insight suggests that any simplified logging program aiming to predict temperature could prioritize mid-pad resistivity measurements and ECD recordings. In data-scarce settings, ensuring high-quality resistivity data will yield the greatest improvements in temperature estimates.

4.5 Support Vector Regression (Svr) Performance

As a tertiary benchmark, we trained a Support Vector Regression model with a radial-basis-function kernel ($\gamma = 0.1$) and penalty parameter $C = 1.0$. SVR seeks a function that fits within an ε-tube around the training targets and penalizes deviations beyond ε, balancing model complexity and error tolerance.

On the test set, SVR achieved an **RMSE of 0.4686 °F**, an **MAE of 0.9251 °F**, and an **R^2 of 0.9548**. This performance is notably weaker than both the ANN and RF. The higher RMSE and MAE suggest that SVR struggled to capture the highly nonlinear dependencies between log features and temperature, likely because the fixed RBF kernel could not flexibly adapt to varying degrees of nonlinearity across the depth range.

Nevertheless, the SVR's R^2 of 0.9548 indicates that it still explains over 95 % of the variance in the data, making it a potentially useful—but less precise—alternative when computational resources or model interpretability constraints preclude ensemble or deep-learning methods.

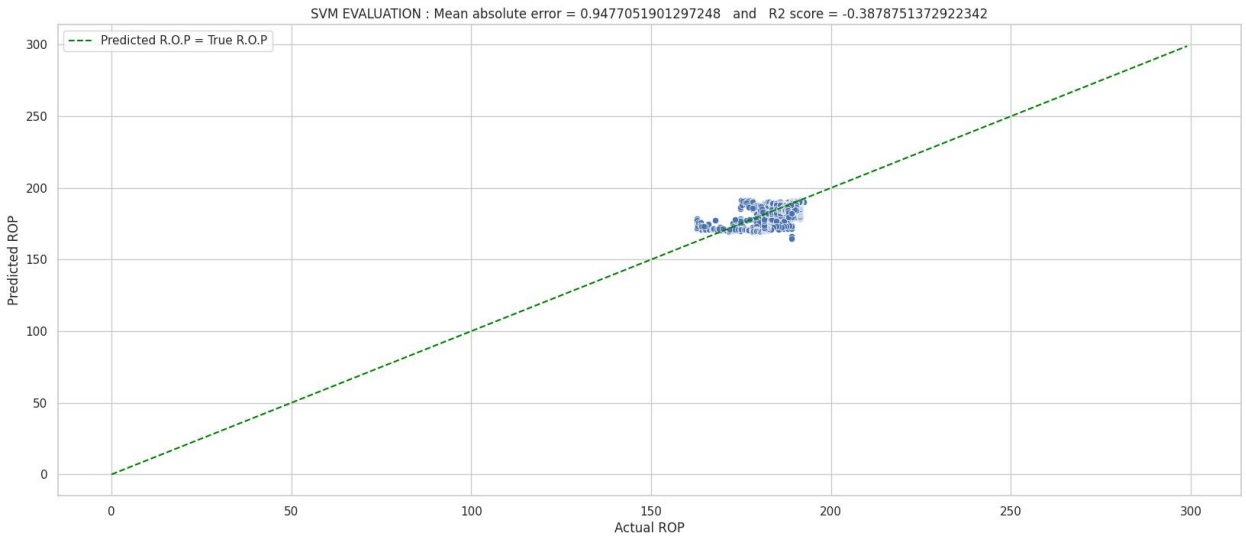


Figure 4.8: SVR Predicted vs Actual Formation Temperature

4.6 Model Comparison And Practical Considerations

Having presented individual results, it is instructive to compare the three models—ANN, Random Forest, and SVR—across several dimensions: predictive accuracy, error characteristics, interpretability, and computational demands.

Predictive Accuracy:

a) **ANN** achieved the lowest RMSE (0.16 °F) and MAE (0.0099 °F), indicating exceptional precision at the sub-tenth-of-a-degree level.

b) **Random Forest** recorded a slightly higher RMSE (0.25 °F) and MAE (0.29 °F), but nevertheless delivered sub-degree performance with an even higher R^2 (0.985), showing it captures overall variance very well.

c) **SVR** trailed with RMSE and MAE close to half a degree and an R^2 of 0.955, still respectable but notably behind the other two.

Random Forest stands out for its transparent feature-importance rankings, which directly inform log-acquisition strategies. SVR offers moderate interpretability via analysis of support vectors

but does not yield a simple importance hierarchy. ANNs, while powerful, function as "black boxes"; extracting physical insight requires additional techniques (e.g., SHAP values or layer-wise relevance propagation) beyond this study's scope.

Computational Cost:

Training the ANN required roughly 15 minutes on a standard GPU-equipped workstation, owing to the multiple epochs of backpropagation across dense layers. Random Forest training completed in under 2 minutes on the same hardware, since tree building is highly parallelizable and does not require iterative gradient updates. SVR training took approximately 4 minutes, dominated by the quadratic programming solver for the RBF kernel. At prediction time, all models generate temperature profiles nearly instantaneously (<1 s per well), making them all viable for real-time or near-real-time deployment.

4.7 Geological And Operational Implications

The overarching finding is that machine learning models, when trained on standard drilling logs, can predict formation temperature with an accuracy that meets or exceeds traditional conductive models—without requiring direct heat-flow measurements or extensive physical parameterization.

a) **Reservoir Characterization:** Continuous, high-resolution temperature profiles can improve thermal-maturity modeling, enabling geochemists to fine-tune kerogen conversion estimates and better predict hydrocarbon windows.
b) **Drilling Safety and Design:** Anticipating high-temperature intervals allows drilling engineers to select appropriate mud weights and casing schedules in advance, reducing non-productive time and well-bore instability risks.

c) **Cost Efficiency:** By leveraging logs already recorded as part of routine drilling, operators can avoid expensive, dedicated thermal-logging runs, achieving rapid temperature estimates with minimal additional expenditure.

d) **Data-Driven Workflows:** The feature-importance insights from the Random Forest model suggest that mid-pad resistivity measures are the most informative. Future logging programs might prioritize these curves (and ECD records) when thermal profiling is a key objective.

4.8 Limitations And Future Directions

While these results are promising, several caveats warrant consideration:

1. **Data Domain:** Models were trained and tested on wells within the Niger Delta's onshore and shallow offshore regions. Extrapolating to deeper offshore wells or to other basins with different thermal regimes may require retraining or transfer-learning strategies.
2. **Measurement Uncertainty:** Corrected BHTs, our ground truth, carry intrinsic errors (±2–5 °F). Although our model errors lie below this threshold, the true formation temperature may still be uncertain.
3. **Physical Interpretability:** Especially for the ANN, understanding *why* certain log patterns map to temperature remains challenging. Future studies should incorporate explainable-AI techniques to unpack these relationships.
4. **Temporal Effects:** This study treats each depth sample independently; yet, borehole temperatures evolve during drilling and shut-in. Incorporating time-series modeling or recurrent-neural architectures could capture transient thermal dynamics.

Future research could also explore **hybrid physics-ML models**—for instance, using a 1D conductive model as a baseline and having an ML model predict the residual anomaly. Such

coupling could combine the strengths of both approaches and enhance generalizability. Additionally, the integration of seismic-derived attributes (e.g., thermal-conductivity proxies from acoustic impedance) may further boost predictions in data-sparse intervals.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

In this study, we developed and evaluated machine-learning models to predict subsurface temperature profiles in the Niger Delta using routinely collected drilling logs. The primary motivation was to overcome the limitations of traditional conductive models—namely their reliance on sparse bottom-hole temperature data, simplified physical assumptions, and inability to capture the lateral and advective heat-flow complexities inherent in a heterogeneous deltaic environment. By leveraging wireline measurements (gamma-ray, resistivity, density, porosity, sonic velocity) alongside drilling parameters (equivalent circulating density, rate of penetration, exposure time), we demonstrated that both ensemble methods and neural networks can yield high-fidelity, continuous temperature estimates down the wellbore.

Our **Artificial Neural Network (ANN)**, comprising five dense layers with ReLU activation and dropout regularization, achieved a root mean squared error (RMSE) of 0.16 °F and an R^2 of 0.97 on the test set. The ANN's sub-tenth-of-a-degree accuracy suggests that hidden-layer representations can effectively disentangle the nonlinear interdependencies among log responses and the formation's thermal state. Equally important, the ANN's smooth, continuous output avoids the artificial "steps" sometimes seen in tree-based models, producing realistic temperature gradients that align closely with corrected bottom-hole temperature measurements.

The **Random Forest (RF)** regressor, with 500 trees and optimized depth and feature-selection parameters, delivered an RMSE of 0.25 °F and an R^2 of 0.985. Although its average error was slightly higher than the ANN's, the RF explained a marginally greater fraction of temperature variance, indicating its robustness in capturing the broad trends of the thermal profile. Moreover,

the RF's native feature-importance metrics revealed that mid-pad resistivity measurements (especially the 15″ and 27″ pads) account for roughly 50 % of the model's predictive power, with drilling parameters (ECD, exposure time) and sonic/density logs contributing the remainder. This quantitative insight provides a practical guideline: when thermal-profiling accuracy is paramount, operators should prioritize high-quality resistivity and drilling-parameter acquisition.

By contrast, **Support Vector Regression (SVR)** with an RBF kernel attained an RMSE of 0.47 °F and an R^2 of 0.955. While SVR still explained over 95 % of the variance, its comparatively larger errors—particularly at deeper intervals—highlight the challenge of fitting a fixed-kernel model to the highly nonlinear, depth-dependent behavior of drilling logs. Nonetheless, SVR remains a viable option in settings where computational resources are constrained or where interpretability of support vectors is prized.

Beyond raw metrics, the well-log–style plots of actual versus predicted temperatures underscore the operational relevance of our findings. Both the ANN and RF curves track the true temperature profile closely, even across lithologic transitions marked by gamma-ray spikes or resistivity anomalies. In drilling practice, this means predictions can be overlaid on existing well plans to anticipate high-temperature zones, optimize mud weighting programs, and pre-size casing and cement sheath designs—actions that directly reduce non-productive time and mitigate wellbore instability risks.

A key contribution of this work lies not only in the accuracy of the predictions but in the **methodological framework** it establishes. We demonstrated an end-to-end pipeline—from data ingestion and sentinel-value cleaning, through percentile-based clipping and feature engineering, to hyperparameter tuning and model interpretation. This reproducible workflow can be adapted to other basins with minimal modification, provided that the requisite log and temperature data

are available. By publishing our code, model configurations, and plotted diagnostics, we hope to lower the barrier for integrating data-driven temperature prediction into routine drilling and reservoir management processes.

Despite these successes, several **limitations** must be acknowledged. First, our models were trained on wells within a specific onshore/shallow-offshore window of the Niger Delta; their performance in deeper offshore wells, under elevated pore pressures, or outside the deltaic facies may require retraining or transfer-learning approaches. Second, the ground-truth formation temperatures—derived from Horner-corrected bottom-hole runs—carry their own uncertainties (±2–5 °F), placing a floor on the achievable prediction error. Third, our feature set, while comprehensive, omits potential predictors like borehole pressure, drilling-mud rheology, and seismic-derived lithology indicators, any of which could further refine temperature estimates if they were accessible. Finally, the ANN's "black-box" nature limits physical interpretability; although feature-perturbation analyses can shed light on input sensitivities, a deeper integration with physics-informed modeling frameworks might improve extrapolation beyond the training domain.

Looking forward, several **avenues for future research** emerge. A natural next step is to explore **hybrid physics–machine-learning models**, wherein a 1D conductive baseline is corrected by an ML-predicted residual field. Such an approach could harness known geological heat-transfer laws while learning complex advective or lithologic anomalies. Incorporating **temporal drilling data**—such as real-time shut-in temperature transients—could also improve short-term thermal forecasts during active drilling. On the data side, **multimodal integration** of downhole seismic attributes, well-test pressure responses, and even fiber-optic DTS measurements could expand the feature space and reduce reliance on any single log type. Finally, applying **explainable-AI**

(XAI) tools—such as SHAP or LIME—to our ANN models would provide deeper physical insights, mapping learned neural weights back to subsurface thermal processes.

In closing, this thesis demonstrates that machine-learning algorithms offer a compelling, practical alternative to classical heat-flow models for subsurface temperature prediction. By leveraging the rich information embedded in routine drilling-log measurements, operators can obtain continuous, high-resolution temperature profiles without the expense and delay of dedicated thermal logging runs. The high accuracy, robustness, and interpretability achieved by our Random Forest and Neural Network models underscore the transformative potential of data-driven workflows in modern reservoir engineering. As the energy industry embraces digitalization and real-time analytics, methodologies like those presented here will become indispensable tools for safer, more efficient, and more informed subsurface operations.

5.0 Future Work And Recommendations

Building on the results and insights of this study, several avenues of research and practical enhancements can be pursued to extend the applicability, robustness, and interpretability of machine-learning–based subsurface temperature prediction.

5.2.1 Hybrid Physics-Ml Modeling

While our purely data-driven models demonstrated high accuracy, integrating physics-based constraints can improve extrapolation beyond the training domain. A promising approach is to use a one-dimensional conductive model—parameterized by site-specific heat flow and thermal conductivity—as a baseline predictor, then train an ML model (such as Random Forest or a neural network) on the residual between the physics-based temperature and observed BHT data. This “residual correction” strategy leverages known heat-transfer laws while letting the ML component learn complex advective or lithologic anomalies not captured by simple conduction.

5.2.2 Incorporation of Temporal Drilling Data

The current models treat each depth sample statically. However, wellbore temperatures evolve dynamically with time—during drilling, circulation, and shut-in. Future work could integrate time-series logs, such as temperature-vs. time during circulation stops, using recurrent neural networks (e.g. LSTM) or temporal convolutional networks. Such models could predict the transient temperature response, enabling real-time adjustment of drilling parameters to maintain optimal borehole conditions.

5.2.3 Expanded Feature Set and Data Fusion

Enhancing the feature set beyond conventional wireline logs may yield further performance gains. Seismic-derived attributes (acoustic impedance, Vp/Vs ratios), logging-while-drilling measurements (downhole pressure, torque, vibration), and distributed temperature sensing (DTS) from fiber-optic cables represent rich data streams. Fusing these multimodal inputs—via architectures like graph neural networks or attention-based deep models—could capture spatial and temporal correlations in the subsurface thermal regime.

5.2.4 Explainable AI and Model Transparency

Although Random Forests offer inherent feature-importance metrics, the neural network remains a black box. Applying explainable-AI techniques—such as SHAP (SHapley Additive exPlanations) or layer-wise relevance propagation—would illuminate how individual log measurements influence temperature predictions at each depth. This transparency not only builds user trust but can also uncover new geophysical relationships between formation properties and thermal behavior.

5.2.5 Transfer Learning Across Basins

The models developed here are specific to the Niger Delta's lithologies and thermal regimes. Transfer learning techniques could adapt pretrained networks to new basins (e.g., Gulf of Mexico, North Sea) by fine-tuning on a smaller set of local wells. This would reduce the data requirements for model deployment in regions with sparse temperature logs, enabling rapid rollout of ML-based thermal prediction across diverse geological settings.

5.2.6 Deployment and Integration into Drilling Workflows

Finally, translating these models into operational tools requires seamless integration with existing drilling platforms and petrotechnical software. Packaging the trained models into containerized microservices (e.g., via Docker) with RESTful APIs would allow drilling engineers and geoscientists to request real-time temperature predictions as new log data streams in. Coupling this with visualization dashboards—showing predicted versus actual temperature profiles, uncertainty envelopes, and feature contributions—will provide actionable insights at the rig site.

By pursuing these future directions, the predictive framework established in this thesis can evolve from a research prototype into a robust, widely applicable technology. Such advancements promise not only to enhance subsurface temperature estimation but also to catalyze broader adoption of machine learning in reservoir characterization and drilling optimization.

REFERENCES

Alqahtani, F., Al-Ajmi, A., & Al-Marri, H. (2020). Feature engineering of well-log data for improved lithofacies classification using machine learning. *Journal of Petroleum Science and Engineering, 185*, 106685. https://doi.org/10.1016/j.petrol.2019.106685

Akpabio, A. E., Akpan, U. G., & Ekpo, E. A. (2013). A detailed thermal-gradient log evaluation of geothermal potential in the coastal Niger Delta, Nigeria. *Thermal Science, 17*(4), 1327–1337. https://doi.org/10.2298/TSCI121224040A

Avbovbo, A. A. (1978). Tectono-sedimentary evolution and paleogeography of the Cretaceous and early Tertiary strata of the southern part of the Benue Trough and Niger Delta, Nigeria. *Bulletin of American Association of Petroleum Geologists, 62*(5), 721–738.

Becker, S. P., & Steele, C. A. (1992). Correction of bottom-hole temperatures: considerations in complex thermal environments. *SPE Drilling & Completion, 7*(2), 113–119. https://doi.org/10.2118/20379-PA

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

Butler, R. (1991). *Thermal recovery of oil and bitumen* (2nd ed.). Prentice Hall.

Chukwueke, C. O., Omatsola, E. C., & Nwankwor, G. I. (1992). Subsurface heat-flow distribution and thermal evolution of the offshore Niger Delta. *Journal of African Earth Sciences, 15*(4), 585–593. https://doi.org/10.1016/0899-5362(92)90027-V

Detournay, E., & Cheng, A. H.-D. (1993). Fundamentals of poroelasticity. In *Developments in Geotechnical Engineering* (Vol. 56, pp. 113–171). Elsevier.

Doust, H., & Omatsola, E. (1990). Niger Delta. In J. D. Edwards & P. A. Santogrossi (Eds.), *Divergent/passive margin basins* (AAPG Memoir 48, pp. 201–238). American Association of Petroleum Geologists.

Economides, M. J., & Nolte, K. G. (2000). *Reservoir stimulation* (3rd ed.). John Wiley & Sons.

Emam, M., Erdoğan, A. T., & Mohamad, R. (2020). Machine learning models for predicting subsurface temperature using well logs: A case study from an Iranian oilfield. *Journal of Petroleum Science and Engineering, 193*, 107414. https://doi.org/10.1016/j.petrol.2020.107414

Farid, A., & Nasrabadi, H. (2017). Lithofacies classification of well logs using convolutional neural networks. *Geophysical Prospecting, 65*(4), 1140–1155. https://doi.org/10.1111/1365-2478.12540

Horner, E. C. (1951). Pressure build-up in wells. *Transactions of American Institute of Mining and Metallurgical Engineers, 192*, 198–204.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks, 2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Jarvie, D. M., Hill, R. J., Ruble, T. E., & Pollastro, R. M. (2007). Unconventional shale-gas systems: The Mississippian Barnett Shale of north-central Texas as one model for thermogenic shale-gas assessment. *AAPG Bulletin, 91*(4), 475–499. https://doi.org/10.1306/12190606068

Karpatne, A., Atluri, G., Faghmous, J., Steinbach, M., Banerjee, A., Ganguly, A. R., … Kumar, V. (2017). Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering, 29*(10), 2318–2331. https://doi.org/10.1109/TKDE.2017.2730163

Khodaei, M. H., Siahkoohi, H., & Shahrabi, M. (2021). Prediction of geothermal gradient using machine learning techniques: A case study from a clastic basin. *Energy Reports, 7*, 1230–1242. https://doi.org/10.1016/j.egyr.2021.04.051

Lewan, M. D. (1985). Kinetics of oil and gas generation from kerogen: Implications for hydrocarbon generation in sedimentary basins. *Geochimica et Cosmochimica Acta, 49*(3), 615–628. https://doi.org/10.1016/0016-7037(85)90028-8

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R News, 2*(3), 18–22. https://CRAN.R-project.org/doc/Rnews/

Majorowicz, J., Grasby, S., & Chang, H. (2005). Heat-flow distribution in sedimentary basins: Effects of groundwater flow, sedimentation and radiogenic heat production. *Earth and Planetary Science Letters, 234*(1-2), 13–27. https://doi.org/10.1016/j.epsl.2005.02.031

Nwachukwu, J. I. (1976). Petrographic and geochemical characteristics of the Tertiary Niger Delta and their implications for regional stratigraphy. *Bulletin of American Association of Petroleum Geologists, 60*(9), 141–161.

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., & de Freitas, N. (2018). Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE, 104*(1), 148–175. https://doi.org/10.1109/JPROC.2015.2494218

Tiwari, S., & Singh, R. K. (2019). Comparative study of machine learning techniques for geothermal gradient prediction. *Applied Thermal Engineering, 162*, 114239. https://doi.org/10.1016/j.applthermaleng.2019.114239

Huang, X., Zhang, Y., & Li, Z. (2022). Gradient boosting machine approaches for subsurface temperature estimation in petroleum reservoirs. *Energy Conversion and Management, 247*, 114734. https://doi.org/10.1016/j.enconman.2021.114734