PROCEEDINGS, 49th Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 12-14, 2024
SGP-TR-227

An Integrated Data and Physics-based Temperature Model for Real-time Estimation of Bottomhole Temperature for Downhole Tool Failure Prevention

Sadjad Naderi and Naveen Velmurugan.

Virgil Dynamics SAS, 45-47 avenue Carnot, 94230 Cachan, France.

naveen.velmurugan@virgildynamics.com

Pragna Nannapaneni, Michael Yi, and Pradeepkumar Ashok.

Intellicess, Inc., 7000 North Mopac Expressway #200, Austin, Texas 78731, USA.

pradeepkumar.ashok@intellicess.com

Keywords: Geothermal Wells, Drilling, Real-time Prediction, Hydrothermal Model, Physics-Informed Machine Learning, and Data Assimilation.

ABSTRACT

Efficient drilling of geothermal wells hinges on a thorough comprehension of temperature distribution along the wellbore, vital for preventing tool failures and reducing non-productive time (NPT). This involves ascertaining the requisite mud cooling at the surface and optimizing circulation flow rates to extract heat effectively. While extensive theoretical work exists in temperature profiling, practical implementations have been limited. This paper introduces a field-deployable model addressing these challenges through the implementation of a Physics-Informed Machine Learning (PIML) method. This innovative approach distinguishes itself by integrating a physics-based model with real-time data, resulting in a notable enhancement in temperature prediction accuracy. The refined predictions facilitate precise determination of required wellbore cooling and mud flow rates, consequently aiding in extending the lifespan of downhole tools during drilling operations.

The physics-based temperature model is developed, considering critical factors such as heat transfer in both axial and radial directions, as well as friction between the fluid and casings/ drill pipes. This construction harnesses the 3D finite difference method (FDM). A dynamic simulation framework seamlessly integrates essential data for development and simulation assimilation, encompassing wellbore survey data, bottom hole assembly (BHA) data, mud properties, and additional contextual information. Real-time data, continuously streamed every second from the drilling rig, undergoes discrete event simulation (DES). Subsequently, this data undergoes processing utilizing an FDM-based engine, generating a dynamic temperature profile of the fluid every five-minute interval at a minimum, if not quicker due to state change. The assimilation of the FDM model is accomplished through the calibration of the geothermal gradient using Ensemble Kalman Filter (EnKF). Observed data, including inlet mud temperature (utilized as FDM input) and outlet mud temperature (compared to FDM output), is incorporated into the assimilation process. This comprehensive approach ensures an accurate representation of downhole temperature dynamics, facilitating proactive measures for enhanced downhole tools functionality and drilling optimization.

After constructing the model, validation was conducted using the Utah FORGE dataset to showcase the predictive capability of the FDM model in estimating outlet mud temperature over specific time intervals through assimilation. The assimilated temperature history simulated closely mirrored the actual field data. This algorithm not only showcased proficiency in the temperature prediction but also introduced an innovative approach for calibrating and identifying other input parameters, such as fluid properties in response to potential temperature variations and geothermal gradient. Essentially, the proposed technique leverages the FDM model in inverse, serving as a soft sensor. This approach provides valuable insights for control, planning, and understanding by elucidating parameters that are typically challenging to determine.

1. INTRODUCTION

Geothermal drilling, a crucial element of clean energy infrastructure, poses a distinctive challenge: the intense temperatures encountered during the exploration of deep underground reservoirs. The primary focus here is on maintaining an optimum temperature for the drilling mud, a vital component in geothermal wells. This fluid serves a dual purpose by facilitating smooth drilling operations through the removal of cuttings and lubrication of the drill string, while also managing formation pressure. However, as the drilling delves deeper, the temperature of the mud sharply rises. This thermal escalation not only jeopardizes the essential rheological properties of the mud, potentially impeding drilling efficiency, but also poses a threat to critical downhole tools. This leads to an accelerated deterioration of bearings, seals, and other delicate electronic components (Mitchell & Miska, 2011). The core of the challenge lies in the precise measurement and control of this formidable adversary. Traditional mud pulse telemetry (MPT)-based thermometry encounters difficulties in the demanding borehole environment, primarily due to an increased risk of downhole tool failures (Kruszewski & Wittig, 2018).

Predicting the temperature profile within a geothermal well is a complex balancing act, akin to navigating a thermal maze. Researchers have tackled this challenge with a diverse toolbox of modeling techniques, each bringing unique strengths and limitations. Numerical simulations like finite element models (FEMs) and finite difference methods (FDMs) excel at capturing intricate heat transfer, but their

computational burdens often limit their real-time application (Chen & Novotny, 2003). Analytical (Hasan & Jang, 2021) and semi-analytical (Polat, 2022) models offer swift predictions, but their accuracy hinges on simplifying assumptions that can overlook crucial factors like complex wellbore geometries and transient thermal behavior during the drilling operation. Data-driven approaches, powered by machine learning, hold promise for mining historical data and adapting to changing conditions, but their effectiveness relies heavily on data quality and their interpretability can be challenging (Kshirsagar & Sanghavi, 2022). Despite these advancements, a critical gap remains between theoretical models and practical needs. Extensive hydraulic-thermal models exist, but their cumbersome nature and static input requirements make them impractical for real-time decision-making during a dynamic drilling operation (Gonzalez Angarita, 2020) where the rig states and operational parameters (mud flow rate, drill string rotation, location of bit with respect to hole depth, etc.) change quite frequently. This is where our proposed Physics-Informed Machine Learning (PIML) model comes in, offering a new approach to bridge this gap. This innovative technique marries the power of a physics-based FDM model (Hasan, Kabir, & Sarica, 2018) with real-time data, yielding an unprecedented level of accuracy in predicting mud temperature distribution throughout the wellbore.

FDM have established themselves as a workhorse in simulating heat transfer within wellbores. Their strength lies in discretizing the wellbore geometry into a grid, allowing for the precise evaluation of heat conduction, convection, and radiation across each individual point (Holmes & Swift, 1970). This granular approach captures a multitude of factors critical to geothermal drilling, including fluid flow dynamics, formation properties, and frictional heating (Zhang, Xiong, & Guo, 2018). The resulting temperature profiles offer detailed insights into the thermal landscape within a wellbore (Zhou, 2013). However, this integration comes with its own set of uncertainties. FDM models rely heavily on accurate input parameters, including rock and material thermal properties, fluid rheology (Merlo, Maglione, & Piatti, 1995), and wellbore geometry (Hasan, Kabir, & Wang, 2009). Unfortunately, in the often-murky depths of geothermal reservoirs, obtaining precise data can be a Sisyphean task. Limited downhole measurements, complex geological formations, and dynamic fluid behavior contribute to a significant degree of uncertainty, introducing potential discrepancies between model predictions and real-world temperature profiles. Calibrating these models further intensifies the challenge and involves a laborious process of fine-tuning individual input parameters, commonly based on iterative trial-and-error method, until the simulated temperature profile aligns with limited observations, like outlet mud temperature (Tekin, 2010). However, with numerous and often poorly constrained parameters, this approach can be akin to searching for a needle in a thermal haystack. The potential for overfitting becomes a looming threat, as adjustments to one parameter might simply compensate for inaccuracies in another, masking underlying issues in the model or data. This conundrum of uncertainties and calibration hurdles underlines the critical need for a new systematic calibration method.

Traditionally, FDM model calibration has been a static process, relying on limited data points to adjust key parameters. This approach, however, fails to capture the dynamic nature of geothermal wells while drilling, where fluid flow, formation properties, wellbore geometry and heat transfer constantly evolve. To address this limitation, we propose a real-time data assimilation framework that integrates live measurements into the FDM model, enabling continuous calibration and adaptation. At the heart of this framework lies the concept of feedback. Real-time data, such as inlet and outlet mud temperatures, is directly fed back into the model, allowing it to adjust its predictions as the drilling progresses. This continuous feedback loop acts as a perpetual calibration mechanism, mitigating uncertainties and preventing overfitting. Discrepancies between model predictions and observations are promptly addressed, ensuring the model remains faithfully in tune with the actual downhole thermal dynamics. One of the most powerful tools for this data assimilation is the Ensemble Kalman Filter (EnKF) (Tureyen, Onur, ve Dogal, & Bol, 2011). Such methods have shown advantages in calibrating physics-based simulations across various scientific and engineering disciplines (Roth, Hendeby, Fritsche, & Gustafsson, 2017). Unlike traditional single-point adjustments, EnKF works with an ensemble of model realizations, each representing a plausible scenario within the realm of uncertainties. By comparing this ensemble's predictions with real-time data, EnKF statistically updates the entire ensemble, effectively guiding the model towards a more accurate representation of the true downhole temperature profile. This probabilistic approach not only provides robust and realistic predictions but also quantifies the remaining uncertainties, offering valuable insights for informed decision-making throughout the drilling process.

The subsequent sections of this manuscript unfold a comprehensive framework for advancing temperature prediction in geothermal drilling. All the codes presented were developed using Python. We commence by detailing the development of FDM models dedicated to temperature prediction. Subsequently, our focus shifts to the implementation of an FDM-based engine, facilitating dynamic simulation by capturing the evolving thermal dynamics within the wellbore. The integration of a Discrete Event Simulation (DES) framework further enhances our methodology, harmonizing it with real-time data assimilation for continuous calibration. The efficacy of this framework will be rigorously evaluated through the calibration of field data, leveraging the Utah FORGE dataset (McLennan, Nash, Moore, Skowron, & Woolsey, 2021). Results and discussions will provide insights into the model's performance, culminating in a nuanced understanding of its strengths and limitations. The manuscript concludes by outlining avenues for future work, aiming to continually refine and extend the proposed framework for enhanced accuracy and adaptability in geothermal drilling applications.

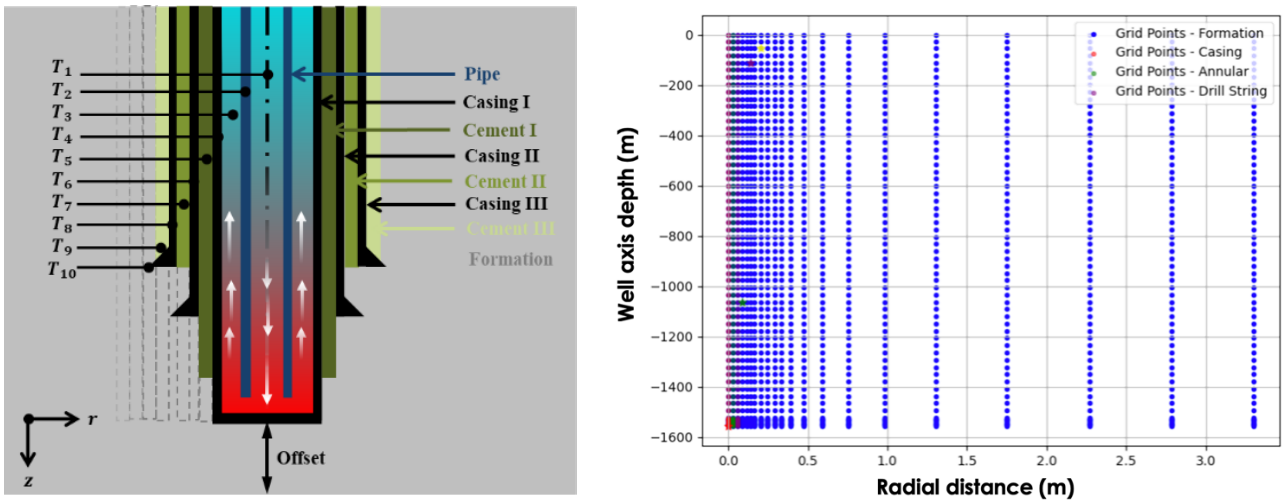
2. HYDROTHERMAL MODELING WITH FDM

2.1 Geometry Model and Domain Discretization

The simulation considers a coaxial well represented by concentric cylinders in 3D with mud flowing through the drill pipe and rising up to the surface through the annular region between drill pipe and casing. **Figure 1**a illustrates the schematic of the geometry model based on the Utah FORGE Dataset, specifically referencing wellbore 16A(78)-32. The drillstring components and casing program for sections using BHA 9 are detailed in **Table 1**. The simulation encompasses a directional well trajectory, incorporating both vertical and horizontal sections. The model is treated as a symmetric problem around the well axis. The Borehole Heat Exchanger (BHE) is modeled as several 1D line components, including inside the drill string/pipe, the annular and external pipe wall, casing, and cement layers. The surrounding rock is simulated in both radial and axial directions. Utilizing a 1D line source to represent the BHE requires fewer nodes and less computational time, a critical consideration for this model. While this assumption introduces some degree of simplification, its impact on

model accuracy will be appropriately addressed in the assimilation phase. The model accounts for an axial offset into the rock formation beyond the bottom of the well. In the radial direction, the length of the formation is sufficiently large – 30 times the wellbore radius. This design ensures that the temperature at the outer radial boundary remains constant and mirrors the surrounding rock, especially far from the wellbore. Such considerations contribute to a more accurate representation of thermal behavior in both axial and radial dimensions.

For radial discretization, the domain grid is generated with initial dimensions following the diameter and thickness specifications outlined in **Table 1** for the wellbore components. As the grid extends into the rock formation, the spacing gradually increases outward, reaching a maximum value of 0.5 m before maintaining a constant spacing. In the axial direction, discretization is determined by data points along the well trajectory, ensuring alignment with the bit motion during drilling. This approach establishes a consistent and accurate representation of the model, harmonizing with the dynamic nature of the drilling process.



(a) **(b)**

Figure 1: **(a) 2D schematic of a coaxial borehole heat exchanger, and (b) schematic of a 2D cross section through 3D model of the discretization of FDM grid.**

Table 1: Basic data of drill string assembly and casing program for wellbore 16A(78)-32.

Wellbore	Parameter	Inner diameter (mm)	Outer Diameter (mm)	Depth (m)
BHA 9	Drill Pipe	110	127	Variable
	First Casing	220	240	1062.5
	Second Casing	320	340	116.4
	Third Casing	480	510	39.3

2.2 Governing Equations

The borehole is modeled to capture thermal interactions between the wellbore and the surrounding rock. The model is built on the following key assumptions: (i) the drilling fluid is treated as incompressible, maintaining constant density, specific heat capacity, and thermal conductivity; and (ii) the velocity of the drilling fluid in the drilling string and annulus is considered solely in the axial direction, while the radial velocity is disregarded. Governing equations **(1)** to **(5)** describe the heat flow in the fluid inside the drill string, the drill string pipe, the fluid in the annular space, the first layer of casing, and the casing/cement/surrounding formation, respectively. While these equations form the backbone of the model, it is essential to acknowledge that they have been extensively covered in the literature (Zhang, Xiong, & Guo, 2018). To maintain brevity, we highlight key aspects of the modeling process, directing readers to the provided references for a more comprehensive understanding of the equations and their derivations.

$$-\frac{\partial(\rho_1 c_1 v_p T_1)}{\partial z}+\frac{2h_1(T_2-T_1)}{r_1}+\frac{q_p}{\pi r_1^2}=\frac{\partial(\rho_1 c_1 T_1)}{\partial t} \tag{1}$$

$$\frac{\partial}{\partial z}\left(\frac{\lambda_p \partial T_2}{\partial z}\right)+\frac{2h_2(T_3-T_2)r_2}{r_2^2-r_1^2}-\frac{2h_1(T_2-T_1)r_1}{r_2^2-r_1^2}=\frac{\partial(\rho_2 c_2 T_2)}{\partial t} \tag{2}$$

$$\frac{\partial(\rho_3 v_a c_3 T_3)}{\partial z}+\frac{2h_3(T_4-T_3)r_3}{r_3^2-r_2^2}-\frac{2h_2(T_3-T_2)r_2}{r_3^2-r_2^2}+\frac{q_a}{\pi(r_3^2-r_2^2)}=\frac{\partial(\rho_3 c_3 T_3)}{\partial t} \tag{3}$$

$$\frac{\partial}{\partial z}\left(\lambda_4\frac{\partial T_4}{\partial z}\right)+\frac{2h_3(T_5-T_4)}{r_4^2-r_3^2}-\frac{2h_3(T_4-T_3)r_3}{r_4^2-r_3^2}=\frac{\partial(\rho_4 c_4 T_4)}{\partial t} \tag{4}$$

$$\frac{\partial}{\partial z}\left(\lambda_i \frac{\partial T_i}{\partial z}\right) + \frac{1}{r}\frac{\partial}{\partial r}\left(\lambda_{effective}^{i,i+1} r \frac{\partial T}{\partial r}\right) = \frac{\partial(\rho_i c_i T_i)}{\partial t} \tag{5}$$

Here, r and z denote radial and axial coordinates, respectively. In Equations **(1)** to **(4)**, the subscript numbers follow the layer order illustrated in Figure 1a, while the subscript i in Equation **(5)** denotes the layer number in casing/cement/surrounding formation. In the context of the equations, T, h, λ, c and ρ represent temperature, convective heat transfer coefficient, thermal conductivity, specific heat capacity, and density, respectively. The conductivity coefficient between layer i and $i+1$, denoted as $\lambda_{effective}^{i,i+1}$, is calculated using the thermal-electrical analogy (Brown, Cassidy, Egan, & Griffiths, 2021), employing the following formulation:

$$R_{effective} = \frac{1}{\lambda_{effective}}, \quad R_{effective} = R_i + R_{i+1} = \frac{1}{\lambda_i} + \frac{1}{\lambda_{i+1}} \tag{6}$$

This expression signifies that the thermal resistivity of the effective layers can be considered analogous to electrical resistivity in series. The next section details the calculation of q_p and q_a, representing friction-induced heat source in drilling pipe and annular, respectively.

2.3 Calculation of Heat Source Terms

The calculation of the friction-induced heat source involves considerations of fluid dynamics within the drilling pipe and the annulus (Zhang, Xiong, & Guo, 2018). Numerous empirical and analytical models exist for this purpose. Here, the Reynolds number (Re) as a crucial parameter is defined as (Durst & Arnold, 2008):

$$Re = \frac{\rho v D_h}{\mu} \tag{7}$$

where μ is the dynamic viscosity of the mud, and D_h and v are the hydraulic diameter and fluid velocity respectively. It should be noted that for the annular flow, $D_h = 4\,(d_3^2 - d_2^2)/(d_3 + d_2)$, where d_3 represents the diameter of the wellbore, and d_2 is the outer diameter of the drill string. For laminar flow ($Re \leq 2000$), the coefficient of friction resistance (f) is calculated as:

$$f = \frac{16}{Re} \tag{8}$$

In turbulent flow ($Re > 2000$), f is determined using the empirical Darcy friction factor equation:

$$\frac{1}{\sqrt{f}} = 4\log\left(Re\sqrt{f}\right) - 03.95 \tag{9}$$

The frictional resistance pressure drop in the drill string is then computed using:

$$\frac{\Delta P}{\Delta L} = \frac{2 f \rho v_p^2}{d_1} \tag{10}$$

where L denotes the flow distance of drilling fluid. Similarly, in the annular region, the pressure drop is determined by:

$$\frac{\Delta P}{\Delta L} = \frac{2 f \rho v_a^2}{d_3 - d_2} \tag{11}$$

Finally, the heat source term (q) in Equations **(1)** and **(3)** is quantified by:

$$q = \frac{\Delta P Q}{\Delta L} \tag{12}$$

where Q is the volume flow rate of drilling fluid.

2.4 Boundary and Initial Conditions

The initial temperature in the formation is determined by a linearly increasing geothermal gradient (G). Although a linear, continuous relationship is assumed, adjustments like including stepwise changing G can be made in the assimilation process, as discussed later in the following sections.

The inlet temperature is known and can be measured directly on the ground:

$$T_1(z = 0, t) = T_{in} \tag{13}$$

At the bottom of the well, the annulus drilling fluid temperature is equal to the drilling fluid temperature inside the drill string and the drill string temperature:

$$T_1(z = z_f, t) = T_2(z = z_f, t) = T_3(z = z_f, t) \tag{14}$$

where z_f is the vertical depth at the bottom of the well. There is no heat exchange between the formation and the atmosphere at the surface:

$$\frac{\partial T}{\partial z}|_{z=0} = 0 \tag{15}$$

The formation temperature far away from the wellbore is undisturbed and is equal to the initial formation temperature. This condition is described at r_f, the end of the rock domain in the radial direction:

$$\frac{\partial T}{\partial r}|_{r \to r_f} = 0 \tag{16}$$

2.5 Numerical Solution and implementation

The FDM employed in this work for solving the governing equations utilizes a first-order forward difference scheme for temporal discretization and a central difference scheme for spatial discretization. These choices ensure first-order stability, which is crucial for explicit methods. In our framework, stability concerns are further amplified due to the model's real-time update using live data received at per-second intervals. To address this, an adaptive time step approach is implemented. This dynamically adjusts the time step based on CFL (Courant-Friedrichs-Lewy) condition (Moura & Kubrusly, 2012) dependent on grid spacing, and data frequency, guaranteeing both computational stability and efficiency while remaining consistent with the underlying physical processes. While the specific details of FDM implementation have been extensively documented in the literature, the novelty lies in our new time step adaptation strategy specifically tailored for real-time data assimilation within the geothermal wellbore context.

3. FDM-BASED ENGINE FOR DYNAMIC SIMULATION

3.1 Overview of Engine Structure

The FDM-based engine for dynamic simulation serves as the conductor in our geothermal temperature prediction orchestra, blending static and dynamic data seamlessly to deliver an accurate and real-time symphony of downhole behavior. As schematically illustrated in **Figure 2**, this engine comprises two distinct compositions:

1. **Static Data-Based Model:** This foundational component establishes the computational groundwork for the initiation of dynamic model. The time-step determination for this segment is independent from the data frequency. It is worth mentioning that the choice between explicit and implicit methods should align with the system's overall dynamism.
2. **Dynamic Data-Based Model:** Extending from the surface to the wellbore's bottom, this section integrates live measurements into its simulation. Data assimilation, a critical process fusing observations with model predictions, operates exclusively within the targeted wellbore, focusing on real-time dynamics. The time step for this dynamic model aligns directly with the data frequency (1 second) of the thermal-hydro model, ensuring continuous synchronization with real-time conditions.

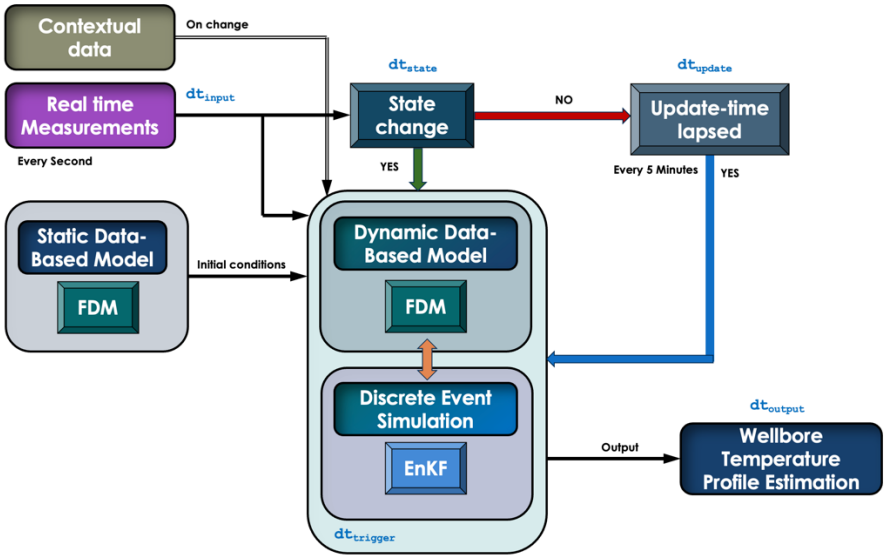


Figure 2: Schematic representation of the integrated framework, illustrating the dynamic assimilation of real-time data into the forward-marching explicit numerical solution.

This dynamic core operates through a dedicated Assimilation Process (AP), occurring exclusively within the targeted wellbore to capture dynamic simulation aspects. This targeted approach ensures that real-time data influences the section critical for accurate predictions. Establishing initial temperatures for both static and dynamic models involve careful consideration: For the static model, the initial temperature across the wellbore and surrounding formation typically aligns with the initial formation temperature. For the dynamic model, two options exist for initial conditions – either inherit the temperature field from the static model, creating a smooth transition between foundational and real-time components or known temperatures, if available, can be directly prescribed as the starting point for the dynamic simulation.

Figure 2 illustrates the interaction among various components in our framework, orchestrating a harmonious integration of data and model predictions. At the core of this workflow is a two-fold FDM engine. The first module acts as the foundational conductor, employing static data to establish the initial thermal landscape of the wellbore and surrounding formation. This initial temperature profile transitions to the second FDM module, where it forms the basis for real-time simulations. In this dynamic phase, real time measurements, a vital component of our approach, is continuously input into the FDM model at a rate of one data point per second, denoted as dt_{input}. The real-time influx of information is integrated into the *forward-marching explicit numerical solution*, ensuring the model remains in perfect harmony with the evolving subsurface conditions. Concurrently, when there are changes in drilling contextual data and/or rig state and operational parameters or merely on the lapse of update-time dt_{update} (here five-minute intervals), a subset of this observed data is directed to the ML module, facilitating the assimilation process, considered occurring at time intervals $dt_{trigger}$. State changes such as connection and tripping or change in mud flowrate are identified using a robust methodology utilizing Bayesian network and their duration is identified as dt_{state}. This continuous feedback loop enhances the synergistic capabilities of FDM and ML, progressively refining the representation of downhole thermal dynamics whose data are stored at time intervals dt_{output}. Although dt_{update} and dt_{output} may be user defined, and dt_{input} is determined by practical limitations of rig data acquisition systems, following conditions are helpful to understand the working of this framework: $dt_{trigger}$ is the minimum of dt_{state} and dt_{update}. Also, the maximum value of dt_{output} is $dt_{trigger}$.

3.2 Input Data Integration

As shown in the **Figure 3**, the input data is composed of three main modules: (i) raw field data (observed data) including real-time drilling data, drilling survey and temperature data at inlet and outlet; (ii) primary inputs including well geometry, drillstring information, physical properties of material, operational parameters, etc., and ; (iii) secondary inputs including calculation of friction-induced heat source based on Reynold Number for the fluids inside the drill string and the annulus. In terms of how they are linked: the raw data needs to be pre-processed including filtering (if necessary), re-constructing data format, unit conversion, etc. Then, the data will populate into the primary data. Based on this information some additional parameters will be calculated such radii, conductivity coefficient, equivalent resistance, etc. The key point in the primary data set is that some of them can be treated as dynamic live data as required while the other static dataset parameters remain unchanged during assimilation. The structure allows a better data integrate-ability and management in terms of code development.

Figure 3 illustrates the coordination of input data within our framework, outlining the constraints and composition of essential information. Three distinct modules integrate, each contributing vital data:

- **Raw Field Data (Observed Data):** This initial step involves the unprocessed wellbore data, including drilling survey data and real-time data. Prior to integration into the predictive model, pre-processing procedures are applied to cleanse the data, eliminate noise, and ensure compatibility with other components.
- **Primary Inputs:** Serving as the foundational elements, this core module accommodates critical parameters such as wellbore geometry, physical properties of the formation, operational details, and constants. Dynamic live data continuously refreshes the simulation with the latest subsurface information, while static parameters remain unaltered during assimilation, imparting stability to the evolving thermal predictions. Notably, in **Figure 3**, selected parameters are labeled as "Live Data" for illustrative purposes, demonstrating how this approach can be implemented.
- **Secondary Inputs:** Expanding on the primary data, this module introduces calculated values derived from primary input data. Formulas, such as the calculation of friction-induced heat source based on Reynolds number for both drill string and annular fluid flows, contribute additional complexity to the data landscape.

This organized structure of input data ensures efficient integration and management within the framework's code, facilitating a clear and coordinated performance.

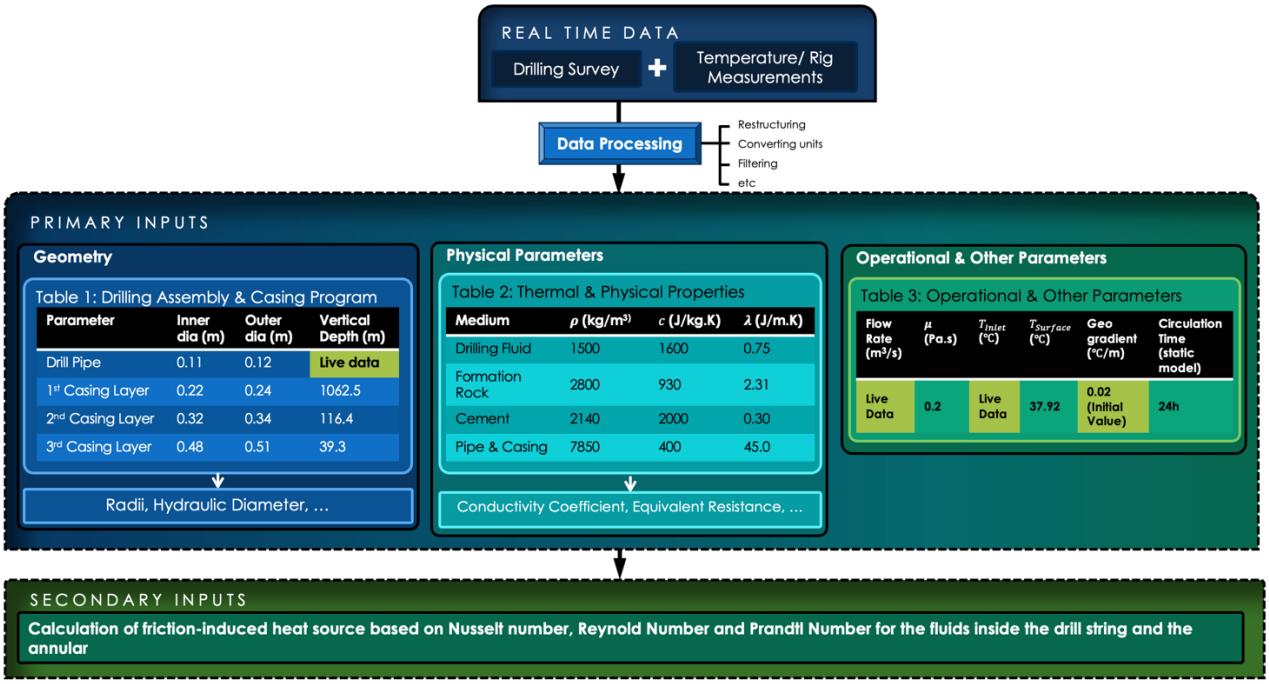


Figure 3: A schematic overview of the input data integration framework.

4. DES FRAMEWORK INTEGRATED WITH DATA ASSIMILATION FOR CALIBRATION

Discrete Event Simulation (DES) stands as a computational modeling approach designed to analyze and comprehend the intricate dynamics of complex systems. Unlike continuous simulations, where changes unfold continuously over time, DES focuses on events that occur at distinct points in time, capturing the discrete nature of operations and interactions within a system. This is particularly relevant in well construction where rig state change and operational parameter changes constitute discrete events. At its core, DES operates on the principle of simulating individual events and their effects on the system. These events represent occurrences that prompt changes in the system's state or trigger specific actions. Here, events include triggering calibration/assimilation event at least once every five minutes based on observed data. In a DES framework, several key components play integral roles in shaping the system's dynamics. Central to this structure are *Events*, defining essential occurrences that exert influence on the system. These events span the spectrum from initiating drilling operations to identifying anomalies in temperature readings. *Entities*, representing the various objects or components within the system, respond to these events. *Queues*, acting as waiting lines for entities not immediately processed, play a significant role in managing the workflow. Finally, the *Clock* discretely advances, symbolizing the progression of time within the system. Events unfold at precise moments, synchronized with the progress of this simulation clock, providing a temporal framework for the dynamic interactions within the system.

The integration of DES with the EnKF introduces a dynamic dimension to real-time system analysis (Budgaga, et al., 2016). DES, proficient in capturing discrete events and interactions, synergizes with the probabilistic and ensemble-based characteristics of EnKF (Lang, Kuetgens, Reichardt, & Reggelin, 2021). This integration enhances the representation of the geothermal drilling environment, enabling the modeling of uncertainties and variations in a manner that aligns with the stochastic nature of drilling operations, such as during connection and tripping operations. At the core of our integrated framework is the EnKF, a multivariate, sequential Monte Carlo method designed to reconcile live measurements with the predictions of our physics-based model. Departing from traditional single-point adjustments, our EnKF methodology integrates with DES, fostering a probabilistic approach through an ensemble of model realizations. This ensemble-based approach, within the DES framework, encapsulates a spectrum of plausible scenarios, providing a statistically robust foundation for model calibration. The synergy between this ensemble-based approach and the dynamic nature of DES enables the model to adapt to real-time events and changes within the geothermal system. The integration ensures effective capture of uncertainties and variability intrinsic to dynamic drilling operations, collectively enhancing the accuracy and reliability of our model's predictions.

The assimilation process for the geothermal gradient (G) unfolds through a series of steps:

1. **Ensemble Initialization:** The algorithm initiates a loop, iterating for a specified number of cycles. Within each cycle, a set of simulations with an ensemble size is generated, representing diverse and plausible system states.
2. **Forward Model Simulation:** For each ensemble member, the algorithm performs a forward model simulation using the physics-based FDM. Simulated mud temperatures are generated based on the current system state.
3. **Observation Integration:** Simulated ensemble states are compared with observed real-time mud temperatures. The innovation, depicting the difference between simulated and observed temperatures, is calculated.
4. **Kalman Gain Calculation:** The Kalman gain, a vital statistical parameter, is computed using the covariance between geothermal gradients and simulated temperatures. Normalized by the variance of simulated temperatures, this gain quantifies the optimal balance between trusting model predictions and incorporating observed data.

5. **Ensemble Update:** The ensemble of geothermal gradients undergoes adjustments based on the Kalman gain and innovation, refining the representation of geothermal gradients within the ensemble.
6. **Convergence Check:** The algorithm verifies convergence based on a specified threshold for the mean of the innovation. If the convergence criterion is satisfied, the iteration halts, signaling a refined estimate of the system's current state.

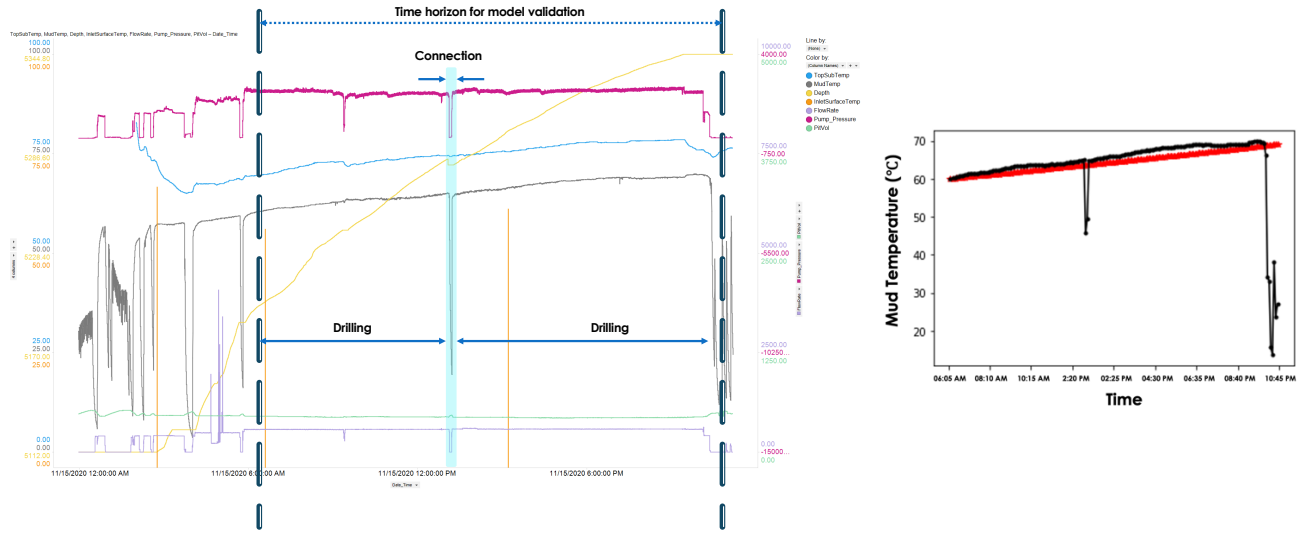
The iterative nature of this assimilation process allows the ensemble to adapt and learn from real-time data, progressively improving the accuracy of the model predictions in capturing the dynamics of the system. It is important to highlight that the ensemble state can encompass various parameters, including fluid properties or other input parameters for the FDM, allowing for a comprehensive calibration of the FDM by considering multiple variables.

5. EVALUATION AND CALIBRATION WITH UTAH FORGE DATASET

In configuring the model for evaluation and calibration using the Utah FORGE dataset, the static segment of the framework was defined from the surface down to the section corresponding to BHA 9 – the region chosen for assimilation performance assessment. Raw data underwent trimming and filtering procedures to enhance data quality. The envisioned drilling scenario encompasses a continuous operation transitioning from the static to dynamic phases. This approach ensures fluid progression, with the static model furnishing initial conditions for the dynamic component. The circulation time for the static model was established at 24 hours, a duration sufficiently ample to establish initial conditions consistent with the assimilation phase's temperature starting point.

It is noteworthy that the model can be configured for more complex scenarios, and the DES framework can be tailored to accommodate various events beyond calibration, such as drilling pauses or failures. However, this work focuses on presenting the fundamental aspects of the methodology rather than delving into an exhaustive array of potential applications, opting for a simplified setup for clarity. **Figure 3** provides essential data, encompassing geometry, physical parameters, operational details, and other relevant factors. For this benchmark case, fixed values of 0.035 m³/s for the flow rate and 37.92 °C for the temperature of the inlet mud were set for the sake of simplicity.

In **Figure 4**a the measurements from the wellbore 16A(78)-32 within BHA 9 section is presented and the time interval for which the DES framework is validated. During this period, two phases of drilling operation is separated by a connection operation. In **Figure 4**b the assimilated results for outlet mud temperature are presented, providing a comparative analysis with observed temperatures. Simulated data points, plotted at 5-minute intervals, offer a detailed overview of outlet mud temperature predictions. The figure provides insights into the efficacy of our data assimilation framework. The assimilated results closely align with the observed temperature throughout the entire time series, highlighting the framework's ability to predict dynamic temperature changes. This successful assimilation process indicates the framework's capability to refine model predictions by incorporating real-world data, resulting in temperature estimates that accurately capture the dynamic nature of the geothermal wellbore during drilling.



(a) **(b)**

Figure 4. (a) Realtime measurement of well 16A(78)-32 within BHA 9 section highlighted for model validation, and (b) Comparison of the output mud temperature calibrated by PIML to the value measured in the field.

As observed in **Figure 4**b, upon closer examination of the dynamic behavior, the assimilated temperatures exhibit reduced fluctuations compared to the observed data. While this initially suggests improved predictability, it also unveils the model's potential limitations in adapting to real-time changes. These limitations manifest in two distinct forms:

1. **Local Nonlinearity (within ±5°C):** Real-time fluctuations within a range of less than 5°C, potentially representing variations in mud and formation properties during drilling, are smoothed out by the assimilation process. This suggests the model struggles to capture these subtle, non-linear changes in the wellbore environment. Addressing this limitation could involve: (i) enhancing the ensemble state with additional parameters sensitive to mud and formation properties; and (ii) incorporating local nonlinearities in the physics-based model such as nonlinear geothermal gradient relationship.

2. **Drastic Temperature Changes (exceeding ±20°C):** The model tends to underestimate or overestimate temperature shifts exceeding ~ 20°C, likely indicative of interruptions in the drilling operation, here a connection. Mitigating this limitation could involve: (i) including a broader range of potential events within the DES framework to better capture the drilling operational changes and disruptions; and (ii) modifying the data assimilation algorithm to be more responsive to significant deviations.

While these limitations highlight areas for improvement, they also underscore the valuable role of a physics-based model and real time data within the data assimilation process. By acting as a constraint, the model offers guidance to the machine learning algorithms, preventing potentially erroneous interpretations and overfitting of the observed data. This synergy between physics and machine learning paves the way for robust and reliable real-time predictions in such complex systems.

It is worthy to note the DES framework considered here, powered by FDM and EnKF, was run on a MacBook with an Apple M1 chip and 16GB RAM with an average execution time for each iteration (here, 5 mins of actual operational time) of less than a minute. However, a tradeoff between computational effort and continuous temperature estimation must be made during edge deployment on a rig – tuned using dt_{input} and $dt_{trigger}$.

6. CONCLUSIONS AND FUTURE WORK

- Traditional approaches to wellbore temperature modelling often grapple with computational inefficiencies or an inability to adapt to real-time changes. This study addresses these limitations through the development of a pioneering framework that seamlessly integrates a precise FDM model with the EnKF within a DES architecture.
- The DES framework, marrying the accuracy of FDM with EnKF's data assimilation capabilities, excels in capturing dynamic temperature changes across the wellbore, from surface to bottom, in real-time. This innovation overcomes the challenges faced by conventional methods and offers a more responsive and adaptable solution for wellbore temperature modelling during drilling operations.
- Using geothermal gradient calibration as a representative case, we showcased the efficacy of the DES framework in augmenting the accuracy of the FDM model. The integration of live data through EnKF resulted in significantly improved temperature predictions compared to traditional methodologies. This successful demonstration underscores the potential for broader applications in diverse scenarios.
- Our DES framework, beyond enhancing accuracy, demonstrates promising potential to act as a "soft sensor" by analyzing calibrated parameters like fluid and rock properties, providing valuable insights into downhole conditions. The real-time prediction capabilities of the framework also open avenues for integration into semi- or fully-automated drilling systems, thereby enhancing operational efficiency and safety.
- Our investigation revealed two primary limitations within the DES framework. Firstly, the current set of input variables for EnKF may be insufficient to comprehensively capture the FDM model's response in specific scenarios. Secondly, the DES framework's simulation engine could benefit from a more diverse range of drilling scenarios to ensure responsiveness across different contexts. Future research will focus on addressing these limitations by:
  - Expanding the ensemble state with additional parameters sensitive to fluid properties and rock thermal properties.
  - Enriching the DES with a wider variety of drilling events and data patterns.
  - Enhancing the FDM model by incorporating additional physics or modifying governing equations to improve accuracy and realism.
  - Refining the geothermal gradient relationship through investigation and adjustments.
  - Optimizing the EnKF algorithm by exploring parameter adjustments or refining the algorithm itself. This may involve investigating the incorporation of previously calibrated parameters into the ensemble state for continuous assimilation.
- Future work will also involve the validation of the enhanced framework with more complex drilling trajectories, including inclined and horizontal sections. Opportunities for additional validation studies will explore the robustness of the developed framework under different geological conditions or wellbore configurations.
- Proposals for collaboration with industry partners or participation in data-sharing initiatives will be pursued to enhance model validation and broaden the dataset. Collaborative efforts can provide diverse scenarios and data inputs that further strengthen the framework's applicability and reliability.

ACKNOWLEDGEMENT

The authors acknowledge the Utah FORGE project for the open dataset used in this study. The authors would like to thank Virgil Dynamics and Intellicess for granting permission to publish the results of their collaborative project in commissioning innovative solutions such as digital twin, cutting-edge simulation techniques, parameter estimation and system automation to enhance drilling performances in the geothermal industry.

NOMENCLATURE

Abbreviations

BHA	Bottom Hole Assembly
BHE	Borehole Heat Exchanger
CFL	Courant-Friedrichs-Lewy
DES	Discrete Event Simulation

EnKF	Ensemble Kalman Filter
FEM	Finite Element Model
FDM	Finite Difference Method
ML	Machine Learning
MPT	Mud Pulse Telemetry
NPT	Non-Productive Time
PIML	Physics-Informed Machine Learning
3D	3 Dimensional
1D	1 Dimensional

Variables

R	Analogical thermal-electrical resistance
h	Convective heat transfer coefficient
ρ	Density
Z	Depth
d	Diameter
f	Friction resistance coefficient
ΔP	Frictional resistance pressure drop
G	Geothermal gradient
q	Heat source term
D_h	Hydraulic diameter
r	Radius
Re	Reynolds number
c	Specific heat capacity
T	Temperature
λ	Thermal conductivity
t	Time
dt	Timestep
ν	Viscosity

Units

°C	Degree Celsius
K	Kelvin
kg	Kilogram
m	Meter
mm	Millimeter
Pa	Pascal
s	Second

REFERENCES

Chen, Z., & Novotny, R. J. (2003). Accurate prediction wellbore transient temperature profile under multiple temperature gradients: finite difference approach and case history. *SPE Annual Technical Conference and Exhibition.* (pp. SPE-84583) SPE .

Mitchell, R. F., & Miska, S. (2011). *Fundamentals of drilling engineering.*

Hasan, A. R., & Jang, M. (2021). An analytic model for computing the countercurrent flow of heat in tubing and annulus system and its application: Jet pump. *Journal of Petroleum Science and Engineering*, 203, 108492.

Polat, C. (2022). A semi-analytical solution technique for predicting circulating mud temperatures. *Journal of Natural Gas Science and Engineering*, 106, 104754.

Kshirsagar, A., & Sanghavi, P. (2022). Geothermal, oil and gas well subsurface temperature prediction employing machine learning. *47th workshop on geothermal reservoir engineering.*

Gonzalez Angarita, J. C. (2020). *Integrated Modelling and Simulation of Wellbore Heat Transfer Processes through High-level Programming, Sensitivity Analysis and Initial Approach with Machine Learning Predictive Models.* Norway: University of Stavanger.

Hasan, A. R., Kabir, C. S., & Sarica, C. (2018). *Fluid flow and heat transfer in wellbores.* Richardson, Texas, Richardson, Texas: Society of Petroleum Engineers.

Holmes, C. S., & Swift, S. C. (1970). Calculation of circulating mud temperatures . *Journal of petroleum technology*, 22(06), 670-674.

Zhang, Z., Xiong, Y., & Guo, F. (2018). Analysis of wellbore temperature distribution and influencing factors during drilling horizontal wells. *Journal of Energy Resources Technology*, 140(9), 092901.

Zhou, F. (2013). *Research on heat transfer in geothermal wellbore and surroundings.* Berlin.

Hasan, A. R., Kabir, C. S., & Wang, X. (2009). A robust steady-state model for flowing-fluid temperature in complex wells. *SPE Production & Operations*, 24(02), 269-276.

Merlo, A., Maglione, R., & Piatti, C. (1995). An innovative model for drilling fluid hydraulics. *SPE Asia Pacific Oil and Gas Conference and Exhibition* (pp. pp. SPE-29259). SPE.

Tekin, S. (2010). *Estimation of the formation temperature from the inlet and outlet mud temperatures while drilling geothermal formations* . Master's thesis, Middle East Technical University.

Tureyen, O. I., Onur, M., ve Dogal, I. M., & Bol, G. M. (2011). Investigation of the use of the ensemble Kalman filter (EnKF) for history matching pressure data from geothermal reservoirs. *36th Workshop on Geothermal Reservoir Engineering.*

Roth, M., Hendeby, G., Fritsche, C., & Gustafsson, F. (2017). The Ensemble Kalman filter: a signal processing perspective. *EURASIP Journal on Advances in Signal Processing*, 1-16.

McLennan, J., Nash, G., Moore, J., Skowron, G., & Woolsey, S. (2021). *Utah FORGE: Well 16A(78)-32 Drilling Data.* United States: USDOE Office of Energy Efficiency and Renewable Energy (EERE), Renewable Power Office. Geothermal Technologies Program (EE-4G).

Brown, C. S., Cassidy, N. J., Egan, S. S., & Griffiths, D. (2021). Numerical modelling of deep coaxial borehole heat exchangers in the Cheshire Basin, UK. *Comuters & Geosciences*, 152, 104752.

Durst, F., & Arnold, I. (2008). *Fluid mechanics: an introduction to the theory of fluid flows.* Berlin: Springer.

Moura, C. A., & Kubrusly, C. S. (2012). *The Courant-Friedrichs-Lewy (CFL) condition: 80 years after its discovery.* . Basel: Birkhäuser.

Budgaga, W., Malensek, M., Pallickara, S., Harvey, N., Breidt, F. J., & Pallickara, S. (2016). Predictive analytics using statistical, learning, and ensemble methods to support real-time exploration of discrete event simulations. *Future Generation Computer Systems*, 56, 360-374.

Lang, S., Kuetgens, M., Reichardt, P., & Reggelin, T. (2021). Modeling production scheduling problems as reinforcement learning environments based on discrete-event simulation and openai gym. *IFAC* (pp. 54(1), 793-798). IFAC-PapersOnLine.

Kruszewski, M., & Wittig, V. (2018). Review of failure modes in supercritical geothermal drilling projects. *Geothermal Energy*, 6(1), 28.