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

Time Series Analysis for Long-term Monitoring and Forecasting of Subsurface Temperatures for a Campus-scale Geothermal Exchange Field

Shubham Dutt Attri[1], Evan Heeg[2], James M. Tinjum[3], Dante Fratta[4], David J. Hart[5], Orhun Aydin[6]

[1]Teaching Faculty, Biological Systems Engineering, College of Agriculture and Life Sciences, University of Wisconsin-Madison, Madison, WI, USA, sattri@wisc.edu,

[2]Graduate Student, Geological Engineering, University of Wisconsin-Madison, Madison, WI, USA, epheeg@wisc.edu,

[3]Associate Professor, Geological Engineering, Civil and Environmental Engineering, University of Wisconsin–Madison, Madison, WI, USA, email: jmtinjum@wisc.edu,

[4]Associate Professor, Geological Engineering, Civil and Environmental Engineering, University of Wisconsin–Madison, Madison, WI, USA, email: fratta@wisc.edu,

[5]Professor, Wisconsin Geological and Natural History Survey, Madison, WI, USA, email: djhart@wisc.edu,

[6]Assistant Professor, Department of Earth and Atmospheric Sciences, Saint Louis University, St. Louis, MO, USA, email: orhun.aydin@slu.edu

Keywords: district-scale geothermal, ground heat exchange, machine learning, exponential smoothing, regression analysis, subsurface temperature predictions

ABSTRACT

We investigate the application of autoregressive time-series models in predicting the subsurface temperature of a low-temperature, geothermal heat exchange (GHX) system. We use subsurface temperature data from 2596 152-m-deep boreholes in a 280 m by 360 m, cooling-dominated, district-scale GHX field in the Midwest region of the United States. We monitored the temperature for over seven years via the deployment of fiber-optic distributed temperature sensing (FO-DTS). This study aims to impute a two-year gap in temperature measurements using the first three years of data and test the out-of-sample performance of three forecasting models. We use autoregressive time series forecasting models (including ARIMA and Holt-Winters Triple Exponential Smoothing) to forecast subsurface temperature using previously observed time series patterns. To predict subsurface temperature, we define the forecast model with three exogenous variables—air temperature, humidity, and the energy exchanged for heating and cooling the campus buildings. We observe that the best MSE value among all the models is 0.0100 $^{o}C^{2}$ for a prediction horizon of one month and 0.0665 $^{o}C^{2}$ for a horizon of six months using the Holt-Winters smoothing. Findings showcase a gradual, seasonal rise in subsurface temperature and offer valuable insights for designing more efficient GHX systems, conducting improved energy balance assessments, and creating long-term ground-temperature change models. We demonstrate the potential of autoregressive time-lag models in forecasting geothermal heat exchange.

1. INTRODUCTION

Geothermal heat exchange (GHX) systems offer a sustainable alternative to traditional fossil-based heating and cooling technologies by tapping into the Earth's stable underground temperatures for space conditioning and domestic hot water production (Walker et al. 2015). In simple terms, GHX systems use the Earth as a leaky thermal battery, harvesting heat from the earth when needed, and injecting heat back to ground when necessary. GHX systems are engineered in various configurations, such as open-loop systems that utilize natural water sources and closed-loop systems that recirculate a heat-conducting fluid through subterranean pipes. The versatility of GHX systems allows for their application across different scales, from individual homes to large districts. Enhancing our grasp of the factors that influence the performance of GHX systems and how these variables impact the system's efficacy over time can significantly bolster our ability to manage, monitor, and forecast the system's behavior throughout its operational lifespan (McDaniel et al. 2018a). Monitoring can be further used to assess the associated environmental and working costs (Thomas et al. 2020), evaluate the system's performance more accurately (Walker et al. 2015), and help make better decisions for the viability and sustainability of future projects; particularly when GHX fields are anticipated to become unbalanced over time (Florea et al. 2017). Fiber-optic Distributed Temperature Sensing (FO-DTS) is an emerging technology in this domain, offering high-resolution, accurate temperature data crucial for assessing and optimizing GHX systems' functionality (Herrera et al. 2018). FO-DTS captures Raman scattering within fiber optic cables, using laser pulses to calculate the real-time temperature data along the fiber route in location and time (Suárez et al., 2011). The strategic application of FO-DTS, underpinned by a carefully selected calibration methodology, enables a detailed spatial and temporal analysis of the geothermal system's performance, aiding in more informed decision-making for future geothermal projects.

This research exploits the sophisticated analytical abilities of ML to unravel and decode patterns found in geothermal datasets. The ML-geothermal energy integration enhances existing processes and extends to the identification of novel exploration techniques and optimization strategies. Researchers and practitioners extract, process, and analyze data more effectively than ever by applying techniques such as neural networks, support vector machines, and ensemble methods. This leads to more accurate predictions of geothermal system behavior and a deeper understanding of the subsurface dynamics that govern the availability and sustainability of geothermal energy resources. Some significant studies include those by Tut Haklidir and Haklidir (2020), who predicted reservoir temperatures with hydrogeochemical data using Linear Regression, Support Vector Machines, and Deep Neural Network, and Holtzman et al. (2018), who

studied seismic source spectra in Geysers geothermal field using ML techniques. In Rezvanbehbahani et al. (2017), Gradient Boost Regression Trees was used to develop a new Greenland geothermal heat flux map with 15% accuracy. In the research conducted by Tang et al. (2022), the authors employed four ML algorithms to examine the correlation between Heat Extraction Rate (HER) with average surrounding temperature, groundwater level, soil type, grout thermal conductivity, and other factors. They analyzed the annual HERs from 400 thermal performance tests using twelve key factors in a validated numerical simulation framework. Yang and Dashdondov (2023), developed ML models for predicting ground temperature (GT) on an hourly basis at depths up to 30 cm, and Guo et al. (2018) developed ML energy demand prediction models for building heating systems.

Time series analyses use systematic observations recorded at consistent intervals within a specified timeframe to examine a specific variable or event (Burtiev et al. 2013). The goal of such analyses is to identify recurring patterns within the variable of interest and to develop models that effectively quantify and forecast the dynamics among the observed phenomena. For example, Cadenas et al. (2016), compared single and multi-variable models for predicting wind speeds in Mexico. They studied how exogenous variables impact wind speed predictions. In contrast, Burtiev et al. (2013) analyzed wind speed and temperature data from 1945–2009, averaged monthly, using the autoregressive integrated moving average (ARIMA) and Winters models—which are discussed later in this paper—and predicted values for 2010 and 2011. Papaioannou et al. (2019) utilized multivariate time series analysis and the Hybrid Manifold Learning Principal Component Model to evaluate load management in the Greek electricity market. In contrast, Elamin and Fukushige (2018) used the Seasonal ARIMA with exogenous variables, or SARIMAX, to forecast hourly load using temperature, humidity, and seasonal dummy variables. Qiu et al. (2021) used a long short-term neural network (LSTM), a deep learning technique, to predict river water temperatures and how climate and dam construction changes can cause variations in the river's thermal regime over time. Numerous instances substantiate the use of ML and deep learning across various temperature domains. Yet, their integration within low-enthalpy geothermal systems remains in a nascent stage of advancement.

In this study, we propose applying ML algorithms for imputing a two-year gap in critical time-series data. The subsequent section will outline the data calibration and collection and a detailed description of the chosen field site. This is followed by an exposition of ML methodologies applied in the analysis and projection of the time series. In the third section, we compare the effectiveness of various ML strategies, present the results and insights derived from them, and discuss the utility and possible enhancements of these outcomes. We conclude the study by offering insights into potential expansions of this research and the application of more advanced methods and technologies in this field.

2. METHODOLOGY

2.1 Measuring Methodology and Field Site

FO-DTS systems utilize laser pulses transmitted through fiber-optic cables, with the scattered light returning to the system due to variations in the fiber's composition. This technology enables the measurement of temperatures along optical fibers spanning several kilometers, achieving a resolution of less than a meter (Tombe et al. 2020). The need for a single cable, high-sensor density and resistance to environmental interference make FO-DTS an excellent alternative to conventional temperature sensors (i.e., thermocouples and thermistors) for measuring environmental temperatures. A FO-DTS interrogator emits brief laser pulses along an optical fiber during operation. The laser pulses interact with imperfections in the silica, creating scattering photons that return to the interrogator. While most photons scatter elastically, maintaining the pulse's original frequency, a fraction undergoes non-elastic Raman scattering events, shifting to higher (anti-Stokes) or lower (Stokes) frequencies than the original pulses. The intensity of both Stokes and anti-Stokes scattering varies with the light's intensity. Notably, anti-Stokes scattering intensity changes with temperature changes along the fiber. Consequently, the temperature profile along the fiber's path is derived from the adjusted ratio of anti-Stokes to Stokes scattering amplitudes. Furthermore, the location is of the scattering events is determined by monitoring the time and scattering event returns and knowing the speed of light in the fiber, enabling detailed temperature distribution measurements (Suárez et al. 2011, van de Giesen et al. 2012). Fiber-optic cables can be arranged in various configurations, including simple single-ended, duplex single-ended, and double-ended setups to allow the calibration of the optical measurements and determining the temperature along the fiber (McDaniel et al. 2018a). At any given time, t and position, z along the fiber, the power of the Stokes (Ps) and anti-Stokes (Pas) signals are mathematically converted into temperature readings as (Herrera et al. 2018; McDaniel et al. 2018; McDaniel et al. 2018, Tombe et al. 2020, van de Giesen et al. 2012):

$$T(z,t) = \frac{\gamma}{ln\frac{P_S(z,t)}{Pas(z,t)} + C - \int_0^z \Delta\alpha \,.\, z'\, dz'} \quad (1)$$

where γ is the energy difference between the incoming and backscattered Raman photons, C is a dimensionless coefficient that is determined by the characteristics of the input laser in the interrogator, and Δα is the differential attenuation between the anti-Stokes and Stokes backscattered signals within the fiber. Calibration processes rely on cable segments at known temperatures to determine γ, C, and Δα. Our approach employs a dynamic, long-term double-ended calibration method within an intricate network of FO cables embedded in a low-temperature geothermal exchange field (Dutt Attri et al. 2022, McDaniel et al. 2018a, McDaniel et al. 2018b). The long-term data analysis allows us to observe temporal temperature fluctuations in the field and their possible implications. Additionally, we examined variations in the temperatures over time, potential sources of error and noise in the calibration process, and the effects of cooling-dominated loads on ground temperature and the system's overall efficiency.

We deployed a FO-DTS array on the campus of large software company in the Midwest region of the United States. This campus has 19 major office buildings and accommodates over 13,000 employees. The site operations generate significant heat, necessitating a robust cooling system (Dutt Attri et al. 2022, Tinjum et al. 2022). The company aims to achieve energy neutrality by utilizing wind and solar power complemented by four geothermal borefields and a supplemental cooling pond. Collectively, the campus manages 6,172 U-pipe geothermal heat exchangers (GHXs) across four borefields to deliver 48.5 MW of cooling power. With additional closed-loop exchanges

in a 2.2-ha stormwater management pond and an 8.5-ha, 14-m-deep Quarry Lake, the total capacity reaches 92 $MW_{thermal}$. The largest of the borehole fields, borefield 4, is responsible for over half of the total ground-based cooling capacity and is equipped with temperature monitoring wells (TMWs) that house fiber-optic cables extending to their base. A sketch of the borefield is presented in Figure 1(a). The red circles indicate the location of sentry wells containing FO loops for ground temperature readings. Blue squares mark the position of piezometric wells containing fiber optic loops and are connected to a piezometer screen in shallow and deep aquifers, allowing for water pressure measurements.

The geology of the Borefield 4 is documented in Figure 1(b) (Clayton 1989, Cline 1965, Hart et al. 2022, Özdoğan Dölçek et al. 2017). Sediments and rock layers begin with a 10-m top layer of gravel and sand overlaying the Prairie du Chien, a hard dolomite rock with chert nodules and oolites. Next, there is a 10-m layer of clean quartz sandstone known as the Jordan formation. Underneath the Jordan is the St. Lawrence formation (a marine sandstone and dolomite group), followed by the Tunnel City Formation, which is 35 m of sandstone with clay and feldspathic cement, reducing its thermal conductivity. Beneath this is the Wonewoc sandstone, a cleaner quartz-rich sandstone. At the base is the Eau Claire Formation, a group consisting of dolomitic shale with low thermal conductivity, making it the least heat-conductive layer in borefield 4 (Herrera et al. 2018, McDaniel et al. 2018a).

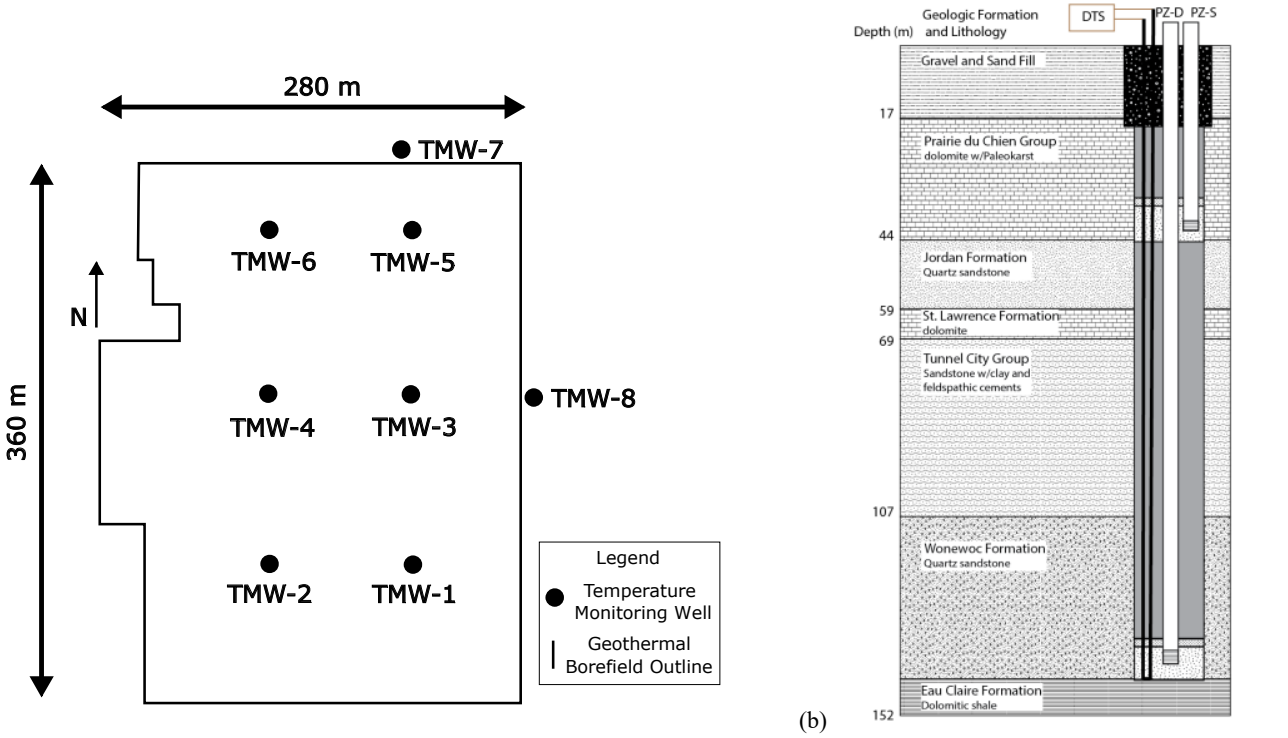


Figure 1: (a) Design for borefield 4 (2596 GHX wells) showing fiber-optic temperature monitoring well locations; (b) Borefield 4 geologic profile (after Hart et al. 2020 and Heeg et al. 2024.)

2.2 ARIMA and Seasonal ARIMA Models

To interpret the temperature data from Borefield 4, we deployed several time series analysis techniques. The ARIMA model represents a group of models that are widely applied for fitting and predicting time series data (Chaturvedi et al. 2022). In the case of stationary time series, the ARIMA model functions as a linear regression equation. This model predicts a value (the dependent variable) using both lagged terms of the forecasted values and lagged terms of the residuals (Noureen et al. 2019). The ARIMA model category is often described using the notation (p, d, q), encompassing three key elements, where p is the autoregressive (AR) term, d is the integrated (or differencing) term, and q is the moving average (MA) term. In these models, the extent of temporal correlation in the time series influences the AR and MA components, while the differencing aspect converts nonstationary series into stationary ones (Lai and Dzombak 2020). The ARIMA model is represented in equations (2) and (3) (Chaturvedi et al. 2022), where y_t is the observation at time t.

$$\boldsymbol{y'_t = c + (\phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p}) + (\theta_1 \varepsilon'_{t-1} + \cdots + \theta_q \varepsilon'_{t-q}) + \varepsilon_t} \quad (2)$$

$$\left(1 + \phi_1 B + \cdots + \phi_p B^p\right)(y_t - \mu) = \left(1 + \theta_1 B + \cdots + \theta_q B^q\right)\varepsilon_t \quad (3)$$

Here, c is a constant (the drift term when d=1), $(\phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p})$ is the AR part with order p, whereas $(\theta_1 \varepsilon'_{t-1} + \cdots + \theta_q \varepsilon'_{t-q})$ represents the MA with order q; ϕ and θ are the regression weights for the lagged observations and errors terms, ε_t is an error term for random background noise at time t, and y'_t is the differencing series. The first order differencing is shown in equation (4):

$$y'_t = y_t - y_{t-1} \quad (4)$$

Now, equation (3) is just a neatly rearranged form of equation (2), where B is called a backshift operator that represents the differenced part such that the first order difference can be written as $y_t' = (1 - B)y_t$, an hence, $B^p y_t = y_{t-p}$ and μ is the mean of all y_t'.

The limitation of the ARIMA model is its inability to represent seasonality, which limits it in its use for time series where seasonality effects are significant. An extension of ARIMA, the SARIMA model (Alharbi and Csala 2022), is used to jointly model seasonal and non-seasonal data for a univariate time series. This advancement allows the handling of both seasonal and non-seasonal data jointly for a univariate time series analysis. The SARIMA model is expressed as the SARIMA (p, d, q) (P, D, Q, s) model, where 'P' is the order of the seasonal autoregressive component, 'D' is the order of seasonal differencing, 'Q' represents the order of the seasonal moving average component, and 's' indicates the total number of observations within a time frame. The seasonal components of SARIMA (P, D, Q, s) mirror the non-seasonal elements found in ARIMA, but they apply to shifts corresponding to the seasonal cycle. Similarly, the equation for seasonal ARIMA or SARIMA can be represented as (Noureen et al. 2019):

$$\left(1 + \phi_1 B + \cdots + \phi_p B^p\right)\left(1 + \Phi_1 B^S + \cdots + \Phi_P B^{PS}\right)(y_t - \mu) = \left(1 + \theta_1 B + \cdots + \theta_q B^q\right)\left(1 + \Theta_1 B^S + \cdots + \Theta_Q B^{QS}\right)\varepsilon_t \quad (5)$$

We implement the following workflow to assess the feasibility of SARIMA model for our data, determine initial model parameters, optimize, and lastly, validate the model fit (Noureen et al. 2019, Minaravesh and Aydin 2023):

I. *Data Visualization:* We plot the data to identify any patterns of seasonality and trend. This visual analysis is crucial as it helps in understanding the underlying patterns in the time series.

II. *Differencing for Seasonality and Trend:* If seasonality or trend is present in the data, we perform differencing. For seasonality, use differencing with a lag that corresponds to the seasonal period (lag S). For trend, we typically use a first-order differencing. If both seasonality and trend are present, we may need to do both types of differencing.

III. *Analyzing Differenced Data with Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF):* We use the ACF and PACF to analyze the differenced time series. These tools help in determining the order of the MA and AR terms for both seasonal and non-seasonal components.

IV. *Model Estimation:* We estimate the SARIMA model parameters. This involves identifying the best combination of p, d, q terms in ARIMA and the seasonal P, D, Q terms in SARIMA.

V. *Residual Analysis:* We analyze the residuals of the fitted model to ensure that it adequately captures the information in the data. The residuals should be independent and identically distributed with a mean of zero and constant variance.

VI. *Model Validation and Forecasting:* Finally, we validate our model by checking its performance on unseen data and use it for forecasting future values.

This process is iterative and may require several rounds of model fitting and evaluation to find the best SARIMA model for your data. After establishing the appropriate orders and estimating the coefficients for a specific time series by adapting the model to historical data; thus, generation of both point and interval forecasts becomes possible.

2.3 Holt Winters' Exponential Smoothing

Exponential smoothing in time series analysis prioritizes recent data by allocating higher weights to newer observations and progressively smaller weights to older ones. This method exponentially diminishes the influence of older data, thereby relying more heavily on recent observations for forecasting (Marera Supervisor and Beichelt 2016). The simple exponential smoothing is typically employed for short-term forecasting, often up to a month. Equation (6) shows the simple exponential smoothing governing formula with only one smoothing parameter 'α ' is to be determined. determines the emphasis between recent and historical data. A higher α (closer to 1) gives more weight to recent observations, while a lower α (closer to 0) increases the weight of older data. This setting adjusts the model's sensitivity to changes in the time series (Lai and Dzombak 2020).

$$S_i = \alpha x_i + (1 - \alpha)S_{i-1} \quad where, 0 < \alpha < 1 \quad (6)$$

This method operates under the assumption that the time series exhibits a stable mean, indicating the absence of trends or seasonality in the data. The approach employs a specific formula to apply this smoothing technique, focusing on maintaining a consistent mean value throughout the series. To include the trend component of the time series along with the level, the double exponential smoothing, or Holt's method can be utilized (equations (7)):

$$\text{Exponentially smoother series: } S_i = \alpha x_i + (1 - \alpha)\ (S_{i-1} + T_{i-1})\ \ where, 0 < \alpha < 1 \quad (7.1)$$

$$\text{Trend estimate: } T_i = \beta(S_i - S_{i-1}) + (1 - \beta)T_{i-1} \quad where, 0 < \beta < 1 \quad (7.2)$$

The level (S_t) in is computed as the end-of-period smoothed value, representing the series' current state. Concurrently, the trend component (T_t) is derived as the smoothed average rate of change at each period's conclusion. This model incorporates two distinct parameters, one each for estimating the level and trend, enabling a nuanced adjustment to both the immediate value and the directional movement of the time series data (Dalimunthe and Aldila 2019, Lai and Dzombak 2020). To improve upon on the Holt's methodology and incorporate the seasonal trends in the time series analytics, Winters added a third parameter to deal with seasonality (Marera Supervisor and Beichelt

2016). Holt-Winters models, established for time series forecasting, come in two seasonal variations: additive and multiplicative. The additive version is best for time series with a consistent seasonal variation, where the size of the seasonal pattern does not change over time. On the other hand, the multiplicative version fits time series where the seasonal pattern grows or shrinks in proportion to the data's level, adapting to changes in the scale of the time series (Almazrouee et al. 2020). This study utilizes additive seasonality and the governing equations to calculate smoothed series are shown in equations (8.1 to 8.3). Subsequently, equation (8.4) allows us to forecast 'h' steps ahead from time 't' using additive seasonality (Lai and Dzombak 2020, Marera Supervisor and Beichelt 2016).

$$\text{Exponentially smoother series: } S_i = \alpha(x_i - p_{i-k}) + (1-\alpha)\,(S_{i-1} + T_{i-1}) \;\; where, 0 < \alpha < 1 \tag{8.1}$$

$$\text{Trend estimate: } T_i = \beta(S_i - S_{i-1}) + (1-\beta)T_{i-1} \quad where, 0 < \beta < 1 \tag{8.2}$$

$$\text{Seasonal factors: } p_i = \gamma(x_i - S_i) + (1-\gamma)p_{i-1} \quad where, 0 < \gamma < 1 \tag{8.3}$$

And the future predictions after 'h' time steps can be made as $\hat{x}_{i+h}$:

$$\hat{x}_{i+h} = S_i + h\,T_i + p_{i-k+h} \tag{8.4}$$

where 'k' represents the seasonality or the length of one seasonal period, for example, for quarterly data k = 4, $'\gamma'$ is the smoothing coefficient of the seasonal term. The additive model is optimal when seasonal variations are consistent and unchanging over time. This model assumes that seasonal effects are added to the trend level, remaining constant in magnitude irrespective of the level of the time series. Further sections discuss the utilization of these methods on a temperature time series and how the results vary.

3. RESULTS AND DISCUSSION

3.1 Temperature Data

Our methodology uses FO-DTS, leveraging dynamic, double-ended calibration for continuous field monitoring to compute temperature profiles. Figure (2) illustrates the temperature distribution for Temperature Monitoring Well (TMW-1) depicted on a depth-time plot alongside the geological strata of the site. Malfunctions and operational disruptions caused the missing data intervals during the COVID-19 pandemic. The data reveal subsurface temperatures' relative stability compared to surface readings, which are susceptible to atmospheric variations. Notably, seasonal effects are discernible underground, influenced by the GHX system's operational dynamics, in which temperatures vary between summer and winter. Our system is cooling-dominated, meaning more heat is stored in the reservoir during the hot summers than is extracted in cold weather on an annual basis. The impact of this can be seen as an overall increase in temperatures over the period of more than seven years. A notable cooler zone at around 30 m deep corresponds to a water flow in the karst within the dolomite formation (Hart et al. 2022). We study the temperature variations and the impact of various factors on the subsurface temperatures, too. In this paper, we are more focused on reconstructing a two-year interruption in temperature recordings from 2019 to 2021 using advanced regression techniques, specifically ARIMA and Holt-Winters triple exponential smoothing models.

The analysis commenced at an 80-m depth within TMW-1, a strategic depth that not only represents the median depth range but also parallels the thermal variations observed at other strata. This is illustrated in Figure 3, which plots the temperature changes over time at this depth. The time series analysis indicates a minor downtrend in temperatures leading up to 2019, followed by an uptrend from 2021 onwards. A contributing factor to this temperature shift may be the expansion of the campus workforce and infrastructure, growing from 10,000 employees in 2015 to approximately 13,000 in 2023, with the majority stationed on-site. It is also noted that data variability was more pronounced before 2016, which is attributable to field startup and the formative stages of the FO-DTS technology and fine-tuning of its calibration process. Therefore, the data preceding 2016 was excluded from our study to ensure a more accurate and reliable analysis. Daily average temperatures were also calculated to study the seasonality effect on the data every year.

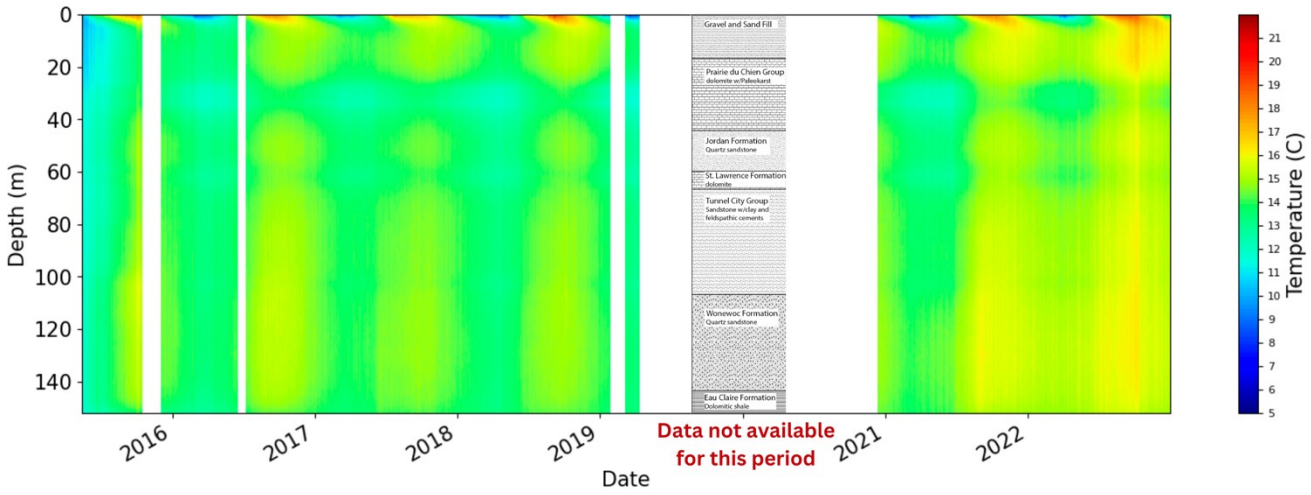


Figure 2: Borehole temperature variation for TMW-1 with depth from 2015-2022

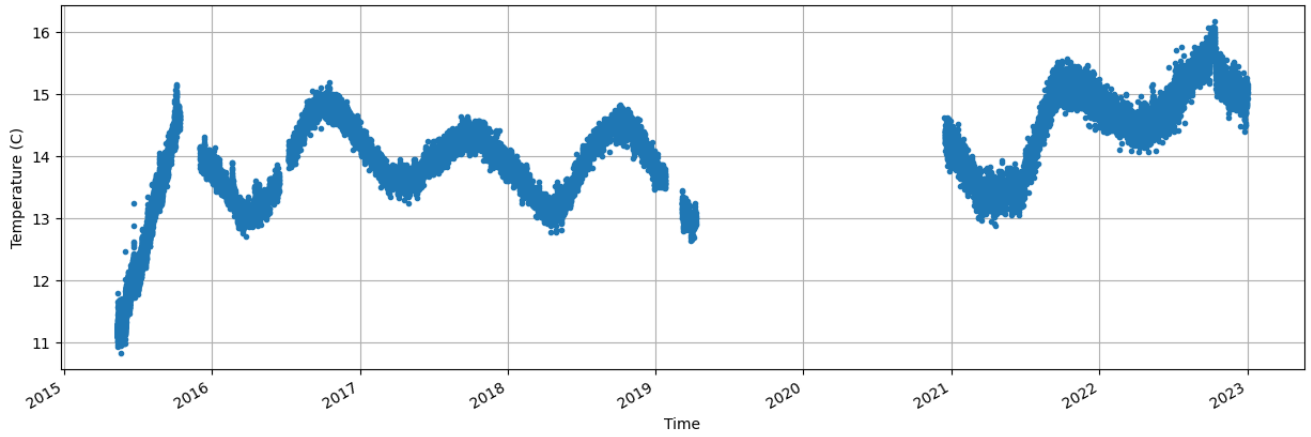


Figure 3: Borehole temperature variation at 80 m depth for TMW-1 from 2015-2022

3.2 Exploratory Statistical Analysis

We utilized an exploratory approach to uncover significant characteristics within the dataset. For the daily average temperatures from the year 2016 to 2022, Figure 4 shows the quantile-quantile (Q-Q) plot. This plot assesses whether the distribution of a dataset is approximately normally distributed. This graphical tool plotted the quantiles of the time series data against the quantiles of a standard normal distribution. The data in Figure 4 were not normally distributed or had more outliers than a normal distribution would have had. To check the stationarity of our time series, the Augmented Dickey-Fuller (ADF) test was done. The test focuses on testing the null hypothesis that a unit root is present in an autoregressive model of the time series, which implies non-stationarity (Noureen et al. 2019). The ADF test statistics and the corresponding critical values for different significance levels were as follows:

- Test Statistic: -1.569
- Number of Lags Used: 4
- Critical Values: -3.434 (1%), -2.863 (5%), -2.568 (10%)
- p-value: 0.499

The test statistic of -1.569 failed to fall below the critical values of -3.434, -2.863, and -2.568 for the 1%, 5%, and 10% levels, respectively. Furthermore, the p-value of 0.499 was significantly above the conventional threshold of 0.05. These findings support the null hypothesis, indicating that the data is likely non-stationary, and therefore, any analyses conducted on the raw data without addressing this issue may yield unreliable inferences. To proceed with ARIMA modeling or any other time series forecasting methods that assume stationarity, differencing, or other data transformations, it will be necessary to stabilize the mean of the time series across time.

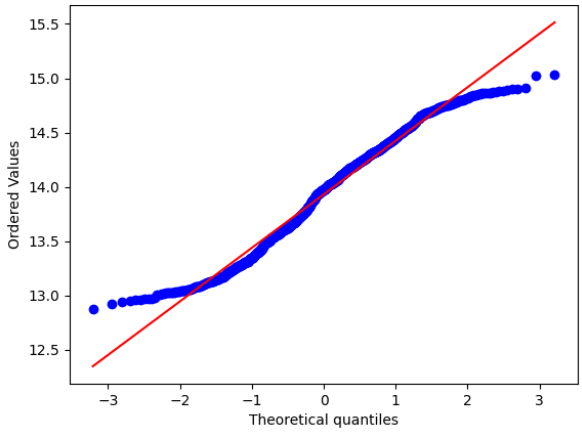


Figure 4: The Quantile-Quantile (Q-Q) plot for temperatures at 80-m depth

In the context of our study, the non-stationarity in the time series data will prompt us first to apply different techniques to achieve stationarity before fitting a predictive model. Hence, we went on to take the first-order difference of our series, shown in Figure (5). The ADF test statistics for the differenced series provided strong evidence that the time series is stationary, as the test statistic value of -3.897 was lower than the critical values, and the p-value was 0.00205 (<<0.05), indicated a high likelihood that the results are not due to random chance.

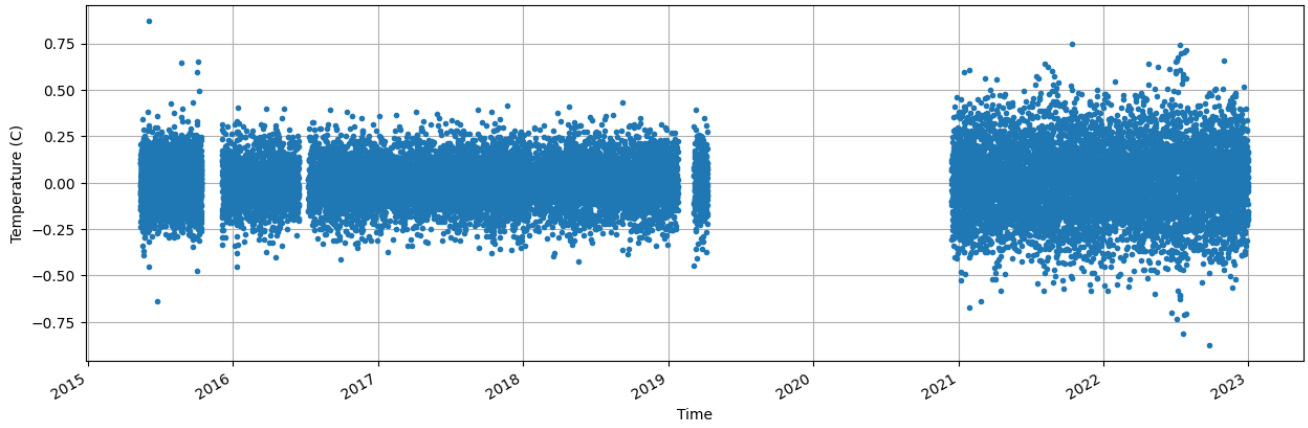
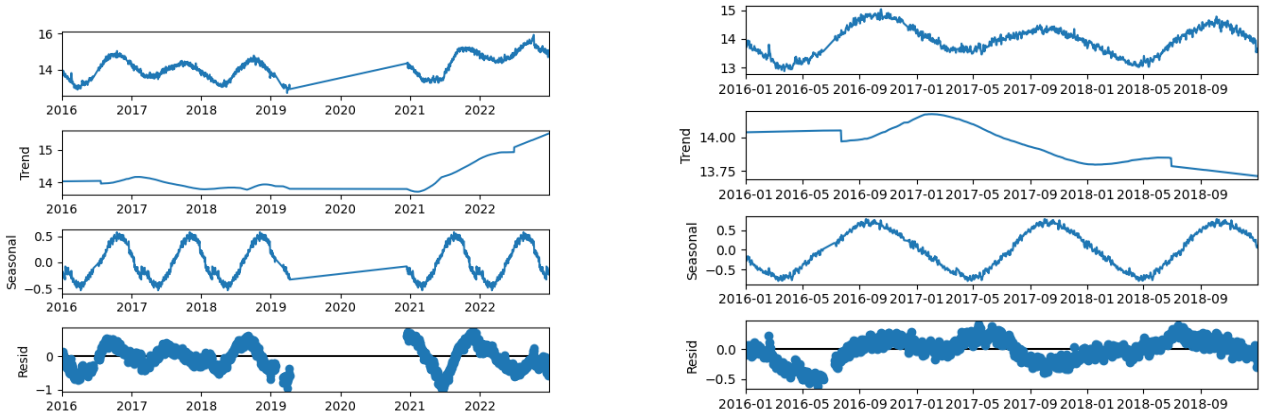


Figure 5: First order differenced series for borehole temperatures for TMW 1 at 80 m depth

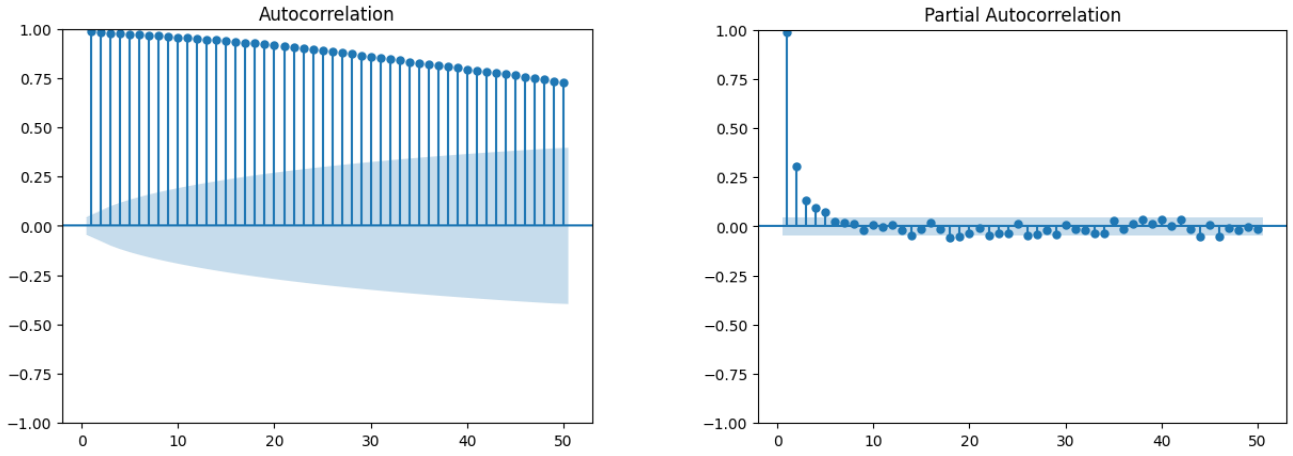
Figure 6a represents the decomposition of the TMW-1 80-m-depth temperature time series averaged daily. The trend for this well temperature shows an increase over the years. The seasonal component exhibits a consistent yearly pattern, which likely reflects the influence of seasonal weather and campus loads. Meanwhile, the residuals show some variability, particularly around 2021 when sporadic influences are not captured by the trend or seasonal components. The residuals suggest that occasional anomalies or other unexplained variations are affecting the well's temperature beyond the regular seasonal fluctuations and the general upward trend. Subsequently, Figure 6b shows the seasonal decomposition of the series only from January 2016 to December 2018. We used this portion of the data set to model and predict for the missing period; i.e., years 2019–2020. The trend for this period was fluctuating but eventually decreasing over time, whereas seasonal patterns were evident and repeated annually, as seen in the third panel, reflecting expected periodic changes in temperature. The residual component was consistent, hovering around zero, suggesting that the trend and seasonal components have accounted for most of the data's variability, leaving only minor unexplained noise.



(a) **(b)**

Figure 6: Seasonal decomposition of the daily averaged temperature values, (a) Jan 2016 – Dec 2022, (b) Jan 2016 – Dec 2018

We then studied the autocorrelation among the data points and Figure 7 demonstrates the autocorrelation (ACF) and partial-autocorrelation (pACF) graphs for our data. While we could see a strong autocorrelation among the subsequent data points (Figure 7(a)), leading us to plot the pACF, which measures the correlation between the time series and its lags after eliminating the variations explained by the intervening lags. The pACF plot (Figure 7(b)) had a significant spike at the first lag and then cut off, which typically suggested an AR process, meaning a first-order autoregressive model might be a good fit for this time series. The subsequent lags were within the confidence interval, indicating that they did not provide additional explanatory power beyond the first lag. In summary, the ACF graph indicated a strong seasonal pattern in the well temperature time series, while the pACF graph suggested that an autoregressive model of order 1 might be suitable for modeling the data after accounting for seasonality. These insights were particularly useful when building predictive models for this time series data, as they helped in selecting appropriate terms for ARIMA modeling.



(a) **(b)**

Figure 7: (a) Autocorrelation for the TMW-1 temperature values (2016–2022), (b) Partial Autocorrelation for the TMW-1 temperature values (2016-2022)

3.3 ARIMA Analysis

We used the data from January 2016 to December 2018 to model with ARIMA and test as a base model for our study. The ARIMA model is based on the lowest Akaike Information Criteria (AIC) values, and Augmented Dicky Fuller (ADF) test to determine the order of differencing. AIC evaluates models based on accuracy and model complexity, favoring simpler models that achieve a better fit with fewer features (Chaturvedi et al., 2022). The ADF test is commonly employed to assess the stationarity of a time series to examine the presence or absence of a unit root in the characteristic equation, indicating non-stationarity or stationarity, respectively (Noureen et al. 2019). The best model was identified as ARIMA (1,1,4) with the lowest AIC values among the tested models and this combination of parameters provided the best balance between model complexity and fit for your data. The model specified as ARIMA (1,1,4) indicated one autoregressive term, one differencing pass to ensure stationarity, and four moving average terms. The fit of the model was quite good, as evidenced by a log likelihood of 1038.975. Figure 8 presented is a compilation of diagnostic plots for an ARIMA model, used to evaluate the quality of the model fit. The plot in Figure 8(a) is of standardized residuals and indicates no apparent patterns, suggesting a good model fit. Figure 8(b) shows residuals aligned with a kernel density estimate and a standard normal distribution, indicating that the

residuals are approximately normal, an assumption of the ARIMA model. This is important for the reliability of the model's confidence intervals and p-values. The Normal Q-Q plot in Figure 8(c) compares the residuals to a perfect normal distribution. The points largely follow the reference line, except for the tails, which may indicate outliers or heavy-tailed distribution of residuals, a common occurrence in real-world data. Lastly, the bottom-right correlogram or plot (d) assesses the autocorrelation of the residuals at different lags. The plot shows that the autocorrelations are within the confidence interval, implying that the model has successfully captured the time series' patterns and that the residuals are essentially random, which is ideal for a well-specified model. Overall, the diagnostics suggest a satisfactory fit, with minor concerns about residual normality.

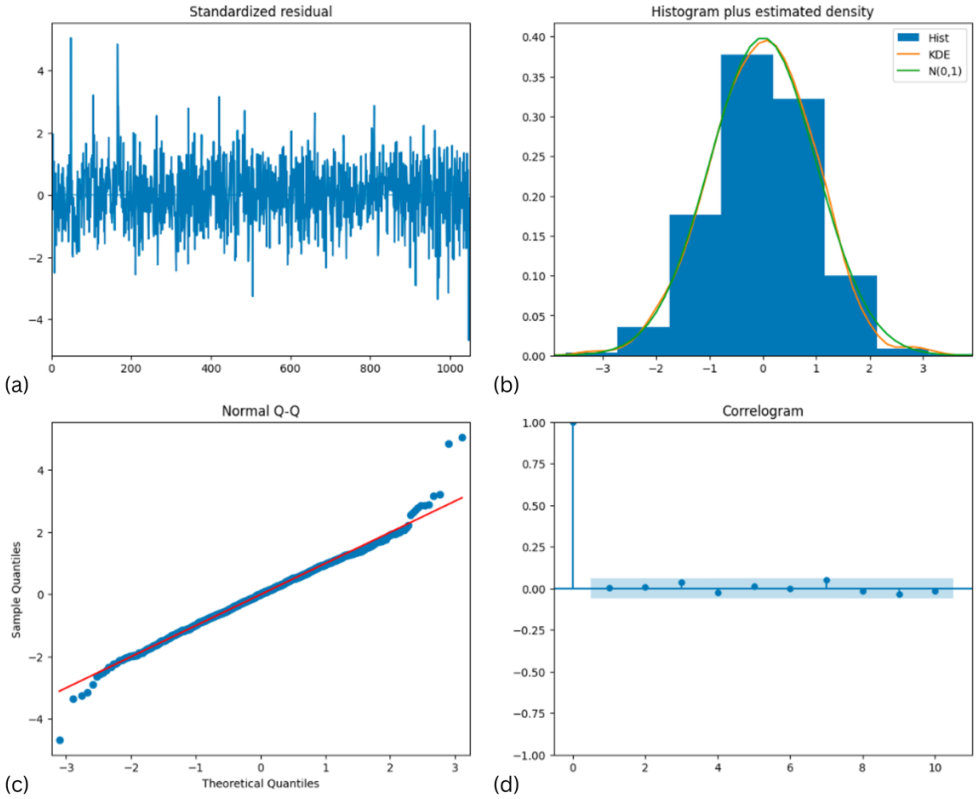


Figure 8: Diagnostic plots for the ARIMA (1,1,4)

The model was trained on the training set from January 2016 to November 2018, and then predictions were made using rolling forward method using the previously predicted values. Figures 9 and 10 show temperatures for the whole dataset and predicted temperatures for the test set for TMW-1. The root mean square error (RMSE) for these predictions was 0.197, which is good enough for a rolling forward prediction using ARIMA.

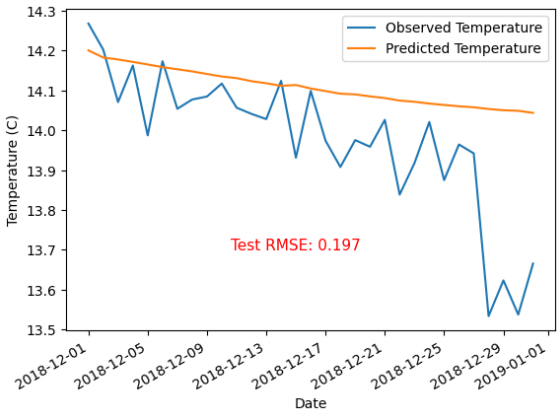


Figure 9: Predicted and observed temperature values for TMW-1 at 80-m depth using ARIMA from Oct 2018 to Dec 2018

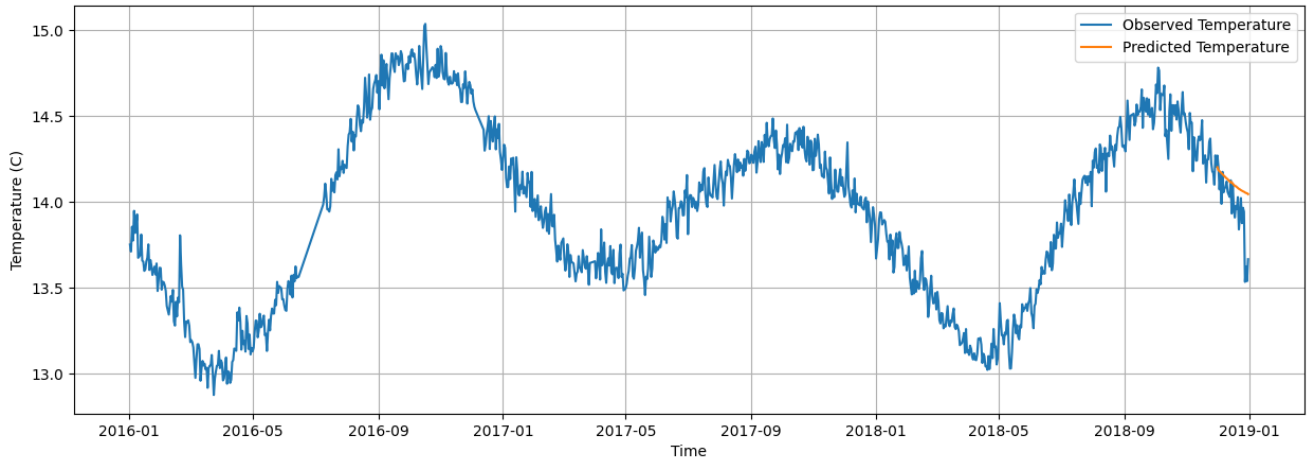


Figure 10: Predicted and observed temperature values for TMW-1 at 80-m depth using ARIMA from Jan 2016 to Dec 2018

Since the RMSE value for predictions and the model looked promising, we then forecast the temperatures for the year 2019. Figure 11 shows the predicted values starting from January 2019 until December 2019. The forecasted values followed the trend, especially for the first few months and then did not really follow the pattern expected. This made sense since our ARIMA (1,1,4) model does not include seasonality. Since our temperature measurements are averaged for daily observations, the potential seasonality component would be 365, which we use as our seasonality parameter. We decompose and remove seasonality by subtracting the average seasonal pattern from the original time series. Using these averages, the seasonal adjustment is performed by subtracting the seasonal signal. The ARIMA model is applied to model correlations at different spatial lags on the seasonality-adjusted data for analysis and forecasting. The seasonality component is superimposed on the autoregressive forecasts defined using ARIMA.

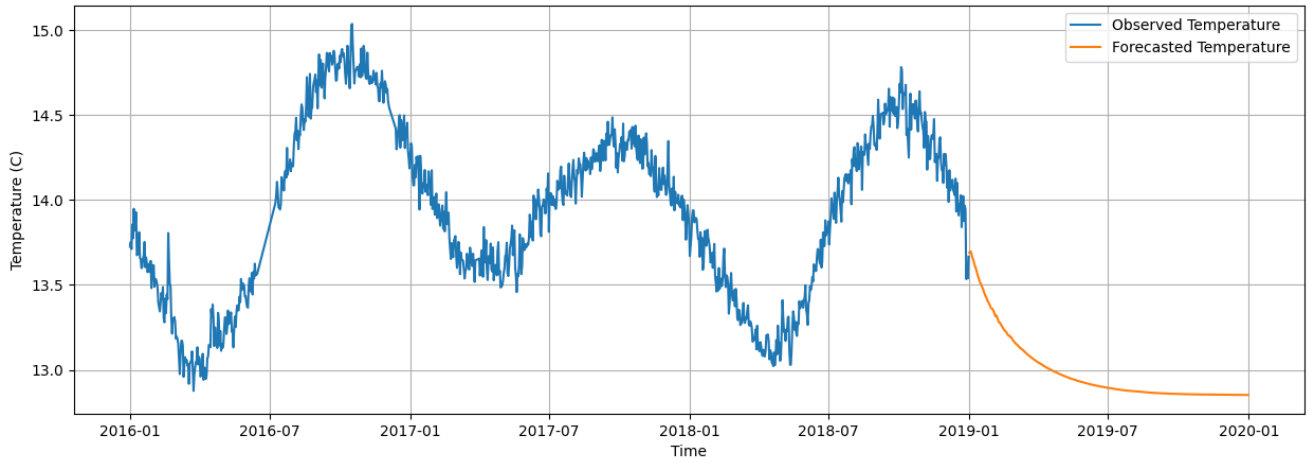


Figure 11: Temperature predictions for 2019 using ARIMA (1,1,4) for TMW-1 at 80-m depth.

3.4 Holt Winters' Smoothing

The exponential smoothing (section 2.3) is a technique that models the time series data as a linear combination of time lags with smaller lags (more recent time steps) having higher weight than the larger lags. by giving more weight to recent observations and progressively less weight to older ones. There are Three primary variations of exponential smoothing techniques exist, namely: single, double, and triple exponential smoothing. Single smoothing is applicable to time series homoscedastic time series that do not have a trend component. Double smoothing models the impact of trend, while triple smoothing has a lag, trend, and season component. Thus, for our dataset, triple exponential smoothing, also known as Holt Winter's Smoothing, is the most appropriate.

We split the dataset was taken from 2016 to 2018 and we divided our dataset into training and testing sets, parting where we used the on November 30th, 2018, time step to split the data. We initially utilized Simple Exponential (Single) Smoothing, premised on the notion that which assumes that the time series oscillates around a constant mean, indicating stationarity, and lacking any trend or seasonal variations. The single smoothing result is depicted in Figure (12) as a weighted mean using optimal smoothing coefficients inferred from using the training data. Single smoothing defines the forecast as a single value for the entire forecast horizon, as within the 80% CI in Figure 12.

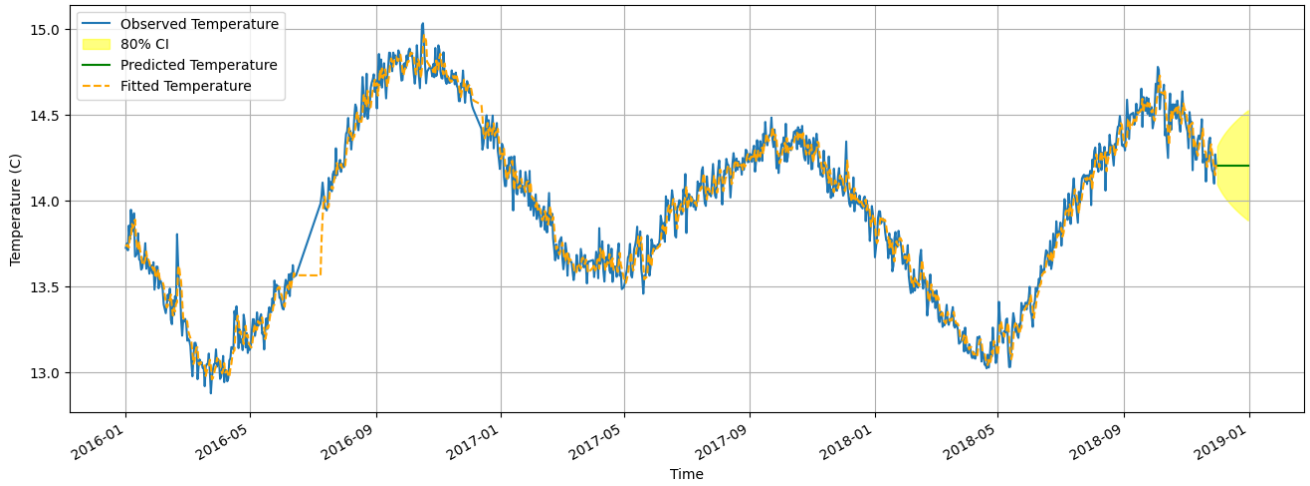


Figure 12: Predicted and observed temperature values for TMW-1 at 80-m depth using Simple Exponential Smoothing from Jan 2016 to Dec 2018

We define a forecast model with triple exponential smoothing—Holt-Winter's exponential smoothing, HW-ES—to explicitly model the trend and seasonality components to our predictions. Figure 13 and 14 demonstrates the predictions for our test periods using the HW-ES method where the impact of seasonality is observed. The RMSE value for the predictions was 0.168, which shows that the addition of the trend and seasonality components improves forecast accuracy significantly compared to the ARIMA model.

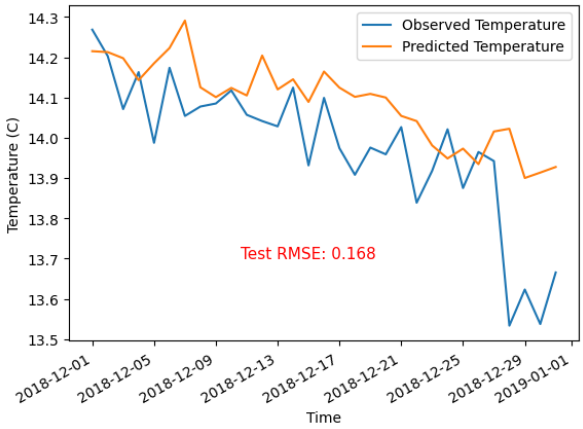


Figure 13: Predicted and observed temperatures for TMW-1 at 80-m depth using the HW-ES method from Oct 2018 to Dec 2018

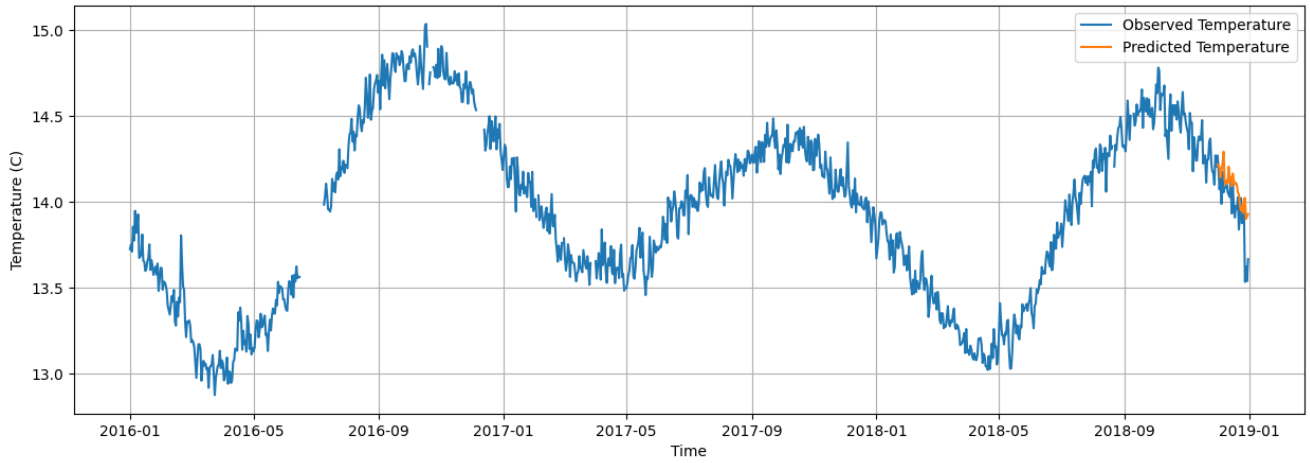


Figure 14: Predicted and observed temperature values for TMW-1 at 80-m depth using HW-ES from Jan 2016 to Dec 2018

We define the HW model to impute the missing temperature measurements between 2019 and 2021. The result in Figure 15 shows that the model captures seasonality correctly. The mismatch for 2021 is observed due to the non-linearity of the temperature increase that is dictated by what we believe to be campus operational changes early in the COVID-19 pandemic rather than natural variation under constant operating conditions.

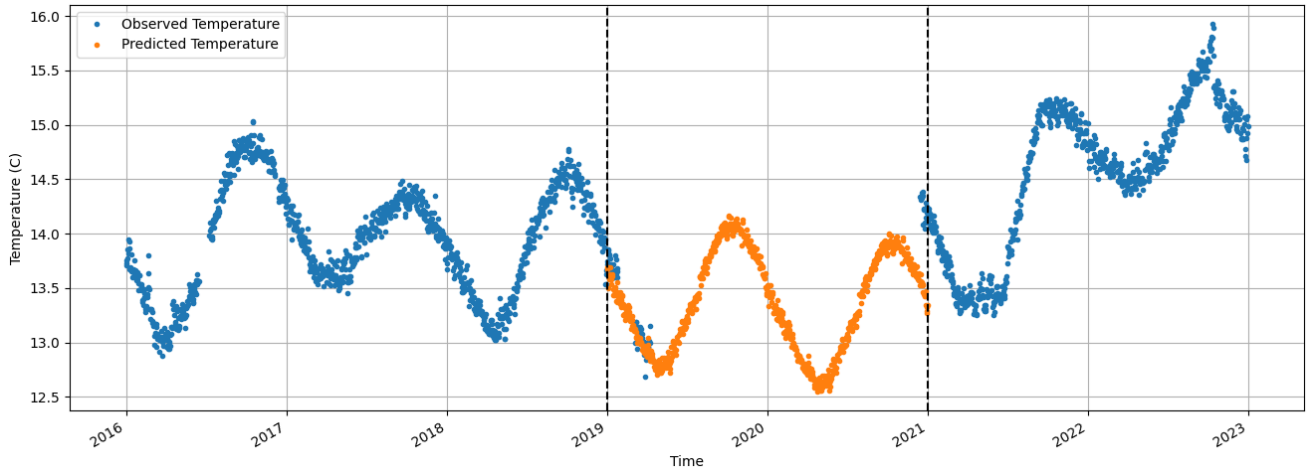


Figure 15: TMW-1 temperature data at 80-m depth from 2016 to 2022 with imputed values for 2019 and 2020 using the HW-ES method.

4. CONCLUSIONS AND FUTURE WORK

In conclusion, our results demonstrate some significant insights into the subsurface temperature dynamics of a geothermal heat exchange system. We have imputed a two-year data gap and tested the efficacy of ARIMA, and HW-ES models for forecasting. Our results indicate that the Holt-Winters smoothing model yielded more accurate predictions, with the lowest root mean squared error, adeptly capturing the gradual seasonal rise in subsurface temperatures. Both the models incorporated the trend components of the time series in the predictions, but including the seasonal component was much easier and computationally faster with the HW-ES method. Our work underscores the versatility and robustness of the time series models in forecasting geothermal heat exchange. While this provides a strong base to utilize regression analysis and machine learning for data imputation, the results can be further improved. Moving forward, the research opens avenues for further exploration into the integration of additional exogenous variables and the application of more sophisticated machine learning or deep learning algorithms to enhance predictive accuracy. The methodologies and insights gained from this study will serve as a cornerstone for future work aimed at optimizing the operation and performance of geothermal exchange fields.

ACKNOWLEDGEMENT

The authors would like to acknowledge Epic Systems Corporation for their generous contributions of time, talent, and resources in the ongoing development of this research. The Morse Company, Hooper Corporation, Salas O'Brien, JP Cullen, and Teel Plastics all donated time, materials, and/or expertise, without which, this research would not have been possible. We thank Mr. Adam McDaniel (Westwood Professional Services) for preparing the initial MATLAB scripts to process and calibrate FO DTS data. The National Science Foundation supported undergraduate research opportunities during the early phases of this project under Award # 1156674. Dr. Aydin acknowledges Taylor Geospatial Institute Award #PROJ-000323. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations.

REFERENCES

Alharbi, F.R., and Csala, D.A.: Seasonal Autoregressive Integrated Moving Average with Exogenous Factors (SARIMAX) Forecasting Model-Based Time Series Approach. Inventions, 7(4), (2022).

Almazrouee, A.I., Almeshal, A.M., Almutairi, A.S., Alenezi, M.R., Alhajeri, S.N., and Alshammari, F.M.: Forecasting of electrical generation using prophet and multiple seasonality of holt–winters models: A case study of Kuwait. Applied Sciences (Switzerland), 10(23), (2020).

Burtiev, R., Greenwell, F., and Kolivenko, V.: Time Series Analysis of Wind Speed and Temperature in Tiraspol, Moldova. Environmental Engineering and Management Journal, 12(1), (2013), 23–33.

Cadenas, E., Rivera, W., Campos-Amezcua, R., and Heard, C.: Wind Speed Prediction Using a Univariate ARIMA Model and a Multivariate NARX Model. Energies, 9(2), (2016).

Chaturvedi, S., Rajasekar, E., Natarajan, S., and McCullen, N.: A Comparative Assessment of SARIMA, LSTM RNN and Fb Prophet Models to Forecast Total and Peak Monthly Energy Demand for India. Energy Policy, (2022), 168.

Clayton, L.: Geology of Juneau County, Wisconsin Information Circular 66, Wisconsin Geological and Natural History Survey, (1989).

Cline, D. R:. Geology and Ground-Water Resources of Dane County, Wisconsin, Contributions to the Hydrology of the United States, (1965).

Dalimunthe, D.Y., and Aldila, H.: Projection and Analysis of National Energy Consumption Levels on Indonesia's Economic Growth Rate through Exponential Smoothing Approach. IOP Conference Series: Earth and Environmental Science, 353(1), (2019).

Dutt Attri, S., Heeg, E., Yilmaz, M., Tinjum, J.M., Fratta, D., and Hart, D.: Long-term Temperature Monitoring of a Campus-Scale Geothermal Exchange Field using a Fiber-Optic Sensing Array, 48th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, CA (2023), SGP-TR-224.

Elamin, N., and Fukushige, M.: Modeling and Forecasting Hourly Electricity Demand by SARIMAX with Interactions. Energy, 165, (2018), 257–268.

Florea, L.J., Hart, D., Tinjum, J.M., and Choi, C.: Potential Impacts to Groundwater from Ground-Coupled Geothermal Heat Pumps in District Scale. Groundwater, 55(1), (2017), 8–9.

Guo, Y., Wang, J., Chen, H., Li, G., Liu, J., Xu, C., Huang, R., and Huang, Y.: Machine Learning-based Thermal Response Time Ahead Energy Demand Prediction for Building Heating Systems. Applied Energy, 221, (2018), 16–27.

Hart, D.J., Tinjum, J.M., Fratta, D., Thomas, L.K., and Carew, E.L.: Radiators or Reservoirs: Heat Budgets in District-Scale Ground-Source Geothermal Exchange Fields. 47th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, CA (2022), SGP-TR-223.

Herrera, C., Nellis, G., Reindl, D., Klein, S., Tinjum, J.M., and McDaniel, A.: Use of a Fiber Optic Distributed Temperature Sensing System for Thermal Response Testing of Ground-coupled Heat Exchangers. Geothermics, 71, (2018), 331–338.

Holtzman, B. K., Paté, A., Paisley, J., Waldhauser, F., and Repetto, D.: G E O L O G Y Machine Learning Reveals Cyclic Changes in Seismic Source Spectra in Geysers Geothermal Field, (2018).

Lai, Y., and Dzombak, D.A.: Use of the Autoregressive Integrated Moving Average (Arima) Model to Forecast Near-term Regional Temperature and Precipitation. Weather and Forecasting, 35(3), (2020), 959–976.

Marera Supervisor, D.-H., and Beichelt, F.: An Application of Exponential Smoothing Methods to Weather Related Data Declaration, (2016).

McDaniel, A., Fratta, D., Tinjum, J.M., and Hart, D.J.: Long-term District-scale Geothermal Exchange Borefield Monitoring with Fiber Optic Distributed Temperature Sensing. Geothermics, 72, (2018a), 193–204.

McDaniel, A., Tinjum, J.M., Hart, D.J., and Fratta, D.: Dynamic Calibration for Permanent Distributed Temperature Sensing Networks. IEEE Sensors Journal, 18(6), (2018b), 2342–2352.

Minaravesh, B., and Aydin, O.: Environmental and Demographic Factors Affecting Childhood Academic Performance in Los Angeles County: A Generalized Linear Elastic Net Regression Model. Remote Sensing Applications: Society and Environment, 30, (2023).

Özdoğan Dölçek, A., Atkins, I., Harper, M.K., Tinjum, J.M., and Choi, C.Y.: Performance and Sustainability of District-Scale Ground Coupled Heat Pump Systems. Geotechnical and Geological Engineering, 35(2), (2017), 843–856.

Papaioannou, G.P., Dikaiakos, C., Dramountanis, A., and Papaioannou, P. G.: Analysis and Modeling for Short-to Medium-Term Load Forecasting Using a Hybrid Manifold Learning Principal Component Model and Comparison with Classical Statistical Models (SARIMAX, Exponential Smoothing) and Artificial Intelligence Models (ANN, SVM): The Case of Greek Electricity Market. Energies, 9(8), (2019).

Qiu, R., Wang, Y., Rhoads, B., Wang, D., Qiu, W., Tao, Y., and Wu, J.: River Water Temperature Forecasting Using a Deep Learning Method. Journal of Hydrology, (2021), 595.

Rezvanbehbahani, S., Stearns, L.A., Kadivar, A., Walker, J.D., & van der Veen, C.J.: Predicting the Geothermal Heat Flux in Greenland: A Machine Learning Approach. Geophysical Research Letters, 44(24), (2017), 12,271-12,279.

S. Noureen, S. Atique, V. Roy and S. Bayne: Analysis and Application of Seasonal ARIMA Model in Energy Demand Forecasting: A Case Study of Small Scale Agricultural Load, 2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS), Dallas, TX, USA, (2019), 521-524

Suárez, F., Aravena, J.E., Hausner, M.B., Childress, A.E., and Tyler, S.W.: Assessment of a Vertical High-resolution Distributed-Temperature-sensing System in a Shallow Thermohaline Environment. Hydrology and Earth System Sciences, 15(3), (2011), 1081–1093.

Tang, F., Nowamooz, H., Wang, D., Luo, J., Wang, W., and Sun, X.: Heat Exchange Capacity Prediction of Borehole Heat Exchanger (BHE) From Infrastructure Based on Machine Learning (ML) Methods. IEEE Transactions on Intelligent Transportation Systems, 23(11), (2022), 22409–22420

Thomas, L., Tinjum, J. M., and Holcomb, F. H. Environmental Life Cycle Assessment of a Deep Direct-use Geothermal System in Champaign, Illinois. Proceedings, 45th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford. CA (2020), SGP-TR-216.

Tinjum, J.M., Yilmaz, M., Heeg, E., Fratta, D., Hart, D.J., and Attri, S.D. Energy Efficiency and Life Cycle Assessment of a District-Scale Geothermal Exchange Field. 48th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, CA (2023), SGP-TR-224.

Tombe, B. des, Schilperoort, B., and Bakker, M.: Estimation of Temperature and Associated Uncertainty from Fiber-optic Raman-spectrum Distributed Temperature Sensing. Sensors (Switzerland), 20(8), (2020).

Tut Haklidir, F.S., and Haklidir, M.: Prediction of Reservoir Temperatures Using Hydrogeochemical Data, Western Anatolia Geothermal Systems (Turkey): A Machine Learning Approach. Natural Resources Research, 29(4), (2020), 2333–2346.

van de Giesen, N., Steele-Dunne, S.C., Jansen, J., Hoes, O., Hausner, M.B., Tyler, S., and Selker, J.: Double-ended Calibration of Fiber-optic Raman Spectra Distributed Temperature Sensing Data. Sensors (Switzerland), 12(5), (2012), 5471–5485.

Walker, M.D., Meyer, L.L., Tinjum, J.M., and Hart, D.J.: Thermal Property Measurements of Stratigraphic Units with Modeled Implications for Expected Performance of Vertical Ground Source Heat Pumps. Geotechnical and Geological Engineering, 33(2), (2015), 223–238.

Yang, J.W., and Dashdondov, K.: In-Depth Examination of Machine Learning Models for the Prediction of Ground Temperature at Various Depths. Atmosphere, 14(1), (2023).