Scientific Research Publishing

Open Journal of Applied Sciences, 2025, 15(9), 2820-2845
https://www.scirp.org/journal/ojapps
ISSN Online: 2165-3925
ISSN Print: 2165-3917

Hourly Soil Temperature Variability and Estimation at Multiple Depths from Meteorological Data Using Multiple Linear Regression in a Tropical Semi-Arid Region of Burkina Faso

François Dabilgou*, Soumaila Gandema, Marcel Bawindsom Kébré*, Zamantakonè Guillaume Ki, Zakarie Koalaga

Laboratoire de Matériaux et Environnement (LAME), Université Joseph KI-ZERBO, Ouagadougou, Burkina Faso
Email: *dabilgoufrancois2016@gmail.com, gandema.s@gmail.com, *marcel.kebre@ujkz.bf, zamantakoneky@gmail.com, zacharie.koalaga@ujkz.bf

How to cite this paper: Dabilgou, F., Gandema, S., Kébré, M.B., Ki, Z.G. and Koalaga, Z. (2025) Hourly Soil Temperature Variability and Estimation at Multiple Depths from Meteorological Data Using Multiple Linear Regression in a Tropical Semi-Arid Region of Burkina Faso. *Open Journal of Applied Sciences*, **15**, 2820-2845.
https://doi.org/10.4236/ojapps.2025.159188

Received: August 21, 2025
Accepted: September 21, 2025
Published: September 24, 2025

Abstract

Accurate estimation of soil temperature is essential for understanding land–atmosphere interactions and supporting agricultural management in semi-arid regions. This study focuses on the analysis of hourly soil temperature variability and the performance of Multiple Linear Regression (MLR) models at depths of 10, 20, 30, and 40 cm in a tropical semi-arid soil of Burkina Faso. Using *in situ* meteorological data collected at the Tanghin (Saaba, Kadiogo) experimental site, the results reveal a progressive attenuation of temperature fluctuations with depth, confirming the buffering effect and thermal inertia of subsurface layers. A clear phase lag between soil and air temperatures was observed, increasing with depth and particularly evident beyond 20 cm, which reflects delayed heat propagation within the soil profile. MLR models performed reasonably well at shallow depths, especially when driven by a minimal set of meteorological predictors (Ta, BP, WS, RH), achieving R^2 values up to 0.75 at 10 cm. However, their predictive ability declined with increasing depth and under wet-season conditions, where soil temperature becomes increasingly decoupled from atmospheric drivers. Moreover, the models struggled to reproduce full diurnal cycles, underestimating nighttime cooling and overestimating during the rainy season, likely due to limited sensitivity to soil moisture dynamics. These findings emphasize both the potential and the limitations of MLR in capturing subsurface thermal variability.

Keywords

Soil Temperature, Subsurface Modeling, Diurnal Variability, Multiple Linear Regression (MLR), Semi-Arid Climate

1. Introduction

Soil temperature is a key variable governing the energy and water fluxes at the land-atmosphere interface. It plays a central role in the regulation of evaporation, soil respiration, microbial activity, seed germination, and root development, and thus significantly influences hydrological and agricultural processes [1]-[3]. In dry tropical regions such as Burkina Faso, located in the Sahel, accurate monitoring and forecasting of subsurface soil temperature is essential for sustainable land management, irrigation scheduling, and climate adaptation strategies. Moreover, in the context of increasing climate variability, capturing soil thermal dynamics becomes even more crucial for understanding feedback mechanisms in agro-hydro-meteorological systems [3]. In particular, having access to high-resolution soil temperature data across multiple depths is vital for estimating the components of the surface energy balance, especially the ground heat flux [4] [5]. Furthermore, with the growing frequency of heatwave events in recent years, soil temperature has become a critical variable for studying urban heat island effects and radiative processes influencing near-surface microclimates [6] [7].

Various methods have been used to monitor soil temperature, ranging from *in situ* sensors to satellite-based remote sensing. Ground-based measurements, typically obtained using thermocouples, thermistors, or advanced sensors, offer high temporal resolution and reliable data near the surface [8] [9]. However, the implementation of such ground monitoring networks remains limited due to high installation and maintenance costs, especially in resource-constrained regions like the Sahel. In many low-income countries, including Burkina Faso, the national meteorological agencies rarely monitor soil temperature as part of their routine observations. Instead, a few agrometeorological institutions attempt to measure it sporadically, often on a limited spatial scale, typically tied to research projects or specific field trials. On the other hand, remote sensing approaches have enabled the retrieval of land surface temperature (LST) and, more recently, subsoil temperature estimates at increasing depths using microwave or thermal infrared sensors. Products such as those from MODIS, SMOS, or SMAP allow soil temperature profiling at regional to global scales [5] [10]. Nevertheless, these satellite-derived datasets often suffer from low spatial resolution and reduced accuracy over heterogeneous terrains, particularly in arid and semi-arid zones like the Sahel. The absence of dense and reliable ground observations further limits the validation and bias correction of satellite products, compromising their usefulness for local-scale applications. This reinforces the need to develop and validate numerical estimation methods that can provide accurate and continuous soil temperature estimates across different depths and environmental conditions. Given these limi-

tations of direct measurement methods, especially in data-scarce regions like the Sahel, it becomes essential to rely on numerical estimation techniques to obtain continuous and spatially representative soil temperature data. In this context, various techniques have been developed to estimate soil temperature.

The most physically grounded approach involves solving the one-dimensional heat conduction equation derived from Fourier's law, with surface temperature and soil properties as boundary conditions [5] [11]-[13]. While these methods are accurate under well-defined conditions, they require detailed soil thermal properties and often fail to account for complex surface-atmosphere interactions or heterogeneous field conditions [9]. This method is also the simplest among physically based approaches. The more comprehensive and complete involves solving the coupled heat and mass transfer equations, which include both liquid and vapor water flow (using Richards' equation for unsaturated flow) and heat transfer mechanisms, often under non-equilibrium conditions with phase change between liquid and vapor [14] [15]. Advanced simulation tools such as HYDRUS (2D/3D) implement these coupled processes with high accuracy [16] [17]. However, this modeling framework requires numerous parameters, such as hydraulic conductivity, water retention curves, vapor diffusivity, and thermal conductivity, which are difficult to measure *in situ*, especially in data-scarce regions where instrumentation is limited and soil characterization is lacking. As a result, the application of such comprehensive models remains challenging in semi-arid and tropical contexts like the Sahel. To overcome these limitations, statistical models such as Multiple Linear Regression (MLR) have traditionally been employed to estimate soil temperature by establishing linear relationships between meteorological predictors and the target variable.

MLR models are computationally efficient and interpretable, making them valuable for applications with limited data or infrastructure. Asadzadeh *et al.* [8] compared three statistical and machine learning models, SARIMA, Multiple Linear Regression (MLR), and Artificial Neural Networks (ANN), to predict daily soil temperature at multiple depths (5, 10, 20, 50, and 100 cm) at the Riley Station (Oregon, USA). They showed that although MLR performed less accurately than ANN, it provided a simple, robust, and cost-effective modeling approach using easily accessible meteorological variables (e.g., air temperature, infrared radiation). In a similar vein, Cemek *et al.* [18] assessed various soil temperature estimation techniques (depths: 5 to 100 cm) in the Tokat region, a semi-arid area of Turkey, by comparing MLR with fuzzy logic models (MFIS, ANFIS) and various machine learning algorithms (RF, XGB, MLP, KNN, SVR). MLR was used as a baseline model and, although less accurate than ANFIS or RF, it remained useful for straightforward and interpretable predictions in precision agriculture. In a subtropical grassland during the summer season, Progga *et al.* [9] evaluated soil temperature at various depths (5, 10, 20, 30, and 50 cm) in a loamy soil under arid tropical conditions in Bangladesh, using meteorological parameters (minimum/maximum air temperatures, relative humidity, pressure, dew point, precipitation, sunshine duration, wind speed, evaporation) through an MLR model. At 5, 10, and 20 cm depths, the MLR model outperformed the HYDRUS-1D model, with R^2

values of approximately 0.96, 0.94, and 0.88, and MAPE values of 0.81%, 0.87%, and 1.05%, respectively. At 50 cm depth, the R^2 dropped to 0.55 and MAPE increased to about 1.25%. Their findings suggest that MLR is reliable for modeling soil temperature up to ~20 cm depth. Overall, these studies confirm that MLR remains a robust and interpretable method for estimating soil temperature in subtropical and semi-arid regions, particularly up to ~20 cm depth. More generally, it often serves as a baseline model, being simpler yet less accurate than some advanced algorithms. It was specifically selected for this study due to its interpretability, low computational cost, and straightforward implementation, which are particularly valuable attributes for potential applications in resource-limited contexts.

This study investigates hourly soil temperature variability and estimation at multiple depths (10, 20, 30, and 40 cm) in a tropical semi-arid soil of Burkina Faso using a statistical regression approach (Multiple Linear Regression, MLR) based on *in situ* meteorological data. While most research emphasizes daily or monthly scales and surface layers, subsurface hourly modeling and diurnal-nocturnal variability remain underexplored. By leveraging high-frequency meteorological observations in a data-scarce tropical context, this work evaluates the performance and limitations of MLR to identify suitable approaches for capturing subsurface thermal dynamics and supporting land surface modeling, precision agriculture, and climate resilience.

2. Materials and Methods

2.1. Site Characteristics and Data Gathering Techniques

The study was conducted in Tanghin, a village located in the rural municipality of Saaba in the Kadiogo Province of Burkina Faso (12˚22'5"N, 1˚25'0"W). The site lies in the Sudano-Sahelian climatic zone, characterized by two distinct seasons: a dry season from November to May and a rainy season from June to October. The average annual rainfall ranges from 600 to 900 mm, and the average temperature varies between 22.6˚C and 35.5˚C [19].

For ground data collection, an excavation was made to a depth of 80 cm to facilitate sensor installation and ensure accurate soil measurements. Within this soil profile, two distinct layers were identified, a subsurface soil up de 30 cm depth. The physical properties of the soil at 10 - 30 cm and 30 - 80 cm depths are summarized in Table 1. Both layers exhibit a sandy loam texture, with nearly identical porosity (≈0.41) and bulk density (~1560 kg·m^{-3}), suggesting homogeneity in the soil's water retention and thermal characteristics within the measured horizon. Vegetation was periodically cleared to minimize interference with surface sensors, particularly during the rainy season.

Table 1. Physical soil properties by profile.

Depth (cm)	Sand (%)	Clay (%)	Silt (%)	Porosity (–)	Bulk Density (kg·m^{-3})
0 - 30	68.21	11.66	20.13	0.410	1563
30 - 80	68.43	11.51	20.05	0.411	1562

The meteorological and soil data measurements were collected using instruments from Campbell Scientific Inc. (Logan, UT, USA) and METER Group. The ClimaVUE50 autonomous weather station was used to record air temperature, relative humidity, vapour pressure, barometric pressure, wind speed and direction, solar radiation, sunshine duration, precipitation, and lightning activity. A CR350-CELL215 datalogger facilitated data collection and transmission via a 4G modem. Additional atmospheric parameters derived from direct measurements include Dew Point Temperature, Wet Bulb Temperature, Actual Vapor Pressure, Saturation Vapor Pressure, Vapour Pressure Deficit (VPD), and Evapotranspiration were calculated from raw measurements using internal algorithms. To monitor some soil data, TEROS21 and CS655 sensors were installed at multiple depths (10, 20, 30, and 40 cm). These instruments enabled the continuous collection of soil temperature, volumetric water content, and matric potential. All variables were logged at three temporal resolutions: every 10 minutes, hourly, and daily. Figure 1 displays the station setup.

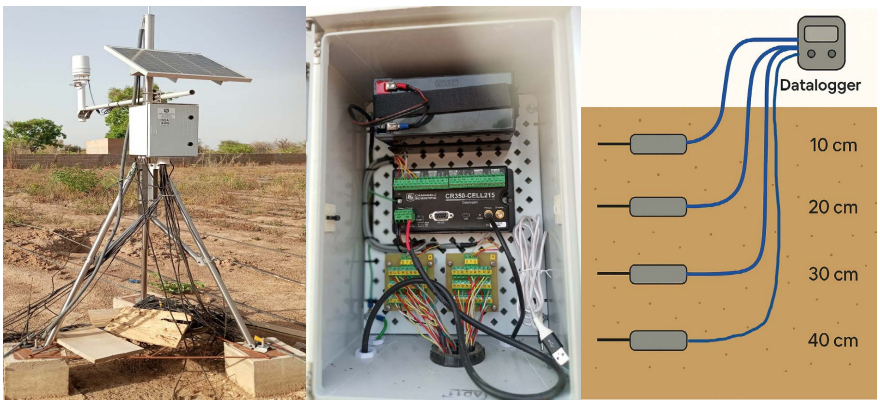

Figure 1. Field instrumentation used for soil and meteorological data acquisition at the Tanghin experimental site (Saaba, Burkina Faso). (Left) The ClimaVUE50 autonomous weather station, mounted on a metallic frame and powered by a solar panel system, records key atmospheric variables. (Center) Internal view of the control unit containing the CR350-CELL215 datalogger and wiring for real-time data transmission. (Right) Cross-sectional schematic of the sensor layout showing TEROS21 & CS655 probes deployed at 10, 20, 30, and 40 cm depths, each individually connected to the datalogger.

For the present study, we selected hourly data for the following meteorological variables, denoted by their abbreviations: air temperature (Ta, ˚C), relative humidity (RH, %), dew point temperature (DPT, ˚C), wind speed (WS, m/s), wind direction (WD, ˚), barometric pressure (BP, hPa), precipitation (RAIN, mm), solar radiation (RAD, W/m^2), and vapour pressure deficit (VPD, hPa). Soil temperature at various depths is denoted by STxy, where x indicates the depth in cm (10, 20, 30, 40) and y specifies the sensor type: t for TEROS21 and c for CS655 (e.g., ST10t = soil temperature at 10 cm from TEROS21).

The data used in this study were collected from 22 September 2023 to 08 March 2025, covering a total period of 534 days. Measurement instrument failures occurred between 01 April 2024 and 14 May 2024, and between 03 and 28 July 2024,

resulting in missing data. In total, the dataset comprises 464 days and 11,136 hours (corresponding to 11,136 data rows) for this study. Missing data entries were deleted prior to analysis. Indeed, since the MLR model itself does not account for or rely on any temporal structure, this deletion method cannot introduce a time-dependent bias in the model's estimates or performance.

2.2. Methodology

To estimate soil temperature from in situ meteorological inputs, we adopted a statistical approach based on multivariate linear regression (MLR). This classical method is widely used to model the linear relationship between a dependent variable, soil temperature in this study, and several independent predictors [8]. MLR relies on the least squares principle to minimize the sum of squared errors between observed and predicted values. For the model to yield reliable and interpretable coefficients, it is crucial that the predictors exhibit low intercorrelation, thereby avoiding redundancy and reducing the variance of estimators.

2.2.1. Variable Selection and Multicollinearity in MLR Baseline Model

To handle potential multicollinearity among the predictors [20], a two-step procedure was implemented. First, a stepwise variable selection based on the Akaike Information Criterion (stepAIC) was used to retain the most relevant variables. Then, the Variance Inflation Factor (VIF) was computed to quantify the extent of multicollinearity. Variables with VIF > 5 were treated with caution, while those exceeding a threshold of 10 were considered highly collinear; in such cases, only one representative variable was retained [9] [21]. This strategy ensures a parsimonious and stable regression model suitable for explanatory and predictive purposes.

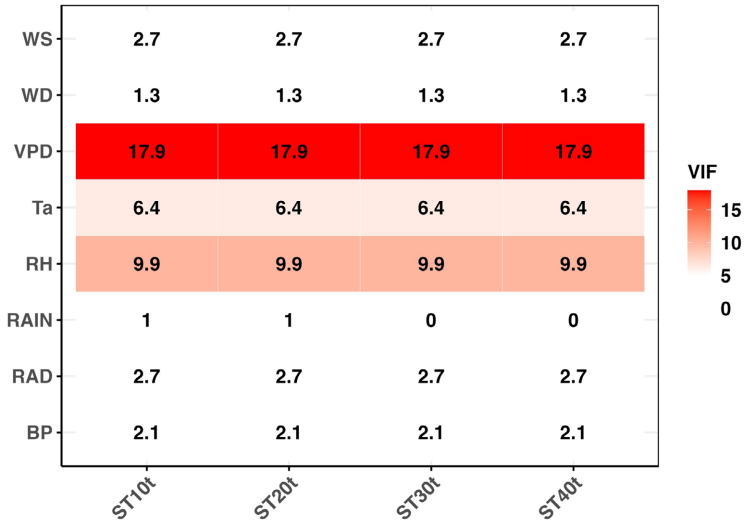

Figure 2. Variance Inflation Factor (VIF) values for soil temperature at different depths with respect to the meteorological predictors.

Figure 2 displays the VIF values associated with each meteorological variable

included in the multivariate linear regression models for soil temperature at four depths (ST10t to ST40t). The variables VPD (Vapour Pressure Deficit) and RH (Relative Humidity) exhibit very high VIFs (>10), indicating strong multicollinearity, especially for VPD (~17.9), which could severely bias regression estimates. Air temperature (Ta) also shows moderate multicollinearity (VIF ≈ 6.36). To maintain model reliability, variables with high VIFs were either excluded or carefully selected based on correlation structure and predictive relevance.

2.2.2. Development of Regression Scenario

To apply multivariate linear regression (MLR) across all soil depths, we established a set of input scenarios based on combinations of meteorological predictors, following a strategy similar to that proposed [3] for tree-based hybrid Data Mining models. For each depth, air temperature (Ta), a variable that is easy to measure and broadly available, is used as the baseline predictor. Additional variables were progressively added to account for measurement complexity and to minimize multicollinearity, assessed using Variance Inflation Factor (VIF) values after an initial selection based on stepwise AIC (stepAIC). For example, in resource-limited settings with fewer ground-based instruments, relative humidity (RH) may be favored over vapor pressure deficit (VPD), which requires two separate measurements.

Based on this approach, six (06) regression scenarios were designed to support the implementation of MLR models. These scenarios reflect a balance between data availability, predictor relevance, and acceptable levels of multicollinearity. Identical sets of predictors were retained for the shallow soil layers (10 and 20 cm) and, similarly, for the deeper layers (30 and 40 cm), following the outcomes of stepAIC selection and VIF screening. **Table 2** and **Table 3** summarize these scenarios and provide the rationale for each variable's inclusion. The two tables differ in the final variable included: S6 adds rainfall (RAIN) for shallow layers (ST10 - ST20), while it adds wind direction (WD) for deeper layers (ST30 - ST40), reflecting the differing relevance of surface vs. subsurface meteorological influences.

Table 2. Description and rationale of regression scenarios applied to shallow soil layers (10 cm and 20 cm depths).

Scenario	Selected Variables	Estimated Max VIF	Justification
S1	~Ta	6.4	Simple baseline; Ta is easy to measure.
S2	~Ta + WS	6.4	Adds wind speed (WS), easy to measure; moderate VIF (~2.7).
S3	~Ta + WS + BP	6.4	Adds barometric pressure (BP); low VIF and widely available sensors.
S4	~Ta + WS + BP + RH	~10	RH is easily measured, but has high VIF (~10).
S5	~Ta + WS + BP + RH + RAD	~10	Solar radiation is thermally relevant; VIF still acceptable.
S6	~Ta + WS + BP + RH + RAD + WD + RAIN	~10	Highest complexity before excluding VPD due to excessive VIF (~18).

Table 3. Description and rationale of regression scenarios applied to deeper soil layers (30 cm and 40 cm depths).

Scenario	Selected Variables	Estimated Max VIF	Justification
S1	~Ta	6.36	Simple baseline; Ta is directly and continuously measurable.
S2	~Ta + WS	6.36	Wind speed (WS) is easy to measure with anemometers; low VIF (~2.6).
S3	~Ta + WS + BP	6.36	Adds atmospheric pressure (BP); commonly measured with moderate VIF (~2.1).
S4	~Ta + WS + BP + RH	~9.9	RH is easier to obtain than VPD; VIF is high but still acceptable (~10).
S5	~Ta + WS + BP + RH + RAD	~10	Solar radiation is relevant for deeper soil heat flux; sensors are available.
S6	~Ta + WS + BP + RH + RAD + WD	~10	Wind direction (WD) adds value without significantly increasing multicollinearity (very low VIF).

2.3. Criteria for Model Evaluation

To evaluate and compare the performance of the different MLR scenarios tested in this study, we used a set of classical regression metrics, including the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). Additional information criteria such as the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) were computed to assess model parsimony and penalize overfitting based on the number of input variables retained. These indicators quantify both the explanatory power and the predictive accuracy of the models across soil depths

The formulation of each evaluation metric is presented as follows:

- **Mean Absolute Error (MAE)**

$$\mathrm{MAE}=\frac{\sum_{i=1}^{N}\left|y_i-\hat{y}_i\right|}{N} \tag{1}$$

- **Coefficient of Determination (R^2)**

$$\mathrm{R}^2=1-\frac{\sum_{i=1}^{N}\left(y_i-\hat{y}_i\right)^2}{\sum_{i=1}^{N}\left(y_i-\bar{y}\right)^2} \tag{2}$$

- **Root Mean Square Error (RMSE)**

$$\mathrm{RMSE}=\sqrt{\frac{\sum_{n=1}^{N}\left(y_i-\hat{y}_i\right)^2}{N}} \tag{3}$$

where N is the total number of observations, y_i is the observed value at time i, $\hat{y}_i$ is the predicted value, and $\bar{y}$ is the mean of the observed values.

- **Akaike Information Criterion (AIC)**

$$\mathrm{AIC}=-2\cdot\log\left(L\right)+2k \tag{4}$$

- **Bayesian Information Criterion (BIC)**

$$\mathrm{BIC} = -2 \cdot \log(L) + \log(n) \cdot k \tag{5}$$

where n is the number of observations, k is the number of model parameters, and L is the maximized value of the likelihood function of the model.

3. Results and Discussion

3.1. Overview of Soil Temperature Variability

Figures 3-5 present the variability of hourly soil temperature measurements during the study period. Figure 3 shows the overall temporal evolution of the different depths soil temperature measured by TEROS21 across the entire period. It highlights hourly fluctuations and missing data intervals (shaded areas), as an illustration among other recorded values. A higher level of noise is observed near the surface, while measurements beyond 20 cm appear smoother. This is due to instantaneous variability in surface-layer processes such as incoming shortwave radiation, surface-emitted longwave radiation, and rapid fluctuations in atmospheric conditions (wind, air temperature, humidity).

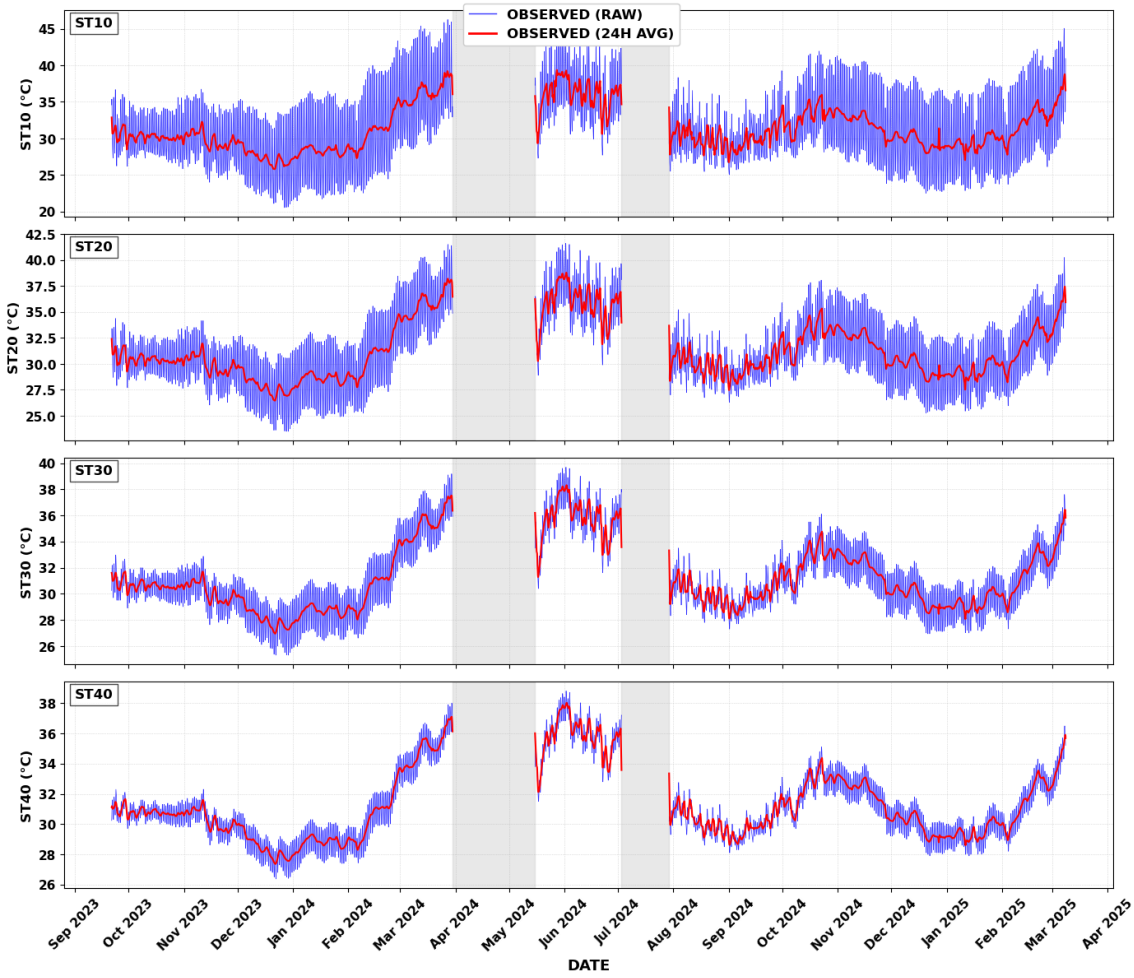

Figure 3. Hourly Soil Temperatures at the four depths with daily moving average and missing data intervals (shaded areas).

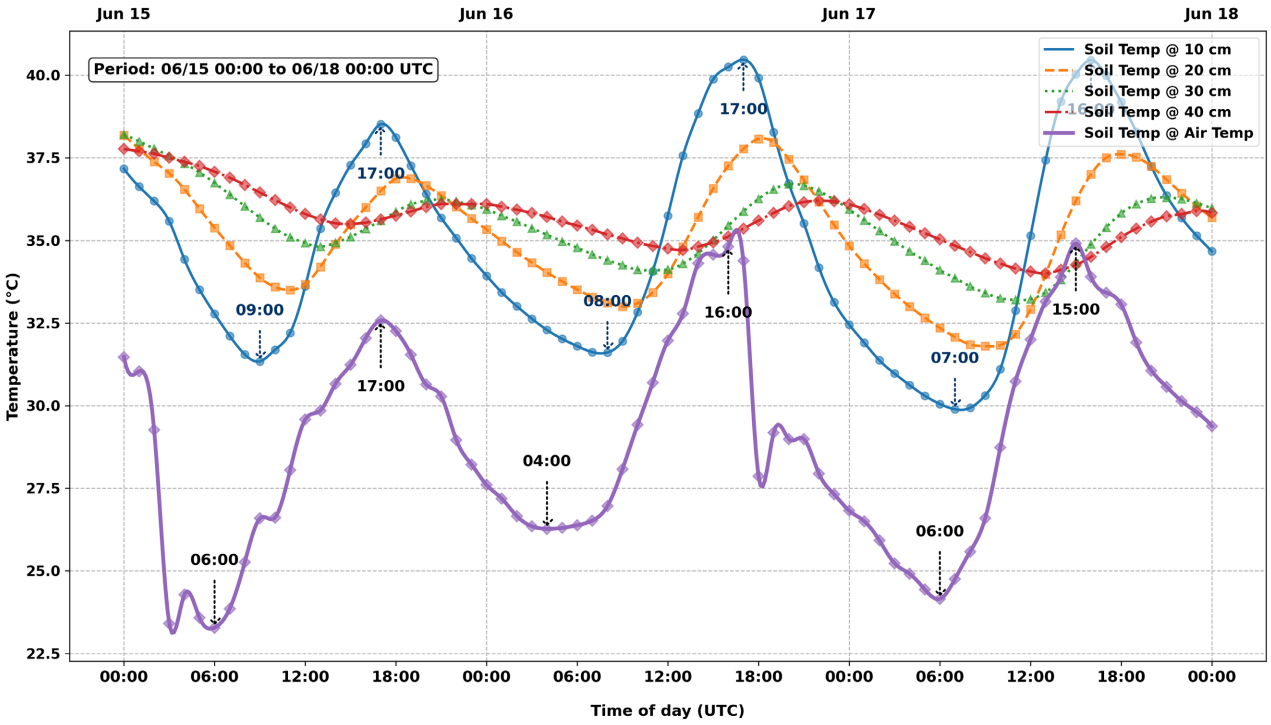

Figure 4. Multi-day diurnal dynamics of soil and air temperatures (15-18 June 2024).

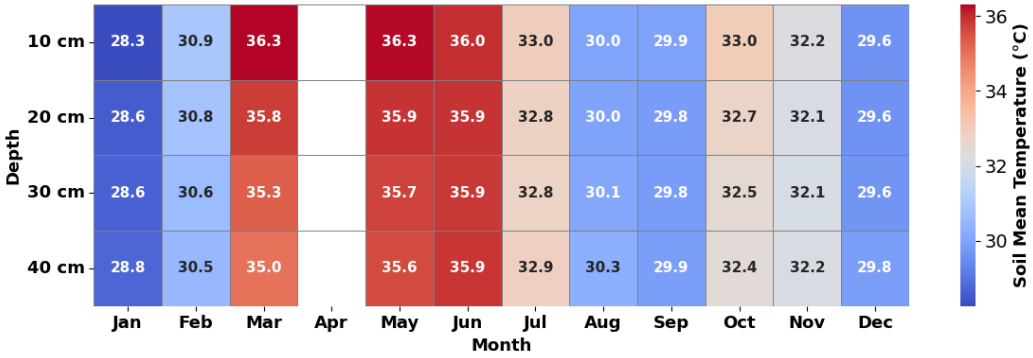

Figure 5. Monthly Mean Soil Temperatures in 2024 by depth.

Figure 4 reveals a clear attenuation of temperature fluctuations with increasing soil depth, highlighting the buffering role of deeper layers against diurnal variations. These results show that the soil temperature is higher than the air temperature at all depths, although this difference is more pronounced in the near surface layers. The comparison with air temperature confirms the soil's thermal inertia, particularly beyond 20 cm. Interestingly, during the morning hours, soil temperature continues to decrease while air temperature begins to rise. For the 10 cm depth, the timing of minimum temperature lags behind air temperature by more than two hours, as annotated in Figure 4. This lag increases with depth; for instance, at 40 cm, the soil temperature reaches its minimum almost simultaneously with the peak in air temperature, highlighting the strong thermal inertia of deeper layers during the rainy season.

This shift likely results from the delay in heat propagation through the soil profile, governed by radiative fluxes and energy exchanges at the surface. A more detailed investigation is needed to understand how this phase shift varies across the rainy and dry seasons, particularly in the context of arid tropical environments. Conversely, maximum temperatures occur almost simultaneously across layers, suggesting faster surface heating during daytime.

Figure 5 summarizes the monthly mean soil temperatures using a heatmap, providing insight into seasonal variations with depth.

These visualizations provide insight into vertical and seasonal thermal fluctuations and introduce the subsequent analyses focused on distribution patterns using violin plots and correlations with explanatory variables.

3.2. Exploration of Meteorological and Soil Temperature Distributions

To further explore the distribution and variability of the key environmental variables recorded at the Saaba station, violin plots were generated for air and soil temperatures as well as major meteorological parameters.

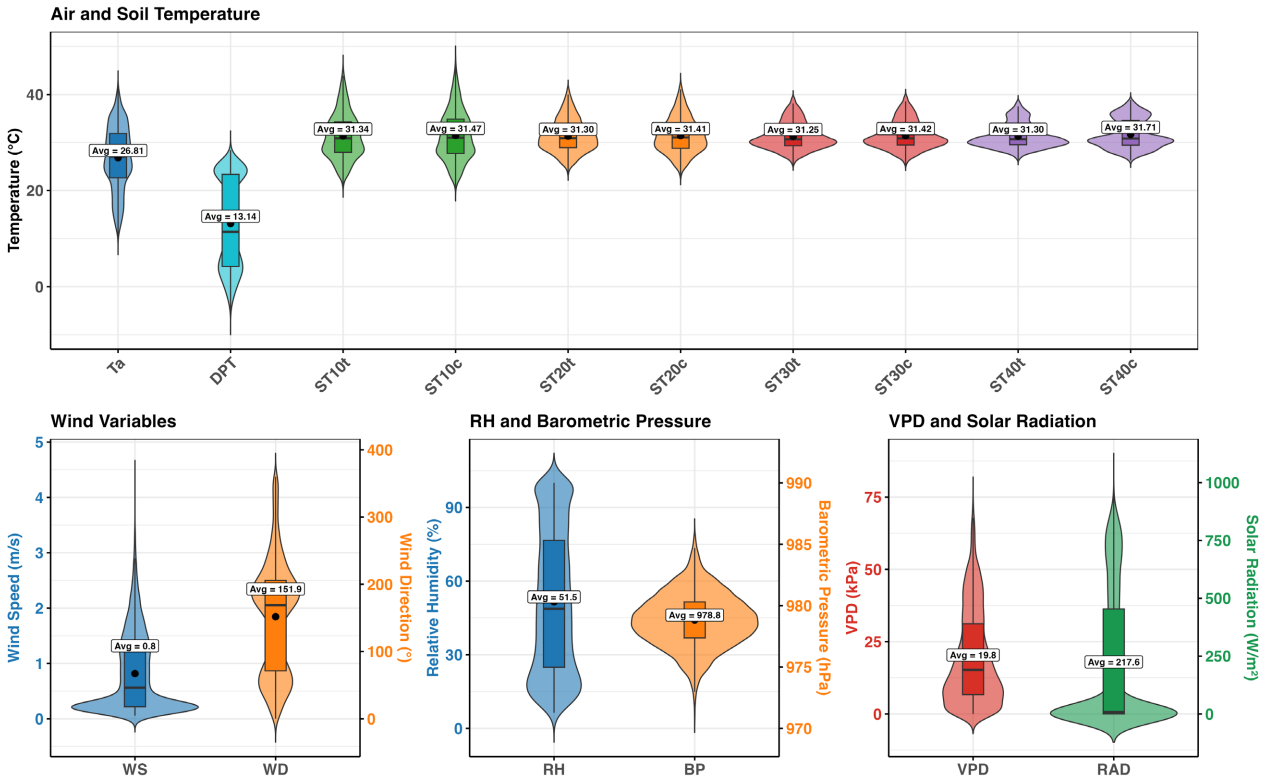

Figure 6. Distribution of meteorological and soil variables over the observation period. The top panel shows air temperature (Ta), dew point temperature (DPT), and soil temperature (ST) at 10, 20, 30, and 40 cm depths measured with TEROS21 sensors (t) and CS655 sensors (c). The bottom panels illustrate (i) wind speed (WS) and wind direction (WD), (ii) relative humidity (RH) and barometric pressure (BP), and (iii) vapor pressure deficit (VPD) and solar radiation (RAD). Mean values are displayed for each distribution.

Figure 6 provides visual summaries of the central tendencies and variability of

the environmental variables over the observation period. The top panel shows the distributions of air temperature (Ta), dew point temperature (DPT), and soil temperatures (ST) at 10, 20, 30, and 40 cm depths, measured with TEROS21 (t) and CS655 (c) sensors. Air and soil temperatures exhibit more stable distributions. Air temperature (Ta) averages 26.8˚C, while dew point temperature (DPT) is significantly lower at 13.3˚C, indicating a marked vapor pressure gap. Soil temperatures are more stable and consistently average around 31˚C - 32˚C across depths and sensor types, with slight increases in the TEROS21 sensors at 10 and 30 cm. The close agreement between TEROS21 and CS655 measurements supports their interchangeability, although TEROS21 data will be preferred for further modeling.

The bottom panels display the distributions of key meteorological drivers. Wind speed (WS) shows a narrow distribution centered around 0.8 m/s, suggesting weak air movement, whereas wind direction (WD) is broadly distributed with a central tendency around 151.9˚, reflecting variable wind patterns. Relative humidity (RH) exhibits a bimodal distribution with substantial variability (mean ≈ 51.5%), while barometric pressure (BP) follows a more symmetrical distribution centered near 978.8 hPa. Vapor pressure deficit (VPD) spans a wide range with a positively skewed distribution, indicating frequent episodes of atmospheric dryness. Solar radiation (RAD) also shows a right-skewed pattern, with a mean of approximately 217.6 W/m^2, reflecting the expected diurnal variability in tropical settings.

3.3. Correlation Analyses between Meteorological Predictors and Soil Temperatures

To complement the distribution analysis, a correlation matrix was constructed to explore the relationships between meteorological variables and soil temperatures at multiple depths using the two types of sensors (TEROS21 and CS655). **Figure 7** summarizes the pairwise Pearson correlation coefficients, offering insights into the interactions driving subsurface thermal dynamics.

The correlation matrix reveals strong positive associations between air temperature (Ta) and soil temperatures across all depths, especially at shallower levels (e.g., ST10t: $r = 0.83$), highlighting the dominant influence of surface conditions. Other meteorological variables such as vapor pressure deficit (VPD) and solar radiation (RAD) also show moderate to weak positive correlations, suggesting their role in surface energy balance.

Barometric pressure (BP), in contrast, exhibits strong negative correlations with soil temperature, indicating an inverse relationship possibly linked to broader atmospheric processes rather than local heating. Dew point temperature (DPT) gains relevance with increasing depth, surpassing Ta in explanatory strength beyond 30 cm.

The soil temperature variables demonstrate remarkably high internal consistency, both within and between sensor types (e.g., $r > 0.98$ between ST30c and ST40c, and $r = 0.99$ between ST10t and ST10c), confirming the robustness of measurements and justifying the use of TEROS21 data for subsequent analyses.

These results underline the stability of vertical soil temperature profiles and support the representativeness of selected metrics.

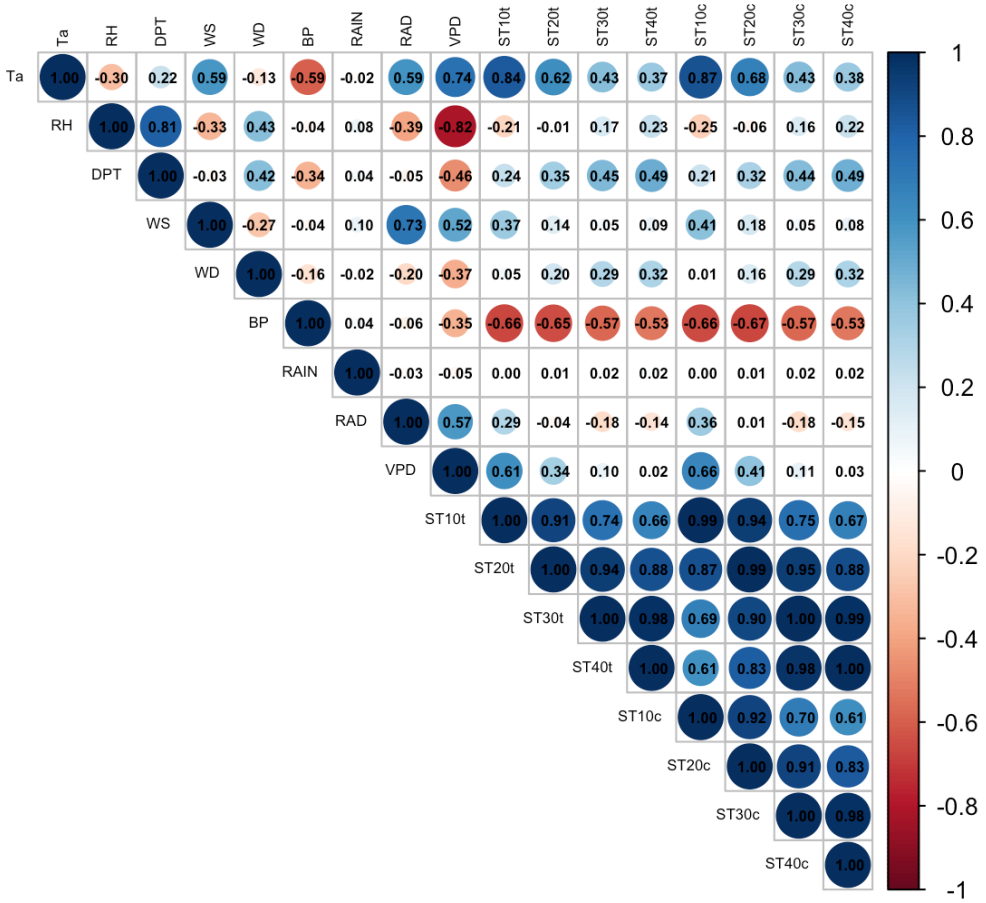

Figure 7. Pearson correlation matrix between meteorological variables and soil temperatures measured by TEROS21 (ST10t to ST40t) and CS655 (ST10c to ST40c) sensors.

To further dissect these relationships, **Figure 8** presents the Pearson correlation coefficients between soil temperature at four depths (ST10t to ST40t, measured by TEROS21) and key meteorological variables. This bar plot format allows for a clearer view of how the strength and direction of associations evolve with soil depth. As observed, air temperature (Ta) and barometric pressure (BP) remain the most influential factors across all depths. Ta shows a strong positive correlation that progressively weakens with depth (from r = 0.83 at 10 cm to r = 0.58 at 40 cm), highlighting the decreasing influence of surface heating. Conversely, BP maintains a consistently strong negative correlation ($r < -0.65$), suggesting broader atmospheric control mechanisms. A plausible physical explanation is that high-pressure systems are typically associated with clearer skies, which enhance nocturnal radiative cooling of the land surface, and with the advection of cooler, drier air masses. This combination of factors leads to a more pronounced drop in soil temperature, hence the observed strong negative correlation. Dew point temperature (DPT) gains importance at deeper layers, surpassing Ta from 30 cm down-

ward, thus reflecting latent heat and moisture effects in subsurface processes. Other meteorological variables, including vapor pressure deficit (VPD), relative humidity (RH), and wind parameters, exhibit weak to moderate correlations, with no clear patterns across depths. These findings emphasize the progressive dampening of meteorological influence with increasing soil depth, and reinforce the selection of key predictors in subsequent modeling efforts.

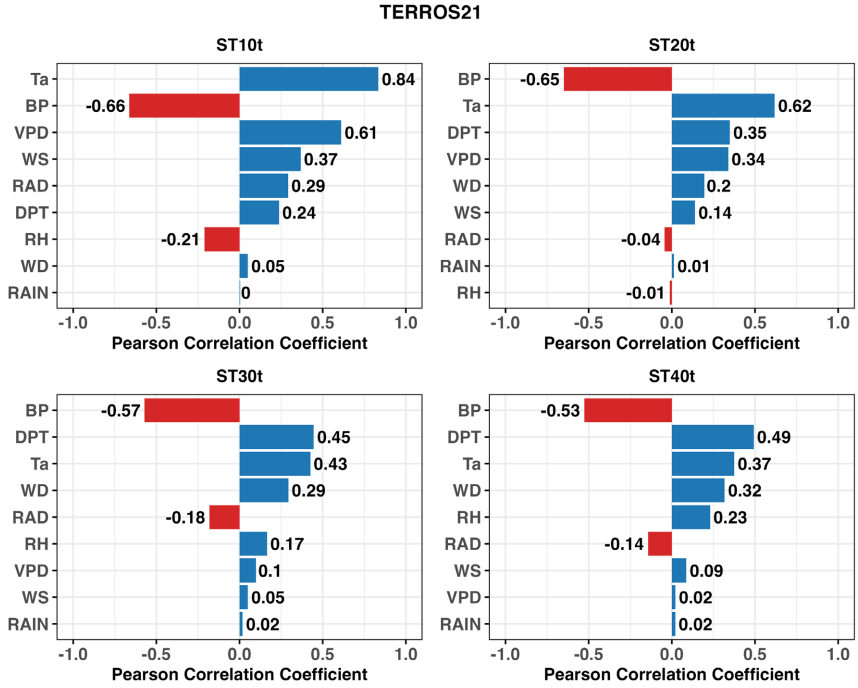

Figure 8. Pearson correlation coefficients between soil temperature at 10, 20, 30, and 40 cm (measured by TEROS21 sensors) and selected meteorological variables.

3.4. Comparison of Observed and MLR-Predicted Distributions of Hourly Soil Temperature Across Depths

The results obtained for each regression scenario are presented using violin plots, which allow simultaneous comparison of both the distribution shape and the central tendency (mean and median) of the measured and estimated soil temperatures. This visual representation helps assess the accuracy and consistency of the multivariate linear regression models across different soil depths. Then, **Figure 9** compare the observed and predicted soil temperatures across all the four depths using the six MLR scenarios.

The distribution shapes offer more insight. At shallow depths (10 and 20 cm), adding variables slightly improves the resemblance between predicted and observed distributions, reducing the asymmetry and narrowing the spread. At deeper layers (30 and 40 cm), the gain is less evident, and the predicted distributions remain narrower than the measured ones, suggesting that the models capture the central tendency well but may underestimate the full variability. Overall, increasing the number of predictors improves the representation of the distribution's

shape, particularly at surface levels, but has diminishing returns at depth due to the smoothing effect of the soil profile.

3.5. Performance Evaluation of Tested MLR Models Using Scatter Plots, Regression Metrics, and Fitted Equations

To evaluate the predictive performance of the multivariate linear regression (MLR) scenarios, scatterplots of predicted versus observed soil temperatures were generated for each scenario and soil depth. Key performance metrics, including the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), and model equations, were used to assess model accuracy and goodness of fit.

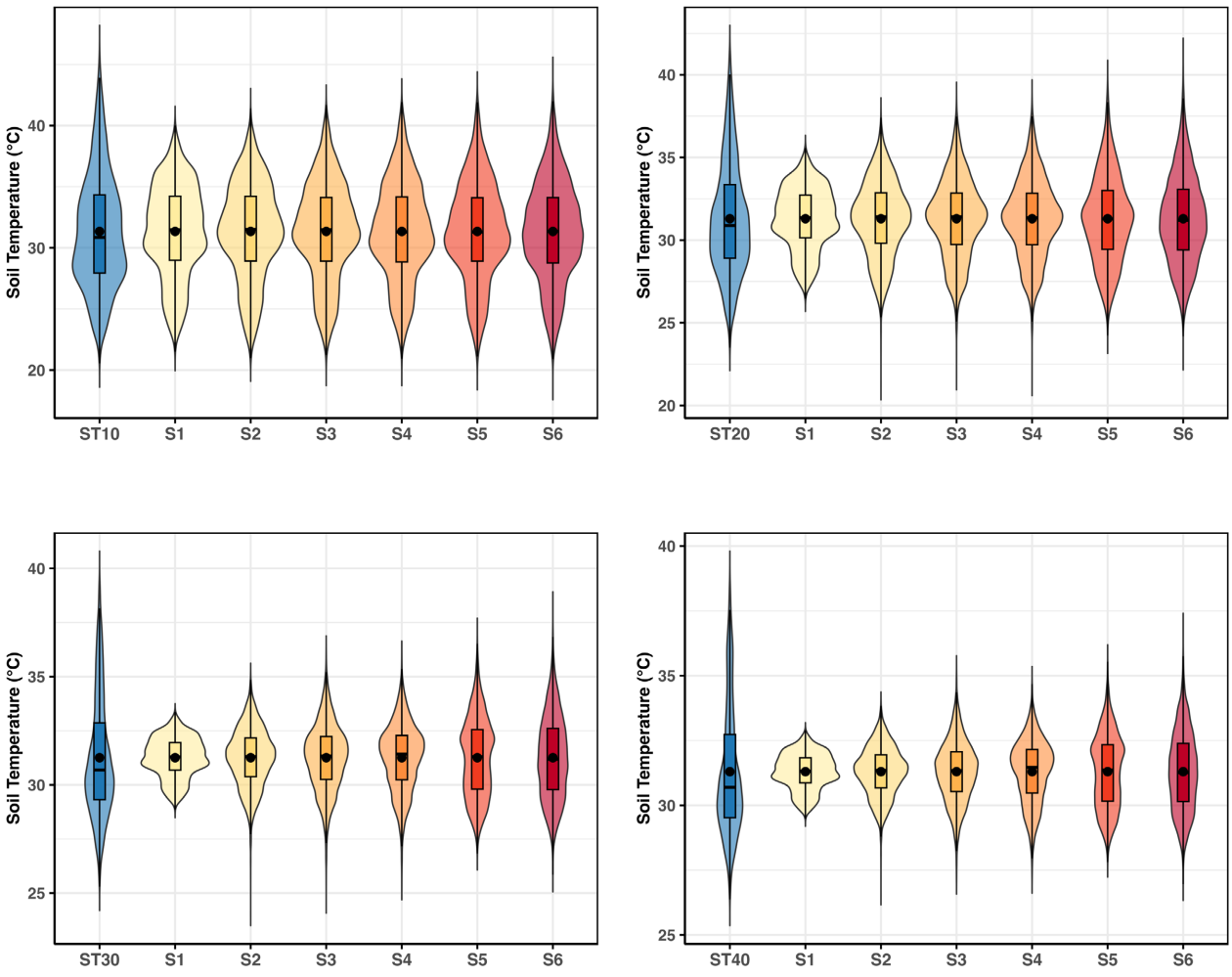

Figure 9. Distribution comparison of observed vs. predicted hourly soil temperature using MLR models under six scenarios, at four soil depths: (a) 10 cm; (b) 20 cm (c) 30 cm and (d) 40 cm.

Figures 10-13 summarize the results across the six scenarios for estimating hourly soil temperature at four depths (10, 20, 30, and 40 cm), by comparing observed and predicted values in each case by scatterplot.

For 10 cm-depth (ST10), performance steadily improves with model complex-

ity. The R^2 increases from 0.645 (S1) to 0.745 (S6), while RMSE decreases from 2.77˚C to 2.34˚C. This indicates that adding predictors like wind speed, pressure, and solar radiation enhances the fit without overfitting. The S6 scenario achieves the best compromise between accuracy and model complexity. At 20 cm-depth (ST20), the models also improve as predictors are added. R^2 rises from 0.301 (S1) to 0.627 (S6), and RMSE drops from 2.81˚C to 2.05˚C. Notably, adding radiation and wind direction from scenario S5 onwards significantly boosts performance, suggesting greater influence of these factors at this depth.

For 30 cm-depth (ST30), predictive power remains moderate. R^2 starts low at 0.109 (S1) but reaches 0.475 (S6). RMSE decreases from 2.60˚C to 2.01˚C. Although the accuracy improves with more variables, the dispersion remains more compact, indicating that deeper soil temperature dynamics are harder to capture using surface meteorological inputs alone. At 40 cm-depth (ST40), similar trends are observed. The basic model (S1) yields a low R^2 of 0.077 and RMSE of 2.38˚C, while the most complex scenario (S6) improves to R^2 = 0.361 and RMSE = 1.98˚C. However, gains from S4 to S6 are marginal, highlighting the diminishing influence of surface variables with increasing depth.

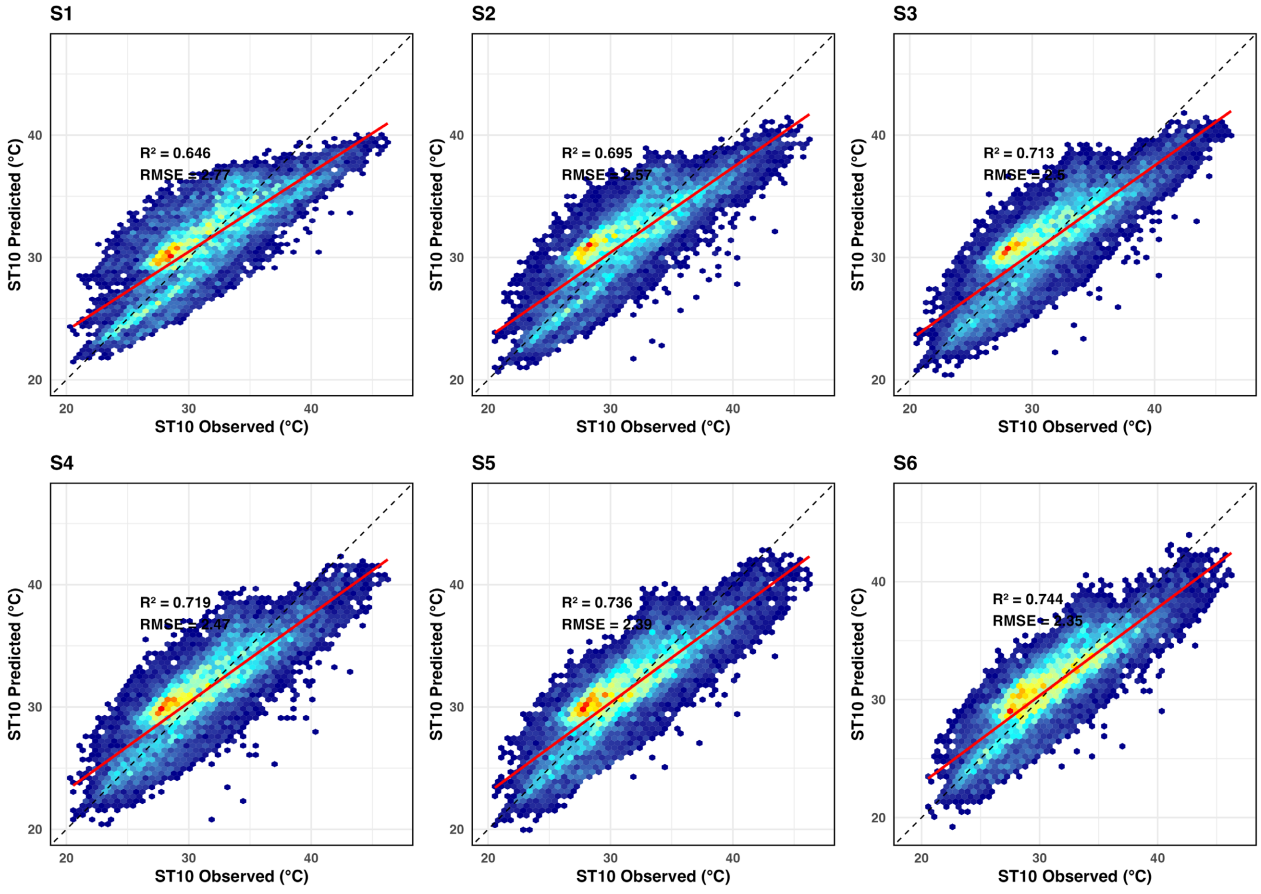

Figure 10. Observed vs. predicted hourly soil temperature using MLR models under six scenarios at 10 cm soil depth.

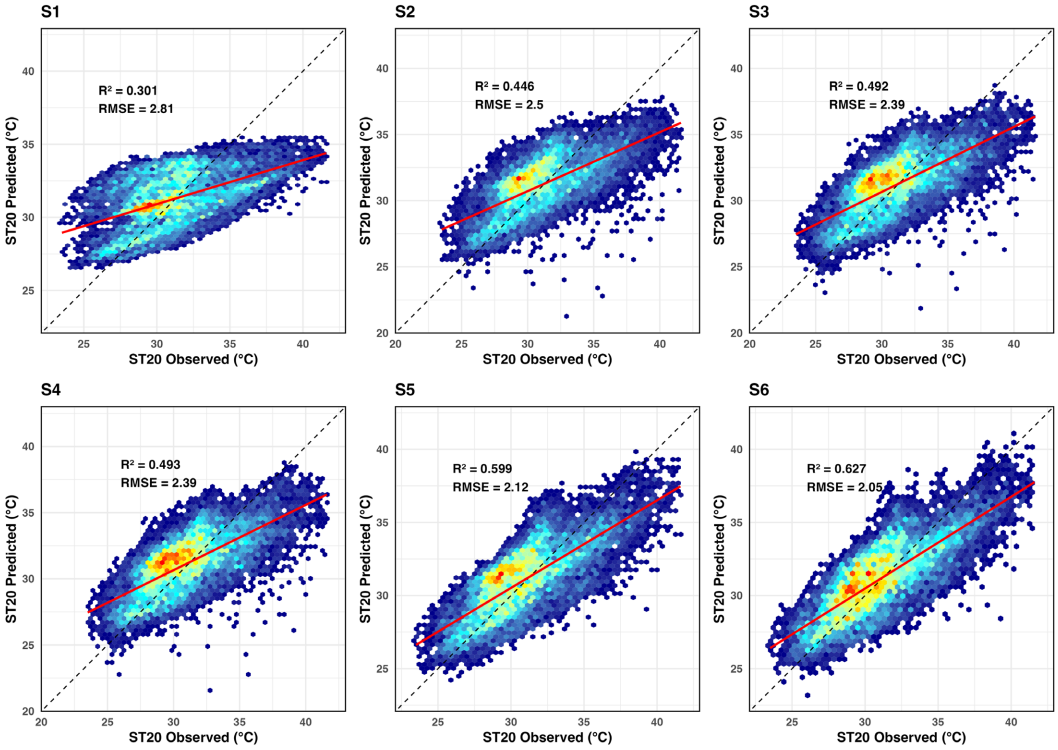

Figure 11. Observed vs. predicted hourly soil temperature using MLR models under six scenarios at 20 cm soil depth.

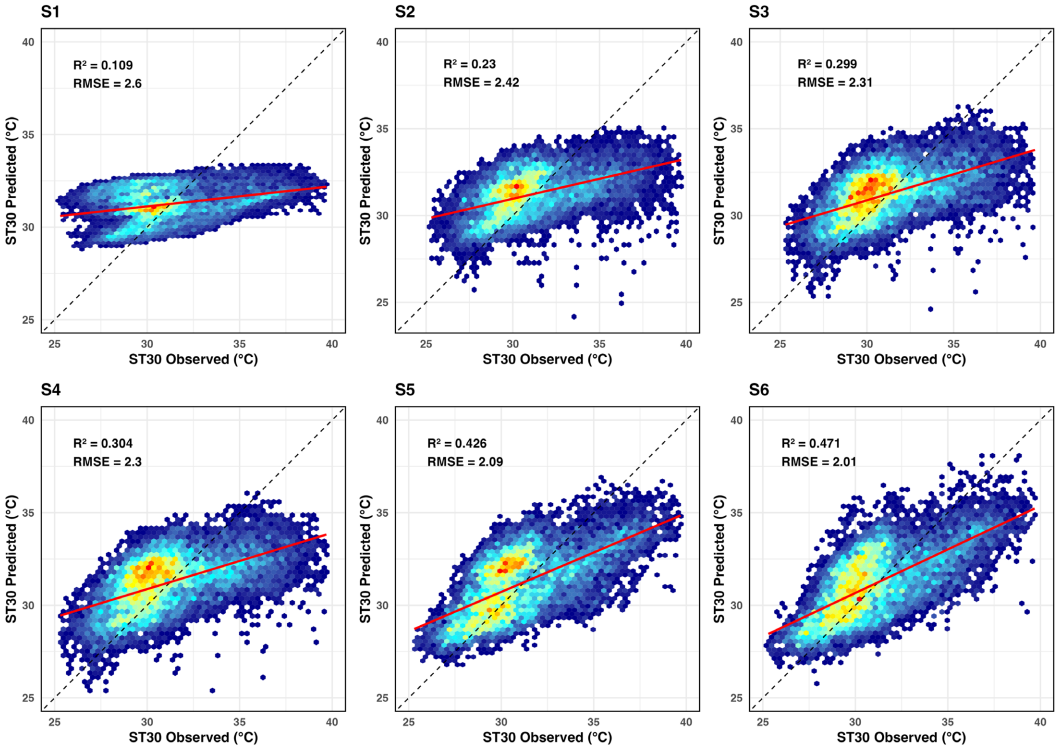

Figure 12. Observed vs. predicted hourly soil temperature using MLR models under six scenarios at 30 cm soil depth.

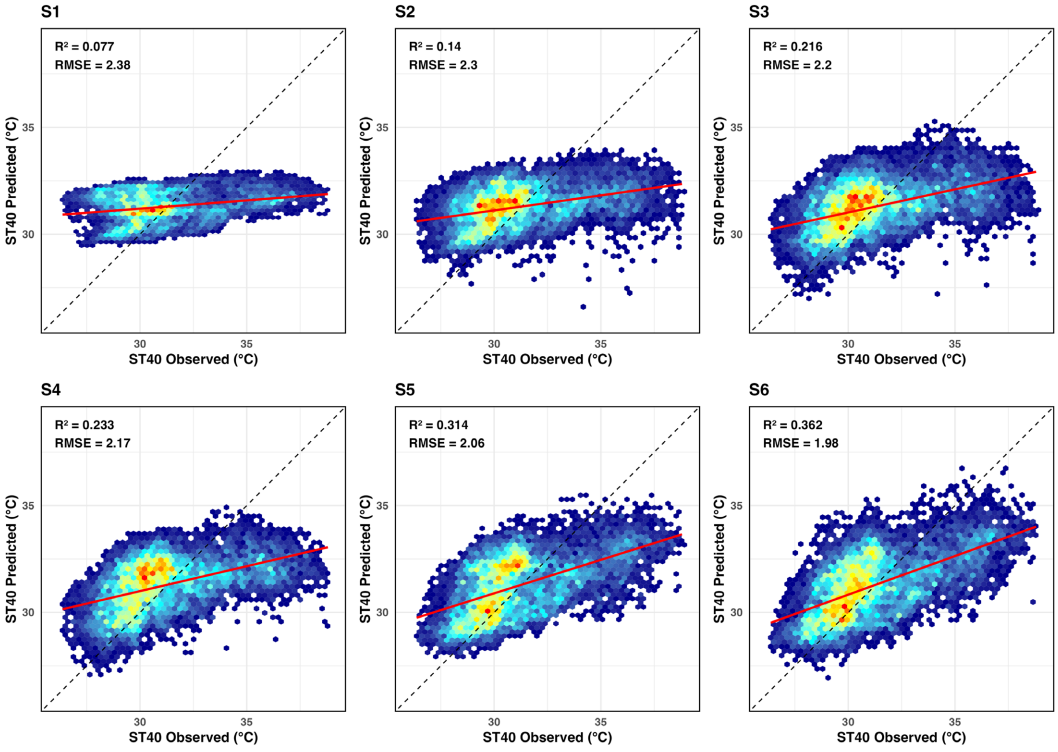

Figure 13. Observed vs. predicted hourly soil temperature using MLR models under six scenarios at 40 cm soil depth.

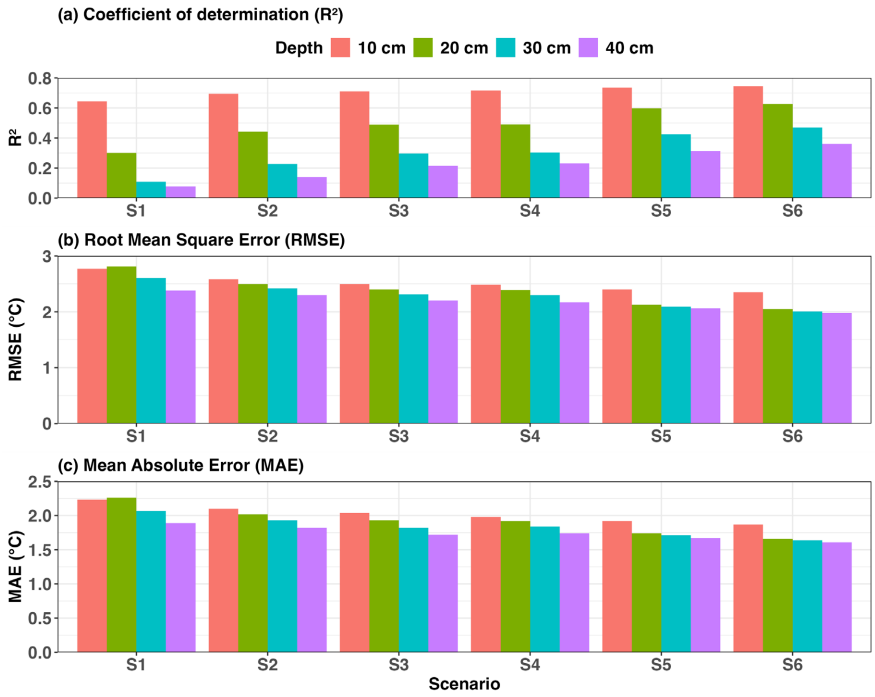

Figure 14. Performance metrics (R^2, RMSE, MAE) of multivariate linear regression models across six scenarios and four soil depths (10, 20, 30, and 40 cm).

Figure 14 presents the performance metrics, coefficient of determination (R^2),

root mean square error (RMSE), and mean absolute error (MAE), for the MLR models across the six scenarios and the four soil depths. These results support the scatterplot analysis, highlighting that the inclusion of additional meteorological predictors improves model accuracy, especially for the surface layers (10 - 20 cm), where R^2 exceeds 0.6 in the most complete scenarios while maintaining acceptable error levels (RMSE < 2˚C, MAE < 1.7˚C).

We summarize in Tables 4-7 the performance metrics of the MLR models across six scenarios for estimating hourly soil temperatures at four different depths (10, 20, 30, and 40 cm). For each scenario, we add also Akaike information criterion (AIC), Bayesian information criterion (BIC), and the fitted regression equation.

Table 4. Performance metrics and model equations for MLR across six scenarios for predicting hourly soil temperature at 10 cm depth.

Scenario	R^2	RMSE (˚C)	MAE (˚C)	AIC	BIC	Equation
S1	0.645	2.77	2.23	53,543	53,565	$y = 16.19 + 0.57 \times Ta$
S2	0.694	2.58	2.10	51,940	51,969	$y = 14.57 + 0.68 \times Ta - 1.81 \times WS$
S3	0.711	2.50	2.04	51,284	51,320	$y = 381.88 + 0.58 \times Ta - 0.37 \times BP - 1.33 \times WS$
S4	0.717	2.48	1.98	51,058	51,102	$y = 412.88 + 0.56 \times Ta - 0.40 \times BP - 1.47 \times WS - 0.01 \times RH$
S5	0.735	2.40	1.92	50,362	50,414	$y = 316.98 + 0.63 \times Ta - 0.31 \times BP + 0 \times RAD - 0.78 \times WS - 0.01 \times RH$
S6	0.745	2.35	1.87	49,951	50,017	$y = 270.65 + 0.64 \times Ta - 0.02 \times RH - 0.70 \times WS + 0.01 \times WD - 0.26 \times BP + 0.19 \times RAIN + 0 \times RAD$

Table 5. Performance metrics and model equations for MLR across six scenarios for predicting hourly soil temperature at 20 cm depth.

Scenario	R^2	RMSE (˚C)	MAE (˚C)	AIC	BIC	Equation
S1	0.301	2.81	2.26	53,798	53,820	$y = 23.84 + 0.28 \times Ta$
S2	0.443	2.50	2.02	51,307	51,337	$y = 21.84 + 0.42 \times Ta - 2.24 \times WS$
S3	0.490	2.40	1.93	50,343	50,380	$y = 451.19 + 0.30 \times Ta - 0.44 \times BP - 1.67 \times WS$
S4	0.491	2.39	1.92	50,326	50,370	$y = 459.79 + 0.30 \times Ta - 0.44 \times BP - 1.71 \times WS + 0 \times RH$
S5	0.598	2.13	1.74	47,719	47,770	$y = 287.89 + 0.42 \times Ta - 0.27 \times BP - 0.01 \times RAD - 0.48 \times WS - 0.01 \times RH$
S6	0.627	2.05	1.66	46,910	46,976	$y = 228.37 + 0.43 \times Ta - 0.01 \times RH - 0.33 \times WS + 0.01 \times WD - 0.21 \times BP + 0.10 \times RAIN - 0.01 \times RAD$

Table 6. Performance metrics and model equations for MLR across six scenarios for predicting hourly soil temperature at 30 cm depth.

Scenario	R^2	RMSE	MAE	AIC	BIC	Equation
S1	0.109	2.6	2.07	52,144	52,166	$y = 27.56 + 0.14 \times Ta$
S2	0.228	2.42	1.93	50,569	50,598	$y = 26.06 + 0.25 \times Ta - 1.69 \times WS$
S3	0.298	2.31	1.82	49,538	49,575	$y = 454.57 + 0.13 \times Ta - 0.44 \times BP - 1.12 \times WS$
S4	0.303	2.3	1.84	49,462	49,506	$y = 437.7 + 0.14 \times Ta - 0.42 \times BP - 1.04 \times WS + 0.01 \times RH$
S5	0.426	2.09	1.71	47,335	47,386	$y = 286.84 + 0.24 \times Ta - 0.27 \times BP - 0.01 \times RAD + 0.04 \times WS + 0.01 \times RH$
S6	0.47	2.01	1.64	46,445	46,503	$y = 225.2 + 0.26 \times Ta + 0 \times RH + 0.23 \times WS + 0.01 \times WD - 0.21 \times BP - 0.01 \times RAD$

Table 7. Performance metrics and model equations for MLR across six scenarios for predicting hourly soil temperature at 40 cm depth.

Scenario	R^2	RMSE	MAE	AIC	BIC	Equation
S1	0.077	2.38	1.89	50,227	50,249	$y = 28.5 + 0.1 \times Ta$
S2	0.139	2.3	1.82	49,466	49,495	$y = 27.53 + 0.17 \times Ta - 1.09 \times WS$
S3	0.215	2.2	1.72	48,451	48,487	$y = 432.2 + 0.06 \times Ta - 0.41 \times BP - 0.56 \times WS$
S4	0.232	2.17	1.74	48,212	48,256	$y = 404.2 + 0.08 \times Ta - 0.38 \times BP - 0.43 \times WS + 0.01 \times RH$
S5	0.313	2.06	1.67	46,995	47,046	$y = 294.15 + 0.15 \times Ta - 0.27 \times BP + 0 \times RAD + 0.36 \times WS + 0.01 \times RH$
S6	0.361	1.98	1.61	46,201	46,259	$y = 236.68 + 0.17 \times Ta + 0.01 \times RH + 0.54 \times WS + 0.01 \times WD - 0.22 \times BP + 0 \times RAD$

Across all depths, the AIC and BIC values decrease gradually from Scenario S1 to S6, indicating a better model fit as more explanatory variables are added. However, the variations in AIC and BIC between successive scenarios, especially between S4, S5, and S6, are relatively modest. This implies that while model accuracy slightly improves, the added complexity may not always be justified, particularly for operational use in data-sparse regions. Intermediate models (such as S3 or S4) may offer an optimal trade-off between performance and parsimony.

Overall, and to summarize the MLR models in estimating hourly soil temperatures across depths, we can say that predictive accuracy is strongly influenced by the number and relevance of meteorological predictors. At shallow depths (10 cm and 20 cm), the models show substantial improvement as more variables are included, with R^2 values increasing from ~0.65 (baseline with Ta alone) to ~0.75 (for ST10) and ~0.63 (for ST20) in the most complete scenarios. In contrast, deeper layers (30 cm and 40 cm) remain less predictable with R^2 values not exceeding 0.48 and 0.36, respectively, regardless of model complexity. This confirms that subsurface temperatures are more insulated from short-term atmospheric variability.

Our results are consistent with those of [9] who also found that at soil depths greater than 30 cm, both models failed to accurately estimate soil temperature using daily data. Either method, however, could satisfactorily predict soil temperature at depths between 0 and 20 cm. In their study, model performances were notably higher than in ours, with R^2 values of 0.96, 0.94, and 0.88 at 5, 10, and 20 cm depths, respectively, using daily data and 13 input variables including minima and maxima. By contrast, our hourly model yielded R^2 values of 0.745 and 0.627 at 10 and 20 cm, respectively, using only seven meteorological parameters with multicollinearity excluded. This difference in performance may be attributed to the higher noise level in hourly data compared to daily data.

So, for practical estimation with regression methods of soil skin temperature (ST10) at hourly scale with a good approximation (targeting $R^2 \approx 0.75$), a minimum set of four variables is sufficient: air temperature (Ta), barometric pressure (BP), wind speed (WS) and relative humidity (RH). These variables, commonly measured in standard weather stations, capture the essential surface energy exchanges that drive soil temperature variations. Adding further variables like solar

radiation (RAD) and wind direction (WD) slightly improves model precision.

3.6. Temporal Consistency and Variability Assessment of Predictions in Soil Temperature Series

While standard performance metrics (e.g., R^2, RMSE, MAE), scatter plots, and distribution-based visualizations such as violin plots offer valuable insights into model accuracy and global behavior, they often fail to reveal local temporal inconsistencies in time series data. In the case of soil temperature estimation, which involves spatio-temporal sequences governed by cyclic thermal processes, it is crucial to verify whether the predicted values follow the natural variability and dynamics of the observed series, including the correct reproduction of local extrema (minima and maxima) across different depths. This aspect is particularly important from a physical standpoint, as soil temperature responds to radiative forcing, surface energy balance, and subsurface heat propagation, which collectively induce characteristic oscillations in the signal. A model that performs well globally but fails to replicate these key temporal patterns may lead to misleading conclusions in process-based applications. Therefore, we systematically compare the temporal structure of the predicted and observed series with a specific focus on the preservation of diurnal and sub-daily variability. This temporal fidelity analysis complements the global evaluation metrics and ensures that models not only approximate values accurately but also respect the physical dynamics underlying soil temperature variations.

So, we present in **Figure 15** the temporal variability of measured and predicted soil temperatures under different MLR scenarios (S1, S4, S6), along with their 24-hour rolling means and 95% confidence intervals. Scenario S1 was selected because it relies solely on air temperature (Ta), a variable that is easily and widely measured. Scenario S4 incorporates additional commonly available meteorological variables: wind speed (WS), barometric pressure (BP), and relative humidity (RH), while scenario S6 includes all meteorological predictors, with variable selection adjusted to account for multicollinearity. The analysis focuses exclusively on the 10 cm soil depth, where the MLR predictions yielded the highest statistical performance ($R^2 > 0.60$ across all scenarios), unlike deeper layers where the predictive accuracy of MLR models was markedly lower.

The MLR-based predictions demonstrate good agreement with observations during the dry season (January-February-March, JFM), but tend to overestimate soil temperature during the rainy season (June-September) and underestimate it toward the end of the wet season (October-November-December, OND). These deviations suggest that the models lack sensitivity to soil moisture dynamics, a key variable influencing thermal behavior through evaporative cooling and heat capacity modulation. Specifically, higher moisture content increases the soil's thermal conductivity, promoting heat transfer, while simultaneously raising its heat capacity, requiring more energy to change its temperature. The absence of this variable likely explains the model's error in simulating the amplitude and phase

of diurnal temperature cycles, particularly during wet periods. Incorporating soil water content as an additional predictor could therefore enhance model performance, especially under conditions of high atmospheric humidity and rainfall variability, as highlighted in Taheri *et al.* [10].

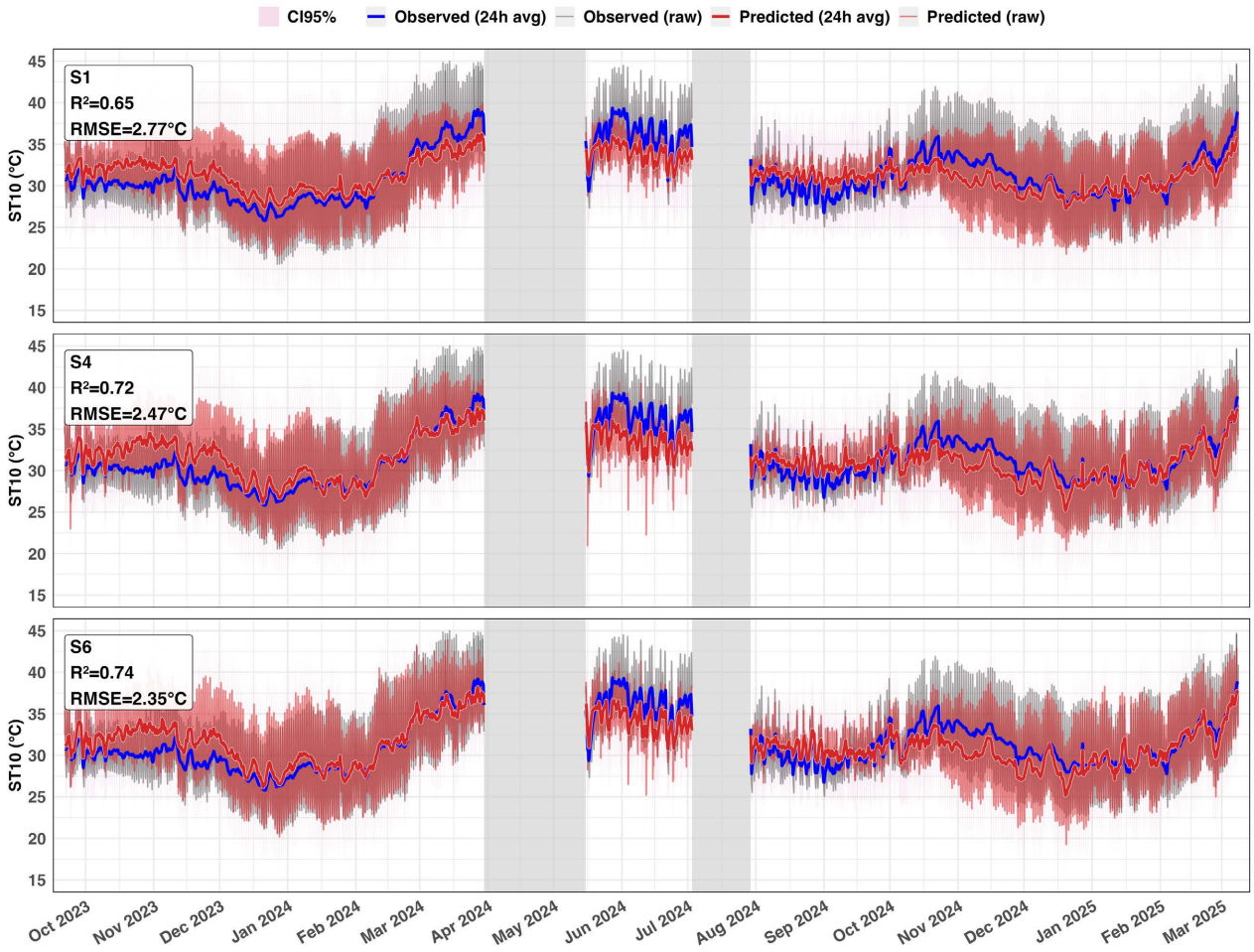

Figure 15. Temporal variability of observed and predicted soil temperature at 10 cm depth under three MLR scenarios (S1, S4, S6), with 24-hour rolling averages and 95% confidence intervals.

To further investigate the ability of the models to reproduce fine-scale temporal fluctuations, we conducted a focused comparison over a short time window (15-18 June 2024), which covers a full diurnal cycle over three days. Figure 16 presents the observed and predicted soil temperatures at 10 cm depth under three MLR scenarios (S1, S4, S6), during 15-18 June 2024, with emphasis on the timing and amplitude of daily extrema.

The short-term validation highlights clear performance differences between the modeling approaches. The predictions systematically underestimate the observed temperatures, especially during the rainy season (June), and fail to capture the full diurnal cycle. This further confirms the limited ability of MLR models to generalize under high humidity and soil moisture conditions typical of the wet season.

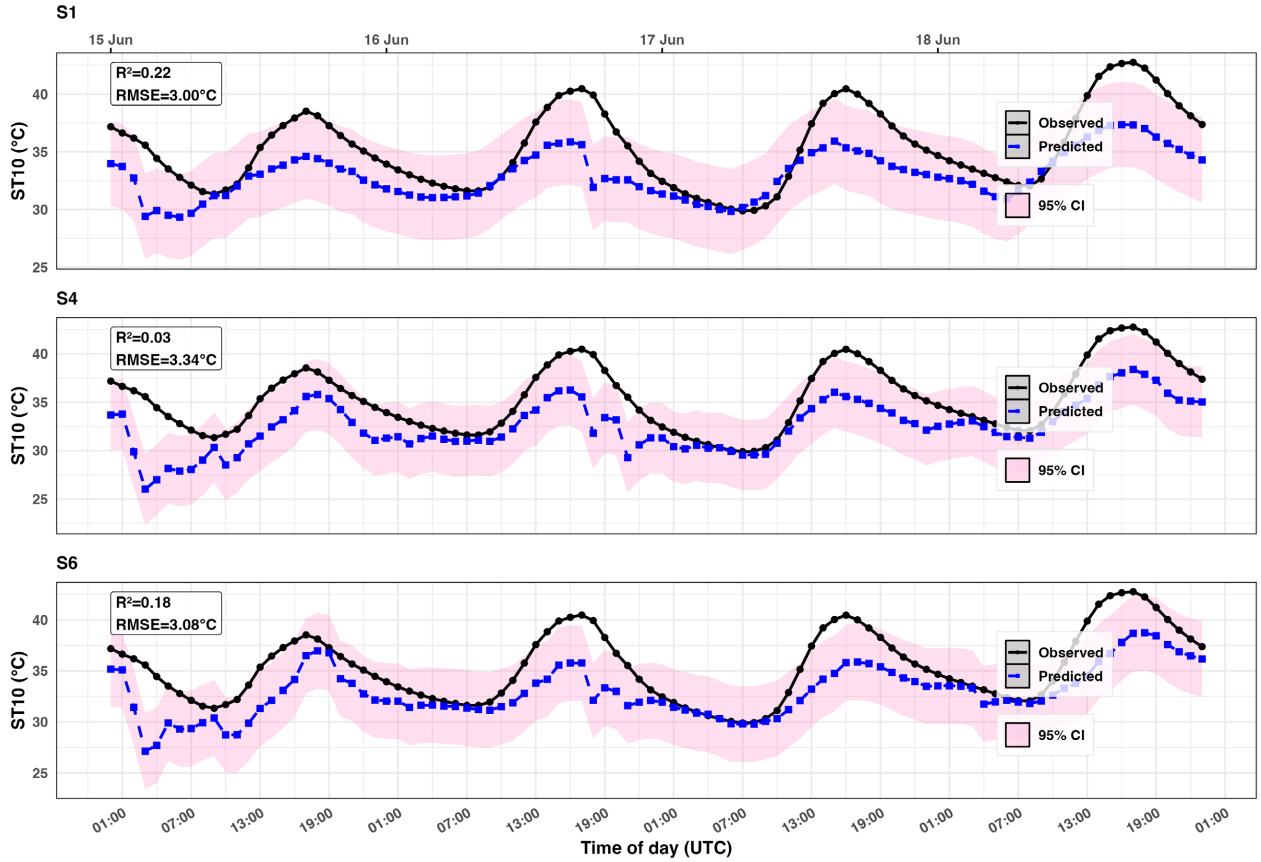

Figure 16. Short-term validation of multiple linear regression (MLR) predictions at 10 cm depth for three scenarios (S1, S4, S6) from 15 to 18 June 2024.

4. Conclusions

This study presents a comprehensive comparative analysis of statistical MLR approaches to estimate hourly soil temperature at different depths of 10, 20, 30, and 40 cm in a semi-arid region of Burkina Faso using *in situ* meteorological data. Conducted at the Tanghin (Saaba, Kadiogo) experimental site, the study highlights the challenges and opportunities associated with modeling subsurface thermal dynamics in low-instrumentation contexts and under high temporal resolution.

Analysis of the variability further revealed a clear attenuation of temperature fluctuations with depth, confirming the buffering effect and thermal inertia of subsurface layers. The observed phase lag between soil and air temperatures, which increases with depth, particularly beyond 20 cm, emphasizes the delayed heat propagation within the soil profile. A more detailed investigation is needed to understand how this phase shift varies across the rainy and dry seasons, particularly in the context of arid tropical environments.

MLR models demonstrate acceptable performance at shallow depths when fed with a minimal set of meteorological variables (Ta, BP, WS, RH), achieving R^2 values up to 0.75 at 10 cm. However, their predictive ability sharply declines in

deeper layers, where the thermal signal becomes increasingly decoupled from surface atmospheric forcing. In addition, the models struggled to accurately reproduce the full diurnal cycles, as evidenced by a consistent underestimation of the diurnal temperature range. This is manifested through an underestimation of nighttime cooling and, to a lesser extent, daytime warming. Furthermore, the models systematically underestimated temperatures during the rainy season. These combined biases reflect a limited sensitivity to soil moisture dynamics, which critically modulate thermal behavior through evaporative cooling and heat capacity effects.

Despite its scope, this study has limitations. Data from a single site may restrict transferability to other contexts with different soil textures, vegetation cover, or climatic regimes. Furthermore, the absence of continuous soil moisture measurements limited the ability to account for its critical role in subsurface thermal regulation.

Future research should focus on extending these models across multiple agroecological zones, integrating soil moisture and thermal properties into prediction frameworks, and exploring hybrid physics-informed machine learning methods to enhance robustness and interpretability. Coupling soil temperature modeling with crop growth, irrigation scheduling, and land surface models would provide valuable tools for precision agriculture and climate resilience in semi-arid regions.

Acknowledgements

The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Mampitiya, L., Rozumbetov, K., Rathnayake, N., Erkudov, V., Esimbetov, A., Arachchi, S., *et al.* (2024) Artificial Intelligence to Predict Soil Temperatures by Development of Novel Model. *Scientific Reports*, **14**, Article No. 9889. https://doi.org/10.1038/s41598-024-60549-x

[2] Elsayed, S., Gupta, M., Chaudhary, G., Taneja, S., Gaur, H., Gad, M., *et al.* (2023) Interpretation the Influence of Hydrometeorological Variables on Soil Temperature Prediction Using the Potential of Deep Learning Model. *Knowledge-Based Engineering and Sciences*, **4**, 55-77. https://doi.org/10.51526/kbes.2023.4.1.55-77

[3] Sattari, M.T., Avram, A., Apaydin, H. and Matei, O. (2020) Soil Temperature Estimation with Meteorological Parameters by Using Tree-Based Hybrid Data Mining Models. *Mathematics*, **8**, Article 1407. https://doi.org/10.3390/math8091407

[4] Hsieh, C., Chiu, C., Huang, I. and Visessri, S. (2023) Estimating Ground Heat Flux from Net Radiation. *Atmosphere*, **14**, Article 1778. https://doi.org/10.3390/atmos14121778

[5] Holmes, T.R.H., Owe, M., De Jeu, R.A.M. and Kooi, H. (2008) Estimating the Soil Temperature Profile from a Single Depth Observation: A Simple Empirical Heatflow Solution. *Water Resources Research*, **44**. https://doi.org/10.1029/2007wr005994

[6] Edmondson, J.L., Stott, I., Davies, Z.G., Gaston, K.J. and Leake, J.R. (2016) Soil Surface Temperatures Reveal Moderation of the Urban Heat Island Effect by Trees and Shrubs. *Scientific Reports*, **6**, Article No. 33708. https://doi.org/10.1038/srep33708

[7] Shi, B., Tang, C., Gao, L., Liu, C. and Wang, B. (2012) Observation and Analysis of the Urban Heat Island Effect on Soil in Nanjing, China. *Environmental Earth Sciences*, **67**, 215-229. https://doi.org/10.1007/s12665-011-1501-2

[8] Asadzadeh, F., Emami, S., Elbeltagi, A., Akiner, M.E., Rezaverdinejad, V., Taran, F., *et al.* (2025) Investigating the Impact of Meteorological Parameters on Daily Soil Temperature Changes Using Machine Learning Models. *Scientific Reports*, **15**, Article No. 19988. https://doi.org/10.1038/s41598-025-04605-0

[9] Progga, J.F., Khan, M.N.H. and Amin, M.G.M. (2023) Meteorological Parameters-Soil Temperature Relations in a Sub-Tropical Summer Grassland: Physically-Based and Data-Driven Modeling. *Ataturk University Journal of Agricultural Faculty*, **54**, 48-56. https://doi.org/10.5152/auaf.2023.23126

[10] Taheri, M., Schreiner, H.K., Mohammadian, A., Shirkhani, H., Payeur, P., Imanian, H., *et al.* (2023) A Review of Machine Learning Approaches to Soil Temperature Estimation. *Sustainability*, **15**, Article 7677. https://doi.org/10.3390/su15097677

[11] Ríos-Arriola, J., Gómez-Arias, E., Zavala-Guillén, I., Velázquez-Limón, N., Bojórquez-Morales, G. and López-Velázquez, J.E. (2023) Numerical Modeling of Soil Temperature Variation under an Extreme Desert Climate. *Geothermics*, **112**, Article ID: 102731. https://doi.org/10.1016/j.geothermics.2023.102731

[12] Tabari, H., Hosseinzadeh Talaee, P. and Willems, P. (2014) Short-Term Forecasting of Soil Temperature Using Artificial Neural Network. *Meteorological Applications*, **22**, 576-585. https://doi.org/10.1002/met.1489

[13] Smits, K.M., Sakaki, T., Limsuwat, A. and Illangasekare, T.H. (2010) Thermal Conductivity of Sands under Varying Moisture and Porosity in Drainage-Wetting Cycles. *Vadose Zone Journal*, **9**, 172-180. https://doi.org/10.2136/vzj2009.0095

[14] Kébré, M.B., Ouédraogo, F., Naon, B., Cherblanc, F. and Zougmoré, F. (2021) Analysis of Non-Equilibrium Phase Change in Transfers at Low Water Content by Considering the Film Flow. *American Journal of Environmental Sciences*, **17**, 1-13. https://doi.org/10.3844/ajessp.2021.1.13

[15] Novak, M.D. (2019) Validity of Assuming Equilibrium between Liquid Water and Vapor for Simulating Evaporation. *Water Resources Research*, **55**, 9858-9872. https://doi.org/10.1029/2019wr025113

[16] Šimůnek, J., Brunetti, G., Jacques, D., van Genuchten, M.T. and Šejna, M. (2024) Developments and Applications of the Hydrus Computer Software Packages since 2016. *Vadose Zone Journal*, **23**, e20310. https://doi.org/10.1002/vzj2.20310

[17] Zhang, Y., Li, X., Šimůnek, J., Chen, N., Hu, Q. and Shi, H. (2024) Evaluating the Effects of Different Irrigation Water Sources on Soil Temperature Using HYDRUS (2D/3D) and Considering the Coupled Movement of Water and Heat. *Soil and Tillage Research*, **244**, Article ID: 106259. https://doi.org/10.1016/j.still.2024.106259

[18] Cemek, B., Kültürel, Y., Cemek, E., Küçüktopçu, E. and Simsek, H. (2025) Modeling Soil Temperature with Fuzzy Logic and Supervised Learning Methods. *Applied Sciences*, **15**, Article 6319. https://doi.org/10.3390/app15116319

[19] Ouoba Nebie, B.A., Lawane, A., Hema, C. and Siroux, M. (2025) Energy Efficiency in the Building Sector in Burkina Faso: Literature Review, SWOT Analysis, and Recommendations. *Energies*, **18**, Article 2689. https://doi.org/10.3390/en18112689

[20] Daoud, J.I. (2017) Multicollinearity and Regression Analysis. *Journal of Physics: Con-*

ference Series, **949**, Article ID: 012009.
https://doi.org/10.1088/1742-6596/949/1/012009

[21] Citakoglu, H. (2016) Comparison of Artificial Intelligence Techniques for Prediction of Soil Temperatures in Turkey. *Theoretical and Applied Climatology*, **130**, 545-556.
https://doi.org/10.1007/s00704-016-1914-7