ATATÜRK ÜNİVERSİTESİ YAYINLARI
ATATURK UNIVERSITY PUBLICATIONS

Research Article *Araştırma Makalesi* DOI: 10.5152/AUAF.2023.23126

Meteorological Parameters–Soil Temperature Relations in a Sub-Tropical Summer Grassland: Physically-Based and Data-Driven Modeling

Subtropikal Bir Yaz Çayırında Meteorolojik Parametreler-Toprak Sıcaklığı İlişkileri: Fiziksel Tabanlı ve Veriye Dayalı Modelleme

Jannatul Ferdaous PROGGA[1]
Mohammad Nur Hossain KHAN[2]
M.G. Mostofa AMIN[1]

[1]Department of Irrigation and Water Management, Bangladesh Agricultural University, Mymensingh, Bangladesh
[2]Department of Electrical and Electronic Engineering, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh.

Received/Geliş Tarihi: 15.03.2023
Accepted/Kabul Tarihi: 14.04.2023
Publication Date/Yayın Tarihi: 18.05.2023

Corresponding Author/Sorumlu Yazar: Jannatul Ferdaous PROGGA
E-mail: progga189@gmail.com

Cite this article as: Progga, J. F., Khan, M. N. H., & Amin, M. M. (2023). Meteorological parameters–soil temperature relations in a sub-tropical summer grassland: physically-based and data-driven modeling. *Research in Agricultural Sciences*, 54(2), 48-56.

ABSTRACT

The knowledge of soil temperature dynamics at different depths is paramount for the agricultural industry because soil temperature impacts the physical, chemical, and biological processes in soil. A relationship between meteorological parameters and temperature at different depths in silt loam soil was assessed by using a physically based HYDRUS-1D model and a linear regression model. Soil temperature at 5, 10, 20, 30, and 50 cm soil layers, minimum and maximum air temperature, air pressure, relative humidity, dew point, rainfall, sunshine duration, wind speed, and evaporation data collected at a weather station were used. The correlation sensitivity for the input combinations was investigated. The quantitative evaluation based on mean absolute percentage error and R^2 showed that the predictions of both linear regression model and HYDRUS-1D models were satisfactory. The R^2 values at 5, 10, and 20 cm depths were 0.96, 0.94, and 0.88 for linear regression model, and 0.85, 0.86, and 0.78, for HYDRUS-1D model, respectively. Similarly, the mean absolute percentage error values for linear regression model were 0.81%, 0.87%, and 1.05%, whereas 3.44%, 2.87%, and 3.73% at 5, 10, and 20 cm depths for HYDRUS-1D model, respectively. Generally, the accuracy of the models diminished with increasing the soil depth. At >30 cm soil depth, both models failed to estimate soil temperature accurately. The R^2 and mean absolute percentage error values at 50 cm depth for linear regression model were 0.55% and 1.25% and 0.51% and 4.13% for HYDRUS-1D, respectively. The linear regression model performed better than the HYDRUS-1D model. Five independent variables (mean air temperature, maximum humidity, rainfall, wind speed, and evaporation) were found to significantly affect the summer-time soil temperature. Either of the methods can be used satisfactorily to predict soil temperature at 0–20 cm soil depth.

Keywords: Evaporation, humidity, HYDRUS-1D, linear regression model, wind speed

ÖZ

Toprak sıcaklığı topraktaki fiziksel, kimyasal ve biyolojik süreçleri etkilediğinden, farklı derinliklerdeki toprak sıcaklığı dinamikleri bilgisi tarım endüstrisi için çok önemlidir. Bu çalışmada siltli tın tekstür sınıfına ait toprakların farklı derinliklerinde meteorolojik parametreler ile sıcaklık arasındaki ilişkiler, fiziksel tabanlı HYDRUS-1D modeli ve bir doğrusal regresyon modeli (LRM) kullanılarak değerlendirilmiştir. Çalışma alanında 5, 10, 20, 30 ve 50 cm derinliğindeki toprak katmanlarının sıcaklık değerleri ile meteoroloji istasyonundan alınan en düşük ve en yüksek hava sıcaklığı, basınç, çiğ oluşum noktası, yağış, güneşlenme süresi, rüzgar hızı verileri kullanılmıştır. Girdi kombinasyonları için korelasyon hassasiyeti araştırılmıştır. Çalışma sonucunda ortalama mutlak yüzde hatası (OMYH) ve R2'ye dayalı kantitatif değerlendirmelerin hem LRM hem de HYDRUS-1D modellerinden elde edilen tahminlerin tatmin edici olduğunu göstermiştir. LRM modelinde 5, 10 ve 20 cm derinlik katmanlarındaki R2 değerlerinin sırasıyla 0,96, 0,94 ve 0,88 olduğu, HYDRUS-1D modelinde ise 0,85, 0,86 ve 0,78 olduğu tespit edilmiştir. Benzer şekilde OMYH değerleri 5, 10 ve 20 cm derinlik kademelerinde LRM için %0,81, %0,87 ve %1,05 iken, HYDRUS-1D modeli için %3,44, %2,87 ve %3,73 olarak hesaplanmıştır. Genel olarak modellerin doğruluğu toprak derinliğinin artmasıyla azalmış ve 30cm'den daha derin katmanlarda her iki modelin de toprak sıcaklığını doğru bir şekilde tahmin edemediği belirlenmiştir. Toprak derinliğinin 50 cm olduğu katmanlarda R2 ve OMYH değerleri LRM modeli için 0,55 ve %1,25, HYDRUS-1D modeli için ise 0,51 ve %4,13 olmuştur. Çalışma sonucunda ayrıca LRM modelinin HYDRUS-1D modelinden daha iyi performans gösterdiği, beş bağımsız değişkenin (ortalama hava sıcaklığı, maksimum nem, yağış, rüzgar hızı ve buharlaşma) yaz mevsimindeki toprak sıcaklığını önemli ölçüde etkilediği, her iki yönteminde 0-20 cm'lik toprak derinliğinde toprak sıcaklığını tahmin etmek için tatmin edici bir şekilde kullanılabileceği belirlenmiştir.

Anahtar Kelimeler: Buharlaşma, nem, HYDRUS-1D, lineer regresyon modeli, rüzgar hızı

Introduction

An essential variable of the land surface scheme that controls the energy and moisture exchange in the atmospheric continuum of soil and plants is soil temperature. Soil temperature plays a critical role in ecosystems, from deserts to forests (Jebamalar et al., 2012). Soil evaporation, soil aeration, soil microbial activity, and many other soil biological, chemical, and physical processes are controlled by soil temperature. Furthermore, plant growth, seed germination, and nutrient uptake by plants also depend on soil temperature (Amin et al., 2021; Yadav et al., 2020). Therefore, knowledge of ground surface and subsurface temperature at various depths is important for agricultural practices (Yilmaz et al., 2009) and for a better understanding of climate change impacts (Kourat et al., 2021; Wu et al., 2013). Studies have shown that soil temperature depends on various meteorological variables, such as air temperature, atmospheric pressure, relative humidity, wind speed, rainfall, solar radiation, and sunshine duration (Kisi et al., 2015). To determine soil temperature, solar radiation and air temperature are the main driving forces, but soil texture, moisture content, and the type of soil covering (plant canopy, crop residue, snow, etc.) also influence soil temperature variably (Amin et al., 2021; Karnieli et al., 2010).

There are numerous laboratory and field procedures for estimating soil hydraulic and thermal characteristics. The necessity of capturing the site-specific variations in soil temperature is apparent. However, it is always a difficult and time-consuming process to continuously measure soil temperature at various soil depths. Predictions by simulation models and machine learning algorithms can be the most viable alternatives for overcoming this problem to offer data on the temperature of a soil profile. A large number of meteorological stations only observe the variables above the ground surface or install sensors within the station areas to measure soil temperature, instead of installing them in the experimental fields, which creates a chance to make the observed data unrepresentative. To avoid this problem and make this work more accessible, scientists have emphasized on mathematical models that can determine soil temperature using meteorological variables and other factors which affect soil temperature. Simple models can perform poorly, but more complex models can provide better predictions. Three major types of soil temperature prediction models are mechanistic models, statistical relationships, and coupled empirical and mechanistic models (Sandor & Fodor, 2012). These models describe the atmosphere–soil–plant system with the help of mathematical tools and simulate those using computers.

A mathematical model called HYDRUS-1D (Šimůnek et al., 2005) can be used for the simulation of soil temperature dynamics. Richards' equation for saturated–unsaturated water flow and Fickian-based convection–dispersion equations for heat and solute transport are both numerically solved by HYDRUS-1D. Shein et al. (2019) performed research using HYDRUS-1D to validate the program's efficiency for predicting soil moisture and temperature dynamics and found out that the efficiency of prediction was high at surface soil depths (0–15 cm). Kanzari et al. (2018) also performed a comparison between a thermal dispersion model and the HYDRUS-1D model for the simulation of the variation of the water content and the temperature in 30 cm topsoil, which showed that thermal dispersion performed similarly to HYDRUS-1D. Besides HYDRUS-1D simulation, the linear regression model (LRM) has also been used to analyze soil temperature dynamics at various depths. Over the recent decades, the efficiency of data-driven models for simulating complicated nonlinear input–output relationships has been reported by many researchers. Several studies have been done for estimating soil temperature from meteorological parameters through linear or nonlinear methods, for example, multivariable linear regression, artificial neural network, and artificial neural fuzzy inferential system models (Bilgili, 2010; George, 2001; Kim & Singh, 2014; Tabari et al., 2010; Wu et al., 2013). Recently, Delbari et al. (2019) compared linear regression with support vector regression (SVR) in modeling soil surface temperature over diverse climate conditions, which showed that MLR can give poor results at depths over 30 cm, while SVR performs better than MLR at a deeper layer of the soil. Hossein and Ahmed (2017) used extreme machine learning for a similar task and showed that this method and linear regression gave a satisfactory result in predicting temperature at the topsoil (0–30 cm), but the accuracy diminished in the deeper soil layers. Machine learning algorithms, such as LRM, have been utilized to predict various soil physical and hydraulic properties for different types of soil in different regions. However, this type of study using either a machine learning algorithm or physically based simulation is scarce for the region of Bangladesh. It is not wise to use a model that was calibrated or developed based on information from other regions because soil properties vary widely with land topography, organic matter content, crop type, and meteorological parameters.

To find the best and most affordable model, studying multiple methods and comparing their performance are desirable. Also, an efficient study of the soil environment requires multi-depth soil temperature data, and such data are measured only at agro-meteorological stations. A limited number of agro-meteorological stations exist in Bangladesh. To our knowledge, there is no study on the comparison of the HYDRUS model with a machine learning model in predicting soil temperature. For that reason, it is imperative to find out a specific model that can predict the soil temperature of a particular region of Bangladesh that will not only provide accuracy but also reduce the time and cost to avail these data manually. This study illustrates how the temperature of the soil varies with depth with respect to various meteorological parameters and how the changes in meteorological factors affect the soil temperature. In this research, HYDRIS-1D and LRM were used to predict the soil temperature based on various meteorological data. The performances of these two methods were also compared. Therefore, the study has the following two objectives:

(i) To predict soil temperature at different soil depths by using HYDRUS-1D and LRM.
(ii) To quantify the performances of HYDRUS-1D and LRM in predicting the soil temperature values.

Methods

Study Location

The daily meteorological variables used in this study were collected from a weather station of Bangladesh Meteorological Department located at Bangladesh Agricultural University Campus, Mymensingh (24.7196° N, 90.4267° E and 18 m above mean sea level). The study area is under the agro-ecological zone named Old Brahmaputra Floodplain. The land of the study site was covered with local perennial grass of 5–8 cm cut. The proportions of sand, silt, and clay were 42%, 49%, and 10%, respectively, with an organic matter content of 1.1% in the silt loam soil of the location (Amin et al., 2022). The average temperature of June in this area varies from 26.7°C to 32.2°C with a mean relative humidity of 78.13% and an average rainfall of 37 cm. The reference evapotranspiration rate in this region considerably varies in different seasons; 2.9 mm/day in winter, 5.3 mm/day in dry summer, and 4.1 mm/day in wet season (Ali et al., 2005). The depth to water level in the shallow aquifer in this location fluctuates from 3.1 to 6.2 m in different seasons (Amin et al., 2023).

Data Collection and Processing

The measured meteorological and soil physical variables were daily soil temperature at a depth of 5, 10, 20, 30, and 50 cm, air temperature, atmospheric pressure, wind speed, relative humidity, rainfall, evaporation, water temperature, and sunshine duration. Monitoring soil temperatures at various depths was performed using multiple sensors in the field. In this study, data for the month of June 2019 were used (Figure 1).

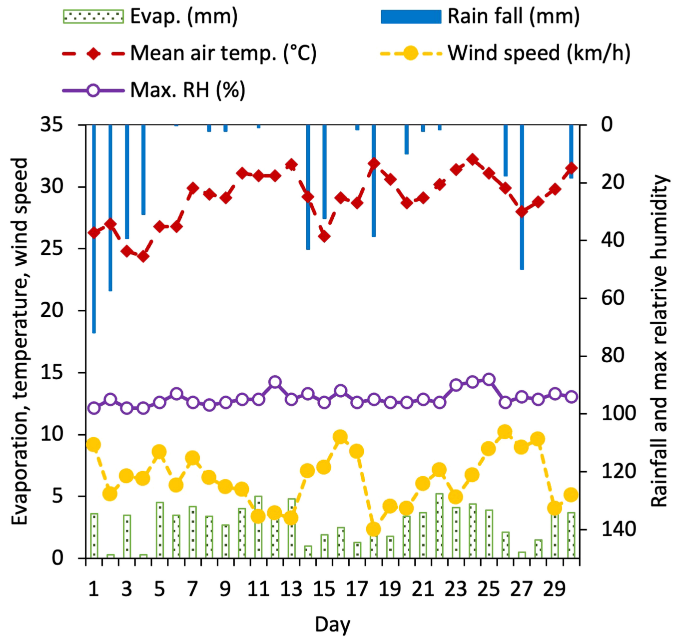

Figure 1.
Daily Variations of Meteorological Parameters in the Month of June 2019.

Prediction by Linear Regression Model

Method Description

Regression analysis enables one to comprehend how the independent variables are changed while the other independent variables are kept constant, and how this alters the typical value of the dependent variable. In this analysis, the conditional expectation of the dependent variable given the independent variables is calculated. The regression function, or estimation objective, is a function of the independent variables in every situation. The linear connection between a scalar dependent variable (Y) and one or more independent variables (X) is known as linear regression. In the case where there is only one explanatory variable, simple linear regression is employed. The procedure is known as multiple linear regression when there are more than one explanatory variable. The general formula for regression is (Menon et al., 2017):

$$Y = aX + c \tag{1}$$

where Y is the measurement of the dependent variable, that is, temperature, X is the independent variables, and c and a are constant. If there are m independent variables and every variable is n dimensional, then X can be written as:

$$X_{n x m} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ X_{21} & X_{22} & \cdots & X_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nm} \end{bmatrix} = \begin{bmatrix} X_{(1)} & X_{(2)} & \cdots & X_{(m)} \end{bmatrix} \tag{2}$$

where X_i are the n-dimensional regression coefficients in the model for the ith variable. The predicted temperature values are obtained by implementing linear regression based on the independent factors. In this study, the least square approach of linear regression has been used.

Least Square Method

In regression analysis, the method of least squares is a common technique for approximating the solution of overdetermined systems by minimizing the sum of the squares of the residuals resulting from each individual equation. The total of squared residuals is reduced in the least-squares sense by the best fit. A function connecting the value of the dependent variable (Y) to the values of an independent variable is found using the conventional formulation. The prediction is given by the following equation (Menon et al., 2017):

$$\check{Y} = aX + c \tag{3}$$

In this equation, the intercept (c) and the slope (a) of the regression line are free variables. The estimate of these parameters, according to the least square approach, is defined as the value that minimizes the sum of squares between the measurements and the model predictions (thus, the name least squares). This amounts to minimizing the expression (Menon et al., 2017):

$$\epsilon = \sum_i \left(Y_i - \check{y}_i\right)^2 = \sum_i \left[Y_i - \left(aX_i + c\right)\right]^2 \tag{4}$$

where ϵ stands for error, which is the quantity to be minimized. Taking the derivative of ϵ with respect to a and c and setting them to zero, we can find the value of a and c. The least square can be extended to more than one independent variable (using matrix algebra) and to nonlinear functions.

Formulation of Linear Regression Model for Soil Temperature Prediction

One of the most critical steps in developing a satisfactory forecasting model is the selection of the input variables. Because these variables determine the structure of the forecasting model and affect the model's weighted coefficient and results, the first step in this analysis is the selection of independent variables. It is known that soil temperature is related to various meteorological variables. Therefore, the daily soil temperature (Y) can be characterized as the function of the air pressure (X_1), maximum air temperature (X_2), minimum air temperature (X_3), average air temperature (X_4), dew point (X_5), maximum humidity (X_6), minimum humidity (X_7), average humidity (X_8), rainfall (X_9), wind speed (X_{10}), sunshine duration (X_{11}), evaporation (X_{12}), and water temperature (X_{13}). The relationship between soil temperature and input variables can be expressed as follows:

$$Y = f\left(X_1, X_2, \ldots X_{13}\right) = a_1 X_1 + a_2 X_2 + \ldots + a_{13} X_{13} + c \quad (5)$$

Here, independent variables must be only included in the model. Because the regression model must be established in a way that the best estimation should be performed using a few independent variables with the maximum possible degree of independence. Cross-correlations between input and output variables were calculated in order to determine the best input structure.

Soil Temperature Prediction by HYDRUS-1D

Model Description

In this study, HYDRUS-1D was implemented to predict the soil temperature at various depths. The HYDRUS-1D model solves the coupled equations governing liquid water, water vapor, and heat transport in the soil, together with the surface water and energy balance for the soil. The code assumes that temperature and pressure gradients work together to drive the movement of liquid and vaporized water in the subsurface. Conduction, convection of sensible heat by liquid water movement, diffusion of latent heat by water vapor, and diffusion of sensible heat by water vapor are all methods for transferring soil heat. Various types of meteorological information can be supplied to solve the surface energy balance at the upper boundary dynamically (Kleissl et al., 2007). Thus, water contents and temperatures of the soil profile can be calculated and coupled to meteorological parameters. In a case study, Saito et al. (2006) showed that soil water dynamics are strongly associated with the soil temperature regime.

Governing Equations

The HYDRUS-1D code for soil heat and water flux has been described in detail by Šimůnek et al. (2005). Convection–dispersion equation is used in HYDRUS for simulating one-dimensional heat transfer modeling. Neglecting the effect of water vapor diffusion on transport, this equation can be expressed as:

$$\frac{\partial C_p\left(\theta\right)T}{\partial t} = \frac{\partial}{\partial x}\left[\lambda\left(\theta\right)\frac{dT}{dx}\right] - C_w\frac{\partial qT}{\partial x} - C_w ST \quad (6)$$

where $\lambda(\theta)$ is the coefficient of the apparent thermal conductivity of the soil, T is temperature, t is time, S is sink term, θ is the volumetric water content, q is the Darcian fluid flux density, and $C_p(\theta)$ and C_w are the volumetric heat capacities of the porous medium and the liquid phase, respectively. C_p (θ) is calculated using the following equation (de Vries, 1963):

$$C_p\left(\theta\right) = C_{nn} + C_{oo} + C_w + C_{av} \quad (7)$$

where θ_n, θ_o, and θ_v are volumetric fraction of solid phase, organic matter, and gas phase, respectively, whereas C_n, C_o, C_a are volumetric heat capacity of solid phase, organic matter, and gas phase, respectively.

The apparent thermal conductivity is defined as (de Marsily, 1986):

$$\lambda\left(\theta\right) = \lambda_o\left(\theta\right) + \beta_t C_w\left|q\right| \quad (8)$$

where β_t is the thermal dispersivity, $\lambda_o\left(\theta\right)$ is the thermal conductivity of the soil defined as (Chung & Horton, 1987):

$$\lambda_o\left(\theta\right) = b_1 + b_2\theta + b_3\theta^{(0.5)} \quad (9)$$

where b_1, b_2, and b_2 are empirical parameters.

Model Parameterization

Soil hydraulic parameters, which were found by van Genuchten–Mualem single porosity model (van Genuchten, 1980) using the known soil texture and bulk density, are shown in Table 1. In this study, the default value for silt-loam soil textures was calculated considering no hysteresis. The value of the heat transport parameters used in this study is given in Table 2.

Q_r is residual soil water content, Q_s is saturated soil water content, K_s is saturated hydraulic conductivity, and I, α, and n are empirical parameters. In this study, the default values for silt loam soil were chosen.

Default values were used for C_n, C_o, and C_w. For water flow parameters calculation, the upper boundary condition was selected as the atmospheric boundary condition with surface runoff, the lower boundary condition as free drainage, and the initial condition as water contents. For heat transport parameters calculation, temperature amplitude was taken as 5°C, and Chung and Horton's method was used for thermal conductivity calculation. The upper and lower boundary conditions for heat transport were selected as the soil temperature boundary conditions.

Prediction Performance Assessment

To evaluate the performance of the linear regression and HYDRUS-1D model, the mean absolute percentage error (MAPE) and the coefficient of determination (R^2) were used to see the convergence between the target values and the output values. Here, MAPE is defined as follows (Melesse & Hanley, 2005):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\text{abs}\left(\frac{o_i - p_i}{o_i}\right)x100 \quad (10)$$

Table 1.
Soil Hydraulic Parameters Used in the Study

Q_r (cm³/cm3)	Q_s (cm³/cm³)	*α* (1/cm)	*n*	k_s (cm/day)	*I*
0.067	0.45	0.02	1.41	10.8	0.5

Table 2.
Soil Heat Transport Parameters

Solid	OM	Disp.	b_1(Wm^{-1}K^{-1})	b_2(Wm^{-1}K^{-1})	b_3(Wm^{-1}K^{-1})	C_n(Jm^{-3}K^{-1})	C_o(Jm^{-3}K^{-1})	C_w(Jm^{-3}K^{-1})
0.55	0.015	5	1.47×10^{19}	1.55×10^{19}	3.16×10^{19}	1.43×10^{14}	1.87×10^{14}	3.12×10^{14}

Disp., dispersion; OM, organic matter.

where o_i is the measured value, p_i is the predicted value, and n is the number of samples.

In addition, the coefficient of determination between the target value and the output value is defined as follows (Bilgili, 2010):

$$R^2 = 1 - \frac{\sum (y_i - \check{y}_i)^2}{\sum (y_i - \bar{y})^2} \tag{11}$$

where y_i is the measured value, and $\check{y}_i$ is the predicted value.

Results

Measured Soil Temperature

The temporal variation of the soil temperature at different depths is shown in Figure 2. Soil temperature at different times of the month varied with the depth of the soil. It was apparent that the soil temperature near the surface was higher than the temperature in deeper soil. Also, the temporal variation in temperature at different depths followed a similar pattern, but the range of fluctuations was diminished with depth. The range of fluctuations at 30 and 50 cm depths was much lower. This occurred due to the high thermal inertia of the soil and the time lag between the temperature fluctuations at the surface and deep in the soil. Therefore, the temperature in the deeper soil was lower than that of the upper soil layers (Kalogirou & Florides, 2004). However, the temperature at the 50 cm layer was higher than that at the 30 cm layer. The 50 cm layer remained relatively warm probably because of the heat conduction from the deeper soil below.

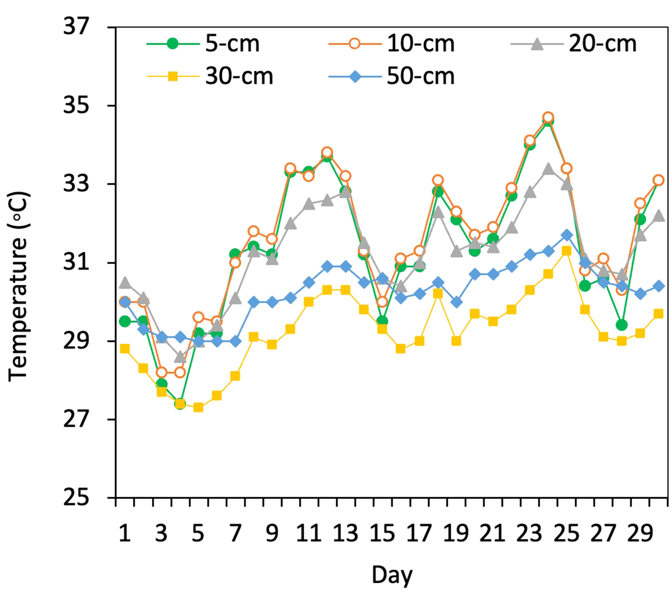

Figure 2.
Variation of the Average Soil Temperature at Different Depths Over the Month of June 2019.

Correlation Between Meteorological Parameters

There was a high rate of correlation between the soil temperature, which is the dependent variable, and the various meteorological variables. The obtained correlation coefficients are shown in Figure 3. Soil temperature has a strong positive correlation with maximum atmospheric air temperature, such as at 5 cm depth the correlation factor was .94 explaining that soil temperature will change proportionally with air temperature, whereas a strong negative correlation with humidity, such as at 5 cm depth the correlation factor was –.63 indicating that the relation of soil temperature is inversely proportional to average humidity, and weak negative correlation with rainfall, the correlation factor was –.47 at 5 cm depth. In addition, there is a weak correlation between soil temperature and air pressure and wind speed. For example, at 5 cm soil depth, the correlation factor for air pressure was .26 and for wind speed was .44.

Linear Regression Model Parameters

In the LRM, the most significant point is to select the predictor variables that provide the best prediction equation for modeling the dependent variable. All independent variables were added to enter the regression model formulated in section 2.3.3 (Formulation of linear regression model for soil temperature prediction), and the following model was obtained for the month of June for the 5 cm soil depth:

$$\begin{aligned} Y_5 &= 0.17 + 0.04X_1 + 1.87X_2 + 1.24X_3 - 2.36X_4 \\ &\quad + 0 * X_5 - 0.20X_6 - 0.15X_7 + 0.01X_8 + 0.04X_9 \\ &\quad - 0.03X_{10} - 0.02X_{11} + 0.12X_{12} + 0.20X_{13} \end{aligned} \tag{12}$$

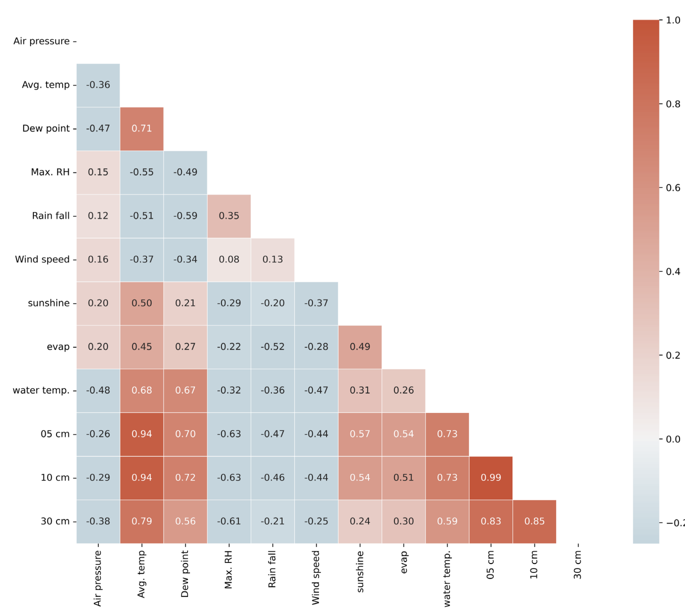

Figure 3.
Correlation Matrix Between Different Meteorological Parameters and Soil Temperature.

where the positive coefficients indicate the proportional relationship of these variables with soil temperature and the negative coefficients refer to an inversely proportional relationship with soil temperature. However, as described earlier and shown in Figure 3, the variables have multiple collinearities. Here, collinearity means some independent variables have a dependency on other independent variables, for example, average temperature and dew point have a high dependency on each other. This collinearity will cause the model to predict inaccurately. This collinearity among variables can also be proved by the variance inflation factor (VIF) (Craney & Surles, 2002). The VIF calculated after the first step of linear regression using all the variables is shown in Table 3. In statistics, the VIF is the ratio of the variance of estimating some parameters in a model that includes multiple other terms by the variance of a model constructed using only one term. In a simple least squares regression analysis, it measures the degree of multicollinearity. If VIF is high, then the variable in the model has high multicollinearity. One can infer that the regression coefficients are inaccurately assessed due to multicollinearity if the VIF is greater than 10 (Miles, 2014).

Table 4 depicts the *p*-values of the independent variables. The *p*-value suggests which variable is statistically significant to predict the soil temperature. Most of the variables have large *p*-value due to high collinearity among the variables.

It is clear from Table 3 that some variables have very high multicollinearity. Therefore, the stepwise regression technique was applied. The VIF values and significant levels (*p*-values) were used to evaluate the estimator performance of the regression model. Thus, the best independent variables were selected for the LRM, and the following model is obtained for the 5 cm soil depth:

$$Y_5 = 0.2 + 0.71X_4 - 0.2X_6 + 0.08X_9 - 0.1X_{10} + 0.14X_{12} \quad (13)$$

For the month of June, five independent variables were used to predict the soil temperature, which are average air temperature, maximum humidity, rainfall, wind speed, and evaporation. Similarly, four other independent equations were found from the LRM for the 10, 20, 30, and 50 cm of soil depth, which are shown as follows, respectively:

$$Y_{10} = 28.38 + 4.82X_4 - 1.22X_6 + 0.44X_9 - 0.72X_{10} + 0.68X_{12} \quad (14)$$

$$Y_{20} = 29.34 + 3.69X_4 - 0.96X_6 + 0.94X_9 - 0.39X_{10} + 0.46X_{12} \quad (15)$$

$$Y_{30} = 27.79 + 2.83X_4 - 1.02X_6 + 1.04X_9 + 0.12X_{10} + 0.18X_{12} \quad (16)$$

$$Y_{50} = 29.31 + 1.79X_4 - 0.8X_6 + 0.54X_9 + 0.46X_{10} + 0.04X_{12} \quad (17)$$

The final VIFs and *p*-values are shown in Table 5. All the VIFs of the final variables are below 5, which infer that the final variables have less collinearity among them. The *p*-values of all the variables are less than .05, which indicates that all the variables are statistically significant in predicting soil temperature.

To evaluate the performance of the LRM, the MAPE and the coefficient of determination (R^2) were used to see the convergence between the target values and the output values. The values of MAPE and R^2 at different soil depths for LRM are given in Table 6.

It is clear that LRM can give very satisfactory results in predicting the temperature at a depth of 5–20 cm. Not only the MAPE values are small, but the R^2 values are also desirable. For soil temperature below 20 cm depth, MAPE were still small, but R^2 values decreased considerably. From the linear regression equation and Figure 3, it is found that the soil temperature at different depths was highly correlated with mean air temperature, evaporation, and relative humidity. However, if the depth of soil increases, the correlation of soil temperature with these meteorological parameters declines.

HYDRUS-1D Predictions

Values of R^2 and MAPE of the HYDRUS-1D predictions of the soil temperature at different depths are shown in Table 7 and it is obvious from the table that the HYDRUS-1D satisfactorily simulated the soil temperature at the shallow depth of the soil (0–20 cm) and less so at the deep soil (30–50 cm). The predictions of HYDRUS-1D matched well with the measured soil temperature values at the depths of 0–20 cm, whereas it overestimated the soil temperature below 20 cm depth (Figures 4 and 5).

Discussion

Previous study shows that soil heat capacity and soil moisture content have a larger impact on soil temperature than meteorological parameter in deeper soil (Bilgili, 2010). In this study, soil

Table 3.
Variance Inflation Factor (VIF) Between Different Meteorological Parameters for All Soil Depths

Mean temperature	Maximum Temperature	Minimum Temperature	Dew Point	Minimum RH	Maximum RH	Average RH	Water temperature	Evaporation	Wind Speed	Air Pressure	Sunshine	Rainfall
47.770	18.270	8773	185	58	54	15	13	10	08	7	5	4

Table 4.
p-Value for Different Meteorological Parameters in Predicting Soil Temperature for All Soil Depths

Mean temperature	Maximum Temperature	Minimum Temperature	Dew Point	Minimum RH	Maximum RH	Average RH	Water Temperature	Evaporation	Wind Speed	Air Pressure	Sunshine	Rainfall
.581	.503	.499	.984	.424	.011	.950	.015	.077	.605	.478	.785	.624

Table 5.
Final Variance Inflation Factor (VIF) and p-Value for Different Parameters for All Soil Depths

Variable	Mean Temperature	Maximum RH	Rainfall	Wind Speed	Evaporation
VIF	4.73	4.95	2.02	3.75	4.55
p	.0	.003	.012	.05	.016

Table 6.
MAPE and R^2 Values at Different Soil Depths for the Linear Regression Model

Depth (cm)	R^2 Value	MAPE (%)
5	0.96	0.81
10	0.94	0.87
20	0.88	1.05
30	0.73	1.30
50	0.55	1.25

MAPE, mean absolute percentage error.

Table 7.
MAPE and R^2 Values for HYDRUS-1D Simulation at Different Soil Depths

Depth (cm)	R^2 Value	MAPE (%)
5	0.85	3.44
10	0.86	2.87
20	0.75	3.73
30	0.58	8.33
50	0.51	4.13

MAPE, mean absolute percentage error.

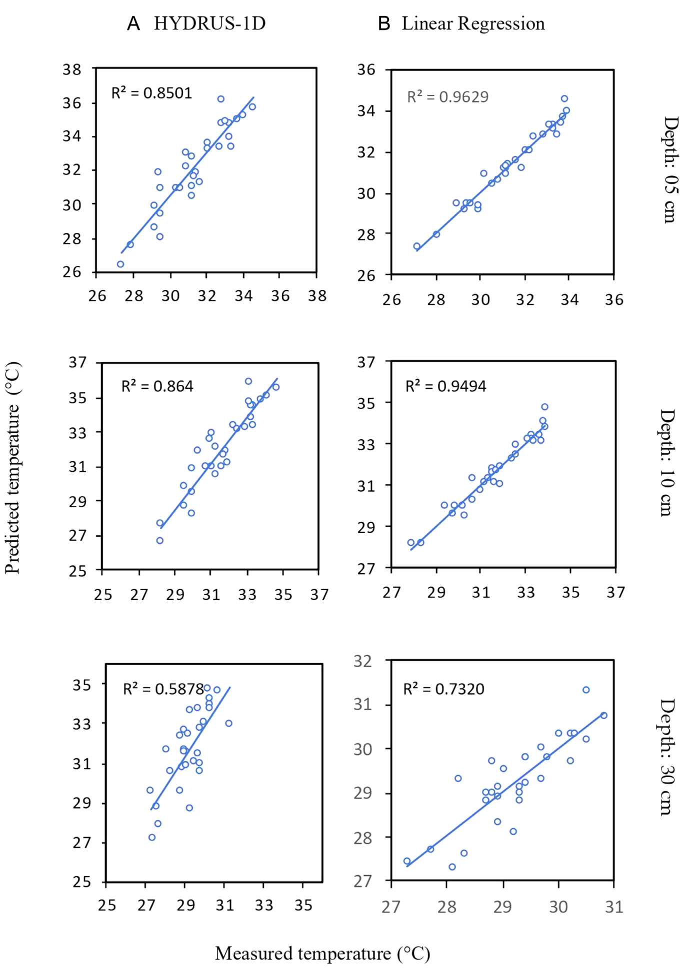

Figure 4.
Comparison of the Measured and Predicted Soil Temperature Values at Different Depths by HYDRUS-1D Simulation and Linear Regression Model.

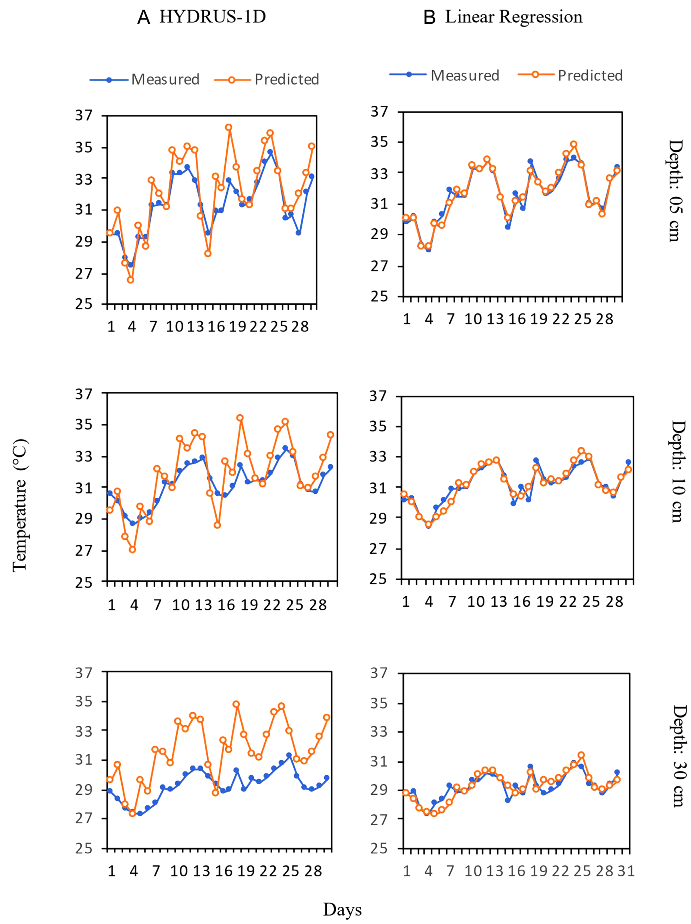

Figure 5.
Measured and Predicted Soil Temperature Values at Different Depths for the Month of June by HYDRUS-1D Simulation and Linear Regression Model.

heat capacity and soil moisture content were not considered in LRM. That is why the predicted values for the deeper soils were not accurate. Although the predicted values for soil depths lower than 20 cm do not vary considerably from the measured value, they do not follow the trend. Therefore, it can be inferred that only the meteorological parameters can predict the soil temperature up to the depth of 20 cm more accurately. Citakoglu (2017) compared the adaptive neural-fuzzy inference system (ANFIS), artificial neural network (ANN), and LRM to predict soil temperatures in Turkey and showed that ANFIS worked better than the other two methods. However, ANFIS and ANN need a lot of data to train the model, whereas LRM can use small data to predict the temperature. Tabari et al. (2010) used relative humidity, air temperature, precipitation, and solar radiation data in ANN and LRM to estimate soil temperatures at different soil depths and found ANN as more suitable than LRM.

However, Tabari et al. (2010) did not consider multicollinearity among independent variables. Sandor and Fodor (2012) compared HYDRUS-1D, CERES, and modified CERES models in predicting soil temperature at different depths in Hungary and showed that the HYDRUS-1D model performed better than the other two models. They also showed that HYDRUS-1D provide acceptable results in deeper soil (40 and 60 cm). However, they calibrated the HYDRUS-1D model using 1 year data before validation and measured soil hydraulic parameters experimentally using different apparatus, whereas in this study, the default values were used for soil hydraulic parameters. Thus, it can be assumed that, to get accurate predictions, the model needs to be calibrated with additional data from different soil layers. The soil's physical properties including soil organic matter content, soil texture, and moisture content can vary in different layers. Sandor and Fodor (2012) stated that the thermal properties of soil of particular regions (e.g., thermal conductivity) are required to predict soil temperature accurately. Since soil temperature is influenced by vegetative growth and soil water balance, the HYDRUS-1D model will give more accurate results if these inputs are integrated in the model.

Comparisons between the measured and predicted values for different models are shown in Figure 4. Also, the fluctuations of the measured and predicted values for different models at different depths over the month are shown in Figure 5. The LRM provided better results than the HYDRUS-1D simulation. Not only the LRM had lower values of MAPE, but it also had higher R^2 values. It means that the LRM predicted the temperature more accurately and captured the trend more precisely than HYDRUS-1D. However, the performance of both models changed with the depth of the soil. The LRM predicted the output based on meteorological parameters given as input to the model. Therefore, it only finds the relationship between the given input and output, and it does not depend on the empirical parameters, such as the thermal properties of soil that need to be measured externally. On the other hand, the HYDRUS-1D model attempts to solve the convection-dispersion equation that requires estimating the thermal conductivity and volumetric heat capacity of the soil (Sandor & Fodor, 2012). The HYDRUS-1D also requires vegetative growth data for better prediction.

Conclusion and Recommaendations

Both HYDRUS-1D and LRMs predicted soil temperature satisfactorily, but their performance varied with the soil depth. The LRM outperformed the HYDRUS-1D model. However, the predicted results of HYDRUS-1D are more mathematically solid and explainable, while the LRM uses a data-driven approach. To get more accuracy from HYDRUS-1D, more accurate calibration with more detailed input data, for example, leaf area index, soil moisture dynamics, etc., would be needed. It appears that five independent variables were found to significantly affect the soil temperature for the month of June, which were mean air temperature, maximum humidity, rainfall, wind speed, and evaporation. Further studies should be conducted to investigate the capability of an LRM in predicting the temperature of more structured soils with higher clay and organic matter contents. The performance of the HYDRUS model can be improved by incorporating more measured values and long-term data. Also, the developed model performance can be validated using more data and for different regions. Additional efforts would be necessary to improve prediction accuracy in the deeper layer of the soil.

Data Availability Statements: All data generated or analyzed during this study are included in this article. The data and materials are available upon request from the corresponding author.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – M.G.M.A., J.F.P.; Design – M.G.M.A., J.F.P.; Supervision – M.G.M.A.; Materials – J.F.P., M.N.H.K.; Data Collection and/or Processing – J.F.P., M.N.H.K.; Analysis and/or Interpretation – M.G.M.A., J.F.P., M.N.H.K.; Literature Search – J.F.P., M.N.H.K.; Writing Manuscript – J.F.P.; Critical Review – M.G.M.A.

Declaration of Interests: The authors declare that they have no competing interest.

Funding: The authors declared that this study has received no financial support.

Veri Kullanılabilirliği Beyanları: Yazarın talebi üzerine veriler ve materyaller mevcut olup, üretilen veya analiz edilen tüm veriler çalışmada yer almaktadır.

Hakem Değerlendirmesi: Dış bağımsız.

Yazar Katkıları: Fikir – M.G.M.A., J.F.P.; Tasarım – M.G.M.A., J.F.P.; Denetleme – M.G.M.A.; Malzemeler – J.F.P., M.N.H.K.; Veri Toplanması ve/veya İşlemesi – J.F.P., M.N.H.K.; Analiz ve/veya Yorum – M.G.M.A., J.F.P., M.N.H.K.; Literatür Taraması – J.F.P., M.N.H.K.; Yazıyı Yazan – J.F.P.; Eleştirel İnceleme – M.G.M.A.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemişlerdir.

Finansal Destek: Yazarlar bu çalışma için finansal destek almadıklarını beyan etmişlerdir.

References

Ali, M. H., Amin, M. G. M., & Islam, A. K. M. R. (2005). Reference crop evapotranspiration (ETo) over Bangladesh and its implication in crop planning. *Journal of Bangladesh Agricultural University*, *3*(1), 139–147. [CrossRef]

Amin, M. G. M., Al Mİnhaj, A. A., Bhowmİk, B., Islam, D., & Islam, M. N. (2022). Nitrogen and phosphorus leaching and vegetative growth of maize as affected by animal manure application. *Eurasian Journal of Soil Science*, *11*(1), 17–24. [CrossRef]

Amin, M. G. M., Lima, L. A., Rahman, A., Liu, J., & Jahangir, M. M. R. (2023). Dairy manure application effects on water percolation, nutrient leaching and rice yield under alternate wetting and drying irrigation. *International Journal of Plant Production*, *17*(1), 95–107. [CrossRef]

Amin, M. G. M., Minhaj, A. A., Islam, D., Bhowmik, B., Hasan, M. M., & Islam, M. N. (2021). Mulch and no-till impacts on nitrogen and phosphorus leaching in maize field under sub-tropic monsoon climate. *Environmental Challenges*, *5*, 5100346. [CrossRef]

Bilgili, M. (2010). Prediction of soil temperature using regression and artificial neural network models. *Meteorology and Atmospheric Physics*, *110*(1–2), 59–70. [CrossRef]

Chung, S. O., & Horton, R. (1987). Soil heat and water flow with a partial surface mulch. *Water Resources Research*, *23*(12), 2175–2186. [CrossRef]

Citakoglu, H. (2017). Comparison of artificial intelligence techniques for prediction of soil temperatures in Turkey. *Theoretical and Applied Climatology*, *130*(1–2), 545–556. [CrossRef]

Craney, T. A., & Surles, J. G. (2002). Model-dependent variance inflation factor cutoff values. *Quality Engineering*, *14*(3), 391–403. [CrossRef]

De Marsily, G. (1986). *Quantitative hydrogeology*. Paris School of Mines.

De Vries, D. A. (1963). Thermal properties of soils. In W. R. van Wijk (Ed.). *Physics of plant environment* (pp. 210–235). North-Holland Publishing Co.

Delbari, M., Sharifazari, S., & Mohammadi, E. (2019). Modeling daily soil temperature over diverse climate conditions in Iran—A comparison of multiple linear regression and support vector regression techniques. *Theoretical and Applied Climatology*, *135*(3–4), 991–1001. [CrossRef].

George, R. K. (2001). Prediction of soil temperature by using artificial neural networks algorithms. *Nonlinear Analysis: Theory, Methods and Applications*, *47*(3), 1737–1748. [CrossRef]

Hossein, Z., & Ahmed, F. F. (2017). Estimating soil temperature from metrological data using extreme learning machine, artificial neural network and multiple linear regression models. *Iranian Journal of Soil and Water Research*, *51*(4), 895–906

Jebamalar, A. S., Raja, S., Thambi, A., & Sunitha, B. S. J. (2012). Prediction of annual and seasonal soil temperature variation using artificial neural network. *Indian Journal of Radio and Space Physics*, *41*(1), 48–57.

Kalogirou, S. A., & Florides, G. A. (2004). *Measurements of ground temperature at various depths*. https://ktisis.cut.ac.cy/bitstream/20.500.14279/2417/3/C55-PRT020-SET3.pdf

Kanzari, S., Mariem, S. B., Nouna, B. B., & Ilahy, R. (2018). Comparison of Hydrus-1D to thermal dispersion model for water-heat transport in a semi-arid region of Tunisia. *Asian Journal of Applied Sciences*, *11*(2), 64–70. [CrossRef]

Karnieli, A., Agam, N., Pinker, R. T., Anderson, M., Imhoff, M. L., Gutman, G. G., Panov, N., & Goldberg, A. (2010). Use of NDVI and land surface temperature for drought assessment: Merits and limitations. *Journal of Climate*, *23*(3), 618–633. [CrossRef]

Kim, S., & Singh, V. P. (2014). Modeling daily soil temperature using data driven models and spatial distribution. *Theoretical and Applied Climatology*, *118*(3), 465–479. [CrossRef]

Kisi, O., Tombul, M., & Kermani, M. Z. (2015). Modeling soil temperatures at different depths by using three different neural computing techniques. *Theoretical and Applied Climatology*, *121*(1–2), 377–387. [CrossRef]

Kleissl, J., Moreno, H., Hendrickx, J. M. H., & Simunek, J. (2007). HYDRUS simulations of soil surface temperatures. In *Detection and Remediation Technologies for Mines and Minelike Targets*, *XII*(6553), 213–224. [CrossRef]

Kourat, T., Smadhi, D., Mouhouche, B., Gourari, N., Mostofa Amin, M. G. M., & Bryant, C. R. (2021). Assessment of future climate change impact on rainfed wheat yield in the semi-arid Eastern High Plain of Algeria using a crop model. *Natural Hazards*, *107*(3), 2175–2203. [CrossRef]

Melesse, A. M., & Hanley, R. S. (2005). Artificial neural network application for multi-ecosystem carbon flux simulation. *Ecological Modelling*, *189*(3–4), 305–314. [CrossRef]

Menon, S. P., Bharadwaj, R., Shetty, P., Sanu, P., & Nagendra, S. (2017). Prediction of temperature using linear regression. In International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEEECCOT) (pp. 1–6). [CrossRef].

Miles, J. (2014). Tolerance and variance inflation factor. *Wiley statsref: Statistics reference*. [CrossRef]

Saito, H., Šimůnek, J., & Mohanty, B. P. (2006). Numerical analysis of coupled water, vapor, and heat transport in the vadose zone. *Vadose Zone Journal*, *5*(2), 784–800. [CrossRef]

Sándor, R., & Fodor, N. (2012). Simulation of soil temperature dynamics with models using different concepts. *TheScientificWorldJournal*, *2012*, 590287. [CrossRef]

Shein, E., Mady, A., & Ili'n, L. (2019). Validation of HYDRUS-1D for Predicting of Soil Moisture Content with Hysteresis Effect. *Biogeosystem Technique*, *6*(1), 59–64.

Šimůnek, J., Van Genuchten, M. T., & Šejna, M. (2005). The HYDRUS-1D software package for simulating the one dimensional movement of water, heat, and multiple solutes in variably-saturated media. *University of California-Riverside Research Reports*, *3*, 1–240.

Tabari, H., Sabziparvar, A., & Ahmadi, A. (2010). Comparison of artificial neural network and multivariate linear regression methods for estimation of daily soil temperature in an arid region. *Meteorology and Atmospheric Physics*, *110*(3), 135–142.

Wu, W., Tang, X. P., Guo, N. J, Yang, C., Liu, H. B., & Shang, Y. F. (2013). Spatiotemporal modeling of monthly soil temperature using artificial neural networks. *Theoretical and Applied Climatology*, *113*(3–4), 481–494.

Yadav, B., Krishnan, P., Shafeeq, P. M., Parihar, C. M., & Aggarwal, P. (2020). Modelling soil thermal regime in wheat using HYDRUS-2D under diversified maize-wheat-mungbean cropping system. *Catena*, *194*, 104765.

Yılmaz, T., Özbek, A., Yılmaz, A., & Büyükalaca, O. (2009). Influence of upper layer properties on the ground temperature distribution. *Journal of Thermal Science and Technology*, *29*(2), 43–51.