GLUE uncertainty analysis of hybrid models for predicting hourly soil temperature and application wavelet coherence analysis for correlation with meteorological variables

Akram Seifi[1*], Mohammad Ehteram[2], Fatemeh Nayebloei[3], Fatemeh Soroush[1], Bahram Gharabaghi[4], Ali Torabi Haghighi[5]

[1]Department of Water Science & Engineering, College of Agriculture, Vali-e-Asr University of Rafsanjan, Iran.

([*]Corresponding author: a.seifi@vru.ac.ir) https://orcid.org/0000-0003-0887-1217

[2]Department of Water Engineering and Hydraulic Structures, Faculty of Civil Engineering, Semnan University, Semnan, Iran.

[3]Department of Irrigation and Drainage Engineering, Tarbiat Modares University, Tehran, Iran.

[4]Professor, School of Engineering, University of Guelph, Guelph, Ontario, NIG 2W1, Canada; bgharaba@uoguelph.ca; DOI: 0000-0003-0454-2811.

[5]Water, Energy and Environmental Engineering Research unit, University of Oulu, P.O. Box 4300, FIN-90014 Oulu, Finland.

Abstract

Accurate prediction of soil temperature (T_s) is critical for efficient soil, water and field crop management. In this study, hourly T_s variations at 5, 10, and 30 cm soil depth were predicted for an arid site (Sirjan) and a semi-humid site (Sanandaj) in Iran. Existing machine learning models have high performance, but suffer from uncertainty and instability in prediction. Therefore, GLUE approach was implemented to quantify model uncertainty, while wavelet coherence was used to assess interactions between T_s and meteorological parameters. Standalone machine learning models (adaptive neuron fuzzy interface system (ANFIS), support vector machine model (SVM), radial basis function neural network (RBFNN), and multilayer perceptron (MLP)) were hybridized

with four optimization algorithms (sunflower optimization (SFO), firefly algorithm (FFA), salp swarm algorithm (SSA), particle swarm optimization (PSO)) to improve T_s prediction accuracy and reduce model uncertainty. For both arid and semi-humid sites, ANFIS-SFO produced the most accurate performance at studied soil depths. At best, hybridization with SFO (ANFIS-SFO, MLP-SFO, RBFNN-SFO, SVM-SFO) decreased RMSE by 5.6%, 18%, 18.3%, and 18.2% at 5 cm, 11.8%, 10.4%, 10.6%, and 12.5% at 10 cm, and 9.1%, 12.1%, 13.9%, and 14.2% at 30 cm soil depth compared with the respective standalone models. GLUE analysis confirmed the superiority of hybrid models over the standalone models, while the hybrid models decreased the uncertainty in T_s predictions. ANFIS-SFO covered 95%, 94%, and 96% observation data at 5, 10, and 30 cm soli depths, respectively. Wavelet coherence analysis demonstrated that air temperature, relative humidity, and solar radiation, but not wind speed, had high coherence with T_s at different soil depths at both sites, and meteorological parameters mostly influenced T_s in upper soil layers. In conclusion, uncertainty analysis is a necessary and powerful technique to obtain an accurate and realistic prediction of T_s. In contrast, wavelet coherence analysis is a useful tool to investigate the most effective variables that strongly affect predictions.

Keywords: Bio-inspired meta-heuristic optimization, Soil depth, Sunflower optimization, Uncertainty analysis, Wavelet coherence analysis.

1- Introduction

Soil temperature (T_s) is an influential and vital parameter in sustainable agriculture and geosciences practices since it greatly influences physical, geological, chemical, and microbiological processes in the soil (Feng et al. 2019; Alizamir et al. 2020). Thus, more research

about T_s is required for soil, plant, and water planning and managing (Zeynoddin et al. 2020), and also is essential for hydrological, meteorological, and environmental modeling (Bonakdari et al. 2019). Soil temperature is dependent on meteorological factors (solar radiation, air temperature, precipitation, wind speed, pressure gradient), and there are relationships between them (Lehnert 2014; Stajkowski et al. 2020).

Direct measurement of hourly T_s using sensors and instruments is expensive in Iran. Therefore, T_s measurements are not easily available for developing countries, such as Iran (Mehdizadeh et al., 2020a). In addition, inaccurate measurement of T_s causes some instability and uncertainty, and thus precise monitoring and managing of agricultural plans will be complex. With this viewpoint, indirect methods, including analytical, numerical, and data-driven models, have been suggested for the accurate prediction of T_s. However, analytical models cannot reproduce the complex relationships between T_s and climatological variables, especially for different climates and global scales (Plauborg, 2002). Numerical models use finite volume and finite element methods to solve heat transfer laws and are highly complex (Samadianfard et al. 2018b; Sofyan et al. 2020).

As an alternative, intelligence machine learning models can predict T_s and describe the interactions between T_s and other predictor variables, such as meteorological parameters (Moazenzadeh and Mohammadi 2019; Feng et al. 2019). Furthermore, current advances in intelligence machine learning models improve their capability for modelling nonlinear relationships (Padarian et al. 2020) to accurately predict processes and quantify the uncertainties.

Recently, machine learning approaches such as adaptive neuron fuzzy inference system (ANFIS) (Abyaneh et al. 2016; Citakoglu 2017; Singh et al. 2018; Penghui et al. 2020), multilayer perceptron (MLP) (Heddam, 2019; Zeynoddin et al. 2019; Sihag et al. 2020), support vector machine (SVM) (Xing et al. 2018; Delbari et al. 2019; Shamshirband et al. 2020), and radial basis

function neural network (RBFNN) (Kisi et al. 2015) have been widely used in geosciences engineering and T_s prediction owning to extra fast training ability of nonlinear and hidden relationships.

However, most machine learning models are prone to membership function learning problems, premature convergence, over-fitting, and low convergence speed (Qasem et al. 2019; Zhao et al. 2019; Mehdizadeh et al. 2020a). Therefore, several studies have proposed hybridizing standalone machine learning approaches using meta-heuristic algorithms to optimize the performance of conventional models and reduce computation cost (Seifi and Soroush 2020; Abualigah et al. 2021b).

Quick data-processing speed, achieving a global optimum, robust generalization (Moazenzadeh and Mohammadi 2019), powerful searchability, and gradient-free mechanisms (Abualigah et al. 2021a; Abualigah and Diabat 2021) are advantages of integrated machine learning-optimization algorithm techniques over conventional models. Samadianfard et al. (2018b) developed a coupled MLP-firefly algorithm (MLP-FFA) model for monthly T_s estimation in the 0-100 cm soil layer in Adana, Turkey, using the climate variables T_a, atmospheric pressure (P_a), and solar radiation (R_s), and found that it performed better than the standalone MLP model, especially at 20 cm soil depth. Moazenzadeh and Mohammadi (2019) estimated daily T_s at six soil depths in Maragheh, Iran, using two standalone models, elman neural network (ENN) and support vector regression (SVR), and four coupled models, (SVR-FFA), SVR-krill herd algorithm (SVR-KHA), ENN-FFA, and ENN-KHA. They found that all models had their highest accuracy at 10 cm soil depth and that SVR-KHA showed the best estimation accuracy based on T_a, relative humidity (RH), and sunshine hours. Mehdizadeh et al. (2020b) hybridized the ENN model with two algorithms, ant colony optimization (ACO) and gravitational search algorithm (GSA), to predict daily T_s at four soil

depths in Rasht and Isfahan, Iran, and found that ENN-GSA produced the most accurate predictions at all depths. Penghui et al. (2020) found that mutation Salp Swarm Algorithm and Grasshopper Optimization Algorithm (ANFIS-mSG) enhanced the estimation accuracy of ANFIS in predicting daily T_s at 10 cm soil depth based on maximum, mean, and minimum air temperature. Hourly evaluation of T_s is essential for high-resolution modelling within hydrology, geosciences, ecology, and crop models as a decision-support system to achieve better crop yield (Qi et al. 2016; Sanikhani et al. 2018). However, since hourly T_s measurement with instruments is difficult, previous studies have generally used daily or monthly datasets in T_s prediction.

To our knowledge, few studies have investigated the precision of conventional machine learning models in predicting T_s at an hourly scale (Araghi et al. 2017; Feng et al. 2019; Li et al. 2020), while none has investigated model hybridization with integrated optimization algorithms to improve hourly T_s estimation in different climates.

The problem of hybrid meta-heuristic models is complex structure that may cause some instability and uncertainty in predictions. Despite high attention to apply hybrid meta-heuristic models in T_s prediction, hybrid model uncertainties have large effects on the accuracy and reliability of predictions (Seifi et al. 2021). Uncertainty analysis is a robust post-processing step to assess the reliability of different hybrid models (Seifi et al. 2020a,b). However, previous studies on T_s prediction are primarily based on standalone machine learning models and consider straightforward prediction, with no detailed investigation of the associated uncertainty and coherence analysis in different climates of Iran.

To the best of our knowledge, the assessment of uncertainty associated with T_s modelling has been disregarded. Hence, more studies are needed to assess the stability of predictions. Uncertainty analysis provides a comprehensive evaluation of hybrid models behaviors and often calculates a

95% confidence interval (95% CI) to reflect estimations covering. With this viewpoint, generalized likelihood uncertainty estimation (GLUE) was applied in this study for uncertainty analysis to compare the validity and reliability of the models. The GLUE approach was widely used for investigating model uncertainty and sensitivity evaluation of parameters in hydrological studies. Previous studies indicated that the GLUE approach has high potential to identify model uncertainty. It is a more suitable technique than Markov Chain Monte Carlo for comparing predictions of models based on uncertainty (Sun et al. 2016).

Inferring the relationships between climate variables and hourly T_s data can better understand soil temperature response to model structure. Monitoring and evaluating various effects of meteorological variables on hourly T_s prediction and uncertainty is informative and can be used as a decision tool.

However, there is limited information about the relationships between hourly T_s and meteorological variables over time in agricultural areas. Furthermore, the effect of meteorological variables on hourly T_s differs with soil depth and time. In this context, it is valuable and important using mathematical computation tools to find the behavior of hourly T_s at different soil depths relative to meteorological variables.

More essentially, there is a need for engineers and predictors to understand which meteorological variable has impressive relation with hourly T_s and to determine variables with a detrimental role in predicting T_s. Recently, statistical techniques are highly applied to investigate the relationship between variables.

Wavelet analysis is a robust non-parametric computational change detection approach in environmental modelling to discover relationships in non-stationary systems (Li et al., 2019). Moreover, wavelet coherence analysis implements two aims of (1) dependencies analysis between

two datasets and (2) indicating the time point of synchronizing and high correlation between two datasets (Damos and Caballero 2021).

Therefore, the application of wavelet coherence analysis as a feature selection technique eliminates unnecessary and uninformative features (Abualigah 2019). Wavelet coherence analysis has been applied in different studies for relations between environmental factors such as winter and spring thaw CO_2 and N_2O fluxes and T_s (Furon et al. 2008), meteorological parameters and soli water content (Li et al. 2019), soil moisture and eco-hydrometric factors of rainfall and evapotranspiration (Lee and Kim 2019), winter precipitation and three oceanic sources (Ebrahimi et al. 2021).

In the present study, wavelet coherence analysis examines the relation between hourly T_s and meteorological variables of T_a, RH, wind speed (U), R_s during the time. To our knowledge, based on a literature review, no previous study has applied wavelet coherence analysis to investigate the influences of meteorological variables on the scale-dependence of hourly T_s in different soil depths.

Despite the popularity of hybrid optimization models for daily and monthly T_s prediction, few studies have compared hybrid model precision in predicting hourly T_s, primarily by analyzing related uncertainty and investigating significant meteorological factors. There have also been few accuracy evaluations of hybrid ANFIS and SVM models in estimating hourly T_s in Iran. Therefore, four machine learning models of ANFIS, MLP, SVM, and RBFNN were hybridized with four bio-inspired meta-heuristic algorithms (sunflower optimization (SFO), salp swarm algorithm (SSA), FFA, particle swarm optimization (PSO)) to predict hourly T_s at three soil depths in two climate zones.

The main and primary objective of this study is to determine the accuracy and feasibility of developed hybrid models for predicting hourly T_s. In particular, according to the before review and explanations, the major contributions of the present study can be expressed as follow.

1) The innovation and critical contributions of this is performing a robust GLUE approach to identify hybrid models uncertainty for predicting hourly T_s at different soil depths.
2) Another significant and novel contribution is determining the relationships between hourly T_s and T_a, U, RH, and R_s at different soil depths, using wavelet coherence analysis.
3) Develop and compare ANFIS, MLP, RBFNN, and SVM hybridized with the SFO, SSA, FFA, and PSO algorithms in predicting hourly T_s at three soil depths in two climate zones.
4) Select best values of optimization algorithm parameters using the Taguchi method.

2- Materials and Methods

2-1- Study area and dataset

Two datasets of measured hourly T_s were chosen to evaluate the hybrid and standalone models. The first dataset was obtained from an automatic monitoring station in a pistachio orchard (55˚82′ N, 29˚30′ E) in Sirjan city, Iran (Fig. 1), where information about T_s at soil depth is essential in disease control, evaporation modelling, irrigation management, and frost protection. Sirjan has an arid climate, with a mean annual air temperature of 27˚C and a mean annual rainfall of 40 mm. Hourly T_s and meteorological parameters (T_a, R_s, U, RH) were monitored from 12 September 2012 to 17 January 2012.

From 1966 data obtained 75% hourly data were selected for training the models, and the remaining 25% hourly data were used for model testing. Different percentages of the total dataset were examined for training and testing subsets to calculate RMSE. Since 75% for training and 25% for

testing had the lowest RMSE, these percentages were selected as desired subsets for soil temperature modelling.

The second T_s dataset was obtained from a synoptic station in Sanandaj (35.33°N, 47.00°E, 1373.4 m height), western Iran (Fig. 1). This region has a semi-humid climate, based on the calculated aridity index using 30-year meteorological data (Zolfaghari et al. 2016), with hot conditions in summer and very cold winters.

The mean annual air temperature measured at Sanandaj station is 13.4˚C, and the mean annual rainfall is 450 mm. Hourly T_s and meteorological parameters (T_a, R_s, U, RH) were recorded from 9 July 2009 to 12 October 2009. A subset from 75% hourly T_s data was selected and used for training the models, while 25% of hourly T_s data were used for testing.

Time series plots of T_a and hourly T_s at three depths (5, 10, and 30 cm) in the soil at Sirjan and Sanandaj revealed that temperature variations had a decreasing trend at all soil depths (Fig. 2). The most significant difference in temperature between initial and final points was at 5 cm soil depth (15˚C) in Sirjan, while in Sanandaj station, it was 21 ˚C at 10 cm and 19˚C at 30 cm soil depth. Statistical characteristics of the measured meteorological and T_s data are given in Table 1.

2-2- Generalized likelihood uncertainty estimation (GLUE)

The GLUE analysis was performed to investigate predictive model reliability. The analysis has three important levels (Sun et al. 2016):

Level 1: Random sampling is used to create many sampling sets from prior distribution of input data.

Level 2: Likelihood value is calculated from model runs and compared with a particular threshold value to evaluate each input parameter as behavioral (value above the threshold likelihood) or non-

behavioral (value below the threshold likelihood). Behavioral parameters are retained to judge the models. The likelihood is calculated as:

$$L(\theta_i|Y) = exp\left[\frac{-N\sigma_i^2}{\sigma_{obs}^2}\right] \quad (1)$$

where N is an adjustable parameter, σ_i^2 is the error variance for the i^{th} model, σ_{obs} is the variance of observations, θ_i is the parameter set, and $L(\theta_i|Y)$ is the likelihood measure for the i^{th} model calculated with the observations Y.

Level 3: Simulation weights for behavioral parameter sets are rescaled, and the cumulative weighted distribution of estimations is used in quantile estimation for uncertainty prediction:

$$\vartheta_i = \frac{L(\theta_i)}{\sum_{k=1}^{n} L(\theta_k)} \quad (2)$$

where ϑ_i is a likelihood weight, and n is the number of data.

Two indices, p and r, are used to quantify model uncertainty, where p is the percentage of bracketed observations at 95% prediction uncertainty (95PPU), and r is the mean 95PPU range, separated using the standard deviation of observations (Seifi et al. 2020a):

$$p = 100\frac{count\left(T_s|T_{s_L} \le T_s \le T_{s_U}\right)}{N} \quad (3)$$

$$r = \frac{1}{N \times \sigma}\sum(T_{s_L} - T_{s_U}) \quad (4)$$

where N is the number of observations, σ is the standard deviation of observations, and T_{s_L} and T_{s_U} are the lower and upper boundary of the 95% prediction uncertainty, respectively. Lower values of r and higher values for p indicate lower uncertainty. In this study, the GLUE technique's spectral responses were used to assess variations in r and p.

2-3- Gamma Test

Before training the standalone and hybrid models, the best input combination was selected using Gamma Test (GT) for two stations. Recently, GT approach has been widely used in different research fields such as rainfall runoff modelling (Singh et al., 2018), suspended sediment load prediction (Panahi et al., 2021), optimal design of groundwater monitoring networks (Azadi et al., 2020), predicting groundwater level (Sharafati et al., 2020), and predicting evapotranspiration (Seifi and Riahi, 2020).

Han et al. (2010) illustrated that the GT approach effectively reduces model development workload over trial and error procedures. In this approach, the main important input variables can be ranked by calculating V_{ratio} index as the following steps (Sharafati et al. 2020):

(1) Determine output (y_i) and input (x_i) values where $1 \leq i \leq M$.

(2) Compute the delta function as $\delta_M(k) = \frac{1}{M}\sum_{i=1}^{M}\left|x_{[i,k]} - x_i\right|^2 \quad (1 \leq k \leq p)$, where |…| is the distance of Eulirean and $x_{[i,k]}$ is the k^{th} nearest neighbors for each x_i.

(3) Calculate the gamma function as $\gamma_M(k) = \frac{1}{2M}\sum_{i=1}^{M}\left|y_{N[i,k]} - y_i\right|^2 \quad (1 \leq k \leq p)$, where $y_{[i,k]}$ is the k^{th} nearest neighbors for each y_i.

(4) Consider regression line for points of $[\delta_M(k), \gamma_M(k)]$ to calculate gamma statistic (Γ) as: $\gamma = A\delta + \Gamma$.

(5) Calculate the V_{ratio} index as $V_{ratio} = \frac{\Gamma}{\sigma^2(y)}$, where $\sigma^2(y)$ indicates the target outputs variance. The Γ and V_{ratio} values close to 1 show a poor model performance.

2-4- Wavelet coherence analysis in interaction analysis

Wavelet coherence analysis of hourly T_s and the four meteorological variables dataset are used to determine the scale effects of meteorological variables in different soil depths. T_s depends on different meteorological variables such as T_a, RH, U, and R_s. Wavelet coherence analysis calculates the coherence value of cross-wavelet transform among two time series in the time-frequency domain (Lee and Kim 2019). The continuous wavelet transformation (CWT) of a factor over time ($Y_i, i = 1,2, \dots, N$) and uniform time step (δt) is presented as:

$$W_i^Y = \sqrt{\frac{\delta t}{s}} \sum_{j=1}^{n} Y_j \varphi \left[(j-1) \frac{\delta t}{s} \right] \quad (5)$$

where $\varphi[]$ is the basic wavelet function and s is the scale. $W_i^Y(s)$ is defined as $a + ib$; where a is the real component and b is the imaginary component of $W_i^Y(s)$.

The Morlet wavelet is defined as:

$$\varphi(\eta) = \pi^{-1/4} e^{i\omega\eta - 0.5\eta^2} \quad (6)$$

where ω is the dimensionless frequency and η is the dimensionless time with equation of $\eta = s/t$. The cross-wavelet transform between hourly T_s and meteorological variables over time can be written as:

$$|W_i^{XY}(s)| = \left| W_i^X(s) \bar{W}_i^Y(s) \right| \quad (7)$$

where W_i^X and W_i^Y are the wavelet coefficients of T_s and meteorological parameters in the time series X and Y, respectively, $W_i^{XY}(s)$ is the cross-wavelet power spectrum of X and Y, and $\bar{W}_i^Y(s)$ is the complex conjugate of $W_i^Y(s)$.

In the present case, wavelet coherence between T_s and meteorological variables at each scale and occasion is calculated as (Lee and Kim 2019):

$$Coherence = \frac{\left|S\left(s^{-1}W_i^{XY}(s)\right)\right|^2}{S(s^{-1}|W_i^X(s)|^2)S(s^{-1}|W_i^Y(s)|^2)} \quad (8)$$

where S is the smoothing operator and can be defined as:

$$S(W) = S_{scale}\left(S_{time}(W(s,\tau))\right) \quad (9)$$

where τ is the occasion, S_{scale} is smoothing along with the wavelet scale axis, and S_{time} is smoothing in time.

The normalized real Morlet wavelet can be written as:

$$\frac{1}{2\sqrt{2\pi}} exp\left(-\frac{\tau^2}{2s^2}\right) \quad (10)$$

Thus, the smoothing across occasions and smoothing across the scales can be written as:

$$S_{time}(W(s,\tau)) = \sum_{k=1}^{N}\left(W(s,\tau)\frac{1}{s\sqrt{2\pi}} exp\left(-\frac{(\tau - t_k)^2}{2s^2}\right)\right) \quad (11)$$

$$S_{scale}(W(s_k,t)) = \frac{1}{2m+1}\sum_{j=k-m}^{k+m}\left(S_{time}\left(W(s_j,t)\right)\prod(0.6s_j)\right) \quad (12)$$

where m is the number of terms on each symmetrical half of the window, and Π is the rectangle function.

2-5- Machine learning and optimization algorithms description

In the present study, four machine learning models, namely ANFIS, MLP, RBFNN, and SVM, were applied to predict hourly T_s. Also, four meta-heuristic optimization algorithms, including the SFO, FFA, SSA, and PSO were hybridized with standalone machine learning models.

2-4-1- Adaptive neuron fuzzy inference system (ANFIS)

ANFIS model combines fuzzy logic and neural networks while benefiting from both methods' advantages (Najafi-Ghiri et al. 2019). In addition, ANFIS model uses the Takagi-Sugeno inference approach to generate fuzzy "if-then" rules, from input to output domains (Penghui et al. 2020).

$$Rule1: if\ x\ is\ A_1 and\ y\ is\ B_1, then\ f_1 = p_1x + q_1y + r_1 \quad (13)$$

$$Rule2: if\ x\ is\ A_2 and\ y\ is\ B_2, then\ f_2 = p_2x + q_2y + r_2 \quad (14)$$

where A_1, A_2, B_1, and B_2 are membership functions and p_1, q_1, r_1, and p_2, q_2, r_2 are consequence parameters.

The ANFIS structure contains five layers with different inputs and one output. The structure of ANFIS model is summarized in the following steps:

Layer 1: each node adjusts to a function parameter and produces a value of membership degree using the bell membership function.

$$O_{1i} = \mu_{Ai}(x), \quad i = 1,2 \quad (15)$$

$$O_{1i} = \mu_{B_{i-2}}(y), \quad i = 3,4 \quad (16)$$

$$\mu(x) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i}\right)^2\right]^{b_i}}, i = 1,2 \quad (17)$$

where a_i, c_i, and b_i are the membership values.

Layer 2: the output of each node is defined that is input signals production to the node and shows a firing strength for each rule.

$$O_{2i} = \mu_{Ai}(x) \times \mu_{B_{i-2}}(y) \quad (18)$$

Layer 3: the output of this layer calculates the ratio of strength for i^{th} rule to the sum strength of all rules.

$$O_{2i} = \bar{w}_i = \frac{\omega_i}{\sum_{i=1}^{2} \omega_i} \quad (19)$$

Layer 4: the adaptive nodes are calculated in this layer.

$$O_{4i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad (20)$$

Layer 5: the network output is computed in this layer.

$$O_5 = \sum_i \bar{w}_i f_i \quad (21)$$

2-4-2- Multilayer perceptron (MLP) model

MLP is a feedforward supervised neural network that has been applied successfully for complex and nonlinear problems. The backpropagation learning algorithm is commonly used for training MLP models, but it may get trapped in a local optimum (Pouladi et al. 2019). The MLP model uses multiple layers with a nonlinear activation function to learn the relationship between input and output datasets. The first layer (input layer) contains the inputs to the MLP model, while the middle layers (hidden layers) have several neurons. Each neuron performs a weighted summation of inputs. The activation functions are used to calculate the inner product of input parameters and adjustable weight vectors of synapses (Pouladi et al. 2019). More details about MLP can be found in Kisi et al. (2015).

2-4-3- Radial basis function neural network (RBFNN)

The RBFNN model is a particular type of ANN model that has been widely used for modelling hydrological variables such as streamflow, runoff, temperature, drought, and groundwater level. The main difference between RBFNN and feedforward multilayer ANN is the transfer function properties employed in the hidden layers (Walczak and Massart 2000). RBFNN contains three

layers in its structure. The first layer receives the input vector data, where the hidden layers include several nodes and a nonlinear transfer function (Tayebi et al. 2019). Then, the Euclidean norm (distance) associated with the hidden layer's input vector and center is computed by the activation function related to each hidden neuron. Commonly, the hidden layers contain the nonlinear radial basis Gaussian function and another activation function placed in the last layer. More details about RBFNN can be found in Kisi et al. (2015).

2-4-4- Support vector machine (SVM)

SVM was initially introduced by Cortes and Vapnik (1995) and has been widely used for both regression and classification analysis due to its advantage in minimizing model complexity and estimation error simultaneously (Zheng et al. 2020). The SVM model uses different kernel functions to estimate the regressions, implicitly converting inputs to high-dimensional feature space using a hyper-plane (Xing et al. 2018). The SVM equation is:

$$f(X_i) = \omega^T . \varphi(X_i) + b \tag{22}$$

where $f(X_i)$ is a deterministic function, $\varphi(x_i)$ is a nonlinear transferring the input vector, ω is a vector of weight coefficients, and b is the bias.

The b and ω values are determined by minimizing risk structure-function. To reduce the complexity and obtain a more robust model, slack variables ξ_i, ξ_{i^*} can be used in the risk structure-function equation:

$$Minimize \left(\frac{1}{2}\right) \|\omega\|^2 + C_0 \sum_{i=1}^{M} (\xi_i + \xi_{i^*}) \tag{23}$$

$$subject\ to = \begin{bmatrix} y_i - (\langle \omega, \psi(X_i) \rangle + b) \leq \varepsilon + \xi_i \\ (\langle \omega, \psi(X_i) \rangle + b) - y_i < \varepsilon + \xi_{i^*} \end{bmatrix}$$

where ε is intensive loss value, $y_i$is the i^{th} output, and $C_0$is the cost constant.

Lagrange multipliers (α_i and α_{i^*}) can be applied for solving eq. 23, with the solution written as follows using kernel function ($K(X_i, X)$):

$$f(X_i) = \sum_{i=1}^{M} (\alpha_i - \alpha_{i^*})\, K(X_i, X) + b \tag{24}$$

The radial basis function ($K(X_i, X) = exp(-\gamma \|X_i - X\|^2)$) is widely used for SVM models.

In this study, four optimization algorithms of SFO, FFA, PSO, and SSA are applied to integrate with standalone machine learning models and enhance hourly T_s prediction accuracy.

2-4-5- Sunflower optimization algorithm (SFO)

SFO is a new optimization algorithm proposed by Yang (2012) inspired by the solar tracking of sunflower heads to enhance pollination. SFO uses inverse square law radiation (ISLR) for optimization, based on the simple assumption that each flower only generates one gamete of pollen and reproduces individually (Qais et al. 2019). When the sunlight falls more obliquely, the sunflower receives less heat from the sun, expressed by the ISLR as:

$$Q_i = \frac{P}{4\pi r_i^2} \tag{25}$$

where P is the power of source and r_i is the distance between the current and the best i plant.

The redirection of position by the sunflower is computed as:

$$d_i = \lambda \times P_i(\|X_i + X_{i-1}\|) \times \|X_i + X_{i-1}\| \tag{26}$$

where λ is the inertial displacement of the sunflower and is constant, and X_i is the current location of the plant, and $P_i(\|X_i + X_{i-1}\|)$ is the pollution probability, where i^{th} sunflower cross-pollinates with another near $i\text{-}1^{th}$ sunflower to generate a new individual in an updated location.

It is essential to limit the maximum step of individuals by:

$$d_{\max} = \frac{\left\| X_{\max} - X_{\min} \right\|}{2 \times N_{pop}} \quad (27)$$

where d_{max} is the maximum step, X_{max} is the upper bound value, X_{min} is the lower bound value, and N_{pop} is the number of sunflowers in the overall population.

Finally, the next population is updated as:

$$\vec{X}_{i+1} = \vec{X}_i + d_i \times \vec{s}_i \quad (28)$$

$$\vec{s}_i = \frac{X^* - X_i}{\|X^* - X_i\|} \quad (29)$$

where $\vec{X}_{i+1}$is the position of sunflower $i+1$ and $X^*$is the best location of the sunflower.

The principal steps of the SFO model can be written as:

(1) initialize algorithm parameter of population size. The initial location of the sunflower is initialized in the initial matrix of population. Each location shows the initial value of sunflower.

(2) evaluate the objective function for solutions. The sun's best solution is where the sun guides the solutions towards the best location during the optimization process. The other solutions modify their orientation towards the sun.

(3) redirect positions using eq. 25.

(4) update the position of each search solution using eq. 29.

(5) investigate convergence criterion and finally, determine the optimal value.

2-4-6- Firefly algorithm (FFA)

Yang (2010) first developed FFA, inspired by the light emission capability of fireflies. The attractiveness of one firefly to another is related to its brightness to neighboring fireflies, where a less bright firefly is attracted to a brighter one (Alor et al. 2019; Riahi-Madvar et al. 2020). The

excellent information-sharing mechanism is one of the advantages of FFA. The firefly mechanism is formulated as:

$$I(r_{ij}) = I_0 e^{-\gamma r_{ij}^2} \tag{30}$$

$$\beta(r_{ij}) = \beta_0 e^{-\gamma r_{ij}^2} \tag{31}$$

where $\beta(r_{ij})$ is the attractiveness of a firefly, β_0 is the attractiveness of firefly at $r=0$, γ is the absorption coefficient in the range of 0 and 1, $I(r_{ij})$ is the light intensity, I_0 is the light intensity at $r=0$, and the distance between two fireflies i, j at locations of x_i and x_j can be defined as $r_{ij} = \|x_i - x_j\|$. The firefly i is attracted by firefly j and updates its location as:

$$\Delta x_i = \beta_0 e^{-\gamma r^2}\left(x_j^t - x_i^t\right) + \varsigma\mu \tag{32}$$

$$x_i^{t+1} = x_i^t + \Delta x_i$$

where ς is the randomization parameter, and μ is a vector of random parameters.

The implementation steps of the FFA can be described as follows:

(1) generate the initial population of fireflies to give the initial location.

(2) evaluate the objective function to identify brighter firefly.

(3) calculate the attractiveness using eq. 31.

(4) update the location of firefly (eq. 32) by moving a firefly i towards other brighter fireflies.

(4) move a firefly i towards other brighter fireflies, the position is updated by Equation.

(5) evaluate solutions and update brightness.

(6) terminate the optimization process if the stop criterion is met.

2-4-7- Salp swarm algorithm (SSA)

SSA is a new meta-heuristic optimization algorithm inspired by the collective behavior of salps (sea squirts), introduced by Mirjalili et al. (2017). Adaptability, robustness, and scalability are the most important advantages of SSA. Based on the individual's location in the chain, salps are divided into leaders or followers in SSA. The followers follow the leader to guide them in their movements. The SSA starts by initializing the salp population. Then, the leader position is updated as:

$$x_i^l = \begin{bmatrix} y_i + r_1\left((ub_i - lb_i)r_2 + lb_i\right) \leftarrow r_3 \geq 0 \\ y_i + r_1\left((ub_i - lb_i)r_2 + lb_i\right) \leftarrow r_3 < 0 \end{bmatrix} \quad (33)$$

where x_i^l is the leader's position, y_i is the food source, lb_i and ub_i are the lower and upper bounds, respectively, r_3 and r_2 are random numbers, and r_1 is computed as:

$$r_1 = 2e^{-\left(\frac{4l}{L}\right)} \quad (34)$$

where l is the existing iteration, and L is the maximum number of iterations. SSA uses the parameter r_1 to increase stability in exploration and exploitation capability. For each follower, the position is updated as:

$$x_j^i = \frac{1}{2}\left(x_j^i + x_j^{i-1}\right) \quad (35)$$

where x_j^i is the location of j^{th} salp in the i^{th} dimension.

The implementation steps of the SSA can be described as follows

(1) initialize a population of salps to give the initial location of salps

(2) compute the objective solution for each solution

(3) update the location of a best salp (leader)

(4) update the location of followers by eq. 35.

(5) terminate the optimization process to reach the best values of decision variables.

2-4-8- Particle swarm optimization (PSO)

PSO is a well-established stochastic/random search approach related to the swarm and inspired by the particle's social behavior. Each particle in the swarm protects the updating of search behavior according to all other particles' learning experiences. In each generation, information is integrated by particles to set the velocity on every dimension. The position and velocity of particles are updated as:

$$v_{i,t+1} = \eta * v_{i,t} + c_1 * rand * (p_{i,t} - x_{i,t}) + c_2 * rand * (p_{g,t} - x_{i,t}) \quad (36)$$

$$x_{i,t+1} = x_{i,t} + v_{i,t+1} \quad (37)$$

where $v_{i,t+1}$ is the velocity of i^{th} particle at iteration $t+1$, η is the weight inertia, c_1 and c_2 are acceleration coefficients, $x_{i,t+1}$ is the location of i^{th} particle at iteration $t+1$, $p_{i,t}$ is the optimal location experienced by particles, and $p_{g,t}$ is the optimal location experienced by any particle.

The principal steps of the PSO model can be written as:

(1) initialize initial location and velocity of particles

(2) compute objective function during each iteration

(3) update the velocity of particles

(4) update the location of particles by eq. 37.

(5) continue the optimization process until the stop criterion is met.

2-4-9- Hybridizing ANFIS, MLP, RBFNN, and SVM with optimization algorithms

In this study, SFO, FFA, SSA, and PSO were used to fine-tune the parameters of the ANFIS, MLP, RBFNN, and SVM models and to improve the convergence rate. By applying GT, the best input

combination was selected for both stations and the training of models was performed using the selected best input combination.

The input data were trained using optimization algorithms with a random selection of agents (particles, fireflies, salps, and sunflowers) and determined each individual's position. The position of agents shows the values of the premise (a_i, b_i, c_i) and consequent (p_1, q_1, r_1, p_2, q_2, r_2) parameters in the ANFIS model.

The values of bias and weight connections for the MLP model; the width values and the hidden neuron center in the RBFNN model; and the values of width (γ), penalty (C), epsilon (ε), and kernel function parameters in the SVM model. The algorithms used their operators to update the position of each agent.

The initial values of the parameters were regarded as the initial positions of agents. The objective function of root mean square error (RMSE) was applied to verify model accuracy. The hybrid ANFIS, MLP, RBFNN, and SVM models continued until the minimum value of RMSE was detected, and algorithms were converged toward the optimal solutions (Abualigah et al., 2021b). The optimal values of model parameters were found in each update of the agents' position.

The robust design of random parameters included in the optimization algorithms is important to enhance model performance. In a novel approach, Taguchi search was used to select the best values of these random parameters. It is a powerful advanced technique that uses orthogonal array and signal to noise ratio (S/N) to minimize the number of experiments and greatly decrease the time, cost, and effort in finding optimal parameters of algorithms (Zhang et al. 2015).

An orthogonal array table is created by calculating the total degree of freedom (DOF) based on the combined degree of freedom of all parameters (Canbolat et al. 2019; Bademlioglu et al. 2020). Based on the number of levels (L) and the number of parameters (NV), the total number of

experiments is computed. The minimum number of experiments (N) is calculated as $N = 1 + NV(L-1)$. Thus, e.g., for the four parameters with four levels in the PSO algorithm, at least 13 experiments should be conducted to discover the optimal values of PSO parameters, while without the Taguchi method, the total number of experiments would be 4^4. Thus, the Taguchi orthogonal array of L16 (4^3) was established.

The S/N ratio is calculated as:

$$\frac{S}{N} = -10log\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{y_i^2}\right) \quad (38)$$

where *n* is the number of the case, and *y* is the performance characteristic performance value (the higher the S/N value, the better the ratio). S/N was used as a performance characteristic of each parameter in this study.

The framework and flowchart of the work in this study are presented in Fig. 3, which also shows the general framework of the 16 hybrid models used to predict hourly T_s (ANFIS-SFO, ANFIS-FFA, ANFIS-SSA, ANFIS-PSO, MLP-SFO, MLP-FFA, MLP-SSA, MLP-PSO, RBFNN-SFO, RBFNN-FFA, RBFNN-SSA, RBFNN-PSO, SVM-SFO, SVM-FFA, SVM-SSA, and SVM-PSO). Finally, a schematic diagram representing the general procedures of the proposed method is given in Fig. 4.

2-6- Evaluation criteria

Coefficient of determination (R^2), RMSE, Nash-Sutcliffe efficiency (NSE), mean absolute error (MAE), and percentage bias (PBIAS) were used to assess the accuracy of the standalone and hybrid models (Moriasi et al. 2007; Moriasi et al., 2015):

$$R^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{(\sum_{i=1}^{n}(x_i - \bar{x})^2(\sum_{i=1}^{n}(y_i - \bar{y})^2))}} \qquad optimal\ value: 1 \tag{39}$$

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(x_i - y_i)^2}{N}} \qquad optimal\ value: 0 \tag{40}$$

$$MAE = \sum_{i=1}^{i=N}\frac{|x_i - y_i|}{N} \qquad optimal\ value: 0 \tag{41}$$

$$NSE = 1 - \frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N}(x_i - \bar{x}_i)} \qquad optimal\ value: 1 \tag{42}$$

$$PBIAS = \left[\frac{\sum_{i=1}^{N}(x_i - y_i) * 100}{\sum_{i=1}^{N}(x_i)}\right] \qquad optimal\ value: 0 \tag{43}$$

where x_i is the i^{th} observed value, y_i is the i^{th} predicted value, $\bar{x}$ and $\bar{y}$ are the mean value of observed and predicted values, respectively.

3- Results and Discussion

Wavelet coherence results for investigating the relationship and coherency between T_s and meteorological variables are summarized in section 3.1. Extracting the best values for optimization algorithm parameters in training the hybrid models using Taguchi search are presented in section 3.2. In section 3.3, model's performance for predicting T_s and the results obtained at different sites and soil depths are compared. Finally, the uncertainty associated with the simulations, determined using the GLUE approach, is presented in section 3.4.

3-1- Selection of the best input combination by Gamma Test

To evaluate the potential of developed models in predicting hourly T_s, the employed dataset is divided into the training and testing sets equal to 75% and 25% of total data, respectively. The GT

was used to construct and examine different input combinations by four meteorological variables of T_a, RH, U, and R_s. Table 2 shows the five optimal input combinations for any soil depth. The optimal input combination was selected using Γ and V_{ratio} indices.

As observed in Table 2, the input combination of mean air temperature, relative humidity, wind speed, and solar radiation had the least values of Γ and V_{ratio} at soil depth of 5 cm in Sirjan station. The results indicated that the input combination of mean air temperature, relative humidity, wind speed, and solar radiation was the best input scenario at depths 10 and 30 cm in Sanandaj station. Removing mean air temperature and solar radiation led to the high values of Γ and V_{ratio} at soil depths of 5, 10, and 30 cm in both stations. The results indicated that the Γ and V_{ratio} indices for four input combinations were lower than those of the other three input combinations.

3-2- Wavelet coherence analysis

To assess the importance of input vector selection in T_s simulation was considered using wavelet coherence analysis. The wavelet coherence between T_s and other climate variables (T_a, U, R_s, and RH) at 5, 10, and 30 soil depths is shown in Figures 5 and 6 for the period 0-1944 hours at 5 cm soil depth and 0-1000 hours at 10 and 30 cm soil depth.

Five dominant scales (periodicity) of <4 h, 4-8 h, 8-16 h, 16-32 h, and 32-64 h were considered following the coherencies between T_s and meteorological variables. There was high coherence between T_a and T_s at scale 4-8 hours in the time domain 486-709 hours at 5 cm soil depth at Sirjan (Fig. 5).

The coherence value between T_a and T_s at the mentioned scale and time was almost 0.8. Overall, T_a showed the highest correlation with T_s at 5 cm soil depth, confirming findings by Kisi et al. (2015) and Nanda et al. (2020). Analysis of the diagram indicated there was a few occasions with

high coherences between U and T_s, and thus there was no significant coherence between U and T_s at different time.

However, there was a significant relationship between T_s and R_s at scale of 28-34 hours, and between T_s and RH at scale of 28-34 hours, in the time domain 1252-1944 hours (Fig. 5). Overall, T_a, R_s, and RH showed variable coherence signals, while no significant coherence signal was observed at different time periodicities for U. Using WNN models, Samadianfard et al. (2018a) demonstrated that T_a and R_s have unquestionable effects on T_s prediction at 5, 10, 20, 30, 50, and 100 soil depth at a synoptic station in Iran.

To further analyze the results of wavelet coherence, the coherence values were plotted for different dominant scales of T_s. Fig. 6 shows the coherence at 10 and 30 cm soil depth for Sanandaj station. The maximum coherence between T_s at 10 and 30 cm soil depth and T_a was 0.8 and 0.7, respectively, at periodicity 8-16 hours in the time domain 0-1000 hours.

There were no significant coherences between T_s and U in different periodicities at 10 and 30 cm soil depth. The maximum coherence between T_s and R_s at periodicity 4-8 hours in the time domain 0-1000 hours was 0.9 and 0.7 at 10 and 30 cm depths. The highest coherence found equal to 0.9 at 10 cm soil depth between RH and T_s at periodicity 8-16 hours in the time domain 0-1000 hours. These results indicate that wavelet coherence analysis can provide more quantitative evidence on periodicity and interaction of T_s with meteorological variables at different soil depths.

3-3- Optimal model parameters derivation by Taguchi search

The optimization parameters of the meta-heuristic algorithms of SFO, SSA, FFA, and PSO at different levels, and the associated S/N values, are shown in Table 3. The values of population size, c_1, c_2, inertia weight, and γ in optimization algorithms were derived based on the Taguchi

method. After constructing an orthogonal array and calculating S/N ratio for each algorithm parameter, the optimal values were chosen (highest S/N ratio) (Table 3).

The effect of each optimization algorithm process parameter was divided into four levels. For the PSO algorithm, the mean value of RMSE for population size varied from 1.12 to 1.54, inertia weight from 1.23 to 1.76, c_1 from 1.12 to 1.45, and c_2 from 1.37 to 1.45, and the best conditions were obtained at a value of 2, 4, 1, and 1, respectively. The greatest variation in S/N values was seen for inertia weight, i.e. PSO was mainly affected by inertia weight, which should be the main focus in model development.

The most critical parameter for SSA, SFO, and FFA was r_3, population size, and γ, respectively. The optimal value of r_3, r_2, and population size for SSA, determined by Taguchi search, was 0.5, 0.8, and 400, respectively. For SFO, population size of 400 was most effective, while for FFA, γ had an optimal value of 0.8. These results indicate that Taguchi search can effectively and systematically provides robustness values in selection optimization parameters, replacing the trial-and-error approach. It can be used in future studies of meta-heuristics to improve the ability and applicability of models.

3-4- Performance analysis of models

The performance of standalone ANFIS, MLP, RBFNN, and SVM and the hybrid models was evaluated based on statistical criteria and by comparison the predictions against measured hourly T_s at 5, 10, and 30 cm soil depth (Tables 4 and 5). Soil temperature was used as the target parameter, and corresponding T_a, U, R_s, and RH as input data for all models. At the training level, integration the SFO with ANFIS, SVM, RBFNN, and MLP led to better results.

The standalone ANFIS model outperformed the other standalone SVM, RBFNN, and MLP models. All models provided the highest accuracy at 10 cm soil depth, and the lowest at 5 cm soil depth. However, the prediction accuracy in terms of RMSE, MAE, NSE, and PBIAS for the training level was lower for SVM-PSO than for the other hybrid ANFIS, SVM, MLP, and RBFNN models.

It should be noted that the measured data at 5 cm depth were from the arid Sirjan site and the measured data at 10 and 30 cm soil depths from the semi-humid Sanandaj site. Table 5 shows the performance of the models in predicting hourly T_s in terms of RMSE (°C), MAE (°C), PBIAS (%) (optimal value=0), and NSE (optimal value=1) in the testing phase. Scatter plots and correlation comparisons between simulated and measured T_s' values at 5, 10, and 30 cm soil depths are presented in Fig. 7, 8, and 9, respectively.

At 5 cm soil depth, ANFIS-SFO had the lowest error among all hybrid models (RMSE=1.18, MAE=1.05, PBIAS=7). ANFIS-SFO also had the lowest error among all models at 10 cm depth (RMSE=0.824, MAE=0.822, PBIAS=1, NSE=0.98) and at 30 cm depth (RMSE=0.911, MAE=0.905, PBIAS=2, NSE=0.97) (Table 5).

The ANFIS model was the most accurate between standalone models, with RMSE=0.978 ^{o}C, MAE= 0.971 ^{o}C, NSE=0.85, and PBIAS=14% in the best case (at 10 cm soil depth). The ANFIS model has been shown previously to be superior in T_s prediction. For example, Citakoglu (2017) reported that ANFIS was better (MAE=1.09, RMSE=1.99, R^2=0.98) than the ANN model in estimating hourly T_s when using monthly minimum temperature, monthly maximum temperature, calendar month number, soil depth, and monthly precipitation as input data.

As can be seen from Table 4, the SVM model had the highest error (highest MAE, RMSE, and PBIAS values, lowest NSE values) compared with all models at all soil depths. Nanda et al. (2020)

also found that SVM had the lowest precision compared to all machine learning models tested in hourly and half-hourly T_s prediction. The main reason for the low accuracy of SVM may be non-linearity between several analyzed parameters.

Comparison of the results at 10 and 30 cm depth indicated that the error in T_s prediction increased with depth, indicating that meteorological parameters mostly influenced soil temperature in the upper layers (Kazemi et al. 2018; Behmanesh and Mehdizadeh et al. 2017). Similarly, Nahvi et al. (2016) found that the precision of machine learning models in estimation of T_s in soils in Turkey and Iran declined from 10 cm to 100 cm depth but increased from 5 to 10 cm depth.

Behmanesh and Mehdizadeh (2017) also found that the accuracy of machine learning models increased between 5 and 10 cm depth. In the present study, the highest accuracy was observed at 10 cm depth for all hybrid and standalone models, as also found by Behmanesh and Mehdizadeh (2017) and Kisi et al. (2017). Among the models they studied, Kisi et al. (2017) found that ANFIS achieved the highest precision at 10 cm depth (RMSE=1.29) based on 25-year monthly dataset at 10, 50, and 100 cm depth.

Based on the results in Table 5, in this study the SFO algorithm increased the accuracy of standalone models by decreasing RMSE by 5.6% for ANFIS-SFO and 18.3% for RBFNN-SFO, both at 5 cm depth. The SSA algorithm decreased RMSE by 6.2% at 10 cm soil depth for MLP-SSA and 15% for RBFNN-SSA at 5 cm soil depth. The FFA algorithm decreased RMSE by 3.3% at 10 cm depth for MLP-FFA and 10.9% at 5 cm depth for SVM-FFA. Similarly, Samadianfard et al. (2018b) found that integrated MLP-FFA and SVM-FFA models had significantly higher T_s prediction accuracy than the standalone MLP and SVM models at 5, 10, and 20 cm soil depth.

Almost all predictions fell on the 1:1 line in the scatterplots, with R^2 values >0.99 for all models (Figs. 7-9). The R^2 value increased from 0.9958 to 0.9998, 0.9934 to 0.9993, and 0.9950 to 0.9998

for all hybrid and standalone models at 5, 10, and 30 cm depth. All models had the most accurate performance for 10 cm soil depth and the worst for 5 cm soil depth (Figs. 7-9).

This reflects the complexity and rapid changes in T_s at shallow depths. Overall, the ANFIS-SFO predictions were closest to the observed data, with the highest R^2 and lowest error at each soil depth, confirming the suitability of ANFIS-SFO for estimating hourly T_s in regions with arid (Sirjan) and semi-humid (Sanandaj) climates. Abyaneh et al. (2016) presented that the artificial neural network models and co-active neuro-fuzzy inference system (CANFIS) had the high abilities for predicting soil temperature.

Table 5 and Figures 7-9 show that the use of all meta-heuristic algorithms (SFO, SSA, FFA, PSO) improved the accuracy of the corresponding standalone models. This may be due to more accurate searching and finding the best solution in the local and global spaces. This is in agreement with findings by Mehdizade et al. (2020a), Riahi-Madvar et al. (2020), and Shamshirband et al. (2020) that hybrid models show higher accuracy than conventional models.

Samadianfard et al. (2018b) coupled MLP model with FFA method to estimate soil temperature. They showed that the hybrid MLP models performed better than the standalone MLP models. The error of MLP models at depth 20 cm was lower than those of MLP models at depth 50 cm. The current research results confirmed that the models had more accuracy at the depths 5 and 10 cm.

Based on results from the present study, the ANFIS model integrated with a meta-heuristic optimization algorithm is highly recommended for estimating T_s at different soil depths in different climates. The developed hybrid ANFIS and MLP models in the current study provided lower RMSE and MAE values than previous studies. The results of the present study at 30 cm soil depth were more accurate than the results of Alizamir et al. (2020).

In several prediction cases by conventional machine learning methods, the search domain become wider while the convergence rate decreases and then it may get trapped in local optimum. Therefore, optimizing the weights of conventional machine learning models (ANFIS, SVM, MLP, and RBFNN) is helpful to overcome this issue (Penghui et al. 2020).

Hence, the CPU time and number of functional evaluation (NFE) were obtained in this study to compare the convergence rate of models. Table 6 shows the values of CPU time and NFE to reach the best solution for developed models. As observed in Table 1, the CPU time of ANFIS, MLP, RBFNN, and SVM model hybridized with SFO was lower than those of other hybrid models. It means that the SFO could converge earlier than other optimization algorithms and conventional machine learning models.

The SSA, FFA, and PSO had the following ranks based on the CPU time. The ANFIS, MLP, and RBFNN models had the longest CPU time. It means that the standalone models of ANFIS, MLP, SVM, and RBFNN need a longer time for training. Table 6 reports the minimum NFE required to achieve the optimal solution. Hybrid models of ANFIS-SFO, MLP-SFO, RBFNN-SFO, SVM-SFO had the lowest values of NFE, indicating a faster convergence rate of the SFO algorithm to find the best solution. Thus, the SFO outperformed the other algorithms for converging and finding the optimal solution.

3-5- Uncertainty analysis of models using GLUE

Spectral representations of models uncertainty values in testing level for 5, 10, and 30 cm soil depth determined using the GLUE approach and the results are shown in Fig. 10. At 5 cm soil depth, the *r* value for the testing data for ANFIS-SFO ranged from 0.10 to 0.15 with 486-972 data

points and from 0.15 to 0.20 with 1458-1944 data points. The *p* value at 5 cm soil depth for ANFIS-SFO was found to be 0.95 with 486-972 data points.

For 10 and 30 cm soil depth, the *p* range for ANFIS-SFO was 0.90-0.94 and 0.92-0.96, respectively. The uncertainty analysis showed *p* values in the range 0.86-0.90 and 0.86-0.90 for MLP-SFO and RBFNN-SFO, respectively, with 600-1000 data points. ANFIS-PSO, MLP-PSO, RBFNN-PSO, and SVM-PSO had higher *p* and lowest *r* than other hybrid models. ANFIS-SFO and MLP-SFO had the highest bracketed observed values in the testing level, while the lowest unbracketed observations were obtained with the standalone SVM and RBFNN models.

This indicates that using meta-heuristic algorithms such as SFO adjusts specific parameters in AI models while reduces the case-specificity and improving the generality of the models. In the present case, hybridization of the traditional models, decreased soil temperature predictions and provided more accurate and reliable analysis for soil temperature in depths.

In the uncertainty analysis, an increase in data points increased the *r* values for the different models (Fig. 10). In general, SVM had the highest *r* and lowest *p*, indicating a high level of uncertainty. The standalone and hybrid MLP models produced more accurate predictions than the standalone and hybrid RBFNN models. Overall, the optimization algorithms improved the performance and reliability of the standalone models in terms of *r* and *p* values. SFO outperformed the other optimization algorithms in reducing uncertainty in the estimation of hourly T_s at all soil depths.

In general, the evaluation criteria values and GLUE uncertainty analysis of models confirmed that the hybrid ANFIS-SFO model had the most accurate performance in predicting hourly Ts based on the Ta, RH, U, and Rs semi-humid and arid climates of Iran. Thus, the results demonstrated the suitability of SFO for hybridizing machine learning models.

4- Conclusions

Soil temperature strongly affects soil biological processes, irrigation scheduling, plant growth, the environment, and water resource management, but measuring soil temperature is time-consuming and costly. Accurate, reliable, and stable soil temperature prediction using models is necessary in agricultural, environmental, and geosciences practices.

The scale-specific coherency between soil temperature at three soil depths of 5, 10, and 30 cm and meteorological variables of air temperature, relative humidity, wind speed, and solar radiation was investigated using wavelet coherence analysis.

The soil temperature time series data at three soil depths were mainly affected by air temperature, relative humidity, and solar radiation, especially in upper layers. In addition, coherence wavelet analysis indicated weak effects of wind speed on soil temperature.

The relationships between soil temperature and most effective variables of air temperature and relative humidity differed with scale (periodicity). High coherency with air temperature was generally observed at periodicity of 4-8 h at 5 cm and 8-16 h at 10 and 30 cm soil depths, while strong coherency with relative humidity obtained at periodicity of 16-64 h at 5 cm and 4-8 h at 10 and 30 cm soil depths. Therefore, soil temperature can be predicted with high accuracy at recommended periodicity as an application of information for scales.

In this study, four meta-heuristic optimization algorithms (SFO, FFA, PSO, SSA) were used to hybridize and optimize the standalone models of ANFIS, MLP, RBFNN, and SVM for predicting hourly soil temperature at various soil depths, based on meteorological data from an arid site (Sirjan) and a semi-humid site (Sanandaj) in Iran.

The uncertainty of conventional and hybrid models was evaluated using GLUE approach to investigate the stability of the prediction. Input combination of air temperature, relative humidity, wind speed, and solar radiation was trained by models to predict soil temperature.

Parameter selection of the meta-heuristic models was optimized by Taguchi search. Taguchi search has a high potential to overcome the excellent effort to find the optimal values of optimization parameters and considerably reduce the number of experiments.

ANFIS was the most accurate standalone model, while SVM was the least accurate for both sites and soil depths (5, 10, 30 cm). The optimization algorithms SFO and SSA were best in enhancing the performance of all standalone models. The performance criteria proved that the hybrid ANFIS-SFO model had a strong correlation with observed data and predicted soil temperature with low error at all soil depths.

Uncertainty results of GLUE approach indicated that the bracketed observed values of the best model, ANFIS-SFO, were in the range 94-96%. The wider uncertainty bounds were obtained by standalone models, indicating some instability in soil temperature prediction and the small accuracy.

The developed approach in this study proved reliable and accurate for hybrid meta-heuristic models. It applies in regions with similar climate conditions to the study sites. Future studies should consider the effects of climate scenarios on soil temperature for a future period. ANIFS, SVM, RBFNN, and MLP can also be hybridized with multi-objective optimization algorithms to determine the optimal values of hyperparameters and appropriate inputs to the models.

The grey models are one of the most important models for predicting hydrological variables. In the following studies, the abilities of grey models and decision tree models can be evaluated for soil temperature estimation.

Also, this study performed a comprehensive uncertainty analysis to evaluate the accuracy of standalone and hybrid machine learning models that can be used to evaluate other estimation models.

Ethics declarations

Ethical approval

Not applicable.

Competing interests

The authors declare that they have no conflict of interest.

Research involving human participants and/or animals

Not applicable.

Consent to participate

All authors have given consent to their contribution.

Consent to publish

All authors have agreed with the content and all have given explicit consent to publish.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Code availability

The code is available from the corresponding author on reasonable request.

Funding

Not applicable.

Contributions

Conceptualization: Akram Seifi, Mohammad Ehteram; Methodology: Akram Seifi, Mohammad Ehteram, Formal analysis and investigation: Akram Seifi, Mohammad Ehteram, Fatemeh Soroush, Fatemeh Nayebloei , Writing original draft preparation: Akram Seifi, Fatemeh Soroush, Fatemeh Nayebloei, Bahram Gharabaghi, Ali Torabi Haghighi

References

Abualigah, L. M. Q. (2019). *Feature selection and enhanced krill herd algorithm for text document clustering* (pp. 1-165). Berlin: Springer.

Abualigah, L., & Diabat, A. (2021). Advances in sine cosine algorithm: a comprehensive survey. *Artificial Intelligence Review*, 1-42.

Abualigah, L., Diabat, A., Mirjalili, S., Abd Elaziz, M., & Gandomi, A. H. (2021a). The arithmetic optimization algorithm. *Computer methods in applied mechanics and engineering*, *376*, 113609.

Abualigah, L., Yousri, D., Abd Elaziz, M., Ewees, A. A., Al-qaness, M. A., & Gandomi, A. H. (2021b). Aquila Optimizer: A novel meta-heuristic optimization Algorithm. *Computers & Industrial Engineering*, *157*, 107250.

Abyaneh, H. Z., Varkeshi, M. B., Golmohammadi, G., & Mohammadi, K. (2016). Soil temperature estimation using an artificial neural network and co-active neuro-fuzzy inference system in two different climates. *Arabian Journal of Geosciences*, *9*(5), 377.

Alizamir, M., Kim, S., Zounemat-Kermani, M., Heddam, S., Shahrabadi, A. H., & Gharabaghi, B. (2020). Modelling daily soil temperature by hydro-meteorological data at different depths using a novel data-intelligence model: deep echo state network model. *Artificial Intelligence Review*, 1-28. DOI: 10.1007/s10462-020-09915-5.

Alor, A., Mota, D., Olmos-Sánchez, K., & Rodas-Osollo, J. (2019). An Order-Picking Model Associated With Hospital Components and Solved by a Firefly Algorithm. In Handbook of Research on Metaheuristics for Order Picking Optimization in Warehouses to Smart Cities (pp. 173-188). IGI Global.

Araghi, A., Mousavi-Baygi, M., Adamowski, J., Martinez, C., & van der Ploeg, M. (2017). Forecasting soil temperature based on surface air temperature using a wavelet artificial neural network. *Meteorological Applications*, *24*(4), 603-611.

Azadi, S., Amiri, H., Ataei, P., & Javadpour, S. (2020). Optimal design of groundwater monitoring networks using gamma test theory. *Hydrogeology Journal*, 1-14.

Bademlioglu, A. H., Canbolat, A. S., & Kaynakli, O. (2020). Multi-objective optimization of parameters affecting Organic Rankine Cycle performance characteristics with Taguchi-Grey Relational Analysis. *Renewable and Sustainable Energy Reviews*, *117*, 109483.

Behmanesh, J., & Mehdizadeh, S. (2017). Estimation of soil temperature using gene expression programming and artificial neural networks in a semiarid region. *Environmental Earth Sciences*, *76*(2), 76.

Bonakdari, H., Moeeni, H., Ebtehaj, I., Zeynoddin, M., Mahoammadian, A., & Gharabaghi, B. (2019). New insights into soil temperature time series modeling: linear or

nonlinear?. *Theoretical and Applied Climatology*, *135*(3-4), 1157-1177. DOI: 10.1007/s00704-018-2436-2.

Canbolat, A. S., Bademlioglu, A. H., Arslanoglu, N., & Kaynakli, O. (2019). Performance optimization of absorption refrigeration systems using Taguchi, ANOVA and Grey Relational Analysis methods. *Journal of Cleaner Production*, *229*, 874-885.

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

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

Damos, P., & Caballero, P. (2021). Detecting seasonal transient correlations between populations of the West Nile Virus vector Culex sp. and temperatures with wavelet coherence analysis. *Ecological Informatics*, *61*, 101216.

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.

Ebrahimi, A., Rahimi, D., Joghataei, M., & Movahedi, S. (2021). Correlation Wavelet Analysis for Linkage between Winter Precipitation and Three Oceanic Sources in Iran. *Environmental Processes*, 1-19.

Feng, Y., Cui, N., Hao, W., Gao, L., & Gong, D. (2019). Estimation of soil temperature from meteorological data using different machine learning models. *Geoderma*, *338*, 67-77.

Furon, A. C., Wagner-Riddle, C., Smith, C. R., & Warland, J. S. (2008). Wavelet analysis of wintertime and spring thaw CO2 and N2O fluxes from agricultural fields. *Agricultural and forest meteorology*, *148*(8-9), 1305-1317.

Han, D., Yan, W., & Nia, A. M. (2010, July). Uncertainty with the Gamma Test for model input data selection. In *The 2010 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-5). IEEE.

Heddam, S. (2019). Development of air–soil temperature model using computational intelligence paradigms: artificial neural network versus multiple linear regression. *Modeling Earth Systems and Environment*, *5*(3), 747-751.

Kazemi, S. M. R., Minaei Bidgoli, B., Shamshirband, S., Karimi, S. M., Ghorbani, M. A., Chau, K. W., & Kazem Pour, R. (2018). Novel genetic-based negative correlation learning for estimating soil temperature. *Engineering Applications of Computational Fluid Mechanics*, *12*(1), 506-516.

Kisi, O., Sanikhani, H., & Cobaner, M. (2017). Soil temperature modeling at different depths using neuro-fuzzy, neural network, and genetic programming techniques. *Theoretical and Applied Climatology*, *129*(3-4), 833-848.

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.

Lee, E., & Kim, S. (2019). Wavelet analysis of soil moisture measurements for hillslope hydrological processes. *Journal of Hydrology*, *575*, 82-93.

Lehnert, M. (2014). Factors affecting soil temperature as limits of spatial interpretation and simulation of soil temperature. *Acta Universitatis Palackianae Olomucensis–Geographica*, *45*(1), 5-21.

Li, Q., Hao, H., Zhao, Y., Geng, Q., Liu, G., Zhang, Y., & Yu, F. (2020). GANs-LSTM model for soil temperature estimation from meteorological: A new approach. *IEEE Access*, *8*, 59427-59443.

Li, X., Xu, X., Liu, W., He, L., Xu, C., Zhang, R., ... & Wang, K. (2019). Revealing the scale-specific influence of meteorological controls on soil water content in a karst depression using wavelet coherency. *Agriculture, Ecosystems & Environment*, *279*, 89-99.

Mehdizadeh, S., Behmanesh, J., & Khalili, K. (2017). Evaluating the performance of artificial intelligence methods for estimation of monthly mean soil temperature without using meteorological data. *Environmental Earth Sciences*, *76*(8), 325.

Mehdizadeh, S., Fathian, F., Safari, M. J. S., & Khosravi, A. (2020a). Developing novel hybrid models for estimation of daily soil temperature at various depths. *Soil and Tillage Research*, *197*, 104513.

Mehdizadeh, S., Mohammadi, B., Pham, Q. B., Khoi, D. N., & Linh, N. T. T. (2020b). Implementing novel hybrid models to improve indirect measurement of the daily soil temperature: Elman neural network coupled with gravitational search algorithm and ant colony optimization. *Measurement*, *165*, 108127.

Mirjalili, S., Gandomi, A. H., Mirjalili, S. Z., Saremi, S., Faris, H., & Mirjalili, S. M. (2017). Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Advances in Engineering Software*, *114*, 163-191.

Moazenzadeh, R., & Mohammadi, B. (2019). Assessment of bio-inspired metaheuristic optimisation algorithms for estimating soil temperature. *Geoderma*, *353*, 152-171.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885-900.

Moriasi, D. N., Gitau, M. W., Pai, N., & Daggupati, P. (2015). Hydrologic and water quality models: Performance measures and evaluation criteria. *Transactions of the ASABE*, *58*(6), 1763-1785.

Nahvi, B., Habibi, J., Mohammadi, K., Shamshirband, S., & Al Razgan, O. S. (2016). Using self-adaptive evolutionary algorithm to improve the performance of an extreme learning machine for estimating soil temperature. *Computers and Electronics in Agriculture*, *124*, 150-160.

Najafi-Ghiri, M., Mokarram, M., & Owliaie, H. R. (2019). Prediction of soil clay minerals from some soil properties with use of feature selection algorithm and ANFIS methods. *Soil Research*, *57*(7), 788-796.

Nanda, A., Sen, S., Sharma, A. N., & Sudheer, K. P. (2020). Soil Temperature Dynamics at Hillslope Scale—Field Observation and Machine Learning-Based Approach. *Water*, *12*(3), 713.

Padarian, J., Minasny, B., & McBratney, A. B. (2020). Machine learning and soil sciences: A review aided by machine learning tools. *Soil*, *6*(1), 35-52.

Panahi, F., Ehteram, M., & Emami, M. (2021). Suspended sediment load prediction based on soft computing models and Black Widow Optimization Algorithm using an enhanced gamma test. *Environmental Science and Pollution Research*, 1-21.

Penghui, L., Ewees, A. A., Beyaztas, B. H., Qi, C., Salih, S. Q., Al-Ansari, N., ... & Singh, V. P. (2020). Metaheuristic optimization algorithms hybridized with artificial intelligence model for soil temperature prediction: Novel model. *IEEE Access*, *8*, 51884-51904.

Plauborg, F. (2002). Simple model for 10 cm soil temperature in different soils with short grass. *European Journal of Agronomy*, *17*(3), 173-179.

Pouladi, N., Jafarzadeh, A. A., Shahbazi, F., & Ghorbani, M. A. (2019). Design and implementation of a hybrid MLP-FFA model for soil salinity prediction. Environmental earth sciences, 78(5), 159.

Qais, M. H., Hasanien, H. M., & Alghuwainem, S. (2019). Identification of electrical parameters for three-diode photovoltaic model using analytical and sunflower optimization algorithm. *Applied Energy*, *250*, 109-117.

Qasem, S. N., Ebtehaj, I., & Riahi Madavar, H. (2017). Optimizing ANFIS for sediment transport in open channels using different evolutionary algorithms. Journal of Applied Research in Water and Wastewater, 4(1), 290-298.

Qi, J., Li, S., Li, Q., Xing, Z., Bourque, C. P. A., & Meng, F. R. (2016). A new soil-temperature module for SWAT application in regions with seasonal snow cover. *Journal of Hydrology*, *538*, 863-877.

Riahi-Madvar, H., Dehghani, M., Parmar, K. S., Nabipour, N., & Shamshirband, S. (2020). Improvements in the Explicit Estimation of Pollutant Dispersion Coefficient in Rivers by Subset Selection of Maximum Dissimilarity Hybridized With ANFIS-Firefly Algorithm (FFA). IEEE Access, 8, 60314-60337.

Samadianfard, S., Asadi, E., Jarhan, S., Kazemi, H., Kheshtgar, S., Kisi, O., ... & Manaf, A. A. (2018a). Wavelet neural networks and gene expression programming models to predict short-term soil temperature at different depths. *Soil and Tillage Research*, *175*, 37-50.

Samadianfard, S., Ghorbani, M. A., & Mohammadi, B. (2018b). Forecasting soil temperature at multiple-depth with a hybrid artificial neural network model coupled-hybrid firefly optimizer algorithm. *Information Processing in Agriculture*, *5*(4), 465-476.

Sanikhani, H., Deo, R. C., Yaseen, Z. M., Eray, O., & Kisi, O. (2018). Non-tuned data intelligent model for soil temperature estimation: A new approach. *Geoderma*, *330*, 52-64.

Seifi, A., & Riahi, H. (2020). Estimating daily reference evapotranspiration using hybrid gamma test-least square support vector machine, gamma test-ANN, and gamma test-ANFIS models in an arid area of Iran. *Journal of Water and Climate Change*, *11*(1), 217-240.

Seifi, A., & Soroush, F. (2020). Pan evaporation estimation and derivation of explicit optimized equations by novel hybrid meta-heuristic ANN based methods in different climates of Iran. Computers and Electronics in Agriculture, 173, 105418.

Seifi, A., Ehteram, M., & Dehghani, M. (2021). A robust integrated Bayesian multi-model uncertainty estimation framework (IBMUEF) for quantifying the uncertainty of hybrid meta-heuristic in global horizontal irradiation predictions. *Energy Conversion and Management*, *241*, 114292.

Seifi, A., Ehteram, M., & Soroush, F. (2020a). Uncertainties of instantaneous influent flow predictions by intelligence models hybridized with multi-objective shark smell optimization algorithm. Journal of Hydrology, 124977.

Seifi, A., Ehteram, M., Singh, V. P., & Mosavi, A. (2020b). Modeling and Uncertainty Analysis of Groundwater Level Using Six Evolutionary Optimization Algorithms Hybridized with ANFIS, SVM, and ANN. Sustainability, 12(10), 4023

Shamshirband, S., Esmaeilbeiki, F., Zarehaghi, D., Neyshabouri, M., Samadianfard, S., Ghorbani, M. A., ... & Chau, K. W. (2020). Comparative analysis of hybrid models of firefly optimization algorithm with support vector machines and multilayer perceptron for predicting soil

temperature at different depths. *Engineering Applications of Computational Fluid Mechanics*, *14*(1), 939-953.

Sharafati, A., Asadollah, S. B. H. S., & Neshat, A. (2020). A new artificial intelligence strategy for predicting the groundwater level over the Rafsanjan aquifer in Iran. *Journal of Hydrology*, *591*, 125468.

Sihag, P., Esmaeilbeiki, F., Singh, B., & Pandhiani, S. M. (2020). Model-based soil temperature estimation using climatic parameters: the case of Azerbaijan Province, Iran. *Geology, Ecology, and Landscapes*, *4*(3), 203-215.

Singh, A., Malik, A., Kumar, A., & Kisi, O. (2018). Rainfall-runoff modeling in hilly watershed using heuristic approaches with gamma test. *Arabian Journal of Geosciences*, *11*(11), 1-12.

Singh, V. K., Singh, B. P., Kisi, O., & Kushwaha, D. P. (2018). Spatial and multi-depth temporal soil temperature assessment by assimilating satellite imagery, artificial intelligence and regression based models in arid area. *Computers and electronics in agriculture*, *150*, 205-219.

Sofyan, S. E., Hu, E., Kotousov, A., & Riayatsyah, T. M. I. (2020). A new approach to modelling of seasonal soil temperature fluctuations and their impact on the performance of a shallow borehole heat exchanger. *Case Studies in Thermal Engineering*, *22*, 100781.

Stajkowski, S., Kumar, D., Samui, P., Bonakdari, H., & Gharabaghi, B. (2020). Genetic-algorithm-optimized sequential model for water temperature prediction. Sustainability, 12(13), 5374. https://doi.org/10.3390/su12135374.

Sun, M., Zhang, X., Huo, Z., Feng, S., Huang, G., & Mao, X. (2016). Uncertainty and sensitivity assessments of an agricultural–hydrological model (RZWQM2) using the GLUE method. *Journal of hydrology*, *534*, 19-30.

Tayebi, H. A., Ghanei, M., Aghajani, K., & Zohrevandi, M. (2019). Modeling of reactive orange 16 dye removal from aqueous media by mesoporous silica/crosslinked polymer hybrid using RBF, MLP and GMDH neural network models. *Journal of Molecular Structure*, *1178*, 514-523.

Walczak, B., & Massart, D. L. (2000). Local modelling with radial basis function networks. *Chemometrics and Intelligent Laboratory Systems*, *50*(2), 179-198.

Xing, L., Li, L., Gong, J., Ren, C., Liu, J., & Chen, H. (2018). Daily soil temperatures predictions for various climates in United States using data-driven model. *Energy*, *160*, 430-440.

Yang, X. S. (2010). Firefly algorithm, stochastic test functions and design optimisation. *International journal of bio-inspired computation*, *2*(2), 78-84.

Yang, X. S. (2012, September). Flower pollination algorithm for global optimization. In *International conference on unconventional computing and natural computation* (pp. 240-249). Springer, Berlin, Heidelberg.

Zeynoddin, M., Ebtehaj, I., & Bonakdari, H. (2020). Development of a linear based stochastic model for daily soil temperature prediction: One step forward to sustainable agriculture. *Computers and Electronics in Agriculture*, *176*, 105636.

Zhang, F. B., Wang, Z. L., & Yang, M. Y. (2015). Assessing the applicability of the Taguchi design method to an interrill erosion study. *Journal of Hydrology*, *521*, 65-73.

Zhao, X., Wang, C., Su, J., & Wang, J. (2019). Research and application based on the swarm intelligence algorithm and artificial intelligence for wind farm decision system. *Renewable energy*, *134*, 681-697.

Zheng, G., Zhang, W., Zhang, W., Zhou, H., & Yang, P. (2020). Neural network and support vector machine models for the prediction of the liquefaction-induced uplift displacement of tunnels. *Underground Space*.

Zolfaghari, H., Masoompour, J., Yeganefar, M., & Akbary, M. (2016). Studying spatial and temporal changes of aridity in Iran. *Arabian Journal of Geosciences*, *9*(5), 375.

Table 1: Statistical characteristics of climatic parameters and soil temperature at studied stations

Dataset	Factor	Sirjan station					Sanandaj station					
		T_a	R_s	U	RH	T_s @ 5 cm	T_a	R_s	U	RH	T_s @ 30 cm	T_s @ 10 cm
All data	Min	-9.31	5	0.15	2	5.2	-10.1	4	0.19	5	5.9	7.4
	Max	25.05	1151	9.92	79	31.77	32.23	1132	7.85	69	24.69	29.12
	Mean	24.23	625	5.23	65.12	19.23	21.48	567.1	4.55	56	18.12	19.1
	Skewness	0.86	4.12	1.12	2.43	2.46	2.45	3.45	1.1	2.4	1.78	2.1
Train set	Min	-7.12	7	0.14	4	6.7	-9.98	6	0.86	7	6.12	8.12
	Max	24.87	1146	8.45	65	29.87	30.23	985	8.22	65	24.01	28.12
	Mean	26.12	678	6.12	56.25	18.23	22.12	567.1	5.12	45	17.65	17.99
	Skewness	1.32	1.67	1.56	2.14	2.01	4.55	3.11	3.12	2.2	1.45	1.23
Test set	Min	-6.12	4	0.15	2	6.8	-8.14	4	0.98	6	5.14	7.12
	Max	24.78	1056	8.23	69	28.78	29.87	987	6.72	59	24.42	28.12
	Mean	29.78	614	6.45	58.12	16.54	24.56	589.1	4.23	44	16.56	18.12
	Skewness	1.25	1.28	1.45	2.44	2.12	5.56	3.22	2.44	2.8	1.43	2.12

Table 2: Optimal input combinations obtained using the GT approach.

Input combination	5 cm		10 cm		30 cm	
	Γ	V_{ratio}	Γ	V_{ratio}	Γ	V_{ratio}
T_a, RH, U, R_s	0.02345	0.0033	0.065	0.0083	0.0551	0.007
T_a, U, R_s	0.03567	0.0051	0.078	0.0100	0.0671	0.008
T_a, RH, R_s	0.0672	0.0095	0.082	0.0105	0.072	0.009
T_a, RH, U	0.0721	0.01037	0.089	0.0114	0.084	0.0106
RH, U, R_s	0.082	0.01123	0.093	0.0119	0.0923	0.0112

Table 3: Optimal value of random parameters optimization algorithms

Optimization algorithm	Level	1	2	3	4
	Parameters				
PSO	Population size (S/N)	100 (1.12)	**200 (1.54)**	300 (1.34)	400 (1.46)
	Inertia weight (S/N)	0.30 (1.23)	0.50 (1.32)	0.70 (1.50)	**0.90 (1.76)**
	c_1 (S/N)	**1.40 (1.45)**	1.60 (1.30)	1.80 (1.28)	2 (1.12)
	c_2 (S/N)	**1.40 (1.45)**	1.60 (1.39)	1.80 (1.42)	2 (1.37)
SSA	Population size (S/N)	100 (1.16)	200 (1.14)	300 (1.19)	**400 (1.21)**
	r_2 (S/N)	0.5 (1.19)	0.6 (1.12)	0.7 (1.15)	**0.8 (1.21)**
	r_3 (S/N)	**0.5 (1.22)**	0.6 (1.12)	0.7 (1.17)	0.8 (1.21)
SFO	Population size (S/N)	100 (1.02)	200 (1.07)	300 (1.12)	**400 (1.14)**
FFA	Population size (S/N)	**100 (1.26)**	200 (1.22)	300 (1.21)	400 (1.26)
	γ	0.50 (1.27)	0.6 (1.29)	0.70 (1.32)	**0.80 (1.36)**

Table 4: Statistical results of models performance in training phase.

Model	5 cm				10 cm				30 cm			
	RMSE (°C)	MAE (°C)	NSE	PBIAS (%)	RMSE (°C)	MAE (°C)	NSE	PBIAS (%)	RMSE (°C)	MAE (°C)	NSE	PBIAS (%)
ANFIS-SFO	1.12	1.01	0.95	5	0.95	0.91	0.97	2	0.98	0.97	0.97	4
MLP-SFO	1.20	1.10	0.93	7	1.02	1.00	0.95	3	1.09	1.06	0.95	5
RBFNN-SFO	1.23	1.16	0.90	10	1.12	1.03	0.93	5	1.14	1.10	0.93	8
SVM-SFO	1.32	1.19	0.89	12	1.14	1.08	0.91	7	1.17	1.12	0.91	9
ANFIS-SSA	1.14	1.10	0.94	14	1.02	1.00	0.96	3	1.08	1.07	0.96	10
MLP- SSA	1.25	1.18	0.91	16	1.04	1.02	0.94	4	1.10	1.09	0.94	6
RBFNN- SSA	1.27	1.20	0.89	18	1.14	1.10	0.93	8	1.18	1.14	0.93	9
SVM- SSA	1.39	1.37	0.88	21	1.19	1.17	0.91	10	1.23	1.21	0.91	11
ANFIS-FFA	1.17	1.14	0.92	15	1.05	1.03	0.95	4	1.07	1.03	0.95	7
MLP-FFA	1.35	1.20	0.90	18	1.09	1.08	0.93	5	1.14	1.12	0.93	11
RBFNN-FFA	1.42	1.25	0.89	20	1.15	1.12	0.92	12	1.19	1.18	0.92	14
SVM-FFA	1.45	1.39	0.87	22	1.20	1.18	0.90	14	1.24	1.22	0.90	15
ANFIS-PSO	1.18	1.7	0.90	21	1.10	1.08	0.94	9	1.12	1.14	0.94	12
MLP-PSO	1.36	1.32	0.88	23	1.12	1.10	0.92	12	1.18	1.19	0.92	15
RBFNN-PSO	1.45	1.40	0.86	25	1.16	1.14	0.90	15	1.19	1.14	0.90	17
SVM- PSO	1.48	1.42	0.84	27	1.19	1.17	0.89	17	1.25	1.18	0.89	19
ANFIS	1.23	1.20	0.86	29	1.19	1.18	0.88	12	1.16	1.19	0.88	17

MLP	1.42	1.39	0.85	31	1.21	1.19	0.87	14	1.23	1.20	0.87	19
RBFNN	1.51	1.41	0.83	33	1.29	1.23	0.86	19	1.26	1.24	0.86	22
SVM	1.62	1.55	0.81	35	1.32	1.30	0.85	20	1.29	1.28	0.85	25

Table 5: Statistical results of models performance in testing phase.

Model	5 cm				10 cm				30 cm			
	RMSE (°C)	MAE (°C)	NSE	PBIAS (%)	RMSE (°C)	MAE (°C)	NSE	PBIAS (%)	RMSE (°C)	MAE (°C)	NSE	PBIAS (%)
ANFIS-SFO	1.18	1.05	0.93	7	0.824	0.822	0.98	1	0.911	0.905	0.97	2
MLP-SFO	1.23	1.12	0.92	9	0.835	0.832	0.97	3	0.914	0.909	0.96	4
RBFNN-SFO	1.25	1.18	0.89	11	0.845	0.843	0.95	5	0.915	0.912	0.94	6
SVM-SFO	1.35	1.20	0.87	14	0.856	0.854	0.94	6	0.917	0.914	0.93	7
ANFIS-SSA	1.17	1.12	0.92	16	0.832	0.828	0.97	2	0.914	0.911	0.96	9
MLP- SSA	1.29	1.19	0.90	18	0.839	0.837	0.95	5	0.919	0.918	0.94	10
RBFNN- SSA	1.30	1.22	0.86	20	0.849	0.842	0.94	8	0.933	0.932	0.91	12
SVM- SSA	1.42	1.39	0.85	22	0.860	0.857	0.93	9	0.935	0.931	0.90	14
ANFIS-FFA	1.19	1.16	0.82	23	0.878	0.876	0.93	3	0.955	0.951	0.92	4
MLP-FFA	1.37	1.22	0.81	24	0.882	0.880	0.91	5	0.962	0.958	0.90	6
RBFNN-FFA	1.44	1.27	0.84	26	0.893	0.891	0.90	7	0.969	0.967	0.89	8
SVM-FFA	1.47	1.42	0.83	28	0.896	0.892	0.88	10	0.976	0.972	0.87	15
ANFIS-PSO	1.22	1.26	0.88	29	0.901	0.899	0.92	5	0.980	0.976	0.91	7
MLP-PSO	1.39	1.35	0.87	30	0.923	0.912	0.89	8	0.983	0.981	0.88	9
RBFNN-PSO	1.48	1.43	0.84	32	0.935	0.932	0.88	9	0.985	0.983	0.86	11
SVM- PSO	1.51	1.45	0.82	33	0.945	0.935	0.86	11	0.990	0.986	0.84	12
ANFIS	1.25	1.23	0.81	35	0.934	0.931	0.90	6	1.002	1.001	0.89	8
MLP	1.50	1.45	0.83	37	0.932	0.930	0.88	9	1.040	1.030	0.87	10
RBFNN	1.53	1.47	0.80	39	0.945	0.941	0.87	12	1.063	1.055	0.86	15
SVM	1.65	1.52	0.80	40	0.978	0.971	0.85	14	1.069	1.054	0.83	17

Table 6: Criteria for convergence rate of developed models

Model	CPU time (s)			Number of functional evaluation (NFE): Population size* number of iterations		
	5 cm	10 cm	30 cm	5 cm	10 cm	30 cm
ANFIS-SFO	208	212	214	12222	12341	12124
MLP-SFO	212	216	217	12453	12562	12455
RBFNN-SFO	224	218	223	12654	12673	12876
SVM-SFO	229	221	232	13786	13784	13457
ANFIS-SSA	230	224	234	14569	13812	15678
MLP- SSA	232	234	236	14987	14913	17899
RBFNN- SSA	234	237	238	15123	15454	18244
SVM- SSA	246	241	239	16223	15675	19215
ANFIS-FFA	267	243	240	17124	15926	20202
MLP-FFA	312	245	242	17225	16222	21003
RBFNN-FFA	314	267	243	17346	17223	22002
SVM-FFA	324	278	245	17678	18224	24002
ANFIS-PSO	328	289	246	17567	19225	26002
MLP-PSO	332	310	247	18223	20123	28001
RBFNN-PSO	335	312	251	18344	20224	32002
SVM- PSO	337	321	253	18565	20263	34003
ANFIS	338	324	255	19227	20352	35004
MLP	339	325	256	20128	20591	38005
RBFNN	340	326	267	22129	22212	40006
SVM	345	329	267	22243	22453	42007

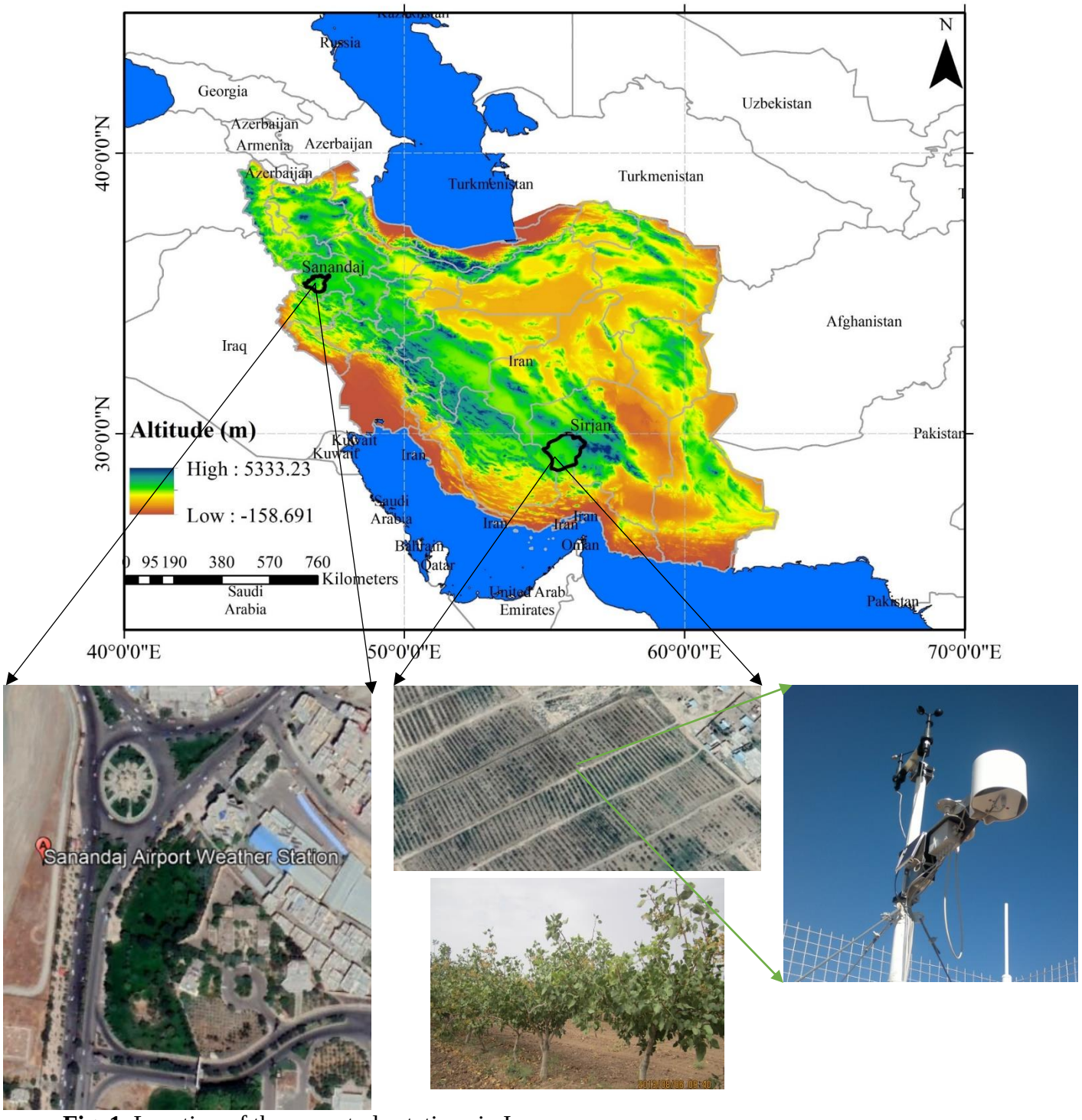
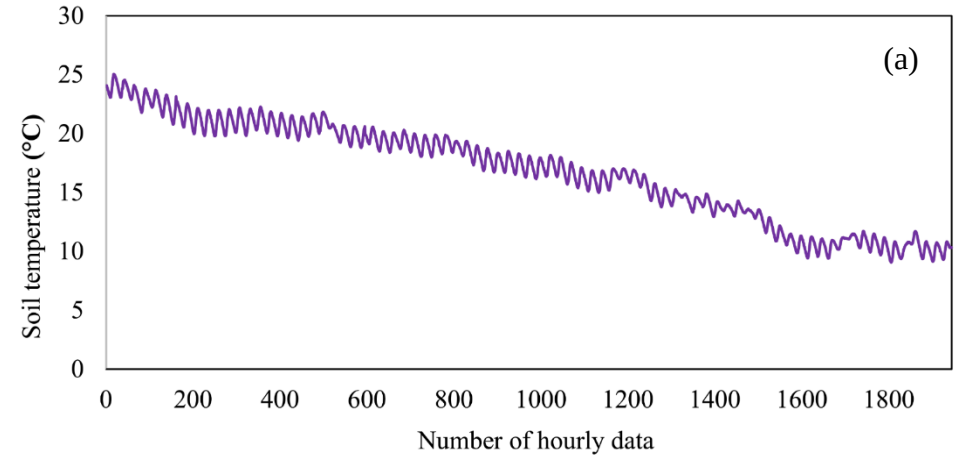
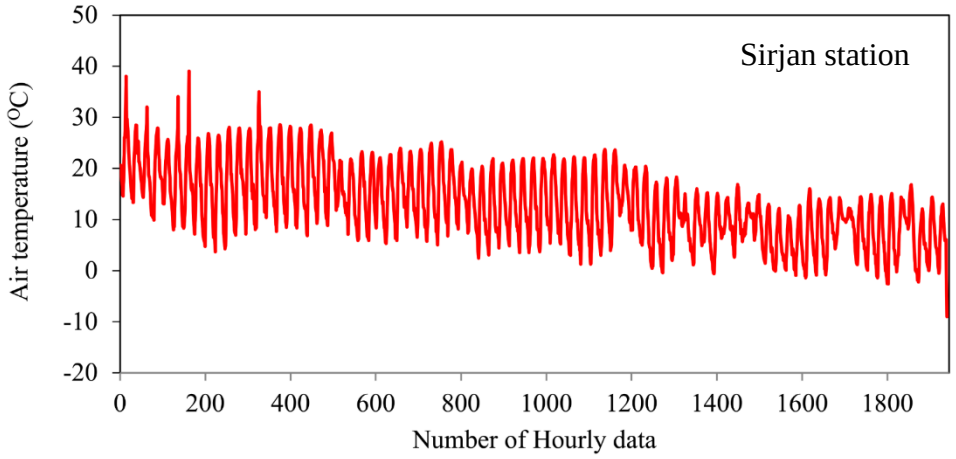
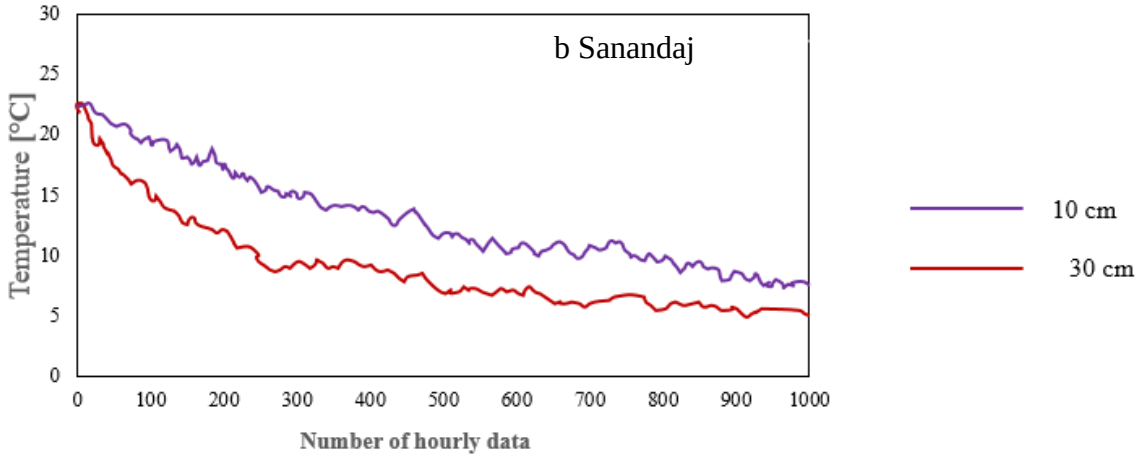


Fig. 1. Location of the case study stations in Iran


(a)
Soil temperature (°C)
Number of hourly data


Sirjan station
Air temperature (°C)
Number of Hourly data


b Sanandaj
Temperature [°C]
Number of hourly data
10 cm
30 cm

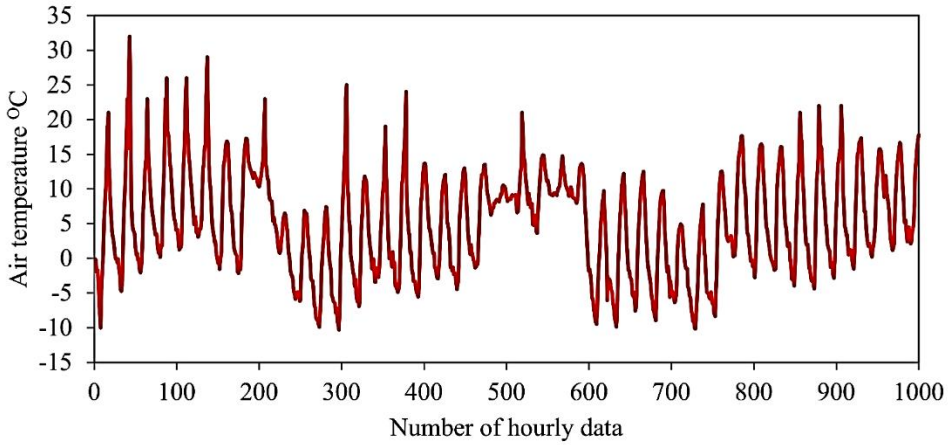


Fig. 2 Time series of (1): soil temperature at (a): 5 cm depth for Sirjan and (b): 10 and 30 cm depths for Sanandaj stations, and (2) air temperature

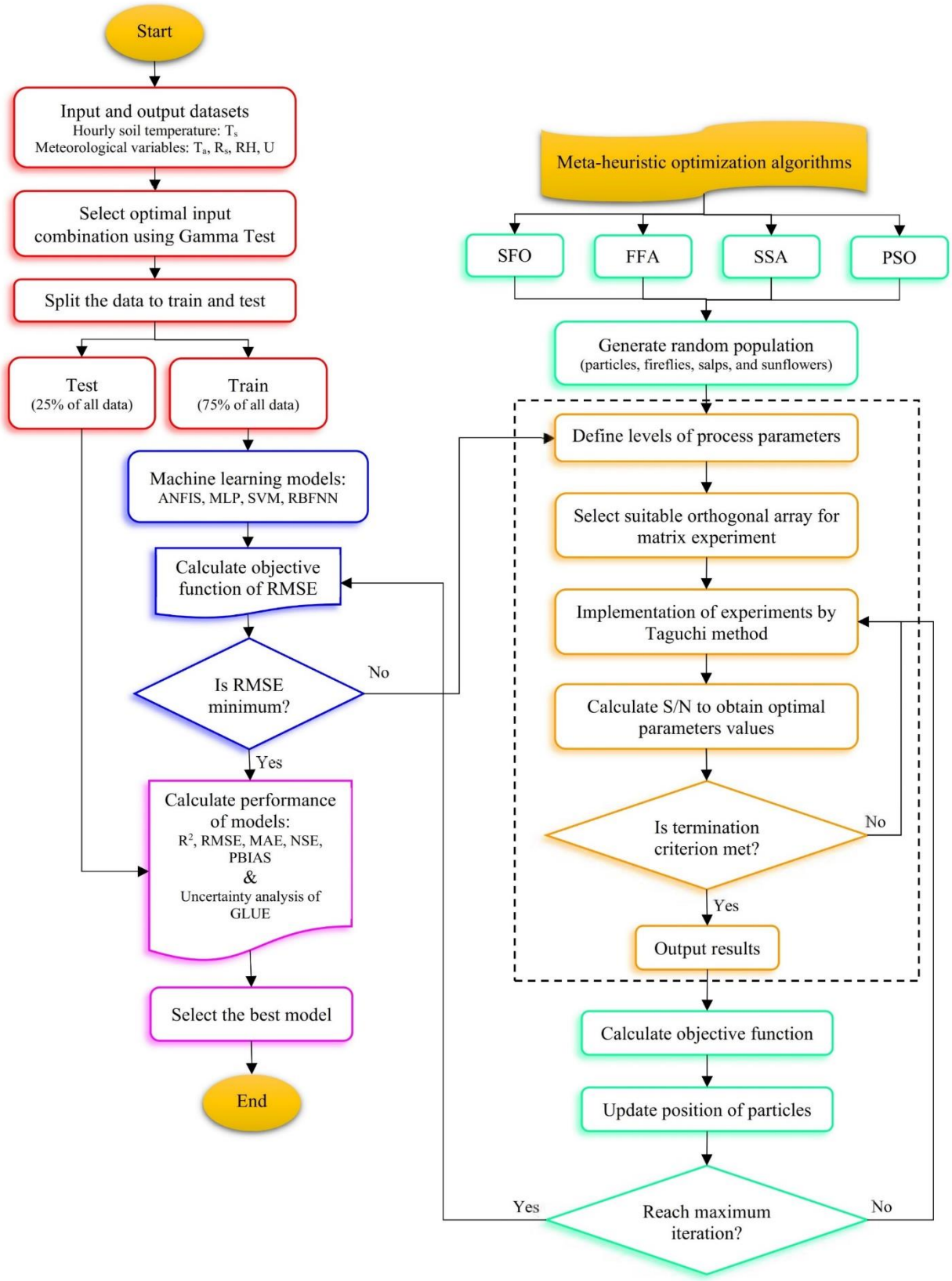


Fig. 3 Framework for predicting hourly soil temperature using hybrid models.

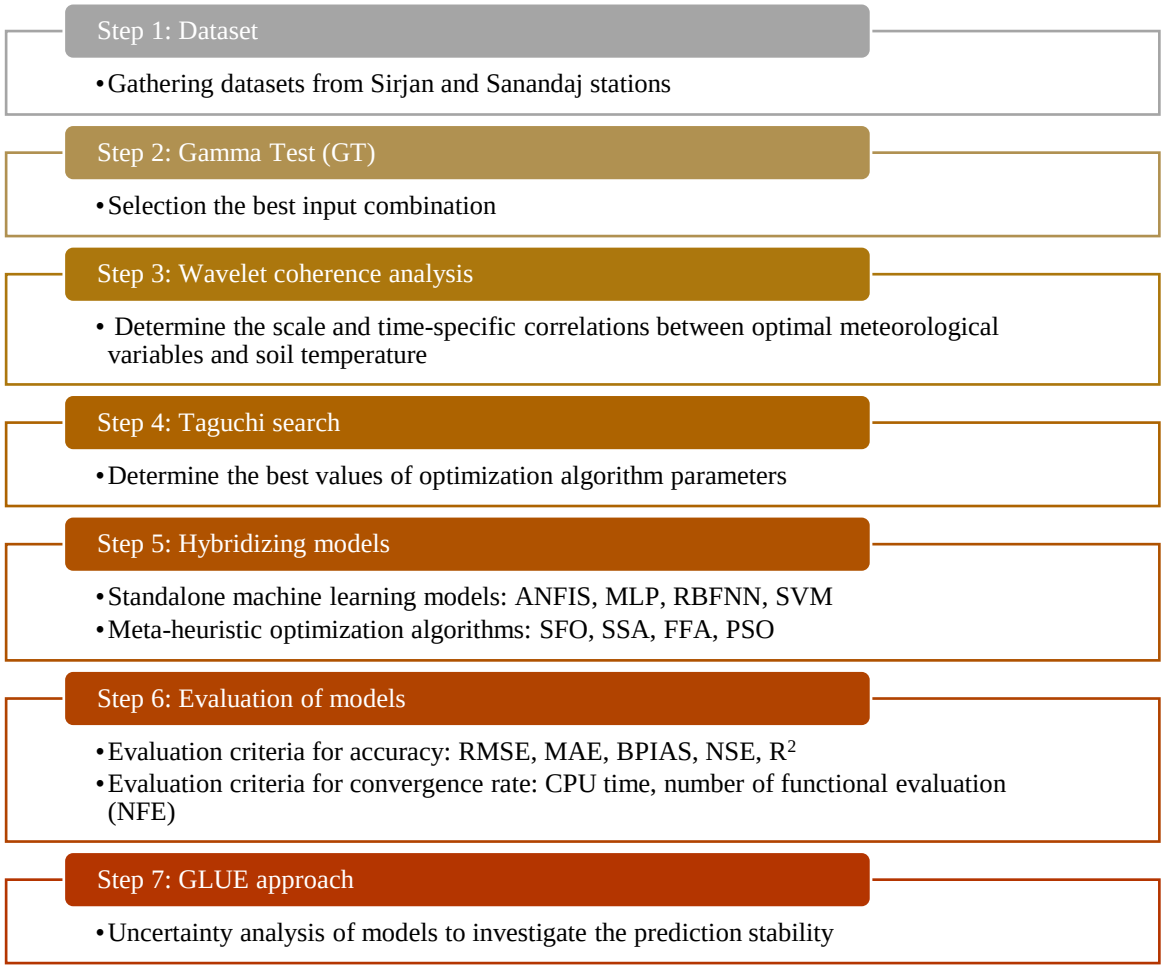


Fig. 4 Schematic diagram representing the general procedures of the proposed method

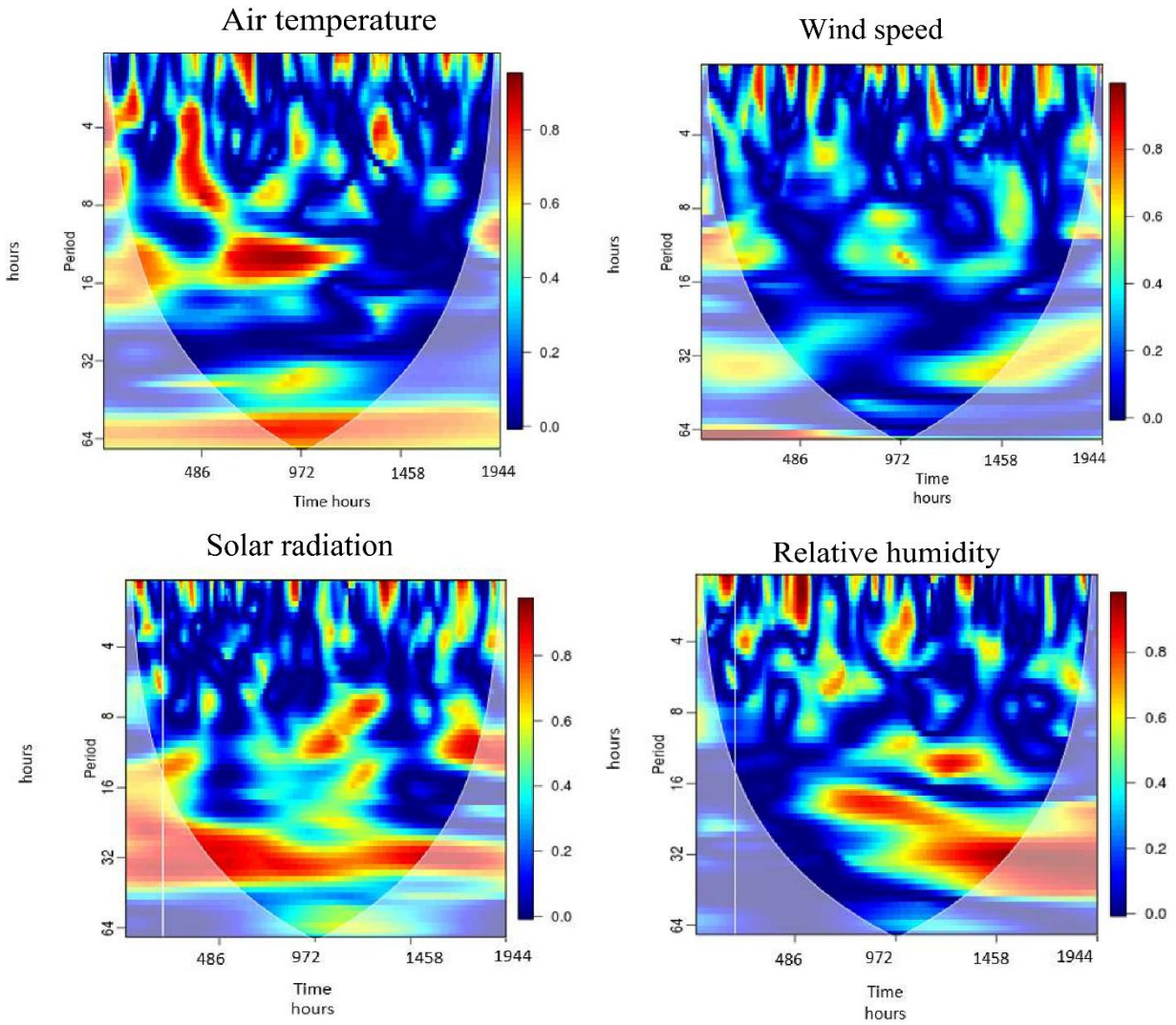


Fig. 5 The coherence value for meteorological parameters in Sirjan at 5 cm soil depth

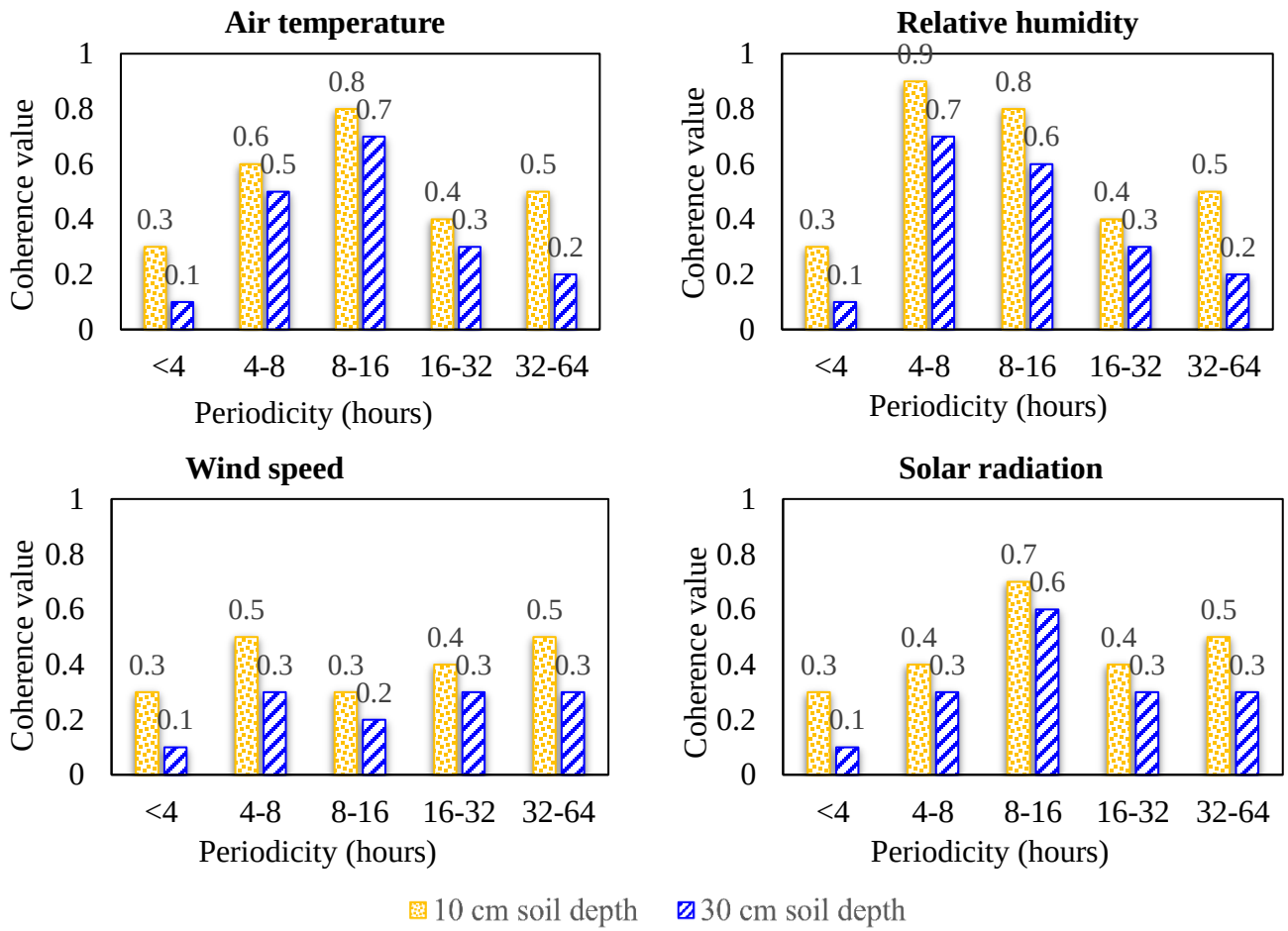


Fig. 6 The coherence value for meteorological parameters in Sanandaj station at 10 and 30 cm soil depths

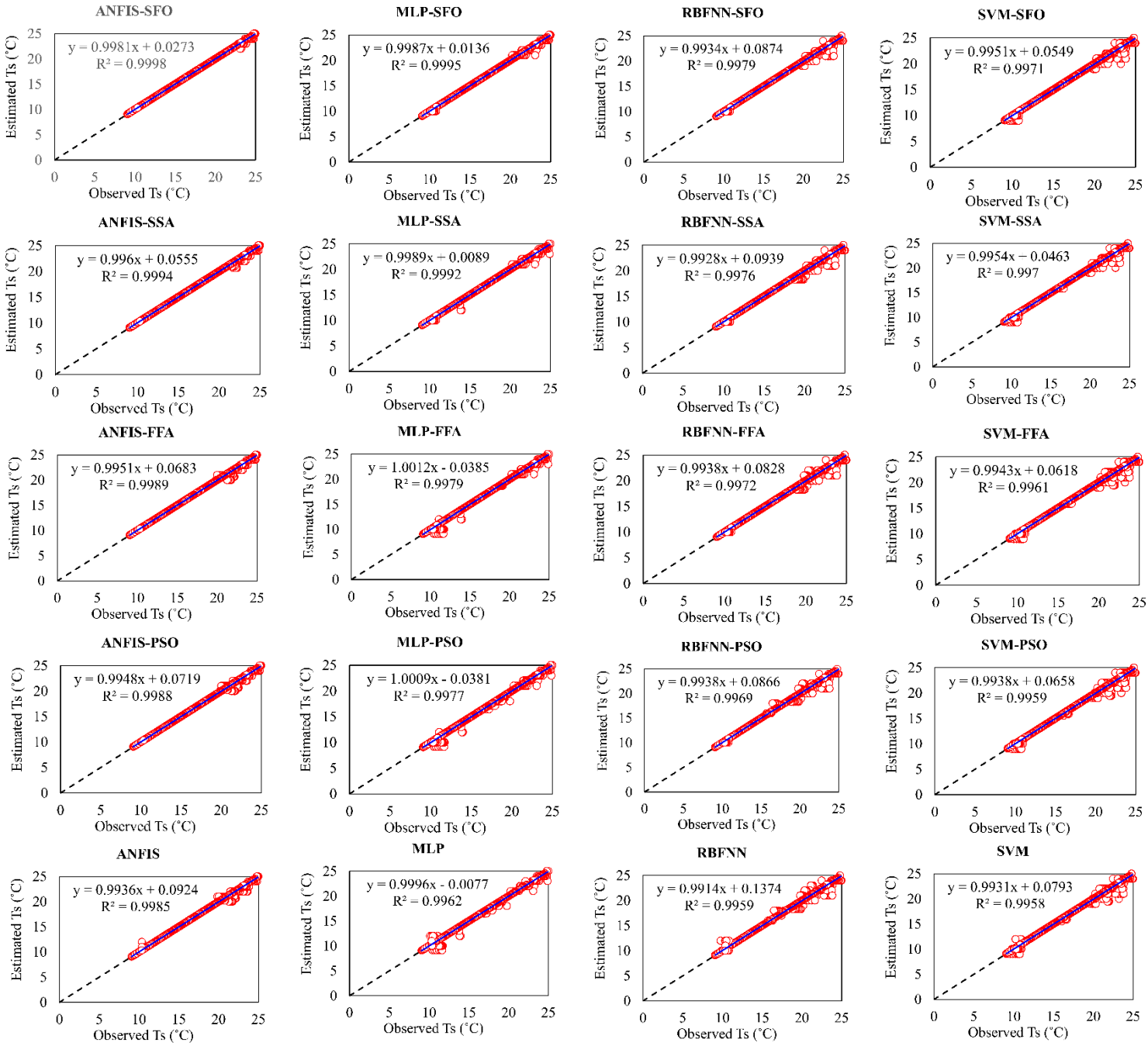


Fig. 7 Soil temperatures estimated by developed hybrid and standalone models versus the measured values at 5 cm soil depth in Sirjan station

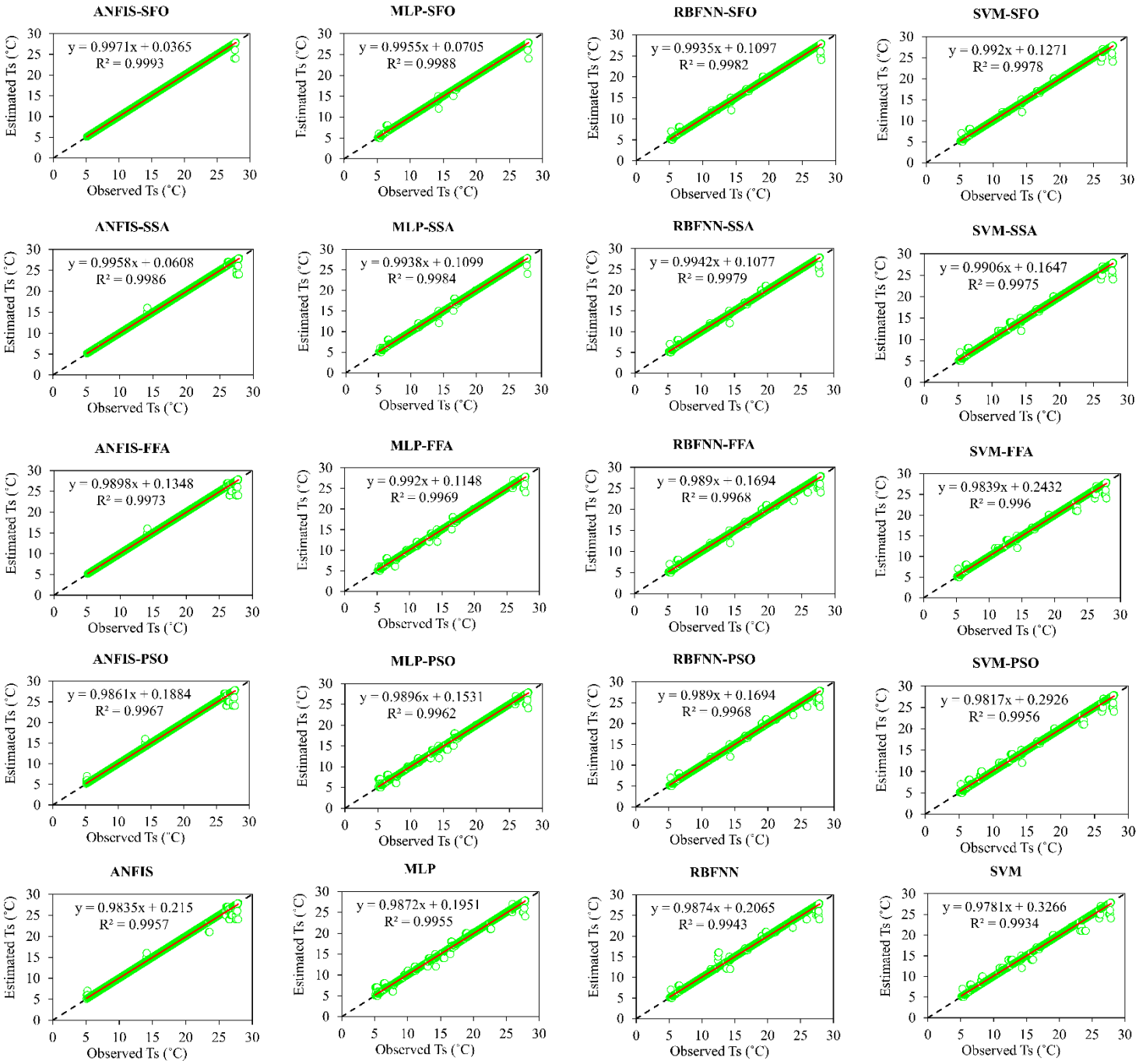


Fig. 8 Soil temperatures estimated by developed hybrid and standalone models versus the measured values at 10 cm soil depth in Sanandaj station

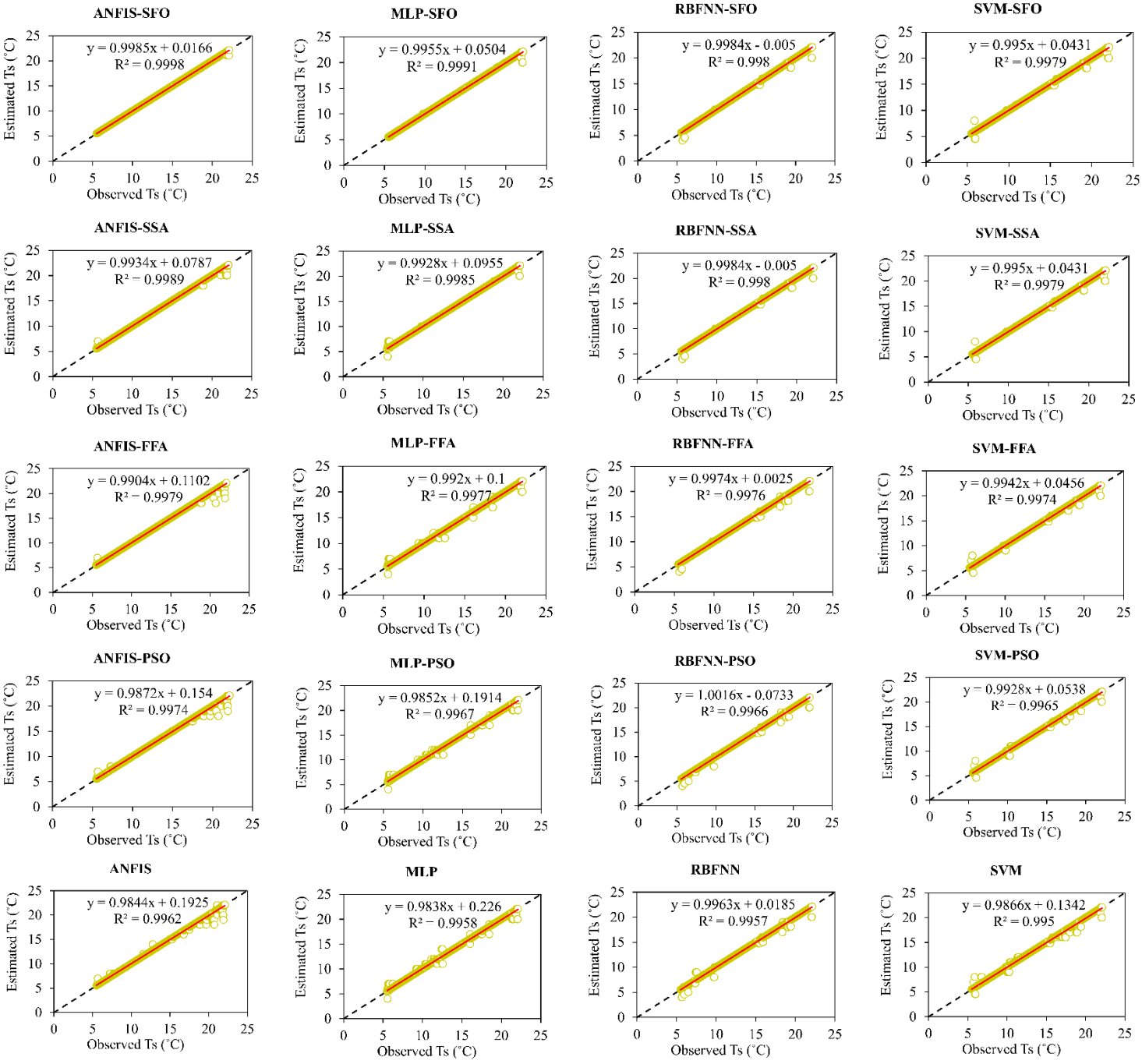
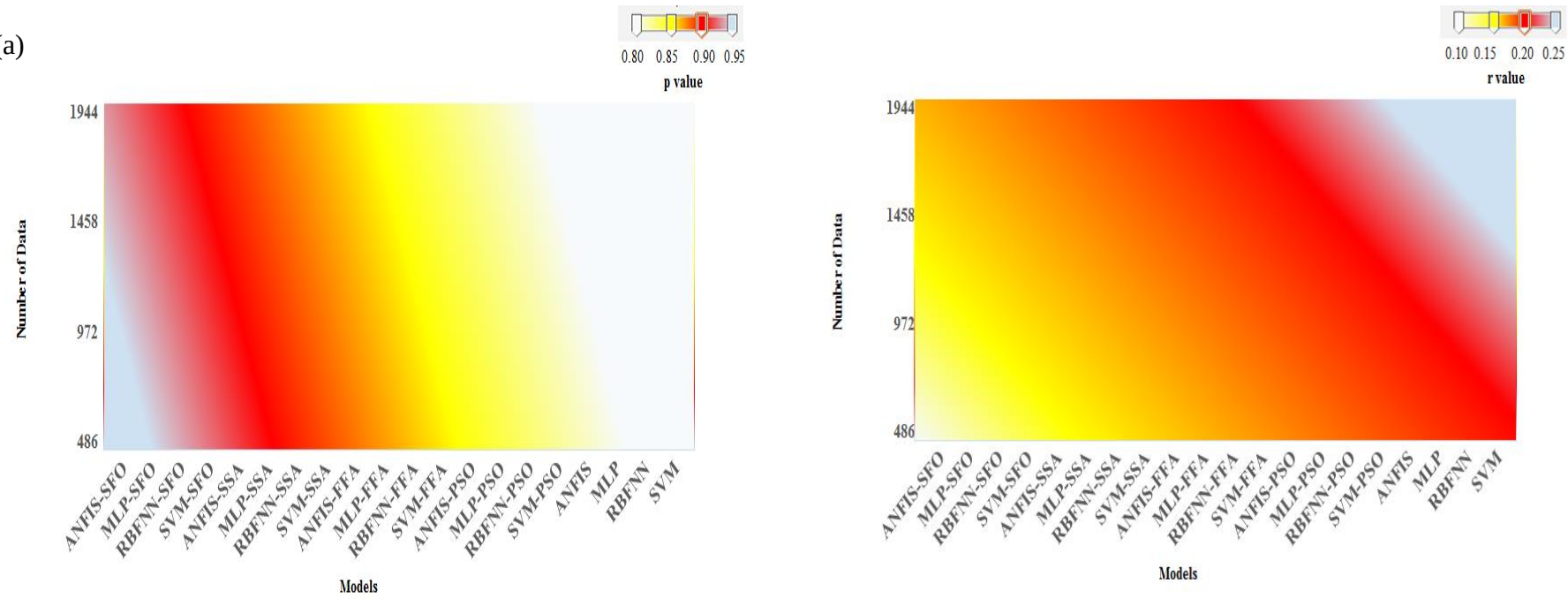
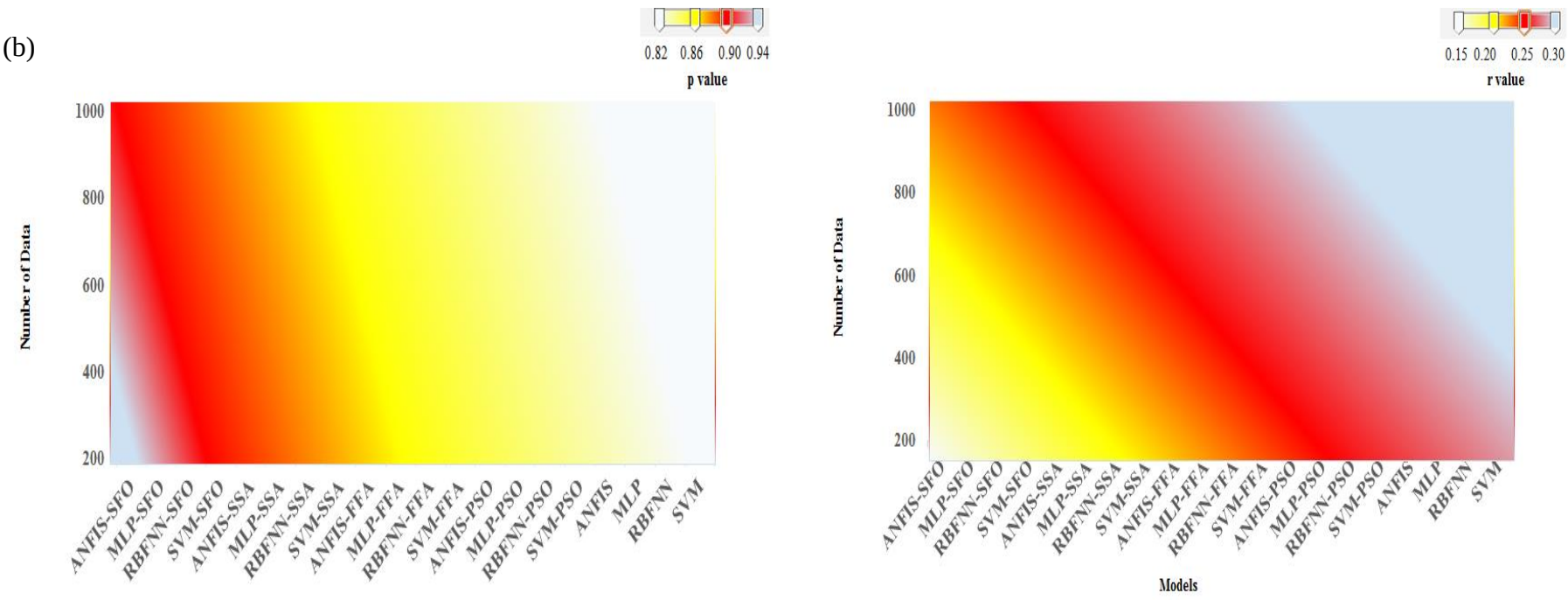


Fig. 9 Soil temperatures estimated by developed hybrid and standalone models versus the measured values at 30 cm soil depth in Sanandaj station

(a)



(b)



(c)

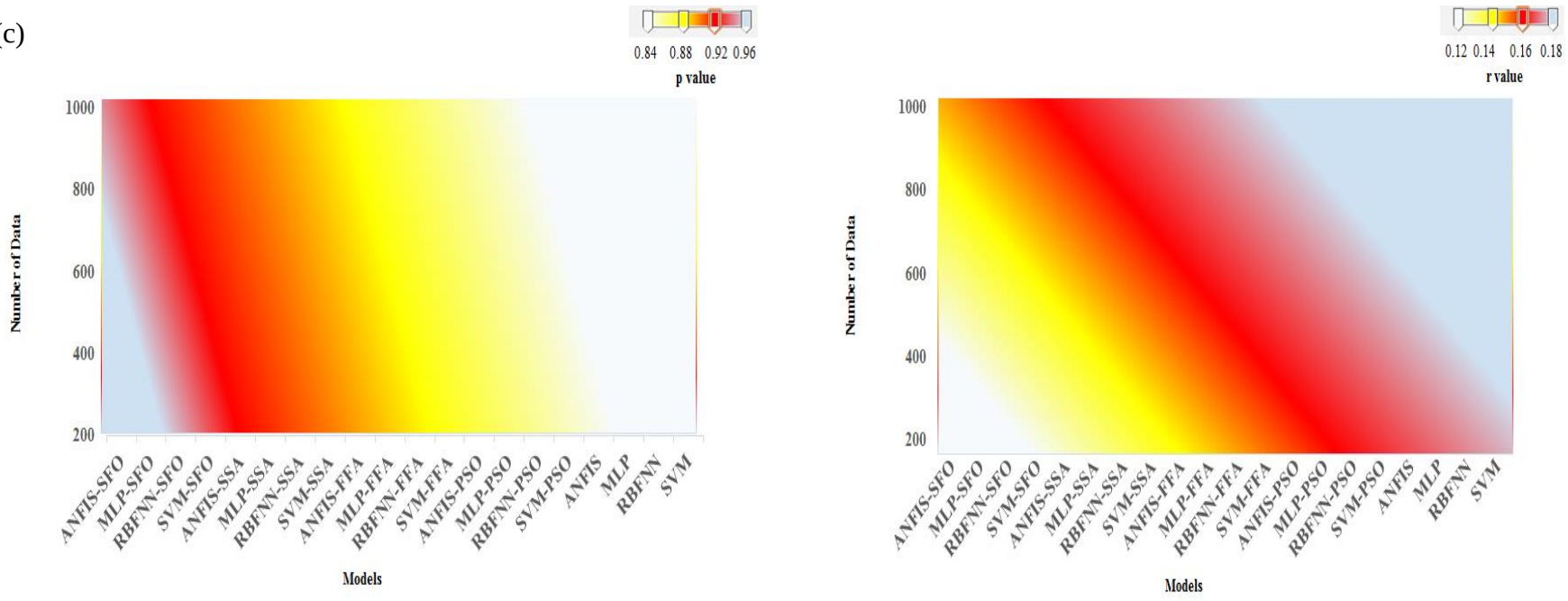


Fig. 10 The uncertainty analysis based on 1: r and 2: p values at a: 5 cm, b: 10 cm and c: 30 cm soil depth in studied stations.