Soil moisture estimation using novel bio-inspired soft computing approaches

Roozbeh Moazenzadeh, Babak Mohammadi, Mir Jafar Sadegh Safari & Kwok-wing Chau

Published online: 22 Mar 2022.

Submit your article to this journal ☐

Article views: 517

View related articles ☐

View Crossmark data ☐

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Check for updates

Soil moisture estimation using novel bio-inspired soft computing approaches

Roozbeh Moazenzadeh [a], Babak Mohammadi [b], Mir Jafar Sadegh Safari [c] and Kwok-wing Chau [d]

[a]Department of Water Engineering, Faculty of Agriculture, Shahrood University of Technology, Shahrood, Iran; [b]Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden; [c]Department of Civil Engineering, Yaşar University, Izmir, Turkey; [d]Department of Civil and Environmental Engineering, Hong Kong Polytechnic University, Hong Kong, People's Republic of China

ABSTRACT

Soil moisture (*SM*) is of paramount importance in irrigation scheduling, infiltration, runoff, and agricultural drought monitoring. This work aimed at evaluating the performance of the classical ANFIS (Adaptive Neuro-Fuzzy Inference System) model as well as ANFIS coupled with three bio-inspired metaheuristic optimization methods including whale optimization algorithm (ANFIS-WOA), krill herd algorithm (ANFIS-KHA) and firefly algorithm (ANFIS-FA) in estimating *SM*. Daily air temperature, relative humidity, wind speed and sunshine hours data at Istanbul Bolge station in Turkey and soil temperature values measured over 2008–2009 were fed into the models under six different scenarios. ANFIS-WOA (*RMSE* = 1.68, *MAPE* = 0.04) and ANFIS (*RMSE* = 2.55, *MAPE* = 0.07) exhibited the best and worst performance in *SM* estimation, respectively. All three hybrid models (ANFIS-WOA, ANFIS-KHA and ANFIS-FA) improved *SM* estimates, reducing RMSE by 34, 28 and 27% relative to the base ANFIS model, respectively. A more detailed analysis of model performances in estimating moisture content over three intervals including [15–25], [25–35] and $\geq 35\%$ revealed that ANFIS-WOA has had the lowest errors with *RMSEs* of 1.69, 1.89 and 1.55 in the three *SM* intervals, respectively. From the perspective of under- or over-estimation of moisture values, ANFIS-WOA (*RMSE* = 1.44, *MAPE* = 0.03) in under-estimation set and ANFIS-KHA (*RMSE* = 1.94, *MAPE* = 0.05) in over-estimation set showed the highest accuracies. Overall, all three hybrid models performed better in the underestimation set compared to overestimation set.
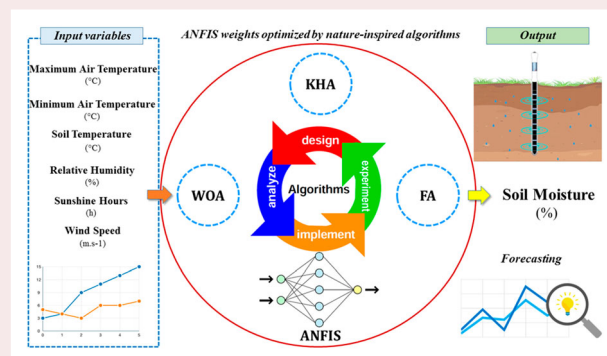
1. Introduction

Moisture is one of the most important soil properties and plays a substantial role in energy partitioning between topsoil and atmosphere (Kornelsen & Coulibaly, 2014; Maroufpoor et al., 2019; Wang et al., 2011). This parameter is dramatically affected by the interactions occurring within the soil profile – such as changes in physical and chemical properties of the soil – or by such external factors as evaporation, irrigation, precipitation and vegetation (Elshorbagy & Parasuraman, 2008; Gaur & Mohanty,

2013; Lee et al., 2019). It is one of the main parameters fed into soil-vegetation-atmosphere transfer (SVAT) models (Ridler et al., 2012). *SM* is also a very important indicator in agricultural drought monitoring (Koley & Jeganathan, 2020; Shahabfar et al., 2012; Zormand et al., 2017).

SM data are quite vital in several areas such as runoff estimation, climate modeling, flood control, geotechnical engineering, slope failure prediction and water quality assessment. It is considered one of the principal factors

in irrigation scheduling, and determines irrigation water requirement and irrigation timing. Over recent years, the idea of precision agriculture on the farm scale has been very popular among researchers and farmers. One of the approaches in precision agriculture is determination of irrigation water requirement and the appropriate irrigation time, which has been realized with the use of moisture sensors and by accurate determination of *SM*. However, spatial and temporal variability, which is a characteristic feature of *SM*, complicates its monitoring (Dobriyal et al., 2012).

The significance of *SM* in irrigation scheduling becomes much more apparent on the small scale of a farm. Accordingly, several studies have been conducted to estimate *SM* at field capacity (*FC*) and permanent wilting point (*PWP*). Application of moisture sensors for exact irrigation timing based on real-time *SM* measurement has been also developed under the concept of 'precision agriculture' and with the aim of optimizing water consumption and minimizing the water stress imposed on plants (Seneviratne et al., 2010; Zhou et al., 2019). The role of this parameter is obvious on a large scale, as in catchments, where it plays a part in controlling such processes as infiltration, runoff and evapotranspiration and eventually the water balance of catchments.

Gravimetric and volumetric methods, neutron probes and time domain reflectometers (TDR) are some of the methods and tools used for measuring *SM*. Capacitance-based sensors, which measure the moisture content in real-time and allow measurements to be recorded and stored by data loggers, have also attracted the attention of researchers in recent years (Jimenez et al., 2019). Although these methods provide measurements close to actual values – apart from measurement and instrumentation errors – they do have limitations such as being time consuming, labor intensive, expensive, or complex.

In the unsaturated phase, moisture is a function of matric potential and soil hydraulic properties and varies by soil depth and time. One of the best equations incorporating these variations is Richard's equation, which can be solved either numerically or analytically with certain boundary conditions (Abbasi et al., 2003; Sadeghi et al., 2011; Sadeghi & Jones, 2012). Two general approaches can be used for estimating *SM*. In the first approach, *SM* is regarded as a function of matric potential and estimated by pedotransfer functions. In parametric functions, the parameters of an equation describing $(\theta - h)$ relation are expressed in the form of pedotransfer functions (Javanshir et al., 2020). Point transfer functions are a class of functions which explicitly estimate moisture content for a particular suction (Dobarco et al., 2019), most importantly at -33 kPa and -1500 kPa (Ghorbani et al., 2017;

Vaheddoost et al., 2020). In the second approach, *SM* is estimated regardless of the magnitude of suction by various techniques and using the soil physical, chemical and hydraulic parameters.

Two approaches can be adopted for improving *SM* estimation. The first is to better understand the physics governing the relationships which occur in the soil profile. This, however, appears to be very difficult and complex due to the considerable diversity of the variables controlling the soil environment; although researchers who work in the pure branch of soil science have always shouldered this task (Amente et al., 2000; Chanasyk & Neath, 1996; Elliott & Price, 2020; Gao et al., 2013; Jung et al., 2019; Pan et al., 2015; Pan & Nieswiadomy, 2016; Pardossi et al., 2009; Pires et al., 2005). The second approach, which has been taken by engineers from different disciplines, seeks improvement of the tools employed and application of more efficient techniques, so that soil properties can be better approximated with the least complexity (Zhang, Shao, et al., 2017; Maroufpoor et al., 2019; Moazenzadeh & Mohammadi, 2019; Pezij et al., 2019; Penghui et al., 2020).

SM is also one of the pivotal components used in agricultural drought assessment, and various SM-based indicators have been developed for drought monitoring (e.g. normalized difference moisture index or NDMI, soil adjusted vegetation index or SAVI, moisture stress index or MSI, and soil wetness deficit index or SWDI). The use of satellite images and various algorithms which have enabled *SM* monitoring on different temporal and spatial scales (Kolassa et al., 2018; Lees et al., 2021; Mishra et al., 2021; Moran et al., 2004; Sadeghi et al., 2015; Sadeghi et al., 2020; Schnur et al., 2010; Vergopolan et al., 2020; Wang et al., 2011; Yang et al., 2019), application of AI-based models (Ahmad et al., 2010; Srivastava et al., 2015; Im et al., 2016; Karandish & Simnek, 2016; Brandhorst et al., 2017; Zhang, Zhang, et al., 2017 ; Han et al., 2018; Shin et al., 2018), or simultaneous application of both options have led to significant improvements in *SM* estimation studies. An AI-based model with meteorological variables including air temperature, wind speed, relative humidity and solar radiation as inputs was developed by Tsang and Jim (2016) to determine irrigation time and depth based on *SM*. Coopersmith et al. (2016) estimated *SM* at a depth of 5 cm by a machine learning technique using measured moisture contents at 10 cm depth and antecedent precipitation in 2010 and 2013, and reported *RMSEs* of 0.0173 and 0.0215 $m^3.m^{-3}$, respectively. Using dynamic variables (albedo, NDVI (normalized difference vegetation index), and LST (land surface temperature)) extracted from satellite images as well as steady-state variables (DEM, longitude, and latitude), Liu et al. (2020) investigated the capability of six machine learning

algorithms for *SM* spatial downscaling in four regions in North America, Spain, China, and Australia. Their findings showed that the RF (random forest) model has had the best performance and has been mostly affected by DEM, LST and NDVI variables. Effectiveness of AI-based models has been also proved in various areas of study including soil temperature estimation (Moazenzadeh & Mohammadi, 2019), water resources engineering (Deo et al., 2017) and solar energy management (Feng et al., 2018; Moazenzadeh et al., 2022; Qin et al., 2018; Qin et al., 2019; Qiu et al., 2022; Wang et al., 2016; Wang et al., 2019).

The *SM* content in the rhizosphere is important from the viewpoint of irrigation scheduling. On the one hand, the moisture required by a plant is mainly absorbed from topsoil and on the other, topsoil is the layer from which evaporation occurs, and it plays an important role in energy exchange with the atmosphere. A review of the literature shows that *SM* estimation is impacted by complexities associated with nonlinear changes in both space and time. Meteorological parameters, physico-chemical parameters of the soil and interactions that occur in this environment will control soil moisture variations. Although pedotransfer functions (at specific matric potentials) and satellite imagery have been successfully used for *SM* estimation, the need for a database of soil physical and chemical properties in the former, and the constraints of obtaining appropriate images and complexity of algorithms in the latter case do limit the application of these methods. Therefore, using the main meteorological parameters (which are measured at most stations) and AI models (which possess a remarkable ability to simulate nonlinear phenomena) can be an effective solution worth examining. As a novel contribution in terms of applied algorithms, this study scrutinizes the applicability of three bio-inspired metaheuristic optimization algorithms for hybridization with ANFIS to construct robust models for *SM* computation using meteorological variables as model inputs.

2. Materials and methods

2.1. Study area and data collection

Istanbul is the largest city in the north-west of Turkey and the Strait of Istanbul (the Bosphorus) bridges Asia and Europe. Its climate is a combination of the Black Sea and Mediterranean climate types, with warm summers and cold winters. According to the European Commission and European Soil Bureau Network (2006), four soil types including calisols, cambisols, leptosols and fluvisols are found in Turkey. Istanbul Bolge weather station is located in the Kartal district of Istanbul at 40°54' N, 29°09' E.

Most of the agricultural soils in the region are classified as leptosols. With an area of about 5540 km^2 and 815 mm of precipitation per year (Cuceloglu et al., 2017), Istanbul has approximately 74100 hectares of cultivated land. Wheat, barley, oats, paddy, corn, broad beans, chickpeas, beans, vetch, sugar beet, sunflower, potato and alfalfa are produced in Istanbul. Wheat, sunflower and barley are the most important plant products in the region. In addition, although the production of vegetables such as tomatoes, lettuce, beans, watermelon and spinach is dominant in Istanbul, almost all other vegetables are also grown such as cucumber, lettuce and spring onion. According to the Ministry of Agriculture and Forestry (2017) report, between 1975 and 2015, 1209 floods took place in Turkey which caused almost $100 million economic losses annually. As Istanbul is the most populated city in Turkey, it is a high-risk region in terms of flood hazard. The flood taken place in Marmara region of Istanbul is 2009 resulted in a death of at least 31 people in Istanbul and considerable economic losses (Altunkaynak & Bizimana, 2020; Ekmekcioğlu et al., 2021). Therefore, the study area is important both in terms of agriculture (from the perspective of irrigation scheduling and drought monitoring in the region) and water balance management at the catchment level (through controlling the infiltration component and surface runoff) as well as in terms of flood control.

In the present study, *SM* content and soil temperature were measured once a day at a depth of 20 cm in different parts of the study area at various time intervals totaling 490 days between 2008 and 2009. SM was measured by the volumetric method using a Campbell Scientific CS616 device which measures the volumetric water content of porous media from zero to saturation by vertically inserting its probe rods into the soil. This device gives an indication of the water content in the upper 30 cm of soil. In this study, measurements were performed at 20 cm below the soil surface. The main meteorological variables including air temperature (minimum and maximum), relative humidity, wind speed and sunshine hours measured at Istanbul Bolge station over these same time intervals were also obtained from the Turkish State Meteorological Service, a governmental bureau authorized with collecting climatic and meteorological data throughout the country. The study area is shown in Figure 1.

Statistical indices for measured parameters (soil temperature and moisture at 20 cm depth as well as meteorological data) are given in Table 1, and distribution of moisture contents of the soil samples is plotted in Figure 2. As can be seen in Figure 2, the highest density of *SM* contents falls in the range above 35% (by volume). Moisture contents of about 16% of samples are in the range
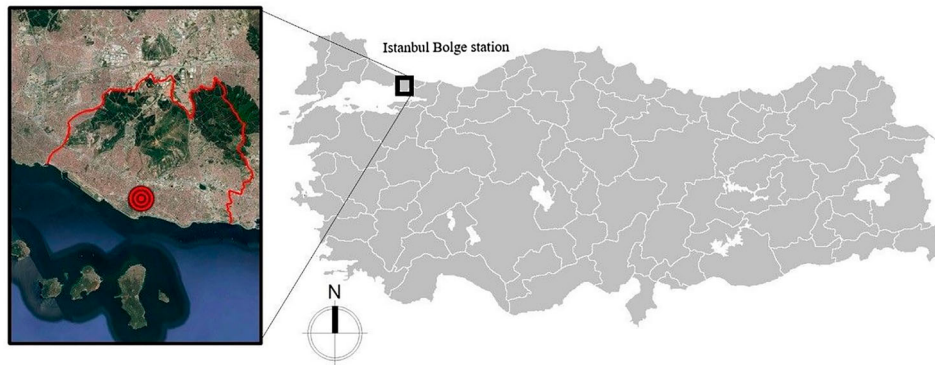
Figure 1. Location of the study area.

Table 1. Statistical profiles of the data utilized for *SM* estimation.

Variables	Min	Max	Mean	SD	SV	Kurtosis	Skewness
SM (%)	18.00	43.00	32.35	6.29	39.58	−0.67	−0.70
ST (°C)	4.40	30.90	15.39	8.29	68.66	−1.16	0.53
RH (%)	39.20	96.30	69.74	10.23	104.60	−0.26	−0.12
U_2 (m/s)	0.40	5.70	2.04	0.80	0.64	1.91	1.11
Tmax (°C)	1.20	38.10	17.73	8.51	72.49	−0.99	0.32
Tmin (°C)	−5.10	25.10	10.07	7.03	49.43	−0.97	0.43
SSH (h)	0.00	10.80	4.69	3.65	13.31	−1.54	−0.004

SD: Standard deviation, SV: Sample variance.

[15–25), about 31% of samples have moisture contents in the range [25–35), and about 53% have moisture contents higher than 35%.

2.2. Estimating SM

2.2.1. ANFIS model

Introduced by Jang (1993), ANFIS is a combination of ANNs and fuzzy inference systems (FISs). This intelligence system uses a non-linear process from input to output and has a multi-layer form with three sections connected with a number of nodes (Leski & Czogala, 1999). In the first part, a rule-based approach including fuzzy rules is selected. The second part contains the database. A learning algorithm with the compatible membership functions (MFs) is used to create a set of fuzzy if–then laws from certain input-output subsets. For this purpose, ANFIS receives the primary FIS and modifies it with a propagation algorithm. Finally, in the last section, a logical inference system is used to remove it from the rules and input data to reach an acceptable output. The most famous models used in fuzzy systems are Mamdani, Sugeno and Takagi models. Figure 3 illustrates an ANFIS structure with two inputs (x_1, x_2) and one output (F) in which a circle represents a stable node and a square shows an adaptive node (Mohammadi et al., 2020).

2.2.2. Hybrid models

In the present study, classical ANFIS model's ability was improved by three bio-inspired optimization algorithms in the form of hybrid models. The algorithms included Firefly Algorithm (FA), Krill Herd Algorithm (KHA), and Whale Optimization Algorithm (WOA). These algorithms find the optimal ANFIS parameters considering the lowest error rate (Devi & Vijayalakshmi, 2020). Hybrid models were implemented in three steps: (i) Calculating error rates with optimal ANFIS parameters (ii) Computing the series values (iii) Fitting the ANFIS model to the values computed in previous steps.

2.2.2.1. Firefly algorithm (FA). FA was initially presented by Yang (2009), and its main idea was inspired by the real life of fireflies (Yang, 2009). It can be considered a new manifestation of swarm intelligence, which has been widely used as a boosting tool in hydrological modeling (Moazenzadeh & Mohammadi, 2019). Researchers have found that fireflies use flashing lights as a protective mechanism, alerting their conspecifics in the environment. The rhythm or frequency of flashing light, the rate at which light flashes, and the duration of light blinking by fireflies form the different parts of the communication system of these insects (Yang, 2009). Since the attractiveness of a firefly is proportional to the light intensity seen by adjacent (neighboring) fireflies, the attractiveness parameter can be defined by the following equation:

$$\beta(r) = \beta_\circ e^{-\gamma^{rz}} \tag{1}$$

where $\beta_\circ$ is attractiveness of the brighter insect at r = 0. The fireflies find the optimized value of each objective function.
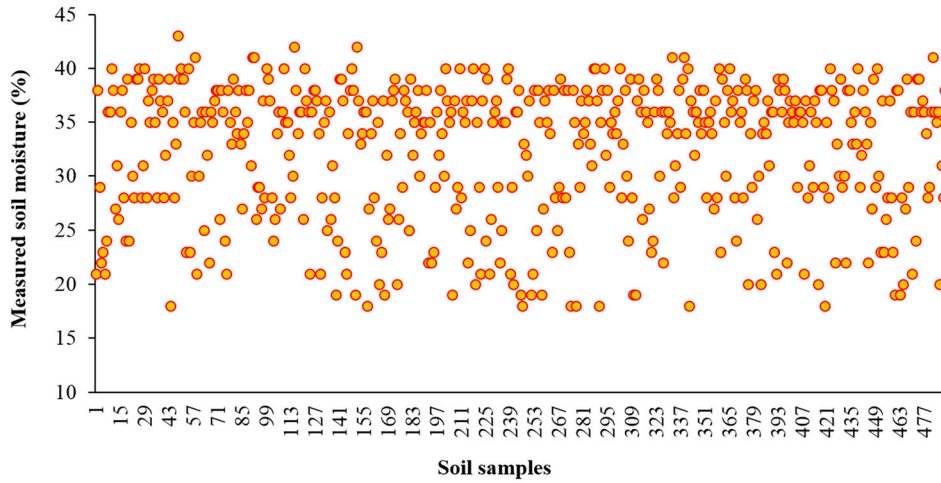
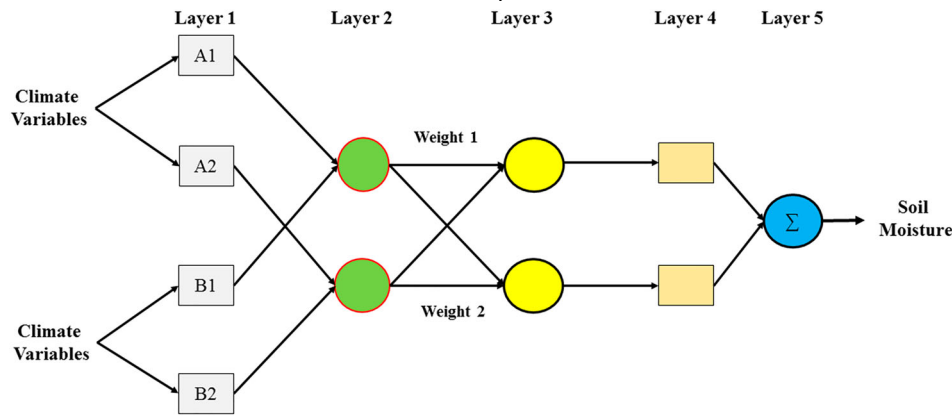**Figure 2.** Distribution of measured moisture contents in the soil samples.

Figure 3. ANFIS structure used in the present study.

2.2.2.2. Krill Herd Algorithm (KHA). KHA was introduced by Gandomi and Alavi (2012) as a solution for optimization problems. It is a metaheuristic population-based algorithm and was inspired by the natural behavior of krill herds. The fitness function of each krill is unique and equals the distance from food to the highest congestion density. Types of krill movement include: (1) Herd movement; (2) Activity to find food; (3) Random displacement. If a herd is attacked by predators, a group of krill will be removed, resulting in a reduction in their density. Krill shrinkage is a multi-purpose process with two main purposes: increasing density and reaching food. Each shrimp moves toward the best possible solution, which is the shortest distance to the highest density and maximum food (Gandomi & Alavi, 2012).

2.2.2.3. Whale Optimization Algorithm (WOA). WOA is another bio-inspired algorithm that mimics the hunting behavior of whales in nature (Mirjalili & Lewis, 2016). It is performed in three phases as follows: (i) Siege hunting; (ii) Operation phase: the method of attacking the net bubble; (iii) Exploration phase: Hunting search. WOA starts with a set of random solutions. In each repetition, the search agents update their position using the three aforementioned operators (Vaheddoost et al., 2020). This algorithm supposes that the best answer at the moment (the best solution) is prey, so it recognizes the prey and then surrounds it. Once the best search agents are identified, other search agents will update their location to the best search agent. This behavior is expressed by Equations (2) and (3) (Mirjalili & Lewis, 2016):

$$\vec{D}' = |\vec{C}.\vec{X}^*(t) - \vec{X}(t)| \tag{2}$$

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A}.\vec{D} \tag{3}$$

where t denotes the present iteration, X is the location vector, X^* is the location vector of the best solution obtained at the present time, and A and C are coefficient vectors. It should be noted that if there is a better answer, then the parameter of X^* should be updated in each iteration. Vectors A and C are determined as follows (Vaheddoost et al., 2020):

$$\vec{A} = 2\vec{a}.\vec{r} - \vec{a} \tag{4}$$
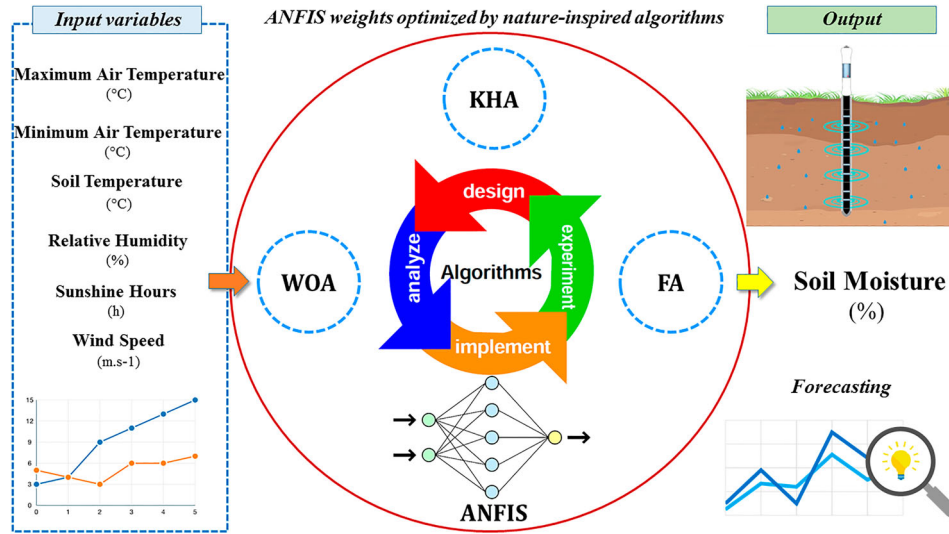
$$\vec{C} = 2.\vec{r} \tag{5}$$

Figure 4. Schematic drawing of *SM* modeling: meteorological variables (five parameters) and soil characteristics (one parameter) used as inputs of AI-based models.
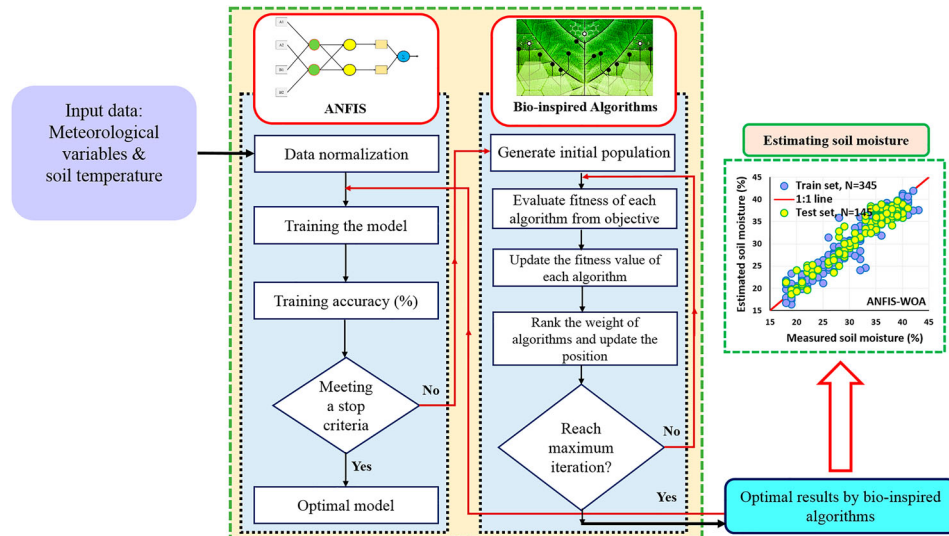
Figure 5. Flowchart of the base (ANFIS) and hybrid (ANFIS boosted by bio-inspired optimization algorithms) models used in the present study for estimating *SM*.

where a decreases linearly from 2 to 0 during iterations (in both exploration and extraction phases), and *r* is a random vector ranging between 0–1.

Figure 4 depicts the general structure used in the present study for estimating SM from meteorological parameters in the form of the ANFIS model and its hybrids with bio-inspired optimization algorithms. Structure of the coupling process between ANFIS and bio-inspired algorithms is shown in Figure 5 (Mohammadi et al., 2020, 2021).

2.3. Scenario definition

In the present work, we first determined the best combination of input variables for the base ANFIS model.

For this purpose, input variables were fed into ANFIS through different scenarios and their performances evaluated according to *SM* estimation error rates in testing set. The best results of each scenario based on the number of input variables are given in Table 2. After selecting the best combination of input parameters fed into the base

Table 2. Best input parameters in each scenario: scenarios with one to six inputs (the best scenario is presented in bold type).

Scenario	Number of inputs	Best input parameters	Output
1	1	*ST*	*SM*
2	2	*ST, Tmin*	*SM*
3	**3**	***ST, Tmin, RH***	***SM***
4	4	*ST, Tmin, RH, Tmax*	*SM*
5	5	*ST, Tmin, RH, Tmax, SSH*	*SM*
6	6	*ST, Tmin, RH, Tmax, SSH, U2*	*SM*

Table 3. Performance analysis of the ANFIS model and ANFIS coupled with bio-inspired optimization algorithms.

Models	Training set					Testing set				
	RMSE	MAPE	RRMSE	NSE	AI	RMSE	MAPE	RRMSE	NSE	IA
ANFIS	2.63	0.07	8.15	0.80	0.95	2.55	0.07	7.85	0.79	0.95
ANFIS-WOA	1.69	0.04	5.22	0.92	0.98	1.68	0.04	5.18	0.91	0.98
ANFIS-KHA	2.01	0.05	6.23	0.89	0.97	1.83	0.05	5.62	0.90	0.98
ANFIS-FA	1.83	0.04	5.67	0.91	0.98	1.87	0.05	5.76	0.90	0.97

ANFIS model (Scenario 3), parameter values were optimized for this model using the optimization algorithms in the form of hybrid models.

2.4. Model evaluation

Five statistical indices including root mean square error (*RMSE*), mean absolute percentage error (*MAPE*), relative root mean square error (*RRMSE*), Nash-Sutcliffe Efficiency (*NSE*) and Index of Agreement (*IA*) were used for assessing model performance as follows:

$$RMSE = \left[\sum_{i=1}^{n} (SM_{(i,obs)} - SM_{(i,est)})^2 \div n \right]^{0.5} \quad (6)$$

$$MAPE = \left[\left(\sum_{i=1}^{n} |((SM_{(i,est)} - SM_{(i,obs)}) \div SM_{(i,obs)})| \right) \div n \right] \quad (7)$$

$$RRMSE = \left[RMSE \div \left(\frac{1}{n} \sum_{i=1}^{n} SM_{(i,obs)} \right) \right] \quad (8)$$

$$NSE = 1 - \left[\sum_{i=1}^{n} (SM_{(i,obs)} - SM_{(i,est)})^2 \right. $$
$$\left. \div \sum_{i=1}^{n} (SM_{i,est} - \overline{SM}_{obs})^2 \right] \quad (9)$$

$$IA = 1 - \left[\sum_{i=1}^{n} (SM_{(i,obs)} - SM_{(i,est)})^2 \right.$$
$$\left. \div \sum_{i=1}^{n} (|SM_{i,est} - \overline{SM}_{obs}| - |SM_{i,obs} - \overline{SM}_{obs}|)^2 \right] \quad (10)$$

where $SM_{i,obs}$, $SM_{i,est}$ and $\overline{SM}_{obs}$ are observed, estimated and average of measured *SM* values, respectively, and *n* refers to the number of data points. Ertekin and Yaldiz (2000) proposed the following categories for rating models according to accuracy: A model is excellent if its *RRMSE* is below 10%, good if 10% < *RRMSE* < 20%, fair if 20% < *RRMSE* < 30%, and poor if *RRMSE* is higher than 30%.

3. Results and discussions

3.1. SM estimation

Figure 6 shows measured *SM* values at 20 cm depth as well as corresponding values estimated using ANFIS as the base model and ANFIS coupled with three bio-inspired optimization algorithms. Model performance indicators are given in Table 3. In training set, application of each of the three optimization algorithms to the ANFIS model improved *SM* estimates, and the best values of all three indices (*RMSE*, *MAPE,* and *RRMSE*) were obtained with ANFIS-WOA. According to the results, the lowest and highest decreases in *RMSE* (*RRMSE*) were approximately 24% (ANFIS-KHA) and 36% (ANFIS-WOA), and minimum and maximum decreases in *MAPE* were about 31% (ANFIS-KHA) and 43% (ANFIS-WOA), respectively. In training phase, the increase in *NSE* from the base ANFIS model (*NSE* = 0.8) to hybrid models ranged from 11% (ANFIS-KHA model) to 15% (ANFIS-WOA model); and the lowest *IA* was 0.95 for the base ANFIS model, whereas all three hybrid models exhibited similar performances with *IA* values of 0.97 and 0.98.

In testing set, all optimization algorithms improved results when coupled with the base ANFIS model. According to the results, the lowest and highest decreases in *RMSE* (*RRMSE*) were about 27% (ANFIS-FA) and 34% (ANFIS-WOA), and the lowest and highest decreases in *MAPE* were about 30% (ANFIS-FA) and 41% (ANFIS-WOA), respectively. In validation phase, the lowest *NSE* (approximately 0.79) was again obtained with the base ANFIS model as expected. ANFIS-KHA and ANFIS-FA showed a similar performance with *NSE* = 0.9 (13.9% higher than ANFIS) and ANFIS-WOA outperformed all models with *NSE* = 0.91 (15.2% higher than ANFIS). In this phase, the *IA* index was improved by about 3% for ANFIS-WOA and ANFIS-KHA and by about 2% for ANFIS-FA relative to the base model, indicating the relatively similar performance of the hybrid models. There are two main reasons for the differences in model performance: (1) Different results due to the differences in training data (input values), (2) Different results due to the differences in learning algorithms or in optimizing method of the learning algorithm; and differences in the models result from the mathematical equations behind them. ANFIS model uses fuzzy
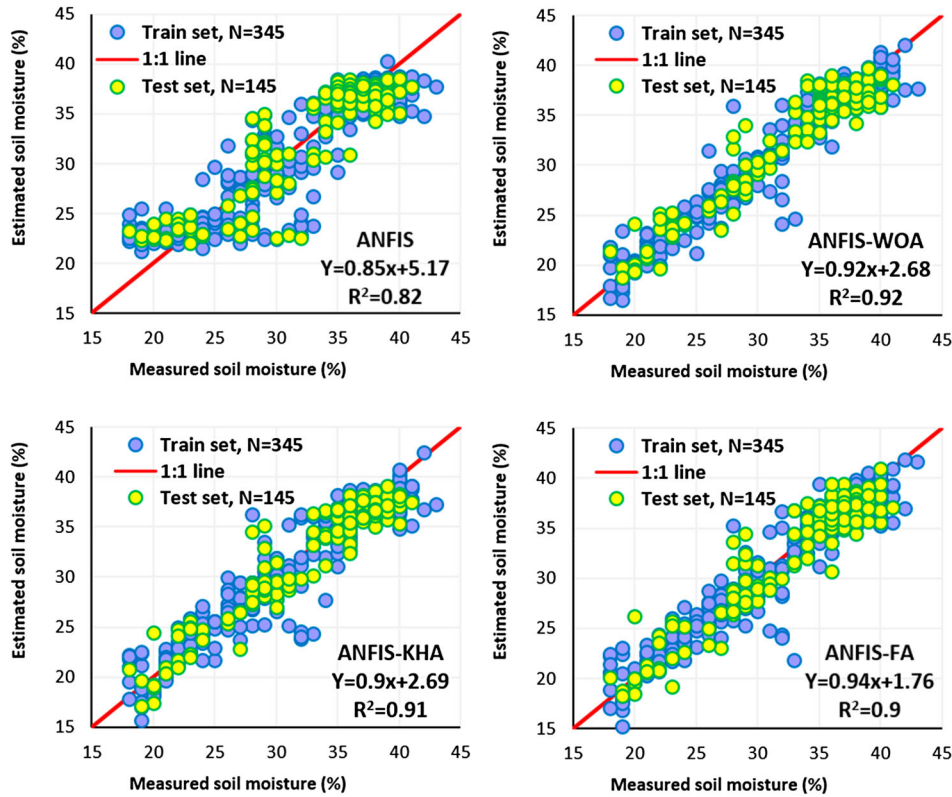
Figure 6. Measured versus estimated *SM* values using ANFIS model, and ANFIS coupled with three bio-inspired optimization algorithms (regression equations and R^2 values are presented for testing sets).

if–then rules by integrating neural networks technique. Then ANFIS finds input-output pairs using its membership functions. Regarding the membership functions of ANFIS in the present study, the differences in ANFIS results are caused by different meteorological variables. The FA algorithm can be implemented for parallel problems and then the system can easily find optimum minimum (Khan et al., 2016). The WOA is a capable method in global search, has a suitable performance in terms of convergence accuracy, and is able to solve complex optimization problems (Ning & Cao, 2021). The KHA considers the neighbor responses around possible optimum response for analyzing all possible optimum values (Gandomi & Alavi, 2012). Therefore, differences in the type and composition of the data used for training of the simulator models and differences in training algorithms, how the parameters of the hybrid model were optimized, and mathematical relationships and simplifications in the course of implementation of each optimization algorithm can be among the most important reasons of the different performances of simulator models.

As one of green water sources, *SM* is a highly important factor in agricultural drought monitoring in any region, to the extent that some drought indicators have

been developed based on it. Also on a large scale (e.g. catchments), *SM* controls the rate of water infiltration and thus can affect one of the main components of water balance, namely surface runoff. All of the above shows that accurate estimation of *SM* can greatly help in proper management of water resources, both on the farm scale by reducing water loss and on the catchment scale by controlling water balance components. Using ERA-interim data, Srivastava et al. (2015) compared performance of the WRF-Noah LSM model (Weather Research and Forecasting model coupled with Noah land surface model) with PDM (Probability Distributed Model) in estimating soil moisture deficit and reported *NSEs* of 0.63 and 0.7 in calibration and validation phases, respectively. Elliott and Price (2020) compared soil moisture values measured by triplicate CS605 TDR probes in different depths and in two profiles with those estimated by three methods of determining van Genuchten model parameters including RETC (SSL), inverse modeling using HYDRUS-1D (TF) and alternate strategy for TF parametrization (ALT) and obtained *NSE* values of −1.4 and 0.24 for SSL, 0.89 and 0.91 for TF, and 0.89 and 0.79 for ALT in the two mentioned profiles. Using a number of soil properties, performance of three AI-based models (ANN, SVR (support vector regression),

and ANFIS) and a hybrid model (ANFIS coupled with Gray Wolf Optimization (GWO)) in simulating *SM* in the northwest of Iran (Sanandaj, Dehgolan plain) was investigated by Maroufpoor et al. (2019). According to an approximately 50% improvement of estimation accuracy of ANFIS-GWO compared to ANFIS and SVR, they proposed ANFIS-GWO as a powerful tool for *SM* estimation. Capability of Random Forest (RF), a type of learning machine with input variables including meteorological, vegetation and soil parameters, was examined for root zone soil moisture estimation in an agricultural catchment by Carranza et al. (2021). With high R^2 values (> 0.75) and low *RMSEs* (< 0.06 m^3.m^{-3}), they proved the accuracy of data-driven models in estimating *SM*. Accuracy of an ANN-based model (with climatic data and rooting depth as inputs) in estimating *SM* during the plant growing season from 2008 to 2010 was examined and compared to that of a process-based model (RZWQM2), and *NRMSEs* of 7.64, 6.37 and 6.66 per cent were reported for the three years of the experiment, respectively (Gu et al., 2021). With *NRMSEs* of 1.27 and 0.02 per cent for 0–15 and 150–200 cm layers, respectively, the ANN-based model proved to be more efficient for evaluating SM dynamics throughout the growth season. Elshorbagy and Parasuraman (2008) compared the performance of a higher-order neural network (HONNs) with that of traditional ANNs and a conceptual model (a site-specific system dynamic watershed, (SDW)) in estimating moisture in peat and till soils of various depths. Inputs of their neural network model included precipitation, air temperature, net radiation and soil temperature at different depths. In a 35 cm deep soil, the proposed HONNs model (with *RMSEs* of 0.05 and 0.01 cm^3.cm^3 in peat layer and till layer, respectively) outperformed SDW (with corresponding *RMSEs* of 0.08 and 0.1, respectively). They concluded that application of neural networks in *SM* estimation is challenging but achievable, and their performance is largely influenced by the structure and formation of the soil covers. Ability of SPoRT-LIS to provide accurate estimates of *SM* in the 0–10 cm layer was examined by McDonough et al. (2018), with *RMSEs* of 0.1009, 0.1043, 0.1136 and 0.1102 at 5 cm depth for Summer, Fall, Winter and Spring, respectively. Corresponding *RMSEs* at 10 cm depth were 0.0879, 0.0917, 0.1071 and 0.1024, respectively.

For a more accurate analysis of the behavior of each simulator model, model performances in estimating moisture over three different intervals including [15–25)% (first interval), [25–35)% (middle interval) and ≥ 35 (last interval) were examined separately in testing set, with the results given in Table 4. Also, the trend of *RMSE* variations over the same three intervals along with *RMSE* values for the entire measured moisture data set ('All data') in testing set are represented in Figure 7.

Examination of the results showed that the lowest and highest *RMSE*, *MAPE* and *RRMSE* values for all models have been obtained in the last interval ($\theta_S \geq 35$) and the middle interval ($25 \leq \theta_S < 35$), respectively. In other words, the trend of variations in error rates between the three intervals has been first upward (from first to middle interval) and then downward (from middle to last interval), so that the minimum increase in error rate is 2.8% (ANFIS-FA) and its maximum is 39% (ANFIS); and the lowest and highest decreases in error rate are 18% (ANFIS-WOA) and 48.9% (ANFIS), respectively. For all models, error reduction from middle interval ($25 \leq \theta_S < 35$) to last interval ($\theta_S \geq 35$) is greater than corresponding increase in error from first interval ($15 \leq \theta_S < 25$) to middle interval ($25 \leq \theta_S < 35$), although this trend is stronger for ANFIS-FA compared to the other models.

Table 4. Performance analysis of the ANFIS model and ANFIS coupled with bio-inspired optimization algorithms over the three discussed intervals, including [15–25], [25–35] and ≥ 35%, as well as the entire testing data set.

| Models	Intervals	RMSE	MAPE	RRMSE%	NSE	IA
ANFIS	[15–25)%	2.53	0.11	11.86	−0.47	0.49
	[25–35)%	3.52	0.09	11.81	0.28	0.68
	≥ 35%	**1.80**	**0.04**	**4.86**	−1.22	0.48
	All data	2.55	0.07	7.85	0.79	0.95
ANFIS-WOA	[15–25)%	1.69	0.06	7.90	0.44	0.79
	[25–35)%	1.89	0.04	6.34	0.68	0.89
	≥ 35%	**1.55**	**0.03**	**4.19**	−0.53	0.65
	All data	1.68	0.04	5.18	0.91	0.98
ANFIS-KHA	[15–25)%	1.78	0.07	8.34	0.51	0.80
	[25–35)%	2.15	0.05	7.19	0.55	0.84
	≥ 35%	**1.63**	**0.04**	**4.41**	−0.23	0.64
	All data	1.83	0.05	5.62	0.9	0.98
ANFIS-FA	[15–25)%	2.06	0.07	9.62	0.34	0.74
	[25–35)%	2.11	0.06	7.08	0.59	0.85
	≥ 35%	**1.66**	**0.03**	**4.48**	−0.18	0.65
	All data	1.87	0.05	5.76	0.90	0.97
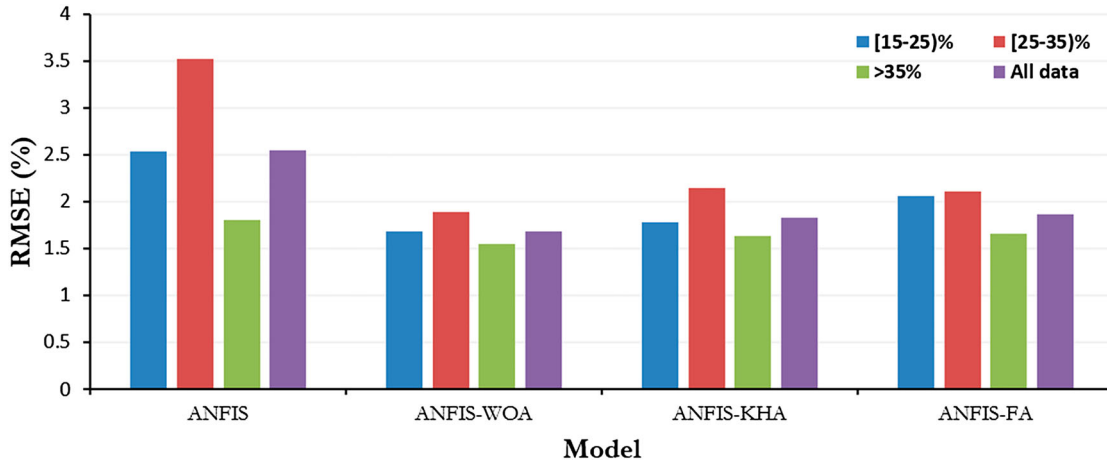
Figure 7. *RMSE* variations for ANFIS and ANFIS coupled with three bio-inspired optimization algorithms over the three discussed intervals as well as the entire testing data set.
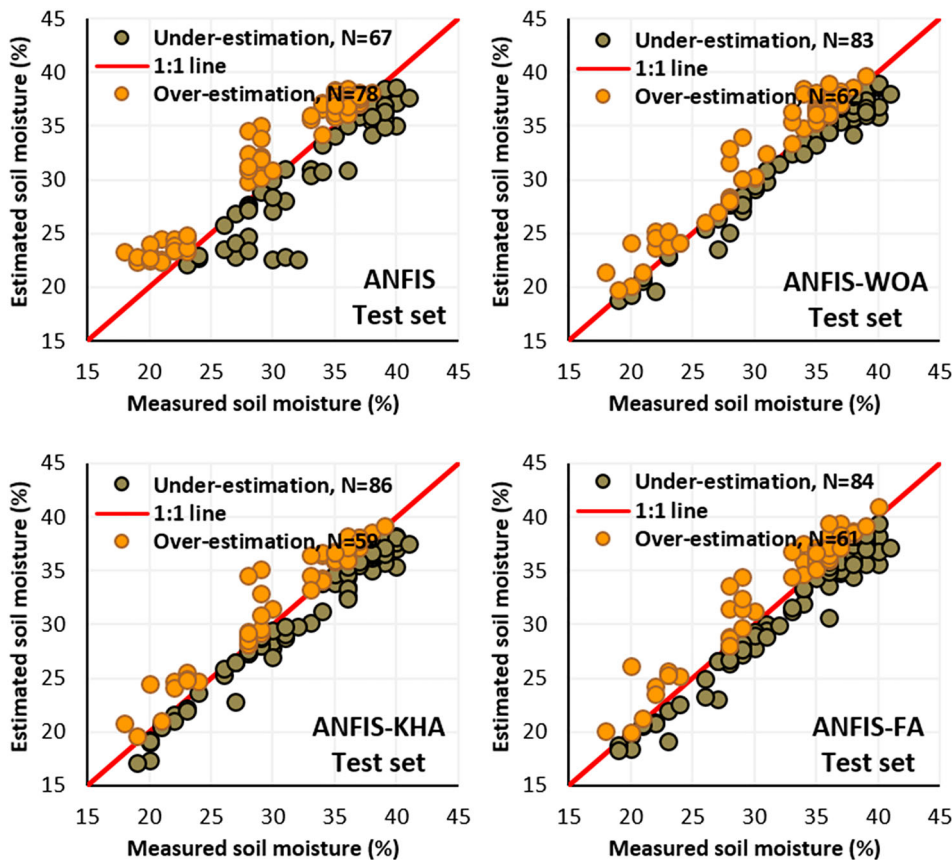

Figure 8. Measured versus estimated *SM* values for all models in under- and over-estimation sets (testing set).

According to the results shown in Table 4, minimum and maximum *RMSE* increments between the last interval ($\theta_S \geq 35$) and the entire testing set (All data) are 8.4% (ANFIS-WOA) and 42% (ANFIS), respectively. The highest difference in estimation error ($\Delta RMSE$) in the first interval ($15 \leq \theta_S < 25$), middle interval ($25 \leq \theta_S < 35$) and last interval ($\theta_S \geq 35$) are also about 0.85, 1.63, and 0.25, respectively, between ANFIS and

ANFIS-WOA in all mentioned intervals. It can be therefore concluded that the differences in performance between the optimization algorithms, and particularly their difference with the base ANFIS model, in the entire testing set (All data) are caused by the differences in how they simulate moisture in the first two intervals ($15 \leq \theta_S < 25$ and $25 \leq \theta_S < 35$), that is the same intervals in which the base ANFIS model had the poorest
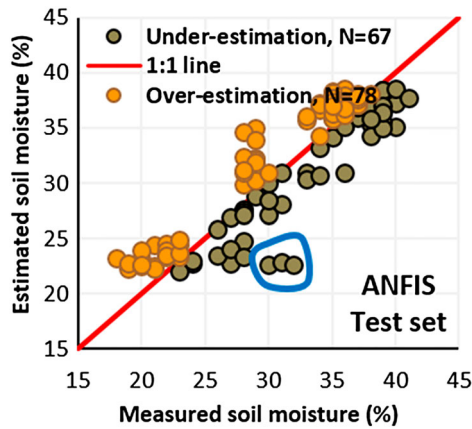
Figure 9. Three particular points of inaccurate soil moisture estimation by ANFIS in under-estimation set.

performance among all models with *RMSEs* of 2.534% and 3.524%, respectively. According to the results (Table 4), the highest decreases in *RMSE* in the first, middle and last intervals relative to the base ANFIS model are 33.4, 46.3 and 13.8 per cent, all of them obtained with ANFIS-WOA.

3.2. Under- and over-estimation

Distribution of measured and estimated *SM* values for all models in testing set and statistical indicators of performance of models in under- and over-estimation sets are given in Figure 8 and Table 5, respectively.

In under-estimation set, ANFIS-WOA has had the best performance with *RMSE*, *MAPE* and *RRMSE* of 1.441, 0.032 and 4.288 per cent, respectively. In over-estimation set, ANFIS-KHA has outperformed the other models with *RMSE* and *RRMSE* of 1.936 and 6.155 per cent, respectively. ANFIS-WOA and ANFIS-KHA share the best *MAPE* value (*MAPE* = 0.049), although it is not much different from that of ANFIS-FA (*MAPE* = 0.053). Comparison of the results shown in Table 5 reveals that apart from the base ANFIS model, all other models have performed better in under-estimation set, with minimum and maximum increases in *RMSE* from under- to over-estimation set being approximately 10.8 and 36 per cent, for ANFIS-KHA and ANFIS-WOA, respectively. One of the main reasons for the different performance of ANFIS

compared to other models (better performance in over-estimation set) has been its moisture estimates at three particular points which are marked in Figure 9. *RMSE* of these three points alone is about 8.35%, while *RMSE* of the whole under-estimated subset of the testing set for the ANFIS model is about 2.72%; and the high error rate in estimating these three particular points has decreased the overall accuracy in under-estimation set.

In under-estimation set and in comparison to the base ANFIS model, minimum and maximum decreases in *RMSE* were 35.7% (ANFIS-KHA) and 47% (ANFIS-WOA), respectively; whereas corresponding values in over-estimation set were 13.8% (ANFIS-FA) and 19.3% (ANFIS-KHA). These figures not only confirm better performance of the models in under-estimation set compared to the over-estimation set, but also show that the percentage of error reduction in estimation of *SM* due to application of each of the optimization algorithms to the base ANFIS model has been higher in under- compared to over-estimation set.

4. Conclusions

Soil moisture plays a key role in irrigation scheduling as the goal of irrigation is to increase current soil moisture until reaching field capacity. Therefore, knowing *SM* content will be very helpful in determining irrigation depth and frequency, which are effective in optimal water consumption and prevention of water stress to plants, respectively. This work examined the capability of bio-inspired optimization algorithms in estimating *SM* over 2007–2008 in Turkey using meteorological variables as model inputs. Hybridization of all bio-inspired algorithms with ANFIS (ANFIS-WOA, ANFIS-KHA and ANFIS-FA) improved *SM* estimates in comparison with the base ANFIS model, although ANFIS-WOA performed best with an *RMSE* of 1.68. All models showed almost the same performance for measured moisture contents higher than 35%, so it can be concluded that the differences in performance among the optimization algorithms and their differences with the base ANFIS model in estimating *SM* has stemmed from the first two intervals, i.e. [15–25) and [25–35)%. In terms of model

Table 5. Statistical indices for the base ANFIS model and ANFIS coupled with bio-inspired optimization algorithms in under- and over-estimation sets (testing set).

	Under-estimation set					Over-estimation set				
Models	RMSE	MAPE	RRMSE	NSE	AI	RMSE	MAPE	RRMSE	NSE	IA
ANFIS	2.72	0.06	7.89	0.78	0.93	2.40	0.07	7.78	0.84	0.96
ANFIS-WOA	1.44	0.03	4.29	0.94	0.98	1.96	0.05	6.32	0.89	0.97
ANFIS-KHA	1.75	0.04	5.26	0.92	0.98	1.94	0.05	6.16	0.87	0.97
ANFIS-FA	1.71	0.04	5.25	0.92	0.98	2.07	0.05	6.41	0.86	0.97

performance from the viewpoint of under- or over-estimation of *SM*, ANFIS-WOA in under-estimation set and ANFIS-KHA in over-estimation set outperformed the other models, with *RMSEs* of 1.44 and 1.94, respectively. All models except ANFIS had lower errors in under-estimation set compared to the over-estimation set. Limitations of the present study are twofold. The first is *SM* simulation in the form of AI models, which, similar to all modeling attempts, may be accompanied by uncertainties. The structure of the ANFIS model is complex, and the number and type of its membership functions can add to this complexity. Metaheuristic optimization algorithms also suffer such disadvantages as parameter setting, high computational complexity, getting trapped in local optimums, and high running time. The second limitation is the constraints that affect the structures defined in this study. For example, the authors tried to estimate *SM* – which is one of the hard-to-measure soil parameters and only measured in rare and special conditions – using the main meteorological parameters which are measured at all stations together with a single soil parameter (soil temperature) as the inputs of AI models. Obviously, the use of other readily available soil parameters along with the parameters used in this study can lead to a better understanding of *SM* dynamics and thus to better *SM* estimates. The basic idea of the present work can be evaluated more rigorously in further studies by employing other AI models and optimization algorithms under various climates, and using remote sensing indices that indirectly indicate the *SM* status and can help in making more accurate *SM* estimates across larger areas.

Availability of data and materials

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

Authors' contributions

Conceptualization, Writing - review & editing, Software, Formal analysis [Roozbeh Moazenzadeh], Methodology, Writing original draft, Software [Babak Mohammadi], Data preparing, Methodology, Writing original draft [Mir Jafar Sadegh Safari], Methodology, Supervision [Kwok-wing Chau].

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Roozbeh Moazenzadeh http://orcid.org/0000-0002-1057-3801
Babak Mohammadi http://orcid.org/0000-0001-8427-5965
Mir Jafar Sadegh Safari http://orcid.org/0000-0003-0559-5261
Kwok-wing Chau http://orcid.org/0000-0001-6457-161X

References

Abbasi, F., Jacques, D., Simunek, J., Feyen, J., & Van Genuchten, M. T. (2003). Inverse estimation of soil hydraulic and solute transport parameters from transient field experiments: Heterogeneous soil. *American Society of Agricultural Engineers*, *46*(4), 1097–1111. https://doi.org/10.13031/2013.13961

Ahmad, S., Kalra, A., & Stephen, H. (2010). Estimating soil moisture using remote sensing data: A machine learning approach. *Advances in Water Resources*, *33*(1), 69–80. https://doi.org/10.1016/j.advwatres.2009.10.008

Altunkaynak, A., & Bizimana, H. (2020). Prediction of flood hazard map based on hybrid fuzzy geographic information system and its application for Ayamama watershed. *International Journal of Fuzzy Computation and Modelling*, *3*(1), 16–32. https://doi.org/10.1504/IJFCM.2020.106092

Amente, G., Baker, J. M., & Reece, F. C. (2000). Estimation of soil solution electrical conductivity from bulk soil electrical conductivity in sandy soils. *Soil Society of America Journal*, *64*(6), 1931–1939. https://doi.org/10.2136/sssaj2000.6461931x

Brandhorst, N., Erdal, D., & Neuweiler, I. (2017). Soil moisture prediction with the ensemble kalman filter: Handling uncertainty of soil hydraulic parameters. *Advances in Water Resources*, *110*, 360–370. https://doi.org/10.1016/j.advwatres.2017.10.022

Carranza, C., Nolet, C., Pezij, M., & Ploeg, M. V. D. (2021). Root zone soil moisture estimation with random forest. *Journal of Hydrology*, *593*, 125840. https://doi.org/10.1016/j.jhydrol.2020.125840

Chanasyk, D. S., & Neath, M. A. (1996). Field measurement of soil moisture using neutron probe. *Canadian Journal of Soil Science*, *76*(3), 317–323. https://doi.org/10.4141/cjss96-038

Coopersmith, E. J., Cosh, M. H., Bell, J. E., & Boyles, R. (2016). Using machine learning to produce near surface soil moisture estimates from deeper in situ records at U.S. Climate Reference Network (USCRN) locations: Analysis and applications to AMSR-E satellite validation. *Advances in Water Resources*, *98*, 122–131. https://doi.org/10.1016/j.advwatres.2016.10.007

Cuceloglu, G., Abbaspour, K. C., & Ozturk, I. (2017). Assessing the water-resources potential of Istanbul by using a soil and water assessment tool (SWAT) hydrological model. *Water*, *814*(9), 1–18. https://doi.org/10.3390/w9100814.

Deo, R. C., Kisi, O., & Singh, V. P. (2017). Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmospheric Research*, *184*, 149–175. https://doi.org/10.1016/j.atmosres.2016.10.004

Devi, S. A., & Vijayalakshmi, C. (2020). Bio inspired optimization algorithms in disaster. *Procedia Computer Science*, *172*, 176–180. https://doi.org/10.1016/j.procs.2020.05.027

Dobarco, M. R., Cousin, I., Le Bas, C., & Martin, M. P. (2019). Pedotransfer functions for predicting available water capacity in French soils, their applicability domain and associated uncertainty. *Geoderma*, *336*, 81–95. https://doi.org/10.1016/j.geoderma.2018.08.022

Dobriyal, P., Qureshi, A., Badola, R., & Hussain, S. A. (2012). A review of the methods available for estimating soil moisture and its implications for water resource management. *Journal of Hydrology*, *458–459*, 110–117. https://doi.org/10.1016/j.jhydrol.2012.06.021.

Ekmekcioğlu, Ö, Koc, K., & Özger, M. (2021). District based flood risk assessment in Istanbul using fuzzy analytical hierarchy process. *Stochastic Environmental Research and Risk Assessment*, *35*(3), 617–637. https://doi.org/10.1007/s00477-020-01924-8

Elliott, J., & Price, J. (2020). Comparison of soil hydraulic properties estimated from steady-state experiments and transient field observations through simulating soil moisture in regenerated sphagnum moss. *Journal of Hydrology*, *582*, 124489. https://doi.org/10.1016/j.jhydrol.2019.124489

Elshorbagy, A., & Parasuraman, K. (2008). On the relevance of using artificial neural networks for estimating soil moisture content. *Journal of Hydrology*, *362*(1–2), 1–18. https://doi.org/10.1016/j.jhydrol.2008.08.012

Ertekin, C., & Yaldiz, O. (2000). Comparison of some existing models for estimating global solar radiation for antalya (Turkey). *Energy Conversion and Management*, *41*(4), 311–330. https://doi.org/10.1016/S0196-8904(99)00127-2

Feng, L., Qin, W., Wang, L., Lin, A., & Zhang, M. (2018). Comparison of artificial intelligence and physical models for forecasting photosynthetically-active radiation. *Remote Sensing*, *10*(11), 1855. https://doi.org/10.3390/rs10111855

Gandomi, A. H., & Alavi, A. H. (2012). Krill herd: A new bio-inspired optimization algorithm. *Communications in Nonlinear Science and Numerical Simulation*, *17*(12), 4831–4845. https://doi.org/10.1016/j.cnsns.2012.05.010

Gao, X., Wu, P., Zhao, X., Zhang, B., Wang, J., & Shi, Y. (2013). Estimating the spatial means and variability of root-zone soil moisture in gullies using measurements from nearby uplands. *Journal of Hydrology*, *476*, 28–41. https://doi.org/10.1016/j.jhydrol.2012.10.030

Gaur, N., & Mohanty, B. P. (2013). Evolution of physical controls for soil moisture in humid and subhumid watersheds. *Water Resources Research*, *49*(3), 1244–1258. https://doi.org/10.1002/wrcr.20069

Ghorbani, M. A., Shamshirband, S., Haghi, D. Z., Azani, A., Bonakdari, H., & Ebtehaj, I. (2017). Application of firefly algorithm-based support vector machines for prediction of field capacity and permanent wilting point. *Soil & Tillage Research*, *172*, 32–38. https://doi.org/10.1016/j.still.2017.04.009

Gu, Z., Zhu, T., Jiao, X., Xu, J., & Qi, Z. (2021). Neural network soil moisture model for irrigation scheduling. *Computers and Electronics in Agriculture*, *180*, 105801. https://doi.org/10.1016/j.compag.2020.105801

Han, J., Mao, K., Xu, T., Guo, J., Zuo, Z., & Gao, C. (2018). A soil moisture estimation framework based on the CART algorithm and its application in China. *Journal of Hydrology*, *563*, 65–75. https://doi.org/10.1016/j.jhydrol.2018.05.051

Im, J., Park, S., Rhee, J., Baik, J., & Choi, M. (2016). Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. *Environmental Earth Sciences*, *75*(15), 1–19. https://doi.org/10.1007/s12665-016-5917-6

Jang, J.-S.R. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics*, *23*(3), 665–685. http://dx.doi.org/10.1109/21.256541

Javanshir, S., Bayat, H., & Gregory, A. S. (2020). Effect of free swelling index on improving estimation of the soil moisture retention curve by different methods. *Catena*, *189*, 104479. https://doi.org/10.1016/j.catena.2020.104479

Jimenez, A., Almeida, C., Junior, J., Morais, J., Almeida, B., & Andrade, F. (2019). Accuracy of capacitive sensors for estimating soil moisture in northeastern Brazil. *Soil & Tillage Research*, *195*, 104413. https://doi.org/10.1016/j.still.2019.104413

Jung, H. C., Getirana, A., Arsenault, K. R., Kumar, S., & Maigary, I. (2019). Improving surface soil moisture estimates in West Africa through GRACE data assimilation. *Journal of Hydrology*, *575*, 192–201. https://doi.org/10.1016/j.jhydrol.2019.05.042

Karandish, F., & Simnek, J. (2016). A comparison of numerical and machine-learning modeling of soil water content with limited input data. *Journal of Hydrology*, *543*, 892–909. https://doi.org/10.1016/j.jhydrol.2016.11.007

Khan, W. A., Hamadneh, N. N., Tilahun, S. L., & Ngnotchouye, J. M. (2016). A review and comparative study of firefly algorithm and its modified versions. *Optimization Algorithms-Methods and Applications*, 281–313. https://doi.org/10.5772/62472.

Kolassa, J., Reichle, R. H., Lio, Q., Alemohammad, S. H., Gentine, P., Aidae, K., Asanuma, J., Bircher, S., Caldwell, T., Colliander, A., Cosh, M., Holifield Collins, C., Jackson, T.J., Martínez-Fernández, J., McNairn, H., Pacheco, A., Thibeaultj, M., & Walker, J.P. (2018). Estimating surface soil moisture from SMAP observations using a neural network technique. *Remote Sensing of Environment*, *204*, 43–59. https://doi.org/10.1016/j.rse.2017.10.045

Koley, S., & Jeganathan, C. (2020). Estimation and evaluation of high spatial resolution surface soil moisture using multi-sensor multi-resolution approach. *Geoderma*, *378*, 114618. https://doi.org/10.1016/j.geoderma.2020.114618

Kornelsen, K. C., & Coulibaly, P. (2014). Root-zone soil moisture estimation using data-driven methods. *Water Resource Research*, *50*(4), 2946–2962. https://doi.org/10.1002/2013WR014127

Lee, Y., Jung, C., & Kim, S. (2019). Spatial distribution of soil moisture estimates using a multiple linear regression model and Korean geostationary satellite (COMS) data. *Agricultural Water Management*, *213*, 580–593. https://doi.org/10.1016/j.agwat.2018.09.004

Lees, K. J., Artz, R. R. E., Chandler, D., Aspinall, T., Boulton, C. A., Buxton, J., Cowie, N. R., & Lenton, T. M. (2021). Using remote sensing to assess peatland resilience by estimating soil surface moisture and drought recovery. *Science of the Total Environment*, *761*, 143312. https://doi.org/10.1016/j.scitotenv.2020.143312

Leski, J., & Czogala, E. (1999). A new artificial neural network based fuzzy inference system with moving consequents in if-then rules and selected applications. *Fuzzy Sets and Systems*, *108*(3), 289–297. https://doi.org/10.1016/S0165-0114(97)00314-X

Liu, Y., Jing, W., Wang, Q., & Xia, X. (2020). Generating high-resolution daily soil moisture by using spatial downscaling techniques: A comparison of six machine learning algorithms. *Advances in Water Resources*, *141*, 103601. https://doi.org/10.1016/j.advwatres.2020.103601

Maroufpoor, S., Maroufpoor, E., Bozorg-haddad, O., Shiri, J., & Yaseen, Z. M. (2019). Soil moisture simulation using hybrid artificial intelligent model: Hybridization of adaptive neuro fuzzy inference system with grey wolf optimizer algorithm. *Journal of Hydrology*, *575*, 544–556. https://doi.org/10.1016/j.jhydrol.2019.05.045

McDonough, K. R., Hutchinson, S. L., Hutchinson, J. M. S., Case, J. L., & Rahmani, V. (2018). Validation and assessment of SPoRT-LIS surface soil moisture estimates for water resources management applications. *Journal of Hydrology*, *566*, 43–54. https://doi.org/10.1016/j.jhydrol.2018.09.007

Mirjalili, S., & Lewis, A. (2016). The Whale optimization algorithm. *Advances in Engineering Software*, *95*, 51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008

Mishra, V., Cruise, J. F., & Mecikalski, J. R. (2021). Assimilation of coupled microwave/thermal infrared soil moisture profiles into a crop model for robust maize yield estimates over southeast United States. *European Journal of Agronomy*, *123*, 126208. https://doi.org/10.1016/j.eja.2020.126208

Moazenzadeh, R., & Mohammadi, B. (2019). Assessment of bio-inspired metaheuristic optimization algorithms for estimating soil temperature. *Geoderma*, *353*, 152–171. https://doi.org/10.1016/j.geoderma.2019.06.028

Moazenzadeh, R., Mohammadi, B., Duan, Z., & Delghandi, M. (2022). Improving generalistion capability of artificial intelligence-based solar radiation estimator models using a bio-inspired optimisation algorithm and multi-model approach. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-021-17852-1

Mohammadi, B., Ahmadi, F., Mehdizadeh, S., Guan, Y., Pham, Q. B., Linh, N. T. T., & Tri, D. Q. (2020). Developing novel robust models to improve the accuracy of daily streamflow modeling. *Water Resources Management*, *34*(10), 3387–3409. https://doi.org/10.1007/s11269-020-02619-z

Mohammadi, B., Guan, Y., Moazenzadeh, R., & Safari, M. J. S. (2021). Implementation of hybrid particle swarm optimization-differential evolution algorithms coupled with multi-layer perceptron for suspended sediment load estimation. *Catena*, *198*, 105024. https://doi.org/10.1016/j.catena.2020.105024

Moran, M. S., Peters-Lidard, C. D., Watts, J. M., & Mc Elroy, S. (2004). Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. *Canadian Journal of Remote Sensing*, *30*(5), 805–826. https://doi.org/10.5589/m04-043

Ning, G. Y., & Cao, D. Q. (2021). Improved Whale optimization algorithm for solving constrained optimization problems. *Discrete Dynamics in Nature and Society*. https://doi.org/10.1155/2021/8832251.

Pan, F., & Nieswiadomy, M. (2016). Estimating daily root-zone soil moisture in snow-dominated regions using an empirical soil moisture diagnostic equation. *Journal of Hydrology*, *542*, 938–952. https://doi.org/10.1016/j.jhydrol.2016.09.063

Pan, F., Nieswiadomy, M., & Qian, S. (2015). Application of a soil moisture diagnostic equation for estimating root-zone soil moisture in arid and semi-arid regions. *Journal of Hydrology*, *524*, 296–310. https://doi.org/10.1016/j.jhydrol.2015.02.044

Pardossi, A., Incrocci, L., Incrocci, G., Malorgio, F., Battista, P., Bacci, L., Rapi, B., Marzialetti, P., Hemming, J., & Balendonck, J. (2009). Root zone sensors for irrigation management in intensive agriculture. *Sensors*, *9*(4), 2809–2835. https://doi.org/10.3390/s90402809

Penghui, L., Ewees, A. A., Beyaztas, B. H., Qi, C., Salih, S. Q., Al-Ansari, N., Bhagat, S. K., Yaseen, Z. M., & Singh, V. P. (2020). Metaheuristic optimization algorithms hybridized With artificial intelligence model for soil temperature prediction: Novel model. *IEEE Access*, *8*, 51884–51904. https://doi.org/10.1109/ACCESS.2020.2979822

Pezij, M., Augustijn, D. C. M., Hendriks, D. M. D., Weerts, A. H., Hummel, S., Velde, R., & Hulscher, S. J. M. H. (2019). State updating of root zone soil moisture estimates of an unsaturated zone metamodel for operational water resources management. *Journal of Hydrology X*, *4*, 100040. https://doi.org/10.1016/j.hydroa.2019.100040

Pires, L., Bacchi, O. S., & Reichardt, K. (2005). Soil water retention curve determined by gamma ray beam attenuation. *Soil & Tillage Research*, *82*(1), 89–97. https://doi.org/10.1016/j.still.2004.06.003

Qin, W., Wang, L., Lin, A., Zhang, M., Xia, X., Hu, B., & Niu, Z. (2018). Comparison of deterministic and data-driven models for solar radiation estimation in China. *Renewable and Sustainable Energy Reviews*, *81*, 579–594. https://doi.org/10.1016/j.rser.2017.08.037

Qin, W., Wang, L., Zhang, M., & Niu, Z. (2019). First effort for constructing a high density photosynthetically active radiation dataset during 1961-2014 in China. *Journal of Climate*, *32*(10), 2761–2780. https://doi.org/10.1175/JCLI-D-18-0590.1

Qiu, T., Wang, L., Lu, Y., Zhang, M., Qin, W., Wang, S., & Wang, L. (2022). Potential assessment of photovoltaic power generation in China. *Renewable and Sustainable Energy Reviews*, *154*, 111900. https://doi.org/10.1016/j.rser.2021.111900

Ridler, M. E., Sandholt, I., Butts, M., Lerer, S., Mougin, E., Timouk, F., Kergoat, L., & Madsen, H. (2012). Calibrating a soil-vegetation-atmosphere transfer model with remote sensing estimates of surface temperature and soil surface moisture in a semi-arid environment. *Journal of Hydrology*, *436–437*, 1–12. https://doi.org/10.1016/j.jhydrol.2012.01.047.

Sadeghi, M., Gao, L., Ebtehaj, A., Wigneron, J. P., Crow, W. T., Reager, J. T., & Warrick, A. W. (2020). Retrieving global surface soil moisture from GRACE satellite gravity data. *Journal of Hydrology*, *584*, 124717. https://doi.org/10.1016/j.jhydrol.2020.124717

Sadeghi, M., Ghahraman, B., Davary, K., Hasheminia, M., & Reichardt, K. (2011). Scaling to generalize a single solution of richards' equation for soil water redistribution. Scientia agricola, (piracicaba. *Braz*, *68*(5), 582–591. https://doi.org/10.1590/S0103-90162011000500011.

Sadeghi, M., & Jones, S. B. (2012). Scaled solutions to coupled soil-water flow and solute transport during the

redistribution process. *Vadose Zone Journal*, *11*(4), 4. https://doi.org/10.2136/vzj2012.0023

Sadeghi, M., Jones, S. B., & Philpot, W. D. (2015). A linear physically-based model for remote sensing of soil moisture using short wave infrared bands. *Remote Sensing of Environment*, *164*, 66–76. https://doi.org/10.1016/j.rse.2015.04.007

Schnur, M. T., Xie, H., & Wang, X. (2010). Estimating root zone soil moisture at distant sites using MODIS NDVI and EVI in a semi-arid region of southwestern USA. *Ecological Informatics*, *5*(5), 400–409. https://doi.org/10.1016/j.ecoinf.2010.05.001

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-science Reviews*, *99*(3-4), 125–161. https://doi.org/10.1016/j.earscirev.2010.02.004

Shahabfar, A., Ghulam, A., & Eitzinger, J. (2012). Drought monitoring in Iran using the perpendicular drought indices. *International Journal of Applied Earth Observation and Geoinformation*, *18*, 119–127. https://doi.org/10.1016/j.jag.2012.01.011

Shin, Y., Mohanty, B. P., & Ines, A. V. M. (2018). Development of non-parametric evolutionary algorithm for predicting soil moisture dynamics. *Journal of Hydrology*, *564*, 208–221. https://doi.org/10.1016/j.jhydrol.2018.07.003

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., O'Neill, P., Islam, T., Gupta, M., & Dai, Q. (2015). Performance evaluation of WRF-Noah land surface model estimated soil moisture for hydrological application: Synergistic evaluation using SMOS retrieved soil moisture. *Journal of Hydrology*, *529*, 200–212. https://doi.org/10.1016/j.jhydrol.2015.07.041

Tsang, S. W., & Jim, C. Y. (2016). Applying artificial intelligence modeling to optimize Green roof irrigation. *Energy and Buildings*, *127*, 360–369. https://doi.org/10.1016/j.enbuild.2016.06.005

Vaheddoost, B., Guan, Y., & Mohammadi, B. (2020). Application of hybrid ANN-whale optimization model in evaluation of the field capacity and the permanent wilting point of the soils. *Environmental Science and Pollution Research*, *27*(12), 1–11. https://doi.org/10.1007/s11356-020-07868-4

Vergopolan, N., Chaney, N. W., Beck, H. E., Pan, M., Sheffield, J., Chan, S., & Wood, E. F. (2020). Combining hyper-resolution land surface modeling with SMAP brightness temperatures to obtain 30-m soil moisture

estimates. *Remote Sensing of Environment*, *242*, 111740. https://doi.org/10.1016/j.rse.2020.111740

Wang, L., Kisi, O., Zounemat-Kermani, M., Salazar, G. A., Zhu, Z., & Gong, W. (2016). Solar radiation prediction using different techniques: Model evaluation and comparison. *Renewable and Sustainable Energy Reviews*, *61*, 384–397. https://doi.org/10.1016/j.rser.2016.04.024

Wang, L., Lu, Y., Zou, L., Feng, L., Wei, J., Qin, W., & Niu, Z. (2019). Prediction of diffuse solar radiation based on multiple variables in China. *Renewable and Sustainable Energy Reviews*, *103*, 151–216. https://doi.org/10.1016/j.rser.2018.12.029

Wang, Q., Li, P., Pu, Z., & Chen, X. (2011). Calibration and validation of salt-resistant hyperspectral indices for estimating soil moisture in arid land. *Journal of Hydrology*, *408*(3-4), 276–285. https://doi.org/10.1016/j.jhydrol.2011.08.012

Yang, X. (2009). Firefly algorithms for multimodal optimization. In O. Watanabe & T. Zeugmann (Eds.), *Stochastic algorithms: Foundations and applications* (pp. 169–178). Berlin: Springer Berlin-Heidelberg. https://doi.org/10.1007/978-3-642-04944-6_14

Yang, X., Yu, Y., & Li, M. (2019). Estimating soil moisture content using laboratory spectral data. *Journal of Forestry Research*, *30*(3), 1073–1080. https://doi.org/10.1007/s11676-018-0633-6

Zhang, D., Zhang, W., Huang, W., Hong, Z., & Meng, L. (2017). Upscaling of surface soil moisture using a deep learning model with viirs rdr. *ISPRS International Journal of Geo-Information*, *6*(5), 130. https://doi.org/10.3390/ijgi6050130

Zhang, S., Shao, M., & Li, D. (2017). Prediction of soil moisture scarcity using sequential Gaussian simulation in an arid region of China. *Geoderma*, *295*, 119–128. https://doi.org/10.1016/j.geoderma.2017.02.003

Zhou, W., Xu, Z., Ross, D., Dignan, J., Fan, Y., Huang, Y., Wang, G., Bagtzoglou, A. C., Lei, Y., & Li, B. (2019). Towards water-saving irrigation methodology: Field test of soil moisture profiling using flat thin mm-sized soil moisture sensors (MSMSs). *Sensors and Actuators B: Chemical*, *298*, 126857. https://doi.org/10.1016/j.snb.2019.126857

Zormand, S., Jafari, R., & Koupaei, S. S. (2017). Assessment of PDI, MPDI and TVDI drought indices derived from MODIS aqua/terra level 1B data in natural lands. *Natural Hazards*, *86*(2), 757–777. https://doi.org/10.1007/s11069-016-2715-0