

JaNarX: Smart Soil Moisture Prediction using Jackal Optimiser Enabled Machine Learning Model in IoT Coupled Agricultural Applications

Seema Jitendra Patil^{1,2*} and B. Ankayarkanni¹

¹Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, 600119 Chennai, Tamil Nadu, India

²School of Computer Science and Engineering, Dr. Vishwanath Karad, MIT World Peace University, 411038 Pune, Maharashtra, India

ABSTRACT

Soil moisture can be used to provide predictive information important for precision farming, irrigation management, and pollution monitoring. Nonetheless, existing approaches often struggle to handle non-linear temporal relationships, environmental uncertainties, and poor real-time integration of sensor data, thereby degrading prediction performance. To overcome these problems, in this article, we introduce JaNarX, a soil moisture prediction framework (SMPF) based on IoT and the NARX model, optimised with the Jackal Apis Optimisation (JAO) algorithm. The approach utilises time-series radar satellite variables, meteorological terms, and continuous in situ sensor data to better capture dynamic changes of soil moisture. The Wazihub Soil Moisture Dataset (WSMD), which is the aggregation of multi-sensor environmental data measured in in-field conditions from real agricultural fields in Senegal, was used for training and validation. JAO was used to speed up convergence and tune NARX hyperparameters for solving the local minima problem, and increase the generalisation performance of the model. Experimental results in MAE, MSE, and RMSE evaluations had MAE = 1.53, MSE = 2.89, and RMSE = 1.70, which outperformed the state-of-the-art base-models such as SVM, LSTM, GLM XGBR, and traditional NARX network. The simulation results confirm that the proposed IOT-based, JAO-optimised NARX structure can achieve more accurate prediction than the traditional method and has good stability and computational efficiency. This paper presents a

scalable, high-accuracy model for the prediction of soil moisture, suitable to enable real-time planning decisions for precision irrigation and sustainable water management in agriculture.

ARTICLE INFO

Article history:

Received: 16 October 2025

Accepted: 12 December 2025

Published: 17 April 2026

DOI: <https://doi.org/10.47836/pjst.34.2.06>

E-mail addresses:

seema.patil@mitwpu.edu.in (Seema Jitendra Patil)

ankayarkanni.cse@sathyabama.ac.in (B. Ankayarkanni)

* Corresponding author

Keywords: Internet of things, Jackal APIs optimisation, machine learning, precision agriculture, soil moisture

INTRODUCTION

Soil moisture is an important factor in the hydrologic, ecological, and agricultural systems, affecting plant growth and nutrient uptake as well as crop yield (Liu et al. 2014). Maintaining adequate soil moisture content is especially crucial in precision agriculture because both drought and waterlogging conditions will adversely affect the development of plants (Cai et al., 2019). In addition to being the frontiers of technology, like IoT devices and automated irrigation systems or sensor networks that promote these developments (Magallanes-Quintanar et al., 2022), smart agriculture has advanced significantly in recent years with soil monitoring capabilities and real-time field assessment. Yet, despite these efforts, soil moisture still poses a difficult prediction and characterisation problem because it exhibits strong spatial and temporal variability, complex environmental interactions, and dependence on a wide variety of physical properties and climate parameters.

Conventional soil moisture estimation means, including the hydrological model, the soil water balance model and empirical equations, need local calibration over a large area or are computationally expensive. Furthermore, they fail to consider nonlinear interactions that dominate between the soil and environment (Liu et al., 2020; Li et al., 2022). Remote sensing-derived methods are suitable for large-scale monitoring but tend to have low temporal resolution and cannot be used in real time (Gao et al., 2021). For example, machine learning and deep learning models have been shown to enhance the predictive accuracy of soil moisture and temperature, but are still limited by issues such as a lack of convergence stability in training, poor generalisation between distinct and heterogeneous soil conditions, sensitivity to noisy IoT sensor-based data, and difficulty with modelling long-term temporal dependencies (Karthikeyan et al., 2021). Moreover, most of the currently available IoT-based frameworks are unable to successfully assimilate multiple environmental parameters, such as air temperature, humidity, wind factors, and atmospheric pressure, into a unified predictive model such that performance is inconsistent in time and space (Economou et al., 2007; Fotis et al., 2007), and there is a high level of uncertainty.

In the literature, many methods have been proposed to predict soil moisture, ranging from hydrological and land surface models (Ming et al., 2022), statistical models such as ARMA or empirical water balance equations (Liu et al., 2020), inversion models based on remote sensing (Gao et al., 2021), and machine learning-type or deep learning-based algorithms like SVM, GLM, XGBoost, LSTM, etc. (Granata et al., 2022; Karthikeyan et al., 2021; Nguyen et al., 2022). While these methods have revealed some merits, they are still facing several limitations: physics-based models require parameter tuning for different crops and suffer from computational insufficiency; The remote sensing approaches cannot operate in real time; ML/DL models fail to learn *nonlinear* temporal relationships rather than inherent convergence problems and a mismatch of patterns between trained data and transfer data across different agricultural conditions. Optimisation-based methods (e.g., BES-enhanced SVM) show the benefits of parameter optimisation, but they do not

model multi-source IoT variables jointly or capture the intricate nonlinear hydrological dynamics.

In comparison, the proposed JaNarX model makes use of memory-intensive NARX network architecture and a hybrid JAO optimisation algorithm that leads to a more efficient nonlinear modelling, enhanced convergence dynamics, less localised minima consequences, and more trustworthy IoT-inspired multi-parameter predictions. This model hybridisation, which not only improves the predictability of the JaNarX model but also provides a much stronger alternative to currently available soil moisture prediction methods.

The literature shows a need for a flexible, robust, and data-driven model that can predict nonlinear soil moisture dynamics and that works well by using networked sensor fusion from diverse sources, as well as fulfilling the necessity of keeping its predictions stable under an uncertain environment. NARX is an effective choice for time-series hydrological forecasting because of its memory structure and ability to capture complex temporal dynamics. Nevertheless, NARX models are highly dependent on learning processes, which consist of trial-and-error setting and tuning of parameters. If not properly trained, NARX may return suboptimal models (e.g., converging to local minima and performing poorly in predictions), indicating that efficient optimisation techniques are necessary for improving the training process of NARX.

In order to facilitate these issues, JaNarX, an IoT-based soil moisture prediction model using Jackal Apis Optimisation (JAO) for the optimisation process, is proposed. JAO integrates the cooperation strategies of the Artificial Bee Colony (ABC) algorithm with the chasing behaviour of Coyote Optimisation (CO), and it has good convergence, global search ability, and parameter setting. The motivation behind incorporating JAO into NARX is to help overcome the limitations of the nonlinear model and make it robust under environmental uncertainty, which leads to better prediction accuracy.

The literature gap is systematically covered through the following research questions:

- RQ1 : How can the accuracy of soil moisture prediction be improved in a nonlinear, uncertain, and time-varying natural environment?
- RQ2 : Whether an IoT-based NARX model trained with JAO can reduce the forecast errors over conventional ML, DL, and non-optimised NARX approaches?
- RQ3 : Do predictions considering temporal multi-parameter IoT sensor data (e.g., soil moisture, air temperature, humidity, atmospheric pressure, wind) perform better than those without integrating them?
- RQ4 : How much is the convergence improved, model parameters optimised, and local minima avoided by the JAO method in the JaNarX framework?
- RQ5 : How well does the proposed JaNarX model perform compared to state of the art models (SVM, LSTM, XGBR, GLM, DT, and LR) and classical NARX in terms of delay-based prediction assessment?

The objective of this study is to propose an IoT-enabled, JAO-tuned NARX model for real-time accurate estimation of SM. Specifically, the study intends to:

- RO1 : Develop a combined model, capturing nonlinear hydrological dynamics more precisely
- RO2 : Combine multi-source environmental sensor data into an integrated forecasting system (IPS).
- RO3 : Use JAO to train NARX and thereby obtain better fitting results and convergence performance and benchmark the proposed approach under different delay patterns versus the state-of-the-art solution.

Several methods for forecasting Soil Moisture have been explored, such as hydrological models (Pal & Olsson, 2004), statistical techniques (Al-Yaari et al., 2016), and empirical water balance equations. Mechanistic hydrological models offer conceptual understanding but are limited by large computational requirements and uncertainty in environmental forcings (Li et al., 2022). They enlarge spatial coverage, although with a compromised temporal resolution and susceptibility to atmospheric noise in remote sensing-derived models (Hegazi et al., 2023). Machine learning techniques, such as SVM, Gaussian regression, XGBoost and LSTM achieve better prediction accuracy (Granata et al., 2022; Nguyen et al., 2022), but they still face the challenge of modelling long-term dependency and performance deterioration for noisy or missing sensor data. A BES-optimised SVM (Huang, 2023) is also a good template to refer to, but it still cannot ensure the stable convergence and well-extended generalisation over wide-ranging farming conditions.

Accordingly, there is a serious demand for an optimisation-based predictive framework IoT-supported that can properly represent the nonlinear features of time-series soil moisture and guarantee convergence to potential users in agriculture. While current soil moisture prediction algorithms, such as hydrological models, statistical time-series methods, ML regressors, and DL architectures, can provide good insights into soil water phenomena, they suffer from the common problems in dealing with nonlinear temporal dynamics, heterogeneous environmental inputs, and noisy IoT sensor readings. Machine learning (ML) and DL approaches (e.g., SVM, XGboost, LSTM) frequently suffer from convergence instability and struggle in capturing long-term dependencies (Granata et al., 2022; Karthikeyan et al., 2021), which precludes them from being generalisable over different soil types and climates (Li et al., 2022). While the optimisation-enhanced methods, such as BES-enhanced SVM, have certainly relieved the parameter adjustment focus, there lacks one real time and consistent answer for multi-sensor fusion in IoT systems.

The JaNarX model overcomes these issues by integrating the NARX structure, which has a large memory requirement and therefore poor nonlinear modelling characteristics, but is complemented with the hybrid Jackal Apis Optimisation (JAO) technique that is

capable of bringing better convergence properties and is immune to temporal shifting and local minima effects. This hybrid system enhances prediction performance without sacrificing real-time application.

This paper has numerous original contributions in terms of soil moisture prediction, agriculture for the IoT-driven analytics and model-based optimisation Hybrid JaNarX forecasting model, by combining an IoT-based NARX system with the Jackal Apis Optimisation (JAO) algorithm to deal with nonlinearity as well as uncertainty in the soil moisture dynamics. The JAO algorithm, a hybrid metaheuristic approach that includes both artificial bees foraging behaviour and cooperative coyote hunting strategies, is utilised to enhance the hyperparameter tuning process while preventing convergence to local minima as well as improving model training speed. Integration of multiparameter IoT sensor data, soil humidity, air temperature and humidity, atmospheric pressure, wind speed and direction, gusts in a single predictive model to improve temporal accuracy and environment adaptability. Performance testing of delay-based derivatives for varying population size and time-lags, with JaNarX exhibiting better generalisation stability and predictability when compared to nine state-of-the-art (LR/DT/GLM/SVM/LSTM/XGBR/NARX/CANAX/ANAX) models. Extensive statistical validation (MAE, MSE, RMSE, and R^2) proves that prediction error is reduced as well as the reliability on the real Wazihub Soil Moisture dataset. Demonstrated practical real-time implementation showing that the JaNarX–JAO framework for precision irrigation can be deployed for better water management, and to account for real-time field variabilities in large-scale farms.

LITERATURE REVIEW

Modern agriculture depends critically on effective management of soil moisture, which influences irrigation efficiency, crop yield, and resource economy. Manual sampling and laboratory analysis are among the time-consuming, labour-intensive traditional soil moisture monitoring techniques that lack real-time adaptation. Furthermore, although being extensively used, sensor-based methods often suffer from data loss, misinterpretation, and errors resulting from environmental conditions (Rezk et al., 2021).

Different research has looked at cutting-edge techniques for irrigation optimisation and soil moisture prediction. Using deep learning methods to examine moisture variance (Chandrappa et al., 2023) presented a soil moisture modelling approach enabled by IoT and machine learning. But the method was limited in that it couldn't make real-time irrigation changes. To schedule irrigation plans, Gao et al. (2021) also developed an IoT-based environmental data collecting system connected with Bidirectional Long Short-Term Memory (Bid-LSTM) networks. Although the study needed more improvements in transfer learning to fit various regional conditions, its expected performance was good.

As noted by Ming et al. (2022), indicated that missing value is still a common problem in soil moisture data, and we also found many missing values during the exploratory analysis. While processing these gaps added computational burdens, it also enhanced the general trustworthiness of the dataset. Similarly, Ainiwaer et al. (2020) demonstrated that soil hyperspectral data can be complex, where noise or redundant information could have a negative impact on model predictive capability. Pandey et al. (2021) presented TOTRAM, a soil moisture retrieval method using the land surface temperature (LST) and vegetation indices (VIs), but this technique is computationally demanding and susceptible to different environmental conditions. Similarly, Huang (2023) further enhanced predictive ability by modifying the support vector machine (SVM) with Bald Eagle Search (BES) in tuning for hyperparameters; however, this technique had poor generalisability across agricultural zones.

To break through the bottlenecks of conventional methods, several researchers have introduced machine learning 17-18, deep learning 19-20, and IoT 21-based techniques, which can provide better efficiency and flexibility for soil moisture prediction and irrigation management. For example, Nguyen et al. (2022) adopted a hybrid XGBoost Regression (XGBR) model with the use of GA to select features, which proved to be capable of an accurate prediction of soil moisture variations at different locations around the world. These examples highlight the AI-driven support that precision agriculture could enjoy thanks to such systems, with specific intelligence and guidance for farmers, municipalities, or water management organisations.

Furthermore, emerging as good substitutes for conventional irrigation systems are IoT-driven precision irrigation models. (Abioye et al., 2021) presented an IoT-based model using predictive analytics to improve real-time soil parameter monitoring and optimise irrigation schedules. To guarantee greenhouse applications, though, a greater connection with adaptive control systems was required. Using real-time soil moisture, temperature, and humidity data, Podder et al. (2021) created an IoT-enabled Smart AgroTech system that dynamically changes irrigation. Although it underlined sensor coverage limits that need optimisation, the study confirmed the dependability of IoT-driven agriculture.

Dealing with missing data (Cordeiro et al., 2022) showed how deep learning models coupled with KNN data imputation methods improve soil moisture predictions while reducing discrepancies. Likewise, Bid-LSTM networks, as used by (Gao et al., 2021), have shown promise in long-term irrigation scheduling; nevertheless, more study is required to increase computational efficiency and adaptability over several farming environments. Overcoming a few restrictions of conventional sensor-based approaches (Ainiwaer et al., 2020), also proposed a multi-source remote sensing method integrating hyperspectral and multispectral data to improve the spatial resolution of soil moisture predictions.

By means of analysis of important influencing elements along with topography, climate, soil type, and vegetation, the proposed JaNarX approach seeks to overcome the

shortcomings of past models by refining soil moisture predictions. JaNarX greatly reduces overfitting problems, speeds convergence, and keeps models from being caught in local minima by using the JAO method for optimisation. Moreover, sophisticated methods of data preparation promise the removal of extraneous elements, thereby enhancing the predictive power and dependability of soil moisture estimations. These methods taken together position JaNarX as a highly scalable, flexible, and efficient precision agricultural solution.

This work presents an advanced soil moisture prediction model based on the JaNarX model to overcome these challenges. The model considered IoT-generated soil moisture data, sent to a cloud server for storage and additional processing. Embedded in a controller, a NARX machine learning-based model forecasts soil moisture levels, enabling effective water management and lowering of waste. To optimise the weights of the NARX model, a hybrid optimisation method combining the hunting behaviour of jackals with the communication strategies of bee optimisation is applied: the JAO method. This optimisation reduces overfitting, improves convergence speed, and helps the model not to get caught in local optima. The proposed system guarantees high detection accuracy, enhanced prediction dependability, and a general increase in agricultural water-use efficiency by combining the JAO-optimised JaNarX model.

MATERIAL AND METHODS

The proposed methodology aims to develop a predictive model for soil moisture to optimise irrigation scheduling based on meteorological variations. It comprises four main components, which are: sensor networks, data aggregation, predictive modelling, and decision support. Wireless sensor nodes collect real-time data on soil moisture, temperature, and humidity, which is transmitted to a central server and preprocessed for accuracy, as shown in Figure 1.

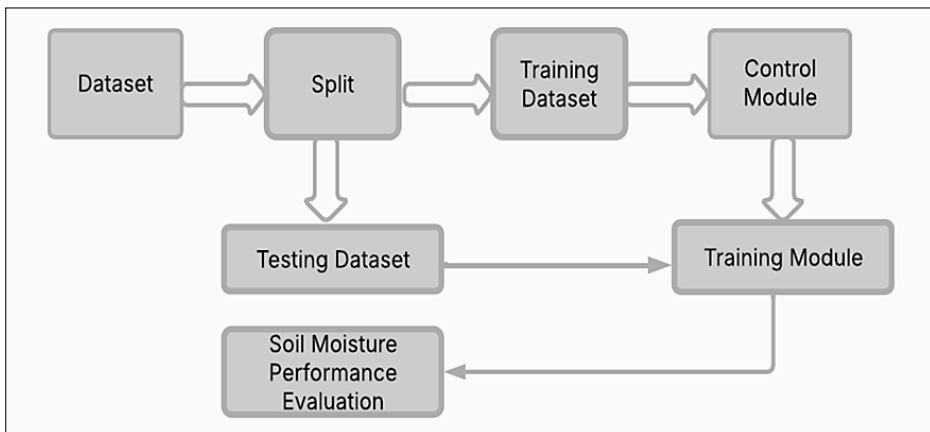


Figure 1. Schematic representation of the proposed soil moisture prediction model

Data Collection

This study utilised the Wazihub soil moisture dataset (Bera et al., 2024), collected over four months from four agricultural plots in Senegal cultivated with peanuts and maize. IoT sensors were deployed in each plot to continuously monitor soil moisture dynamics, enabling precise tracking of hydration levels across different crop environments. To maintain consistent environmental conditions, the plots were positioned within one met of each other. The dataset comprises key parameters, including soil humidity, air humidity, atmospheric pressure, wind speed, air temperature, wind gust, and wind direction-critical variables for comprehensive soil moisture assessment.

IoT-based Soil Moisture Prediction using Jackal Apis Optimised Nonlinear Autoregressive Exogenous (JaNarX) Model

JaNarX is a recurrent dynamic network effective at simulating the nonlinear time series with feedback connections. In addition, the NARX network provides time-delayed feedback on the model's output for the modelling of a nonlinear dynamical system. Considerably, it is intriguing to make use of the NARX neural network's memory function by utilising the historical values of predicted or genuine time series in order to maximise its performance for nonlinear time series prediction. The vector w_r forms the input to the soil moisture prediction, which is forwarded to the Nonlinear Autoregressive Exogenous model to predict the moisture present in the soil. JaNarX networks are an effective class of models well-suited for this kind of challenge when modelling nonlinear systems, especially time series. The mathematical formulation of the NARX model is:- of a nonlinear dynamical system. Considerably, it is intriguing to make use of the NARX neural network's memory function by utilising the historical values of predicted or genuine time series in order to maximise its performance for nonlinear time series prediction. The vector w_r forms the input to the soil moisture prediction, which is forwarded to the Nonlinear Autoregressive Exogenous model to predict the moisture present in the soil. JaNarX networks are an effective class of models well-suited for this kind of challenge when modelling nonlinear systems, especially time series. The mathematical formulation of the NARX model is:

$$\hat{z}(t) = f \left(\begin{array}{c} z(t-1), z(t-2), \dots, z(t-n_z), \\ w_r(t-1), w_r(t-2), \dots, w_r(t-n_w) + e(t) \end{array} \right) \quad [1]$$

The weight of the connection between the i^{th} hidden neuron and the input neuron $w(t-m)$ is indicated by V_{im} a_j is the bias of the i^{th} hidden neuron; f_h is the hidden layer activation function; f_y is the hidden transfer and output function. Further, V_{iq} is the

connection weight between the i^{th} hidden neuron and the output feedback. Combining the output of the hidden layer allows one to get the last prediction as follows:

$$\hat{Z}_j(t) = f_y \left[\sum_{i=1}^{n_h} V_{ji} h_i(t) + a_j \right] \tag{2}$$

where n_h is the number of hidden neurons; V_{ij} is the connection weight between the i^{th} hidden neuron and the j^{th} expected output; a_j is the bias of the j^{th} expected output; and f_y is the activation function for the output layer.

Figure 2 shows the Ja NarX network architecture. Increasing hidden neurons helps the model to perform better, so improving learning capacity and generalisation. By changing weights and biases, Jackal Apis Optimisation (JAO) minimises overfitting and increases prediction accuracy, so refining the network.

$$h_i(t) = f_h \left[\sum_{m=0}^{n_w} V_{im} w(t-m) + \sum_{q=0}^{n_z} V_{iq} z(t-q) + a_i \right] \tag{3}$$

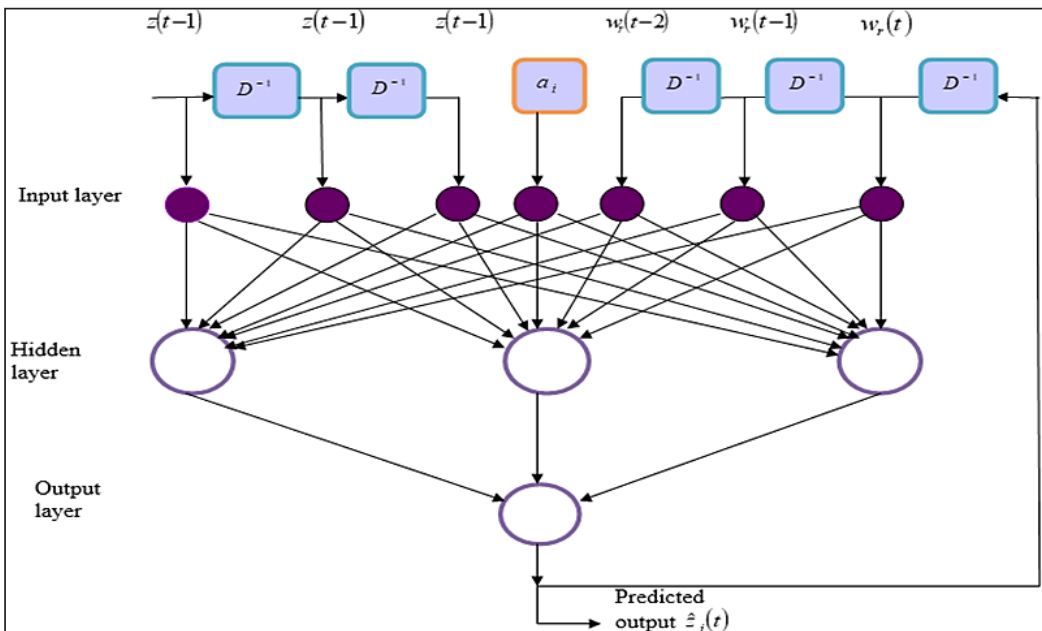


Figure 2. Architecture of the developed JaNarX network

JAO adjusts the hyperparameters of the JaNarX model such that they maximise prediction performance. It hybridises the Artificial Bee Colony (ABC) and Coyote Optimisation Algorithm (COA). Inspired by bee foraging techniques and jackal hunting behaviour, JAO advances quick convergence and effective exploration.

Initialisation of Jackal Coyote Population

Initialising the JAO algorithm as $g_t^{h,\alpha} = [V, \alpha]$ results in a weight and a bias of the JaNarX classifier. In this optimisation, t_e represents the jackal coyote population overall and t_f represents the total number of possible solutions. Based on the social conditions (g) of the jackal coyotes, the jackal apis optimisation is mathematically stated, and these jackal coyotes are considered as the solutions in the search space. As a result, the following equation depicts the social situation of the g^{th} jackal h^{th} coyote pack at that particular moment in time α is given as:

$$g_{tc}^{h,\alpha} = c^{low} + \beta_c (c^{up} - c^{low}) \quad [4]$$

Here, c stands for the dimension of the jackal coyote's search space, while low and up are used to describe the lower and upper bounds of the search space, respectively. β_c stands for a valid random number that is generated between $[0,1]$ and assigned using uniform probability.

Determine Fitness Function Value

The cost of the system is decided by the fitness solution, which is dependent on the social situation determined by the jackal coyote's behaviour in adapting to its environment. Further, the fitness of the model is based on the lower MSE of the solution and is defined as follows:

$$fit(g_t^h) = \min(MSE(g_t^h)) \quad [5]$$

The population of jackal coyotes is initialised as the first phase in the jackal bee optimisation process. According to the following representation, each jackal coyote's social position is initially allocated at random inside the search area.

Exploration Phase

The jackal coyotes are randomly assigned to a group, but occasionally they quit their pack and join another pack or go solitary. The number of members in the pack determines

whether this jackal coyote is expelled, and the possibility that it is expelled from the packs is based on the probabilities listed below:

$$J_t = 0.005 t_f \quad [6]$$

There are typically 14 individuals in each pack; the probability given in the equation above should be higher than for the coyote solution $t_f \leq \sqrt{200}$. This exemplifies the coyote's social or communicative behaviour, which facilitates effective coyote interaction and aids in the development of a global search strategy.

Leader Selection Phase

Only one leader, Jackal Coyote (L) that is most suited to the environment is chosen for this optimisation, and this leader jackal coyote is regarded as the most suited to the environment. The leader of the coyotes or the alpha coyote is considered as the best-adapted coyote, and it can be mathematically given as follows: The leader jackal coyote of the h^{th} pack at a given instant α is indicated in a minimisation problem by:

$$G^{t,\alpha} = \{g_t^{h,\alpha} | \arg_{t=\{1,2,\dots,t_f\}} \min n(g_t^{h,\alpha}) \quad [7]$$

Communicative Phase

The jackal coyote's swarm intelligence demonstrates the ability of the coyote species to share its social conditions with all the packs, illuminating their organisational behaviour and highlighting the importance of each jackal coyote posse in preserving social behaviour. The algorithm considers the life cycle of coyotes, including birth and death. Specifically, the birth of coyotes is a combination of the social behaviour of two parents that are selected randomly within the search space, along with the environmental factor. The pack's whole data is computed, and their social tendency behaviour is represented as:

$$f_c^{h,\alpha} = \begin{cases} U_{\left(\frac{t_f+1}{2}\right),c}^{h,\alpha}; & ; \text{if } t_f \text{ is odd} \\ U_{\left(\frac{t_f}{2}\right),c}^{h,\alpha} + U_{\left(\frac{t_f+1}{2}\right),c}^{h,\alpha} & ; \text{else} \end{cases} \quad [8]$$

Here, $J^{(h,\alpha)}$ stands for the ranked social state of every jackal coyote t in the pack in the range $[1, g]$ of dimension c . The cultural tendency of a single pack is determined using the average social conditions of all coyotes in that pack.

Balancing Probability

The age of jackal coyotes is determined using two biological features, such as birth and death. The social interactions of the other two jackal coyote wolves and the birth of the additional wolves are provided by:

$$Z_c^{h,\alpha} = \begin{cases} g_{b_1,c}^{h,\alpha}; r_c < J_{scatter} & \text{or } c = c_1 \\ g_{b_2,c}^{h,\alpha}; r_c \geq J_{scatter} + J_{associate} & \text{or } c = c_2 \\ B_c; & \text{else} \end{cases} \quad [9]$$

In this case, Z stands for the pub of the jackal coyote among the random coyotes b_1 and b_2 in the random dimensions c_1 and c_2 . The decision variable's random number is selected, and it is symbolised by the value B_c . In addition, r_c designates a random number in the range $[0,1]$ That was produced using a uniform probability distribution. The dispersion and associative probability assigned to jackal coyotes are represented by, $J_{scatter}$ and $J_{associate}$, and these probabilities aid in maintaining social communicative behaviour. Following is a mathematical equation that describes the probabilities:

$$J_{scatter} = \frac{1}{g} \quad [10]$$

$$J_{associate} = \left(1 - \frac{1-J_{scatter}}{2}\right) \quad [11]$$

Exploitation Phase

If a solution does not show any improvement and the number of trials without change exceeds a predefined threshold, known as the boundary, it is considered to have stagnated. In this case, the scout bees will abandon the current solution. After that, the scout bees begin a random search process to explore new potential solutions in the search space. This mechanism helps maintain diversity in the population and prevents the algorithm from getting stuck in local optima:

$$g_{new}^{h,\alpha} = g_{tc}^{h,\alpha} + \delta_{tc} (g_{tc}^{h,\alpha} - g_{tc}^{h,\beta}) \quad [12]$$

where δ_{tc} is a random number in the range $[-1,1]$

Enhanced Convergence

Coyotes can locate their prey and surround them after recognising their location. The jackal coyote generally directs the hunt. However, we are unsure of the exact location of the ideal (prey) in an abstract search space. To mathematically replicate the coyote's hunting behaviour, we suppose that the jackal coyote has a stronger awareness of the likely location of prey. In this scenario, to improve the attacking strategy, the communication behaviour of the bee is hybridised with the jackal coyote; hence jackal bee communicates with the group members and performs hunting in a cooperative manner, which improves the hunting speed and provides faster convergence. The hybridised equation for the solution update is expressed as follows:

$$g_t^{h+1} = 0.5Z_c^{h,\alpha} + 0.5g_{new}^{h,\alpha} \quad [13]$$

$$g_t^{h+1} = 0.5 \left[\begin{array}{l} g_{b1,c}^{h,\alpha}; r_c < J_{scatter} \text{ or } c = c_1 \\ g_{b2}^{h,\alpha}; r_c \geq J_{scatter} + J_{associate} \text{ or } c = c_2 \\ B_c; \text{ else} \end{array} \right] + 0.5 [g_{tc}^{h,\alpha} + \delta_{tc} (g_{tc}^{h,\alpha} - g_{tc}^{h,\beta})] \quad [14]$$

In this case, the hunting operation is directed by the coefficient vectors. \vec{B} and \vec{P} , the determined vector \vec{M} and the vector of the coyote's positions \vec{Q} , which are used to construct a new position for the jackal coyote. Another way to state it is, μ estimates the locations of the prey and the other wolves before updating their new locations arbitrarily close to the prey.

The probability of successfully obtaining food is determined by the fading parameter. u , which is a positive integer number, the distance t between the bee and the other colony members, the targeted communication range, g denotes the neighbour bee node.

Termination

The global solution is declared through the final updated position of the solution, and the verification process ends when the iteration reaches the condition ($t > t_{max}$).

Algorithm: Pseudo Code for Proposed Jackal Apis Optimisation

S.NO	Pseudo Code
1	Input: $g = (g_1, g_2, \dots, g_E)$ Output: g_t^{h+1}
2	Determine fitness: $fit(g_t^h) = \min(MSE(g_t^h))$
3	Exploration phase: $J_t = 0.005t_f$
4	Selection of leader: $G^{h,\alpha} = \{g_t^{h,\alpha} \arg \min_{t=\{1,2,\dots,t_f\}} n(g_t^{h,\alpha})\}$
5	Communicative behaviour: $f_c^{h,\alpha} = \begin{cases} U_{\left(\frac{t_f+1}{2}\right),c}^{h,\alpha}; & \text{if } t_f \text{ is odd} \\ U_{\left(\frac{t_f}{2}\right),c}^{h,\alpha}; & \text{else} \end{cases}$
6	Balancing probability: $Z_c^{h,\alpha} = \begin{cases} g_{b1,c}^{h,\alpha}; & r_c < J_{scatter} \text{ or } c = c_1 \\ g_{b2}^{h,\alpha}; & r_c \geq J_{scatter} + J_{associate} \text{ or } c = c_2 \\ B_c; & \text{else} \end{cases}$
7	Probability: $J_{associate} = \left(1 - \frac{1-J_{scatter}}{2}\right)$ Exploitation Phase: $g_{new}^{h,\alpha} = g_{tc}^{h,\alpha} + \delta_{tc} (g_{tc}^{h,\alpha} - g_{tc}^{h,\beta})$
8	Solution update with Enhanced convergence: $g_t^{h+1} = 0.5 \left[\begin{cases} g_{b1,c}^{h,\alpha}; & r_c < J_{scatter} \text{ or } c = c_1 \\ g_{b2}^{h,\alpha}; & r_c \geq J_{scatter} + J_{associate} \text{ or } c = c_2 \\ B_c; & \text{else} \end{cases} \right] + 0.5 [g_{tc}^{h,\alpha} + \delta_{tc} (g_{tc}^{h,\alpha} - g_{tc}^{h,\beta})]$
9	Termination

RESULTS

Model Evaluation Overview

The IoT-SM-based JaNarX model was evaluated using the Wazihub Soil Moisture Database, with performance assessed through MAE (Mean Absolute Error), MSE (Mean Squared Error), RMSE (Root Mean Squared Error), and R^2 . These metrics collectively capture the magnitude of prediction errors, their variability, and the correlation between predicted and observed values. The results indicate that JaNarX provides reliable predictions with high accuracy and stability across different parameter settings.

Delay Analysis

Delay profile analysis was conducted with a delay of 25 and population sizes ranging from 10 to 50, shown in Figure 3. The results show that increasing population size significantly enhances model performance. Specifically, MAE decreased from 2.87 to 1.70, and MSE decreased from 8.22 to 2.89, reflecting a substantial reduction in prediction errors. Similar improvements were observed for RMSE, while R^2 values remained consistently around 0.93, indicating a strong correlation between predicted and observed soil moisture and confirming the stability of the JaNarX model.

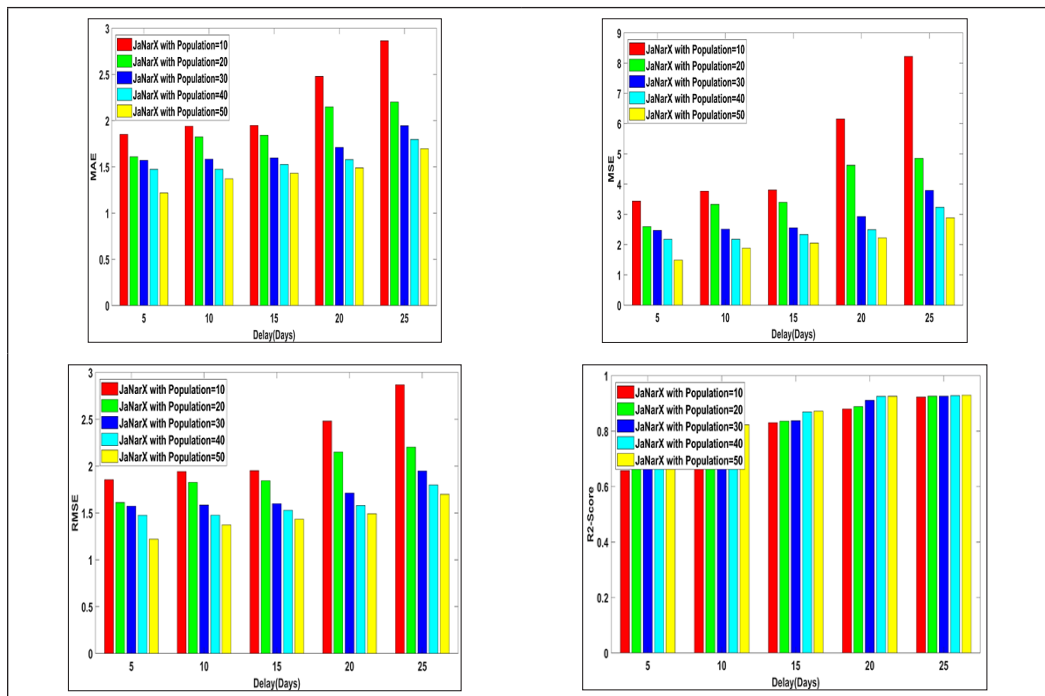


Figure 3. Performance analysis of JaNarX using the Wazihub soil moisture database with MAE, MSE, RMSE, and R^2 metrics

Statistical Evaluation

Statistical analysis included computation of means, variances, and best-case performance (Table 1). The lower mean error and variance observed for JaNarX suggest greater reliability and minimal fluctuation under test conditions compared to benchmark models. These results validate that JaNarX generates soil moisture estimates that are more stable and accurate than competing methods.

Table 1
Statistical summary of Wazihub soil moisture database performance metrics

Models	Best	Mean	Variance	Best	Mean	Variance	Best	Mean	Variance
LR (Rani et al., 2022)	6.99	4.56	3.91	60.35	30.52	587.27	7.77	5.07	4.82
DT (Nikjoo et al., 1997)	5.70	3.40	1.34	40.18	15.90	147.62	6.34	3.77	1.65
XGBR (Nguyen et al., 2022)	4.33	3.01	0.45	25.52	12.22	44.62	5.05	3.39	0.70
GLM (Pekár & Brabec, 2018)	2.96	2.62	0.05	10.86	8.54	2.07	3.29	2.91	0.06
SVM (Huang, 2023)	2.84	2.36	0.09	10.00	6.99	3.16	3.16	2.62	0.11
LSTM (Datta & Faroughi, 2023)	2.84	2.44	0.07	10.00	7.47	2.74	3.16	2.72	0.09
NARX (Siegelmann et al., 1997)	2.84	2.27	0.11	10.00	6.51	3.81	3.16	2.53	0.13
CNaX (Lin et al., 1996)	2.84	2.15	0.16	10.00	5.90	5.29	3.16	2.39	0.20
AnaX (Omkar et al., 2020)	2.77	2.07	0.16	9.51	5.51	4.82	3.08	2.31	0.19
JaNarX	1.53	1.30	0.02	2.89	2.11	0.21	1.70	1.44	0.02

Loss Curve Behaviour

The training loss curve (Figure 4) demonstrates a smooth and continuous decrease in error over epochs, indicating effective learning behaviour. This validates the capability of the JAO algorithm to optimise the NARX network efficiently. The declining loss curve also confirms that the model successfully captures temporal dependencies and reduces uncertainty in predictions.

Comparative Evaluation

Comparative performance under a delay of 25 is shown in Figure 5, where JaNarX is compared against LR, DT, XGBR, GLM, SVM-LSTM, NARX, CNAX, and ANAX

models. JaNarX achieves the lowest error metrics with MAE = 1.53, MSE = 2.88, and RMSE = 1.69, outperforming all other models by a significant margin. Additionally, $R^2 = 0.93$, the highest among the compared models, demonstrates superior ability to capture the nonlinear behaviour of soil moisture dynamics and stronger predictive skill. These results confirm that JaNarX improves both generalisation and accuracy compared to classical ML, DL, and NARX-based approaches.

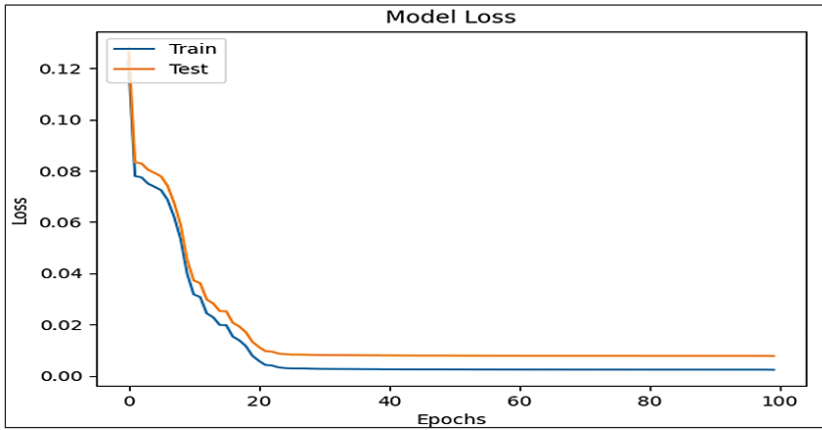


Figure 4. Training loss curve of JaNarX on the Wazihub soil moisture database

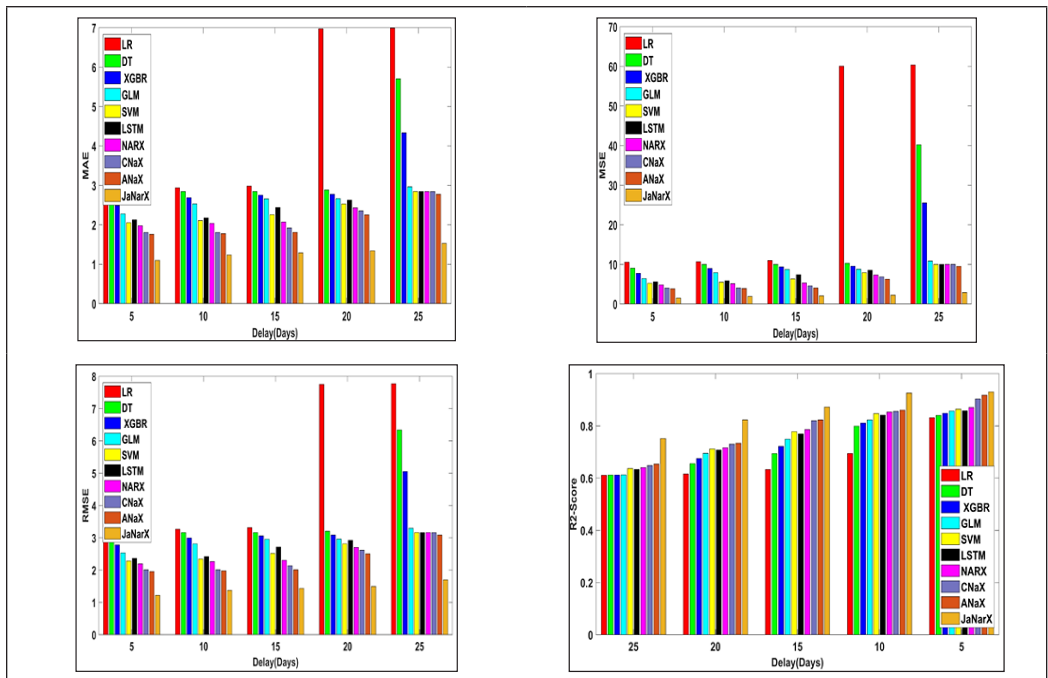


Figure 5. Comparative evaluation of JaNarX and benchmark models on the Wazihub soil moisture database with MAE, MSE, RMSE, and R^2 metrics

DISCUSSION

The results shown in this study indicate that the JaNarX model significantly increases the quality of soil moisture forecasts. Through multi-parameter IoT sensor data fusion, the model can effectively reflect various environmental effects. JaNarX being used with the NARX structure that has a large amount of memory, can learn complex temporal dependencies that traditional models cannot consider. Additionally, the consistently high R^2 , together with decreases in MAE, MSE and RMSE for all training population sizes, again underlines the model's great generalisation ability across diverse scenarios.

The faster behaviour of the loss descent curve also emphasises the convergence-enhancing effectiveness of the JAO algorithm. Smooth and reasonable convergence of training loss verifies that JAO can avoid local minima effectively and allow for a stable optimisation. When compared to traditional hydrological models and statistical predictors, as well as deep learning methods, JaNarX outperforms for robustness and reliability, in particular given noisy IoT sensor inputs and nonlinear temporal trends in soil moisture.

In summary, quantitative comparisons demonstrate that JaNarX achieves better performance than previous methods in term of accuracy and robustness. Earlier models like LR and DT suffer from high prediction errors because they cannot capture nonlinear interactions. Some of the state-of-the-art deep learning-based models, like LSTM, offer superior temporal learning, but their convergence is often not guaranteed, and convergence to noise is a challenge, hence delivering far inferior performance than JaNarX. It is observed that, although CNAX and ANaX are more complex models in terms of model structure, they do not lead to much of a gain in error reduction or correlation.

These results also have potential practical applications for precision agriculture. The improved predictability of JaNarX may aid real-time irrigation scheduling and soil-water management, helping to reduce the possibility of over- or under-application. Moreover, the stability of the model under differential context supports its integrability in a landscape-scale, real-time decision support system.

CONCLUSION

To tackle the long-standing issues of soil moisture prediction, this article has provided an innovative JaNarX model for the NARX framework in the IoT domain, empowered by the JAO algorithm that is particularly designed to minimise the impacts of uncertainty and non-linear perturbations and temporally varying responses. Multi-parameter time-series data was utilised effectively to decrease the ambiguity between input and output data values and represent dynamic hydraulic features for practical agricultural use. The inclusion of climatic variables like soil humidity, air temperature and humidity, atmospheric pressure and wind has allowed JaNarX to provide a complete picture of the variability in soil moisture throughout agricultural regions.

Because of the continuous monitoring through IoT sensors, such a model was able to capture seasonal patterns and aid in data-driven decisions on when to irrigate, how much

water should be applied and how much efficiency will result from it. Due to the adaptive parameter search of JAO and the nonlinear structure of NARX, it allowed for faster convergence, decreased training loss and more stable predictions. Validation on Wazihub Soil Moisture Database showed it to outperform versus benchmark models (MAE = 1.53, MSE = 2.89, RMSE = 1.70).

JaNarX has some limitations despite its predictive power. The model is, however, dependent on the temporal continuity of environmental data, and its generalisation could suffer in the case of sudden or extreme events, which are not aligned to learned temporal patterns, like abrupt rainfall. The short-term reliability of the forecast can also be decreased by noise spikes or missing sensor measurements. Furthermore, predictions based on IoT devices developed in a particular geographic location can have reduced utility in areas with different soil or climate and would need to be recalibrated or retrained.

Possible future directions of the JaNarX framework could be hybrid deep learning architectures that merge NARX's temporal modelling with the feature extraction capability of LSTM, GRU or transformer-based attention mechanisms. These strategies could improve stress tolerance to rapid changes in the environment and field adaptation. Data assimilation or physics-based primitives can also further enhance robustness and promote immediate reaction to malfunctions of sensors. Generalising these findings in different climate, land types, and soil profiles will lend to application in precision agriculture/smart farming ecosystems.

ACKNOWLEDGEMENT

Experiments in this paper are performed on data from the Wazihub Soil Moisture Prediction Challenge, by Zindi. The authors thank Wazihub, Zindi, Microsoft and Université Gaston Berger (Senegal) for the access to this Soil Moisture dataset used in this study. The Wazihub initiative is funded by the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 780229, which allowed making this dataset publicly available for research purposes.

LIST OF ABBREVIATIONS

XGBR	:	XGBoost regressor
GLM	:	Generalised linear model
SVM-LSTM	:	Support vector machine- long short-term memory
NARX	:	Nonlinear AutoRegressive with eXogenous inputs
MAE	:	Mean absolute error
MSE	:	Mean square error
RMSE	:	Root mean squared error
ABC	:	Artificial bee colony

REFERENCES

- Abioye, E. A., Mohd Ali, M. A., Ahmad, M. S., *et al.* (2021). IoT-based monitoring and data-driven modelling of drip irrigation system for mustard leaf cultivation experiment. *Information Processing in Agriculture*, 8(2), 270-283. <https://doi.org/10.1016/j.inpa.2020.06.002>
- Ainiwaer, M., Ding, J., Kasim, N., Wang, J., & Wang, J. (2020). Regional scale soil moisture content estimation based on multi-source remote sensing parameters. *International Journal of Remote Sensing*, 41(9), 3346-3367. <https://doi.org/10.1080/01431161.2019.1701723>
- Bera, S., Dey, T., Mukherjee, A., & De, D. (2024). FLAG: Federated learning for sustainable irrigation in agriculture 5.0. *IEEE Transactions on Consumer Electronics*. Advance online publication. <https://doi.org/10.1109/TCE.2024.3370373>
- Cai, Y., Zheng, W., Zhang, X., Zhangzhong, L., & Xue, X. (2019). Research on soil moisture prediction model based on deep learning. *PLOS ONE*, 14(4), Article e0214508. <https://doi.org/10.1371/journal.pone.0214508>
- Chandrappa, V. Y., Ray, B., Ashwatha, N., & Shresth, P. (2023). Spatiotemporal modelling to predict soil moisture for sustainable smart irrigation. *Internet of Things*, 21, Article 100671. <https://doi.org/10.1016/j.iot.2022.100671>
- Cordeiro, M., Rosário, D., Gomes, D., *et al.* (2022). Towards smart farming: Fog-enabled intelligent irrigation system using deep neural networks. *Future Generation Computer Systems*, 129, 115-124. <https://doi.org/10.1016/j.future.2021.11.029>
- Datta, P., & Faroughi, S. A. (2023). A multihead LSTM technique for prognostic prediction of soil moisture. *Geoderma*, 433, Article 116452. <https://doi.org/10.1016/j.geoderma.2023.116452>
- Ekonomou, L., Fotis, G., Maris, T., & Liatsis, P. (2007). Estimation of the electromagnetic field radiating by electrostatic discharges using artificial neural networks. *Simulation Modelling Practice and Theory*, 15(9), 1089-1102. <https://doi.org/10.1016/j.simpat.2007.07.003>
- Fotis, G., Ekonomou, L., Maris, T., & Liatsis, P. (2007). Development of an artificial neural network software tool for the assessment of the electromagnetic field radiating by electrostatic discharges. *IET Science, Measurement & Technology*, 1(5), 261-269. <https://doi.org/10.1049/iet-smt:20060137>
- Gao, P., Li, H., Wang, Z., Zhao, Y., Sun, W., & Zhang, Y. (2021). Improved soil moisture and electrical conductivity prediction of citrus orchards based on IoT using deep bidirectional LSTM. *Agriculture*, 11(7), Article 635. <https://doi.org/10.3390/agriculture11070635>
- Granata, F., Di Nunno, F., Najafzadeh, M., & Demir, I. (2022). A stacked machine learning algorithm for multi-step ahead prediction of soil moisture. *Hydrology*, 10(1), Article 1. <https://doi.org/10.3390/hydrology10010001>
- Hegazi, E. H., Samak, A. A., Yang, L., Huang, R., & Huang, J. (2023). Prediction of soil moisture content from Sentinel-2 images using convolutional neural network (CNN). *Agronomy*, 13(3), Article 656. <https://doi.org/10.3390/agronomy13030656>
- Huang, Y. (2023). Improved SVM-based soil-moisture-content prediction model for tea plantation. *Plants*, 12(12), Article 2309. <https://doi.org/10.3390/plants12122309>

- Karthikeyan, L., & Mishra, A. K. (2021). Multi-layer high-resolution soil moisture estimation using machine learning over the United States. *Remote Sensing of Environment*, 266, Article 112706. <https://doi.org/10.1016/j.rse.2021.112706>
- Li, Q., Zhu, Y., Shangguan, W., Wang, X., Li, L., & Yu, F. (2022). An attention-aware LSTM model for soil moisture and soil temperature prediction. *Geoderma*, 409, Article 115651. <https://doi.org/10.1016/j.geoderma.2021.115651>
- Lin, T., Horne, B. G., Tino, P., & Giles, C. L. (1996). Learning long-term dependencies in NARX recurrent neural networks. *IEEE Transactions on Neural Networks*, 7(6), 1329-1338. <https://doi.org/10.1109/72.548162>
- Liu, M., Huang, C., Wang, L., Zhang, Y., & Luo, X. (2020). Short-term soil moisture forecasting via Gaussian process regression with sample selection. *Water*, 12(11), Article 3085. <https://doi.org/10.3390/w12113085>
- Liu, Y., Mei, L., & Ooi, S. K. (2014). Prediction of soil moisture based on extreme learning machine for an apple orchard. In *2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems (CCIS)* (pp. 400-404). IEEE. <https://doi.org/10.1109/CCIS.2014.7175760>
- Magallanes-Quintanar, R., Galván-Tejada, C. E., Galván-Tejada, J. I., Méndez-Gallegos, S. d. J., García-Domínguez, A., & Gamboa-Rosales, H. (2022). NARX neural network models for prediction of the standardised precipitation index in Central Mexico. *Atmosphere*, 13(8), Article 1254. <https://doi.org/10.3390/atmos13081254>
- Ming, W., Zhang, Y., Chen, L., He, L., & Liu, Z. (2022). A hybrid triple collocation-deep learning approach for improving soil moisture estimation from satellite and model-based data. *Remote Sensing*, 14(7), Article 1744. <https://doi.org/10.3390/rs14071744>
- Nguyen, T. T., Do, H. H., Nguyen, D. H., Tran, Q. H., & Pham, T. T. (2022). A low-cost approach for soil moisture prediction using multi-sensor data and machine learning algorithms. *Science of the Total Environment*, 833, Article 155066. <https://doi.org/10.1016/j.scitotenv.2022.155066>
- Nikjoo, P. O. N., Goodhead, D. T., & Terrissol, H. M. (1997). Computational modelling of low-energy electron-induced DNA damage by early physical and chemical events. *International Journal of Radiation Biology*, 71(5), 467-483. <https://doi.org/10.1080/095530097143509>
- Omkar, S., Mudigere, D., Senthilnath, J., & Kumar, M. V. (2020). Identification of helicopter dynamics based on flight data using nature-inspired techniques. In *Deep learning and neural networks: Concepts, methodologies, tools, and applications* (pp. 257-273). IGI Global. <https://doi.org/10.4018/978-1-7998-0945-6.ch012>
- Pandey, R., Goswami, S., Sarup, J., & Matin, S. (2021). The thermal-optical trapezoid model-based soil moisture estimation using Landsat-8 data. *Modelling Earth Systems and Environment*, 7(2), 1029-1037. <https://doi.org/10.1007/s40808-020-00975-8>
- Pekár, S., & Brabec, M. (2018). Generalised estimating equations: A pragmatic and flexible approach to the marginal GLM modelling of correlated data in the behavioural sciences. *Ethology*, 124(2), 86-93. <https://doi.org/10.1111/eth.12709>
- Podder, A. K., Chaki, R., Ghosal, R., Chakraborty, S., & Ghosh, S. (2021). IoT-based smart agrotech system for verification of urban farming parameters. *Microprocessors and Microsystems*, 82, Article 104025. <https://doi.org/10.1016/j.micpro.2021.104025>

- Rani, A., Kumar, N., Kumar, J., & Sinha, N. K. (2022). Machine learning for soil moisture assessment. In *Deep learning for sustainable agriculture* (pp. 143-168). Elsevier. <https://doi.org/10.1016/B978-0-323-85214-2.00001-X>
- Rezk, N. G., Hemdan, E. E.-D., Attia, A.-F., El-Sayed, A., & El-Rashidy, M. A. (2021). An efficient IoT-based smart farming system using machine learning algorithms. *Multimedia Tools and Applications*, *80*, 773-797. <https://doi.org/10.1007/s11042-020-09740-6>
- Siegelmann, H. T., Horne, B. G., & Giles, C. L. (1997). Computational capabilities of recurrent NARX neural networks. *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, *27*(2), 208-215. <https://doi.org/10.1109/3477.558801>