Nat. Volatiles & Essent. Oils, 2021; 8(5): 13184 -13205

NVEO Natural Volatiles & Essential Oils

Soil Moisture Quantity Prediction using Optimized Deep Learning Supported model for Sustainable cultivation of Groundnut plant

Sivasankaran S[1], Dr.K. Jagan Mohan[2], Dr.G. Mohammed Nazer[3]

[1]*Research Scholar, Department of Computer Science and Engineering, Annamalai University, Chidambaram, Tamilnadu, India.*

[2]*Associate Professor, Department of Information Technology, Annamalai University, Chidambaram, Tamilnadu, India.*

[3]*Professor and Principal, Department of Computer Science, RAAK Arts and Science College, Villupuram, Tamilnadu, India.*

Abstract

In the hydrological cycle, soil moisture is the predominant component for regulating the evapotranspiration, vegetation production and runoff. The process of soil moisture is directly essential for farmers involved in sustainable agriculture and their related social and economic activities. This soil moisture is the potential soil indicators that aids in determining the conditions of early water deficit in the earlier stage, decision making about crop planning, policy making by government, food security conditions and uncertainty in crop yield. In specific, the soil moisture prediction of groundnut plant is essential as it is the principal crop cultivated by significant number of farmers in the semi-arid regions of Southern India. In this paper, Generative Adversarial Networks-based Deep Learning Model with Sailfish Optimization Algorithm (GAN-DLMSOA) for predicting the quantity of soil moisture. This GAN-DLMSOA is proposed with the potential characteristics of GAN that aids in significant prediction through the comprehensive exploration of exhaustive parameters associated with the problem from the dataset to render accurate prediction process. In specific, Sailfish Optimization Algorithm (SOA) is included to perform remarkable feature optimization that contributes towards accurate soil moisture prediction. SOA is also used to train the GANs with fast convergence degree for accomplishing the task of prediction and facilitate precise results. The data perturbation-based stability analysis of the proposed GAN-DLMSOA is also conducted with different weather conditions to perceive the stability of the incorporated model. The proposed GAN-DLMSOA confirmed with a mean RMSE value of 0.04065 and a maximized accuracy level of 99.66% compared to all the baseline soil prediction schemes.

2. Introduction

2.1 Preamble

The agricultural industry completely relies on raw agricultural supplies with good quality. The yield of crops, vegetables and fruits considered as the principal resource of the agricultural industry plays and anchor role in impacting the economy associated with it. From the viewpoint of ever-increasing food demand, the total yield of crops is determined as the primary key factor that depicts the phenomenal growth of the agricultural industry. In this context, different factors that contributes towards predominant agriculture production need to be clearly understood for potential decision making that could introduce maximized crop yield. The two significant associated with agriculture production are genetic factors and environment factors. The genetic factors determine the growth and productivity of the crops in the presence of some specific genes that inculcate maximized crop yield. On the other hand, environmental features represent the role of radiant energy, moisture, and temperature with its influence over the crop production. In specific, soil moisture pertaining to environmental factor is identified as a potential factor that is highly essential for healthy growth of crops that results in maximized yield. Soil moisture improves the crops' probability of fixing large amount of natural nutrients that in turn introduces its maximized growth. Inadequate amount of soil moisture negatively impacts the cells growth and results in growth reduction of crops during cultivation. However, the growth of crops is also impacted at some specific life stages when the soil moisture is high than the required level. Further, the soil moisture necessitated by each crop for achieving their growth is also different depending on their variety, season, and geographical location in which it is cultivated. Moreover, soil moisture referred as one of the core factors in hydrological cycles and agricultural production need to be precisely predicted to achieve optimal management and rational use of water resources. Thus, accurate prediction of soil moisture is essential for improving the agricultural productions. However, the prediction of soil moisture necessitates the extraction of complex meteorological and structural characteristics which is identified to be a Herculean task. Further the design and development of an ideal mathematical model for predicting soil moisture is highly difficult. Furthermore, the majority of the existing soil moisture prediction models possesses the limitations of prediction performance, multi-feature processing capability, generalization, and prediction accuracy. Thus, these limitations of the existing soil moisture prediction techniques need to be improved through the employment of a suitable and reliable deep learning model to the significant level to achieve maximum prediction accuracy.

2.2 Motivation

In general, mathematical model developed for predicting soil moisture need to measure physical quantities that could be possibly derived from the environment and must derive formulas to depict the actual association between the involved parameters. In this context, methods of empirical modeling have evolved as the predominant development model that has the potentiality of extracting contextual-aware physical quantities related to soil moisture, due to the advancement in computer modelling. This empirical model generally used to select a right model, calibrate the physical quantities through the selected model and validate it to verify its significance in soil moisture prediction. These models need to measure relevant data that comprises of predetermined input parameters and the required output parameters which need to be modelled. In specific, input parameters are selected by experience with the fact that a at least a minimum degree of correlation need to be maintained between them. Majority of the core soil moisture prediction schemes contributed to the literature mainly utilized neural networks, linear regression, machine learning and empirical formulas for the construction of prediction models. At this juncture, deep learning-based soil moisture prediction models is determined to be more potent and reliable in achieving prediction accuracy.

The rapid development of artificial intelligence from the past decades innovated the option of utilizing a potential deep learning models that possesses the merits of multiple hidden layer structure to enhance multiple feature data and big data fitting capability with improved classification degree. The deep learning models are identified to include strong computing capability on par with the classical neural networks and machine learning models. Moreover, the success of deep learning in the fields of stock price predictions, search engines and image recognition motivated the option of utilizing its benefits towards the process of soil moisture prediction. This deep learning models include the potentiality of handling the extremely complex and non-linear characteristics of soil parameters that could be feasibly extracted from a specific region of interest. It is also potent in overcoming the issues of lower prediction accuracy, Thus, deep learning model-based soil moisture prediction approach that aids in constructing and optimizing the features of the soil with its powerful data processing capabilities which can step towards the achievement of precise soil moisture prediction is essential.

2.3 Related Work

Regression Network-based deep learning scheme was proposed for accurate prediction of soil moisture by handling the problem of multiple feature processing capability, generation, and improved prediction performance [11]. This RN-based deep learning model considered the area of Beijing as the object of research with the capability of big data fitting that aided in constructing a ideal prediction model. It adopted Taylor diagram for determining the predictive variables that helped in selecting the parameters of meteorology to facilitate the estimation of weights during moisture prediction. It played an anchor role in clarifying the associated between features considered for soil moisture prediction. It also investigated the time series related to the predictive variables during the integration of dataset. The results of RNDLM confirmed better input characteristics with maximized accuracy due to its inherent generalization capability and good data fitting potentiality. A Dynamic Neural Network-based soil moisture prediction model was proposed by generation and training of one day ahead to enhance degree of accuracy [12]. This prediction model included the quantification parameters of climatic essentialities, past soil moisture content and volumetric soil moisture. This model evaluation was performed in most of sites based on the field data whose value of R^2 is greater than 0.94. It was determined to improve the possibility of generating soil moisture prediction with robustness independent to the data associated with the sites that are used for model training. The simulation of this predictive model is conducted using AQUACROP to determine the level of soil moisture required for growing potato. The results of this model confirmed 20-46% improvement in water saving on par with the other rule-based systems.

Temporal Graphs-based Soil Moisture Prediction (TGSMP) scheme was proposed with the merits of semi-supervised learning for preventing the limitations of the conventional approaches to improve the degree of accuracy [13]. This temporal graph approach used dynamic graph neural network for establishing the dependency between the associated locations on a region considered for predicting soil moisture. It is the first soil moisture prediction approach that utilized temporal graphs since this category of graph modelling was utilized in the field of information and social networks. It adopted dynamic GNN-based t-problem of graph structure to derive the process of end-to-end learning that predicts soil moisture over a region in the entire time duration. It also updated the structure of graph periodically depending on the parameters that could be contextually derived from a specific region of study. The results of this TGSMP confirmed better input characteristics with maximized accuracy. An integrated machine learning techniques-based soil moisture prediction strategy was proposed with Naïve Bayes, PCA, SVM and linear regression to minimize average RMSE independent to the parameters considered for prediction [14]. This integrated machine learning approach was tested over the datasets extracted from thirteen districts of West Bengal, India in which the plants of cauliflower, paddy, mustard, and potato are cultivated. The performance of this integrated machine learning strategy confirmed excellent F1-Score, Accuracy and minimized RMSE compared to the existing approaches.

An integrated KNN, SVM and Wenner's four electrodes method-based soil prediction approach was proposed with the measurement of electrical resistivity which depends on the water retaining capability in the clay and clayey slit regions [15]. It attained soil moisture prediction using dataset that comprised of 162 sample points from which 11 points and 151 points is used for testing and training, respectively. It included the merits of KNN for clustering only the predominant features that helps in better prediction process. The results also confirmed its potential in minimizing the MAD and MAE, compared to the baseline used for analysis. An enhanced deep learning model based on improved SVM was proposed for evaluating the quality of meteorological factors and observing the complex soil feature to estimate the degree of moisture [16]. The input to this model was the collected real time samples determined from the local area sensor network used for investigation. It was proposed with the ability of big data fitting that aids in better prediction of soil quality. It measured the soil quality depending on the appropriate factors of weight. The comprehensive experimental analysis of this improved SVM model confirmed better processing time, accuracy and precision, compared to the compared deep learning models considered for investigation. .

A Support Vector Machine-based soil prediction scheme was proposed in [17] to facilitate the water level in the dry sub-humid tropics with maximized accuracy. This SVMSPS was proposed for improve the performance offered by the multiple regression method. It facilitated prediction using soil physicochemical properties with the dataset comprising of 296 samples. It was proposed with the capability to accurately estimate the moisture content with respect to the parameters of coefficients of determination and RMSE. It included the merits of linear kernel during the process of calibration to improve their capability towards generalization and predictive performance. A Long Short-Term Memory (LSTM)-based soil moisture prediction scheme was proposed to extrapolate the space and time dynamics of soil moisture [18] facilitated temporal dynamics in a predominant manner. It was proposed in such a way that it could be applied for applications that necessitates time-varying soil moisture with respect to anomaly detection and memory analyses. It was proposed with the capability of deriving temporal characteristics of features that helps in accurate estimation of water content in the soil over the period. The results of this LSTM model confirmed better exploration of features and minimized the processing time compared to the competitive approaches used for evaluation.

2.4 Extract of the Literature

The shortcomings of the existing state-of-the art soil moisture prediction schemes contributed to the literature over the recent decade is listed as follows.

i) Majority of the existing soil prediction approaches utilized only empirical methods which were not capable in handling the derivation of extracted automatic features from the dataset.
ii) Most of the existing deep learning approaches were not able to sustain generalization, prediction performance and accuracy as per the requirement, thereby possessing a room of improvement.
iii) The existing deep learning approaches used complete set of all parameters from which some of the parameters may be not potent in a specific context.
iv) The degree of mean RMSE and best RMSE attained by the existing deep learning-based soil moisture prediction approaches also have a room of improvement.

Based on the above-mentioned shortcomings, It is decided to formulate a Generative Adversarial Networks-based Deep Learning Model with Sailfish Optimization Algorithm (GAN-DLMSOA) to facilitate accurate estimation of soil moisture required for cultivating groundnut plants.

2.5 Contributions of the work

The major contributions of the proposed Adversarial Networks-based Deep Learning Model with Sailfish Optimization Algorithm (GAN-DLMSOA) is listed as follows:

i) It is proposed for predicting the degree of soil moisture which is required for better cultivation of groundnut plants in the district of Villupuram, Tamilnadu, India.
ii) It incorporated the significant merits of GAN to perform reliable and automatic extraction of features that aids in better prediction of soil moisture level.
iii) It is proposed for handling the limitations of prediction performance, multi-feature processing capability, generalization, and prediction accuracy that are inherent with the existing methods of prediction.
iv) It particularly utilizes Sailfish Optimization Algorithm (SOA) for optimizing the features and parameters that concentrates towards the attainment of accurate soil moisture prediction.
v) It further adopted SOA to train the GANs such that it can facilitate reliable and rapid convergence degree for accomplishing the task of prediction and facilitate precise results.
vi) Experiments of the proposed GAN-DLMSOA-based soil moisture prediction scheme is conducted in terms of performance metrics that include accuracy, average RMSE, best RMSE, MAE and MAD to evaluate its predominance over the benchmarked schemes used for analysis.
vii) The statistical and stability analysis using data perturbation is also conducted with different weather conditions to perceive the stability of the utilized deep learning model.

The remaining section of the paper is structured as follows. Section 2 presents the comprehensive review of the existing machine learning and deep learning-based soil prediction approaches contributed to the literature over the recent years. Section 3 depicts the detailed view of the proposed GAN-DLMSOA-based soil moisture prediction scheme with the role of GAN and feature optimization approach through sailfish optimization algorithm with suitable justifications. Section 4 demonstrates the experimental results and discussions of the proposed GAN-DLMSOA-based soil moisture prediction scheme with respect to the performance metrics of accuracy, average RMSE, best RMSE, MAE and MAD on par with the benchmarked schemes used for investigation.

3. Proposed Generative Adversarial Networks-based Deep Learning Model with Sailfish Optimization Algorithm (GAN-DLMSOA)-based soil moisture prediction

The proposed Generative Adversarial Networks-based Deep Learning Model with Sailfish Optimization Algorithm (GAN-DLMSOA) is proposed for predicting the available quantity of soil moisture which is essential for attaining sustainable agriculture. This GAN-DLMSOA adopted the significant parameters of relative humidity, air temperature and soil temperature into account during the process of precise soil prediction as depicted in Figure 1.

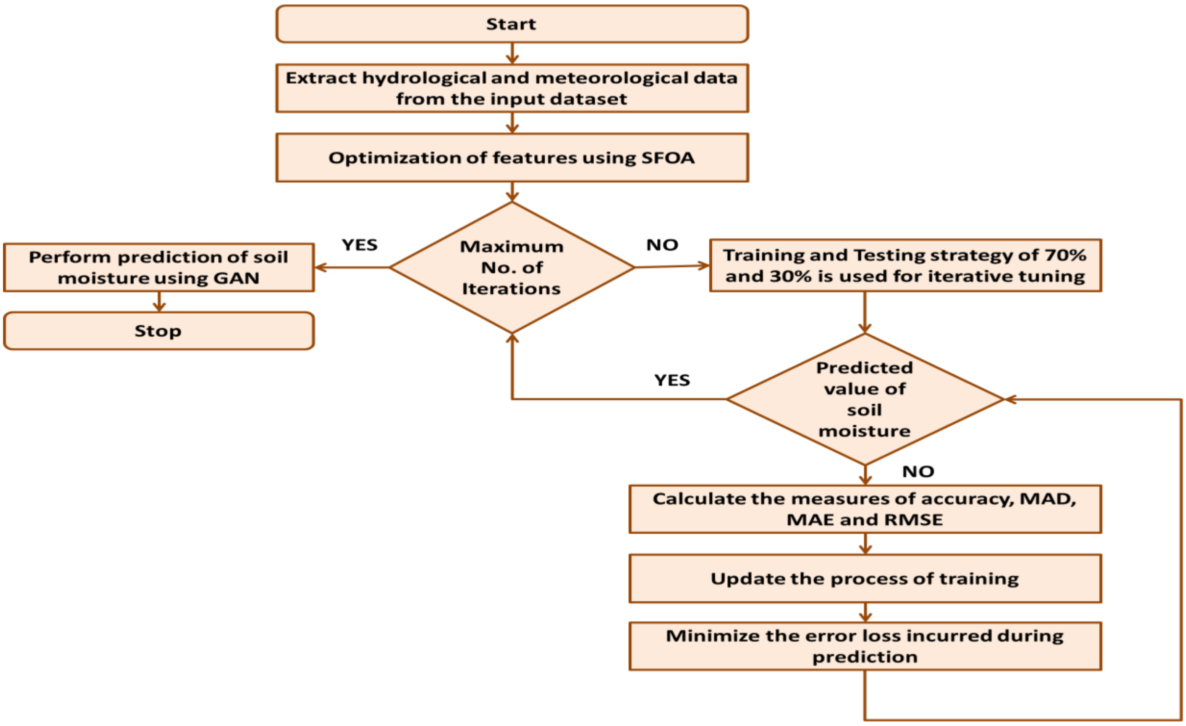

Figure 1: Flow diagram of the proposed GAN-DLMSOA-based soil prediction

It utilized the merits of GAN-based predictive models to realize the non-linear associated that exists between the soil moisture and the above-mentioned significant parameters. GAN is an approach for generating convolutional neural networks like generative models through the methods of deep learning. The generative modelling in machine learning refers to the task of unsupervised learning that aids in automatic learning and discovery of patterns or regularities form the input data. This generative model has the capability of generating or identifying new examples of output that could be plausibly derived from the original input dataset. GANs represents the intelligent strategy used for training the generative model through the process of formulating the problem as a supervised learning problem through the inclusion of two sub models, viz., i) generator model and ii) discriminator model. The generator model is responsible for generating new potential instances of examples, while the discriminator model plays an anchor role in the process of classifying examples as genuine or fake from the real or generated domain.

Deep GAN Network architecture used for soil moisture prediction

The deep GAN network architecture considered for soil prediction comprises of four main entities that includes encoder network (E_{Net}), decoder or generator network (G_{Net}), fusion of external factors (F_{ext}), and discriminator network (D_{Net}). The encoder network E_{Net} inherits a probabilistic encoder which can encode the data space (x_{Input}), into its latent code (x_{LC}). This inference network is responsible for deriving the output parameters to the distribution $D_p(x_{LC} | x_{Input})$. On the other hand, the decoder or generator network (G_{Net}) possesses a probabilistic decoder which is capable for learning the features that are possibly derived from the context of application to learn and reconstruct the input space x_{Input} from the available representation x_{LC}. The output parameters determined from the generative network is provided to the likelihood distribution $ML_{Fn}(x_{LC} | x_{Input})$.. Then, the Adversial process is employed for training the network through which the decoder network acquires knowledge to approximate the real data distribution in the problem space. The discriminative network D_{Net} helps in discriminating or differentiating between samples that are feasibly derived from the real samples and distributions that are generated through. It further possessed the capability of jointly discriminating the latent space to attain better stability in training and learning. This DGAN architecture during the training process uses the method of reconstruction loss that could be possibly estimated between $D_p(x_{LC} | x_{Input})$ and $ML_{Fn}(x_{LC} | x_{Input})$. Moreover, it is useful in computing the Adversial loss of G_{Net} and D_{Net} during the backpropagation process. It also possessed a facility of a generic fusion network that contextually integrated external parameters determined from the different fields with the data. This data fusion entity is given as input to E_{Net} and the generated latent code (x_{LC}) with respect to suitable external factors (F_{ext}) is integrated with the automatically extracted features during the prediction process.

In this DGAN predictive model, E_{Net}, G_{Net} and D_{Net} is developed based on the stacks of 3 dimensional ConvNet elements and Convolutional LSTM. The Convolutional LSTM neural network is potent in capturing the trends of spatio-temporal map sequences over a long term. It possesses the structures of convolution in the state-to-state and input-to-state transitions. On the other hand, 3D ConvNet aids in capturing the dependencies of local spatiality. It furthermore incorporates the capability of modelling the data correlations with the characteristics of cross temporality. It furthermore can better capture the data volume fluctuations and enhance the overall generalization potentiality of the model in the short term. This potential of 3D ConvNet is made possible only by establishing a sustained relationship between the neighbouring input data points that shares weights across different input and spatial temporal locality in the representations of features.

Role of encoder and generator network

The main role of encode and generator network is to concentrate on the process of processing the real time input data through the employment of multiple stacks of 3D ConvNet and ConvLSTM followed by Multi-Layer Perceptron (MLP) for the purpose of generating a condensed vector of features represented through Equation (1)

$$V_{F(x)} = MLP\left(Conv_{Net-3D}\left(\dots\dots, Conv_{LSTM}(x_{Input})\right)\right) \quad (1)$$

$$= MLP\left(Conv_{Net-3D}\left(Conv_{LSTM}^{D}(x_{Input})\right)\right)$$

Where, x_{Input} represents the real data input and as the extracted feature factor $V_{F(x)}$ related to the input x_{Input} with 'D' as the number of layers in the ConvLSTM. Then, a method of variational Bayesian mechanism is adopted with the assumption of multivariate Gaussian variable and the function of variational lower threshold. This estimation of thresholds aids in computing the mean and variance associated with the data input distribution in an explicit manner as presented in Equation (2)

$$V_{F(x)}(r) = \alpha + \beta \cdot \gamma, \text{ with } \gamma \sim N(0,1) \quad (2)$$

Where, γ is considered as the auxiliary random variable that range between 0 and 1 under the independence with $\cdot$ as the element-wise product operation. In particular, the method of Kullback-Leibler (KL) divergence method is utilized as the regularization term for preventing maximum degree of the possible deviation between $D_p(x_{LC}\,|x_{Input})$ and $ML_{Fn}(x_{LC}\,|x_{Input})$, respectively. The similar kind of process is introduced for the complete set of external factors (F_{ext}) to include their impact into the model of prediction. Further, the module of feature extraction is designed with the similar stack of 3D ConvNet and ConvLSTM followed by MLP. Then the auxiliary feature vector with the assumptions of Gaussian data distribution is used for learning as represented as $V_{Aux\,(x)}(r)$. Finally, the two-feature vector such as $V_{F(x)}(r)$and $V_{Aux\,(x)}(r)$ are concatenated and given as input to the decoder network. In other words, the reconstruction of spatio-temporal maps is attained through the combined merits of MLP, 3D ConvNet and ConvLSTM in its original size as presented in Equation (3).

$$V_{F(x)}(E_{Net}) = (\ 〚Conv〛\_(Net-3D)\ (\ 〚Conv〛\_LSTM\ (ML\left(PV_{Cat-F-Aux\,(x)}(r)\right))))$$
(3)

Where, $V_{F(x)}(E_{Net})$ represents the spatio-temporal maps after reconstruction process. In addition, a vector of noise N_{Vector} is forwarder to G_{Net} for the purpose of reconstructing spatio-temporal maps $F-N_{Vector}$

Role of Discriminator network

The primary role of discriminator network focusses on checking whether the generated spatio-temporal maps is determined from the ground truth or G_{Net}. In the D-GAN, the concatenation of the latent code with its generated spatio-temporal map is attained to combinedly learn the data space and latent code. This concatenation aids in achieving high training stability, better learning, and faster convergence. This strategy involved in attaining rapid convergence and better stability is represented in Equation (4)

$$W_{F(x)}(E_{Net}) = \left(V_{F(x)}(E_{Net}), ML\left(PV_{Cat-F-Aux\,(x)}(r)\right)\right) \quad (4)$$

Further, the generation of fake and noise feature vectors are generated based on and

Finally, the similar kind of stacked 3D ConvNet and ConvLSTM layers is used for the implementation of the discriminator network represented in Equation (5)

$$D_{Net\,(Output)} = \sigma\left(Conv_{Net-3D}\left(Conv_{LSTM}\left(W_{F(x)}(E_{Net})\right)\right)\right) \quad (5)$$

Need for feature optimization of GAN

The features considered for optimizing the performance of GAN is listed as follows.

i) The potential features that aid in improving the accuracy of GAN need to be identified with more efficiency.
ii) The features need to be detected with robustness and the false positive rate possible during prediction process need to be minimized.
iii) The problem of overfitting which is the major challenge in the implementation of the proposed deep learning model need to be handled.

Primitives of SailFish Optimization Algorithm (SFOA) and its role in feature optimization of GAN

The SailFish Optimization Algorithm (SFOA) was proposed based on the group hunting behavioral characteristics of sailfish. This SFOA algorithm comprises of two types of population such as sailfish and sardines' population. In specific, sailfish population is responsible for performing the process of intensification to search the neighbourhood region determined to be the best until the current iteration. On the other hand, sardines' population corresponds to the diversification process facilitated over the search space.

3.1 Process of Initialization

In the initialization process, the sailfish is analogical to the search agent that explores the candidate solutions with the variables of the problem representing the sailfish position in the space of search. In particular, the population is generated randomly defined over the solution space. The search agents (sailfish) possess the capability of exploring the single, double, triple, and multi-dimensional space with their associated variable position vectors. The i^{th} variable explored by the k^{th} searching agent in an $d-$dimensional search space represented through the current position $SF_{(i,d)}$ with $1 \leq i \leq n$ is defined based on Equation (7)

$$SF_{Pos(i,d)} = \begin{bmatrix} SF_{Pos(1,1)} & SF_{Pos(1,2)} & \dots\dots & SF_{Pos(1,d)} \\ SF_{Pos(2,1)} & SF_{Pos(2,2)} & \dots.. & SF_{Pos(2,d)} \\ \dots\dots & \dots\dots & \dots.. & \dots\dots. \\ SF_{Pos(n,1)} & SF_{Pos(n,2)} & \dots.. & SF_{Pos(n,d)} \end{bmatrix} \quad (7)$$

The matrix $SF_{(i,d)}$ representing the candidate solutions that are explored by search agents is identified and saved in the memory. This memorization of positions associated with the candidate solution represents the complete set of variables that represents possible dimensions of the complete set of solutions during the optimization process. The fitness associated with search agent (sailfish) is determined based on Equation (8)

$$Fit_{Fn}\left(SF_{(i)}\right) = Fit_{Fn}\left(SF_{(1)}, SF_{(2)}, \dots\dots, SF_{(n)}\ \right) \quad (8)$$

In this context, the matrix exhibiting the fitness value associated with the computations of fitness value for all the search agents$\left(Fit_{Fn}\left(SF_{Pos(i,d)}\right)\right)$ is determined based on Equation (9)

$$Fit_{Fn}\left(SF_{Pos(i,d)}\right) = \begin{bmatrix} Fit_{Fn}\left(SF_{(1,1)}, SF_{(1,2)}, \dots\dots, SF_{(1,d)}\ \right) \\ Fit_{Fn}\left(SF_{(2,1)}, SF_{(2,2)}, \dots\dots, SF_{(2,d)}\ \right) \\ \dots\dots\dots\dots\dots\dots.. \\ Fit_{Fn}\left(SF_{(n,1)}, SF_{(n,2)}, \dots\dots, SF_{(n,d)}\ \right) \end{bmatrix} \quad (9)$$

Where $SF_{Pos(i,d)}$ and ' n 'represents the value associated with the $d-$dimension and the number of search agents (sailfish).

In this SFOA algorithm, the population of another search agents (sardines) is introduced, and its count is equal to the number of sailfish search agents. The position of this search agents (sardines) with its related fitness value is determined based on Equation (10) and (11), respectively.

$$SD_{Pos(i,d)} = \begin{bmatrix} SD_{Pos(1,1)} & SD_{Pos(1,2)} & \dots\dots & SD_{Pos(1,d)} \\ SD_{Pos(2,1)} & SD_{Pos(2,2)} & \dots.. & SD_{Pos(2,d)} \\ \dots\dots & \dots\dots & \dots.. & \dots\dots. \\ SD_{Pos(n,1)} & SD_{Pos(n,2)} & \dots.. & SD_{Pos(n,d)} \end{bmatrix} \quad (10)$$

$$Fit_{Fn}(SD_{Pos(i,d)}) = \begin{bmatrix} Fit_{Fn}(SD_{(1,1)}, SD_{(1,2)}, \dots \dots, SD_{(1,d)}\ ) \\ Fit_{Fn}(SD_{(2,1)}, SD_{(2,2)}, \dots \dots, SD_{(2,d)}\ ) \\ \dots \dots \dots \dots \dots \dots .. \\ Fit_{Fn}(SD_{(n,1)}, SD_{(n,2)}, \dots \dots, SD_{(n,d)}\ ) \end{bmatrix} \quad (11)$$

Where, $Fit_{Fn}(SD_{Pos(i,d)})$ represents the objective function determined for identifying the fitness value based on the application of each search agent (sardines). Further,

Mechanism of Elitism

The mechanism of elitism is utilized in SFOA for preventing the superior solutions from being lost during the process of search agents' updating process. In specific, the position of solutions lost during each updating process may be weaker than the new solutions that are determined before the application of the elitism phenomenon. In specific, elitism plays an anchor role in copying or maintaining the fittest solution from one generation to the successive generation. In SFOA, the search agent (sailfish) determined so far as the best solution is termed as an elite solution. This elite solution is essential for accelerating (applying exploitation) over the population of search agents (sardines) to determine better local best solution.

Method of attack alternation

In this method of attack alternation, the search agents facilitate the phase of exploration and employs the principle of determining the promising solutions that needs significant refinement. This search agent possesses the capability of performing the complete process of exploration depending on the shrinking circle factor through the exhaustive number of parameters that could help in the process of global search. The updated search agent (sailfish) based on the method of attack alternation is determined in the i^{th} iteration based on Equation (12)

$SF\_((i)\ )\ (New) = SF_{(i)}(Elt) - \alpha\_i \times (\ 〚rnd〛\ \_((0,1))\ ) \times ((SF\_((i))\ (Elt) +\ 〚SD〛$
$\_((i))\ (Inj))/2) - SF\_((i))\ (Old)$ (12)

Where, $SF_{(i)}(Elt)$ and $SD_{(i)}(Inj)$ represents the elite sailfish and injured (exploited) sardine solution as determined until the current iteration. Further, α_i and $rnd_{(0,1)}$ depicts the influential coefficient and random number that ranges between 0 and 1. Furthermore, the influential factor α_i is determined based on Equation (13)

$$\alpha_i = (2 \times rnd_{(0,1)} \times Denst_P) - Denst_P \quad (13)$$

Where, $Denst_P$ represents the number of candidate solutions that are exploited during the search of optimal features in the space. This adaptive parameter '$Denst_P$' is determined based on Equation (14)

$$Denst_P = 1 - \frac{SF_{Cnt}}{SF_{Cnt} + SD_{Cnt}} \quad (14)$$

Where, SF_{Cnt} and SD_{Cnt} depicts the number of search agents (sailfish) and search agents (sardines) utilized for performing the process of exploration and exploitation. Moreover, the value of SD_{Cnt} need to be always greater than the value of SF_{Cnt} .

Prey hunting and catching phase

This phase of hunting and catching associated with SFOA is responsible for updating the position of the solution with respect to sardine (search agent) depending on the current best local solution and attack power identified in each iteration based on Equation (15)

$〚SD〛\_((i))\,(New) = (〚rnd〛\_((0,1))) \times (〚SF〛\_((i))\,(Elt) - 〚SD〛\_((i))\,(Old) + P\_attk)$ (15)

At this juncture, P_{attk} represents the power of attack (exploitation) introduced by search agents (sailfish) in each individual iteration as determined in Equation (16)

$$P_{attk} = P_{attk}(Coeff) \times \left(1 - (2 \times Iter_{Curr})\right) \times \beta \qquad (16)$$

Where, $P_{attk}(Coeff)$and β highlights the coefficients whose values decreases linearly from to 0 depending on the value of the power attack. This use of coefficients clearly portray that the attack power reduces depending on the time incurred for exploration and exploitation, which facilitates the phenomenal support towards the search agents' convergence. The search agent (sardines) depending on the parameter of $P_{attk}(Coeff)$ is responsible for updating the global and local best solution based on Equation (17) and (18)

$$\omega = SD_{Cnt} \times P_{attk}(Coeff) \qquad (17)$$

$$\delta = Var_{Cnt} \times P\_attk\ (Coeff) \qquad (18)$$

Where, SD_{Cnt} and Var_{Cnt}specifies the number of search agent (sardines) and number of variables considered in each iteration during the implementation of the utilized SFOA algorithm. When the value of P_{attk} is less than 0.5, the values of ω and δ associated with search agent (sardine) is updated. On the other hand, the solutions determined by the complete set of search agent (sardine) is updated when the value of P_{attk} is greater than and equal to 0.5. In particular, the parameters of and used for introducing random behavior into the process of feature optimization and prevent the issue of stagnation in determining local optima during the execution of all iterations.

As mentioned, the feature optimization process of GAN is achieved through the above-mentioned SFOA algorithm for improving the accuracy involved in soil moisture prediction. Once the feature optimization process is attained through SFOA, then adversarial losses are used for establishing the balance between G_{Net} and D_{Net} during the adversarial training process. In the context, the adversarial loss in the DGAN is determined based on Equation (19)

$D\_GAN\hat{}(A - Loss) = \|(D_{Real}(y) - 1)\|\hat{}2 + \|D\_Expt\ (y) - 1\|\hat{}2 + \|D\_Error\ (y) - 1\|\hat{}2$
(19)

In this case, the method of least function is employed rather than binary cross entropy method for improving the performance of the employed DGAN to determine the difference between the sample data points considered for prediction process. Finally, the complete objective of DGAN-based soil moisture prediction is determined based on Equation (20)

$$Total - D_{GAN}^{A-Loss} = D_{GAN}^{A-Loss}(E_{Net}) + D_{GAN}^{A-Loss}(D_{Net}) + D_{GAN}^{A-Loss}(F_{ext}) \qquad (20)$$

Where, $D_{GAN}^{A-Loss}(E_{Net})$, $D_{GAN}^{A-Loss}(D_{Net})$ and $D_{GAN}^{A-Loss}(F_{ext})$ represents the losses with respect to encoder network (E_{Net}), fusion of external factors (F_{ext}), and discriminator network (D_{Net}), respectively. This value of $Total - D_{GAN}^{A-Loss}$ is highly minimized during SFOA optimization for achieving predominant accuracy in soil moisture prediction process.

4. Simulation Results and Discussion

The simulation experiments of the proposed GAN-DLMSOA scheme and the benchmarked SVMSPSM, TGSMO, RNDLM and AQUACrop schemes are conducted with the Villupuram district soil moisture dataset. This dataset considered for the current investigation is publicly available. The datasets associated with the years of 2018 and 2019 is considered for the experimental investigation. These experiments are conducted for 10 iterations with respect to each of the algorithms with the validation of 10-fold cross utilized for each of the analysis. The performance measures such as precision, recall score, F1-Measure, ROC and Accuracy along with the error measure RMSE, MAD, MAE and R^2 are calculated and a comparison over the baseline models is derived for the proposed GAN-DLMSOA proved through the graphical comparisons. This experimental analysis is conducted using an 8GB Intel i3 processor machine.

The process of training involved in the proposed GAN-DLMSOA scheme is achieved based on the data (Villupuram district soil moisture dataset) determined from the web-enabled geographical information server. This GIS server refers to the India Water Resource Information System (India-WRIS) maintained by the ISRO (Indian Space Research Organization). This India-WRIS plays an anchor role as a single-window system for the objective of determining hydrological data such as soil temperature and soil moisture, and the meteorological data consisting of rainfall level, air humidity and air temperature. The metrological parameters determined for training purpose is determined from the Tamilnadu metrological department of Villupuram district over two years from 1st June 2018 to 15th June 2020. Thus, data pertaining to 730 days is captured for understanding the dynamics existing the full annual cycle. In specific, cosmic ray soil moisture sensors were deployed during the entire duration in the regions of analysis for estimating the soil temperature and volumetric soil moisture. This cosmic ray soil moisture sensors named Model CRS-1000/B possesses the capability of measuring a phenomenal depth of 20 meters and maximized horizontal range of 200 m. Further, resampling of the extracted data is attained in a daily basis by feeding the individual data into GAN-DLMSOA which aids in predicting the volumetric soil content that could exist in the forthcoming days on the sites of investigation. The dataset is divided based on the ratio of 70:30 for preventing the problem of network overfitting. Thus, the data of 511 days is used for training and the remaining 219 days of data is used for testing. This partition of dataset is mainly for validating the potential of the proposed GAN-DLMSOA scheme determined based on the posterior temporal dataset estimated before the period of training.

On the other hand, the model is tested using the dataset of 219 days kept separated from each of the sets of training. In particular, the proposed GAN-DLMSOA scheme takes the mean value of the preceding climatic variables associated with the preceding three days for the objective of assessing the mean volumetric soil moisture for the successive day of investigation.

Consideration of Performance and Error Measures

In this section, the performance and error measures considered for evaluating the potential of the proposed GAN-DLMSOA scheme is defined as follows.

a. Accuracy: It is computed based on the ratio of the complete number of all correct predictions to the total number of a sample dataset as represented below.

$$ACCURACY = \frac{T_P+T_N}{T_P+T_N+F_P+F_N} \quad (21)$$

The best value of accuracy is 1.0 with the worst being 0.0.

The compared accuracy of the proposed GAN-DLMSOA Model with the other baseline models is shown in the figure 2

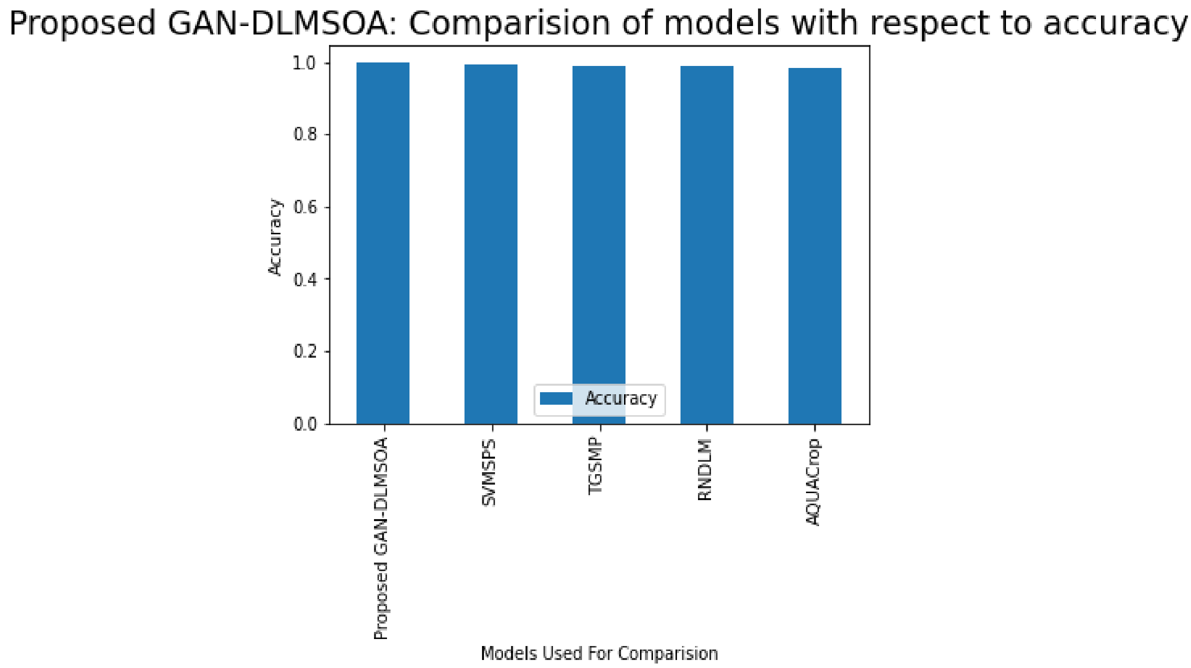

Figure: 2

b. Precision: It is computed as the ratio between the number of positive predictions correctly identified to the cumulative number of positive predictions as represented below.

$$PRECISION = \frac{T_P}{T_P+F_p} \quad (22)$$

The best value of precision is 1.0 with the worst being 0.0.

The compared precision of the proposed GAN-DLMSOA Model with the other baseline models is shown in figure 3

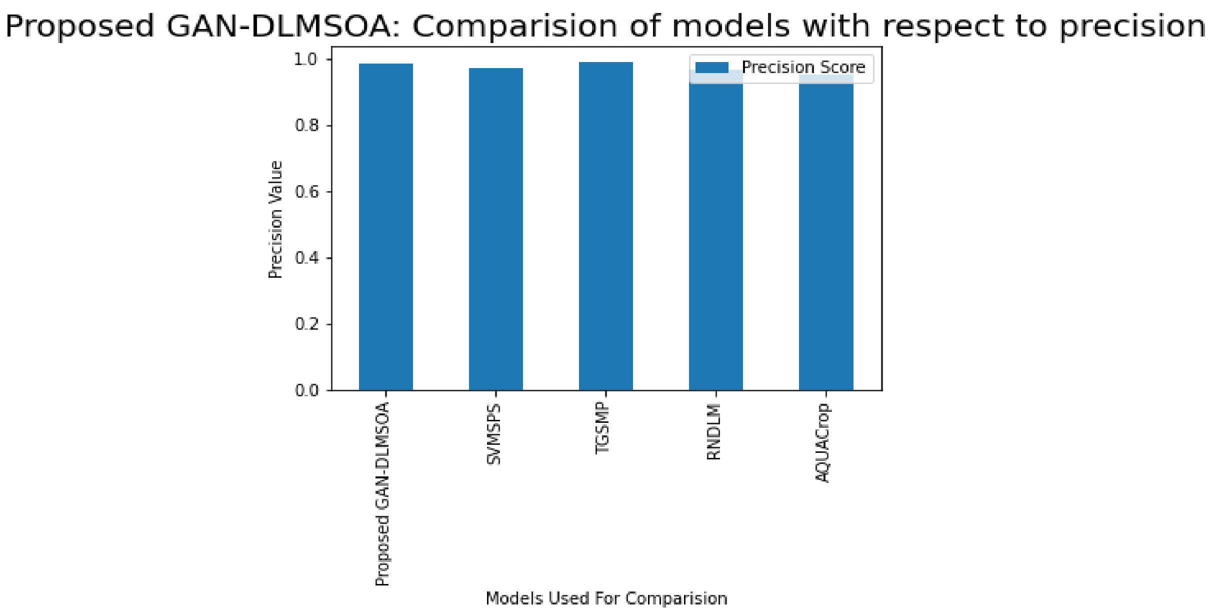

Figure:3

c. Recall:

The recall is defined as the actually predicted true positive values to the overall sum for true predicted positive values and false negative values.

$$\text{RECALL} = \frac{\text{TP}}{TP+FN} \tag{23}$$

Here, TP-True Positive ; FN- False Negative.

d.F1-Measure:

The F1 measure is defined as the harmonic mean of precision and recall

$$F1 = 2 * \frac{PRECISION*RECALL}{PRECISION+RECALL} \tag{24}$$

e. Receiver Operating Curve (ROC)

ROC Curves are used for the process of evaluating the performance of the various baseline models with the proposed model, it has a false positive rate on X-axis and True positive rate on Y-axis.

The table 1 has been derived with the calculated performance measures of all the baseline models in comparison with the proposed GAN-DLMSOA and Figure:4 depicts its corresponding graphical representation of the performance metrics, The comparison clearly concludes that the proposed GAN-DLMSOA has shown a considerable improvement in all its performance measures.

	Proposed GAN-DLMSOA	SVMSPS	TGSMP	RNDLM	AQUACrop
Precision	0.980650	0.972948	0.987545	0.963948	0.953545
Recall Score	0.851250	0.771800	0.722130	0.591180	0.489130
F1-Measure	0.860236	0.828705	0.798412	0.458705	0.558412
ROC	0.892360	0.842360	0.812360	0.542360	0.712360
Accuracy	0.996650	0.995173	0.990019	0.987173	0.985019

Table 1: Comparison on Performance measures of various baseline models with the proposed GAN-DLMSOA Model

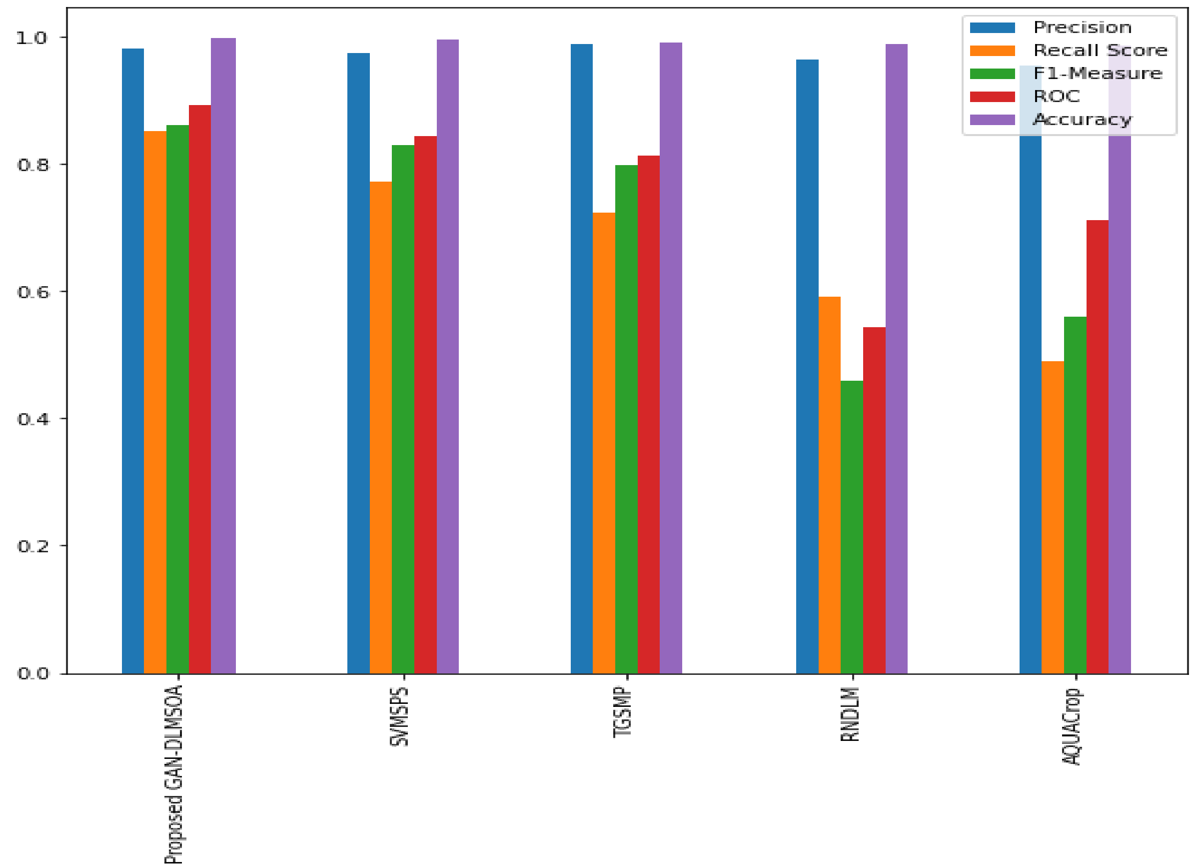

Figure:4

Comparison of Error Measures

RMSE: It is Root of the Mean of the Square of Errors which depicts the square root of the difference between the predicted values of the utilized model and the actual values associated with the variable of study (soil moisture) determined over the total number of observations.

$$RMSE = \sqrt{\frac{ACT_{Val} - PR_{Val}}{Obs_{No}}} \qquad (25)$$

MAE: It is mean of absolute values of the which depicts the difference between the predicted values of the utilized model and the actual values associated with the variable of study (soil moisture) determined over the total number of observations.

$$MAE = \frac{ACT_{Val} - PR_{Val}}{Obs_{No}} \qquad (26)$$

Where, ACT_{Val} and PR_{Val}represent the actual value and predicted value as determined by the proposed model with Obs_{No} as the total number of observations

MAD: The mean absolute deviation of a dataset is the average distance between each data point and the mean. It gives us an idea about the variability in a dataset.

$$\text{MAD} = \frac{\sum Absolute\ values\ of\ deviation\ from\ central\ measures}{Total\ .no.of.Observations} \qquad (27)$$

The table 2 has been derived with the calculated error measures comparing with the all the baseline models with the proposed GAN-DLMSOA and Figure:5 depicts its corresponding graphical representation, The comparison clearly concludes that the proposed GAN-DLMSOA has shown a considerable defending results in all its error measures too. The calculated Coefficient of Determination (R^2) Value of the GAN-DLMSOA model ranges to be greater in comparison with all the baseline models, which signifies how good the coefficient is actually fit with the training dataset values

	Mean RMSE Score	Mean Absolute Deviation	Mean Absolute Error	Coefficient of Determination
Proposed GAN-DLMSOA	0.04065	0.237236	0.010236	0.120254
SVMSPS	0.07118	0.328705	0.021236	0.081548
TGSMP	0.13913	0.358412	0.035236	0.054120
RNDLM	0.14118	0.408705	0.045236	0.075412
AQUACrop	0.23913	0.428412	0.054236	0.010880

Table 2: Comparison of Error measures over various baseline models with the proposed GAN-DLMSOA Model

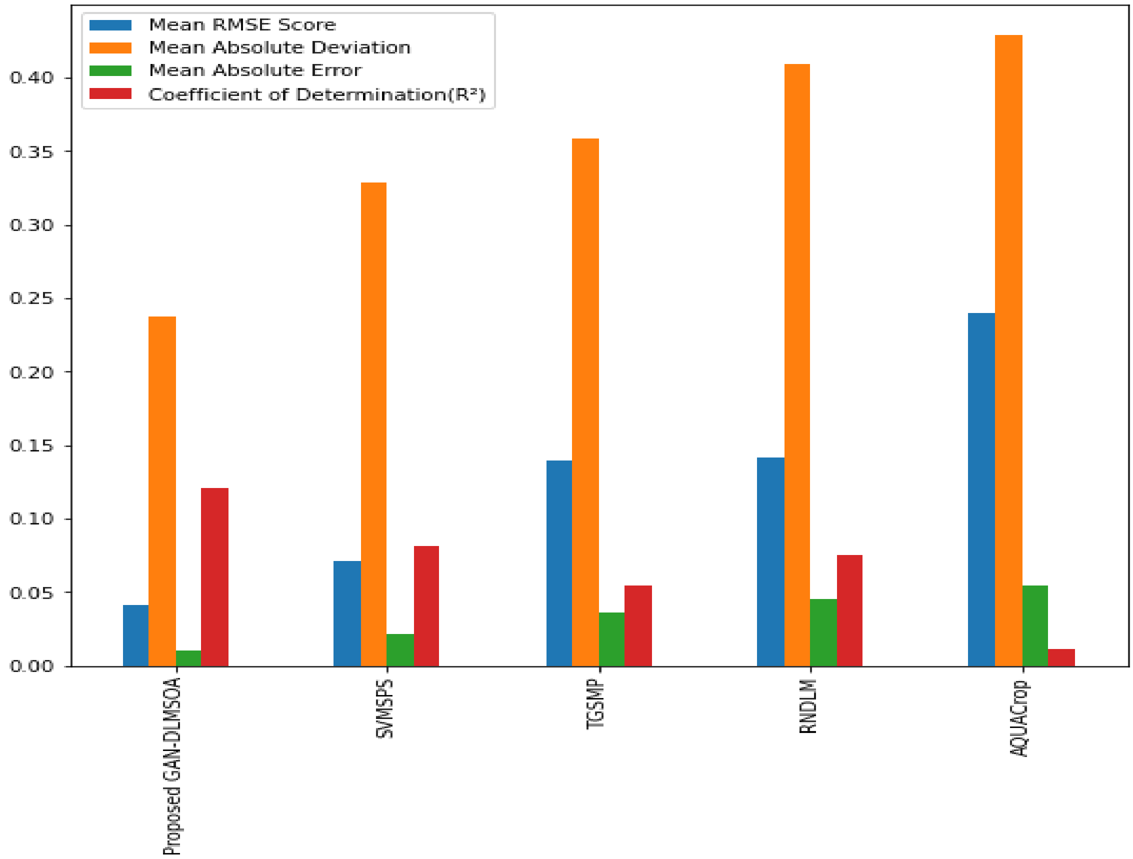

Figure:5

Initially, the prediction efficiency of the proposed GAN-DLMSOA scheme and the benchmarked SVMSPSM, TGSMO, RNDLM and AQUACrop schemes are conducted with the Villupuram district soil moisture dataset using confusion matrix represented in Table 3, 4, 5, 6 and 7. If the model is potent in predicting the moisture with 100% accuracy, then the value of the diagonal elements in the confusion matrix should be 3000. Except the confusion matrix presented in Table 1, the values of the diagonal elements are either higher or lower than 3000 due to the issue of misclassification. However, the value of difference between the predicted and actual value of prediction is comparatively realized to be lower with the proposed GAN-DLMSOA scheme, which is significantly an improved performance over the compared SVMSPSM, TGSMO, RNDLM and AQUACrop schemes.

Table 3: Confusion Matrix for depicting the efficiency of the proposed GAN-DLMSOA

		Predicted Moisture Level					
Actual Moisture Level		Sample A	Sample B	Sample C	Sample D	Sample E	
	Sample A	2995	1	2	0	0	2998
	Sample B	1	2996	2	0	0	2999
	Sample C	0	0	2917	23	1	2941
	Sample D	0	0	56	2904	4	2964
	Sample E	0	0	0	7	3091	3098
		2996	2997	2977	2934	3096	15000

Table 4: Confusion Matrix for depicting the efficiency of the proposed SVMSPSM scheme

Actual Moisture Level	Predicted Moisture Level						
		Sample A	Sample B	Sample C	Sample D	Sample E	
	Sample A	2921	3	5	0	0	2929
	Sample B	3	2936	9	0	0	2948
	Sample C	0	0	2943	29	6	2978
	Sample D	0	0	44	2948	7	2999
	Sample E	0	0	0	158	2988	3146
		2924	2939	3001	3135	3001	15000

Table 5: Confusion Matrix for depicting the efficiency of the proposed TGSMO

Actual Moisture Level	Predicted Moisture Level						
		Sample A	Sample B	Sample C	Sample D	Sample E	
	Sample A	2988	3	6	0	0	2997
	Sample B	2	2965	6	0	0	2973
	Sample C	0	0	2939	31	6	2976
	Sample D	0	0	44	2917	9	2970
	Sample E	0	0	0	7	3077	3084
		2990	2968	2995	2955	3092	15000

Table 6: Confusion Matrix for depicting the efficiency of the proposed RNDLM

Actual Moisture Level	Predicted Moisture Level						
		Sample A	Sample B	Sample C	Sample D	Sample E	
	Sample A	2922	6	8	0	0	2936
	Sample B	4	2931	6	0	0	2934
	Sample C	0	0	2941	12	5	2958
	Sample D	0	0	31	2896	8	2935
	Sample E	0	0	0	9	3226	3235
		2926	2937	2987	2917	3239	15000

Table 7: Confusion Matrix for depicting the efficiency of the proposed AQUACrop Scheme

Actual Moistu	Predicted Moisture Level						
		Sample A	Sample B	Sample C	Sample D	Sample E	
	Sample A	2995	1	6	0	0	3002

	Sample B	1	2996	4	0	0	3001
	Sample C	0	0	2917	23	1	3017
	Sample D	0	0	56	2904	4	2964
	Sample E	0	0	0	7	2995	3202
		2996	2997	2977	2934	3000	15000

The complete experimental process of the proposed GAN-DLMSOA scheme and the benchmarked schemes is achieved in three folds. Initially, the performance of the proposed GAN-DLMSOA scheme and its benchmarked schemes are compared based on the Performance measures which includes (Accuracy, Precision, Recall score,F1-measure and ROC) along with error measures (RMSE, MAD and MAE) with respect to Villupuram district soil moisture dataset. Figure 2 and 3 demonstrates the performance of the proposed GAN-DLMSOA scheme and the benchmarked SVMSPSM TGSMO, RNDLM and AQUACrop schemes based on mean accuracy and precision with Villupuram district soil moisture dataset. The accuracy and precision guaranteed by the implemented proposed GAN-DLMSOA-deep learning architecture was maximal, since the rapidness in deriving robustness features that could possibly influence the prediction of soil moisture was potentially excellent on par with other deep learning models. Thus, accuracy achieved by the proposed GAN-DLMSOA scheme is identified to be potentially enhanced by 10.94%, 11.65%, 13.91% and 15.86%, better than the baseline SVMSPSM TGSMO, RNDLM and AQUACrop schemes. Moreover, the proposed GAN-DLMSOA scheme is also confirmed to improve precision by 11.52%, 13.81%, 15.65% and 16.88%, excellent to the baseline schemes used for investigation.

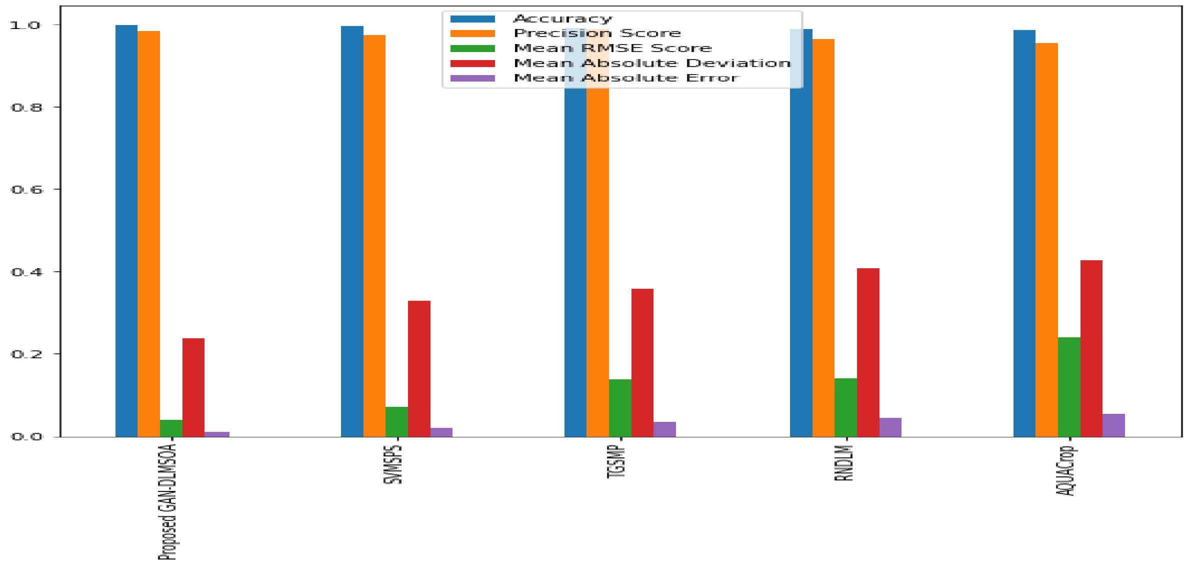

Figure 6: Overall performance comparison of the proposed GAN-DLMSOA

Figure: 4 presents the performance of the proposed GAN-DLMSOA scheme and the benchmarked SVMSPSM TGSMO, RNDLM and AQUACrop schemes based on mean RMSE, MAD , MAE and Coefficient of Determination(R^2) with Villupuram district soil moisture dataset. The RMSE, MAE and MAD value of the proposed GAN-DLMSOA scheme is identified to be significantly minimized and

R^2 value turns to greater on par with the other baseline schemes, since it derives the benefits of SOA algorithm in optimizing the features considered for improving the performance of the utilized GAN. This is mainly due to the process included in the automatic extraction of features that determines the prediction of soil moisture with utmost accuracy and minimal error scores. The figure 6, clearly depicts the Overall efficacy of the proposed GAN-DLMSOA scheme over other compared models.

4.2 Statistical analysis

The statistical investigation of the proposed GAN-DLMSOA scheme and the baseline SVMSPSM TGSMO, RNDLM and AQUACrop schemes is conducted using Wilcoxon rank test with 5% level of significance. In this statistical analysis, the null hypothesis states that “there is no significant difference between the mean of different groups”. At the same time, alternative hypothesis states that “there is significant difference between the mean of different groups”. At this juncture, the threshold considered for the test is determined to be 0.028 depending in the inequality of Bonferroni. Table 1 presents the calculated p-values for GAN-DLMSOA Vs SVMSPSM, GAN-DLMSOA Vs TGSMP, GAN-DLMSOA Vs RNDLM and GAN-DLMSOA Vs AQUACrop, respectively.

Table 6: Proposed GAN-DLMSOA-Wilcoxin rank test-p values

	Algorithms used for Comparison with p values			
	SVMSPSM	TGSMP	RNDLM	AQUACrop
Proposed GAN-DLMSOA	**3.42e-09**	**2.93e-07**	**2.46e-08**	**2.51e-10**

The p-values are determined to be potentially smaller than the value of the threshold as mentioned earlier. This clearly proved that in all the cases, the null hypothesis is identified to be rejected. Moreover, the results achieved through the application of proposed GAN-DLMSOA scheme is not random, thereby identified to be statistically significant.

5. Conclusion

In this paper, GAN-DLMSOA was contributed with the merits of D-GAN to facilitate automatic extraction of features and parameters with optimized capability of SOA to predict the soil moisture level convenient for the cultivation of groundnut plants in the district of Villupuram, Tamilnadu, India. It was proposed with the capability of better prediction, multi-feature processing potential, generalization, and prediction accuracy, which is a significant improvement over the existing soil moisture prediction strategies. It adopted SOA and facilitated features optimization contextually that helped in achieving accurate soil moisture prediction. The experimental results of the proposed GAN-DLMSOA-based soil moisture prediction scheme confirmed better accuracy and precision, on an average by 14.82% and 17.63%, compared to the baseline approaches used for comparison. The results proved that the average RMSE and best RMSE attained by the proposed GAN-DLMSOA-based soil moisture prediction scheme is minimized by 15.12% and 18.65%, superior to the benchmarked schemes used for analysis. Moreover, the proposed GAN-DLMSOA-based soil moisture prediction scheme minimized MAE and MAD to the maximized level of 23.18% and 25.48%, excellent over the benchmarked schemes. In addition, the statistical and stability analysis of the proposed GAN-DLMSOA-based soil moisture prediction scheme also confirmed its predominance with respect to the evaluation done with p-test. The results in confusion matrix confirmed that the value of difference between the predicted and actual value of prediction is comparatively realized to be lower with the proposed GAN-DLMSOA scheme, which is definitely an predominant performance over the compared SVMSPSM, TGSMO, RNDLM and AQUACrop schemes. As the part of future research, it is also planned to formulate a CNN and GRU-based soil moisture prediction approach and compare it with the proposed GAN-DLMSOA scheme to determine the best among the two in terms of generalization and prediction accuracy.

References

1) Chauhan, Y. S., Ryan, M., Chandra, S., & Sadras, V. O. (2019). Accounting for soil moisture improves prediction of flowering time in chickpea and wheat. Scientific Reports, 9(1), 67-79.

2) Chatterjee, S., Kumar, S., Saha, J., & Sen, S. (2019). Hybrid regression model for soil moisture quantity prediction. 2019 International Conference on Opto-Electronics and Applied Optics (Optronix), 2(1), 45-56.

3) Adab, H., Morbidelli, R., Saltalippi, C., Moradian, M., & Ghalhari, G. A. (2020). Machine learning to estimate surface soil moisture from remote sensing data. *Water*, *12*(11), 3223.

4) Mohamed, E., Ali, A., El-Shirbeny, M., Abutaleb, K., & Shaddad, S. M. (2020). Mapping soil moisture and their correlation with crop pattern using remotely sensed data in arid region. *The Egyptian Journal of Remote Sensing and Space Science*, *23*(3), 347-353.

5) Liu, H., Yang, Y., Wan, X., Cui, J., Zhang, F., & Cai, T. (2021). Prediction of soil moisture and temperature based on deep learning. 2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), 2(1), 67-79.

6) Veres, M., Lacey, G., & Taylor, G. W. (2015). Deep learning architectures for soil property prediction. 2015 12th Conference on Computer and Robot Vision, 2(2), 45-56.

7) Chen SY, Zhang XY, Shao LW, Sun HY, Niu JF. A comparative study of yield, cost-benefit and water use efficiency between monoculture of spring maize and double crops of wheat-maize under rain-fed condition in the North China Plain. Chinese Journal of Eco-Agriculture,2015,23(05):535–543.

8) Martínez FJ, González ZA, Sánchez N, Gumuzzio A, Herrero JCM. Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index. Remote Sensing of Environment, 2016, 177: 277–286.

9) Shu SF, Qian HF, Qiu XW. Soil moisture forecast model based on meteorological factors in Jinhua City. Chinese Journal of Agrometeorology, 2009,30(02):180–184.

10) Ronghua, J., Shulei, Z., Lihua, Z., Qiuxia, L., & Saeed, I. A. (2017). Prediction of soil moisture with complex-valued neural network. 2017 29th Chinese Control And Decision Conference (CCDC), 2(2), 360-366.

11) 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), e0214508.

12) Adeyemi, O., Grove, I., Peets, S., Domun, Y., & Norton, T. (2018). Dynamic neural network modelling of soil moisture content for predictive irrigation scheduling. *Sensors*, *18*(10), 3408.

13) Bandyopadhyay, S., Aggarwal, M., & Murty, M. N. (2020). Self-supervised hierarchical graph neural network for graph representation. 2020 IEEE International Conference on Big Data (Big Data), 2(2), 45-56.

14) Prakash, Shikha et al. "Soil Moisture Prediction Using Machine Learning." 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT) (2018): 1-6.
15) Cussed, V. C. (2021). Estimation of soil moisture and earth resistivity using Wenner's method and machine learning. *ECS Meeting Abstracts*, *MA2021-01*(54), 1331-1331.

16) Sumathi, P., Subramanian, R., Karthikeyan, V., & Karthik, S. (2021). Soil monitoring and evaluation system using EDL-ASQE: Enhanced deep learning model for IoI smart agriculture network. International Journal of Communication Systems, 2(2), 67-79.

17) Kaingo, J., Tumbo, S. D., Kihupi, N. I., & Mbilinyi, B. P. (2018). Prediction of soil moisture-holding capacity with support vector machines in dry subhumid tropics. *Applied and Environmental Soil Science*, *2018*(2), 1-10.

18) O., S., & Orth, R. (2021). Global soil moisture data derived through machine learning trained with in-situ measurements. Scientific Data, 8(1). https://doi.org/10.1038/s41597-021-00964-1.

19) Aggarwal, A., Mittal, M., & Battineni, G. (2021). Generative adversarial network: An overview of theory and applications. International Journal of Information Management Data Insights, 1(1), 100004.

20) Shadravan, S., Naji, H., & Bardsiri, V. (2019). The sailfish optimizer: A novel nature-inspired metaheuristic algorithm for solving constrained engineering optimization problems. *Engineering Applications of Artificial Intelligence*, *80*(2), 20-34.

21) https://indiawris.gov.in/wris/#/soilMoisture