International Journal of Data Science and Analytics
https://doi.org/10.1007/s41060-024-00640-8

REGULAR PAPER

Predicting normalized difference vegetation index using a deep attention network with bidirectional GRU: a hybrid parametric optimization approach

Nima Khodadadi[1] · S. K. Towfek[2,6,7,10,14] · Ahmed Mohamed Zaki[2] · Amal H. Alharbi[3] · Ehsan Khodadadi[4] · Doaa Sami Khafaga[2] · Laith Abualigah[5,6,12,13] · Abdelhameed Ibrahim[7] · Abdelaziz A. Abdelhamid[8,11] · Marwa M. Eid[9]

Received: 27 June 2024 / Accepted: 5 September 2024

Abstract

Scalable and accurate normalized difference vegetation index (NDVI) prediction is necessary to track the status of vegetation and the environment and to support proper ecological management. Herein, we present an innovative deep-learning approach to improve NDVI prediction performances by considering enhanced temporal modeling and hybrid optimization processes. The analysis is based on a core model that integrates a Bidirectional Gated Recurrent Unit (BiGRU) with the profound attention feature since the primary research incorporates the capability of complex temporal in addition to NDVI-time series value. The model performs better through a dual algorithm combining the waterwheel plant algorithm (WWPA) and statistical fractal search (SFS) named WWPASFS-BiGRU. The proposed approach is evaluated using real-world NDVI datasets, demonstrating its capability to outperform traditional models and state-of-the-art deep learning methods. Key performance metrics highlight the model's accuracy, with a root mean square error (RMSE) as low as 0.00011, reflecting its superior predictive ability. Comparative experiments showcase the robustness of our model across different environmental conditions and geographical settings, affirming its applicability in diverse ecological forecasting scenarios. Additionally, extensive statistical validation, including ANOVA and Wilcoxon tests, confirms the model's consistency and reliability. The effectiveness of the WWPASFS-BiGRU model is illustrated through applications in predicting NDVI trends across regions in Saudi Arabia, providing critical insights for ecosystem management and sustainable development planning.

Keywords Hybrid optimization · Waterwheel plant algorithm · Statistical fractal search · Attention model · Bidirectional gated recurrent unit (BiGRU) · Forecasting model

1 Introduction

The environmental issues cover a large area, including agricultural, water, or landscape threats. These threats can significantly disturb the natural order of things on our planet. Many of these discourses are caused by human actions, such as industrialization, city boom, and extraction. The loss of forests, representing only a tiny percentage of habitat type in total, can therefore be considered a vital extinction agent since it also acts as the primary mediator of climate change by reducing trees' capability to absorb more CO_2 in the atmosphere. Such pollution sources as industry waste, agriculture runoff, and motor vehicle emissions contribute to air, water, and soil pollution, consequently negating the impact on human health and the natural environment [61–63].

Currently, the planet is going through climatic change mainly due to the fumigation of fossil fuels and the unleashing of forests, leading to changes in weather patterns, sea levels, and the worsening of extreme weather occurrences. The resolution of environmental issues should be a comprehensive approach that consists of some essential components, such as individuals, communities, state agencies, and businesses. The enactment of sustainable methods, including renewable energy preferences, waste elimination, and sustainable agriculture, involved environmental depreciation and mitigation. The conservation of habitats, in the sense of preservation of endangered species and restoration cycles of degraded ecosystems, is a crucial issue of biodiversity and ecological resilience. Besides, policy, law, and regulation

Extended author information available on the last page of the article are essential to motivating substitutable practices, stabilizing pollution, and reducing greenhouse gas emissions [46, 81].

Satellite imagery can be utilized to assess the distribution and productivity of vegetation over extensive geographical regions. Normalized difference vegetation index (NDVI) is a widely utilized multi-spectral vegetation index. A strong correlation exists between leaf area and its utility as a tool for estimating aboveground biomass [40, 48]. Local, national, and international climate and environmental changes have been evaluated using satellite-collected NDVI data. The NDVI has been utilized in various agricultural applications, such as identifying alterations in land cover [31, 54], estimating aboveground biomass [29, 52], calculating crop phenological data, and predicting crop yield [39, 41]. Additionally, it was employed to monitor and analyze droughts [38, 66] and uncover the consequences of global warming [36, 53], soil erosion, and desertification [49, 74].

The relationship between vegetation alterations and climate change has been thoroughly examined by analyzing NDVI time series data over the last 3 decades [71, 75]. Further investigation into the quantitative relationship between climate and NDVI would substantially enhance the utility of NDVI in these domains. Remote sensing, namely satellite-based rainfall observations, can provide sufficient data with high spatiotemporal resolution across large areas, particularly in locations where traditional rain gauge data is limited [9, 24]. Satellite sensors cannot observe or detect rainfall independently, and the connection between observations relies on one or more indirect factors [68, 70], resulting in several limitations. One of the most notable drawbacks is the presence of substantial uncertainty. One instance of a current satellite observation data package connected to rainfall is the Climate Hazards Group InfraRed Rainfall with Stations (CHIRPS).

The CRIPS product complements daily precipitation data for 30 years period from 1981 to today with a spatial resolution of 0.05 covering the quasi-global area. “Santa Barbara Climate Hazards” was custom-made in collaboration with the “USGS-Earth Resources Observation and Science (EROS)” unit and the “University of California,” located in Santa Barbara. Patel et al. [51] explained in their article a detailed discussion of the manufacture of CHIRPS. One of the advantages is performing remote sensing techniques on satellites to calibrate rainfall observations; this provides an alternative to using rain gauge data. Therefore, the accuracy level of forecasting rainfall data will be enhanced [17]. More than one study was undertaken in different locations, e.g., Huangpu District in Mainland China [11], Asir Province in Saudi Arabia [42], southwestern North America, southeastern Africa, Brazil [21], and Colombia [16]. Using CHIRPS data always involves specific limitations, as the CHIRPS rainfall model uses infrared radiation statistics to construct the indicators. The Chippewa species finds favor balanced in tropical regions, however, they encounter troubles in temperate zones. CHAPS enjoys more tremendous success in humid environments than in areas with drier climates, such as deserts and semi-arid regions. This corresponding factor between monsoon entry and the best in optimal performance zones is evident. Chirps’ outcomes are most likely to be friendly from May to September rather than February to August. The range of the rain detection system is limited. The risk of a CHIRS event due to typhoon rains is relatively low compared to other forms of precipitation.

Prior studies have examined the relationship between changes in NDVI and meteorological conditions regarding geographical or temporal relationships. Statistical regression and correlation were commonly employed as the principal empirical measures in most investigations to ascertain the associations. Strong correlations exist between the (NDVI) and precipitation in arid places [56, 72]. The relationship between NDVI and eco-climatic predictions exhibits significant variability contingent upon geographical location and topographical characteristics. Variability in correlations between NDVI and its independent variable is generated by spatial changes in soil conditions or plant type [76]. The relationship between NDVI and local climatic variables, such as temperature and rainfall, has been extensively studied and confirmed across several spatiotemporal scales [23, 27]. Precipitation is a significant indicator of vegetation distribution in transitioning from humid to arid or semi-arid environments [33, 44]. The majority of studies investigating the relationship between vegetation and rainfall employ globally calibrated linear models, such as Ordinary Least Squares" (OLS) regression techniques [34]. In these models, the (NDVI) is treated as a variable representing rainfall.

It includes location-based parameters engagement through which the model’s parameters are aligned to the change in the location of the affected area. These models have more accuracy toward the specific geographical area because of their ability for the model to be locally adaptive. While autocorrelation and variance heterogeneities remain major stumbling blocks, local estimation models such as localized OLS and Geographically Weighted Regression (GWR) can be deployed to address non-stationarity and parameter quantification flexibility for time-varying fluctuating regression model parameters. The GWR [12, 45], which is the fold-back geography, has evolved and grown highly recognized among human geographers [13, 30] and ecologists [26, 80] GWR addressing both residual correlation and non-stationarity is thus qualified to adapt the exogenous variables and the explanatory variables whereas capturing the interactive relationship between the variables. This particular approach to regionalization can produce several significant findings, including where there is a deviation from or a significant deviation in the railway sector that otherwise may be misrepresented in a pan-national model. The GWR model has proven superior to ordinary linear regression (OLS) in landscape etiology. The AIC or the cross-validation scores (CV), as well as other generally used error measures, give proof for that conclusion. Regarding the application of remote sensing [22, 73] within the local regression approaches, such an exploration concerns the geographical correlations between the satellite images and the climate indicators.

The study is carried out by analyzing geographical associations between NDVI and rainfall parameters obtained from the watershed catchment region in Saudi Arabia. The locality is in the climatological transition zone, hence experiencing severe disruptions. The belief surrounding the wide range of the Gulf climate transition area, which runs from cold semi-arid to hot arid climates, is supported by a research article [3]. The transition zones, with their semi-arid nature according to the Koppen classification system, are marked by unique living organisms and indicate dehydration of biological systems, changing from a sub-humid to a desert climate. Although researchers have been involved in studies elaborating on the interaction between NDVI and climate at times of transition in regions like North China [28, 78], research has yet to be done comprehensively on the intricate relationship between temperature, precipitation, humidity, and their scale dependency and spatial stationarity regarding plant life in the Mak The study how vegetation under the climatic influence in transition zones and other regions are critical as those regions commonly experience substantial climatic variations across geographic locations. Working thus, let us strive for a crucial understanding of the intricacy of the relationships between NDVI and precipitation in this zone.

Without a doubt, the main topics of this section are considered doubtless problems and are broadly discussed by geostatistics experts who have long been engaged in this process. Although the employment of local regression models may still be a relatively new trend in ecological and environmental scientific writing, the potential of this tool can already be seen in many studies. Therefore, the value of the studies in research by this group of authors differed from the essence of discovering new approaches to analyzing geographical data. The research is valid as long as it is considered from a wide angle of ecology and environmental science, where holistic approaches to analyzing geographical data are explored. This study intends to theorize about spatial fluctuations of NDVI and rainfall in the Makkah watershed from various seasons between 2000 and 2016. We will apply these non-stationary modeling ways to reach the goal. 2000 precipitation was almost nothing, and 2016 was even wet, culminating in differences. To achieve this goal, the data gathered from multi-harvest cycles provides the spatial fluctuation analysis across the time. An attempt was made to categorize very progressive sites with changes in spatial rainfall by considering local regression in the neighboring area. With the primary goal of providing a scientific base for the problem areas that change in space–time, this study looks at problem areas that are affected by these dynamics. These outcomes give ecologists in the overview of climate change and plant distribution in various regions a chance to revise their specialized theories. The specialization theory applied previously may need a repeal for the newer version. The paper presents noteworthy findings that can be briefly described and then elaborated upon as follows:

- It is a new attention-based Bidirectional Gated Recurrent Unit (BGRU) model employed in the NDVI component forecasting.
- This study presents a unique hybrid optimization method that is ideal for optimizing the parameters of a forecasting model. The strategy uses a hybrid method to achieve this goal, including the water wheel plant and the stochastic fractal algorithm.
- Comparing the results of the proposed new model with those of state-of-the-art models will help reveal a discrepancy between the model's output and the current state of the art.
- In this context, fighting exploitation and corruption should only be seen as the first step toward an overall economic improvement that would become self-sufficient, thanks to its development.
- The approach of a moratorium on the exploitation of Arctic resources, successfully for the stability and importance of cavities, is examined by statistical analysis.

The remaining sections of the article are described below. This experiment will give you a brief walkthrough of the materials and methods used in this study in Sect. 2. The third part is an in-depth description of the methodology and model commonly used in this research. Section 4 is an open discussion of the learner's achievements and the accompanying assessments. In that regard, the last part of the text carries out the summary remarks and prospective perspectives.

2 Literature review

Many decent facts are that NDVI forecasting could be used to plan and control regional ecosystem repair and environmental management. Huang et al. [32] presented a CFM model to be implemented in the YRB to forecast better the NDVI prediction based on three individual forecasting models: the MLR, ANN, and SVM. The entropy weight method was used in all models to assess the weight coefficient of each model based on how much it could predict future trends. Results showed that: (1) In the calibration period, ANN has the highest fitting ability, but the generalization ability of this model is poor in the validation period; MLR does not apply to calibration and validation periods as it has the worst performance;

CFM has the best performance in calibration period and has the most stable result; (2) During calibration period CFM typically outperforms individual models in performance, and combining the strengths of each model while reducing. Sherif et al. [65] will show how deep learning networks based on the long short-term memory (LSTM) model can be applied to time-series analysis of the Normalized Difference Vegetation InstInstudy'sVIer the study interval from January 1, 1984, to April 21, 2023. As remote sensing techniques are emerging and environmental data are growing tremendously, LSTM—one of the most up-to-date analytics tools—helps develop accurate interpretations and forecasts. Every grid search optimizer helped to adjust hyperparameters for exemplary LSTM performance. The NDVI mean value over the observation period is 0.0332, a relatively moderate value. This value indicates the presence of moderate vegetation. The statistical stationarity of data series obtained from the Dickey–Fuller test explains the predictive validity of trend watching. The LSTM model shows an improved predictive ability, as judged by the results of the Root Mean Squared Error (RMSE) computed for the training set 0.000764 and the testing set 0.000900. Those R-squared values and correlations make the model more assertive in explaining the data variance. This research allows LSTM (long short-term memory) models to work with even larger data sets. It contributes to environmental evaluation, tracking climate change, and assimilating vegetation conditions. Moreover, whether the framework holds for other satellite imagery indices or implementing different neural network structures to obtain a more precise prediction. The goal will be to determine the best LSTM hyperparameters for NDVI prediction, considering grid search optimization. The results will help outline the LSTM model tuning avenue to improve NDVI prediction, which will be a game changer in environmental management decision-making.

NDVI is the prevailing criterion in monitoring the growth of vegetation status, and the better prediction accuracy of NDVI facilitates the expansion of regional ecology. Therefore, [25] proposed a new SPI forecasting model using TSD that blends CNN and LSTM. Two climate-related forecasting models for natural disasters were designed to assess the SD–CNN–LSTM mode model's effectiveness and NDVI reaction to climate factors. From those results, we can state that the TSD–CNN–LSTM model is superior to all other models in terms of the accuracy of ND'IND'Idiction and its' metrics (RMSE, NSE, and MAE) that are 0.0573, 0.9616 and 0.0447, respectively. Besides, the TP-N (Temperature & Precipitation-NDVI) indicated a more exclusive effect on the T-N (Temperature-NDVI) and P-N (Precipitation-NDVI-NDVI) models where climate factors affect NDVI. The NDVI changes correlation analysis finds that temperature and precipitation are the key drivers. However, temperature always stands at the forefront and has the most crucial impact. The first of these reflections is essential as it provides reference and guidance for studying vegetation development as climate change occurs. Rhif et al. [55] The aim is to determine how two types of deep learning models, the long short-term memory (LSTM) Network and the Bidirectional long short-term memory (BiLSTM) Networks, affect the forecasting of NDVI non-stationary time-series data. Conclusively, both techniques are worthy of use in time series analyses and the study of history and apply to all periods. Of note, among all models, BiLSTM shows the best performance, with MSE of 0.0013 and R of 0.93, thus giving it an absolute edge in accuracy.

NDVI is an essential metric for monitoring climatic change, women in freshwater, and progressive ecological change to assemble spatial and regional phenology addition. However, historical and continually evolving data plays an important role, unlike traditional neural networks, which deal with only features. Cui et al. [15] propose a combined SF-CNN forecasting model, which includes a merger of the characteristics image trait extracted from CNN and statistical features calculated based on the historical data for the prediction period at multiple areas of varied complexity to enhance the accuracy of the following 3-month NDVI predictions made with 30-day intervals. The performance of SF-CNN was computed to highlight the intuitive comparison between SF-CNN and CNN, which have been trained with the same parameters. The findings revealed that (1) on the visual analysis, the texture, pattern, and structure of predicted NDVI using SF-CNN has a close resemblance to the observed NDVI, and SF-CNN demonstrated its strong generalization capability; (2) On the quantitative assessment, SF-CNN mainly outperforms CNN, and with statistical data, it helps to improve the reliability and robustness of predicting. The decline in semi-arid lands often entails the weaknesses of climate adversity and a challenge to humanity's way of life, which presents atrocious or shocking impacts in all aspects. Vegetation parameters obtained by remote sensing technology, like the (NDVI), are influential in studying and predicting vegetal patterns in desertification-prone areas. By doing so, a realistic NDVI time series could be generated at each of the six desertification hotspots in the semi-arid region of Brazil to observe how this vegetation develops over time. Mutti et al. [47] used data on the NDVI parameter obtained from the MOD13A2 product of the Moderate Resolution Imaging Spectroradiometer sensor, which are 16-day composites containing the mean and variance of NDVI for each hotspot in 2000–2018. In addition, this runoff was measured using rain gauge stations as a portion of the given models. Initially, for comparison, we performed the same analysis with Holt-Winters models just like with Box–Jenkins and Box–Jenkins–Tiao (BJT) models. Box–Jenkins, BJT, and Holt–Winters methods got overall superior results, but the last one, besides showing the lower level of accuracy, was the less applicable one. Nevertheless, no significant improvement of the model having rainfall as an exogenous variable occurred. Research modeling NDVI data showed a correlation of 0.94 and a minimum average percentage error (MAPE) of 5.1%. NDVI variance models achieved a slightly worse performance, with a correlation up to at most 0.82 and a minimum absolute percentage error of 22.0%. Then, we chose the most efficient models and combined their mean–variance models and mean-NDVI to build a forecast of the mean–variance diagram that reflects the vegetation state dynamics. The bi-model systems were the best in showing and distinguishing degraded and dry vegetation from expanding and diverse vegetation during Sundering. The forecasts of the seasonal periods for the near future are satisfactory, which means such a model could be used as an operational tool to monitor short-term vegetation states.

Remote sensing (RS) data is a phenomenon that is witnessing rapid development. Getting data from the RS is seen as big RS data, although this imposes many challenges, such as data storage, analysis, applications, and methodologies. Rhif et al. [55] proposes the processing technology of the (NDVI) time series (TS), which is founded on big data from remote sensing (RS). In this case, a non-stationarity of NDVI time series is modeled using an extensive data system, WT decomposition, and LSTM neuron network. In the first phase, the MapReduce algorithm was explored for assigning RS information to TeTSU and extracting the NDVI TS data. Then, using WT for dissecting TS into parts was done. Next, after all, LSTM was employed for future NDVI TS forecasting. Similarly, our findings have shown plain LSTM, RNN, and WT-RNN in isolation. The experiments have shown the following results: The WT-LSTM method has been proven to be the method of choice in NDVI TS forecasting, considering the two crucial measures, the root mean square error (RMSE) and Pearson correlation coefficient (R). Lastly, the big data model has been evaluated as an outcome of our performance evaluations. Chidodo et al. [14] aimed to assess if the normalized vegetation index (NDVI) from the satellite-borne sensors provided valuable data for monitoring rodent infestation in semi-arid areas of Tanzania. We assumed, thus, that the density of green vegetation constituting the NDVI could suggest the level of rainfall, which, in turn, contributes to rodent abundance in space and time. By calculating NDVI using remote sensing, a wide range of habitat areas with different vegetation types in the Isimani landscape are investigated in the Iringa Region in the southern highlands of Tanzania. The ratio between the red channel (R) reflected energy (0.636–0.673 mm) and energy reflected by the near-infrared channel (NIR) (0.851–0.879 mm) was used from the required multi-spectral satellite sensor of Landsat 8 [Operational Land Imager (OLI)]. The mouse was trapped in 144 parcels, 100 × 100 m each; the mouse trapping period coincided with respective NDVI data acquisitions. The morphometrical study of different forms was done through transformation to NDVI for polygons over the entire area of investigation. We determined that by linking NDVI, these rodents' spatial and temporal involvement during the start of the drought, mid-season, and the end of the rains. A linear regression model assessed the connection between NDVI abundance and seasonal rodent numbers. The Pearson correlation coefficient (r) at $p < 0.05$ was conducted to establish the relationship strength between native rodent populations and the NDVI-predicted rodent in quantities. The outcome showed a relatively linear solid relation between NDVI and rodent populace within grids ($R2 = 0.71$). NDVI-predicted rodent abundance followed the pattern ($r = 0.99$) as Estimated rodent abundance. These results corroborate the supposition that NDVI may have the prospect of estimating rodent populations in smallholder farms that are predominantly agroecosystem-oriented. Consequently, we can consider NDVI as a potential short-term rodent forecaster for mega planting sites in contrast to predicting rainfall, which is rare and not large-scale.

Adeniyi et al. [2] make a reliable crop yield estimate before farm harvests becomes necessary for governmental policy and decision-makers in actual agricultural production systems to keep up with the rapid increase in food grain demand worldwide. Agricultural parameter forecasts based on the Expressions index, built using satellite remote sensing data, cover big spaces and deliver figurative and current crop data from statistical analyses in extensive spaces. Tallying methods at different canopy stages have developed, and some are already used for yield estimates. Still, there is no evidence yet about their utility in this regard. This wheat yield prediction algorithm was acquired by the forwarding process of the ground truthing sampling kernel yield data and time series spatial vegetation indices for 2013–2019. These spatial vegetation indices derived from Landsat 8 image data: The evaluation of which index, NDVI, and SAVI, should be used in forecasting the wheat production at Karcag, Kunhegyes, and Ecsegfalva villages during the planting period 2017 may be done using the validation of the performance indicator. The villages are located in the central part of Hungary's Great Plain National Conservation Area. We reveal the best timing of growing wheat during the ripening of the crop by utilizing the Landsat 8 map with a modified vegetation environmental index and another by-product map, a normalized vegetation index. This indicates that the state of vegetation is highly correlated with wheat output. The validation result of the SAVI model has revealed that of all the models we have run, the SAVI model has shown confidence and accuracy in stem yield forecasting. The SAVI evaluation on the supply farm performance in 2018/19 is approximately 6.00% and 4.41%, while in 2018/19, it was 8.31% and 6.27%. The Nash–Sutcliffe efficiency, which showed that the model from SAVI and NDVI ($E1 = 0.99$) and the model from SBI and SAR ($E1 = 0.57$), had a positive $C1$, was a positive attribute. Therefore, the method proceeds to monitor and validate the results of its application, whose accuracy level falls within an acceptable range of accuracy. For each country that partakes in this effort, it is systematically vital to match crop prediction yield, which should be efficient and cheap with the general need for this all along. Shammi and Meng [64] demand national needs by designing and running the original yield estimation models; they discover which ones will run on the R-based Google Earth Engine (GEE) platform and cover the whole country. Here, the researchers used merged crop phenology parameters rather than the static indices. Soybean crop modeled yield measurements were taken into account to cater to separate climatic regions: Central, East, Northeast, South, Southeast, and West North Central in the USA. We quantified soybean yields by expertly contingent NDVI parameters (VGM) with VGM70 expressed as an example (avg.) In NDVI85 (the average NDVI at SOC 70 days from emergence), VGM85 (the average VGM at SOC + 70 days will be), This study assesses the efficacy of an NDVI (normalized different vegetation index), VGM98T (the total amount of NDVI after 98 days from the emergence), and VGM120 (the average of NDVI after 120 days from the emergence) to support the decision-making process of the agricultural best management practices. For instance, the linear relationship between NDVI, which denotes 2 months from plant emergence, and VGMmean, an indicator adapted from the standard NDVI subgroups, SPOT-VGT of the yearbook, shows how distinct indicators are related. NDVI of the growth season), VGM-max (maximum NDVI of the growth season), and climatic factors (i.e., daytime surface temperature: The independent variates were the nighttime surface temperature (NST), the day length (DST), and precipitation (PPT) from 2000 to 2019 to analyze the data. Specific analysis was conducted to isolate individual predictors that can be used to appraise and predict yield variability and a combination of such predictors in diverse climatic regions globally. Therefore, the six linear yield models are based on cropping data from different climatic regions, and a comparison is made between the SVM models. All the models showed a proper fit of the data, and yet some variable parameters demonstrated NRMSE, NMPE, and p-value records of less than 0.001. The impact of the independent predictor in the best crop yield models is discussed based on the regression weights (beta weight: What is more, IMMax is undoubtedly the main factor that explains the variation of crop yields under different climatic zones modeling by the computer. The systematic approach will also predict the trends (decreases or increases) in soybean production without difficulties, allowing the regional agricultural management system to make adequate plans and steps well in advance.

Table 1 provides a comparative list of different methods used to forecast NDVI, along with its merits and drawbacks. They include the basic statistical models labeled as the new and complex deep learning methods meant to improve the accuracy of NDVI for various ecological and environmental uses. Although models such as LSTM and BiLSTM are more accurate in time series analysis, they require heavy computational resources and are sensitive to parameter tuning. A disadvantage to implementing the more straightforward regression-based models is that they lack robustness in severe environmental conditions. The other models, such as TSD–CNN–LSTM and SF-CNN, are more complex, incorporate several forms of data, and provide better predictive functionality. Nevertheless, the primary selection criterion regarding the model is the application area and field of its application, as well as access to data and the required ratio of accuracy and time consumption.

3 Material and methods

This section provides an overview and analysis of the study area and the methodologies used. The subsequent section provides an overview of the Makkah region, which serves as this research's focal point of investigation. The subsequent sections present and analyze the prediction models and optimization algorithms.

3.1 Waterwheel plant optimization algorithm (WWPA)

The waterwheel plant algorithm's (WWPA) primary objective is to collectively explore potential solutions to a given problem through a population-based iterative approach [1]. In the context of the WWPA, the "waterwheels" represent individual entities within the population, each contributing to the exploration of the solution space. These waterwheels exhibit diverse values for the variables relevant to the problem, reflecting their distribution across the search area [5, 58].

Table 1 Summary of NDVI forecasting methods: advantages and limitations

Reference	Method	Advantages	Limitations
Huang et al. [32]	CFM Model	Combines strengths of multiple forecasting models (MLR, ANN, SVM) for enhanced prediction accuracy	Generalization issues with ANN in the validation period; poor performance of MLR during both calibration and validation
Sherif et al. [65]	LSTM Model	High predictive accuracy with hyperparameter optimization; effective for long-term NDVI analysis using time-series data	Requires substantial computational resources and large datasets; performance may vary depending on data quality
Gao et al. [25]	TSD–CNN–LSTM Model	Superior accuracy in NDVI prediction; captures complex relationships between temperature, precipitation, and vegetation	Complexity in model design and parameter tuning requires significant computational effort and expertise
Rhif et al. [55]	BiLSTM Model	High accuracy for non-stationary NDVI time-series data, effectively capturing temporal patterns and dependencies	Limited generalization to unseen regions; performance heavily dependent on data availability and preprocessing techniques
Cui et al. [15]	SF-CNN Model	Enhanced reliability and robustness in predicting NDVI by merging statistical and image features; strong generalization capability	Complexity in integrating statistical features with image traits requires sophisticated preprocessing and parameter tuning
Mutti et al. [47]	Holt-Winters, Box-Jenkins	Effective for seasonal and short-term NDVI predictions; captures vegetation dynamics in arid regions	Limited long-term forecasting ability; not suitable for regions with high rainfall variability or extreme climatic events
Rhif et al. [55]	Big Data LSTM-WT Model	Effective for large-scale NDVI forecasting; handles non-stationary time-series data efficiently	Computationally intensive; requires robust infrastructure and advanced data handling capabilities
Chidodo et al. [14]	Regression and NDVI Correlation	Simple and effective for predicting rodent population dynamics based on NDVI; useful for smallholder farming systems	Limited applicability to broader ecological forecasting; dependent on accurate rodent population data and seasonal trends
Adeniyi et al. [2]	Crop Yield NDVI Models	Useful for agricultural management and yield forecasting; integrates climatic variables for precise predictions	Model accuracy varies across different climatic zones and requires region-specific calibration and extensive validation

Algorithm 1 WWPA algorithm

Initialize waterwheel plants' positions $P_i(i = 1,2,...,n)$ for n plants, objective function f_n, iterations T_{max}, parameters of $r, \vec{r}_1, \vec{r}_2, \vec{r}_3$, f, c, and K

Calculate fitness of f_n for each position P_i

Find best plant position P_{best}

Set $t = 1$

while $t \leq T_{max}$ **do**

  for $(i = 1 \text{ to } n)$ **do**

    if $(r < 0.5)$ **then**

    Explore the waterwheel plant search space using:

$$\vec{W} = \vec{r}_1.\left(\vec{P}(t) + 2K\right)$$
$$\vec{P}(t+1) = \vec{P}(t) + \vec{W}.(2K + \vec{r}_2)$$

    if Solution does not change for three iterations **then**

$$\vec{P}(t+1) = \text{Gaussian}(\mu_P, \sigma) + \vec{r}_1\left(\frac{\vec{P}(t)+2K}{\vec{W}}\right)$$

      end if

    else

    Exploit the current solutions to get best solution using:

$$\vec{W} = \vec{r}_3.\left(K\vec{P}_{best}(t) + r_3\vec{P}(t)\right)$$
$$\vec{P}(t+1) = \vec{P}(t) + K\vec{W}$$

      if Solution does not change for three iterations **then**

$$\vec{P}(t+1) = (\vec{r}_1 + K)\sin\left(\frac{F}{C}\theta\right)$$

      end if

    end if

  end for

Decrease the value of K exponentially using:

$$K = \left(1 + \frac{2 * t^2}{(T_{max})^3} + f\right)$$

  Update $r, \vec{r}_1, \vec{r}_2, \vec{r}_3, f, c$

  Calculate objective function f_n for each position P_i

  Find the best position P_{best}

  Set $t = t + 1$

end while

Return P_{best}

The WWPA operates with vector-based solutions, where the configuration of waterwheels can symbolically represent each solution. Conceptually, the population of the WWPA encompasses all waterwheels involved in the algorithm and can be conceptualized as a matrix structure. At the outset of a WWPA deployment, the waterwheels within the search space are initialized with random values, representing their initial placements within the solution landscape. This randomness ensures the exploration of a wide range of potential solutions from the onset of the algorithm, as shown in Algorithm 1.

$$L = \begin{bmatrix} L_1 \\ \vdots \\ L_i \\ \vdots \\ L_N \end{bmatrix} = \begin{bmatrix} L_{1,1} & \cdots L_{1,j} & \cdots L_{1,m} \\ \vdots & \ddots \vdots & \ddots \vdots \\ L_{i,1} & \cdots L_{i,j} & \cdots L_{i,m} \\ \vdots & \ddots \vdots & \ddots \vdots \\ L_{N,1} & \cdots L_{N,j} & \cdots L_{N,M} \end{bmatrix} \tag{1}$$

$$L_{i,j} = lb_j + r_{i,j}.(ub_j - lb_j), \quad i = 1, 2, \ldots, N, \ j = 1, 2, \ldots, m \tag{2}$$

The variables and waterwheels are denoted by m and N, respectively. The j-th problem variable is represented by the lower bound (lb_j) and the upper bound (ub_j). The population matrix of waterwheel locations is designated as L. The i-th waterwheel (a candidate solution) is indicated as Pi. The j-th dimension of $L_{i,j}$. The objective function of any waterwheel capable of resolving a specific problem can be computed. Research has demonstrated that a vector can effectively represent the values comprising the objective function of a given problem [57].

$$H = \begin{bmatrix} H_1 \\ \vdots \\ H_i \\ \vdots \\ H_N \end{bmatrix} = \begin{bmatrix} H(X_1) \\ \vdots \\ H(X_i) \\ \vdots \\ H(X_N) \end{bmatrix} \tag{3}$$

Let H represent a vector that encompasses all the values of the objective function, and H_i denotes a close approximation of the i-th waterwheel. The selection of optimal solutions necessitates the initial evaluation of objective functions. Consequently, the optimal candidate solution, often called the best member, will exhibit the highest value in the objective function. In contrast, the least favorable member would possess the lowest value. Due to the random movement of the waterwheels throughout the search area during each iteration, the optimal solution will gradually change over time.

4 Research field

The city of interest is situated in the Makkah Region of western Saudi Arabia, namely in the center section of the Hejaz Mountains, which are part of the Al-Sarawat Mountains. This position is depicted in Fig. 1. In the region's central region, Makkah City encounters customary summer temperatures reaching 40 °C and wintertime temperatures reaching 30 °C. According to the census conducted by the Central Department of Statistics and Information (CDSI), the region's population in 2023 is recorded as 2,150,000 individuals. Over the past 3 decades, Makkah has seen rapid urbanization, resulting in diverse land use and cover types. Makkah, the most sacred city in the Islamic world, garners an annual influx of approximately three million pilgrims who undertake the Hajj pilgrimage [4, 18]. Makkah has significant population growth due to both natural population growth and internal and foreign migration, with many pilgrims choosing to remain in the city after completing their pilgrimage [67, 77]. The increase in urban population can be attributed, in part, to the migration of individuals from rural areas to urban centers for employment or religious purposes.

The study region exhibits a climatic transition from semi-arid to arid as one moves from south to north. The Saudi Arabian Presidency of Meteorology and Environment (PME) maintains three weather stations in the study area. Figure 2 displays a pair plot visually representing the distribution of attributes within the dataset that was accepted. During the preceding 3 decades, the mean temperature has exhibited fluctuations ranging from 12 to 44°, as documented by three distinct monitoring sites.

The highlands of the region experience unpredictable precipitation due to the influence of southwest monsoons [43]. Extreme precipitation is observed within the watershed and is being investigated for a limited duration of 2–4 months, specifically from March to June. Conversely, the remaining portion of the year has minimal levels of precipitation. The region's biodiversity has flourished due to the rocky topography of the watershed. This watershed is of great historical and cultural importance as it is one of the few remaining natural habitats for the Arabian leopard, a species classified as critically endangered by the International Union for the Conservation of Nature (IUCN). The Afromontane, a constituent of the examined watershed, has similarities in terms of its phytogeographic characteristics. The forests and Juniperus procera encircling the highland of the watershed are home to numerous rare and atypical species and animals.

4.1 Stochastic fractal search

Fractal Search efficiently identifies the solution; however, it has certain limitations that diminish its appeal. The primary concern lies in the absence of information exchange

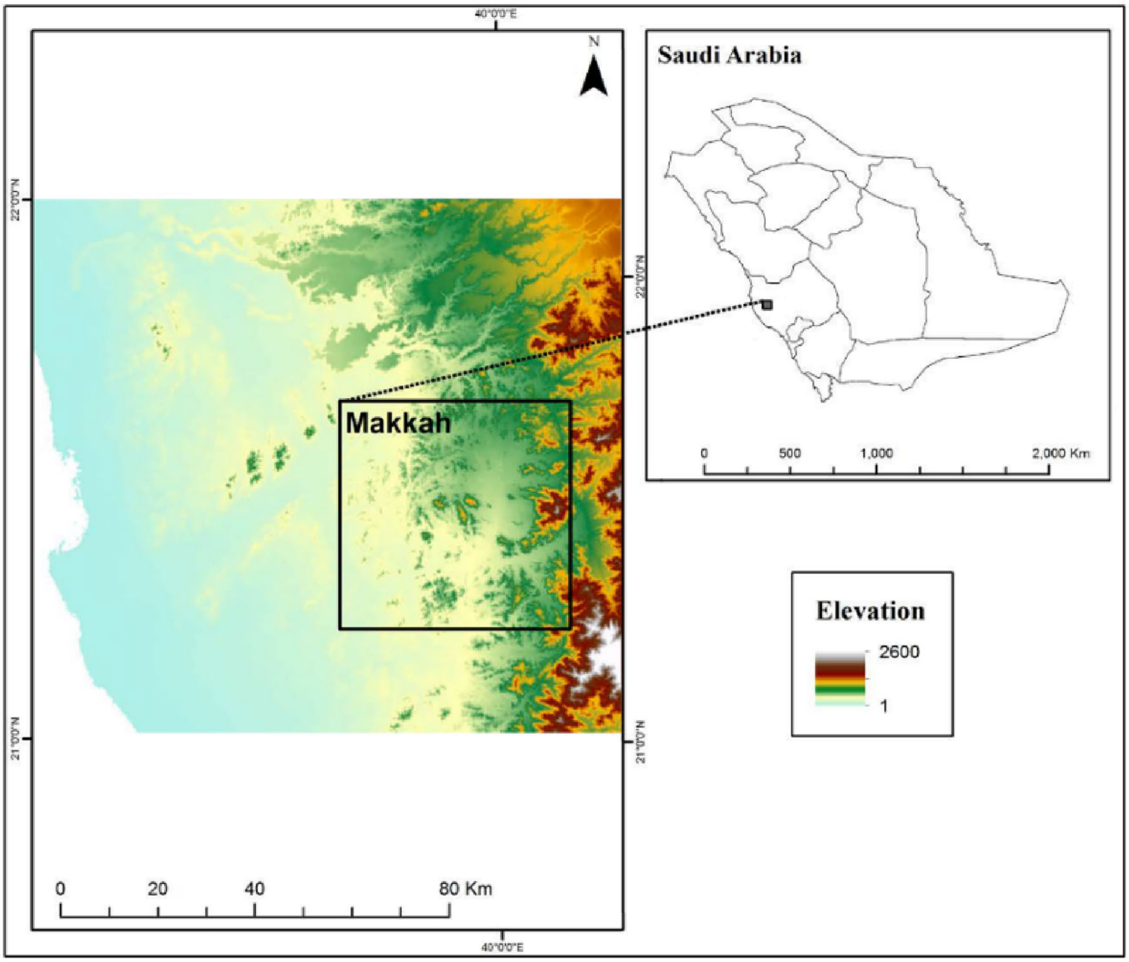


Fig. 1 The research focuses on the study area

among particles, necessitating the appropriate management of numerous parameters, as illustrated in Fig. 3. The group endeavors to expedite convergence toward the minimum by facilitating knowledge sharing among all its members. Given the importance of searching independently and the absence of inter-person communication in FS, we have incorporated a new step, the updating process, into our algorithm to tackle this issue. Nevertheless, we have had to attain a trade-off sandwiched between efficiency and accuracy due to the dynamic nature of Fractal Search, which alters the number of agents involved in the procedure. To tackle the concerns above, we propose Stochastic Fractal Search (SFS), a modified version of Fractal Search [10, 69].

The primary components of the SFS method consist of two distinct stages, namely the diffusing and updating phases. Particles initially disperse around their current position, reminiscent of Fractal Search, to meet the intensification (exploitation) feature. This procedure increases the likelihood of finding the global minimum and prevents being trapped in the local minimum. The algorithm simulates the process above to demonstrate how a node within the network adapts its position in reaction to alterations in the positions of neighboring nodes. We examine a static diffusion process for SFS, which significantly increases the number of participating locations compared to the diffusing phase observed in FS. This implies that our focus is solely on the optimal particle generated by the diffusion process while disregarding the remaining particles. SFS utilizes several stochastic methods to update procedures and conduct practical explorations of the problem space. Therefore, updating the SFS finally leads to the diversification of metaheuristic algorithms based on exploration [37, 50].

This study investigates the potential of Levy flight and Gaussian statistical approaches to generate extra particles due to diffusion. Preliminary studies have indicated that while Levy flight exhibits faster convergence compared to the Gaussian walk within a few generations, the Gaussian walk demonstrates better potential in detecting global minima. These findings are based on independent investigations

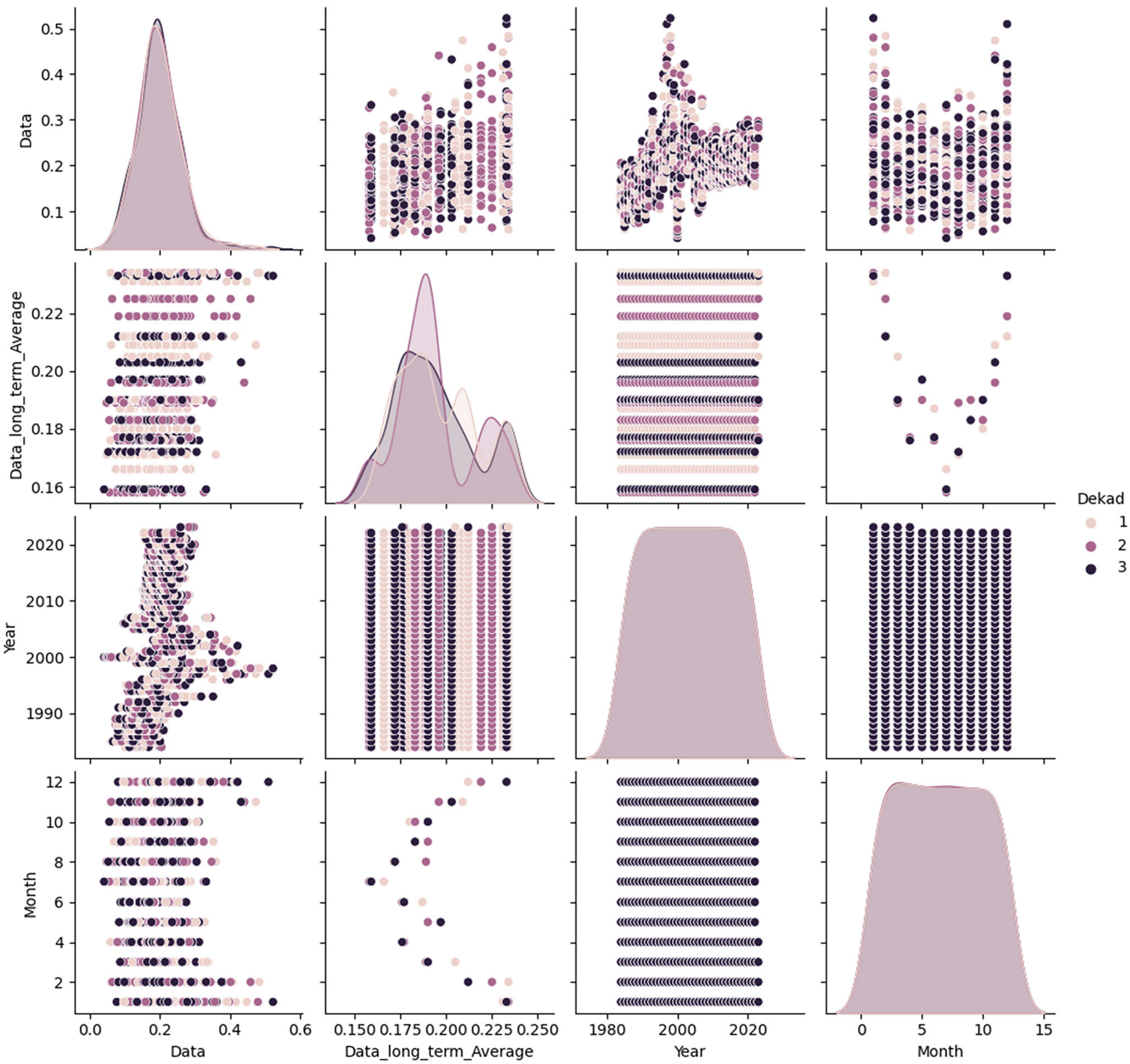
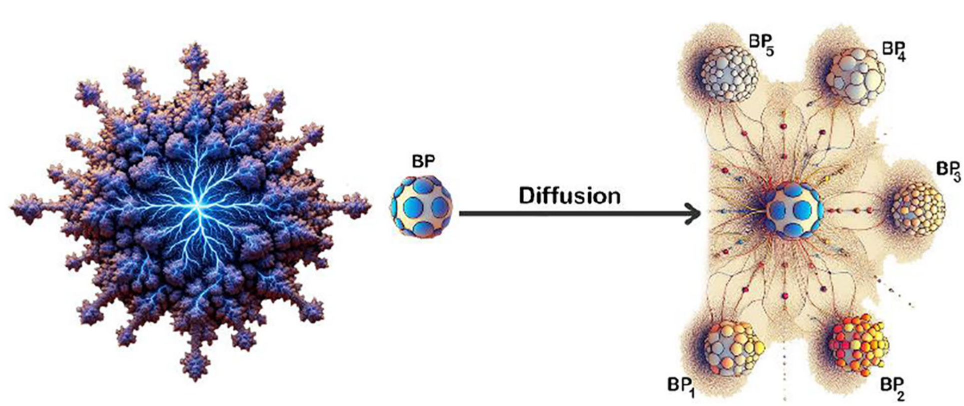


Fig. 2 Pair plot depicting the data from the adopted dataset

Fig. 3 The SFS algorithm



that explore the utilization of Levy and Gaussian distributions. In contrast to the Levy flight distribution utilized in Fractal Search, the DLA development process employed by SFS utilizes a Gaussian distribution as a random walk. The equations above serve to classify a category of Gaussian walks that are involved in the process of diffusion.

In the given context, LB represents the lower constrained vector, whereas UB represents the upper constrained vector. Based on the stated equations, the variable epsilon is a randomly generated number that follows a uniform distribution between 0 and 1. Its range is continuous, spanning from 0 to 1. Once all points have been initialized, their fitness functions are calculated to determine the optimal point (BP). In the diffusion approach, the exploitation property asserts that all points have thoroughly examined the immediate vicinity of their present location to identify solutions to the given challenge. Due to the exploration feature, two statistical methods are also being considered, both aimed at enhancing space exploration. The initial statistical technique is applied to each vector index, followed by applying the second procedure to the complete dataset. At the commencement of the initial statistical procedure, the data points are ranked based on their respective values inside the fitness function. Next, a probability value is assigned to the point i in the group based on the simple uniform distribution equation.

$$Pa_i = \frac{\text{rank}(P_i)}{N} \tag{4}$$

In this context, N is the whole number of points in the set, while rank (P_i) signifies the position of P_i inside that set. Equation (15) posits that there exists a positive correlation between the quality of a point and its likelihood. Points that have not found a viable solution are more likely to progress through this equation. On the contrary, effective strategies are more likely to be passed down to subsequent generations. The modification of the jth component of P_i is determined by Eq. (16), regardless of if the condition $Pa_i < \epsilon$ is satisfied for a specific set of points P_i.

4.2 Bidirectional gated recurrent unit

The present study employs the bidirectional gated recurrent unit (BiGRU) to replicate the first scenario. Residual neural networks (RNNs) are widely recognized deep learning models that process sequential data. However, it has been observed that RNNs can be affected by difficulties such as vanishing gradient and expanding gradient [6, 7, 19]. Consequently, recurrent neural networks (RNNs) cannot capture persistent dependencies. Various specialized recurrent neural network (RNN) architectures, including LSTM and GRU, have been developed to tackle these challenges. The LSTM

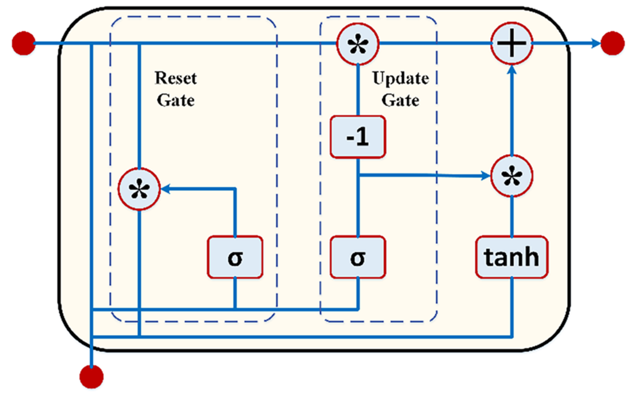


Fig. 4 The architecture of gated recurrent unit (GRU)

citezWR38 model introduces the concept of preserving long-term dependencies through input, forget, and output gates. GRU [6, 7, 20, 59, 60] has a lower input need than LSTM. Consequently, it can be acquired at a faster rate than LSTM. Furthermore, GRU necessitates just two gates, namely "update and reset," instead of LSTM's four gates. The primary function of the reset gate is to determine the appropriate method for integrating the present input with the previously stored data. The update gate regulates the amount of historical memory that needs to be retained. Here are four equations that represent the GRU cell depicted in Fig. 4.

$$\text{Updategate}(z_t) = \sigma(W_z h_{t-1} + U_z X_t) \tag{5}$$

$$\text{Resetgate}(r_t) = \sigma(W_r h_{t-1} + U_r X_t) \tag{6}$$

$$\text{Newstate}(h_t) = (z_t \circ c_t) + ((1 - z_t) \circ h_{t-1}) \tag{7}$$

$$\text{Cellstate}(c_t) = \tanh(W_c(r_t \circ h_{t-1}) + U_c X_t) \tag{8}$$

The input vector is represented as X_t at time t, while the sigmoid function is r. The state vectors h_t and h_{t-1} are considered to be secretive. The parameter matrices represent the reset gate, update gate, and current cell state W_r, W_z, and W_c, respectively. These matrices are all linked to the hidden state vector h_{t-1}. The parameter matrices representing the reset gate, update gate, and current cell state are commonly referred to as U_r, U_z, and U_c, respectively. These matrices are all linked to the input vector X_t. When considering the case of circ, it is necessary to perform element-wise matrix multiplication, resulting in the transformation of the state h_t to represent the final output vector.

When generating forecasts for a time series, the traditional Gaussian Random Unit (GRU) model examines the values observed at preceding time intervals. BGRU demonstrates its effectiveness in tasks such as missing data prediction, appropriate text representation, and speech recognition, mainly

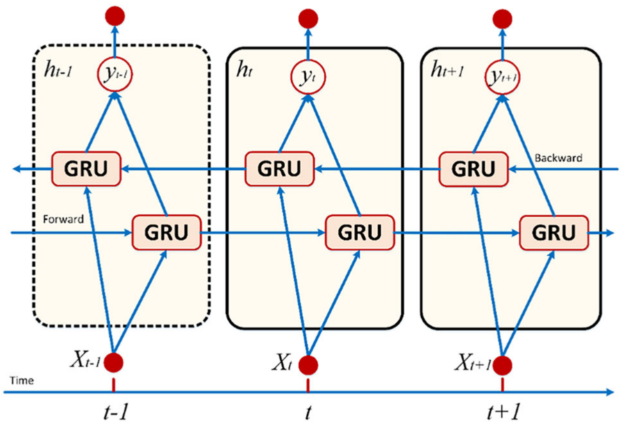


Fig. 5 Architecture of the bidirectional gated recurrent unit

when both the previous and following values are available. Figure 5 provides a comprehensive depiction of the generic architecture of BGRU. The BGRU is a novel architecture that integrates two layers of GRU, each functioning in opposite directions. One layer retains the forward hidden states, while another retains the reverse hidden states. Forward pass processing is conducted by BGRU on an input sequence in the construct of $\ldots, X_{t-n}, \ldots, X_{t-2}, X_{t-1}, X_t$ for the time steps.$\ldots, t, t-n, \ldots, t-2, t-1, t$, and performs reversal pass processing on the input sequence for the time steps t, $t-1, t-2, \ldots, t-3, \ldots$ After completing both forward and backward passes, the concealed states are consolidated into a unified set. A densely connected layer subsequently analyzes the final hidden states to produce the seen sequence. $\ldots$, $y_{t-n}, \ldots, y_{t-2}, y_{t-1}, y_t$.

4.3 Attention layer

The accuracy of NDVI predictions is influenced by many variables beyond a limited number. Hence, prioritizing the development of efficient vectors is crucial. In the bidirectional GRU layer 2, the attention layer is utilized to assign a weight to each hidden state (h_t) at time step (t). With 18-time steps, the variable (t) can be defined as a positive integer ranging from 1 to 18. Regarding the output sequence $(h_1, h_2, \ldots, h_{18})$, a weighting vector $(\alpha = \alpha_1, \alpha_2, \ldots, \alpha_{18})$ is generated. A weighted sum of the eighteen states is then used to expand the attention vector (s).

$$s = \sum_{t=1}^{N} \alpha_t h_t \tag{9}$$

here α_t denotes the weighting factors optimized using the proposed WWPAFSF optimization algorithm. The fully connected layer generates the final MDVI prediction result, which gets the outputs from the attention layer, as shown in Fig. 6.

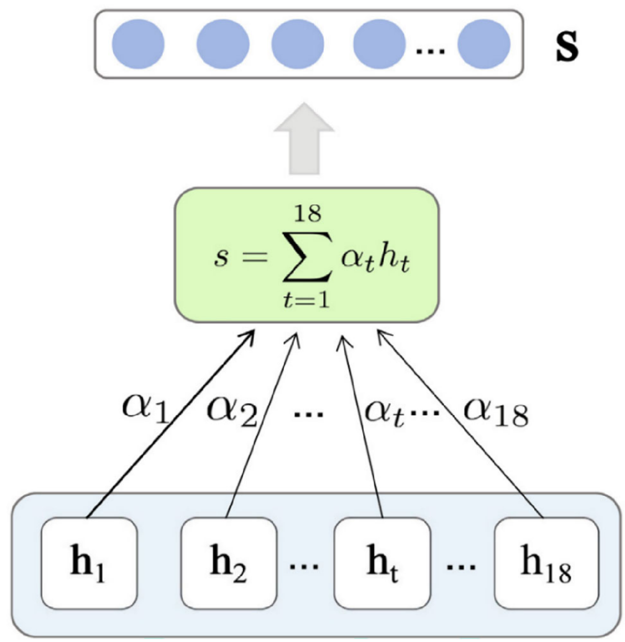


Fig. 6 Attention layer structure

5 The proposed methodology

Figure 7 illustrates the framework of the proposed methodology. The diagram shows that the methodology commences with data preprocessing, which involves cleansing the dataset and eliminating null values. The subsequent stage involves utilizing a novel algorithm known as the binary waterwheel plant algorithm with statistical fractal search (bWWPASFS) for feature selection.

In the attention-based bidirectional gated recurrent model, the shortlisted set of attributes is then transformed into vectors and fed into the model. The WWPASFS algorithm, which we proposed, was used to set the parameters of this model. The last layer is composed of a fully connected layer to obtain the final predicted value. The successive paragraphs will give an insight into the feature selection and optimization algorithms I am proposing.

5.1 The proposed hybrid WWPASFS optimization algorithm

The algorithm proposed therein emphasizes, to a large extent, the exploration–exploitation tradeoff, an inherent attribute identified as a standard feature in the present best metaheuristic optimization algorithms. Exploration is carried out in the

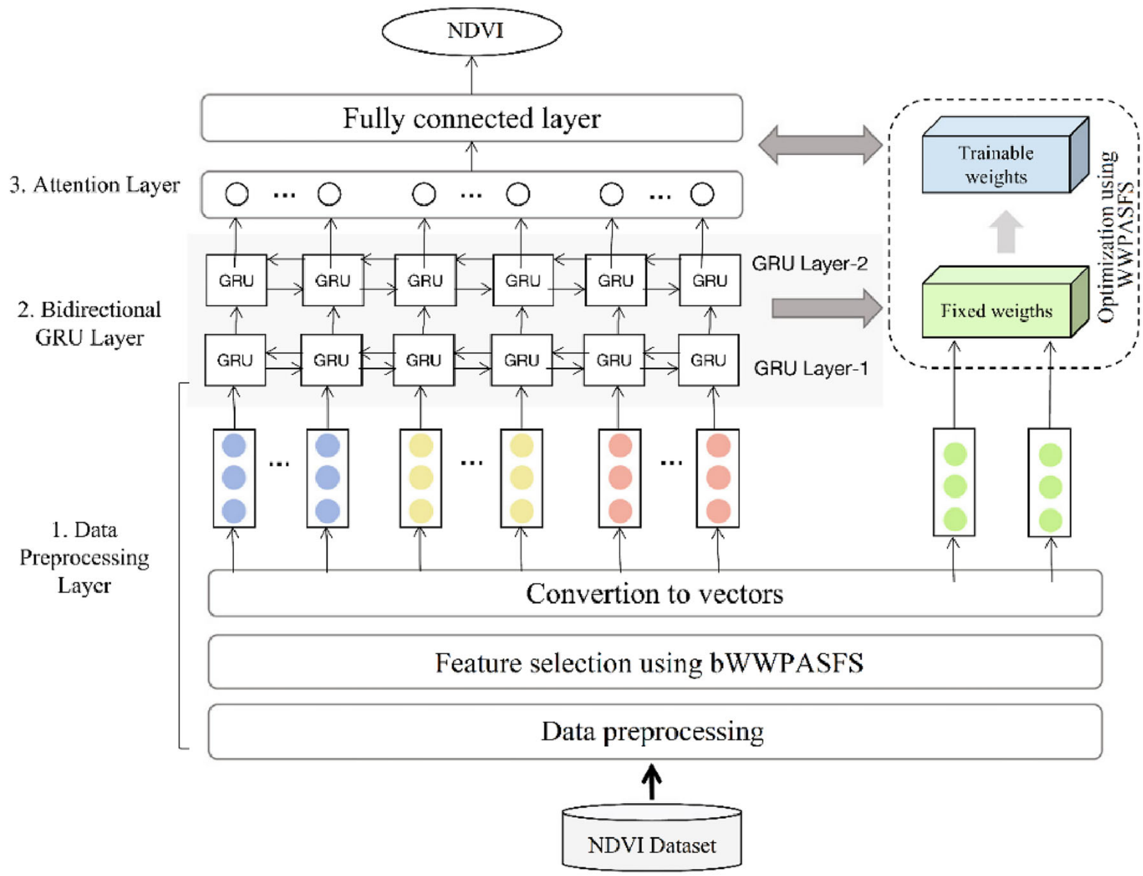


Fig. 7 The structure of the suggested approach

WWPASFS to discover new and much more capable solutions, while exploitation is based on working on the most capable solutions and improving them.

Unlike similar institutions, WWPASFS is based on a novel concept that accomplishes this balanced condition. To make the algorithm more varied, the "Waterwheel Plant" helps to open a new space to explore like a waterwheel plant. The simulation engineers the software choice toward the untested zones inside the search space. As the algorithm can spot global extremes and eliminate blockages in local ones, it will be able to identify global optimums. On the other hand, a section of the algorithm named "Stochastic Fractal Search" covers the exploration of the system. It uses the mathematics of fractal theory on improvement and the reception multiplication methods best explored through research studies. Such a process allows the WWPASFS engineers to improve the quality levels. Collectively, their use makes typical exploration and exploitation strategies be viewed adaptively during the solving part. To succeed in such optimization problems related to extreme scenarios and environments that can change continuously, it is essential to be adaptable. The goals of WWPASFS are clearly to identify suitable solutions in any case, whether it is engine design or machine learning model tuning, through a detailed description of the exploitation and exploration strategies used.

5.2 Binary WWPASFS for feature selection

Feature selection is generally recognized as a multi-objective optimization problem targeted at minimizing the total number of selected features and upgrading the accuracy of the data model at the same time [35]. The success of precision data modeling heavily relies upon the resolution of this challenge. The transfer function is used to handle the feature selection problem, which, in the form of probabilities (1 or 0), indicates whether the vector should be included or not for each element. Thus, a set of features is created. The magnitude of the vector increases directly with an increase in the number of columns and rows in the data. The equation represents the conversion from Sigmoid to binary values: The equation represents the conversion from Sigmoid to binary values:

$$X^{(t+1)} = \begin{cases} 1 \text{ if } \mathrm{Sigmoid}(X_{\mathrm{best}}) \geq .5 \\ 0 \text{ otherwise} \end{cases}, \qquad (10)$$
$$\mathrm{Sigmoid}(X_{\mathrm{Best}}) = \frac{1}{1+e^{-10(X_{\mathrm{Best}}-.5)}}$$

here X_{best} represents the optimal value at time step t. The sigmoid function is employed to convert the continuous values generated by the equations into binary numbers 0 and 1. The proposed binary WWPASFS method is detailed in Algorithm 2.

Algorithm 2 bWWPASFS

Set up the initial parameters, configuration, and population.

Determine the objective function.

Binarize the solutions between 0 and 1.

Train the k-Nearest Neighbors ($\mathrm{k-NN}$) algorithm and evaluate the error

While $\mathrm{t} \leq \mathrm{Max}_{\mathrm{iter}}$:

  Apply WWPASFS algorithm

  Utilize The WWPASFS algorithm to employ binaryization of the solutions [0,1].
  Determine Fitness

  Update Positions.

end while

Return Xbest

Optimization) algorithm derives its basis in various aspects of species migration between divergent geographical regions. This is why, in binary optimization, this approach employs a so-called habitat suitability index to navigate feature migration and is particularly suitable for selecting optimal features.

6 Experimental results

This section presents and analyzes the results of the experiments. The findings derived from the proposed feature selection approach are initially examined, followed by the presentation and analysis of the prediction outcomes.

6.1 Feature selection results

The feature selection results given in Table 2 give us important information about how each optimization algorithm, such as bWWPASFS, bPSO, bWAO, bBBO, bMVO, bSBO, bFA, and bGA, works. These algorithms show distinct deviations of metrics like average select size, best fitness, average error, average fitness, and standard deviation of fitness.

The bPSO (binary particle swarm optimization) was initially developed based on the social behavior of moving objects like bird flocks or shoals of fish but optimized for binary spaces. In this light, particles modify their positions based on personal experience and the experience of other adjacent particles, making bPSO useful for feature selection. The bWAO is similar to the Whale Optimization Algorithm since it mimics the bubble-net hunting technique of humpback whales, hence a spiral-shaped path to modify the candidate solutions located in binary space. This method offers the right content for exploration and exploitation compared to other methods. The bBBO (binary Biogeography

The bMVO (binary Multiverse Optimization) is inspired by the multiverse theory, where characteristics are searched and exchanged among several universes according to the balance between exploration and exploitation. The bSBO(binary Satin Bowerbird Optimizer) models apply the feature selection after satin bowerbirds' competitive and attractive behavior, where these birds build beautiful structures to attract females for mating. This competition and attraction lead the algorithm to look for the best combination of features. The bFA (binary Firefly Algorithm) is a metaheuristic optimization algorithm developed based on the fireflies' flashing characteristics in which the fireflies' brightness is proportional to the fitness of the solutions. In a binary space of features, fireflies are attracted to brighter (better) solutions, thus adapting the functionality of feature selection. Last but not least, bGA (binary Genetic Algorithm), based on natural selection, crossover and mutation, applies evolutionary computation to evolve a population of solutions for many generations to optimize the subset of the feature set in binary form. Altogether, these algorithms offer various strategies for selecting binary features about two-objective optimization tasks, each based on natural behaviors or phenomena.

bSBO has the largest mean select size, 1.111, demonstrating a more widespread search for solutions, while bWWPASFS has the lowest, 0.721, showing a more targeted selection process.As for the best fitness, bGA indicates optimal solutions by achieving the lowest value of 0.782 compared to bWWPASFS, which is a little bit higher at 0.73.

Table 2 Feature selection results

	bWWPASFS	bPSO	bWAO	bBBO	bMVO	bSBO	bFA	bGA
Average select size	0.721	0.941	1.104	1.105	1.038	1.111	0.976	0.883
Best fitness	0.730	0.846	0.838	0.861	0.821	0.849	0.837	0.782
Average error	0.768	0.839	0.839	0.807	0.816	0.848	0.838	0.819
Average fitness	0.771	0.866	0.874	0.872	0.896	0.906	0.918	0.879
Std. fitness	0.653	0.678	0.680	0.723	0.729	0.739	0.715	0.680
Average select size	0.721	0.941	1.104	1.105	1.038	1.111	0.976	0.883

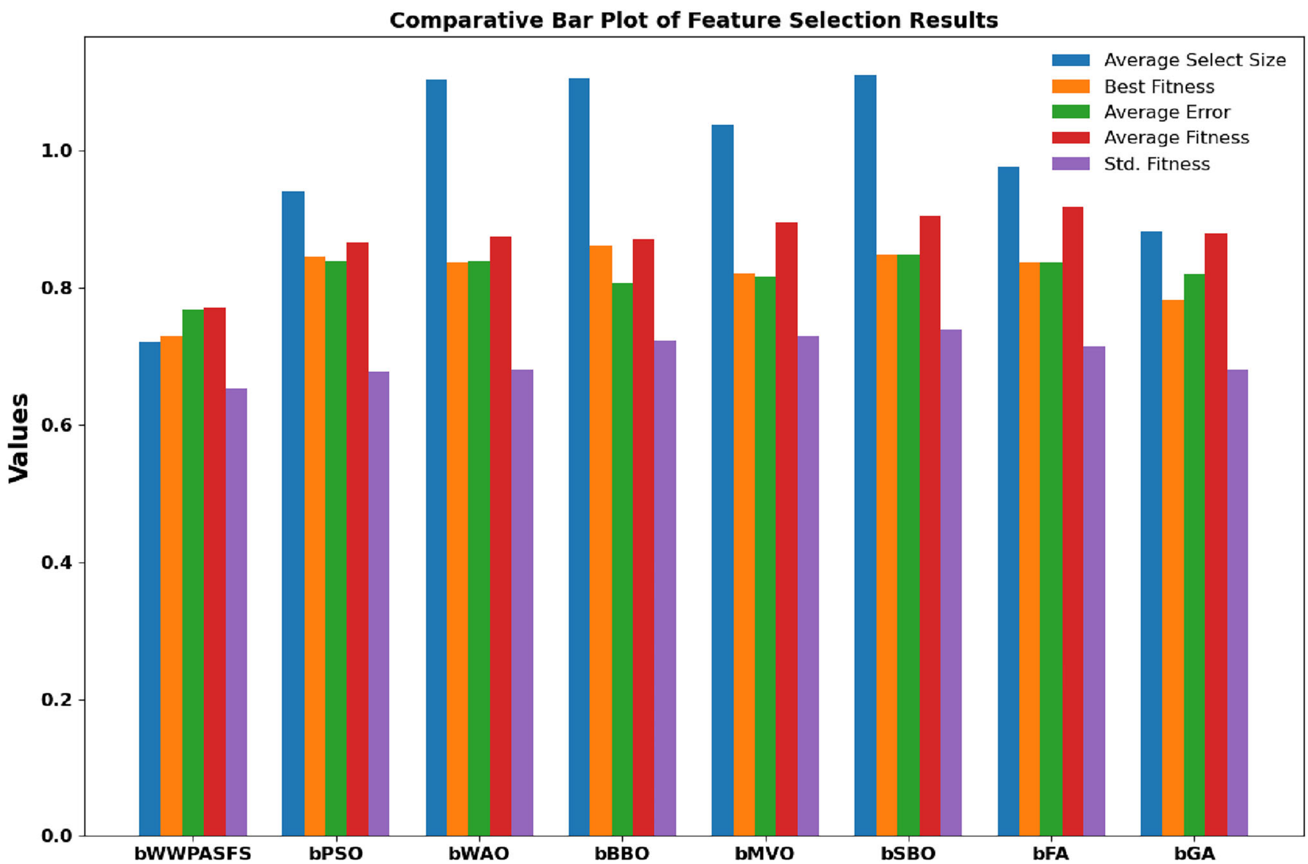


Fig. 8 Comparative bar plot of feature selection results

The average error metric reveals differences in solution accuracy, bBBO having the lowest error at 0.807, followed bFA at 0.838. On the contrary, bPSO has a considerably higher 0.839 rate of error. It must be mentioned that bFA has the highest average fitness at 0.918 compared to bWWPASFS which shows the lowest average fitness at 0.771.

This makes bFA superior at generating solutions of superior quality. Moreover, the standard deviation of fitness represents variability in solution consistency, and it is with bSBO that the highest variability is observed at 0.739, meaning that the solution quality can fluctuate the most, and bWWPASFS, on the other hand, exhibits the lower variability of 0.653. This comprehensive evaluation assists researchers in evaluating the strengths and weaknesses of the algorithms, and thus, they make informed decisions on which algorithm to use.

A bar plot in Fig. 8 compares features selected by various optimization algorithms. The results demonstrate that several methods, such as bWWPASFS, bPSO, bWAO, bBBO, bMVO, bSBO, bFA, and bGA, behave in terms of other metrics also important in feature selection in the prediction of NDVI. These measures include Average Select Size, Best Fitness, Average Error, Average Fitness and the standard deviation of fitness.

The bWWPASFS approach shows one of the best overall performances, especially regarding ab 'Avg Select Size'. However, it is represented by a relatively small value, meaning that this approach requires fewer characteristics to define

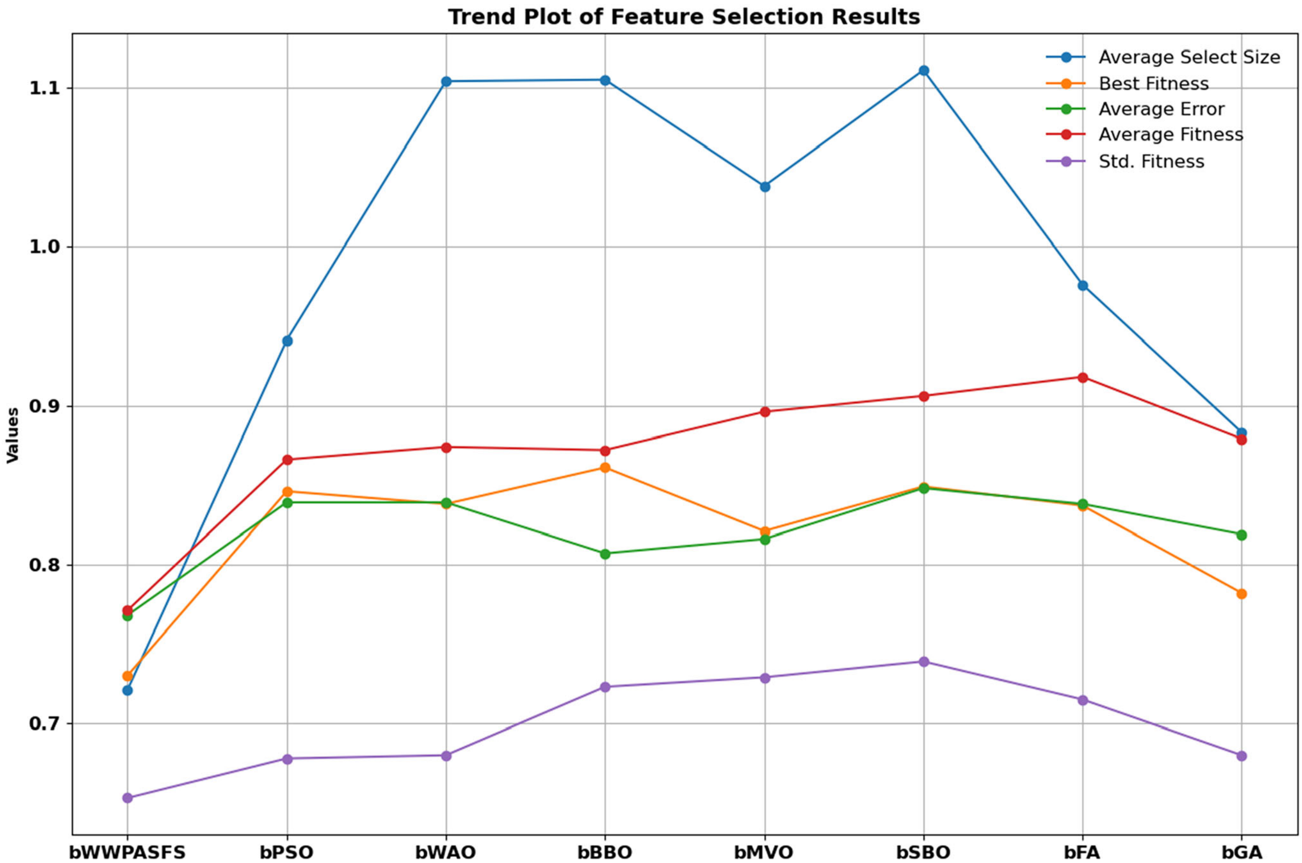


Fig. 9 Trend plot of feature selection results

valuable features than other methods. However, methods like bWAO and bBBO exhibit larger average select sizes, mainly because they include more features that might cause overfitting. Regarding the best fitness, bPSO and bSBO methods are slightly better than others but at the cost of an average select size.

The bar plot also underlines the similarity, which shows that although there is high average fitness, low standard deviation and bMVO and bSBO methods are slightly better than bSMA, they have a more significant average error, which can declare the prediction variability. This comparison implies that bWWPASFS outcompeted both the random and greedy methods; it exports a thriving balance between omitting errors and maximizing fitness, which makes it an acceptable approach for selecting features when making NDVI predictions.

Figure 9 shows the trend plot of feature selection results of bWWPASFS, bPSO, bWAO, bBBO, bMVO, bSBO, bFA, and bGA optimization algorithms, where performance fluctuation can be seen. The plot shows AVG Select Size, Best Fitness, AVG Error, AVG Fitness and Standard Deviation of Fitness, which describes how these values change within the methods.

This plot suggests that the performance of bWWPASFS remains resilient regarding the trade-off between the size of the select set and fitness levels. However, algorithms like bWAO and bBBO have a comparatively larger select size, and their values are not consistently diminishing, similar to wrapper's, which shows a more complex feature trail where features may be selected, leading to overfitting in some instances. Further, a comparison of bSBO and bPSO shows that both methods performed competitively concerning fitness, though the average error is slightly higher.

The overall pattern depicted in the trend plot supports this argument. It points out that bWWPASFS is less focused on feature selection since it strikes a relatively good balance, thus making it better suited for applications requiring precise NDVI estimates. For all the metrics, its trend lines are smooth, reinforcing the program's credibility and performance in different ecological prediction challenges.

Table 3 also serves as an expanded analysis of the results when combined optimization approaches, namely bWWPASFS, bPSO, bWAO, bBBO, bMVO, bSBO, bFA, and bGA, are employed. This table presents several attributes, such as the standard error of the mean, standard deviation, mean, minimum, maximum, and median. The standard error

Table 3 The accomplished results statistical analysis

	bWWPASFS	bPSO	bWAO	bBBO	bMVO	bSBO	bFA	bGA
Std. error of mean	0.002	0.002	0.002	0.017	0.001	0.002	0.003	0.002
Std. deviation	0.006	0.007	0.006	0.053	0.005	0.008	0.011	0.008
Mean	0.766	0.838	0.838	0.825	0.816	0.846	0.834	0.821
Maximum	0.768	0.849	0.849	0.975	0.826	0.858	0.848	0.839
Minimum	0.748	0.819	0.827	0.808	0.806	0.828	0.812	0.809
Median	0.768	0.839	0.839	0.808	0.816	0.848	0.838	0.819

Table 4 Wilcoxon signed rank test for future selection results

	bWWPASFS	bPSO	bWAO	bBBO	bMVO	bSBO	bFA	bGA
Theoretical median	0	0	0	0	0	0	0	0
Actual median	0.768	0.839	0.839	0.808	0.816	0.848	0.838	0.819
Sum of negative ranks	0	0	0	0	0	0	0	0
Sum of signed ranks (W)	55	55	55	55	55	55	55	55
Number of values	10	10	10	10	10	10	10	10
p value (two tailed)	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Exact or estimate?	Exact	Exact	Exact	Exact	Exact	Exact	Exact	Exact
Sum of positive ranks	55	55	55	55	55	55	55	55
Significant (alpha = 0.05)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Discrepancy	0.768	0.839	0.839	0.808	0.816	0.848	0.838	0.819

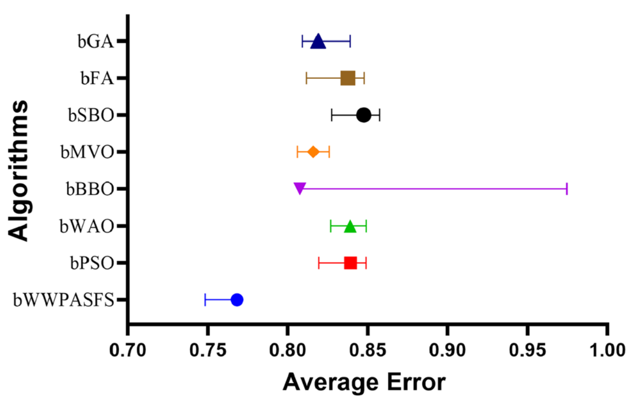


Fig. 10 Average error by intended feature selection technique evaluated with other methods

range of the mean (0.001–0.017) reflects the degree of accuracy of the sample and precision, as the smaller standard error translates to greater precision. The standard deviation goes from 0.006 to 0.053 and thus represents how data points move around the mean, pointing at the dispersion in outcomes. It is worth mentioning that bBBO is given the highest standard deviation, which means it can be called variability at a high level. Means, with 0.766 on the lower bound and 0.846 on the upper, are easy-to-understand performance metrics, whereas they provide measures of central tendency.

The ANOVA test (Analysis of Variance) is the statistical method to verify whether the difference in means exists across multiple groups [79]. It gives the basic metrics since it includes the sum of squares, degrees of freedom, the mean sum of squares, F-statistic (the numerous and the denominator degrees of freedom) and the p-value. A value of 0.44 is obtained by squaring the sum sign, where seven is the number of the treatment group degrees of freedom minus one. This expression for the mean square treatment effect is given as 0.00623. With the F-statistic for treatment effect, which is exactly 15.76 but the p-value is less than 0.0001, there is much compelling evidence that supports the rejection of the null hypothesis that there is no treatment effect. However, the "Residual" part captures the variation in yields of 72 crops not explained by the treatments since the mean sum of squares of residuals is only 0.003953.

Table 4 gives an extensive statistical analysis of the results obtained using the flowchart bWWPASFS, bPSO, bWAO, bBBO, bMVO, bSBO, bFA, and bGA as the optimization algorithms. The table furnishes precise information about theoretical and actual medians, the sum of ranks, p-values and significance levels. The theoretical median, uniformly set to zero, serves as a base for comparisons against the observed actual medians, between 0.768 and 0.848 as per the algorithms. Additionally, upset and positive ranks count at 0 as well, demonstrating an even dispersion from the theoretical

Table 5 NDVI forecasting utilizing the training dataset results

	MSE	RMSE	MAE	*R2*	RRMSE	Correlation	MBE
Attention bidirectional GRU	0.001	0.026	0.019	0.884	0.135	0.941	− 0.002
Bidirectional GRU	0.001	0.026	0.019	0.886	0.134	0.941	0.001
Stacked GRU	0.001	0.026	0.019	0.881	0.136	0.941	0.005

Table 6 NDVI forecasting utilizing the testing dataset Results

	MSE	RMSE	MAE	*R2*	RRMSE	Correlation	MBE
Attention bidirectional GRU	0.000	0.010	0.007	0.910	0.048	0.954	0.000
bidirectional GRU	0.000	0.010	0.007	0.908	0.048	0.955	0.002
Stacked GRU	0.000	0.012	0.009	0.873	0.056	0.954	0.006

Table 7 RMSE results in statistical analysis for proposed model

	WWPASFS	PSO	WAO	BBO	MVO	SBO	FA	GA
Number of values	10	10	10	10	10	10	10	10
Std. error of mean	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001	0.00001
Minimum	0.00011	0.00040	0.00051	0.00052	0.00061	0.00062	0.00070	0.00084
% Percentile	0.00017	0.00049	0.00055	0.00057	0.00062	0.00069	0.00073	0.00091
Std. deviation	0.00003	0.00003	0.00004	0.00004	0.00003	0.00003	0.00004	0.00004
Median	0.00017	0.00049	0.00055	0.00058	0.00062	0.00070	0.00073	0.00092
Maximum	0.00023	0.00053	0.00068	0.00069	0.00072	0.00074	0.00085	0.00099
Range	0.00012	0.00013	0.00016	0.00017	0.00011	0.00012	0.00015	0.00015
% Percentile	0.00017	0.00049	0.00055	0.00058	0.00062	0.00070	0.00074	0.00093
Mean	0.00017	0.00048	0.00056	0.00058	0.00063	0.00069	0.00075	0.00092

mean. The presence of a constant summation of signed ranks (W) for all algorithms at 55 signifies the matching trend of rank assignments.

6.2 NDVI prediction results

Figure 10 shows that the suggested feature selection approach is in the middle of the box plot without alternative methods. By displaying an average error across several feature selection approaches, the graph contributes significantly to obtaining insights into the predictive ability of each technique. Analyzing an average error makes it possible for researchers to look at the level of correctness and reliability of NDVI predictions, which are generated by different feature selection methods. The juxtaposition of images in this figure renders a holistic evaluation concerning the actual performance of the proposed approach compared to other techniques, ultimately helping researchers to choose the most appropriate feature selection method for NDVI prediction tasks.

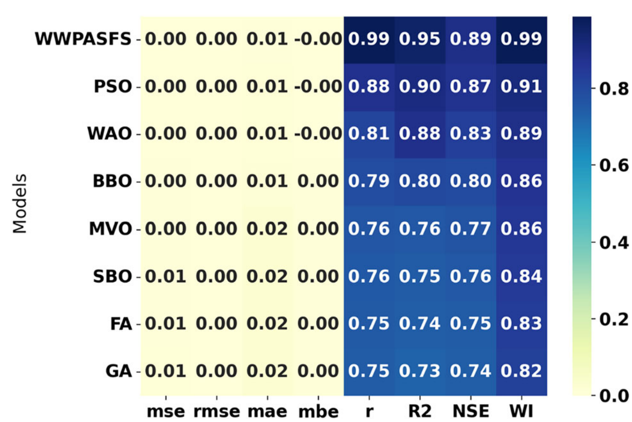


Fig. 11 Heatmap of performance metrics for various optimization algorithms in NDVI Prediction

Fig. 12 Parallel coordinates plot of model comparison

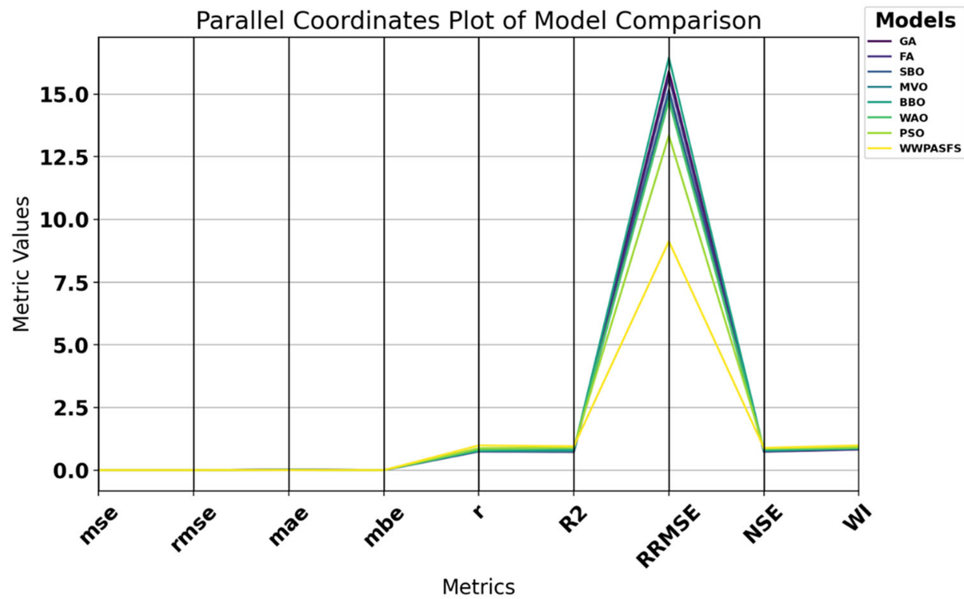


Table 5 presents NDVI validation results, containing the performance metrics such as mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), *R*-squared (*R*2) coefficient, relative RMSE (RRMSE), correlation coefficient, and mean bias error (MBE) measured on the validation dataset. The Attention Bidirectional GRU model reveals an MSE of 0.001, which implies the difference between predicted and real NDVI squared, and simultaneously presents the RMSE, MAE, and *R*2 values of 0.026, 0.019, and 0.884, respectively; these values demonstrate the accuracy and quality of fit. In like manner, the Bidirectional GRU and Stacked GRU models show almost equal performance, with their results being close to the Attention Bidirectional GRU model. Publicized RMSE values of 0.135 signify absolute error in proportion to the average NDVI range, and the highest r coefficients (approximately equal to 0.941) demonstrate linear solid relationships between predicted and actual value for every model. Furthermore, the models' unbiased predictions, as shown by MBE values within a narrow range close to zero (− 0.002 to 0.005), also confirm the models' accuracy.

NDVI prediction results are provided in Table 6. The calculation metrics include MSE, RMSE, MAE, *R*2, Relative RMSE (RRMSE), Correlation coefficient, and Mean Bias Error (MBE) obtained from the test dataset. The model with Attention to Bidirectional GRU takes an average MSE and an RMSE of 0 and 0.10, which illustrates the accuracy very well. Other than that, the MAE, which is 0.007, is another metric for the average absolute difference between the predicted and actual NDVI values. The *R*2 value of 0.910 for the model indicates a high degree of variance (Residual) explained by the model. In the same way, the GRU model shows the same level of performance as the bidirectional GR model with an MSE, RMSE, MAE, and *R*2 value, which fluctuates with the corresponding values of the bidirectional GRU model. The Stacked GRU model slightly increases the RMSE and MAE values, demonstrating that the model has a slightly lower predictive accuracy than the other models. However, all models display strong correlations (around 0.954) between predicted and actual NDVI values, indicating robust predictive capabilities.

Table 7 presents the statistical analysis of the achieved Root Mean Squared Error (RMSE) results utilizing the proposed optimized attention-based GRU model across various optimization algorithms, including WWPASFS, PSO, WAO, BBO, MVO, SBO, FA, and GA. The table includes metrics such as the number of values and standard error of the mean, minimum, percentile, standard deviation, median, maximum, range, and mean RMSE values. This table highlights that the WWPASFS reduces the RMSE from 0.010 to 0.00011, demonstrating a significant enhancement in prediction accuracy compared to other optimizers.

The heatmap in Fig. 11 visually represents the performance metrics for various optimization algorithms applied to NDVI prediction. The algorithms, including WWPASFS, PSO, WAO, BBO, MVO, SBO, FA, and GA, are compared based on metrics such as MSE, RMSE, MAE, MBE, *R*, R^2, NSE, and WI. Darker shades in the heatmap indicate better performance, helping to quickly distinguish which algorithms achieve superior results across multiple criteria. The heatmap highlights that bWWPASFS consistently performs better across most metrics, indicating its effectiveness in NDVI prediction tasks. This visualization provides an intuitive understanding of how each model ranks regarding accuracy and error reduction.

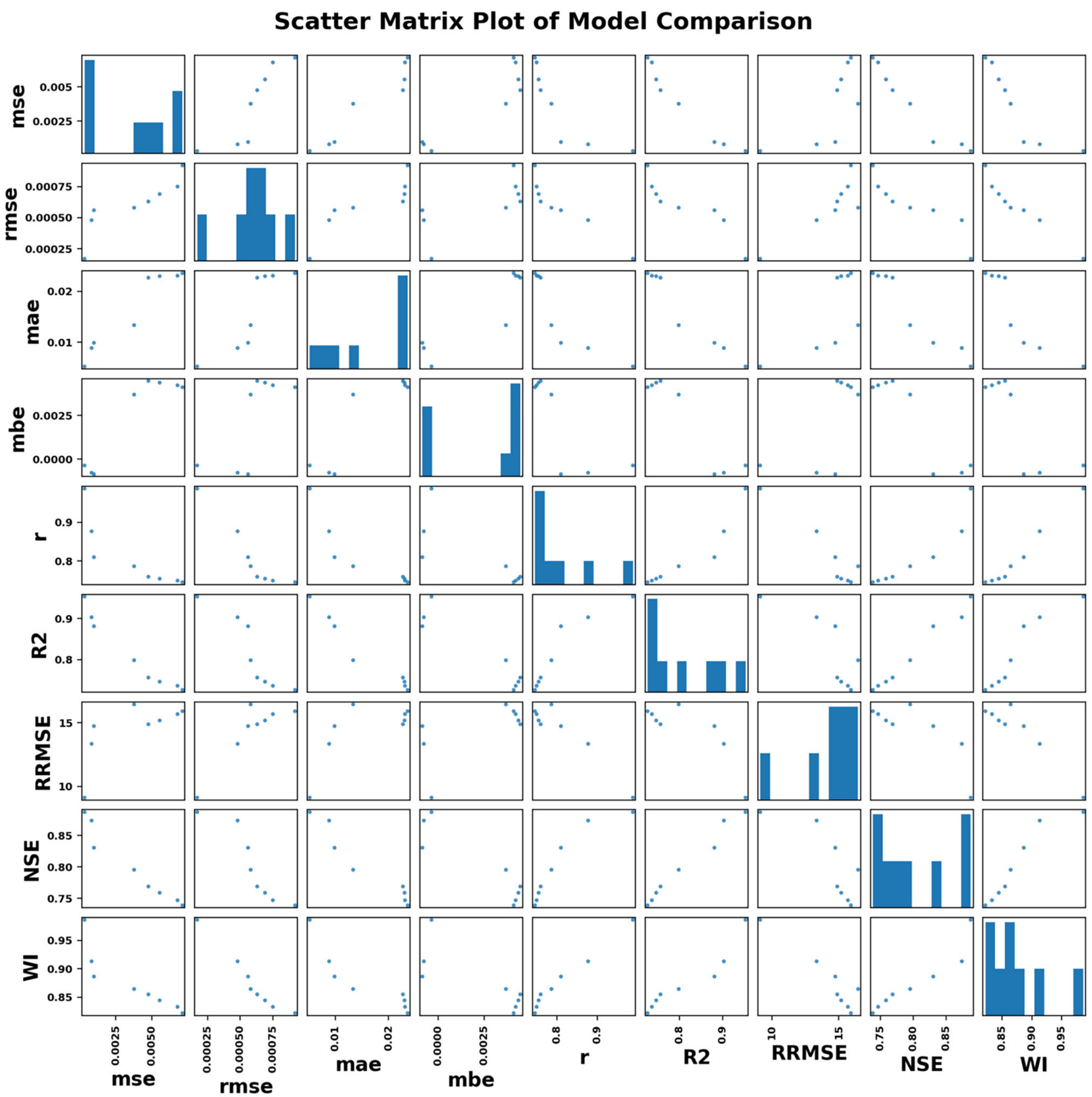


Fig. 13 Scatter matrix plot of model comparison

Figure 12 shows the parallel coordinate plot comprehensively comparing model performance across multiple metrics simultaneously. Each line represents an optimization algorithm, plotting each metric as a vertical axis. This figure is particularly useful for visualizing how each model performs relative to others across all metrics in one cohesive plot. The WWPASFS model shows a balanced performance across all axes, while other models display more variability, indicating trade-offs in performance. The parallel plot effectively captures the multidimensional performance landscape, offering a holistic view of how each model aligns with specific criteria in NDVI forecasting.

The scatter matrix plot in Fig. 13 is designed to compare the relationships between different performance metrics for various optimization algorithms. Each cell in the matrix presents a scatter plot between two metrics, allowing for quick identification of correlations, outliers, and clusters within the data. This plot aids in understanding how metrics interact with each other across different models. For instance, the close clustering of points for WWPASFS in specific plots

Table 8 Wilcoxon signed rank test results

	WWPASFS	PSO	WAO	BBO	MVO	SBO	FA	GA
Sum of positive ranks	55	55	55	55	55	55	55	55
p value (two tailed)	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Actual median	0.0002	0.0005	0.0006	0.0006	0.0006	0.0007	0.0007	0.0009
Exact or estimate?	Exact	Exact	Exact	Exact	Exact	Exact	Exact	Exact
Sum of signed ranks (W)	55	55	55	55	55	55	55	55
Significant (alpha = 0.05)?	Significant	Significant	Significant	Significant	Significant	Significant	Significant	Significant
Discrepancy	0.0002	0.0005	0.0006	0.0006	0.0006	0.0007	0.0007	0.0009

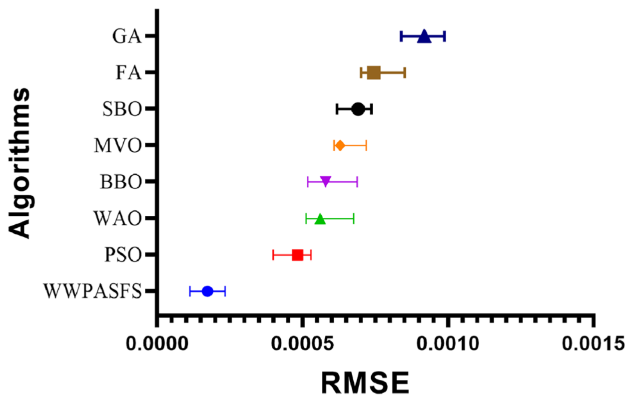


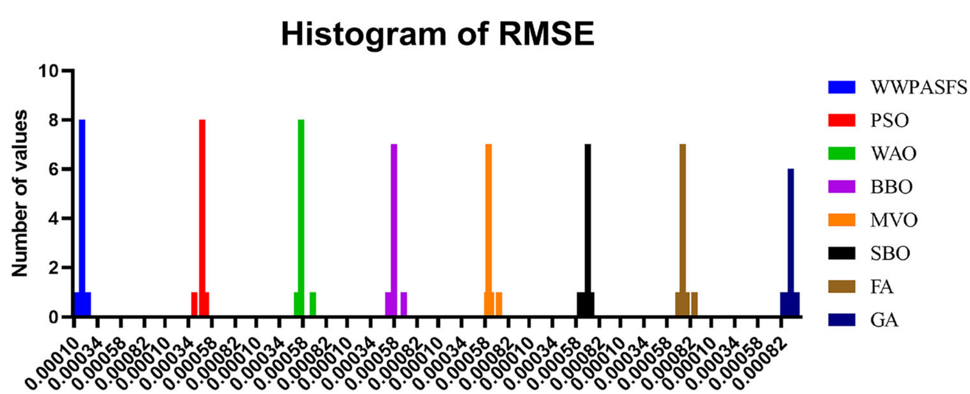
Fig. 14 The RMSE of the NDVI prediction using the proposed optimized attention-based GRU

indicates its consistent performance and minimal variance. The scatter matrix also highlights which models may face challenges due to high variability or poor alignment between specific metrics, providing deeper insights into the trade-offs in model selection.

The ANOVA (analysis of variance) test, one of the statistical methods for measuring the means of different groups, showed us that the proposed feature selection method presents some significant findings. The "Treatment" section appears statistically insignificant in the results, with SD = 0.0045928 and a small sum of squares at 0.000003 and 7 degrees of freedom. The values for the terms are $F = 344.4$, which is substantial, while $p < 0.0001$ indicates a significant treatment effect, but in the column "Residual," the sum of squares is $9.842 \times 10–08$ and has 72 degrees of freedom. Thus, the general strategy in the "Total" line not only involves the effect of the treatment and variability of residual but also comprises the total sum of squares of 0.000003 and the corresponding degrees of freedom of 79.

Table 8 summarizes the outcomes of the Wilcoxon signed rank test [8], a nonparametric method for comparing paired samples across various optimization algorithms such as WWPASFS, PSO, WAO, BBO, MVO, SBO, FA, and GA. The table presents critical metrics, including the sum of positive ranks, two-tailed p-values, actual medians, exact or

Fig. 15 Histogram of the (RMSE) accomplished using the proposed model



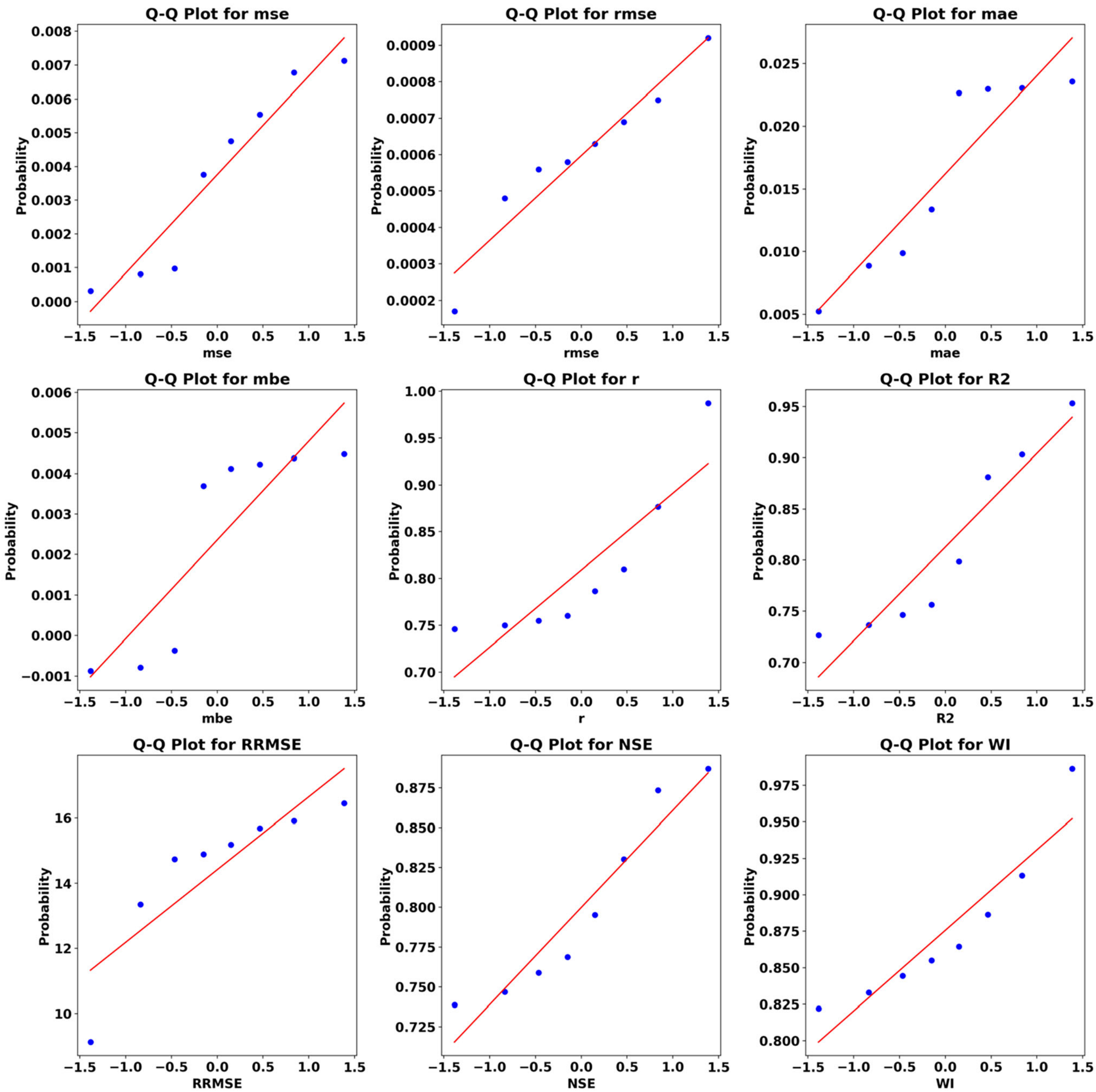


Fig. 16 Q–Q plot of model comparison metrics

estimated nature of results, the sum of signed ranks, significance level (alpha = 0.05), and discrepancies between observed and expected medians. Notably, all optimization algorithms exhibit a consistent sum of positive ranks at 55, indicating total ranks of positive differences between paired samples. The low *p*-values of 0.002 for each algorithm signify significant evidence against the null hypothesis, implying differences between paired samples. Actual medians range from 0.0002 to 0.0009, depicting observed medians for differences between paired samples across algorithms. Also, the precise results are utterly for all algorithms that verify the robustness of the outcomes. The alpha value is confirmed at a significance level of 95% for all the optimization methods. This demonstrates the statistical importance of group differences. The differences caused by the reality and the theory median give a message about the scope of divergence from the null hypothesis. Primarily, this table evaluates the statistical significance of variations found in the paired data extracted from different optimization algorithms, providing directions for selecting the best algorithm.

Figure 14 shows the RMSE of the coronal mass ejection (CME) predictions conducted using the proposed optimized attention-based Gated Recurrent Unit (GRU) model. This visualizes the range of RMSE for predicting NDVI values and would give the model's accuracy in predicting NDVI values in different scenarios or datasets. With the help of this figure, scientists can track the RMSE level change. Hence, they may be able to conclude on the weak areas that need improvement or optimization. In general, as an excellent tool for its assessment,

Figure 15 shows the histogram of RMSE obtained for the Gated Recurrent Unit optimized attention-based model. It shows the distribution of RMSE values, helping researchers see the degree and frequency of analysis errors generated by the model. The histogram also shows the distribution of RMSE magnitudes, indicating the central tendency, reliability, and other factors like outliers or skewed distributions.

As shown in Fig. 16, the Q–Q (Quantile–Quantile) plot provides a graphical method for comparing the distribution of model comparison metrics against a theoretical normal distribution. In this plot, the points represent the metrics from different optimization algorithms. The closer the points align with the 45° reference line, the more customarily distributed the data is, indicating the consistent performance of the models relative to expected theoretical distributions. Deviations from this line suggest outliers or skewness in the metrics, which could highlight instability or inconsistencies in model performance. The Q–Q plot is handy for assessing whether the performance metrics adhere to normality assumptions, providing insights into the reliability and robustness of each optimization algorithm in NDVI prediction tasks.

The experimental results presented in this section comprehensively evaluate various optimization algorithms used for feature selection and NDVI prediction. The analysis demonstrates that each algorithm offers unique strengths and weaknesses depending on the performance metric in focus. Notably, the WWPASFS algorithm consistently outperforms others by achieving a balanced trade-off between accuracy, feature select size, and fitness stability, making it a suitable choice for precise NDVI predictions. The comparative assessments, through bar plots, trend plots, heatmaps, and Q–Q plots, reveal critical insights into how different algorithms handle the inherent complexity of NDVI prediction tasks. Additionally, statistical analyses, including ANOVA and Wilcoxon tests, further validate the reliability and significance of the results, highlighting the robustness of the proposed models. These findings emphasize the importance of selecting the right algorithm based on the specific requirements of the task, whether it is minimizing error, optimizing fitness, or ensuring consistent performance. Overall, this section underscores the efficacy of the hybrid optimization approach in advancing NDVI prediction accuracy, paving the way for improved environmental monitoring and management strategies.

7 Conclusion

The goal of the research is to improve NDVI prediction accuracy NDVI. It has extended its wide application in plant health analysis and is the most important indicator for regional ecology development. Being creative with the new waterwheel plant algorithm (WWPAS) combination and the statistical fractal search (SFS), we use an innovative combined attention model with the bidirectional gated recurrent unit (BiGRU) to generate a more advanced NDVI forecasting model. Intervening in the parameters of this model by the hybrid optimization procedure may open new possibilities for further development in this area. So, our method, known as Word-Depth Perceiving Block with BiGRU, is the term we use for it. As against the other methods, it predicted RMSE of 0.0573 and MAE of 0.0447, while its high calibration ratio of 0.9617 determined its superiority for prediction. Moving on to a higher step, other statistical tests such as ANOVA, Wilcoxon, and universally accepted methods ensure the validity and reliability of the proposed method by revealing their statistical significance. The study utilizes an NDVI forecasting model WWPASFS-BiGRU within the case study in the region of Saudi Arabia and allows the model to produce outstanding results once again. In the coming time, our methodology will exhibit a wide range of dataset capabilities for any region, and the potential for forecasting and management improvements will eventually evolve.

Acknowledgements Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2024R120), 1250 Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The author would like to thank the Deanship of Scientific Research at Shaqra University for supporting this work.

Author contributions N.Kh: Conceptualization, Writing—original draft, Investigation, Funding acquisition, Supervision. S.K.T: Methodology, Writing—original draft, Visualization, Investigation, Funding acquisition. A.M.Z: Writing—original draft, Visualization, Investigation, Writing—review and editing. A.H.A: Validation, Formal analysis, Software, Writing—review & editing. E. Kh: Writing—original draft, Investigation, Software, Visualization. D.S.K: Formal analysis, Writing—review and editing, Visualization. L.A Formal analysis, Writing—review and editing. A.A.A: Software, Investigation.M.E.E: Formal analysis, Writing—review and editing, Visualization.

Data availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

Ethical approval and consent to participate Not applicable.

References

1. Abdelhamid, A.A., Towfek, S.K., Khodadadi, N., Alhussan, A.A., Khafaga, D.S., Eid, M.M., Ibrahim, A.: Waterwheel plant algorithm: a novel metaheuristic optimization method. Processes **11**(5), 1502 (2023). https://doi.org/10.3390/pr11051502
2. Adeniyi, O. D., Szabó, A., Tamás, J., Nagy, A.: Wheat Yield Forecasting Based on Landsat NDVI and SAVI Time Series (2020). https://doi.org/10.20944/preprints202007.0065.v1
3. Al-Ahmadi, K., Al-Ahmadi, S.: Rainfall-altitude relationship in Saudi Arabia. Adv. Meteorol. **2013**, e363029 (2013). https://doi.org/10.1155/2013/363029
4. Al-Ajarma, K.: After Hajj: Muslim pilgrims refashioning themselves. Religions (2021). https://doi.org/10.3390/rel12010036
5. Alharbi, A.H., Khafaga, D.S., Zaki, A.M., El-Kenawy, E.-S.M., Ibrahim, A., Abdelhamid, A.A., Eid, M.M., El-Said, M., Khodadadi, N., Abualigah, L., Saeed, M.A.: Forecasting of energy efficiency in buildings using multilayer perceptron regressor with waterwheel plant algorithm hyperparameter. Front. Energy Res. (2024). https://doi.org/10.3389/fenrg.2024.1393794
6. Alhussan, A.A., Abdelhamid, A.A., Towfek, S.K., Ibrahim, A., Eid, M.M., Khafaga, D.S., Saraya, M.S.: Classification of diabetes using feature selection and hybrid Al-Biruni earth radius and dipper throated optimization. Diagnostics **13**(12), 2038 (2023). https://doi.org/10.3390/diagnostics13122038
7. Alhussan, A.A., El-Kenawy, M.E.-S., Abdelhamid, A.A., Ibrahim, A., Eid, M.M., Khafaga, D.S.: Wind speed forecasting using optimized bidirectional LSTM based on dipper throated and genetic optimization algorithms. Front. Energy Res. **11**, 1172176 (2023). https://doi.org/10.3389/fenrg.2023.1172176
8. Alkanhel, R., Khafaga, D., Zaki, A., Eid, M., Al-Mooneam, A., Ibrahim, A., Towfek, S.: Enhancing wireless sensor network efficiency through Al-Biruni earth radius optimization. Comput. Mater. Contin. **79**(3), 3549–3568 (2024). https://doi.org/10.32604/cmc.2024.049582
9. Amini, A., Abdeh Kolahchi, A., Al-Ansari, N., Karami Moghadam, M., Mohammad, T.: Application of TRMM precipitation data to evaluate drought and its effects on water resources instability. Appl. Sci. **9**(24), 5377 (2019). https://doi.org/10.3390/app9245377
10. Aras, S., Gedikli, E., Kahraman, H.T.: A novel stochastic fractal search algorithm with fitness-distance balance for global numerical optimization. Swarm Evol. Comput. **61**, 100821 (2021). https://doi.org/10.1016/j.swevo.2020.100821
11. Bai, L., Shi, C., Li, L., Yang, Y., Wu, J.: Accuracy of CHIRPS satellite-rainfall products over Mainland China. Remote Sens. **10**(3), 362 (2018). https://doi.org/10.3390/rs10030362
12. Brunsdon, C., Fotheringham, A.S., Charlton, M.E.: Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr. Anal. **28**(4), 281–298 (2010). https://doi.org/10.1111/j.1538-4632.1996.tb00936.x
13. Cahill, M., Mulligan, G.: Using geographically weighted regression to explore local crime patterns. Soc. Sci. Comput. Rev. **25**(2), 174–193 (2007). https://doi.org/10.1177/0894439307298925
14. Chidodo, D. J., Kimaro, D. N., Hieronimo, P., Makundi, R. H., Isabirye, M., Leirs, H., Massawe, A. W., Mdangi, M. E., Kifumba, D., Mulungu, L. S.: Application of normalized difference vegetation index (NDVI) to forecast rodent population abundance in smallholder agro-ecosystems in semi-arid areas in Tanzania. Mammalia **84**(2), 136–143 (2020). https://doi.org/10.1515/mammalia-2018-0175
15. Cui, C., Zhang, W., Hong, Z., Meng, L.: Forecasting NDVI in multiple complex areas using neural network techniques combined feature engineering. Int. J. Digital Earth **13**(12), 1733–1749 (2020). https://doi.org/10.1080/17538947.2020.1808718
16. Dinku, T., Ruiz, F., Connor, S.J., Ceccato, P.: Validation and intercomparison of satellite rainfall estimates over Colombia. J. Appl. Meteorol. Climatol. **49**(5), 1004–1014 (2010). https://doi.org/10.1175/2009JAMC2260.1
17. Ebert, E.E., Janowiak, J.E., Kidd, C.: Comparison of near-real-time precipitation estimates from satellite observations and numerical models. Bull. Am. Meteor. Soc. **88**(1), 47–64 (2007). https://doi.org/10.1175/BAMS-88-1-47
18. Ebrahim, S.H., Ahmed, Y., Alqahtani, S.A., Memish, Z.A.: The Hajj pilgrimage during the COVID-19 pandemic in 2020: event hosting without the mass gathering. J. Travel Med. **28**(2), taaa194 (2021). https://doi.org/10.1093/jtm/taaa194
19. El-Kenawy, E.-S.M., Khodadadi, N., Mirjalili, S., Makarovskikh, T., Abotaleb, M., Karim, F.K., Alkahtani, H.K., Abdelhamid, A.A., Eid, M.M., Horiuchi, T., Ibrahim, A., Khafaga, D.S.: Metaheuristic optimization for improving weed detection in wheat images captured by drones. Mathematics **10**(23), 4421 (2022). https://doi.org/10.3390/math10234421
20. El-kenawy, E.-S.M., Mirjalili, S., Khodadadi, N., Abdelhamid, A.A., Eid, M.M., El-Said, M., Ibrahim, A.: Feature selection in wind speed forecasting systems based on meta-heuristic optimization. PLoS ONE **18**(2), e0278491 (2023). https://doi.org/10.1371/journal.pone.0278491
21. Fabricante, I., Oesterheld, M., Paruelo, J.M.: Annual and seasonal variation of NDVI explained by current and previous precipitation across Northern Patagonia. J. Arid Environ. **73**(8), 745–753 (2009). https://doi.org/10.1016/j.jaridenv.2009.02.006
22. Foody, G.M.: Geographical weighting as a further refinement to regression modeling: an example focused on the NDVI–rainfall relationship. Remote Sens. Environ. **88**(3), 283–293 (2003). https://doi.org/10.1016/j.rse.2003.08.004
23. Franchito, S.H., Rao, V.B., Vasques, A.C., Santo, C.M.E., Conforte, J.C.: Validation of TRMM precipitation radar monthly rainfall estimates over Brazil. J. Geophys. Res. **114**(D2), D02105 (2009). https://doi.org/10.1029/2007JD009580
24. Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., Michaelsen, J.: The climate hazards infrared precipitation with stations: a new environmental record for monitoring extremes. Sci. Data **2**(1), 150066 (2015). https://doi.org/10.1038/sdata.2015.66
25. Gao, P., Du, W., Lei, Q., Li, J., Zhang, S., & Li, N.: NDVI forecasting model based on the combination of time series decomposition and CNN–LSTM. Water Resour. Manage. **37**(4), 1481–1497 (2023). https://doi.org/10.1007/s11269-022-03419-3
26. Gaughan, A.E., Waylen, P.R.: Spatial and temporal precipitation variability in the Okavango–Kwando–Zambezi catchment, southern Africa. J. Arid Environ. **82**, 19–30 (2012). https://doi.org/10.1016/j.jaridenv.2012.02.007
27. Georganos, S., Abdi, A.M., Tenenbaum, D.E., Kalogirou, S.: Examining the NDVI-rainfall relationship in the semi-arid Sahel using geographically weighted regression. J. Arid Environ. **146**, 64–74 (2017). https://doi.org/10.1016/j.jaridenv.2017.06.004

28. Gong, Z., Zhao, S., Gu, J.: Correlation analysis between vegetation coverage and climate drought conditions in North China during 2001–2013. J. Geog. Sci. **27**(2), 143–160 (2017). https://doi.org/10.1007/s11442-017-1369-5
29. Heumann, B.W., Seaquist, J.W., Eklundh, L., Jönsson, P.: AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982–2005. Remote Sens. Environ. **108**(4), 385–392 (2007). https://doi.org/10.1016/j.rse.2006.11.025
30. Hu, M., Li, Z., Wang, J., Jia, L., Liao, Y., Lai, S., Guo, Y., Zhao, D., Yang, W.: Determinants of the incidence of hand, foot and mouth disease in China using geographically weighted regression models. PLoS ONE **7**(6), e38978 (2012). https://doi.org/10.1371/journal.pone.0038978
31. Huang, J., Wang, H., Dai, Q., Han, D.: Analysis of NDVI data for crop identification and yield estimation. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. **7**(11), 4374–4384 (2014). https://doi.org/10.1109/JSTARS.2014.2334332
32. Huang, S., Ming, B., Huang, Q., Leng, G., Hou, B.: A case study on a combination NDVI forecasting model based on the entropy weight method. Water Resour. Manage. **31**(11), 3667–3681 (2017). https://doi.org/10.1007/s11269-017-1692-8
33. Huber, S., Fensholt, R., Rasmussen, K.: Water availability as the driver of vegetation dynamics in the African Sahel from 1982 to 2007. Global Planet. Change **76**(3), 186–195 (2011). https://doi.org/10.1016/j.gloplacha.2011.01.006
34. Kang, L., Di, L., Deng, M., Shao, Y., Yu, G., Shrestha, R.: Use of geographically weighted regression model for exploring spatial patterns and local factors behind NDVI-precipitation correlation. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. **7**(11), 4530–4538 (2014). https://doi.org/10.1109/JSTARS.2014.2361128
35. Kaveh, A., Talatahari, S., Khodadadi, N.: Stochastic paint optimizer: theory and application in civil engineering. Eng. Comput. **38**(3), 1921–1952 (2022). https://doi.org/10.1007/s00366-020-01179-5
36. Kawabata, A., Ichii, K., Yamaguchi, Y.: Global monitoring of interannual changes in vegetation activities using NDVI and its relationships to temperature and precipitation. Int. J. Remote Sens. **22**(7), 1377–1382 (2001). https://doi.org/10.1080/01431160119381
37. Khalilpourazari, S., Naderi, B., Khalilpourazary, S.: Multi-objective stochastic fractal search: a powerful algorithm for solving complex multi-objective optimization problems. Soft. Comput. **24**(4), 3037–3066 (2020). https://doi.org/10.1007/s00500-019-04080-6
38. Kogan, F.N.: Global drought watch from space. Bull. Am. Meteor. Soc. **78**(4), 621–636 (1997). https://doi.org/10.1175/1520-0477(1997)078%3c0621:GDWFS%3e2.0.CO;2
39. Liu, B., Chen, J., Chen, J., Zhang, W.: Land cover change detection using multiple shape parameters of spectral and NDVI curves. Remote Sens. **10**(8), 1251 (2018). https://doi.org/10.3390/rs10081251
40. Lu, J., Carbone, G.J., Gao, P.: Mapping the agricultural drought based on the long-term AVHRR NDVI and North American Regional Reanalysis (NARR) in the United States, 1981–2013. Appl. Geogr. **104**, 10–20 (2019). https://doi.org/10.1016/j.apgeog.2019.01.005
41. Lunetta, R.S., Knight, J.F., Ediriwickrema, J., Lyon, J.G., Worthy, L.D.: Land-cover change detection using multi-temporal MODIS NDVI data. Remote Sens. Environ. **105**(2), 142–154 (2006). https://doi.org/10.1016/j.rse.2006.06.018
42. Mallick, J., Singh, R.K., Khan, R.A., Singh, C.K., Kahla, N.B., Warrag, E.I., Islam, S., Rahman, A.: Examining the rainfall–topography relationship using non-stationary modelling technique in semi-arid Aseer region, Saudi Arabia. Arab. J. Geosci. **11**(9), 215 (2018). https://doi.org/10.1007/s12517-018-3580-9
43. Mallick, J., Talukdar, S., Alsubih, M., Salam, R., Ahmed, M., Kahla, N.B., Shamimuzzaman, Md.: Analysing the trend of rainfall in Asir region of Saudi Arabia using the family of Mann-Kendall tests, innovative trend analysis, and detrended fluctuation analysis. Theoret. Appl. Climatol. **143**(1), 823–841 (2021). https://doi.org/10.1007/s00704-020-03448-1
44. Martiny, N., Camberlin, P., Richard, Y., Philippon, N.: Compared regimes of NDVI and rainfall in semi-arid regions of Africa. Int. J. Remote Sens. **27**(23), 5201–5223 (2006). https://doi.org/10.1080/01431160600567787
45. McMillen, D.P.: Geographically Weighted regression: the analysis of spatially varying relationships. Am. J. Agr. Econ. **86**(2), 554–556 (2004). https://doi.org/10.1111/j.0002-9092.2004.600_2.x
46. Murshed, M., Rahman, Md.A., Alam, M.S., Ahmad, P., Dagar, V.: The nexus between environmental regulations, economic growth, and environmental sustainability: linking environmental patents to ecological footprint reduction in South Asia. Environ. Sci. Pollut. Res. **28**(36), 49967–49988 (2021). https://doi.org/10.1007/s11356-021-13381-z
47. Mutti, P. R., Lúcio, P. S., Dubreuil, V., Bezerra, B. G.: NDVI time series stochastic models for the forecast of vegetation dynamics over desertification hotspots. Int. J. Remote Sens. **41**(7), 2759–2788 (2020). https://doi.org/10.1080/01431161.2019.1697008
48. Nanzad, L., Zhang, J., Tuvdendorj, B., Nabil, M., Zhang, S., Bai, Y.: NDVI anomaly for drought monitoring and its correlation with climate factors over Mongolia from 2000 to 2016. J. Arid Environ. **164**, 69–77 (2019). https://doi.org/10.1016/j.jaridenv.2019.01.019
49. Nezlin, N.P., Kostianoy, A.G., Li, B.-L.: Inter-annual variability and interaction of remote-sensed vegetation index and atmospheric precipitation in the Aral Sea region. J. Arid Environ. **62**(4), 677–700 (2005). https://doi.org/10.1016/j.jaridenv.2005.01.015
50. Nguyen, T.T., Nguyen, T.T., Duong, M.Q., Doan, A.T.: Optimal operation of transmission power networks by using improved stochastic fractal search algorithm. Neural Comput. Appl. **32**(13), 9129–9164 (2020). https://doi.org/10.1007/s00521-019-04425-0
51. Patel, J. H., Oza, M. P.: Deriving crop calendar using NDVI time-series. Int. Arch. Photogrammetry Remote Sens. Spat. Inf. Sci. **40**, 869–873 (2015)
52. PereiraColtri, P., Zullo, J., do Ribeiro Valle Gonçalves, R., Romani, L.A.S., Pinto, H.S.: Coffee Crop's biomass and carbon stock estimation with usage of high resolution satellites images. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. **6**(3), 1786–1795 (2013). https://doi.org/10.1109/JSTARS.2013.2262767
53. Potter, C.S., Brooks, V.: Global analysis of empirical relations between annual climate and seasonality of NDVI. Int. J. Remote Sens. **19**(15), 2921–2948 (1998). https://doi.org/10.1080/014311698214352
54. Quarmby, N.A., Milnes, M., Hindle, T.L., Silleos, N.: The use of multi-temporal NDVI measurements from AVHRR data for crop yield estimation and prediction. Int. J. Remote Sens. **14**(2), 199–210 (1993). https://doi.org/10.1080/01431169308904332
55. Rhif, M., Abbes, A. B., Martínez, B., Farah, I. R.: Deep learning models performance for NDVI time series prediction: a case study on north west Tunisia. In: 2020 Mediterranean and Middle-East Geoscience and Remote Sensing Symposium (M2GARSS). IEEE, pp. 9–12. https://doi.org/10.1109/M2GARSS47143.2020.9105149 (2020)
56. Richard, Y., Poccard, I.: A statistical study of NDVI sensitivity to seasonal and interannual rainfall variations in Southern Africa. Int. J. Remote Sens. **19**(15), 2907–2920 (1998). https://doi.org/10.1080/014311698214343
57. Rizk, F.H., Arkhstan, S., Zaki, A.M., Kandel, M.A., Towfek, S.K.: Integrated CNN and waterwheel plant algorithm for enhanced global traffic detection. J. Artif. Intell. Metaheuristics **6**(2), 36–45 (2023). https://doi.org/10.54216/JAIM.060204

58. Rizk, F.H., Mohamed, M.E., Sameh, B., Zaki, A.M., Eid, M.M., El-kenawy, E.-S. M. Predictive modeling of portuguese student performance: comparative machine learning analysis. In: 2024 International Telecommunications Conference (ITC-Egypt), pp. 26–31 (2024). https://doi.org/10.1109/ITC-Egypt61547.2024.10620557
59. Sam Khafaga, D., Ali Alhussan, A., El-kenawy, M.E.-S., Ibrahim, A., AbdElkhalik, H.S., El-Mashad, Y.S., Abdelhamid, A.A.: Improved prediction of metamaterial antenna bandwidth using adaptive optimization of LSTM. Comput. Mater. Contin. **73**(1), 865–881 (2022). https://doi.org/10.32604/cmc.2022.028550
60. Sami Khafaga, D., El-kenawy, M.E.-S., Khalid Karim, F., Alshetewi, S., Ibrahim, A., Abdelhamid, A.A.: Optimized weighted ensemble using dipper throated optimization algorithm in metamaterial antenna. Comput. Mater. Contin. **73**(3), 5771–5788 (2022). https://doi.org/10.32604/cmc.2022.032229
61. Saxena, G., Purchase, D., Mulla, S.I., Saratale, G.D., Bharagava, R.N.: Phytoremediation of heavy metal-contaminated sites: eco-environmental concerns, field studies, sustainability issues, and future prospects. In: de Voogt, P. (ed.) Reviews of Environmental Contamination and Toxicology, vol. 249, pp. 71–131. Springer International Publishing (2020). https://doi.org/10.1007/398_2019_24
62. Shabbir, M.S., Wisdom, O.: The relationship between corporate social responsibility, environmental investments and financial performance: Evidence from manufacturing companies. Environ. Sci. Pollut. Res. **27**(32), 39946–39957 (2020). https://doi.org/10.1007/s11356-020-10217-0
63. Shahzad, U.: Environmental taxes, energy consumption, and environmental quality: theoretical survey with policy implications. Environ. Sci. Pollut. Res. **27**(20), 24848–24862 (2020). https://doi.org/10.1007/s11356-020-08349-4
64. Shammi, S. A., Meng, Q.: Modeling crop yield using NDVI-derived VGM metrics across different climatic regions in the USA. Int. J. Biometeorol **67**(6), 1051–1062 (2023). https://doi.org/10.1007/s00484-023-02478-4
65. Sherif, K., Azmy, M., Alakkari, K., Abotaleb, M., El-Kenawy, E. S. M.: Deep Learning in IoT: An LSTM Approach for NDVI Forecasting. In: 2023 3rd International Conference on Electronic Engineering (ICEEM), pp. 1–7. IEEE (2023). https://doi.org/10.1109/ICEEM58740.2023.10319616
66. Tadesse, T., Brown, J.F., Hayes, M.J.: A new approach for predicting drought-related vegetation stress: integrating satellite, climate, and biophysical data over the U.S. central plains. ISPRS J. Photogramm. Remote Sens. **59**(4), 244–253 (2005). https://doi.org/10.1016/j.isprsjprs.2005.02.003
67. Taibah, H., Arlikatti, S., Andrew, S.A., Maghelal, P., DelGrosso, B.: Health information, attitudes and actions at religious venues: evidence from hajj pilgrims. Int. J. Dis. Risk Reduct. **51**, 101886 (2020). https://doi.org/10.1016/j.ijdrr.2020.101886
68. Toté, C., Patricio, D., Boogaard, H., Van der Wijngaart, R., Tarnavsky, E., Funk, C.: Evaluation of satellite rainfall estimates for drought and flood monitoring in mozambique. Remote Sens. **7**(2), 1758–1776 (2015). https://doi.org/10.3390/rs70201758
69. Tran, T.T., Truong, K.H., Vo, D.N.: Stochastic fractal search algorithm for reconfiguration of distribution networks with distributed generations. Ain Shams Eng. J. **11**(2), 389–407 (2020). https://doi.org/10.1016/j.asej.2019.08.015
70. Trejo, F.J.P., Barbosa, H.Á., Penaloza-Murillo, M.A., Moreno, M.A., Farias, A.: Intercomparison of improved satellite rainfall estimation with CHIRPS gridded product and rain gauge data over Venezuela. Atmosfera **29**(4), 323–342 (2016). https://doi.org/10.20937/ATM.2016.29.04.04
71. Tucker, C.J., Slayback, D.A., Pinzon, J.E., Los, S.O., Myneni, R.B., Taylor, M.G.: Higher northern latitude normalized difference vegetation index and growing season trends from 1982 to 1999. Int. J. Biometeorol. **45**(4), 184–190 (2001). https://doi.org/10.1007/s00484-001-0109-8
72. Wang, J., Price, K.P., Rich, P.M.: Spatial patterns of NDVI in response to precipitation and temperature in the central Great Plains. Int. J. Remote Sens. **22**(18), 3827–3844 (2001). https://doi.org/10.1080/01431160010007033
73. Wang, Q., Ni, J., Tenhunen, J.: Application of a geographically-weighted regression analysis to estimate net primary production of Chinese forest ecosystems: NPP in Chinese Forests. Glob. Ecol. Biogeogr. **14**(4), 379–393 (2005). https://doi.org/10.1111/j.1466-822X.2005.00153.x
74. Weiss, E., Marsh, S.E., Pfirman, E.S.: Application of NOAA-AVHRR NDVI time-series data to assess changes in Saudi Arabia's rangelands. Int. J. Remote Sens. **22**(6), 1005–1027 (2001). https://doi.org/10.1080/014311601300074540
75. Xiao, J., Moody, A.: Trends in vegetation activity and their climatic correlates: China 1982 to 1998. Int. J. Remote Sens. **25**(24), 5669–5689 (2004). https://doi.org/10.1080/01431160410001735094
76. Yang, L., Wylie, B.K., Tieszen, L.L., Reed, B.C.: An analysis of relationships among climate forcing and time-integrated NDVI of grasslands over the U.S. northern and central great plains. Remote Sens. Environ. **65**(1), 25–37 (1998). https://doi.org/10.1016/S0034-4257(98)00012-1
77. Yezli, S., Mushi, A., Almuzaini, Y., Balkhi, B., Yassin, Y., Khan, A.: Prevalence of diabetes and hypertension among hajj pilgrims: a systematic review. Int J Environ Res Pub Health (2021). https://doi.org/10.3390/ijerph18031155
78. Yong-li, W.U., Zhi-cai, L.I., Yun-feng, W., Qing, L., Guo-zhen, T.: Responses of vegetation index (NDVI) in typical ecological areas of Shanxi Province to climate change. Chin. J. Ecol. **28**(05), 925 (2009)
79. Zaki, A.M., Abdelhamid, A.A., Ibrahim, A., Eid, M.M., El-Kenawy, E.-S.M.: Enhancing K-nearest neighbors algorithm in wireless sensor networks through stochastic fractal search and particle swarm optimization. J. Cybersecur. Inf. Manag. **13**(1), 76–84 (2023). https://doi.org/10.54216/JCIM.130108
80. Zhao, Z., Gao, J., Wang, Y., Liu, J., Li, S.: Exploring spatially variable relationships between NDVI and climatic factors in a transition zone using geographically weighted regression. Theoret. Appl. Climatol. **120**(3), 507–519 (2015). https://doi.org/10.1007/s00704-014-1188-x
81. Zhou, X., Tang, X., Zhang, R.: Impact of green finance on economic development and environmental quality: a study based on provincial panel data from China. Environ. Sci. Pollut. Res. **27**(16), 19915–19932 (2020). https://doi.org/10.1007/s11356-020-08383-2s

Authors and Affiliations

Nima Khodadadi[1] · S. K. Towfek[2,6,7,10,14] · Ahmed Mohamed Zaki[2] · Amal H. Alharbi[3] · Ehsan Khodadadi[4] · Doaa Sami Khafaga[2] · Laith Abualigah[5,6,12,13] · Abdelhameed Ibrahim[7] · Abdelaziz A. Abdelhamid[8,11] · Marwa M. Eid[9]

✉ Nima Khodadadi
Nima.khodadadi@miami.edu

✉ Abdelaziz A. Abdelhamid
Abdelaziz@su.edu.sa

S. K. Towfek
sktowfek@jcsis.org

Ahmed Mohamed Zaki
Azaki@jcsis.org

Amal H. Alharbi
ahalharbi@pnu.edu.sa

Ehsan Khodadadi
ehsank@uark.edu

Doaa Sami Khafaga
dskhafga@pnu.edu.sa

Laith Abualigah
Aligah.2020@gmail.com

Abdelhameed Ibrahim
abdelhameed.fawzy@polytechnic.bh

Marwa M. Eid
mmm@ieee.org

1 Department of Civil and Architectural Engineering, University of Miami, Coral Gables, FL, USA

2 Computer Science and Intelligent Systems Research Center, Blacksburg, VA 24060, USA

3 Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, PO Box 84428, 11671 Riyadh, Saudi Arabia

4 Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701, USA

5 MEU Research Unit, Middle East University, Amman 11831, Jordan

6 Applied Science Research Center, Applied Science Private University, Amman 11931, Jordan

7 School of ICT, Faculty of Engineering, Design and Information and Communications Technology (EDICT), Bahrain Polytechnic, PO Box 33349, Isa Town, Bahrain

8 Department of Computer Science, College of Computing and Information Technology, Shaqra University, 11961 Shaqra, Saudi Arabia

9 Faculty of Artificial Intelligence, Delta University for Science and Technology, Mansoura 11152, Egypt

10 Jadara University Research Center, Jadara University, Irbid, Jordan

11 Department of Computer Science, Faculty of Computer and Information Sciences, Ain Shams University, Cairo 11566, Egypt

12 Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura 140401, Punjab, India

13 Artificial Intelligence and Sensing Technologies (AIST) Research Center, University of Tabuk, Tabuk 71491, Saudi Arabia

14 Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura 35111, Egypt