



OPEN

Design of optimal Elman Recurrent Neural Network based prediction approach for biofuel production

N. Paramesh Kumar¹, S. Vijayabaskar¹, L. Murali² & Krishnaraj Ramaswamy^{3,4}✉

Renewable sources like biofuels have gained significant attention to meet the rising demands of energy supply. Biofuels find useful in several domains of energy generation such as electricity, power, or transportation. Due to the environmental benefits of biofuel, it has gained significant attention in the automotive fuel market. Since the handiness of biofuels become essential, effective models are required to handle and predict the biofuel production in realtime. Deep learning techniques have become a significant technique to model and optimize bioprocesses. In this view, this study designs a new optimal Elman Recurrent Neural Network (OERNN) based prediction model for biofuel prediction, called OERNN-BPP. The OERNN-BPP technique pre-processes the raw data by the use of empirical mode decomposition and fine to coarse reconstruction model. In addition, ERNN model is applied to predict the productivity of biofuel. In order to improve the predictive performance of the ERNN model, a hyperparameter optimization process takes place using political optimizer (PO). The PO is used to optimally select the hyper parameters of the ERNN such as learning rate, batch size, momentum, and weight decay. On the benchmark dataset, a sizable number of simulations are run, and the outcomes are examined from several angles. The simulation results demonstrated the suggested model's advantage over more current methods for estimating the output of biofuels.

The global energy system is based largely on fossil fuels¹. The significance of energy system also its part in politics and economics isn't hidden for everyone. This problem is significant for innovative industrialized countries that are the main energy consumer, as well as significant for oil-rich countries². Since countries must comprehend the fact that fossil fuels resource is constrained resource. Besides the nature of this fuel have polluting substance, the problem of their ended up has provoked the increased attention. Due to the fact that reducing pollution, environmental damage, and non-renewable energy resources the world is turning toward sustainable energy resources³. Still, fossil fuel remains the main energy resource around the world. Highly based fossil fuel has created an energy crisis. Utilizing fossil fuels for economic activity lead to greenhouse gas (GHG) emissions from most region of the world. Because of the environmental concern and the fluctuation and rise in the fossil fuels resource, global interest has moved toward biodiesels, renewable and clean alternatives for fossil fuel⁴. Biofuel is employed indistinct areas for producing energy such as transportation, power, and electricity production.

The importance of energy systems and their role in economics and politics is not hidden for anyone. This issue is not only important for the advanced industrialized countries, which are major energy consumers but is also essential for oil-rich countries. In addition to the nature of these fuels, which contains polluting substances, the issue of their ending up has aggravated the growing concern. Biofuels can be used in different fields for energy production like electricity production, power production, or for transportation⁵. Various scenarios have been written about the estimated biofuels from different sources in the future energy system. The availability of biofuels for the electricity market, heating, and liquid fuels is critical. Accordingly, the need for handling, modeling, decision making, and forecasting for biofuels can be of utmost importance.

Statistical and Mathematical model provides basic data to understand, analyze, and predict biological process, and are essential for optimizing significant parameter for improving system efficiency⁶. Optimization and Modeling of biofuels manufacturing process would participate to understand better of the expenditure procedure for obtaining an optimal efficacy. The primary objective of modelling is to enhance the operation involving in their productions for achieving efficacy development. Artificial intelligence tools had emerged as possible methods to

¹Department of Electrical and Electronics Engineering, P.A. College of Engineering and Technology, Pollachi, India. ²Department of Electronics and Communication Engineering, P.A. College of Engineering and Technology, Pollachi, India. ³Centre for Excellence-Indigenous Knowledge, Innovative Technology Transfer and Entrepreneurship, Dambi Dollo University, Dambi Dollo, Ethiopia. ⁴Department of Mechanical Engineering, Dambi Dollo University, Dambi Dollo, Ethiopia. ✉email: dr.krishnarajdirectorcei@dadu.edu.et

optimize and model bioprocess. Over the previous years, artificial neural network (ANN) was used in nonlinear, multidimensional development and research of bioprocess. They have proved their effectiveness in emerging bioprocesses model lacking of previous data on them etabolic and kinetics flow occurs in cell and cells surrounding⁷.

Moreover, ANN is dependent fully on data, with no previous experience on the event regulate the procedure⁸. The appeal of ANN as modelling tools derive from their exclusive function of processing data i.e., allocated high parallelism, primarily—linearity, and noise and error acceptance—and their ability to generalize and learn. ANN has received more interest from substantial soft computing tool which is constrained only for data analysis and processing, however, could also be employed for solving problem in nonlinear and multifaceted procedures⁹.

In recent times, deep learning and machine learning methods have been widespread in handling, modeling and optimizing the biodiesel consumption, production and its environmental impact by taking into account the effects of parameter on biofuel yields since production of a preferred products need an efficient usage of investigational models. This method provides a self-governing modeling method to the nature of procedure or its arithmetical model as well as capable of modelling the procedure using higher performances¹⁰.

In this research, a new optimal Elman Recurrent Neural Network (OERNN) based prediction model for biofuel prediction is proposed which provides better result when compared with the other existing approaches. The OERNN-BPP technique involves empirical mode decomposition (EMD) based pre-processing and fine to coarse (FTC) based reconstruction model. Besides, ERNN model is employed for the prediction of biofuel productivity. For enhancing the predictive performance of the ERNN model, a hyper parameter optimization process takes place using political optimizer (PO). A comprehensive experimental analysis is carried out on benchmark dataset and the results are examined interms of diverse aspects.

Related works. This section provides numerous research studies that have been focussed based on the production of biofuel. To gain a better understanding of the literal works and the relevant research areas are summarized as follows. The production of biofuel based on spatial distribution was implemented by Elmore et al.¹¹ where this approach utilized the Moderate Resolution Imaging Spectrometer (MODIS). Here, the residue from the rice was employed to produce the biofuel. The accuracy and the flexibility rate were very high in this approach; on the other hand, there occurs a complexity in designing the spatial model.

Chanthawong et al.¹² proposed two different types of approaches namely two stages least square and three stages least square for biofuel production in the Thailand market. This dual-stage approach is developed with minimum cost with very less dynamic model constitution. The accurate biogas prediction was developed by De et al.¹³ that utilized the neural network model namely k-nearest neighbors (KNN) for the effective production of biofuel. Here the forecasting accuracy is very high with improved facility performances. But, there occur a few complexity issues during implementation. Then Dehghani et al.¹⁴ demonstrated a future forecasting model based on the production of biofuel and this approach utilized seven biofuel technologies namely gas turbine, Combined Heat and Power (CHP) turbine, bio-pyrolysis, cellulosic bioethanol, grain bioethanol, torrefaction and biodiesel. Moreover, the execution performance and the accuracy were very high; but the research and development of this approach are not much efficient.

Radivojević et al.¹² introduced the Automated Recommendation Tool (ART), a tool which leverage ML and probabilistic approaches for guiding synthetic biology from the systematic fashion, with no requirement to complete mechanistic kind of biological scheme. The following engineering cycle's group of recommended strains, along with a probabilistic prediction of its production level, are provided by ART using sampling-based optimization.

Elveny et al.¹³ presented a novel Machine Learning (ML) technique dependent upon Extreme Learning Machine (ELM) for modelling this essential value. The real database involving 483 actual datasets has been related to the output forecasted with ELM technique. In Cui et al.¹⁴, distinct ML techniques are estimated to be the primary time for establishing the forecast technique amongst biodiesel composition and cold filter plugging point (CFPP). The decision trees (DT) based techniques are optimum efficient in forecasting CFPP of biodiesel.

Kumar et al.¹⁵ aimed in evolving a new Adaptive Integrated Optimization Network (AION) for attaining optimum biofuel production with maximal accuracy as well as minimal error value rates. Also, the presented AION manner includes 4 important stages as Pre-processing of data, Re-construction of components, Prediction of individuals, and Ensemble predicting. Javed et al.¹⁶ developed a grey predicting technique with optimized the model frame work (data accumulation function and background value generation). The presented predicting technique, Even form of Grey Forecasting model (EGM) $(1,1,\alpha,\theta)$ is a generalization procedure of the even procedure of grey predicting technique and their comparative efficiency turned out that commonly higher than that of the original technique.

Beeravalli et al.¹⁷ search a new manner for classifying feed stock's utilizing secondary works data sources. Also, the maximum reliability of techniques utilized, the study analyzed investigating over 20 parameters of 106 feed stocks. The study established a rating scheme to Multi-Criteria Decision Analysis (MCDA) containing weighting to all parameters dependent upon expert opinion or statistical techniques namely Principal Component Analysis (PCA). The ranking method output afterward is fed as to Multivariate Regression (MVR) and Multilayer perceptron (MLP), for ranking feed stocks for producing the maximum quality maintainable biofuels to a specific place.

Geng et al.¹⁸ resolved the influence of random fluctuation data and weak anti-interference capability in the Markov chain model by proposing a dynamic fuzzy grey-Markov prediction model for biofuel production forecasting, in order to improve the prediction performance of the conventional prediction methods based upon past production levels in conjunction with the factors of economy, governmental policies, and technological developments. Their empirical results demonstrated the superiority of the proposed fuzzy grey-Markov model relative to the benchmark prediction models. However, the biofuel production system is a complex system, which is affected easily by various factors such as the economy, governmental policies, resources, technological

developments and social issues. Thus, the above methods can only provide good prediction results under linear assumptions, being unable to capture the hidden nonlinear features of the biofuel production series. Numerous experiments have demonstrated that the predictive performance may be very poor if one uses these traditional statistical and econometric methods (Mejdoub and Ghorbel¹⁹; Song and Yu²⁰; Weigend²¹). Therefore, the traditional methods are not suitable for predicting biofuel production (Geng et al.¹⁸).

The proposed model. In this study, a new OERNN-BPP technique has been presented to predict the productivity of biofuels. The OERNN-BPP technique follows four major processes namely EMD based processing, FTC based reconstruction, ERNN based prediction, and PO based hyper parameter optimization. The design of PO based hyper parameter tuning process assists in optimally adjusting the learning rate, batch size, momentum, and weight decay.

Figure 1 demonstrates the overall block diagram of OERNN-BPP model. The detailed working of these processes is discussed in the succeeding sections.

EMD based data pre-processing. Initially, a decomposition method called EMD is used in separating raw complex data into relatively simple/uncomplicated data thus decreasing the complication problems. Next, the EMD method is related to other decomposition methods like the Fourier decomposition and wavelet decomposition approaches²². Now, the EMD method, i.e., a type of intuitive, self-adapting, empirical and direct method and is suited well for nonlinear and non-stationary data. Generally, the EMD method employed in decomposing the raw time sequence data to some periodic mode functions I_{MF} contain independent data. Accurately, the IPMF fulfills 2 distinct criteria's which are given below.

$$\text{Criteria : } \begin{cases} N_E \text{ and } N_{ZC} \text{ are equal ; } \forall \text{ Each whole funct (i.e.0, 2, 4 \dots) \\ A_T (T = 1, 2, 3, \dots); \quad \forall \text{ Symmetric function} \end{cases} \quad (1)$$

From Eq.(1), N_E & N_{ZC} denotes the overall amount of extrema and zero crossing correspondingly. The novel data depends on time sequence is given by A_T ; whereas its expressions are signified as follows.

$$A_T = \sum_{I=1}^M B_I + R_M, \quad (2)$$

From Eq. (2), the overall quantity of I_{MF} and there s i due at a time 'T' is denoted by M & R_M , correspondingly.

FTC based reconstruction. The FTC model reconstructs IMF as 2t-testing parts namely minimum and maximum frequency elements. In addition, the t-testing frequency element comprises distinct features involving the information about the centralized traits. So, a simpler structure of FTC model is applied for improving the accuracy and reducing the computational complexity. The FTC model encompasses a 2-stage process. Firstly, the preprocessed IMF attained from the earlier level gets inspected by the use of t-testing. The next level includes the choice of IMF similarity including unrelated divergences at a certain degree of confidence are reconstructed different elements. The classification of maximal and minimal frequency elements is then done using the IMF with closer IMF over t-testing if the IMF resemblance with irrelevant divergence at a certain level of confidence cannot be achieved.

Design of ERNN based predictive model. At this stage, the ERNN model receives the input and predicts the actual production of biofuels. The ERNN has been simple RNN is established by Elman in 1990²³. Already

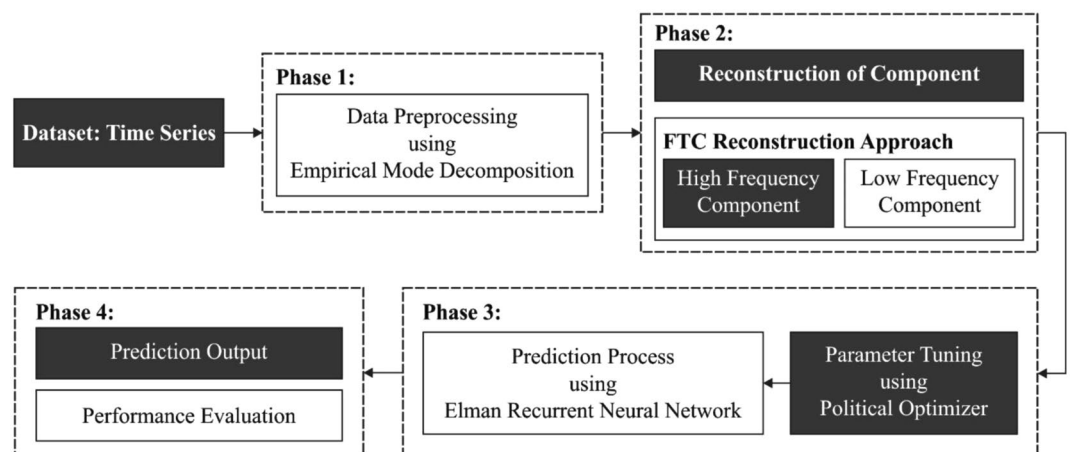


Figure 1. Overall block diagram of OERNN-BPP model.

known, are current network is a few benefits like consuming time series and non-linear forecast abilities, faster convergence, and further accurate mapping capability. They combine Elman neuralnetwork (ENN) with distinct regions to its purpose. During this network, the outcomes of the hidden layers (HL) were permitted to feedback on itself with butter layer is named as recurrent layers (RL). Figure 2 illustrates the architecture of ERNN model.

This feedback enables ERNN to learn, recognise, and produce spatial patterns as well as temporal designs. One RL neuron with a constant weight of one connects all hidden neurons. As a result, the RL almost creates a copy of the HL's previous instantaneous state. Accordingly, the number of recurrent neurons is comparable to the number of hidden neurons. Each layer has several neurons that propagate information from one layer to the next by computing a nonlinear function of the inputs' weighted sum.

The multi-input ERNN technique has been demonstrated, in which the amount of neurons from the input layers are m and during the HL is n and one output unit. Assume that $x_{it}(i = 1, 2, \dots, m)$ represents the group of input vector of neurons at time t , y_{t+1} implies the outcome of networks at time $t + 1$, $z_{jt}(j = 1, 2, \dots, n)$ refers the outcome of HL neuron at time t , and $(j = 1, 2, \dots, n)$ defines the RL neuron. w_{ij} signifies the weight that links the node i from the input layer neuron to node j from the HL. c_j are the weights which link the node j during the HL neurons to node under the RL and output correspondingly. The HL stage is as follows: the inputs of every neuron from the HLs are provided as:

$$x_{jt}(k) = \sum_{i=1}^n w_{ij}x(k-1) + \sum_{j=1}^m c_j u_{jt}(k) \tag{3}$$

$$u_{jt}(k) = z_{jt}(k-1), \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, m.$$

The outcomes of hidden neurons are provided as:

$$z_{jt}(k) = f_H(\text{net}_{jt}(k)) = \left(\sum_{i=1}^n w_{ij}(k) + \sum_{j=1}^m c_j u_j(k) \right) \tag{4}$$

where the sigmoid function from HL has been elected as activation function: $f_H(x) = 1/(1 + e^{-x})$. The outcome of HL is provided as²⁴:

$$y_{t+1}(k) = f_T \left(\sum_{j=1}^n v_j(k) \right) \tag{5}$$

where (x) refers the identity map as activation functions.

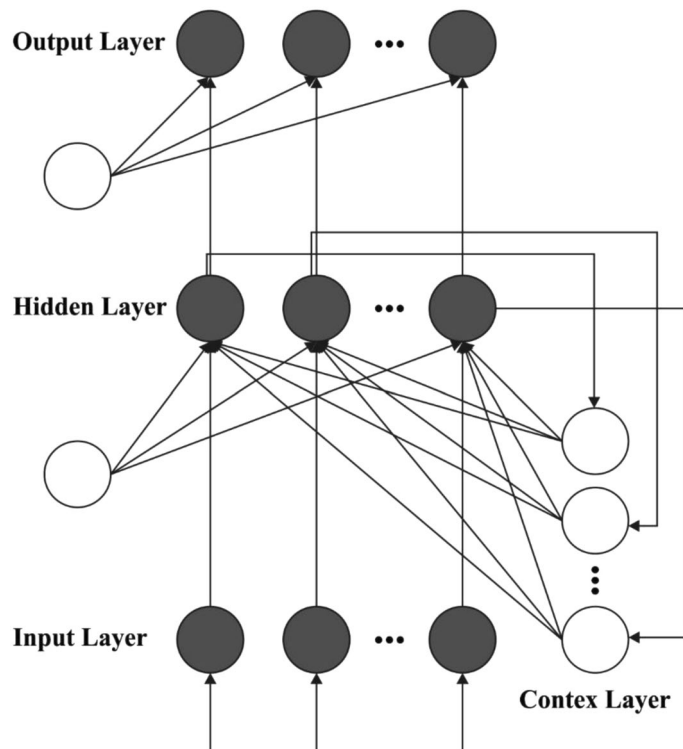


Figure 2. Structure of ERNN.

Design of PO based hyper parameter tuning process. For enhancing the predictive outcome of the ERNN model, the hyperparameter tuning process is carried out using PO. PO is a recently developed Meta heuristic approach that depends on human behaviour and is stimulated from the multiphase PO. But, it has to be mentioned that the presented method isn't primary of these kinds. In PO, the concepts of politics are mapped from a distinct point of view and different from the current politics stimulated algorithm, and it is because of the 4 reasons. Initially, PO tries to model every major step in politics like party development, constituency distribution or party ticket, party transferring, election campaign, parliamentary affairs, and interparty election afterward governments creation.

Next, PO presents a new location upgrading approach named RPPUS. This later represents the learning performance of the politician from the former election. Then, all the individuals' solutions assume a binary task: an election candidate and party member. With these concepts, every solution could be upgraded based on the 2 optimal solutions: constituency winner and party leader. Lastly, in order to enhance the result, intermediary solutions need to communicate and cooperate through a phase called parliamentary affairs. In PO, all the party members are regarded as a candidate solution in which its good will has deliberated the location in the search spaces. Furthermore, the calculation functions are processed in the course of the election stage whereas the numbers of votes attained by all the party members represent the fitness of candidate solutions. PO model is generated using the 5 major stages in the following: constituency allocation, party development, party transferring, election campaign, parliamentary affair, and interparty election²⁵. It has to be noted that the initial stage (constituency allocation and party formation) is performed once for initializing and affects distinct parameters.

Party formation and constituency allocation. At first, the population P is divided into N party, whereas all the parties P_i include N member (possible solutions). Furthermore, all j th members are referred to as P^j and denoted as a d dimension vectors, in which the values d are the amount of input parameters of the processed problems and $P^j_{i,k}$ denote k th dimensions of P^j as above-mentioned, all the members have deliberated as an election candidate as well its role as a member party. Therefore, N constituency is made and have j th members of all contesting parties. Moreover, the leader of the i th party afterward calculating the fitness of each member is stated as P_i^* and the group of each party leader is given as P^* . In contrast, afterward the election, C^* regroup the winner from each constituency called the parliamentarian, whereas C_j^* denote the winners of j th constituencies.

Election campaign. In this stage, party member is trying to improve their chance of being selected by altering their position based on the 3 factors. Firstly, they attempt for learning from prior knowledge with a new location upgrading approach named RPPUS as expressed in Eqs. (6) & (7). Next, all-party members are trying to upgrade their present location based on the party leader. Lastly, candidate position is updated regarding the constituency winners:

$$P^j_{i,k}(t+1) = \begin{cases} m^* + r(m^* - P^j_{i,k}(t)), & \text{if } P^j_{i,k}(t-1) \leq P^j_{i,k}(t) \leq m^* \text{ or } P^j_{i,k}(t-1) \geq P^j_{i,k}(t) \geq m^* \\ m^* + (2r-1) \left| m^* - P^j_{i,k}(t) \right|, & \text{if } P^j_{i,k}(t-1) \leq m^* \leq P^j_{i,k}(t) \text{ or } P^j_{i,k}(t-1) \geq m^* \geq P^j_{i,k}(t) \\ m^* + (2r-1) \left| m^* - P^j_{i,k}(t) \right|, & \text{if } m^* \leq P^j_{i,k}(t-1) \leq P^j_{i,k}(t) \text{ or } m^* \geq P^j_{i,k}(t-1) \geq P^j_{i,k}(t) \end{cases} \quad (6)$$

$$P^j_{i,k}(t+1) = \begin{cases} m^* + (2r-1) \left| m^* - P^j_{i,k}(t) \right|, & \text{if } P^j_{i,k}(t-1) \leq P^j_{i,k}(t) \leq m^* \text{ or } P^j_{i,k}(t-1) \geq P^j_{i,k}(t) \geq m^* \\ P^j_{i,k}(t-1) + r \left(P^j_{i,k}(t) - P^j_{i,k}(t-1) \right), & \text{if } P^j_{i,k}(t-1) \leq m^* \leq P^j_{i,k}(t) \text{ or } P^j_{i,k}(t-1) \geq m^* \geq P^j_{i,k}(t) \\ m^* + (2r-1) \left| m^* - P^j_{i,k}(t-1) \right|, & \text{if } m^* \leq P^j_{i,k}(t-1) \leq P^j_{i,k}(t) \text{ or } m^* \geq P^j_{i,k}(t-1) \geq P^j_{i,k}(t) \end{cases} \quad (7)$$

To balance among exploitation and exploration, a stage named party switching is initiated afterward the election campaign stage. With adaptive parameters, called party switching rate, all party members P^j could be elected and switch to few arbitrarily selected party P_r . Henceforth, it is exchanged with the minimum fit party member P_r .

Election. This stage's aim is to calculate the fitness of each candidate contest in constituency. Afterward, the party leader and constituency winner are upgraded by:

$$\begin{aligned} q &= \operatorname{argmin} \left(P^j_i \right), 1 \leq i \leq N, \\ C_j^* &= P^j_q, \\ P_j^* &= P^j_q \end{aligned} \quad (8)$$

Parliamentary affairs. Afterward defining the party leader and constituency winner (parliamentarian), all the parliamentarians aiming to enhance their performances by mimicking the cooperation and interaction of the winning candidate to manage the governments in the post-election stage. All the parliamentarians C_j^* update

its location regarding arbitrarily selected parliamentarians C_r^* . It has to be pointed out that the movements are used only when the performances of C_r^* are improved.

Figure 3 depicts the flowchart of PO. Initially, the input parameters are initialized. Then the fitness function is calculated for all the individual. For these individuals the party leader and constituency winner is defined, where the position of party leader is updated through election campaign. Afterward through Parliamentary Affairs the party leader is defined and constituency winner, all the parliamentarians aiming to enhance their performances by mimicking the cooperation. In Election stage's it aim is to calculate the fitness of each candidate contest in constituency and update the previous position and fitness which results in performance improvement.

Experimental validation. This section analyses the OERNN-BPP technique's findings using annual bio-fuel production data that was gathered from China between January 2015 and June 2020 (https://apps.fas.usda.gov/newgainapi/api/Report/DownloadReportByFileName?fileName=Biofuels%20Annual_Beijing_China%20-%20People%27s%20Republic%20of_CH2022-0089.pdf). The samples in the dataset are split into training data (80%) and testing data (20%), respectively. The OERNN-BPP technique's outcomes are analysed in a variety of dimensions.

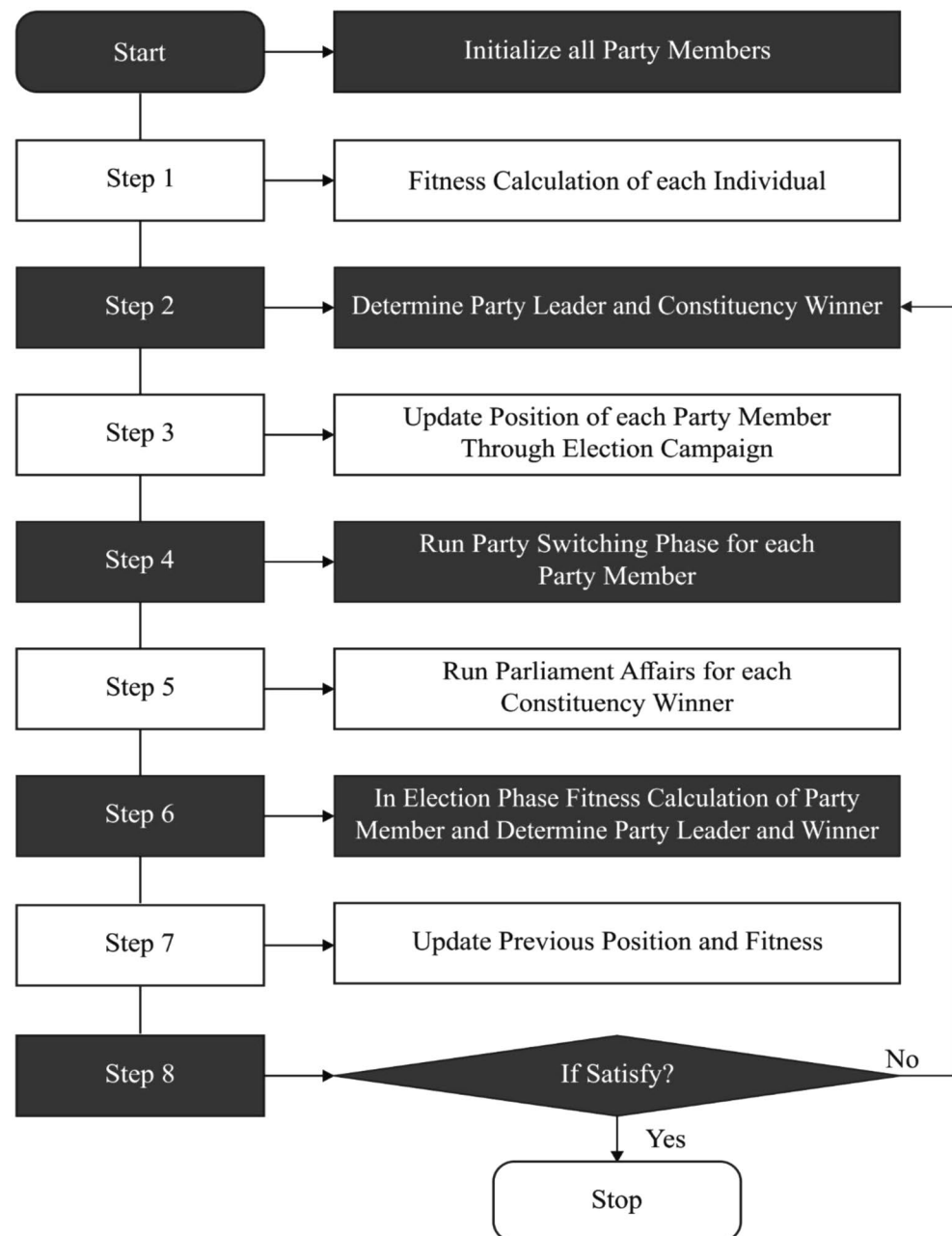


Figure 3. Flowchart of political optimizer.

Root-mean-square error (RMSE). The root-mean-square deviation (RMSD) or root-mean-square error (RMSE) is a frequently used measure of the differences between values (sample or population values) predicted by a model or an estimator and the values observed. The RMSD represents the square root of the second sample moment of the differences between predicted values and observed values or the quadratic mean of these differences.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - \hat{x}_i)^2}{N}} \quad (9)$$

RMSE = root-mean-square deviation, i = variable i , N = number of non-missing data points, x_i = actual observations time series, \hat{x}_i = estimated time series.

Mean absolute percentage error (MAPE). Mean absolute percentage error (MAPE) is a metric that defines the accuracy of a forecasting method. It represents the average of the absolute percentage errors of each entry in a dataset to calculate how accurate the forecasted quantities were in comparison with the actual quantities.

Table 1 shows the OERNN-BPP Model parameter information.

Firstly, a brief investigation of the biofuel production rate of the OERNN-BPP model is investigated for a period of 6 years (2015–2020) in Fig. 4 and Table 2. The results demonstrated the original data, predicted data by OERNN-BPP technique, and divergence rate. The obtained values portrayed that the OERNN-BPP technique has effectively predicted the biofuel production rate and the divergence rate is found to be minimal. Moreover, the divergence at a gets increased with the increase in duration.

Another results analysis of the OERNN-BPP technique takes place interms of biofuel production cost for certain duration in Fig. 5. The figure portrayed that the OERNN-BPP technique has depicted effective performance with as light difference in the actual and predicted data. At the same time, the GWO based LSTM-RNN model has tried to demonstrate reasonable outcomes. However, the OERNN-BPP technique has out performed the existing one with a higher predictive outcome.

Model	Meaning	Value
OERNN-BPP	Number of input layer nodes	8
	Number of hidden layer nodes	15
	Number of output layer nodes	1
	Epochs of training	1000
	Learning rate	0–1
	Momentum coefficient	0–1
	Batch Size	100

Table 1. OERNN-BPP model parameter information.

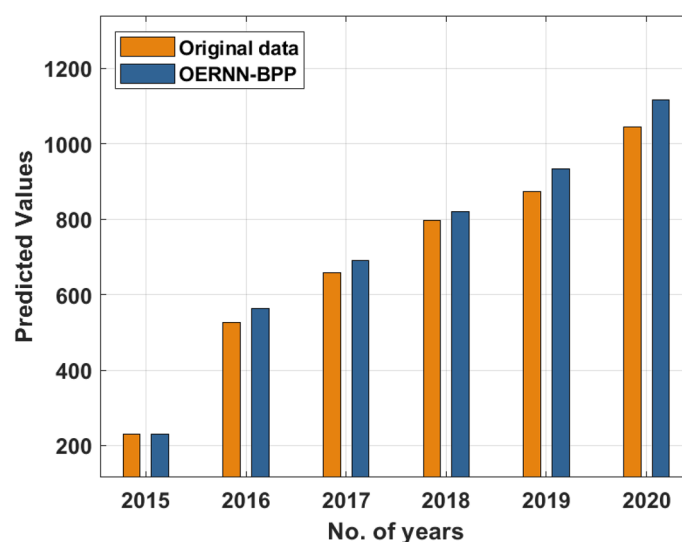


Figure 4. Result analysis of OERNN-BPP model.

No. of years	Original data	OERNN-BPP	Divergence data
2015	231	231	0
2016	527	562.43	35.43
2017	658	691.02	33.02
2018	797	819.33	22.33
2019	874	932.75	58.75
2020	1045	1116.23	71.23

Table 2. Results analysis of proposed OERNN-BPP model.

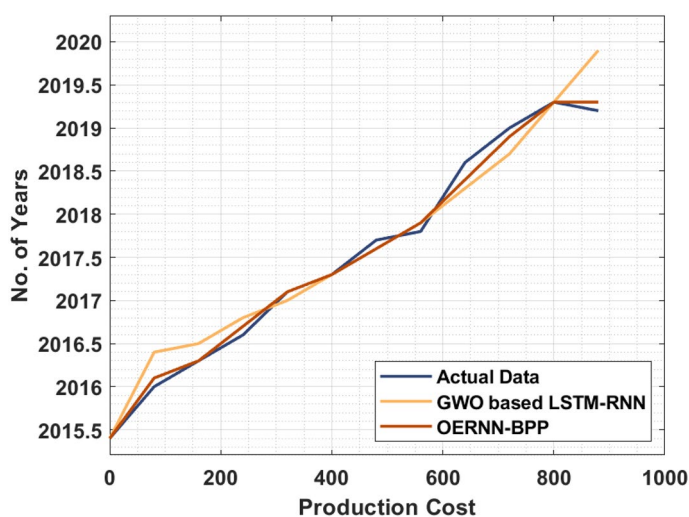


Figure 5. Results of biofuel production.

The existing approaches are empirical mode decomposition (EMD)- (Glow worm Swarm Optimization (GSO)-Long Short-Term Memory (LSTM)—Recurrent Neural Networks (RNNs))- called as EMD-(GWO-LSTM-RNN)-ADD, empirical mode decomposition (EMD)- (Glow worm Swarm Optimization (GSO)-Long Short-Term Memory (LSTM)—Recurrent Neural Networks (RNNs) + adaptive wavelet neural network (AWNN))- called as EMD-(GWO-LSTM-RNN + AWNN)-ADD, and empirical mode decomposition (EMD)-adaptive wavelet neural network (AWNN))-ADD is called as EMD-AWNN-ADD.

A brief comparative root-mean-squared error (RMSE) analysis of the OERNN-BPP technique under varying time duration takes place in Table 3 and Fig. 6. The value of RMSE tends to be minimal for better prediction outcomes. The figure reported that the AWNN technique has appeared as the poor performer with the higher RMSE values. Followed by, the ARIMA model has gained slightly enhanced RMSE value over the AWNN technique whereas the GWO based LSTM-RNN technique has demonstrated moderately RMSE value. In line with, the AION technique has exhibited reasonably reduced RMSE value. However, the proposed OERNN-BPP technique is found to be efficient with the minimum RMSE values under varying time duration.

A detailed comparative MAPE analysis of the OERNN-BPP approach under varying time duration take place in Table 4 and Fig. 7. The value of MAPE tends to be lower for better prediction result. The figure stated that the AWNN manner has demonstrated as the least performer with the maximum MAPE values. Besides, the ARIMA approach has attained some what improved MAPE value over the AWNN technique where as the GWO based LSTM-RNN technique has demonstrated moderately MAPE value. Similarly, the AION method has exhibited reasonably lower MAPE value. However, the proposed OERNN-BPP methodology is established to effectual with the decreased MAPE value under varying time duration.

Finally, a CT analysis of the OERNN-BPP technique with existing techniques is made in Table 5 and Fig. 8. The experimental results highlighted that the GFMP technique has offered worse outcomes with the least CT of 1853.391. Besides, the EMD-LSTM-ELM and EMD-(GWO-LSTM-RNN + AWNN) techniques have tried to show moderate outcomes with the CT of 2101.430 and 2356.358 respectively. However, the proposed OERNN-BPP technique has resulted in superior outcomes with a maximum CT of 3145.152. From the detailed result analysis, it is ensured that the OERNN-BPP technique is found to be an effective tool to predict bio fuel productivity.

Conclusion

In this study, a new OERNN-BPP technique has been presented to predict the productivity of bio fuels. The OERNN-BPP technique encompasses four major processors namely EMD based processing, FTC based reconstruction, ERNN based prediction, and PO based hyper parameter optimization. Once the input data

RMSE	Time in hour				
	ARIMA	AWNN	GWO based LSTM-RNN	AION	OERNN-BPP
1200	12.13438	10.63465	7.635187	6.555381	3.845477
1400	9.134918	7.815155	5.715531	3.495929	1.696047
1600	7.875144	7.335241	5.655542	3.195983	2.496080
1800	8.535026	9.494854	5.955488	3.855865	1.856026
2000	13.69410	13.87407	10.87461	16.21365	9.074907
2200	19.27310	17.53341	14.05404	15.97369	10.87423
2400	19.63304	18.73320	17.35344	16.69356	14.60392
2600	18.67321	19.15312	17.53341	15.25382	13.73389
2800	17.71338	18.55323	16.69356	16.15366	12.96385
3000	16.69356	16.81354	17.05350	16.51360	14.11400
3200	16.63357	17.05350	18.13331	16.63357	12.71390
3400	17.41343	19.09313	17.35344	15.49378	12.78389
3600	17.41343	18.79319	16.99351	15.61376	12.75389
3800	16.75355	16.87353	17.29346	14.23400	12.23398
4000	15.97369	16.39362	16.27364	14.05404	11.37412
4200	15.07385	15.85371	15.67375	12.25436	10.52428
4400	13.87407	16.09367	13.09421	10.81462	9.534470
4600	13.57407	14.83390	12.89421	13.81408	8.934570
4800	13.27407	13.57412	12.69421	16.81354	9.274680

Table 3. RMSE analysis of OERNN-BPP model with existing techniques.

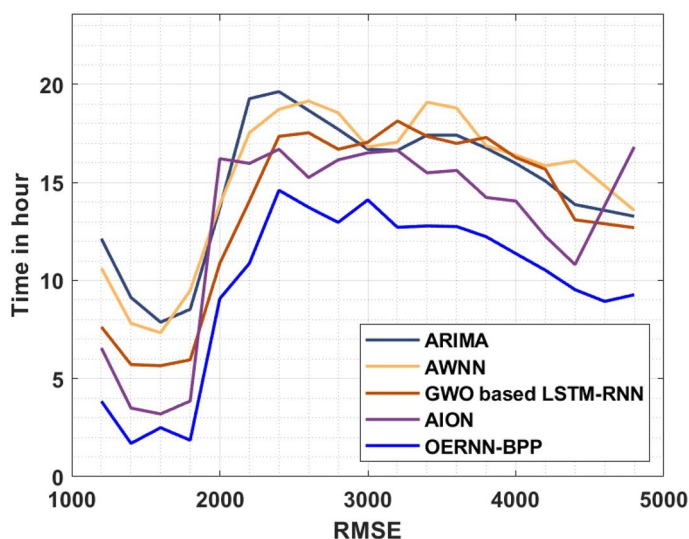


Figure 6. RMSE of different methods.

is pre-processed and reconstructed, the actual prediction process is carried out by the use of ERNN model. Moreover, the hyper parameters of the ERNN model namely learning rate, batch size, momentum, and weight decay are optimally adjusted by the use of PO and thereby raise the predictive outcome to a maximum extent. A thorough simulation analysis is conducted in order to show how the OERNN-BPP technique has improved, and the findings are examined from several angles. The testing results showed that the OERNN-BPP technique outperformed recently published predictive models of biofuel production. It shows that, the EMD-LSTM-ELM and EMD-(GWO-LSTM-RNN + AWNN) techniques have tried to show moderate outcomes with the CT of 2101.430 and 2356.358 respectively. However, the proposed OERNN-BPP technique has resulted in superior outcomes with a maximum CT of 3145.152. Advanced hybrid deep learning architectures may be presented in the future to improve the predicted outcome.

MAPE	Time in hour				
	ARIMA	AWNN	GWO based LSTM-RNN	AION	OERNN-BPP
2400	12.2839	7.602508	11.95422	4.305756	1.166730
2800	9.646495	5.954132	8.921209	3.184860	2.563964
3200	7.998119	5.954132	3.844211	2.393639	1.346484
3600	8.525599	9.184950	3.250795	3.052990	1.780101
4000	9.053080	10.76739	3.910146	3.712340	2.533315
4400	9.580560	12.34983	7.569496	4.371691	1.136730
4800	10.10804	13.93227	12.22885	5.031041	2.720146
5200	10.63552	14.52569	17.03122	6.690392	3.383561
5600	11.16300	17.29496	16.56968	7.349742	2.976976
6000	14.19601	17.62464	17.75651	10.00909	8.695274
6400	19.07521	16.83342	17.09716	14.78943	11.40260
6800	19.47082	16.96529	15.64658	16.10813	14.21788
7200	16.96529	18.08618	16.50374	14.39382	12.90918
7600	16.83342	17.62464	18.02025	14.92130	12.73634
8000	17.49277	17.09716	17.69057	15.77845	13.15195
8400	17.42683	17.42683	16.37187	15.91032	12.16195
8800	16.96529	16.24000	17.22903	15.51471	12.82227
9200	15.97626	15.97626	14.85536	14.39382	11.12292
9600	14.72349	15.97626	14.59162	13.60260	9.820160
10,000	13.99821	14.85536	15.25097	12.34983	10.29674

Table 4. MAPE analysis of OERNN-BPP model with existing techniques.

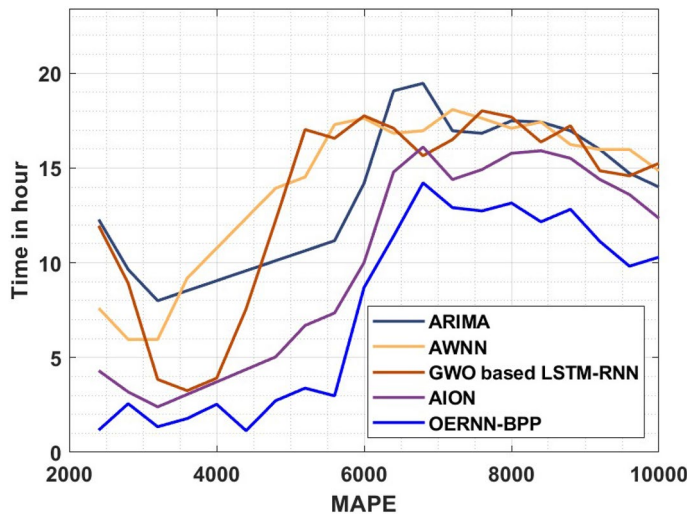


Figure 7. MAPE of different methods.

Methods	Rate of production
GFMP	1853.391
EMD-LSTM-ELM	2101.430
EMD-(GWO-LSTM-RNN + AWNN)	2356.358
OERNN-BPP	3145.152

Table 5. Rate of production on various methods.

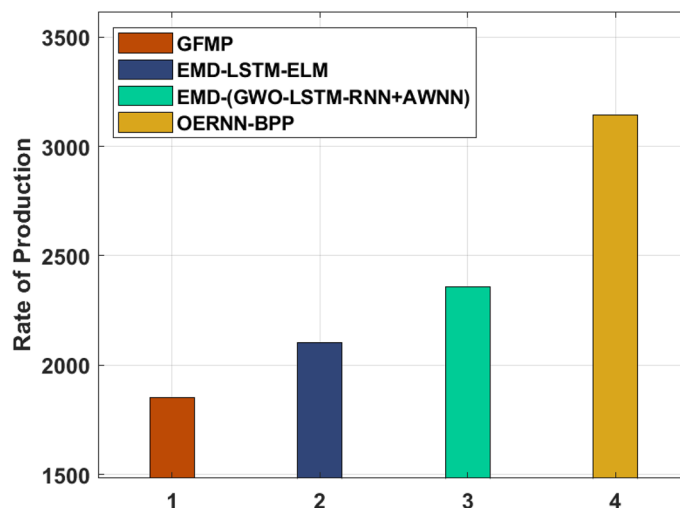


Figure 8. Rate of production on various methods.

Data availability

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

Received: 31 October 2022; Accepted: 7 May 2023

Published online: 26 May 2023

References

- Ardabili, S., Mosavi, A. & Kóczy, A. R. V. *Deep Learning and machine learning models in biofuels research: Systematic review.* (2020).
- Shaosen, S. *et al.* Experimental and artificial intelligence for determination of stable criteria in cyclic voltammetric process of medicinal herbs for biofuel cells. *Int. J. Energy Res.* **43**, 5983–5991 (2019).
- Afsharzade, N. *et al.* Renewable energy development in rural areas of Iran. *Renew. Sustain. Energy Rev.* **65**, 743–755 (2016).
- Wong, K. I. & Wong, P. K. Adaptive air-fuel ratio control of dual-injection engines under biofuel blends using extreme learning machine. *Energy Convers. Manag.* **165**, 66–75 (2018).
- Concu, R., DS cordeiro, M. N., Munteanu, C. R. & González-D'iaz, H. PTML model of enzyme subclasses for mining the proteome of biofuel producing microorganisms. *J. Proteome Res.* **18**, 2735–2746 (2019).
- Smuga-Kogut, M., Kogut, T., Markiewicz, R. & Slowik, A. Use of machine learning methods for predicting amount of bioethanol obtained from lignocellulosic biomass with the use of ionic liquids for pretreatment. *Energies* **14**, 243 (2021).
- Kana, E. B. G., Oloke, J. K., Lateef, A. & Adesiyun, M. O. Modeling and optimization of biogas production on saw dust and other co-substrates using artificial neural network and genetic algorithm. *Renew. Energy* **46**, 276–281 (2012).
- Shi, Y., Gai, G., Zhao, X., Zhu, J. & Zhang, P. Back propagation neural network (BPNN) simulation model and influence of operational parameters on hydrogen bio-production through integrative biological reactor (IBR) treating wastewater. in *2010 4th International Conference on Bioinformatics and Biomedical Engineering* 1–4 (2010).
- Almeida, J. S. Predictive non-linear modeling of complex data by artificial neural networks. *Curr. Opin. Biotechnol.* **13**, 72–76 (2002).
- Anderson, R. *et al.* An integrated modeling framework for crop and biofuel systems using the DSSAT and GREET models. *Environ. Model. Softw.* **108**, 40–50 (2018).
- Yu, L., Liang, S., Chen, R. & Lai, K. K. Predicting monthly biofuel production using a hybrid ensemble forecasting methodology. *Int. J. Forecast.* **38**, 3–20 (2022).
- Radivojević, T., Costello, Z., Workman, K. & Garcia Martin, H. A machine learning automated recommendation tool for synthetic biology. *Nat. Commun.* **11**, 1–14 (2020).
- Elveny, M., Hosseini, M., Chen, T.-C., Lawal, A. I. & Alizadeh, S. M. Estimation of isentropic compressibility of biodiesel using ELM strategy: Application in biofuel production processes. *Biomed. Res. Int.* **2021**, (2021).
- Cui, Z. *et al.* Improving the CFPP property of biodiesel via composition design: An intelligent raw material selection strategy based on different machine learning algorithms. *Renew. Energy* **170**, 354–363 (2021).
- Kumar, N. P., Vijayabaskar, S. & Murali, L. Forecasting biofuel production using adaptive integrated optimization network model. *Fuel* **283**, 118764 (2021).
- Javed, S. A., Zhu, B. & Liu, S. Forecast of biofuel production and consumption in top CO₂ emitting countries using a novel grey model. *J. Clean. Prod.* **276**, 123997 (2020).
- Beeravalli, V. *et al.* Ranking the feedstocks using neural network-based system for biofuel production. in *2021 6th International Conference for Convergence in Technology (I2CT)* 1–5 (2021).
- Geng, R., Indranil, B. & Xi, C. Prediction of financial distress: An empirical study of listed Chinese companies using data mining. *European J. Operat. Res.* **241**(1), 236–247 (2015).
- Mejdoub, H. & Ghorbel, A. The dynamic relationship between oil prices and returns on renewable energy companies. *American J. Fin. Account.* **5**(2), 173–192 (2018).
- Song, W. & Yu, H. Green Innovation Strategy and Green Innovation: The Roles of Green Creativity and Green Organizational Identity. *Corp. Soc. Res. Environ. Manag.* **25**(2), 135–150 (2018).
- Weigend, A. S. *Time series prediction: forecasting the future and understanding the past.* (Routledge, 2018).
- Feng, K., Ying, Z. F., Wu, J. G. & Wan, M. Short-term wind speed forecast based on wavelet packet decomposition and peak-type Markov chain. *J. Nanjing Univ. Sci. Technol* **38**, 639–643 (2014).
- Elman, J. L. Finding structure in time. *Cogn. Sci.* **14**, 179–211 (1990).

24. Wang, J., Wang, J., Fang, W. & Niu, H. Financial time series prediction using elman recurrent random neural networks. *Comput. Intell. Neurosci.* **2016**, (2016).
25. Manita, G. & Korbbaa, O. Binary political optimizer for feature selection using gene expression data. *Comput. Intell. Neurosci.* **2020**, (2020).

Author contributions

Conceptualization, P.K.N, V.S., M.L. and K.R.; Data curation, P.K.N., V.S., M.L. and K.R.; Analysis and Validation, P.K.N., V.S., M.L. and K.R.; Formal analysis, P.K.N., V.S., M.L. and K.R.; Investigation, P.K.N., V.S., M.L. and K.R.; Methodology, P.K.N., V.S., M.L. and K.R.; Project administration, K.R. Resources, P.K.N., V.S., M.L. and K.R.; Software, P.K.N., V.S., M.L. and K.R., Supervision, P.K.N., V.S., M.L. and K.R.; Validation, P.K.N., V.S., M.L. and K.R.; Visualization, P.K.N, V.S, M.L. and K.R.; Writing—original draft, P.K.N, V.S., M.L. and K.R., Data Visualization, Editing and Rewriting, P.K.N., V.S., M.L. and K.R.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to K.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023