SHORT TERM WIND SPEED PREDICTION USING SEARCH AND RESCUE OPTIMIZATION WITH DEEP BELIEF NETWORK ON SCATTEROMETRY DATA

by

Fatimah ALMULHIM[a*], Anis Ben ISHAK[b], Azhari A. ELHAG[c], Hassan M. ALJOHANI[c], and Sayed ABDEL-KHALEK[c]

[a] Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
[b] Department of Quantitative Methods, Higher Institute of Management, University of Tunis, Tunis, Tunisia
[c] Department of Mathematics and Statistics, College of Science, Taif University, Taif, Saudi Arabia

Original scientific paper
https://doi.org/10.2298/TSCI2406097A

Scatterometry is a technique used to transmit radio or microwaves to examine different geophysical properties, wind speed, and direction. Precise and rapid weather predictions become essential in several applications in assisting planning and management in response to weather conditions. At the same time, timely wind speed prediction gains considerable attention in several economical, business, and management areas. With the consideration of wind speed as an arbitrary variable, precise wind speed prediction using machine learning and deep learning models can be established. With this motivation, this study develops a short-term wind speed prediction using search and rescue optimization with deep belief network (STWSP-SRDBN) model. To accomplish accurate wind speed prediction, the STWSP-SRDBN method initially pre-processes the weather data using min-max normalization. Additionally, the STWSP-SRDBN model utilizes DBN model to predict the weather data. Moreover, the SRO algorithm is utilized to fine tune the hyperparameters related to the DBN approach. The presented STWSP-SRDBN method makes use of Spatio-temporal multivariate multi-dimensional historical weather data to learn new representations utilized for wind forecasting. The experimental validation of the STWSP-SRDBN method is tested using a set of weather data and the outcomes are investigated under numerous aspects. The experimental results indicated the enhanced outcomes of the STWSP-SRDBN method over recent state of art methods.

Key words: *wind speed prediction, search optimization, prediction models, time-series forecasting, hyperparameter tuning, deep learning*

Introduction

Scatterometry is a commonly known method under which the transmission of micro or radio waves takes place for measuring several wind speed, directions, and geophysical properties. The waves which are transmitted will spread towards the antenna once after striking a medium or a surface [1]. The amplitude of such backscattered pulses is examined to measure the favorable output units. The usual applications of scatterometry are in wind analysis [2].

* Corresponding author, e-mail: Faalmulhim@pnu.edu.sa

Wind scatterometers may extract information regarding the wind velocity over ocean surfaces. For establishing this, wind scatterometers transfer radar pulses which scatter back upon striking the surface of the ocean [3]. The backscattered signals have variations depending on speed of the wind and roughness of the ocean. Bragg scattering is the main component in the extraction of direction wind and speed of the wind meanwhile the radiation is dispersed in every direction's distinct reflection from flat surfaces. Scatterometers dimensions have to sustain a maximal range of incidence angle and Braggs scattering is reliable on it [4]. Wind scatterometers perform an important role in wave and weather predicting and in the examination of climate models. Future weather patterns might be forecasted with the assimilation of scatterometer data into atmospheric estimating model. There exist several remote sensing methods. Out of these, scatterometry is regarded as a special method and it has the capability of remotely measuring surface of the wind speed and direction through ocean surfaces [5]. This provides predictors worthy information regarding cyclones at the initial phase of development.

In general, the researchers differentiated the prediction methods applied in the wind speed predicting domain into three major groups they are, artificial intelligence methods, physical approaches, and data-driven models depending on statistical theory [6]. The physical methodology is termed the numeric weather prediction technique a common information-driven method that involves a sequence of physical information like humidity, atmospheric pressure, obstacles, roughness, wind direction, temperature*, etc.* [7]. In comparison the physical methodology, the statistical models always had a priority for mining the implicit data which is available in the historical wind speed that becomes very familiar by the right of the minimal information needed. Over the past few years, the speedy advancement of artificial intelligence technologies has drastically brought forth a progression in wind speed estimation [8]. Assuming wind speed as a random parameter, its precise estimation over the duration will earn superior outcomes for the wind turbine operator. Machine learning (ML) has the capability of estimating environmental and hydrological method variables precisely [9]. The ML contribution is very significant in wind speed estimation, and choosing suitable ML model is highly significant in order to get precise outcomes. The ML methods *i.e.* extreme learning machine (ELM), Gaussian process regression (GPR), support vector machine (SVM), fuzzy logic (FL), and ANN are extensively utilized for this purpose. Currently, the implementation of hybrid ML tools for wind speed estimation has gained much more impetus amongst the researchers since it owns the benefits of every method [10]. This article focuses on the design of a short-term wind speed prediction using search and rescue optimization with deep belief network (STWSP-SRDBN) model. The STWSP-SRDBN method firstly pre-processes the weather data by means of min-max normalization. Also, the STWSP-SRDBN method utilizes DBN model to predict the weather data. Moreover, the SRO algorithm is utilized to fine tune the hyperparameters related to the DBN approach. The presented STWSP-SRDBN model makes use of Spatio-temporal multivariate multi-dimensional historical weather data to learn new representations utilized for wind forecasting. The experimental validation of the STWSP-SRDBN method is tested with a set of weather data and the results are investigated under several aspects.

Literature review

Valsaraj *et al.* [11] examined the perception of applying one time trained ML prediction method for wind speed predicting individual location and time. It is discovered in the investigation that ML techniques trained by the preceding wind speed dataset from one position that can predict wind speed efficiently at another location of interest within an extensive geographical area. Hur [12], the authors presented a wind speed predictive technique encompasses primarily two-phases, prediction, and estimation. Firstly, estimation is conducted by an extended kalman filter, *i.e.*, developed according to non-linear rotor and 3-D wind field models. Subsequently, prediction can be carried out in two major stages, ML and extrapolation. Demolli *et al.* [13], the authors performed a long-term wind power prediction on the basis of wind speed dataset with five ML techniques. Then, presented a technique on the basis of ML approaches for forecasting wind power value effectively. Then, carried out different researches for revealing the efficiency of ML approaches.

Li and Jin [14], a hybrid model architecture based on combinatory models has been developed and effectively adapted to build the predictive ranges of the upcoming wind speed. Feature selection method is designed for determining the appropriate mode of real time sequence and the optimum input formation of the algorithm, whereas the optimization predicting model is employed for modelling the sequential wind speed on the basis of the multi-objective optimization and ML techniques, later the compromise solution of Pareto front is designated using *min-max* approach. Zhu *et al.* [15] established a technique for forecasting wind speed with spatiotemporal correlations, that is, the prediction deep convolution neural network (PDCNN). The algorithm is an integrated architecture that integrates convolution neural network (CNN) and a multi-layer perceptron (MLP). At first, the spatial feature is extracted by CNN situated at the bottom of the algorithm. Next, the temporal dependencies amongst this spatial feature extraction are taken using the MLP. Accordingly, the temporal and spatial relationships are taken using PDCNN. At last, PDCNN generates the forecasted wind speed with the learned spatiotemporal relationship. Li [16], a two-phase technique has been designed for stronger spatio-temporal approximations of wind speed at a higher resolution. The presented technique comprises geologically weighted ensemble ML (Phase 1) and downscales according to weather-related re-analysis dataset (Phase 2). The geologically weighted ML technique depends on three basic learners, that is, Autoencoder-based random forest, deep residual network, and XGBoost, also it integrates heterogeneity and spatial autocorrelation increase the ensemble prediction. Luo *et al*. [17], the authors proposed a wind-speed predicting technique based on two kinds of ML methods (multi-objective and decomposition-ensemble optimization) that address the chaotic and non-linearity features of sequential wind-speed.

The proposed model

In this study, a STWSP-SRDBN model has been developed for effective wind speed prediction. The suggested model encompasses a three-stage process. At the initial stage, the *min-max* normalization approach is utilized to pre-process the input weather data. Next, the STWSP-SRDBN model utilizes DBN model to predict the weather data. Finally, the SRO algorithm is utilized to fine tune the hyperparameters related to the DBN model. Figure 1 shows the workflow of proposed model.

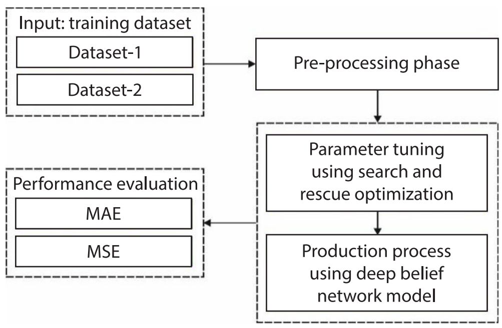

Figure 1. Workflow of STWSP-SRDBN model

Stage I: Data pre-processing

Primarily, the *min-max* normalization approach is utilized to pre-process the input weather data. Data normalization is commonly utilized to attain enhanced results by any ML model. The feature values range from low to high. Therefore, it becomes essential to scale the features to unit variance using min-max normalization approach. It is typically leveraged for computing the similarity degree among points. Consider A as data that should be mapped from dataset ranges from $A_{\min}$ to $A_{\max}$ using:

$$A_{\text{normalized}} = \frac{A - A_{\min}}{A_{\max} - A_{\min}} \tag{1}$$

The utilization of min-max normalization ensured that features are exacted into identical scales.

Stage II: The DBN based prediction model

Once the input data is pre-processed, the STWSP-SRDBN model utilized the DBN model to predict the weather data [18]. The DBN is a kind of deep neural network with large amounts of hidden units and many hidden layers. The standard DBN is equal to stacked restricted boltzmann machine (RBM) model using output layer. The DBN exploits a faster, greedy unsupervised learning technique for RMB training and a supervised fine-tuning technique to alter the network by labeled dataset. Every RBM comprises a hidden layer h and visible layer v, interconnected with undirected weight. For stacked RBM in the DBN, hidden unit of single RBM is regarded the visible unit of the following RBM. The variable set of a RMB is $\theta = (w, b, \text{a})$, where w_{ij} indicates the weight between v_i and h_j. The b_i and b_j represent the bias. A RBM determines its energy:

$$E(v,h \mid \theta) = -\sum_i b_i v_i - \sum_j a_j h_j - \sum_i w_{ij} v_i h_j \tag{2}$$

and the joint likelihood distribution of v and h is represented:

$$p(v,h \mid \theta) = \frac{\exp(-E(v,h \mid \theta))}{\sum_{v,h} \exp(-E(v,h \mid \theta))} \tag{3}$$

and the marginal likelihood distribution of v is given:

$$p(v \mid \theta) = \frac{\sum_h \exp(-E(v,h \mid \theta))}{\sum_{v,h} \exp(-E(v,h \mid \theta))} \tag{4}$$

In order to attain the optimum, θ, for single dataset vector, v, the gradient of log-probability can be computed:

$$\frac{\partial \log p(v|\theta)}{\partial w_{ij}} = \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}} \tag{5}$$

$$\frac{\partial \log p(v|\theta)}{\partial a_j} = \langle h_j \rangle_{\text{data}} - \langle h_j \rangle_{\text{model}} \tag{6}$$

$$\frac{\partial \log p(v|\theta)}{\partial b_i} = \langle v_i \rangle_{\text{data}} - \langle v_i \rangle_{\text{model}} \tag{7}$$

where $\langle\cdot\rangle$ represent the expectation under the distribution determined by the subscript, which can either refer to the data distribution or the model distribution.

As there is no link between the units in the same layer, $\langle\cdot\rangle_{\text{data}}$ is easier to attain by evaluating the conditional likelihood distributions:

$$p(h_j | v,\theta) = \frac{1}{1+\exp\left(-\sum_i w_{ij} v_i - a_j\right)}$$
$$p(v_i | h,\theta) = \frac{1}{1+\exp\left(-\sum_j w_{ij} h_j - b_i\right)} \tag{8}$$

The activation function is sigmoid function. For $\langle\cdot\rangle_{\text{model}}$ the contrastive divergence (CD), learning model was utilized for reconstruction minimalize the dissimilarity of two Kullback-Leibler (KL) divergences. The CD learning is found to be efficiently practical and minimizes computation costs than standard Gibbs sampling model. The weight in DBN layer is trained by unlabeled dataset with the aforementioned faster and greedy unsupervised method. In this work, the fully connected layer was designated to implement as the top layer and uses the sigmoid activation function. For predictive purposes, a supervised layer should be added above the DBN to modify the learned feature by labelled dataset with an up-down fine-tuning method. The structure of DBN model is given in fig. 2.

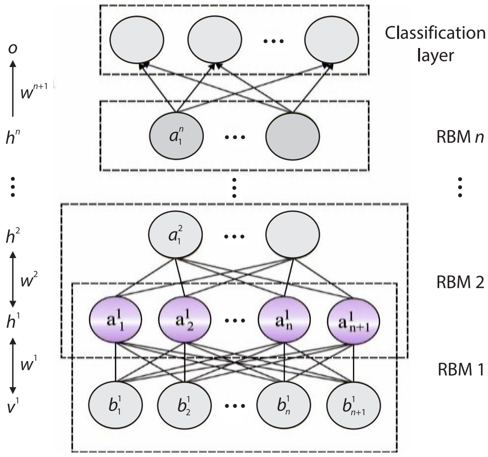

Figure 2. Structure of DBN model

Stage III: The SRO based hyperparameter tuning

In the last stage, the SRO method is utilized to fine tune the hyperparameters related to the DBN model. The location of the lost human is the most important objective of the search and rescue optimization method for optimization problems, and the consequence of the clue originating in this location describes the cost of solution. Now, the best approach discloses the best location with further clues [19]. Individuals search for best possible solution through the searching technique when leaving some hints. However, the search place for the individual is kept in a situation matrix (matrix X) using the equal size of memory matrix and the left clue was kept in a memory matrix (matrix M), n is the individual quantity in a group, and $n \times d$ determines the variable problem:

$$C = \begin{bmatrix} X \\ M \end{bmatrix} = \begin{bmatrix} X_{1,1} & \cdots & X_{1,d} \\ \vdots & \ddots & \vdots \\ X_{n,l} & \cdots & X_{n,d} \\ M_{1,1} & \cdots & M_{1,d} \\ \vdots & \ddots & \vdots \\ M_{n,l} & \cdots & M_{n,d} \end{bmatrix} \tag{9}$$

In eq. (9), consider random clues among the attained clues, the searching direction can be obtained:

$$sd_i = \left(X_j - C_k\right), \quad k \neq i \tag{10}$$

where k indicates random values amongst 1 and $2N$, X, and C_k correspondingly describes the location of i^{th} human and the k^{th} clue. The presented algorithm makes use of binomial crossover operators to employ the limitation. Furthermore, the clue has great consequences than that of existing clue, a region was searched for sp_i direction. Or else, search for the place of the present location in the sp_i is constant [20]. Hence, the novel location of j^{th} parameter can be illustrated as i^{th} human:

$$X_{ij} = \begin{cases} \text{Matrix}\begin{pmatrix} X_{ii} - C_{ki} \\ X_{ii} - C_{ki} \end{pmatrix} \text{ if } f_{\text{otherwise}}\left(C_k\right) > f\left(X_i\right) \text{ if } r_2 < se \text{ or } j = j_r, \; j = 1, \ldots, d \\ X_i \text{ otherwise} \end{cases} \tag{11}$$

$$X_{ij} = \begin{cases} \begin{cases} c_{k,j} + r_1 \times \left(X_{i,j} - C_{k,j}\right) & \text{if } f\left(C_k\right) > f\left(X_i\right) \\ X_{i,j} + r_1 \times \left(X_{i,j} - C_{k,j}\right) & \text{otherwise} \end{cases} & \text{if } r_2 < se \text{ or } j = j_r, \; j - 1, \ldots, d \\ X_i & \text{otherwise} \end{cases} \tag{12}$$

where $c_{k,j}$ is the signifies location of j^{th} variable and the clue kj_r, r_1, and r_2 correspondingly characterizes three arbitrary integer lies in the range of [0, 1], [1, d], and [–1, 1]. The next phase is about individuals. The upgraded position using the i^{th} human is obtained by using:

$$X_i^{'} = X_j + r_3\left(C_k - C_m\right) \tag{13}$$

where r_3 is the uniform distribution value lies in the range of [0, 1], m and k are the two arbitrary integers within [1, $2N$] and $i \neq k \neq m$. When the solution is placed outside the border the subsequent equation can be employed:

$$X_{ij}' = \begin{cases} \dfrac{\left(X_{ij} + X_j^{\max}\right)}{2} & \text{if } X_{ij}' > X_j^{\max} \\ \dfrac{\left(X_{i,j} + X_j^{\min}\right)}{2} & \text{if } X_{ij}' < X_j^{\min} \end{cases} \tag{14}$$

where $j = 1, 2, ..., d$ is the minimal and maximal threshold for the parameter can be expressed as X_j^{miin} and $X_j^{\max}$ correspondingly. Based on the process, the human lost candidate is searched according to the formerly described method. Once the amount of cost functions in $X_i'(f(X_i'))$ scenario is greater than the current one ($f(X_i)$) the prior position (X) could be saved in a random place in the memory matrix, M, or else, the situation is left and the memory hasn't been upgraded:

$$M_n\left(X_i\right) = \begin{cases} X_j \text{ if} & f\left(X_i'\right) > f\left(X_i\right) \\ M_n\left(X_i\right) & \text{otherwise} \end{cases} \tag{15}$$

where n is the random integer that lies in the interval of [1, N] and M_n – the position of n clue numbers in the memory matrix. Once a person during his quest does not discover a prominent clue, it leaves a novel one with the current place. For, usn describes the unproductive searching number we have:

$$usn_i = \begin{cases} usn_i + 1 \text{ if } f\left(X_i'\right) < f\left(X_i\right) \\ 0 \text{ otherwise} \end{cases} \tag{16}$$

If *usn* is higher than *MU*, it moved to another place in the space solution. The SRO algorithm derives a fitness function with the minimization of mean square error (MSE), which can be formulated:

$$MSE = \frac{1}{T}\sum_{j=1}^{L}\sum_{i=1}^{M}\left(y_j^i - d_j^i\right)^2 \tag{17}$$

where *M* and *L* are the define outcomes of the layers and data, correspondingly. Also, y_j^i and d_j^i implies attained and appropriate magnitude for j^{th}. Units in the output layer of network at time, *t*, respectively.

Experimental validation

In this section, the prediction outcomes of the STWSP-SRDBN model is investigated under three distinct cities. The results are assessed in terms of two measures namely mean absolute error (MAE) and MSE. Table 1 and fig. 3 provide a detailed comparative predictive outcome of the STWSP-SRDBN method with recent methods [21] on three cities with 6 hours, 12 hours, 18 hours, and 24 hours prediction on Dataset-1.

Table 1. The MAE Analysis of STWSP-SRDBN model with existing models on Dataset-1

MAE – Dataset-1								
Methods	City-1	City-2	City-3	Average	City-1	City-2	City-3	Average
	6 hours				12 hours			
2-D	1.221	1.182	1.263	1.222	1.626	1.630	1.498	1.585
2-D-Attention	1.197	1.257	1.166	1.207	1.588	1.503	1.565	1.552
2-D-Upscaling	1.241	1.209	1.168	1.206	1.598	1.483	1.560	1.547
3-D	1.265	1.155	1.183	1.201	1.472	1.493	1.562	1.509
Multidimensional	1.195	1.203	1.189	1.196	1.451	1.530	1.471	1.484
STWSP-SRDBN	1.105	1.112	1.125	1.114	1.382	1.421	1.405	1.403
Methods	City-1	City-2	City-3	Average	City-1	City-2	City-3	Average
	18 hours				24 hours			
2-D	1.779	1.791	1.593	1.721	1.888	1.774	1.933	1.865
2-D-Attention	1.673	1.818	1.654	1.715	1.728	1.913	1.943	1.861
2-D-Upscaling	1.574	1.691	1.713	1.659	1.927	1.843	1.806	1.859
3-D	1.588	1.62	1.696	1.635	1.751	1.743	1.944	1.813
Multidimensional	1.605	1.672	1.583	1.620	1.804	1.751	1.741	1.765
STWSP-SRDBN	1.504	1.546	1.516	1.522	1.659	1.678	1.67	1.669

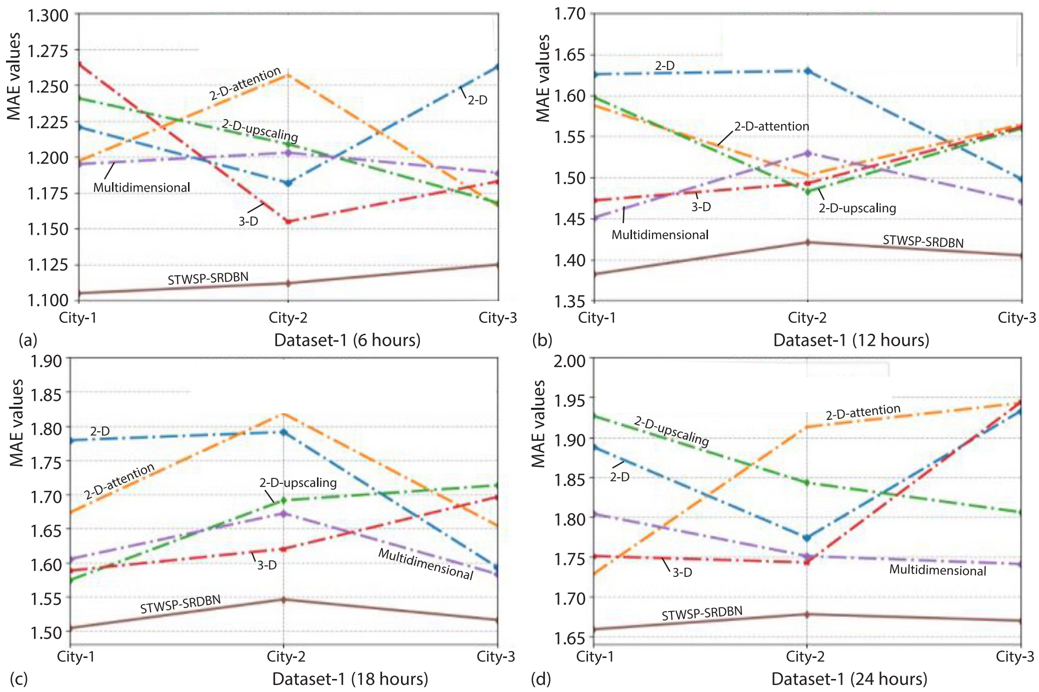

Figure 3. Comparative MAE assessment of STWSP-SRDBN model on Dataset-1; (a) 6 hours, (b) 12 hours, (c) 18 hours, and (d) 24 hours

The results indicated that the STWSP-SRDBN technique has accomplished enhanced performance over other methods. For example, with 6 hours duration and city-1, the STWSP-SRDBN model has offered reduced MAE of 1.105 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional models have reached increased MAE of 1.221, 1.197, 1.241, 1.265, and 1.195, respectively. Similarly, with 12 hours duration and City-1, the STWSP-SRDBN model has provide lower MAE of 1.382 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and Multidimensional models have attained higher MAE of 1.626, 1.588, 1.598, 1.472, and 1.451, respectively.

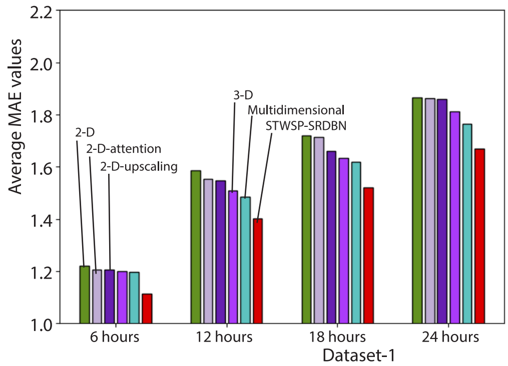

Figure 4. Average MAE assessment of STWSP-SRDBN model on Dataset-1

An average MAE inspection of the STWSP-SRDBN method with existing models on different time durations and Dataset-1 is given in fig. 4. The figure indicated that the STWSP-SRDBN approach has gained lower MAE values under all aspects. For example, with 6 hours prediction, the STWSP-SRDBN method has provided reduced average MAE of 1.114 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional models have obtained increased average MAE of 1.222, 1.207, 1.206, 1.201, and 1.196, respectively. Eventually, with 12 hours prediction, the STWSP-SRDBN model has demonstrated least average MAE of 1.403 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and Multidimensional models have exhibited increased average MAE of 1.585, 1.552, 1.547, 1.509, and 1.484, respectively.

Table 2 and fig. 5 offers an extensive comparative MSE predictive outcomes of the STWSP-SRDBN model on three cities with 6 hours, 12 hours, 18 hours, and 24 hours prediction on Dataset-1. The experimental values demonstrated that the STWSP-SRDBN method has reached improved outcomes over other approaches. For instance, with 6 hours duration and City-1, the STWSP-SRDBN model has exhibited least MSE of 2.187 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional models have accomplished improved MSE of 2.518, 2.659, 2.548, 2.576, and 2.251, respectively. Likewise, with 12 hours duration and City-1, the STWSP-SRDBN model has resulted to minimal MSE of 3.947 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and Multidimensional models have demonstrated maximum MSE of 5.350, 5.327, 5.599, 5.311, and 4.923, respectively.

Table 2. The MSE analysis of STWSP-SRDBN model with existing models on Dataset-1

MSE – Dataset-1								
Methods	City-1	City-2	City-3	Average	City-1	City-2	City-3	Average
	6 hours				12 hours			
2-D	2.518	2.688	2.530	2.579	4.388	4.485	4.215	4.363
2-D-Attention	2.659	2.227	2.550	2.479	4.694	3.995	4.199	4.296
2-D-Upscaling	2.548	2.577	2.301	2.475	4.009	4.554	4.204	4.256
3-D	2.576	2.301	2.274	2.384	4.265	3.915	4.517	4.232
Multidimensional	2.251	2.447	2.291	2.330	4.654	4.116	3.905	4.225
STWSP-SRDBN	2.187	2.158	2.222	2.189	3.947	3.846	3.854	3.882
Methods	City-1	City-2	City-3	Average	City-1	City-2	City-3	Average
	18 hours				24 hours			
2-D	5.350	5.483	5.589	5.474	5.735	5.765	5.681	5.727
2-D-Attention	5.327	5.519	5.330	5.392	5.764	5.482	5.835	5.694
2-D-Upscaling	5.599	5.361	5.102	5.354	5.981	5.560	5.438	5.660
3-D	5.311	5.146	5.362	5.273	5.459	5.539	5.754	5.584
Multidimensional	4.923	5.025	5.434	5.127	5.428	5.512	5.501	5.480
STWSP-SRDBN	4.845	4.956	5.035	4.945	5.362	5.410	5.384	5.385

An average MSE examination of the STWSP-SRDBN model with current models on diverse time durations and Dataset-1 is given in fig. 6. The figure specified that the STWSP-SRDBN model has extended inferior MSE values under all aspects. For instance, with 6 hours prediction, the STWSP-SRDBN model has delivered reduced average MSE of 2.189 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and Multidimensional models have gotten increased average MSE of 2.579, 2.479, 2.475, 2.384, and 2.330, respectively. Finally, with 12 hours prediction, the STWSP-SRDBN model has established minimum average MSE of 3.882 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and Multidimensional models have showed improved average MSE of 4.363, 4.296, 4.256, 4.232, and 4.225, respectively.

A comprehensive comparative MAE predictive results of the STWSP-SRDBN model on 3 cities with 1 hours, 2 hours, 3 hours, and 4 hours prediction on Dataset-2 is demonstrated in tab. 3 and fig. 7. The outcomes denoted that the STWSP-SRDBN model has resulted to

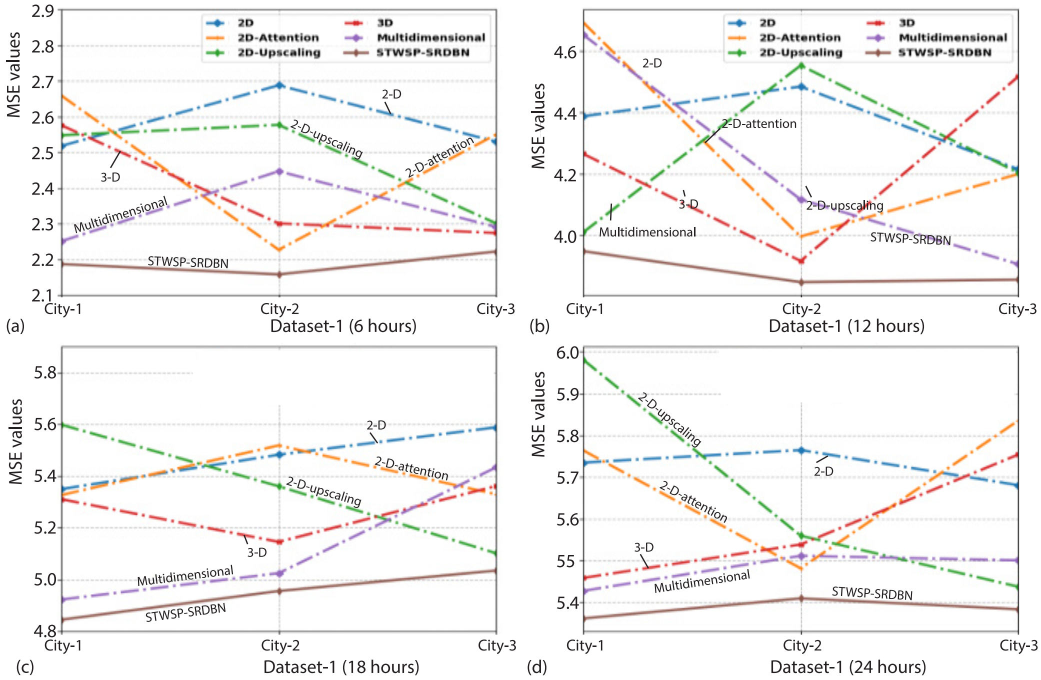

Figure 5. Comparative MSE assessment of STWSP-SRDBN model on Dataset-1; (a) 6 hours, (b) 12 hours, (c) 18 hours, and (d) 24 hours

Table 3. The MAE Analysis of STWSP-SRDBN model with existing models on Dataset-2

MAE – Dataset-2								
Methods	City-1	City-2	City-3	Average	City-1	City-2	City-3	Average
	1 hours				2 hours			
2-D	7.930	7.650	7.990	7.857	8.990	8.150	8.710	8.617
2-D-Attention	7.840	7.510	7.940	7.763	8.260	8.910	8.570	8.580
2-D-Upscaling	7.590	7.870	7.680	7.713	8.510	8.810	8.360	8.560
3-D	7.600	7.510	7.810	7.640	8.550	8.760	8.240	8.517
Multidimensional	7.590	7.710	7.540	7.613	8.140	8.160	8.870	8.390
STWSP-SRDBN	7.533	7.442	7.484	7.486	8.068	8.087	8.173	8.109
Methods	City-1	City-2	City-3	Average	City-1	City-2	City-3	Average
	3 hours				4 hours			
2-D	9.170	9.550	9.640	9.453	10.090	10.070	9.360	9.840
2-D-Attention	9.260	9.690	9.350	9.433	10.340	9.500	9.430	9.757
2-D-Upscaling	9.390	9.120	9.320	9.277	9.080	9.180	10.600	9.620
3-D	9.020	9.150	9.650	9.273	9.070	10.060	9.670	9.600
Multidimensional	9.010	9.370	9.020	9.133	9.250	9.030	9.870	9.383
STWSP-SRDBN	8.951	9.053	8.967	8.990	8.996	8.980	9.301	9.092

better results compared to existing techniques. For instance, with 1 hour duration and City-1, the STWSP-SRDBN model has resulted to lower MAE of 7.533 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional models have gained higher MAE of 7.930, 7.840, 7.590, 7.600, and 7.590, respectively. Furthermore, with 2 hours duration and City-1, the STWSP-SRDBN model has accomplished decreased MAE of 8.068 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and Multidimensional models have obtained increased MAE of 8.990, 8.260, 8.510, 8.550, and 8.140, respectively.

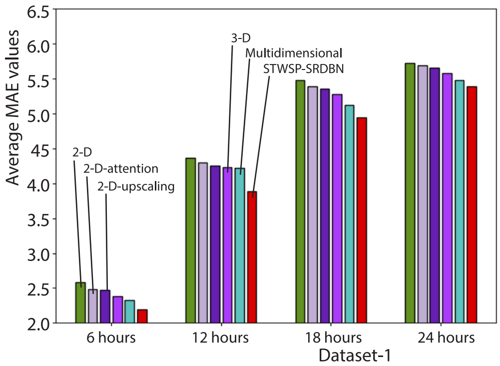

Figure 6. Average MSE assessment of STWSP-SRDBN model on Dataset-1

An average MAE review of the STWSP-SRDBN model with recent approaches on several time durations and Dataset-2 is given in fig. 8. The experimental values denoted that the STWSP-SRDBN method has resulted to least MAE values under all aspects. For instance, with 2 hours prediction, the STWSP-SRDBN model has decreased average MAE of 8.109 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional approaches have reached improved average MAE of 8.617, 8.580, 8.560, 8.517, and 8.390, respectively. At the same time, with 4 hours prediction, the STWSP-SRDBN model has demonstrated least average MAE of 9.092 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional models have showed superior average MAE of 9.840, 9.757, 9.620, 9.600, and 9.383, respectively.

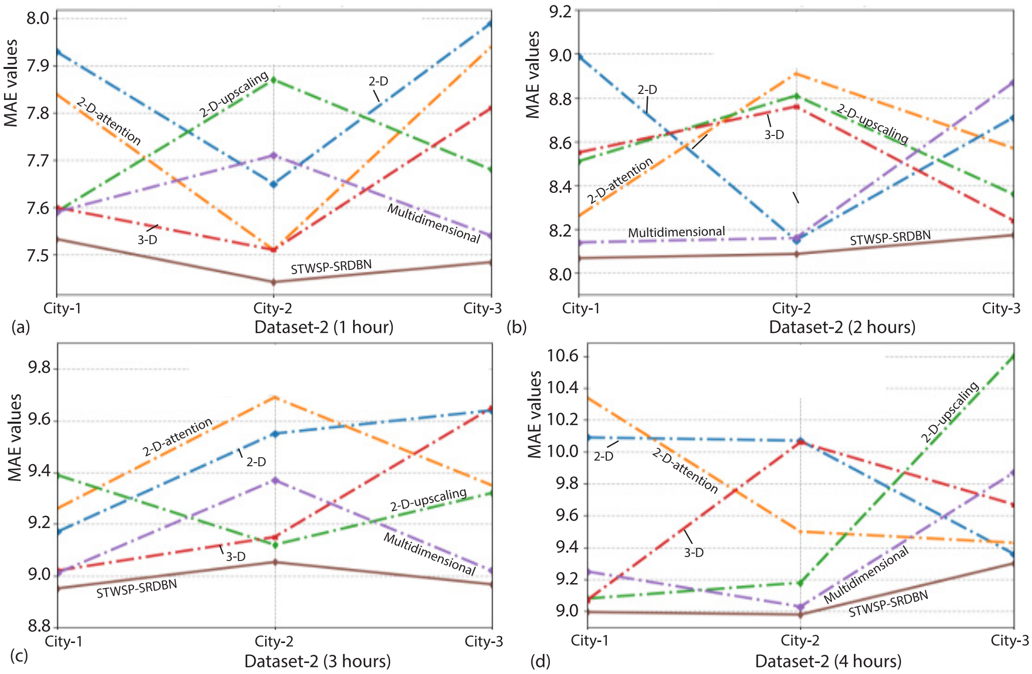

Figure 7. Comparative MAE assessment of STWSP-SRDBN model on Dataset-2; (a) 1 hours, (b) 2 hours, (c) 3 hours, and (d) 4 hours

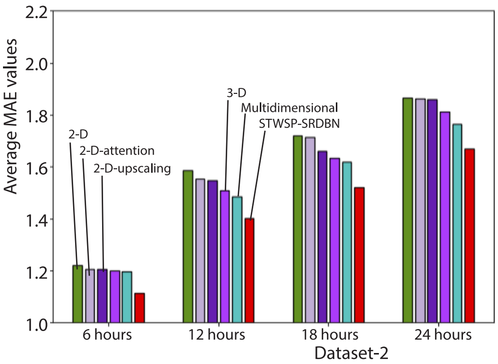

Figure 8. Average MAE assessment of STWSP-SRDBN model on Dataset-2

Table 4 and fig. 9 depicts a brief comparative MSE predictive outcomes of the STWSP-SRDBN model on three cities with 1 hours, 2 hours, 3 hours, and 4 hours prediction on Dataset-2. The obtained outcomes point out that the STWSP-SRDBN model has gotten better outcomes over other methods. For example, with 1 hours duration and city-1, the STWSP-SRDBN model has showed least MSE of 98.810 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional models have accomplished enhanced MSE of 109.000, 108.420, 106.460, 109.250, and 101.090, respectively. Equally, with 2 hours duration and city-1, the STWSP-SRDBN model has resulted to minimal MSE of 142.520 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional models have revealed maximum MSE of 147.340, 147.730, 146.010, 144.200, and 146.580, respectively.

Table 4. The MSE analysis of STWSP-SRDBN model with existing models on Dataset-2

MSE – Dataset-2								
Methods	City-1	City-2	City-3	Average	City-1	City-2	City-3	Average
	1 hours				2 hours			
2-D Model	109.00	108.45	110.85	109.43	147.34	148.78	148.060	148.06
2-D-Att. Model	108.42	114.10	101.62	108.05	147.73	144.25	147.940	146.64
2-D-Upscaling	106.46	110.20	101.59	106.08	146.01	144.81	148.570	146.46
3-D Model	109.25	102.95	104.12	105.44	144.20	146.72	148.370	146.43
Multidimensional	101.09	108.37	102.28	103.91	146.58	145.54	145.500	145.87
STWSP-SRDBN	98.81	101.70	99.57	100.03	142.52	142.11	143.440	142.69
Methods	City-1	City-2	City-3	Average	City-1	City-2	City-3	Average
	3 hours				4 hours			
2-D	182.07	184.18	177.27	181.173	208.62	207.62	204.710	206.98
2-D-Attention	181.01	172.66	188.69	180.787	205.18	207.70	207.290	206.72
2-D-Upscaling	177.01	182.18	182.84	180.677	206.71	207.59	205.410	206.57
3-D	179.75	172.78	187.72	180.083	205.73	208.71	204.990	206.48
Multidimensional	176.11	177.75	176.78	176.880	203.81	203.12	203.430	203.45
STWSP-SRDBN	174.70	170.27	174.31	173.093	202.08	200.89	201.470	201.48

Finally, an average MSE study of the STWSP-SRDBN method with existing models on dissimilar time durations and Dataset-2 is given in fig. 10. The figure portrayed that the STWSP-SRDBN model has decreased MSE values under all aspects. For instance, with 1 hours

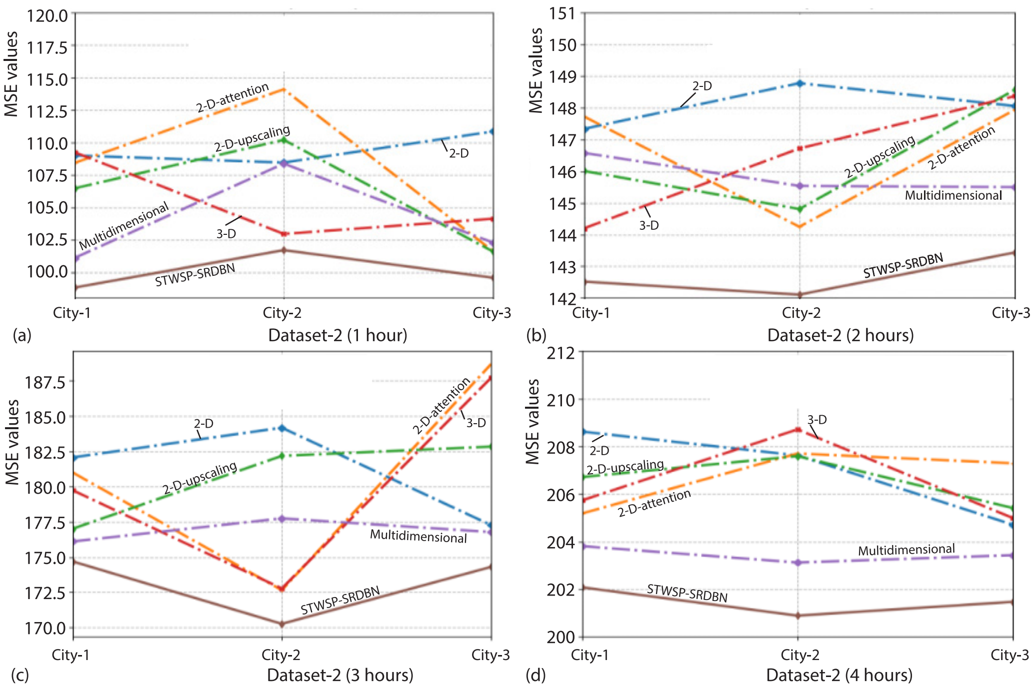

Figure 9. Comparative MSE assessment of STWSP-SRDBN model on Dataset-2; (a) 1 hours, (b) 2 hours, (c) 3 hours, and (d) 4 hours

prediction, the STWSP-SRDBN model has offered minimal average MSE of 100.027 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional models have obtained increased average MSE of 109.433, 108.047, 106.083, 105.440, and 103.913, respectively. At last, with 2 hours prediction, the STWSP-SRDBN model has demonstrated least average MSE of 142.690 whereas the 2-D, 2-D-Attention, 2-D-Upscaling, 3-D, and multidimensional models have exhibited increased average MSE 148.060, 146.640, 146.463, 146.430, and 145.873, respectively.

Figure 11 showcases a detailed predictive result analysis of the STWSP-SRDBN model with 1 hours prediction. The figure specified that the STWSP-SRDBN method has accomplished effectual predictive values under all-time duration. It is noted that the difference between the actual and predictive values are considerably low.

Figure 12 shows an extensive predictive examination of the STWSP-SRDBN technique with 4 hours prediction. The figure stated that the STWSP-SRDBN method has proficient predictive values under all-time durations. It is observed that the variation between the actual and predictive values is significantly low. The aforementioned tables and figures clearly pointed out the supremacy of the STWSP-SRDBN model over recent approaches.

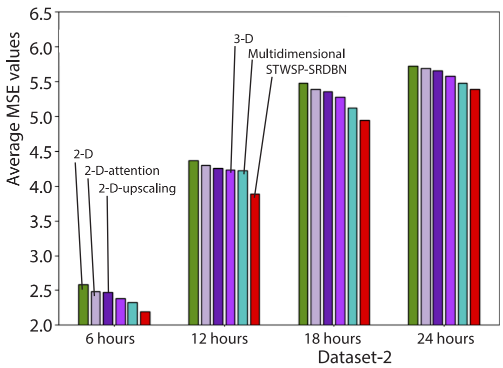

Figure 10. Average MSE assessment of STWSP-SRDBN model on Dataset-2

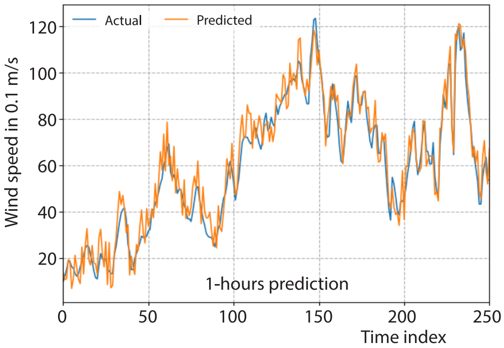

Figure 11. Actual *vs.* predicted values of STWSP-SRDBN model on 1 hour prediction

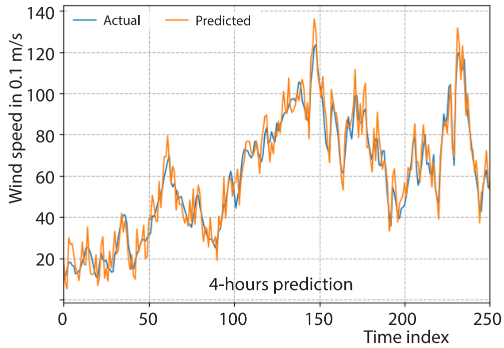

Figure 12. Actual *vs.* predicted values of STWSP-SRDBN model on 4 hours prediction

Conclusion

In this study, a STWSP-SRDBN model has been developed for effective wind speed prediction. The suggested model encompasses a three-stage process. At the initial stage, the *min-max* normalization approach is utilized to pre-process the input weather data. Next, the STWSP-SRDBN model utilizes DBN model to predict the weather data. Finally, the SRO algorithm is utilized to fine tune the hyperparameters relevant to the DBN method. The presented STWSP-SRDBN model makes use of spatio-temporal multivariate multi-dimensional historical weather data to learn new representations utilized for wind forecasting. The experimental validation of the STWSP-SRDBN approach is tested using a set of weather data and the results are investigated under numerous aspects. The experimental outcomes indicated the enhanced outcomes of the STWSP-SRDBN method over recent state of art approaches. As a part of future scope, hybrid deep learning models can be used to increase predictive performance.

Acknowledgment

The authors extend their appreciation to the Deanship of Scientific Research and Libraries in Princess Nourah bint Abdulrahman University for funding this research work through the Research Group project, Grant No. (RG-1445-0035).

References

[1] Nezhad, M. M., *et al.*, Wind Source Potential Assessment Using Sentinel 1 Satellite and a New Forecasting Model Based on Machine Learning: A Case Study Sardinia Islands, *Renewable Energy*, *155* (2020), Aug., pp. 212-224
[2] Liu, W. T., Progress in Scatterometer Application, *Journal of Oceanography*, *58* (2002), Feb., pp. 121-136
[3] Ribal, A., Young, I. R., Calibration and Cross Validation of Global Ocean Wind Speed Based on Scatterometer Observations, *Journal of Atmospheric and Oceanic Technology*, *37* (2020), 2, pp. 279-297
[4] Wang, Z., *et al.*, Inconsistencies in Scatterometer wind Products Based on ASCAT and OSCAT-2 Collocations, *Remote Sensing of Environment*, *225* (2019), May, pp. 207-216
[5] Kumar, S. V. A., *et al.*, Comparative Study of Offshore Winds and Wind Energy Production Derived from Multiple Scatterometers and Met Buoys, *Energy*, *185* (2019), Oct., pp. 599-611
[6] Wang, J., *et al.*, Analysis and Application of Forecasting Models in Wind Power Integration: A Review of Multi-Step-Ahead Wind Speed Forecasting Models, *Renewable and Sustainable Energy Reviews*, *60* (2016), July, pp. 960-981
[7] Senthil, K. P., Improved Prediction of Wind Speed Using Machine Learning, *EAI Endorsed Transactions on Energy Web*, *6* (2019), 157033

[8] Khosravi, A., *et al*., Time-Series Prediction of Wind Speed Using Machine Learning Algorithms: A Case Study Osorio Wind Farm, *Brazil. Applied Energy*, *224* (2018), Aug., pp. 550-566
[9] Luo, X., Short-Term Wind Speed Forecasting Via Stacked Extreme Learning Machine with Generalized Correntropy, *IEEE Transactions on Industrial Informatics*, *14* (2018), 11, pp. 4963-4971
[10] Yu, C., Li, A Novel Framework for Wind Speed Prediction Based on Recurrent Neural Networks and Support Vector Machine, *Energy Conversion and Management*, *178* (2018), Dec., pp. 137-145
[11] Valsaraj, P., *et al*., Spatio-Temporal Independent Applicability of one Time Trained Machine Learning Wind Forecast Models: A Promising Case Study from the Wind Energy Perspective, *International Journal of Sustainable Energy*, *41* (2022), 9, pp. 1164-1182
[12] Hur, S. H., Short-Term Wind Speed Prediction Using Extended Kalman Filter and Machine Learning, *Energy Reports*, *7* (2021), Nov., pp. 1046-1054
[13] Demolli, H., *et al*., Wind Power Forecasting Based on Daily Wind Speed Data Using Machine Learning Algorithms, *Energy Conversion and Management*, *198* (2019), 111823
[14] Li, R., and Jin, Y., A Wind Speed Interval Prediction System Based on Multi-Objective Optimization for Machine Learning Method, *Applied Energy*, *228* (2018), Oct., pp. 2207-2220
[15] Zhu, Q., *et al*., Wind Speed Prediction With Spatio-Temporal Correlation: A Deep Learning Approach, *Energies*, *11* (2018), 705
[16] Li, L., Geographically Weighted Machine Learning and Downscaling for High-Resolution Spatiotemporal Estimations of Wind Speed, *Remote Sensing*, *11* (2019), 1378
[17] Luo, L., *et al*., Design of a Combined Wind Speed Forecasting System Based on Decomposition-Ensemble and Multi-Objective Optimization Approach, *Applied Mathematical Modelling*, *89* (2021), Part 1, pp. 49-72
[18] Jia, Y., *et al*., Traffic Flow Prediction with Rainfall Impact Using a Deep Learning Method, *Journal of Advanced Transportation*, *2017* (2017), ID6575947
[19] Shabani, A., A New Optimization Algorithm Based on Search and Rescue Operations, *Mathematical Problems in Engineering*, *2019* (2019), ID2482543
[20] Shabani, A., *et al*., Search and Rescue Optimization Algorithm: A New Optimization Method for Solving Constrained Engineering Optimization Problems, *Expert Systems with Applications*, *161* (2020), 113698
[21] Trebing, K., Mehrkanoon, S., Wind Speed Prediction Using Multidimensional Convolutional Neural Networks, *Proceeding*, IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, Australia, 2020, pp. 713-720

Paper submitted: July 15, 2024
Paper revised: October 3, 2024
Paper accepted: October 29, 2024