



## Application of four machine-learning methods to predict short-horizon wind energy

Doha Bouabdallaoui<sup>1</sup>, Touria Haidi<sup>1</sup>, Faissal Elmariami<sup>2</sup>, Mounir Derri<sup>1</sup>, El Mehdi Mellouli<sup>3</sup>

1. Laboratory LAGES, Ecole Hassania des Travaux Publics (EHTP), Casablanca 20230, Morocco

2. Laboratory LESE, Ecole Nationale Supérieure d'Electricité et de Mécanique (ENSEM), Casablanca 20100, Morocco

3. Laboratory LESA, School of Applied Sciences, Sidi Mohamed Ben Abdellah University, Fez 30000, Morocco



Scan for more details

**Abstract:** Renewable energy has garnered attention due to the need for sustainable energy sources. Wind power has emerged as an alternative that has contributed to the transition towards cleaner energy. As the importance of wind energy grows, it can be crucial to provide forecasts that optimize its performance potential. Artificial intelligence (AI) methods have risen in prominence due to how well they can handle complicated systems while enhancing the accuracy of prediction. This study explored the area of AI to predict wind-energy production at a wind farm in Yalova, Turkey, using four different AI approaches: support vector machines (SVMs), decision trees, adaptive neuro-fuzzy inference systems (ANFIS) and artificial neural networks (ANNs). Wind speed and direction were considered as essential input parameters, with wind energy as the target parameter, and models are thoroughly evaluated using metrics such as the mean absolute percentage error (MAPE), coefficient of determination ( $R^2$ ), and mean absolute error (MAE). The findings accentuate the superior performance of the SVM, which delivered the lowest MAPE (2.42%), the highest  $R^2$  (0.95), and the lowest MAE (71.21%) compared with actual values, while ANFIS was less effective in this context. The main aim of this comparative analysis was to rank the models to move to the next step in improving the least efficient methods by combining them with optimization algorithms, such as metaheuristic algorithms.

**Keywords:** Wind Energy Prediction; Support Vector Machines; Decision Trees; Adaptive Neuro-Fuzzy Inference Systems; Artificial Neural Networks

## 0 Introduction

Faced with changing energy demand, electricity suppliers are tending to switch to utilizing renewable sources [1-5]. Wind power has become a suitable and promising solution, amidst a myriad of renewable sources available [6-9]. Wind power is increasing in importance. It offers many advantages and is instrumental in the transition to a greener, more ecologically friendly future [10-12].

Received: 10 August 2023/Revised: 17 September 2023/ Accepted: 7 October 2023/ Published: 25 December 2023

✉ Doha Bouabdallaoui  
doha.bouabdallaoui.cedoc@ehtp.ac.ma

Touria Haidi  
touriahaidi60@gmail.com

Faissal Elmariami  
f\_elmariami@yahoo.fr

Mounir Derri  
derrimounir@gmail.com

El Mehdi Mellouli  
mellouli\_elmehdi@hotmail.com

Given its prominence as a reliable source of energy, the ability to predict wind-energy generation is becoming increasingly important to ensure smooth integration into existing distribution networks [13, 14]. Predictive techniques can aid wind-power plant operators in estimating power production variations, enabling better synchronization with alternative energy sources, and improving grid stability [15–18].

Diverse methods such as statistical analysis and artificial intelligence (AI) have been utilized to estimate wind power production. Wind power production has been forecasted by applying a statistical model which relies on both historical and meteorological data. Physical models employ fluid dynamics and atmospheric physics to calculate the motion of winds. As the sophistication of AI increases, neural networks and machine learning are getting more attention. AI-based procedures rely on substantial datasets, meteorological information, and previous wind-energy data to provide more precise forecasts [19–22].

The use of machine-learning approaches to forecast wind generation has been demonstrated in the literature. In [23], a predictive model was designed using artificial neural networks (ANNs). They extracted data from KL University, Andhra Pradesh. The results indicated that increasing the delay reduces errors, while increasing the number of cache layer neurons enhances processing capacity and further decreases predicted output errors. In another study [24] utilized ANN for wind-energy prediction in the Himalayan region. They acquired data consisting of thirty-day inputs of wind velocity, air temperature, and density for predicting wind energy. The ANN proved to be an effective prediction algorithm in this research with a regression coefficient of 0.99. Furthermore, [25] presented a technique using support vector machines (SVMs) to forecast wind turbine generation. The suggested wind power prediction method begins by predicting wind speed, then uses the properties of the turbine generators to predict wind energy. Results indicate the capacity of this approach to be accurate in the very short and short horizon, and to be more efficient than the persistence model. In [26], wind energy was predicted using a mixture of SVM and enhanced dragonfly algorithm. The robustness of this model was validated against actual data collected from the La Haute Borne French wind park. Compared with two other techniques, the proposed model demonstrated superior performance. [27] established a comparison of three methods: ANNs, autoregressive integrated moving average (ARIMA), and adaptive neuro-fuzzy inference system (ANFIS), for predicting wind energy. ARIMA provided more accurate forecasts than the other two methods. The ANFIS yielded the least accurate output. In [28], a hybrid approach of ANFIS and particle

swarm optimization (PSO) was used to predict short-horizon wind energy. With an average value of 5.41%, MAPE performs better than the standard ANFIS and the persistence model, confirming the good performance of the ANFIS. The accuracy of predictions was compared in depth between decision trees and two other regression methods in [29]. The weather variables consisting of wind velocity, wind direction and temperature were ranked according to their dominant influence on wind energy output. The most prominent influence was on the wind speed factor. Regarding the mean absolute error (MAE) and coefficient of determination ( $R^2$ ), the decision tree and nearest-neighbor regression approaches performed more accurately than linear regression techniques. In [30], an exhaustive comparison between ANN, SVM and decision trees was conducted. Results showed that properly tuned ANNs provide accurate predictions of wind-energy production, but the time and processing capacity necessary of finding the optimum hyperparameter configuration is high. The decision-tree model is more transparent than the ANN algorithm, but its performance is inferior.

This study investigated the prediction of wind energy production for an installed wind farm of 54,000 kW in Yalova, Turkey. In this context, four AI techniques were used to compare their performance and determine the most and least accurate models. The selected algorithms were the SVM, decision trees, ANFIS, and ANN. The general structure of the study is shown in Fig. 1.

This study performed a complete comparative assessment of several well-known classical AI methods

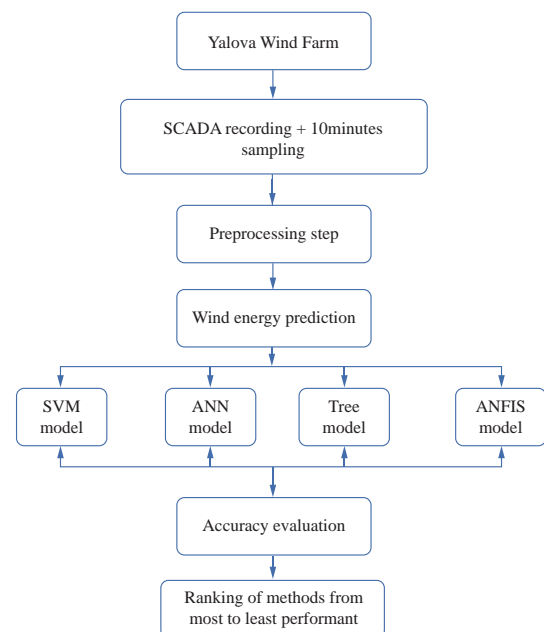


Fig. 1 Global structure of the methodology of the study

addressing the specific task of wind-power production prediction. It goes beyond previous efforts in the literature, which often focused on single methods or advanced methods combined. To this end, this study helps to give an overall concept of the performance of these standard prediction AI methods, and then to classify them, so enabling us to work afterwards on the worst-performing method, trying to improve it by one or more optimization methods.

## 1 Methods

The wind farm in Yalova was commissioned in 2016. It has the following coordinates: latitude:  $40^{\circ} 34' 58.3''$  and longitude:  $28^{\circ} 56' 3''$ . Figure 2, generated by Google Maps, shows the turbines marked in red [31]. An integrated SCADA monitoring system was installed and recorded data for one of the wind generators. The database used can be freely accessed via [32]. It is an open-source dataset which contains 5 data: time, wind power (kW), wind speed (m/s), wind direction ( $^{\circ}$ ), and wind energy (kWh). The record of the data runs from January 1 to December 31, 2018.



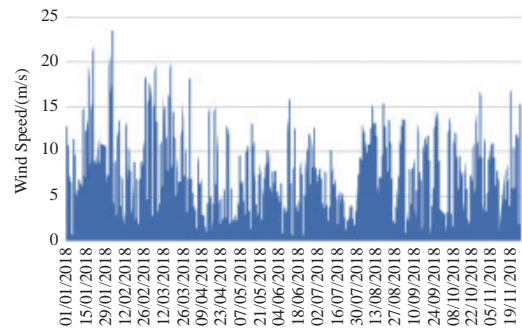
**Fig. 2** Cartographic presentation of the Yalova wind power plant generated by Google Maps

To carry out wind power prediction for Yalova Turkish wind park, SVM, decision trees, ANFIS, and ANN methods are implemented. Specific pre-processing phases were required for each method to guarantee the compatibility and quality of the wind power dataset, notably scalability and normalization, and the handling of missing values and categorical variables.

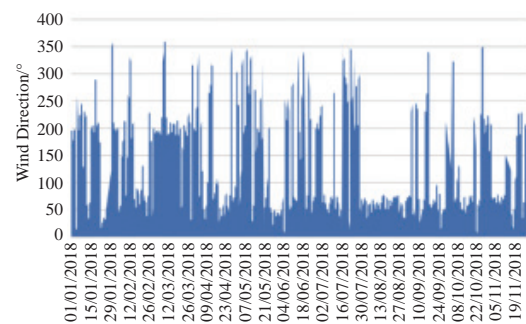
Yalova's wind power database was based on two inputs: wind direction and wind speed, as well as the wind power as a target parameter. A total of 45,507 data samples were utilized for the training phase, which extended from 00:00 on January 1, 2018, to 23:50 on November 26, 2018, as presented in Fig. 3 and Fig. 4.

As for the test phase, 144 samples were used, ranging from 00:00 to 23:50 on November 27, 2018. Prominently, the timescale of the database was defined as 10 min.

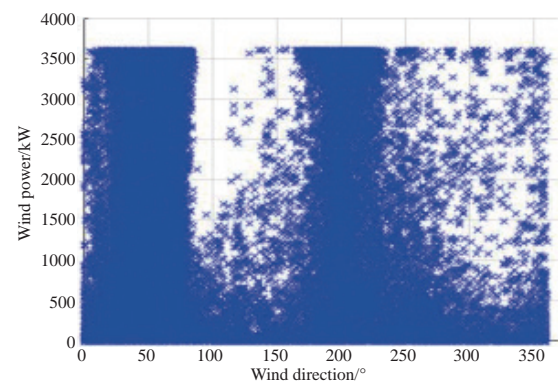
The distribution of the velocity and the angle of wind against wind-energy output is shown in Fig. 5 and Fig. 6,



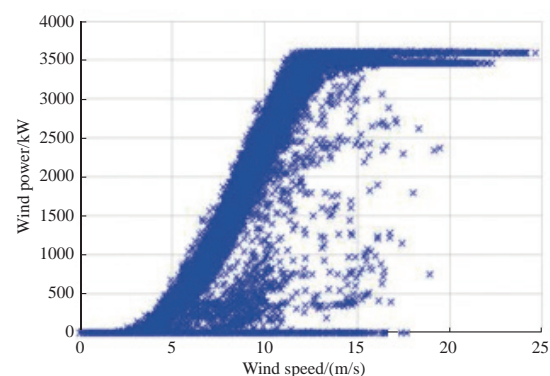
**Fig. 3** Wind speed data input (m/s) for January 1st, 2018 to November 26th, 2018



**Fig. 4** Wind direction data input (m/s) for January 1st, 2018 to November 26th, 2018



**Fig. 5** Wind direction against wind power



**Fig. 6** Wind speed against wind power

respectively. These graphs facilitated data exploration, model selection, reporting of results, and evaluation of their quality. The following Table 1 lists the average and standardized deviation, the minimal and maximal amounts, and the confidence interval for each variable used.

**Table 1** Information about the dataset

|                | Mean   | Standard deviation | Min | Max    | Confidence interval |
|----------------|--------|--------------------|-----|--------|---------------------|
| Wind speed     | 7.50   | 4.17               | 0   | 25.20  | [−0.66 15.66]       |
| Wind direction | 122.57 | 93.23              | 0   | 360    | [−60,14 305,30]     |
| Wind power     | 1562.4 | 1123.3             | 0   | 3618.7 | [−639,26 3764,22]   |

### 1.1 Artificial Neural Networks (ANNs)

ANNs are a category of computing system. Typically, ANNs are formed by neurons interconnected in layers [33], [34]. On one side, data are captured in the input layer, while on the other, the desired prediction is delivered via the output layer. Between these layers, the hidden layers enable the system to evolve and extract features from the input signal. There are three widely employed functions for training data: Levenberg-Marquardt (LM), scaled conjugate gradient (SCG), and Bayesian regularization (BR) [35, 36]. A general flowchart of the ANN design is shown in Fig. 7.

The selected function was the LM algorithm, which has been identified as an effective algorithm for midsize systems. Two independent inputs were included in the model: wind velocity and direction, and the model output

was the wind energy produced. During the modeling process, the number of hidden neurons was increased to 50 to round out the nonlinear ANN architecture. The details of the main model settings are listed in Table 2.

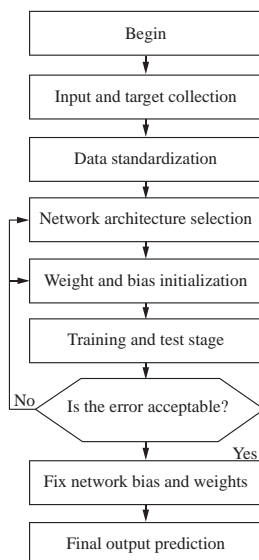
**Table 2** Principal parameters of the ANN model

| Parameter                     | Description   | Value/setting  |
|-------------------------------|---|--|
| Hidden Layer Size             | Number of units in the hidden layers  | 50   |
| Training Function             | Algorithm of training the neural network  | Trainlm  |
| Number of Training Epochs     | Number of training iterations   | 190  |
| Activation function           | Enable nonlinearity in networks   | for the first hidden layer: Sigmoid<br>For the output layer: Purelin |
| Pre/Post-Processing Functions | Removing rows with constant values and then normalizing the remaining data to a specific range. | Remove constant rows<br>Map minmax                                   |
| Train Ratio                   | Ratio of data allocated for training.   | 70%= 31859   |
| Validation Ratio              | Ratio of data allocated for validation  | 15%=6827   |
| Test Ratio                    | Ratio of data allocated for testing   | 15%=6827   |

### 1.2 Adaptive Neuro-Fuzzy Inference System (ANFIS)

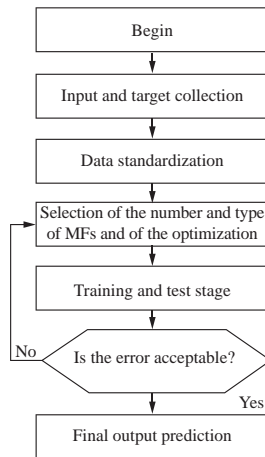
By combining the advantages of fuzzy logic and neural networks, ANFIS provides an efficient hybrid approach for handling complicated forecasting applications. ANFIS involves a two-phase procedure. Initially, fuzzy logic is employed to create a collection of fuzzy rules derived from existing knowledge and training data. These rules establish relationships linking the input and output variables to be predicted. Subsequently, neural networks have been utilized to optimize the settings of the fuzzy rules [37, 38]. This adaptive process enables ANFIS to learn directly from the data and sharpen its prediction ability. By incorporating fuzzy logic and neural networks, ANFIS provides a prediction approach that is versatile and responsive, capable of capturing nonlinear features and addressing the uncertainties inherent in real-time prediction challenges [39, 40]. A general flowchart of the ANFIS model is shown in Fig. 8.

In this case, nine membership functions and two inputs were used. This implies that nine fuzzy rules existed. A



**Fig. 7** General flow chart of an ANN





**Fig. 8 General flow chart of ANFIS**

hybrid training algorithm that associates back-propagation and least-squares methods was used to calculate the membership function parameters. In this configuration, ANFIS uses the Sugeno fuzzy inference system, which employs if-then fuzzy rules with linear consequences. The processing method for antecedent fuzzy sets was “prod,” and “probor” for consequent fuzzy sets. The method employed for defuzzification was “wtaver.” The implication method was “prod,” and the aggregation method was “sum.” Further details are presented in Table 3.

**Table 3 Principal parameters of the ANFIS model**

| Parameter              | Description   | Value/setting             |
|------------------------|---|---------------------------|
| Type                   | Type of ANFIS model   | Sugeno                    |
| And Method             | T-norm method for AND operation                                   | Prod                      |
| Or Method              | T-norm method for OR operation                                    | Probor                    |
| Defuzzification Method | Method for defuzzifying the output                                | Wtaver                    |
| Implication Method     | Implication method  | Prod                      |
| Aggregation Method     | Aggregation method for fuzzy reasoning                            | Sum                       |
| Partitioning Method    | Method used for input space partitioning                          | Grid (Grid Partitioning)  |
| Optimization Method    | Method used for fuzzy inference system (FIS) training             | Hybrid                    |
| Number of Rules        | Total number of rules in the FIS.                                 | 9                         |
| 'in1mf1'               | Membership Functions input 1 (wind speed): Gaussian MF parameters | [2.14, -3.78, 1.71, 8.44] |

continue

| Parameter | Description   | Value/setting                   |
|-----------|---|---------------------------------|
| 'in1mf2'  | Membership Functions input 1 (wind speed): Gaussian MF parameters     | [1.45, 10.23, 2.29, 16.54]      |
| 'in1mf3'  | Membership Functions input 1 (wind speed): Gaussian MF parameters     | [2.926, 20.99, 2.14, 28.99]     |
| 'in2mf1'  | Membership Functions input 2 (wind direction): Gaussian MF parameters | [30.57, -53.99, 30.54, 53.99]   |
| 'in2mf2'  | Membership Functions input 2 (wind direction): Gaussian MF parameters | [30.56, 126.00, 30.612, 234.02] |
| 'in2mf3'  | Membership Functions input 2 (wind direction): Gaussian MF parameters | [30.53, 306.03, 30.57, 413.99]  |
| 'out1mf1' | Membership Functions output (wind power): Constant MF value           | 630.84                          |
| 'out1mf2' | Membership Functions output (wind power): Constant MF value           | 714.79                          |
| 'out1mf3' | Membership Functions output (wind power): Constant MF value           | 904.97                          |
| 'out1mf4' | Membership Functions output (wind power): Constant MF value           | 3521.60                         |
| 'out1mf5' | Membership Functions output (wind power): Constant MF value           | 3497.843                        |
| 'out1mf6' | Membership Functions output (wind power): Constant MF value           | 3334.97                         |
| 'out1mf7' | Membership Functions output (wind power): Constant MF value           | 2274.46                         |
| 'out1mf8' | Membership Functions output (wind power): Constant MF value           | 3574.62                         |
| 'out1mf9' | Membership Functions output (wind power): Constant MF value           | 4531.71                         |

The general expression for a Gaussian membership function (gauss2mf) used in the ANFIS is

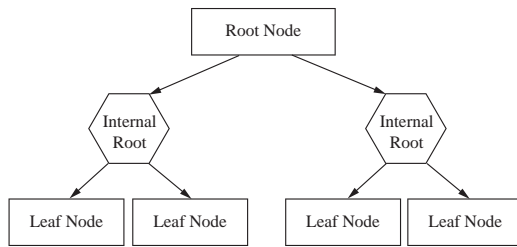
$$\mu(x) = \exp\left(-\left(\frac{x-c}{\sigma}\right)^2\right) \quad (1)$$

where  $\mu(x)$  is the value of membership defined for a given input value  $x$ ,  $c$  is the MF Gaussian center, identifying the

function's peak or center, and  $\sigma$  is the standard deviation, controlling the function's width or range.

### 1.3 Decision Trees

Decision trees are effective forecasting methods [41]. The process consists of recursively splitting data according to attributes or characteristics, thus creating the decision tree structure presented in Fig. 9. A specific characteristic is assessed at every node of the tree, and the data are partitioned into subgroups depending on the value of the characteristic. This process continues until certain cutoff criteria are reached, such as the maximum value or minimum count of data points in the assembly. The final nodes, called the leaf nodes, contain predictions or outcomes. Decision trees can be interpreted, process both numerical and categorical data, and detect nonlinear links between variables [42, 43].



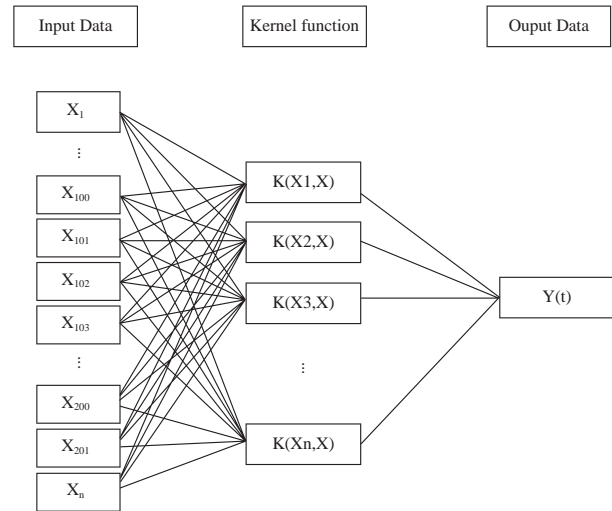
**Fig. 9** General structure of a decision tree

To optimize the efficiency and precision, the coarse tree model employed to predict wind power was parameterized with specific settings. By requiring a minimum of 72 occurrences in the parent node and 36 occurrences in each leaf node, the model guaranteed that the nodes had sufficient data for reliable splitting while avoiding overfitting. The number of splits permitted was set to a maximum of 45,512, meaning that the model could create a complex tree structure to understand the delicate relationships within the wind power data. Pruning was activated to eliminate unnecessary nodes and improve the generalizability of the model. Moreover, the tree technique considered up to ten different categories of class predictors, offering flexibility in the processing of different wind energy features. As predictor selection had to be set to "All Splits," all predictors available were utilized for each node during the tree-building process. Generally, the coarse tree model configuration aimed to balance complexity and generalization, enabling it to predict wind power production efficiently, precisely, and consistently.

### 1.4 Support Vector Machines (SVMs)

SVMs are widespread machine-learning algorithms

designed for predictive applications in a wide range of fields. SVMs operate according to an optimal classification or fitting function that best fits data [44, 45]. SVMs can handle both small and large-scale datasets and operate efficiently in high-dimensional spaces [46, 47]. The basic SVM architecture is shown in Fig. 10.



**Fig. 10** General structure of SVM

The SVM configuration in this study comprised specific parameters and settings. The Gaussian kernel (RBF) is a widely adopted method for capturing complex relationships between data points, and was selected in this case. Epsilon is a hyperparameter that specifies the margin size around the support vectors and has a value of 132.65. The kernel scale determined the impact of each support vector on the boundaries defined by the decision; in this case, it was set to 1.4. The details of the SVM system parameters are listed in Table 4.

**Table 4** Principal parameters of the SVM model

| Parameter          | Description   | Value/setting      |
|--------------------|---|--------------------|
| Sigma ( $\sigma$ ) | Kernel width in the Gaussian (RBF) kernel function                    | [0.1653<br>0.2590] |
| Mu ( $\mu$ )       | Means of the input features, wind direction, and wind speed           | [0.2976<br>0.3405] |
| Bias               | Intercept added to the SVM decision function                          | 1,1099e+03         |
| Alpha or (C)       | Regularization parameter in your SVM model                            | 1.3266e+03         |
| W                  | Weight associated with the support vectors in the SVM model           | 2.1972e-05         |
| SMO                | Sequential minimal optimization (SMO) algorithm used to train the SVM | 1                  |

continue

| Parameter              | Description  | Value/setting |
|------------------------|--|---------------|
| NumIterations          | Number of iterations or steps taken by the SMO algorithm during the training process   | 16911         |
| Epsilon ( $\epsilon$ ) | Parameter used in the SVM model to control the margin of tolerance for support vector identification.                                | 132.6593      |
| Kernel Function        | Gaussian (RBF) kernel commonly used for SVM regression   | Gaussian      |
| Kernel scale           | This parameter controlled the scale or width of the Gaussian kernel  | 1.4           |
| Standardization        | Input features were scaled to have zero mean and unit variance, which is a common preprocessing step in many machine-learning models | 1             |

## 2 Prediction accuracy evaluation criteria

Various measures can be used to assess the efficiency of wind-energy forecasting models. The accuracy of the proposed approach is evaluated using various criteria.

### 2.1 Mean Absolute Percentage Error (MAPE)

This parameter calculates the forecast deviation as a percentage of the actual or target value [48].

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|P_{\text{predicted}} - P_{\text{measured}}|}{|P_{\text{nominal}}|} \quad (2)$$

### 2.2 Determination coefficient ( $R^2$ )

This corresponds to the ratio of the dependent variable variant, which is expressed in the regression formula using one or several independently defined parameters. It can be specified as a value from 0 to 1 or as a percentage [49].

$$R^2 = 1 - \sum_{i=1}^N \frac{(P_{\text{measured}} - P_{\text{predicted}})^2}{(P_{\text{measured}} - P_{\text{empirical mean}})^2} \quad (3)$$

### 2.3 Mean Absolute Error (MAE)

This parameter expresses the average deviation of the predicted values from the target values [50].

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |P_{\text{predicted}} - P_{\text{measured}}| \quad (4)$$

### 2.4 Bias (B)

This parameter is a measure of the systematical error of the forecast model.

$$\text{Bias}(B) = \sum \frac{P_{\text{predicted}} - P_{\text{measured}}}{N} \quad (5)$$

## 3 Results and discussion

Table 5 summarizes the effectiveness of the four forecasting methods using different criteria. The findings presented in Table 6 show the extent to which the SVM system provided more accurate forecasts.

The use of the SVM in the short horizon to predict wind energy achieved relevant feedback, as shown in Fig. 11. The SVM exhibited outstanding performance with a MAPE of 2.42%, indicating its remarkable accuracy in forecasting wind power production. The  $R^2$  of 0.95, as presented in Fig. 12, demonstrated a close correlation between the predicted and observed results. Moreover, compared with the actual values, an MAE of 71.21 demonstrated that the model could achieve a mean prediction deviation of 71.21 units. These convincing results underlined the performance of the SVM method in providing reliable wind-energy forecasts in both the short and very short terms.

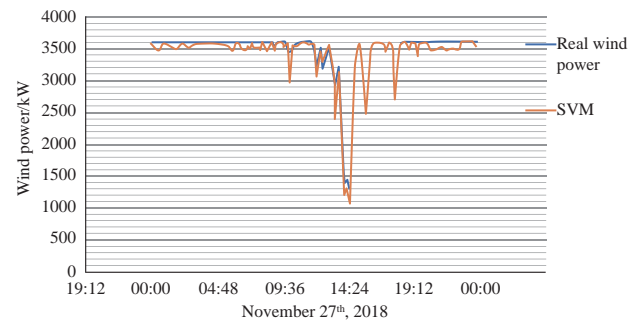


Fig. 11 SVM tracing of predicted and true data on November 27, 2018

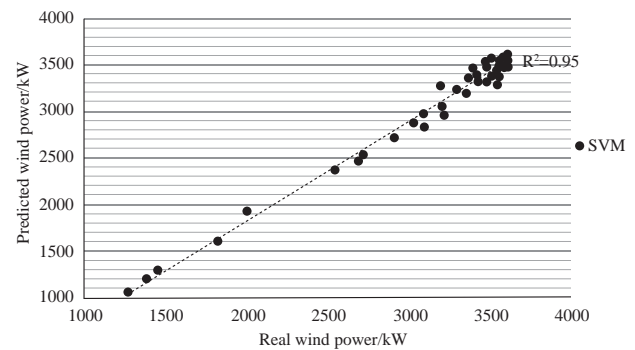
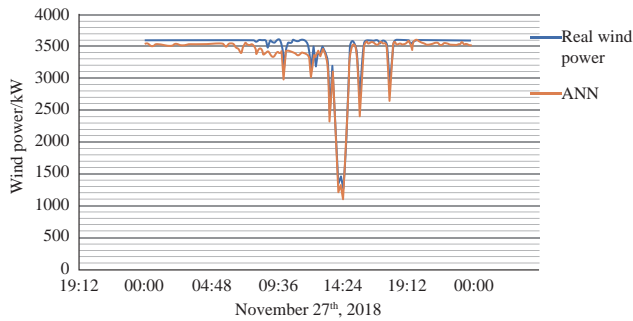


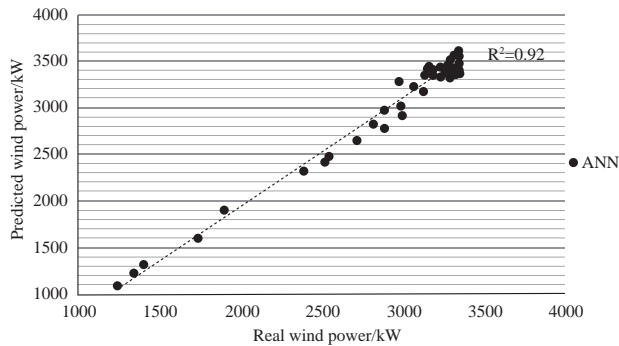
Fig. 12 Scatter plot of real versus forecast data using the SVM

The ANN results for wind-energy prediction revealed considerable promise, as shown in Fig. 13. The model had a MAPE of 3.13%, indicating good accuracy in forecasting wind power production. The  $R^2$  of 0.92, as shown in Fig. 14, indicated a high correspondence between the expected and true values, suggesting that the ANN model effectively captured the underlying patterns of the wind-

energy data. Additionally, the MAE of 96.25 signified that the model's predictions, on average, deviated from the actual values by 96.25 units, which is comparatively modest for the scale of the wind power measures. These results indicated the ability of the ANN model to provide confident predictions of wind power over the short and very short terms, considering that this accuracy can be enhanced by running the model using a large database with multiple data inputs.

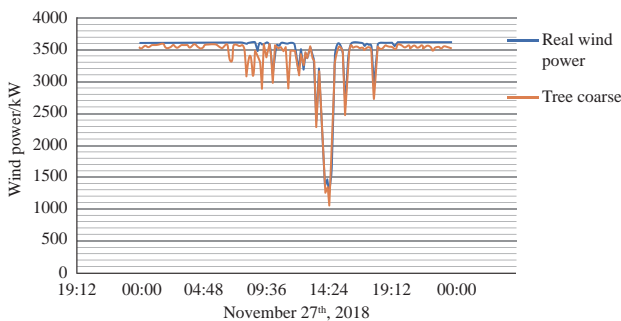


**Fig. 13** ANN tracing of predicted and true data on November 27, 2018



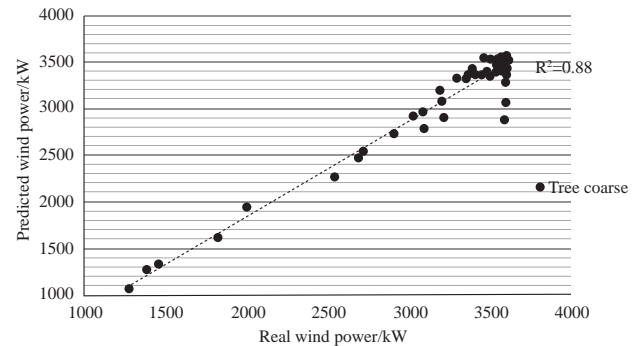
**Fig. 14** Scatter plot of real versus forecast data using the ANN

The use of coarse trees to predict wind energy yielded notable results, as shown in Fig. 15. The coarse tree model had a MAPE of 3.07%, which indicates a reasonable level of forecasting precision for wind energy output. An  $R^2$  of 0.88 (Fig. 16) revealed a modest correlation between the forecast and true values. Meanwhile, the MAE of 92.38 signified that the model's predictions deviated from the



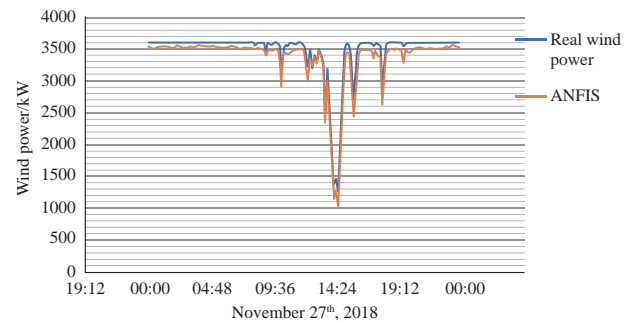
**Fig. 15** Coarse tree tracing of predicted and true data on November 27, 2018

actual values by an average of 92.38 units. These results demonstrate that the coarse-tree model is a viable option for short-horizon wind-energy forecasting.

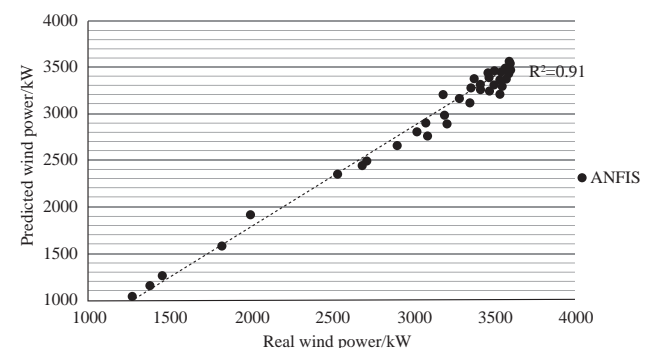


**Fig. 16** Scatter plot of real versus forecast data using the coarse tree

The ANFIS method for wind-energy prediction generated noteworthy results, as shown in Fig. 17. ANFIS achieved a MAPE of 3.53%, reflecting a reasonable precision level in terms of wind power production forecasting. An  $R^2$  of 0.91 (Fig. 18) indicated a close correlation between the predicted and observed values, which means that the ANFIS model captured a considerable amount of the variation present in the wind power data. Furthermore, an MAE of 106.91 meant that, on average,



**Fig. 17** ANFIS tracing of predicted and true data on November 27, 2018



**Fig. 18** Scatter plot of real versus forecast data using ANFIS



the ANFIS model predictions deviated from the actual values by 106.91 units. Overall, these results confirmed the suitability of the ANFIS model for wind-energy prediction.

By comparing the performance metrics of each method, the strongest  $R^2$  value was for the SVM model (0.95), reflecting the greatest correlation and matching with real data, whereas the decision tree model had the smallest  $R^2$  ratio (0.88), which indicated a lower fit. In addition, the SVM model provided the weakest MAPE (2.42%), revealing the smallest deviation percentage compared with the real values overall. In addition, the lowest MAE was achieved by the SVM model (71.21), representing the smallest deviation, on average, relative to the true measured values. In terms of bias, the SVM model yielded the smallest bias (−66.77), which indicated the least underestimation of wind power output consistently. In summary, based on the provided measures, the SVM model was the most successful of the four models, with the greatest  $R^2$ , minimum MAPE, least MAE, and relatively low bias. Its predictive capability was considerable and correlated well with real data.

The different accuracy levels of the wind-energy prediction results were owing to the intrinsic features and capacities of the models used. The inherent qualities of the SVM and ANN models achieve more accurate wind-energy forecasts. In particular, SVMs can handle complex nonlinear data relationships and dynamics, enabling them to effectively understand complex patterns. Conversely, the ability of ANN models to recognize and learn complicated relationships within data is well-known and can be used to uncover patterns and provide precise predictions. However, the less precise results using coarse trees and ANFIS may reflect their limitations in capturing complex relationships and processing large datasets. Fig. 19 presents a histogram of the prediction errors for the four methods used.

This histogram shows that bins ranging from −300 to 0 corresponded to prediction errors in which the model underestimated the target variable, whereas those ranging from 0 to 100 corresponded to prediction errors in which the model overestimated the target variable. The peak at approximately zero indicated that the SVM model had a relatively low bias. This meant that the predictions of the model, on average, were not heavily biased towards over- or under-estimation of the target.

These classical AI methods can perform the prediction task; however, to achieve higher accuracy, a richer database with more inputs would be required and regularization methods should be used to prevent overfitting; moreover, optimization algorithms should be used either as training functions or to optimize the choice of kernel parameters for the models.

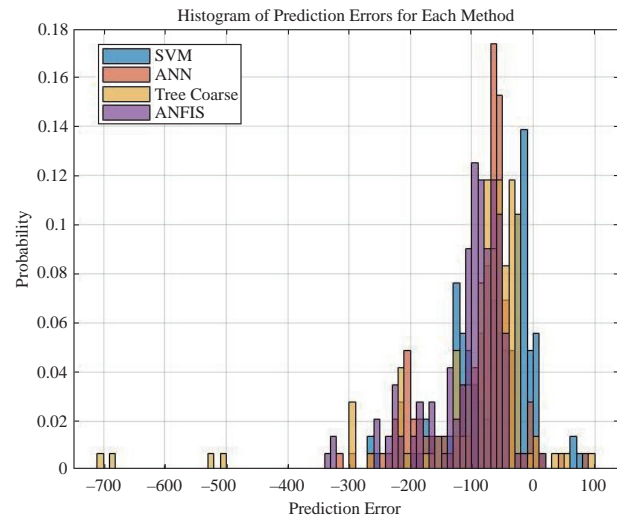


Fig. 19 Histogram of prediction error for the four AI methods used

Table 5 Evaluating model training Performance based on the criteria

| Criterion | ANN    | ANFIS   | Tree decision | SVM    |
|-----------|--------|---------|---------------|--------|
| MAPE      | 3.13   | 3.53    | 3.07          | 2.42   |
| $R^2$     | 0.92   | 0.91    | 0.88          | 0.95   |
| MAE       | 96.25  | 106.21  | 92.38         | 71.21  |
| Bias      | −94.06 | −105.98 | −88.81        | −66.77 |

Table 6 Prediction outputs per hour on November 27, 2018, using the four methods (in kW)

| Date/Time | Real wind power | SVM Output | ANN Output | Tree coarse Output | ANFIS Output |
|-----------|-----------------|------------|------------|--------------------|--------------|
| 00:50     | 3600.805        | 3576.519   | 3521.398   | 3564.180           | 3548.437     |
| 01:50     | 3601.192        | 3489.656   | 3510.045   | 3579.162           | 3524.705     |
| 02:50     | 3601.165        | 3505.474   | 3520.929   | 3545.729           | 3527.886     |
| 03:50     | 3599.978        | 3586.432   | 3540.952   | 3524.950           | 3552.489     |
| 04:50     | 3601.034        | 3586.130   | 3548.233   | 3524.950           | 3556.686     |
| 05:50     | 3601.955        | 3486.206   | 3510.481   | 3566.003           | 3523.576     |
| 06:50     | 3602.585        | 3476.050   | 3463.912   | 3385.166           | 3513.202     |
| 07:50     | 3602.939        | 3514.713   | 3537.613   | 3566.003           | 3529.222     |
| 08:50     | 3599.251        | 3573.373   | 3398.714   | 3473.946           | 3498.994     |
| 09:50     | 3601.128        | 3588.050   | 3388.041   | 3556.343           | 3485.233     |
| 10:50     | 3600.606        | 3571.052   | 3413.540   | 3520.262           | 3464.727     |
| 11:50     | 3581.691        | 3584.343   | 3397.483   | 3473.946           | 3494.710     |
| 12:50     | 3364.996        | 3356.665   | 3355.000   | 3361.936           | 3283.969     |
| 13:50     | 2717.902        | 2538.093   | 2486.597   | 2557.019           | 2495.220     |

continue

| Date/Time | Real<br>wind<br>power | SVM<br>Output | ANN<br>Output | Tree<br>coarse<br>Output | ANFIS<br>Output |
|-----------|-----------------------|---------------|---------------|--------------------------|-----------------|
| 14:50     | 3026.644              | 2874.086      | 2823.200      | 2936.908                 | 2805.878        |
| 15:50     | 3093.568              | 2831.190      | 2779.129      | 2798.683                 | 2762.725        |
| 16:50     | 3590.890              | 3592.018      | 3536.084      | 3528.198                 | 3493.279        |
| 17:50     | 2908.654              | 2716.211      | 2660.837      | 2738.108                 | 2658.477        |
| 18:50     | 3603.378              | 3473.897      | 3532.136      | 3556.357                 | 3513.505        |
| 19:50     | 3602.225              | 3567.426      | 3609.934      | 3569.551                 | 3455.782        |
| 20:50     | 3602.958              | 3502.704      | 3531.662      | 3546.132                 | 3524.929        |
| 21:50     | 3602.929              | 3485.215      | 3548.800      | 3558.892                 | 3511.782        |
| 22:50     | 3602.520              | 3601.406      | 3538.916      | 3545.418                 | 3543.773        |
| 23:50     | 3601.716              | 3561.148      | 3530.554      | 3519.559                 | 3534.572        |

## 4 Conclusion

The shift toward sustainable energy sources has made wind energy a highly desirable alternative. Precise wind generation prediction is the key to optimizing its deployment and integration into the existing power infrastructure. In this study, a wind-energy production forecast for a wind park in Yalova, Turkey, was conducted using artificial intelligence methods. Four AI approaches were used: SVM, decision trees, ANFIS, and ANN. Simulations using the four techniques revealed that the SVM and ANN were accurate predictors for short-term wind power forecasts, recording MAPE=2.42 and  $R^2=0.95$  for SVMs and MAPE=3.13 and  $R^2=0.92$  for the ANN. Nevertheless, coarse trees and adaptive neuro-fuzzy inference systems were relatively less precise, resulting in MAPE=3.07 and  $R^2=0.88$  for coarse trees and MAPE=3.53 and  $R^2=0.91$  for ANFIS. This study aimed to explore and refine forecast models to enhance the precision and robustness of wind-energy predictions. Combining these models with optimization algorithms such as metaheuristic algorithms can generate a more accurate configuration and may be an effective solution for improving the precision of predictive machine-learning methods.

## Declaration of Competing Interest

We declare that we have no conflict of interest.

## References

- [1] Aziz M (2019) Advanced green technologies toward future sustainable energy systems. *Indonesian Journal of Science And*

*Technology*, 4(1): 89

- [2] Haidi T, Cheddadi B (2022) State of wind energy in the world: evolution, impacts and perspectives. *International Journal on Technical and Physical Problems of Engineering*, 14(2): 6
- [3] Lu P, Ye L, Zhao Y, et al. (2021) Review of meta-heuristic algorithms for wind power prediction: Methodologies, applications and challenges. *Applied Energy*, 301: 117446
- [4] Yin R, Li D, Wang Y, et al. (2020) Forecasting method of monthly wind power generation based on climate model and long short-term memory neural network. *Global Energy Interconnection*, 3(6): 571-576
- [5] Malakouti S M (2023) Estimating the output power and wind speed with ML methods: A case study in Texas. *Case Studies in Chemical and Environmental Engineering*, 7: 100324
- [6] Jaadi M El, Haidi T, Bouabdallaoui D (2023) Particle swarm optimization for the optimal layout of wind turbines inside a wind farm. *IAES International Journal of Artificial Intelligence*, 12(3): 1260-1269
- [7] Xiao Y, Zou C, Chi H, et al. (2023) Boosted GRU model for short-term forecasting of wind power with feature-weighted principal component analysis. *Energy*, 267: 126503
- [8] Haidi T, Cheddadi B (2022) Wind energy integration in Africa: development, impacts and barriers. *International Journal of Electrical and Computer Engineering*, 12 (5): 4614
- [9] Jorgensen K L, Shaker H R (2020) Wind power forecasting using machine learning: state of the art, trends and challenges. 8th *International Conference on Smart Energy Grid Engineering*, pp. 44-50
- [10] Abdullah A G, Setiorini A H, Dwitasari N A, et al. (2021) Location suitability analysis for wind farm exploitation using fuzzy analytic hierarchy process. *Indonesian Journal of Science and Technology*, 6(3): 523-534
- [11] Haidi T, Cheddadi B, Mariami F El, et al. (2021) Wind energy development in Morocco: Evolution and impacts. *International Journal of Electrical and Computer Engineering*, 11(4): 2811
- [12] Tumse S, Bilgili M, Sekertekin A, et al. (2023) Comparison and evaluation of machine learning approaches for estimating heat index map in Türkiye. *Neural Computing and Applications*, 35: 1-22
- [13] Idrissi Z E, Haidi T, Elmariami F, et al. (2022) Optimal coordination of docrs using ga with integration of dgs in distribution networks. *International Journal on Technical and Physical Problems of Engineering*, 14(1): 8
- [14] Zheng J, Du J, Wang B, et al. (2023) A hybrid framework for forecasting power generation of multiple renewable energy sources. *Renewable and Sustainable Energy Reviews*, 172: 113046
- [15] Bouabdallaoui D, Haidi T, Jaadi M El (2023) Review of current artificial intelligence methods and metaheuristic algorithms for wind power prediction. *Indonesian Journal of Electrical Engineering and Computer Science*, 29(2): 626
- [16] Atuahene S, Bao Y, Ziggah Y, et al. (2018) Short-term electric power forecasting using dual-stage hierarchical wavelet- particle swarm optimization- adaptive neuro-fuzzy inference system pso-

- anfis approach based on climate change. *Energies*, 11(10): 2822
- [17] Talaat M, Elkholy M, Alblawi A, et al. (2023) Artificial intelligence applications for microgrids integration and management of hybrid renewable energy sources. *Artificial Intelligence Review*, 56(9): 10557-10611
- [18] Bilal B, Adjallah K H, Sava A, et al. (2023) Wind turbine output power prediction and optimization based on a novel adaptive neuro-fuzzy inference system with the moving window. *Energy*, 263: 126159
- [19] Cali U, Sharma V (2019) Short-term wind power forecasting using long-short term memory based recurrent neural network model and variable selection. *International Journal of Smart Grid Clean Energy*, 103-110
- [20] Zedak C, Belfqih A, Berdai A, et al. (2018) Forecasting electrical demand in Morocco using the multiple linear regression and artificial neural networks. *SSRN Electronic Journal*
- [21] Wang H, Lei Z, Zhang X, et al. (2019) A review of deep learning for renewable energy forecasting. *Energy Conversion and Management*, 198: 111799
- [22] Benti N E, Chaka M D, Semie A G (2023) Forecasting renewable energy generation with machine learning and deep learning: current advances and future prospects. *Sustainability*, 15(9): 9
- [23] Jyothi M N, Dinakar V, Teja N S S R, et al. (2015) NARX based short term wind power forecasting model. *Telkomnika Indonesian Journal of Electrical Engineering*, 15(1): 20-25
- [24] Puri V, Kumar N (2022) Wind energy forecasting using artificial neural network in himalayan region. *Modeling Earth Systems and Environment*, 8(1): 59-68
- [25] Zeng J, Qiao W (2011) Support vector machine-based short-term wind power forecasting. *Power Systems Conference and Exposition: IEEE*, pp. 1-8
- [26] Li L L, Zhao X, Tseng M L, et al. (2020) Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. *Journal of Cleaner Production*, 242: 118447
- [27] Chen Q, Folly K A (2018) Comparison of three methods for short-term wind power forecasting. *International Joint Conference on Neural Networks*, pp. 1-8
- [28] Pousinho H M I, Mendes V M F, Catalão J P S (2011) A hybrid PSO-ANFIS approach for short-term wind power prediction in Portugal. *Energy Conversion Management*, 52(1): 397-402
- [29] Eyecioglu O, Hangun B, Kayisli K, et al. (2019) Performance comparison of different machine learning algorithms on the prediction of wind turbine power generation. *International Conference on Renewable Energy Research and Applications*, pp. 922-926
- [30] Buturache A N, Stancu S (2021) Wind energy prediction using machine learning. *Low Carbon Economy*, 12(01): 1-21
- [31] Cartographic presentation of the Yalova wind power plant generated by google Maps. <https://www.google.com/maps/@40.5693315,28.9658398,784a,35y,38.99t/data=!3m1!1e3?entry=ttu> (accessed Sep. 12, 2023)
- [32] Wind turbine scada dataset, <https://www.kaggle.com/datasets/berkerisen/wind-turbine-scada-dataset> (accessed Sep 16, 2023)
- [33] Brahmi A E, Abderafi S, Ellaia R (2021) Artificial neural network analysis of sulfide production in a Moroccan sewerage network. *Indonesian Journal of Science and Technology*, 6(1): 193-204
- [34] Rodríguez F, Florez-Tapia A M, Fontán L, et al. (2020) Very short-term wind power density forecasting through artificial neural networks for microgrid control. *Renewable Energy*, 145: 1517-1527
- [35] Nair K R (2017) Forecasting of wind speed using ANN, ARIMA and Hybrid Models
- [36] Lu W, Duan J, Wang P, et al. (2023) Short-term wind power forecasting using the hybrid model of improved variational mode decomposition and maximum mixture correntropy long short-term memory neural network. *International Journal of Electrical Power & Energy Systems*, 144: 108552
- [37] Xing Y, Lien F S, Melek W, et al. (2022) A multi-hour ahead wind power forecasting system based on a WRF-TOPSIS-ANFIS model. *Energies*, 15(15): 5472
- [38] Tümse S, İlhan A, Bilgili M, et al. (2022) Estimation of wind turbine output power using soft computing models. *Energy Sources Part A: Recovery Utilization and Environmental Effects*, 44(2): 3757-3786
- [39] Eseye A T, Zhang J, Zheng D, et al. (2017) Short-term wind power forecasting using a double-stage hierarchical hybrid GA-ANFIS approach. *International Conference on Cloud Computing and Big Data Analysis*, pp. 499-503
- [40] Çakır S (2023) Renewable energy generation forecasting in Turkey via intuitionistic fuzzy time series approach. *Renewable Energy*, 214: 194-200
- [41] Chaudhary A, Sharma A, Kumar A, et al. (2020) Short term wind power forecasting using machine learning techniques. *Journal of Statistics and Management Systems*, 23(1): 145-156
- [42] Ahmadi A, Nabipour M, Mohammadi-Ivatloo B, et al. (2020) Long-term wind power forecasting using tree-based learning algorithms. *IEEE Access*, 8: 151511-151522
- [43] Pathak R, Wadhwa A, Khetarpal P, et al. (2023) Comparative assessment of regression techniques for wind power forecasting. *IETE Journal of Research*, 69(3): 1393-1402
- [44] Zhang L, Wang K, Lin W, et al. (2019) Wind power prediction based on improved genetic algorithm and support vector machine. *IOP Conference Series: Earth and Environmental Science*, 252: 032052
- [45] Li GM, Yu LY, Zhang Y, et al. (2023) An integrated method with adaptive decomposition and machine learning for renewable energy power generation forecasting. *Environmental Science and Pollution Research*, 30(14): 41937-41953
- [46] Khosravi A, Koury R N N, Machado L, et al. (2018) Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system. *Sustain. Sustainable Energy Technologies and Assessments*, 25: 146-160
- [47] Alkabbani H, Hourfar F, Ahmadian A, et al. (2023) Machine learning-based time series modelling for large-scale regional wind power forecasting: a case study in ontario Canada. *Cleaner Energy Systems*, 5: 100068

- [48] Masrur H, Nimol M, Faisal M, et al. (2016) Short term wind speed forecasting using Artificial Neural Network: A case study. International Conference on Innovations in Science, Engineering and Technology, pp. 1-5
- [49] Fentis A, Bahatti L, Mestari M, et al. (2017) Short-term solar power forecasting using Support Vector Regression and feed-forward NN. International New Circuits and Systems Conference, 405-408
- [50] Maldonado-Correa J, Valdiviezo-Condolo M, Viñan-Ludeña M S, et al. (2021) Wind power forecasting for the Villonaco wind farm. Wind Engineering, 45(5): 1145-1159

## Biographies



**Doha Bouabdallaoui** holds an electrical engineering degree from the Ecole Nationale des Sciences Appliquées de l'Université Mohammed I. She is a PhD student at the Ecole Hassania des Travaux Publics (EHTP) in the Automation, Electrical Systems and Renewable Energies research team in Casablanca, Morocco. She is interested in

renewable energies and artificial intelligence.



**Touria Haidi** is head of the Automation, Electrical Systems and Renewable Energies research team of the Hassania School of Public Works (EHTP) in Casablanca, Morocco. She obtained a diploma in electrical engineering from EHTP, a post-graduate certificate in information processing from Ben M'sick University, a Master's degree in information

systems and processing in EHTP, a PhD, and a Habilitation in University Research (HDR) from ENSEM. She is currently a professor of electrical engineering at EHTP. Her research topics include renewable energy, electrical networks, distributed generation, and electrical machines.



**Faissal Elmariami** is a professor at the Ecole Nationale Supérieure d'Electricité et de Mécanique, Hassan II University, Casablanca, Morocco. He holds a PhD, engineering degree, and Habilitation à diriger des recherches (HDR). He is a team leader of the Electrical Networks and Static Converters research team. He is currently working on power system

stability and smart grids.



**Mounir Derri** completed his doctorate at the Faculty of Science and Technology, Settat, Hassan I University. He is currently a professor at the Hassania School of Public Works (EHTP) and a member of the Automation, Electrical Systems, and Renewable Energies research team in Casablanca, Morocco. His fields of interest include photovoltaics and electric vehicles.



**El Mehdi Mellouli** obtained his doctorate from Sidi Mohamed Ben Abdellah University, Faculty of Science, Fez, Morocco. He is currently a professor at the National School of Applied Sciences, University of Fez, Morocco. His research has focused on modeling and generating intelligent control laws for autonomous systems, robotics, renewable

energy systems, and unknown dynamic systems.

(Editor Yanbo Wang)