Hindawi
Complexity
Volume 2022, Article ID 8451812, 13 pages
https://doi.org/10.1155/2022/8451812

Research Article

Machine Learning as a Downscaling Approach for Prediction of Wind Characteristics under Future Climate Change Scenarios

Abbas Yeganeh-Bakhtiary,[1] **Hossein EyvazOghli**,[1] **Naser Shabakhty**,[1] **Bahareh Kamranzad**,[2,3,4] **and Soroush Abolfathi**[5]

[1]*School of Civil Engineering, Iran University of Science & Technology (IUST), Narmak, Tehran, Iran*
[2]*Hakubi Center for Advanced Research, Kyoto University, Yoshida Honmachi, Sakyo-ku 6068501, Kyoto, Japan*
[3]*Graduate School of Advanced Integrated Studies in Human Survivability (GSAIS), Kyoto University, Yoshida-Nakaadachi 1, Sakyo-ku, Kyoto 6068306, Japan*
[4]*Department of Physics, Faculty of Natural Sciences, Imperial College London, London, SW7 2AZ, UK*
[5]*School of Engineering, University of Warwick, Coventry, CV4 7AL, UK*

Correspondence should be addressed to Abbas Yeganeh-Bakhtiary; yeganeh@iust.ac.ir

Received 18 May 2022; Accepted 2 August 2022; Published 23 August 2022

Academic Editor: Teddy Craciunescu

Assessment of climate change impacts on wind characteristics is crucial for the design, operation, and maintenance of coastal and offshore infrastructures. In the present study, the Model Output Statistics (MOS) method was used to downscale a Coupled Model Intercomparison Project Phase 5 (CMIP5) with General Circulation Model (GCM) results for a case study in the North Atlantic Ocean, and a supervised machine learning method (M5' Decision Tree model) was developed for the first time to establish a statistical relationship between predicator and predicant. To do so, the GCM simulation results and altimeter remote sensing data were employed to examine the capabilities of the M5'DT model in predicting future wind speed and identifying spatiotemporal trends in wind characteristics. For this purpose, three classes of M5′ models were developed to study the annual, seasonal, and monthly variations of wind characteristics. The developed decision tree (DT) models were employed to statistically downscale the Beijing Normal University Earth System Model (BNU-ESM) global climate model output. The *M5′* models are calibrated and successfully validated against the GCM simulation results and altimeter remote sensing data. All the proposed models showed firm outputs in the training section. Predictions from the monthly model with a 70/30 training to test ratio demonstrated the best model performance. The monthly prediction model highlighted the decreasing trend in wind speed relative to the control period in 2030 to 2040 for the case study location and across all three future climate change scenarios tested within this study. This reduction in wind speed reduces wind energy by 13% to 19%.

1. Introduction

Over the past decades, excessive greenhouse gas emissions have resulted in an accelerated rate of global warming and intensified the effects of climate change. The increase in intensity and frequency of extreme climatic events is exacerbated by climate change, leading to natural hazards such as severe floods and erosion in coastal regions [1–11]. The large-scale effects of climate change are beginning to influence several parts of the world by increasing extreme climatic events. A study by Wei et al. [12] showed that the negative impacts of climate change occur at both global and local scales with detrimental consequences on coastal communities, which are at the forefront of battling against the climate change impacts. Given that climate change will also impact major socioeconomic activities and biodiversity in the coastal region, it is vital to have robust predicting frameworks capable of approximating the key climatic parameters in the future considering different climate projection models. Within the context of climate change, wind

climate in offshore and coastal regions is one of the key parameters that influence wave behaviour and hydroclimate [13]. Turki et al. [7] examined the multiscale components of the monthly extreme surges by considering the climatic parameters, e.g., zonal wind, sea surface temperature, and sea-level pressure along the English Channel coasts. Despite the current availability of vast amount of sensing data, high-resolution downscaled models are still needed to investigate the effects of climate variables based on the future climate change projections [14].

Global climate models (GCMs, aka General Circulation Models) have been developed to generate future projections through large-scale spatiotemporal data of climate variables [15]. At present, GCMs are widely used to predict and simulate the large-scale global climate response to increasing temperature at the surface of oceans. Many studies have been carried out using different GCMs to predict the impact of climate change on the variations of wind characteristics in various regions [16–20]. For instance, Segal et al. [21] employed HadCM2 (Hadley Centre coupled model) with coupling local wind data and showed that the availability of daily average wind power reduces within the range of 0–30% by 2050 over most areas of the United States. Later, Breslow and Sailor [22] also used outputs of global coupled global climate models (CGCMs) with a resolution of 3.75° (both latitude and longitude) and HadCM3 with a resolution of 3.75° at 2.5° (longitude and latitude, respectively). The results of the case study in the United States show that climate change reduces the average wind speed by 10% to 15%. This reduction in wind speed resulted in a 30% to 40% reduction in wind power.

Lionello et al. [23] studied the Adriatic Sea region using ECHAM-4 model data, which were downscaled using statistical methods. A comparison between the present and future climate simulations from Lionello et al. [23] study showed that extreme wave height will decrease in the future. In the UK, the 40 yearly records of data show an increase in winter wind speed by 15% to 20%, which can be linked to climate change consequences [24]. For two scenarios of $A2$ (1961–1990) and $B2$ (2071–2100), Lionello et al. [17] studied the seasonal mean of significant wave height (SWH) for the Mediterranean and predicted a reduction in SWH for the B2 period. For a case study of the Bay of Biscay, France, the results of the ARPEGE-Climate model show a reduction in the wind speed and, consequently, the wave height for the summers during the period of 2061 to 2100 [25]. Kamranzad [19] and Kamranzad et al. [26] investigated the capabilities of CGCM3.1 (Canadian Global Coupled Model Version 3.1) in the prediction of wind characteristics in Persian Gulf, which, considering its semienclosed shape, is vastly different from oceans. Kamranzad et al. [26] predicated that by the year 2100, wind speeds and wind energy will decrease across Persian Gulf according to three different emission scenarios ($A2$, $B1$, and $A1B$). Recently, Goharnejad et al. [27] evaluated both the wave characteristics and wave energy extraction potential in Persian Gulf for greenhouse gas concentration trajectory (representative concentration pathway: RCP)-based GCM simulation outputs of RCP4.5 and RCP8.5 future climate change scenarios and showed that the potential wave energy level will be higher in the southern regions.

Given the importance of the North Atlantic in ocean renewable energy and the heavy investment plans for extending the onshore and offshore renewable energy farms, understanding and quantifying the impacts of climate change on wind and wave characteristics across the North Atlantic are of exceptional importance. However, a limited number of studies have focused on the impacts of climate change on wind characteristics across the North Atlantic Ocean. Wang et al. [16] investigated mean and maximum seasonal variations of SWH based on three emission scenarios and showed that for fall and winter seasons during the twenty-first century, at the middle latitudes of the North Atlantic Ocean, the SWH will decrease, while at the southwest regions of the Atlantic the SWH will increase. Using the ECHAM5 model, Hemer et al. [28] predicted up to a 15% decline in SWH across the midlatitudes of the Atlantic and a 10% SWH reduction in the Southern regions of the Atlantic Ocean. The existing studies and modelling data indicate a decreasing trend in the future wind speed across vast regions of the North Atlantic.

Although GCMs are powerful in predicting the main features of the global atmospheric currents, they are often not capable of vigorously approximating local climate details [29]. Hence, there is a need to develop appropriate tools to downscale GCM climate change forecasts to local and regional scales [30]. Previous studies have adopted three downscaling approaches: empirical, semiempirical, and nesting methods (i.e., dynamical downscaling). In the empirical approach, historical climatic conditions are used to present local analogue scenarios. Such studies are attributable to a qualitative conceptual survey, and results from empirical approaches do not generate a climate forecasting model. Semiempirical (statistical) and nested (dynamical) downscaling approaches use large-scale GCM predictions to develop local climate change scenarios [31]. In the dynamical downscaling approach, a regional climate model (RCM) with the target mesh resolution uses large-scale GCM outputs as the boundary condition for the RCM to produce higher resolution outputs [32]. The major drawbacks of dynamical downscaling methods, which limit their applications in climate change impact assessments, are method complexity, high computational cost, and case-sensitive performance [33].

Statistical downscaling methods are divided into three categories [34, 35]: In the Perfect Prognosis (PP) methodology, a relationship is established between large-scale observational data and locally recorded data [36, 37]. The Model Output Statistics (MOS) method is similar to the PP, except that in this approach, a relationship is created between GCM outputs (predictor) and local climate variables (predictands) [38], and in the Stochastic Weather Generator (SWG) category, this relationship is developed by perturbing probably distribution parameters [39]. Considering the many parameters that are involved in the simulations of GCMs and the scenarios that are intended for the future and simulate trends, the use of the MOS method can be proper for downscaling where only limited observational data are

available in the application of the PP approach. For more information on statistical downscaling methods: [35, 40–44].

The statistical downscaling methods are designed based on two assumptions: (i) the empirical relationships between historical large-scale atmospheric predictors modelled by GCMs, and local climate characteristics can be established, and (ii) the obtained empirical relationships are valid under climate change scenarios [18, 45]. The most popular statistical downscaling approach is transfer functions based on fitting a quantitative relationship between large-scale climate variables and local-scale climate variables. In recent years, machine learning techniques have been adopted to determine the required transfer function in statistical downscaling [46, 47].

Due to the nonlinear time-series nature of climatic processes involved in the predication process, the artificial neural network (ANN), as a self-organizing estimator function, is widely adopted in modelling and forecasting wind characteristics (among others, [31, 48–51]). For example, Sailor et al. [31] adopted ANN technique to downscale and forecast surface wind speeds at three locations across the United States with a high potential for future wind power generation over the next 100 years. Under the future climate change scenarios, they estimated that wind power will decrease between 0.9% and 8%. Nourani et al. [49] employed ANN method to downscale climate variables, including temperature and precipitation, at two study locations (Ardabil and Tabriz, Iran) for single and multi-GCM outputs. Nourani et al. [49] showed that the downscaling method using ANN-based multi-GCM outputs leads to more accurate results.

Despite the advantages of ANN in downscaling and predictions of climatic processes, there are several deficiencies, including the probability of error and mismatch, due to not removing irrelevant data and noneffective parameters and the challenges associated with the training process by increasing the size of input time series. Given the generally nonstationary and large temporal scale (from a few minutes to several decades) of climatic data, the increased computational time required to train ANN models limits their applications [50]. Therefore, in recent years clustering methods were proposed as an alternative to robustly predict climatic variables and overcome the difficulties associated with the ANN downscaling approach. The decision tree (DT) is one of the most popular and efficient data mining techniques for clustering and generating regression models [52]. Decision tree (DT) models are clustering-type models with the advantages of setting out the variable choices logically, simultaneously considering potential options and choices, with tangible and easy-to-understand results [53].

Robustness of the DT models in identifying the effective parameters and understanding interdependencies of complex nonlinear climate variables [29, 50, 53–55] makes them a powerful tool for downscaling of GCM outputs and wind speed prediction. In this study, an M5' DT model too was developed to predict the spatiotemporal variations and trends in wind speed across a case study located in North Atlantic Ocean. The capability of M5' DT was examined for the first time for the prediction of wind speed variations in the coming decades, considering a range of climate change scenarios. The variations in the projected wind speed simulated by GCM (i.e., BNU-ESM model by Ji et al. [56]) were investigated as the case study region of the midlatitudes of the North Atlantic Ocean. The study region was chosen based on its importance and potential for current and future offshore wave and wind renewable energy farms. This study investigates the changes in wind speed over North Atlantic for the years 2030, 2035, and 2040, considering three future climate change scenarios outlined by the Intergovernmental Panel on Climate Change (IPCC: www.ipcc-data.org).

2. Materials and Methods

In this study, two sets of GCM and altimeter data were used to develop and verify a DT model to downscale the GCM outputs and predict the wind speed during 2030–2040. The case study location and details of the methodological approach adopted are described in the following section.

2.1. Study Area. Figure 1 shows the geographical extent and location of the case study area. The case study is an area with high wind speed, and the potential for efficient operations of offshore renewable energy projects was selected as the case study location. The boundaries of the study area were selected based on the GCM grids, covering 340.31°E to 357.19°E in longitude and 57.21°N to 62.79°N in latitude. The downscaling and machine learning model developed was applied to the data for the case study area to predict the wind speed variations under future climate scenarios and evaluate the performance of the proposed method in comparison to the measured data.

2.2. General Circulation Models (GCMs) Dataset. The Earth System Model (Beijing Normal University (BNU)-ESM) developed by Ji et al. [56] is based on climate projection models and is widely used to study climate change impacts, ocean-atmosphere interaction mechanisms, and climate-carbon interactions on temporal scales spanning from a month up to a century. The ESM benefits from several submodels, including atmospheric, ocean, sea ice, and land models. The ESM coupling framework is developed based on the Community Climate System Model version 4 (CCSM4) together with the Community Climate System Model and Community Earth System Model (CCSM/CESM) [56]. In the present study, the GCM historical data of 15 years (between 1991 and 2005) were employed as the control period, and the forecast data under three future climate scenarios, including RCP2.6, RCP4.5, and RCP8.5, were used to predict the climate variables for 2030, 2035, and 2040.

2.3. Altimeter Dataset. Historically, wind characteristics have been studied by analysing time-series records of weather stations that provide reliable measurements of temporal variations of wind characteristics in a fixed position. In recent years, remote sensing techniques such as

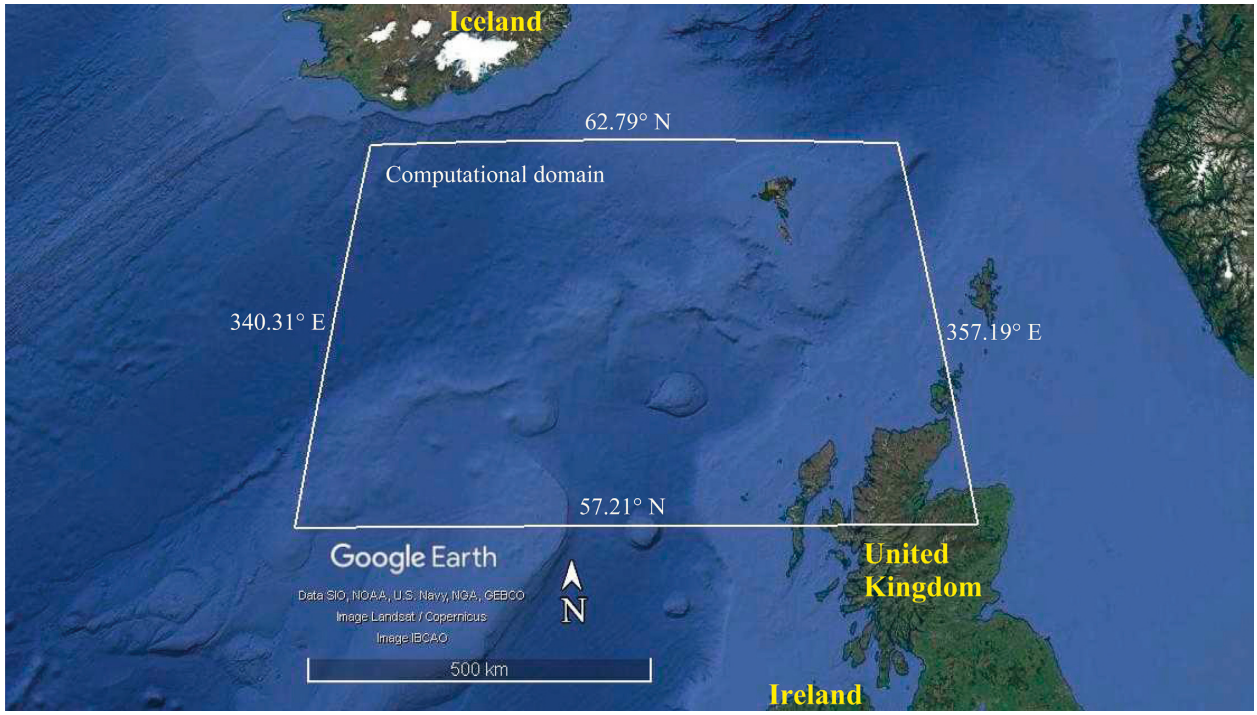

Figure 1: Geographical extent and location of the case study area.

radar measurements and satellite imagery have been widely used as tools to analyse key atmospheric parameters and sea conditions. The spatial distribution of wind characteristics has been analysed and evaluated from remote sensing imageries [57, 58].

Altimeter data are one of the most valuable data sources available that have been employed in previous investigations to verify the performance of a range of climate prediction models (see among others, [59–63]). The altimeter data are recorded along the paths of satellites using remote sensing equipment and provide a firm remote sensing dataset to validate and evaluate the performance of the climate projection models. Altimeter data records are entered in irregular tracks and at different time intervals, which demonstrates the importance of applying appropriate filters to these data to ensure the consistency of the data records. For this purpose, we adopted a 20-minute filter in the time domain and a 1.5-degree filter in the spatial domain to tackle the spatiotemporal nonlinearity of data. The wind speed obtained from altimeter data was used for the downscaling process and validation of the developed DT models. Considering the available period of GCM simulation outputs, the altimeter dataset belonging to the period of 1991–2005 was selected as the control period.

2.4. M5′ Model Tree. Recently, the use of decision tree (DT) algorithms as a robust machine learning technique for prediction of hydroclimatic parameters in coastal and offshore engineering problems has found a growing interest (among others, [51, 54, 55]). In general, the structure of a decision tree is composed of four parts including root, branch, node, and leaves. The root (or first node) is at the top of the tree; also, at the end of the chain of branches and nodes are the leaves (or the last node). Figure 2 shows a schematic of the DT logical structure used as a predictive model. Given that DTs can be classed as a graphical method, interpretations of DT's model outputs are easier compared to other machine learning methods [50].

Nourani et al. [64] investigations on the application of machine learning techniques for predicting climate variables show that as the prediction horizon increases, the accuracy of predictive models is reduced. The reason is due to the nonlinear growth of error propagations in those nonlinear predictive models. However, the issues associated with the error growth are not the case for linear models such as M5′ DTs, as the error in such models will remain constant by increasing the prediction horizon. Thus, the multilinear regression models such as *M5′* model can provide more reliable results in predicting climate variables, compared to nonlinear models. Furthermore, the M5′DT models have superior performance in identifying and selecting the most effective parameters for persistent prediction [50].

2.5. Model Development. Prior to M5′ model development, the control period dataset (historical data) was divided into train and test partitions. The train data were employed for development of the M5′DT models, whereas the test data were utilized for verification and evaluations of the generated M5′ models. Following the data partitioning, the train data were employed to develop three predictive models with M5′ technique. These models were designed with different ratios of test/train data length and three different train/test size ratios of 35/65, 30/70, and 25/75 to evaluate the best predictive model.

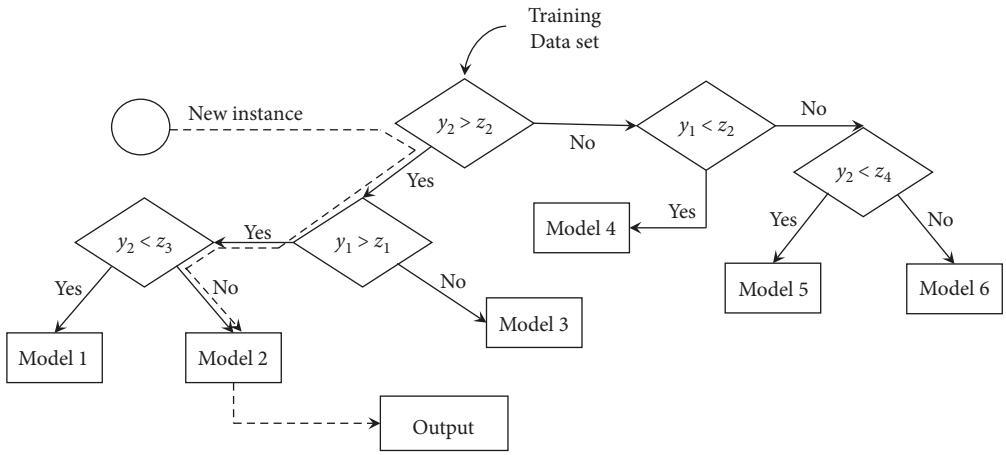

Figure 2: Flow chart of M5′ model prediction procedures [55].

In the first step, the *M5′* models were trained; the process of data selection for training the *M5′* models involved specifying the coordinates of each data instance and determining its corresponding cell from the GCM mesh grid to obtain the wind speed and direction for the specific cell. Based on the instance's time, a time interpolation is performed on the GCM data output. Then, the interpolated wind speeds and direction for a given location and time instance were determined from the GCM cell's corners and introduced to the *M5′* models as an input. In the second step, the trained *M5′* models were verified using the test data from the historical dataset. Then the performance of each model was evaluated using statistical error measures to identify and select the best model for spatial and temporal prediction of wind characteristics across the case study. The final step of the modelling involves prediction of wind speed for the time intervals of 2030 to 2040 based on the three emission scenarios employing the GCM simulation data.

2.6. Performance Evaluation of Predictive Models. Given the nonlinear nature of climatic events and influence of complex marine processes varying at both temporal and spatial scales, predicting the wind characteristics in marine environment is a very difficult and challenging task [65]. Having a strong approach to evaluate and overcome the systematic and random errors of the developed DT models is also crucial. To do so, several statistical assessment criteria, namely correlation coefficient (CC), root mean square error (RMSE), and mean absolute error (MAE), were investigated to evaluate and compare the predictive robustness of the proposed models under future climatic conditions. The correlation coefficient (CC) (1) is adopted to determine the relationship between the predicted wind speed and measured values as:

$$CC = \frac{\sum_{i=1}^{N}\left(\left(O_i - \overline{P}\right)\left(O_i - \overline{O}\right)\right)}{\sqrt{\left(\sum_{i=1}^{N}\left(P_i - \overline{P}\right)^2\right)\left(\sum_{i=1}^{N}\left(O_i - \overline{O}\right)^2\right)}}, \quad (1)$$

where P_i and O_i, respectively, denote the predicted and measured data (observations), $\overline{P}$ is the average of the predicted data, $\overline{O}$ is the average of the measured data, and N is the number of data points. CC only determines the correlation between measured and predicted values and therefore it is not a sufficient measure to provide a comprehensive understanding of the model's performance. Thus, it is critical issue to employ benchmarks that rigorously determine the model's prediction errors. To this end, the present study employed MAE and RMSE statistical measures in equations (2) and (3).

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|P_i - O_i\right|, \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(P_i - O_i\right)^2}. \quad (3)$$

3. Results and Discussion

To obtain each downscaled wind speed, eight parameters from GCM data were inputted to the *M5′* models, including four wind speeds and four wind directions. The purpose of wind directions participating in the wind speed downscaling was to help the projection model to predict wind speeds more accurately at locations such as the shoreline, where wind directions change rapidly. In the following, the results are presented in two parts: results of the control case and results of the prediction case.

3.1. Control Period. In the control case, three types of prediction models were developed and named as: Annual (*A*), Seasonal (*S*), and Monthly (*M*) models. In the *A*-type model, all control period data were used to develop every single *M5′* model. Where the results showed that in the *A*-type model, although the correlation coefficient between the observational data and the outputs of the DT model in the training period was high, in the test period, the correlation coefficient between the measured and predicted data was low, and the prediction error raised considerably. Since the results of the A-type model were not accurate enough, and

Table 1: Developed models in the present study and their description.

Model type	Description
A (Annual)	One DT model was developed using the Train data, validated by the Test data, and adopted for prediction in the forecast section.
S (Seasonal)	The entire data were categorized into seasons, and four decision tree (DT) models were generated alongside together. Each model is developed and validated with the data of the relevant season and used for forecasting in the season for the projection period. The names of these models based on the seasons are as follows: Swinter: winter's model, Sspring: spring's model, Ssummer: summer's model, and Sfall: fall's model.
M (Monthly)	The entire data were divided into months, and 12 DT models were extended beside together. Each model is developed and verified with the data of the relevant month and used for forecasting in the relevant month for the projection period. The names of these models based on the months are MJan to MDec models for January to December, respectively.

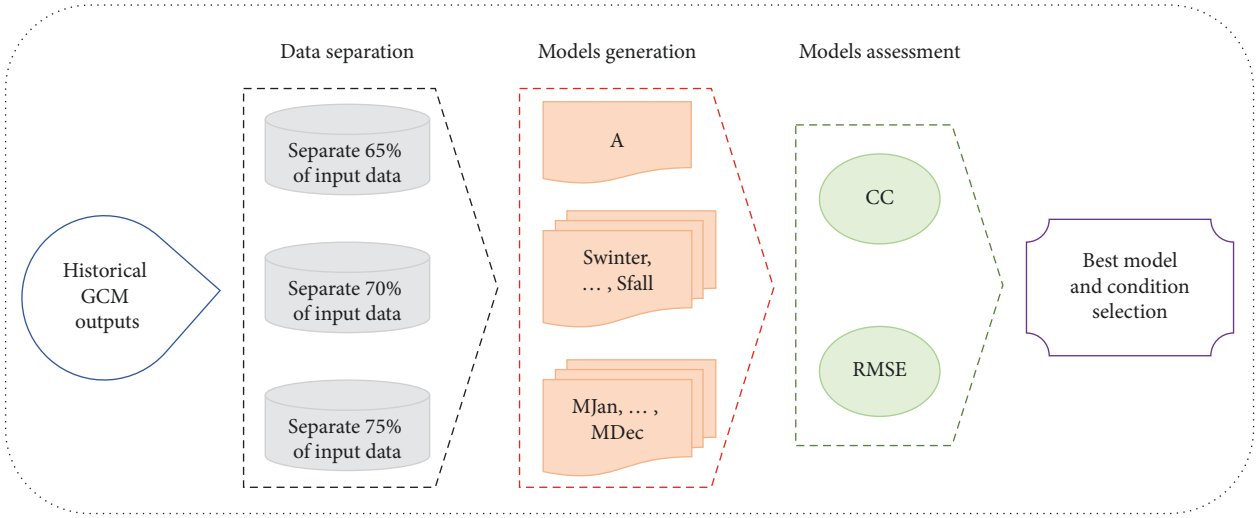

Figure 3: Schematic of the downscaling methodology and predictive modelling processes.

Table 2: Comparison of statistical predictive measures across all the tested scenarios.

Model	Training (70/30)		Verification					
	CC	RMSE	Correlation coefficient			Root mean square error		
			65/35	70/30	75/25	65/35	70/30	75/25
A	0.9028	2.0117	0.3642	0.1237	0.2136	5.7777	17.4996	4.1732
Swinter	0.8944	2.0585	0.0514	0.2003	0.1699	32.723	10.9887	12.8441
Sspring	0.0896	1.8625	0.1296	0.1679	0.6327	15.2014	10.8614	3.0635
Ssummer	0.8934	1.8063	0.0744	0.2707	0.1093	30.627	6.2338	15.2301
Sfall	0.888	2.0465	0.075	0.5276	0.5049	3.9912	3.6316	3.8453
Mean	0.6914	1.9434	0.0826	0.2916	0.3542	20.6356	7.9289	8.7458
MJan	0.8983	2.0922	0.56	0.8054	0.7704	3.7827	2.6749	2.8072
MFeb	0.9046	1.9672	0.5936	0.7791	0.0881	3.6361	2.6831	41.9419
MMar	0.9026	1.9325	0.7804	0.7974	0.6843	2.6535	2.5707	3.0753
MApr	0.8988	1.8653	0.1763	0.7333	0.4716	12.1353	2.6809	4.2891
MMay	0.897	1.7869	0.2977	0.6998	0.0389	5.8888	2.6363	47.1423
MJun	0.8892	1.7279	0.558	0.6979	0.3844	3.0569	2.4745	4.5731
MJul	0.8822	1.6457	0.5116	0.6576	0.6059	2.9252	2.6373	2.6362
MAug	0.8903	1.6815	0.7566	0.7566	0.7392	2.2544	2.2544	2.2925
MSep	0.902	1.9198	0.3506	0.7627	0.0679	5.77	2.6672	40.5325
MOct	0.8815	1.9988	0.0362	0.7565	0.4504	3.86	2.5783	4.0723
MNov	0.889	1.9712	0.0423	0.7174	0.0614	73.324	2.7225	22.6081
MDec	0.8984	2.0211	0.3008	0.7511	0.0912	7.0636	2.7503	24.5093
Mean	0.8945	1.8842	0.4137	0.7429	0.3711	10.5292	2.6109	16.7067

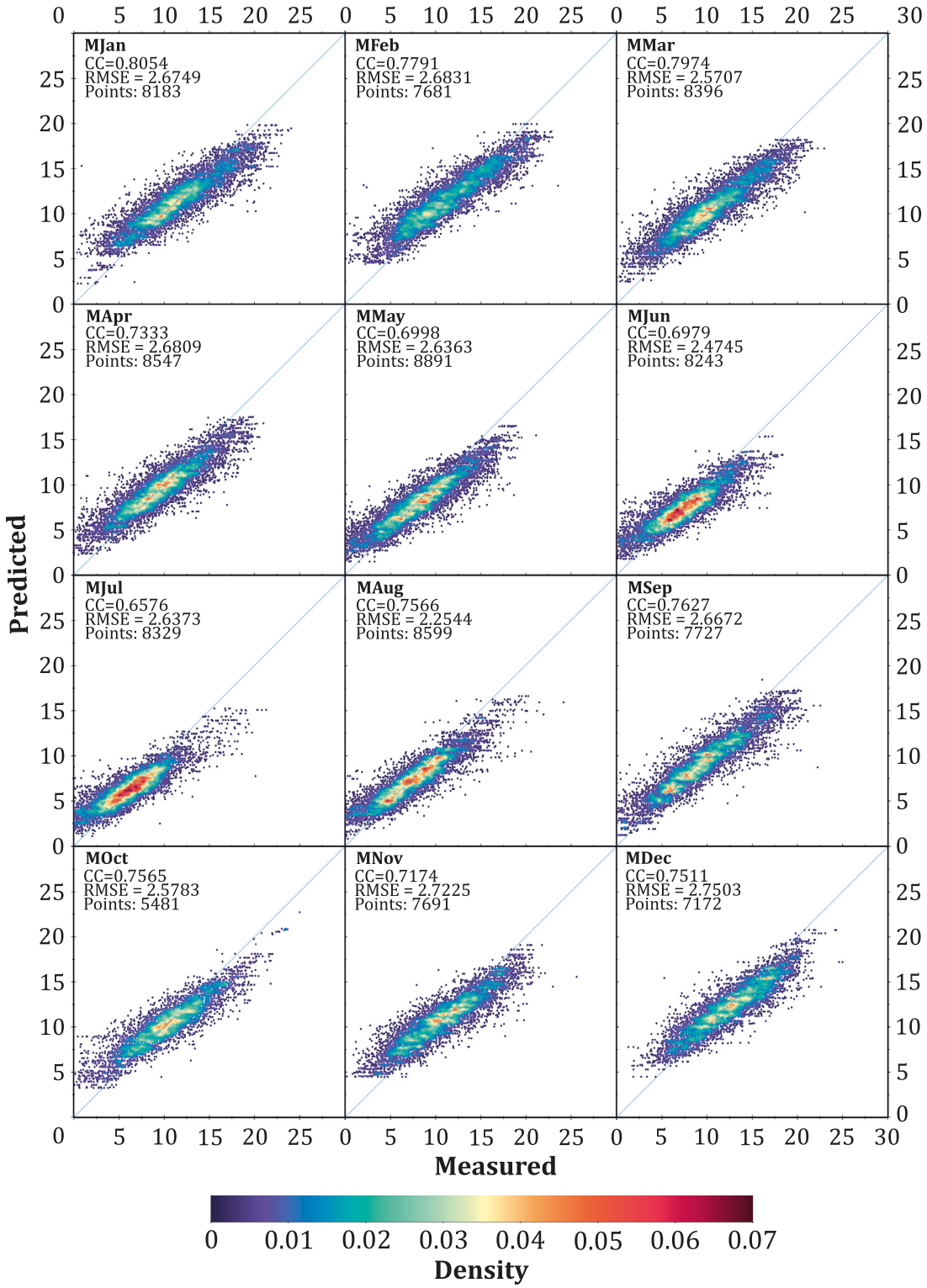

Figure 4: Comparison between the observed data and decision tree (DT) model prediction.

the $M5'$ model acts based on data categories theory, S and M types of the models were developed. Thus, all the data are divided into seasons and months, and the progress of the $M5'$ model development is conducted for each of the categories separately (a summarized description of each model presented in Table 1). Figure 3 presents schematically the downscaling procedure and models' deployment.

Table 1 presents the differences between the models, and as can be seen, the main difference is in the number of models created to downscale GCM's outcomes and predict

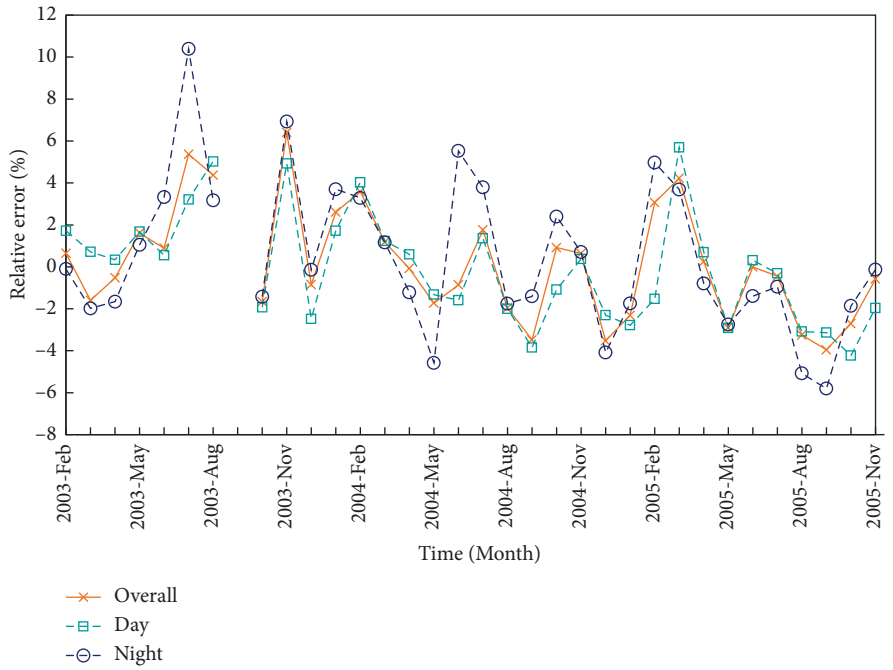

Figure 5: Comparison of the overall, and day and night time prediction errors of M5′ models.

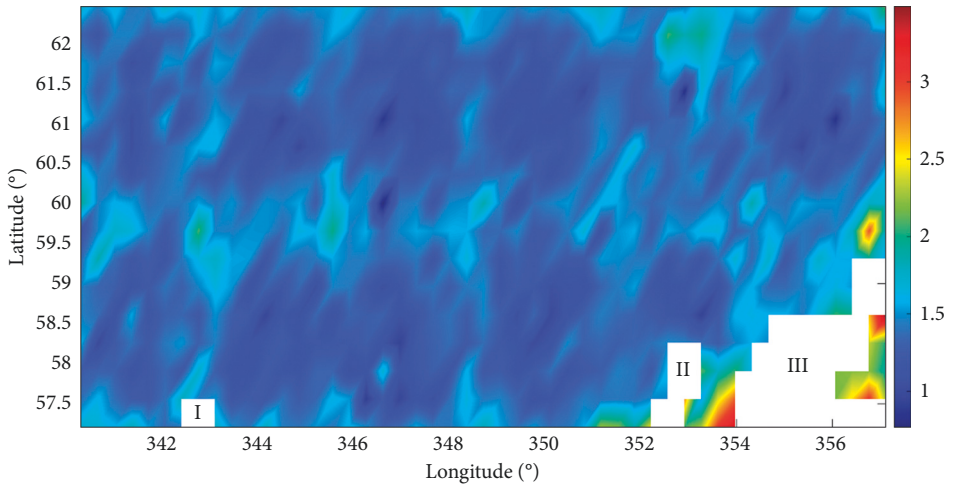

Figure 6: Spatial distribution of mean absolute error (MAE) across the case study region.

the trend of wind changes. The principal idea of the model's development is how the DT model works, which was attempted to prepare more homogeneous input data for each model by dividing the data so that the results reach acceptable accuracy.

Table 2 summarizes the results of the models generated during the training and validation steps. The correlation coefficient (CC) and the root mean square error (RMSE), and the mean of the seasonal and monthly models are reported in this table to compare the generated models' performance. As seen in the training section, the CC for all models is close to 0.9, except for model Sspring, in which CC is less than 0.1. All the models' RMSEs are in the range of 1.6 to 2.0. On the basis of the results, the best model in terms of CC is the *A*-type model with CC = 0.9. Considering the mean of CC, the *A*-type model has also shown to be more efficient. However, considering the RMSE, the annual model performs weakly, and the monthly models often have better performance. In terms of mean values, the *M*-type model with an RMSE of 1.88 has the lowest error, whereas the RMSE of the seasonal and annual models are 1.94 and 2.01, respectively. The annual models have the weakest output among the models. This may be attributed to the high volume of input data leading to the *A*- and *S*-type models, which causes misunderstanding in both classifying the data and establishing a reliable correlation between the input data (wind speed and direction data) and the output data (wind speed data). In contrast, the monthly models were able to detect the trend and establish a proper relationship between the input and output data.

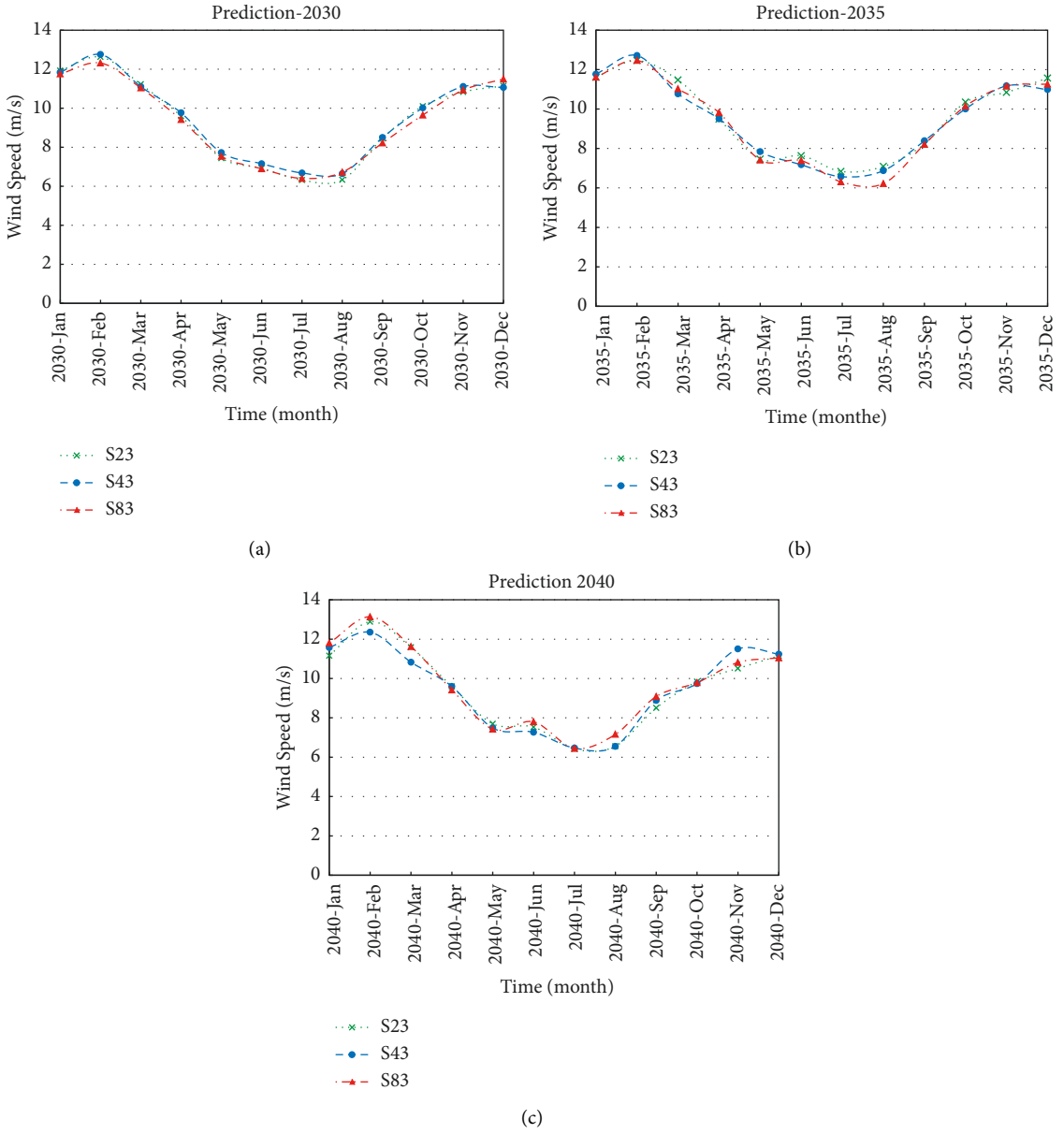

Figure 7: Monthly averaged wind speed predicted based on future climate scenarios (RCP2.6, RCP4.5, RCP8.5) for the year. (a) 2030. (b) 2035. (c) 2040.

In addition, with a close look at Table 2, it can be understood that adopting different historical data for the train and test periods affects the developed model performance and the results. Three modes of data division—65/30, 70/30, and 75/25—and the outcomes of these models' implementation are presented in this paper. As it comes from Table 2, the *A*-type model failed to make accurate predictions in the test section as the best model of *A*-type in 75/25 mode had a CC of 0.21 and RMSE of 4.17. Despite the relative increase of the mean correlation coefficient in the 30/70 and 25/75 modes, *S*-type models have not yet reached an acceptable level, and the errors are still high. The best results belong to the 70/30 mode of *M*-type models with a mean CC of 0.74 and mean RMSE of 2.61 in the test section.

The lack of a persuasive link between the input and output values in the other modes of *M*-type models can be mainly due to the high sensitivity of *M5′* models to the input data and shows the importance of examining various modes of models before adopting them as a prediction model. A detailed appraisal of the selected model (70/30 mode of *M*-type models) is presented in the following. The scatter plots of the measured data versus the models' results are shown to check the projection models' output quality and the correlation between the predicted and observed data (Figure 4). As shown in Figure 4, the *M*-type model outputs had a perfect correlation with the recorded data. It proves the high efficiency of trend recognition and prediction by M-type models.

Table 3: Predicted monthly mean (PMM) of wind speed changes compared to the year 2005.

	Monthly prediction based on Scenarios and Years									
	Historical	RCP2.6			RCP4.5			RCP8.5		
Month	2005	2030	2035	2040	2030	2035	2040	2030	2035	2040
Jan	13.2741	11.93	11.61	11.16	11.84	11.77	11.57	11.76	11.61	11.80
Feb	11.8188	12.63	12.45	12.89	12.75	12.72	12.35	12.32	12.46	13.14
Mar	10.1221	11.22	11.47	11.59	11.10	10.77	10.83	11.04	11.03	11.62
Apr	10.0947	9.590	9.483	9.593	9.763	9.500	9.598	9.416	9.816	9.419
May	9.0151	7.441	7.463	7.694	7.726	7.840	7.480	7.539	7.411	7.422
Jun	7.7712	6.922	7.636	7.530	7.156	7.178	7.268	6.899	7.385	7.797
Jul	7.0615	6.315	6.836	6.414	6.682	6.572	6.464	6.401	6.309	6.447
Aug	8.5289	6.329	7.085	6.555	6.648	6.873	6.550	6.721	6.218	7.160
Sep	9.6920	8.455	8.222	8.503	8.498	8.390	8.884	8.220	8.208	9.083
Oct	11.5291	10.10	10.35	9.839	10.02	10.00	9.741	9.650	10.15	9.788
Nov	11.8781	10.83	10.85	10.51	11.11	11.18	11.50	10.93	11.16	10.81
Dec	11.1440	11.14	11.57	11.15	11.06	10.99	11.24	11.49	11.26	11.05
Mean	10.1608	9.409	9.586	9.452	9.529	9.481	9.457	9.365	9.418	9.628

For a better understanding of the monthly model performance, the average error of wind speed across the entire study area for the predicted values is depicted in Figure 5. As seen from Figures 4 and 5, although the most part of wind speed data vary from around 3 to 15 m/s, the monthly average of predicted values entirely meets the measured values in the test section. In addition to the total relative error (overall error), to get a better picture of the model performance, two relative error curves for day and night periods are also portrayed in Figure 5. As seen, from the Figure, during the daytime, the model prediction is slightly better than during the night time; while, the overall monthly predication of the wind speed was hindcasted precisely. Therefore, it is evident that the monthly model not only shows the proper distribution of wind speed, but it also accurately predicts the wind speed values in the time domain.

In addition to the time domain investigation, it is necessary to evaluate the spatial distribution of prediction errors over the study area. Figure 6 depicts the mean absolute error (MAE) over the entire region as averaged over the validation period. Notably, a filter was applied in portraying the map to be more reliable, and just cells with more than two data were shown in colour. As seen, Figure 6 contains four zones: zone I remained white because there are no observational data in this zone, and therefore no error was calculated. For zones II and III, MAE was also not figured since the relevant cells fully contain land and no observational data are available in these zones. The fourth zone is the coloured area and represents the MAE in the cells with dimensions of 0.4 by 0.4°.

As seen in Figure 6, the MAE in most parts of the study area is less than 2.0 m/s and remains in the acceptable range. It is only near zones II and III that the MAE reaches about 3.0 m/s; the reason for this phenomenon is also the proximity of this part of the region to the land, where the direction and speed of the wind change rapidly. These rapid changes in wind direction disrupt the performance of projection equation and most probably caused significant errors near the coasts.

3.2. Future Projection. In the prediction case, the wind speed is predicted for years of 2030, 2035, and 2040 by implementing the validated monthly model (M-type model in the 70/30 mode) and the GCM prediction data under three future climate change scenarios RCP2.6, RCP4.5, and RCP8.5. Similar to the control stage, the prediction model input was four wind speeds and in four wind directions. In order to simulate as closely as possible to the control period, the time and location of the selected points to anticipate wind speeds in the prediction case are the same as the time and position of the 2004 records, having the highest number of records per year. Thus, the measurement specifications of the 2004 records were used to predict with changing only the components of the "Year" and fixing the other terms (e.g., month, day, hours, minutes, and seconds) and maintaining the longitudes and latitudes.

Figure 7 shows the monthly average wind speed predictions for different future climate change scenarios across the study area. As seen, during 2030 to 2040, the monthly variations of mean wind speed were at a constant range of 6 to 13 m/s, then after reaching its peak in the second month, the decreasing trend began to reach 11 m/s in August. The wind speed trends under the three future climatic scenarios are rather close to each other for those three years. However, the monthly mean wind speed fluctuations increased from 2030 to 2040, and most fluctuations were related to the RCP8.5 scenario.

For further investigation, the predicted monthly mean (PMM) wind speed values are compared with the corresponding wind speed in 2005 (the last year of the GCM historical simulation outputs). Table 3 lists the changes in PMM values. As shown in Table 3, it is only in February and March that the PMM wind speed is higher; in December, on the other hand, there are no significant changes compared to 2005. For the rest of the years, the projected PMM wind speed was declining. Also, the annual mean values of wind speed (the bottom row of Table 3) were reduced comparing to 2005. The reduction ratio of the annual mean wind speed is approximately 7% to 8%, which is in line with the reported results of previous studies (e.g., [16, 28]). Reduction in wind

speed substantially impacts the wind energy due to climate change. Considering the Segal et al. [21] equation for wind power related to the third order of wind speed, the 7% to 8% reduction in annual mean wind speed indicates a 13% to 19% decrease in wind energy.

4. Conclusion

The intensity and frequency of extreme climatic events are exacerbated due to the impacts of climate change in different parts of the world. Predicting the wind characteristics under future climate change scenarios is very vital for evaluating the performance of existing and future marine engineering projects. This study adopts M5′ DT technique for the first time as a MOS multilinear downscaling method for predicting wind speed trends across a case study region in North Atlantic Ocean. The GCM simulation outputs and altimeter remote sensing data were used to train and validate the developed $M5'$ models. Three climate change scenarios of RCP2.6, RCP4.5, and RCP8.5 were adopted for deriving the predictions from $M5'$ models. Three downscaling models, including annual model (A), seasonal models (S), and monthly models (M), were tested for predicting wind speed from DT models. The performance of three $M5'$ models with train-test data ratios of 75/25, 70/30, and 65/35 was investigated to determine the best performing $M5'$ model. Data from the control case were used to choose the best model for predicting wind speed under future climate scenarios. The following conclusion can be drawn:

(i) The detailed analysis of the prediction results from the $M5'$ technique indicates the robustness and appropriateness of the proposed models for assessing wind characteristics under future climatic scenarios. The proposed approach for downscaling has successfully predicted the trend of monthly averaged wind changes in the study area.

(ii) Model performance evaluations were conducted employing appropriate statistical measures. As a result, a detailed analysis of $M5'$ models shows appropriate performance in the training section; whereas, the monthly models provided more reliable predictions in the test section. Moreover, the developed models performed in several modes, and the monthly model in the mode of 70/30 presented the best performance. This diversity in the models' performances in different modes indicated the sensitivity of the $M5'$ DT model to the input data.

(iii) Based on the monthly model predictions, in 2030, 2035, and 2040, the average monthly wind speed values for all the three future climate change scenarios were close. However, from 2030 to 2040, the range of monthly mean wind speed oscillations increased, which in the RCP8.5 scenario was more evident than in the other two scenarios.

(iv) A comparison between the PMM wind speeds and the year 2005 shows that the PMM wind speeds decreased remarkably. Hence, wind energy has also been experimented with a reduction that could be of interest in renewable energy studies and projects in the study area and would challenge the economic exploitation of these resources.

Data Availability

All observed data used during this research are openly available at (ftp://ftp.ifremer.fr/ifremer/cersat/products/swath/altimeters/waves/data/). Also, the data that support the findings of this study are available upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The authors would like to thank Dr. Seyed Mostafa Siadatmousavi, Mr. Mostafa Beyramzadeh, and Mr. Amir Molajou for their helpful technical feedback.

References

[1] S. Dong, M. Salauddin, S. Abolfathi, Z. H. Tan, and J. M. Pearson, "The influence of geometrical shape changes on wave overtopping: a laboratory and SPH numerical study," in *Coasts, Marine Structures and Breakwaters*, pp. 1217–1226, ICE Publishing, London, 2017.

[2] M. Armanfar, H. Goharnejad, M. Z. Niri, and W. Perrie, "Assessment of coastal vulnerability in Chabahar Bay due to climate change scenarios," *Oceanologia*, vol. 61, no. 4, pp. 412–426, 2019.

[3] A. Fitri, R. Hashim, S. Abolfathi, and K. N. Abdul Maulud, "Dynamics of sediment transport and erosion-deposition patterns in the locality of a detached low-crested breakwater on a cohesive coast," *Water*, vol. 11, no. 8, p. 1721, 2019.

[4] S. Dong, S. Abolfathi, M. Salauddin, Z. Tan, and J. Pearson, "Enhancing climate resilience of vertical seawall with retrofitting - a physical modelling study," *Applied Ocean Research*, vol. 103, Article ID 102331, 2020.

[5] M. Salauddin, J. O'Sullivan, S. Abolfathi, S. Dong, and J. Pearson, "Distribution of individual wave overtopping volumes on a sloping structure with a permeable foreshore," *Coastal Engineering Proceedings*, vol. 36v, p. 54, 2020.

[6] S. Dong, S. Abolfathi, M. Salauddin, and J. Pearson, "Spatial distribution of wave-by-wave overtopping at vertical seawalls," *Coastal Engineering Proceedings*, vol. 36v, p. 17, 2020.

[7] I. Turki, N. Massei, B. Laignel, and H. Shafiei, "Effects of global climate oscillations on intermonthly to interannual variability of sea levels along the English channel coasts (NW France)," *Oceanologia*, vol. 62, no. 2, pp. 226–242, 2020.

[8] A. Yeganeh-Bakhtiary, H. Houshangi, and S. Abolfathi, "Lagrangian two-phase flow modeling of scour in front of vertical breakwater," *Coastal Engineering Journal*, vol. 62, no. 2, pp. 252–266, 2020.

[9] H. Gao, B. Liang, and Z. Shao, "A global climate analysis of wave parameters with a focus on wave period from 1979 to 2018," *Applied Ocean Research*, vol. 111, Article ID 102652, 2021.

[10] M. Salauddin, J. J. O'Sullivan, S. Abolfathi, and J. M. Pearson, "Eco-engineering of seawalls—an opportunity for enhanced

climate resilience from increased topographic complexity," *Frontiers in Marine Science*, vol. 2021, Article ID 674630, 2021.

[11] M. Salauddin, J. O'Sullivan, S. Abolfathi, and J. M. Pearson, "Extreme wave overtopping at ecologically modified sea defences," *EGU General Assembly*, vol. 2020, 2020.

[12] J. Wei, A. Hansen, Y. Zhang et al., "Perception, attitude and behavior in relation to climate change: a survey among CDC health professionals in Shanxi province, China," *Environmental Research*, vol. 134, pp. 301–308, 2014.

[13] D. E. Reeve, Y. Chen, S. Pan, V. Magar, D. Simmonds, and A. Zacharioudaki, "An investigation of the impacts of climate change on wave energy generation: the Wave Hub, Cornwall, UK," *Renewable Energy*, vol. 36, no. 9, pp. 2404–2413, 2011.

[14] R. Kabir, H. T. A. Khan, E. Ball, and K. Caldwell, "Climate change impact: the experience of the coastal areas of Bangladesh affected by cyclones sidr and aila," *Journal of Environmental and Public Health*, vol. 2016, Article ID 9654753, 2016.

[15] P. Camus, F. J. Mendez, and R. Medina, "A hybrid efficient method to downscale wave climate to coastal areas," *Coastal Engineering*, vol. 58, no. 9, pp. 851–862, 2011.

[16] X. L. Wang, F. W. Zwiers, and V. R. Swail, "North Atlantic Ocean wave climate change scenarios for the twenty-first century," *Journal of Climate*, vol. 17, no. 12, pp. 2368–2383, 2004.

[17] P. Lionello, S. Cogo, M. B. Galati, and A. Sanna, "The Mediterranean surface wave climate inferred from future scenario simulations," *Global and Planetary Change*, vol. 63, no. 2-3, pp. 152–162, 2008.

[18] B. Kamranzad, A. Etemad-Shahidi, V. Chegini, and S. Hadadpour, "Assessment of CGCM 3.1 wind field in the Persian Gulf," *Journal of Coastal Research*, vol. 65, pp. 249–253, 2013.

[19] B. Kamranzad, "Assessment of the changes in average wind speed in Chabahar, Gulf of Oman, due to climate change," *Journal Of Marine Engineering*, vol. 10, no. 19, pp. 13–20, 2014.

[20] C. W. Zheng, C. Y. Li, and X. Li, "Recent decadal trend in the North Atlantic wind energy resources," *Advances in Meteorology*, vol. 2017, Article ID 7257492, 8 pages, 2017.

[21] M. Segal, Z. Pan, R. W. Arritt, and E. S. Takle, "On the potential change in wind power over the US due to increases of atmospheric greenhouse gases," *Renewable Energy*, vol. 24, no. 2, pp. 235–243, 2001.

[22] P. B. Breslow and D. J. Sailor, "Vulnerability of wind power resources to climate change in the continental United States," *Renewable Energy*, vol. 27, no. 4, pp. 585–598, 2002.

[23] P. Lionello, A. Nizzero, and E. Elvini, "A procedure for estimating wind waves and storm-surge climate scenarios in a regional basin: the Adriatic Sea case," *Climate Research*, vol. 23, no. 3, pp. 217–231, 2003.

[24] G. P. Harrison and A. R. Wallace, "Sensitivity of wave energy to climate change," *IEEE Transactions on Energy Conversion*, vol. 20, no. 4, pp. 870–877, 2005.

[25] E. Charles, D. Idier, P. Delecluse, M. Deque, and G. Le Cozannet, "Climate change impact on waves in the Bay of Biscay, France," *Ocean Dynamics*, vol. 62, no. 6, pp. 831–848, 2012.

[26] B. Kamranzad, A. Etemad-Shahidi, V. Chegini, and A. Yeganeh-Bakhtiary, "Climate change impact on wave energy in the Persian Gulf," *Ocean Dynamics*, vol. 65, no. 6, pp. 777–794, 2015.

[27] H. Goharnejad, E. Nikaein, and W. Perrie, "Assessment of wave energy in the Persian Gulf: an evaluation of the impacts of climate change," *Oceanologia*, vol. 63, no. 1, pp. 27–39, 2021.

[28] M. A. Hemer, J. Katzfey, and C. E. Trenham, "Global dynamical projections of surface ocean wave climate for a future high greenhouse gas emission scenario," *Ocean Modelling*, vol. 70, pp. 221–245, 2013.

[29] R. Schnur and D. P. Lettenmaier, "A case study of statistical downscaling in Australia using weather classification by recursive partitioning," *Journal of Hydrology*, vol. 212-213, pp. 362–379, 1998.

[30] F. Giorgi and L. O. Mearns, "Approaches to the simulation of regional climate change: a review," *Reviews of Geophysics*, vol. 29, no. 2, pp. 191–216, 1991.

[31] D. J. Sailor, T. Hu, X. Li, and J. Rosen, "A neural network approach to local downscaling of GCM output for assessing wind power implications of climate change," *Renewable Energy*, vol. 19, no. 3, pp. 359–378, 2000.

[32] H. J. Fowler, S. Blenkinsop, and C. Tebaldi, "Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling," *International Journal of Climatology*, vol. 27, no. 12, pp. 1547–1578, 2007.

[33] S. Ghosh and C. Misra, "Assessing hydrological impacts of climate change: modeling techniques and challenges," *The Open Hydrology Journal*, vol. 4, no. 1, pp. 115–121, 2010.

[34] D. Maraun, F. Wetterhall, A. M. Ireson et al., "Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user," *Reviews of Geophysics*, vol. 48, no. 3, Article ID RG3003, 2010.

[35] K. Srinivasa Raju and D. Nagesh Kumar, *Impact of Climate Change on Water Resources*, Springer, Singapore, 2018.

[36] R. L. Wilby and T. M. L. Wigley, "Precipitation predictors for downscaling: observed and general circulation model relationships," *International Journal of Climatology*, vol. 20, no. 6, pp. 641–661, 2000.

[37] R. M. Trigo and J. P. Palutikof, "Precipitation scenarios over Iberia: a comparison between direct GCM output and different downscaling techniques," *Journal of Climate*, vol. 14, no. 23, pp. 4422–4446, 2001.

[38] X. C. Zhang, "Spatial downscaling of global climate model output for site-specific assessment of crop production and soil erosion," *Agricultural and Forest Meteorology*, vol. 135, no. 1-4, pp. 215–229, 2005.

[39] J. Chen, F. P. Brissette, D. Chaumont, and M. Braun, "Performance and uncertainty evaluation of empirical downscaling methods in quantifying the climate change impacts on hydrology over two North American river basins," *Journal of Hydrology*, vol. 479, pp. 200–214, 2013.

[40] M. S. Khan, P. Coulibaly, and Y. Dibike, "Uncertainty analysis of statistical downscaling methods," *Journal of Hydrology*, vol. 319, no. 1-4, pp. 357–382, 2006.

[41] J. Chen, F. P. Brissette, D. Chaumont, and M. Braun, "Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America," *Water Resources Research*, vol. 49, no. 7, pp. 4187–4205, 2013.

[42] D. Mullan, J. Chen, and X. J. Zhang, "Validation of non-stationary precipitation series for site-specific impact assessment: comparison of two statistical downscaling techniques," *Climate Dynamics*, vol. 46, no. 3-4, pp. 967–986, 2016.

[43] J. M. Gutiérrez, D. Maraun, M. Widmann et al., "An intercomparison of a large ensemble of statistical downscaling methods over Europe: results from the VALUE perfect

predictor cross-validation experiment," *International Journal of Climatology*, vol. 39, no. 9, pp. 3750–3785, 2019.

[44] X. Zhang, M. Shen, J. Chen, J. W. Homan, and P. R. Busteed, "Evaluation of statistical downscaling methods for simulating daily precipitation distribution, frequency, and temporal sequence," *Transactions of the ASABE*, vol. 64, no. 3, pp. 771–784, 2021.

[45] M. A. Sunyer, H. Madsen, and P. H. Ang, "A comparison of different regional climate models and statistical downscaling methods for extreme rainfall estimation under climate change," *Atmospheric Research*, vol. 103, pp. 119–128, 2012.

[46] D. A. Sachindra, K. Ahmed, M. M. Rashid, S. Shahid, and B. Perera, "Statistical downscaling of precipitation using machine learning techniques," *Atmospheric Research*, vol. 212, pp. 240–258, 2018.

[47] A. Davanlou Tajbakhsh, V. Nourani, and A. Molajou, "Hybrid wavelet-M5 modeling in rainfall-runoff process forecast," *Iran Water Resources Research*, vol. 15, no. 2, pp. 1–10, 2019.

[48] V. Nourani, M. T. Alami, and M. H. Aminfar, "A combined neural-wavelet model for prediction of Ligvanchai watershed precipitation," *Engineering Applications of Artificial Intelligence*, vol. 22, no. 3, pp. 466–472, 2009.

[49] V. Nourani, A. Molajou, S. Uzelaltinbulat, and F. Sadikoglu, "Emotional artificial neural networks (EANNs) for multi-step ahead prediction of monthly precipitation; case study: northern Cyprus," *Theoretical and Applied Climatology*, vol. 138, no. 3-4, pp. 1419–1434, 2019.

[50] V. Nourani, Z. Razzaghzadeh, A. H. Baghanam, and A. Molajou, "ANN-based statistical downscaling of climatic parameters using decision tree predictor screening method," *Theoretical and Applied Climatology*, vol. 137, no. 3-4, pp. 1729–1746, 2019.

[51] D. Avila, G. N. Marichal, I. Padrón, R. Quiza, and A. Hernandez, "Forecasting of wave energy in canary islands based on artificial intelligence," *Applied Ocean Research*, vol. 101, Article ID 102189, 2020.

[52] S. H. Vakili, "Forecasting of monthly precipitation using M5 model tree and classic statistical methods (Case study: oroumieh synoptic station) (Technical note)," *IRAN-WATER RESOURCES RESEARCH*, vol. 13, no. 4, pp. 179–183, 2018.

[53] V. Nourani and A. Molajou, "Application of a hybrid association rules/decision tree model for drought monitoring," *Global and Planetary Change*, vol. 159, pp. 37–45, 2017.

[54] J. Mahjoobi and A. Etemad-Shahidi, "An alternative approach for the prediction of significant wave heights based on classification and regression trees," *Applied Ocean Research*, vol. 30, no. 3, pp. 172–177, 2008.

[55] S. Abolfathi, A. Yeganeh-Bakhtiary, S. M. Hamze-Ziabari, and S. Borzooei, "Wave runup prediction using M5′ model tree algorithm," *Ocean Engineering*, vol. 112, pp. 76–81, 2016.

[56] D. Ji, L. Wang, J. Feng et al., "Description and basic evaluation of beijing normal university Earth system model (BNU-ESM) version 1," *Geoscientific Model Development*, vol. 7, no. 5, pp. 2039–2064, 2014.

[57] S. Lehner and H. Günther, "Extreme wave statistics from radar data sets," in *Proceedings of the 2004 IEEE International Geoscience and Remote Sensing Symposium*, pp. 1880–1883, Anchorage, AK, USA, September 2004.

[58] S. Wei, S. Yang, and D. Xu, "On accuracy of SAR wind speed retrieval in coastal area," *Applied Ocean Research*, vol. 95, Article ID 102012, 2020.

[59] P. A. E. M. Janssen, B. Hansen, and J.-R. Bidlot, "Verification of the ECMWF wave forecasting system against buoy and altimeter data," *Weather and Forecasting*, vol. 12, no. 4, pp. 763–784, 1997.

[60] R. P. da Rocha, S. Sugahara, and R. B. da Silveira, "Sea waves generated by extratropical cyclones in the South Atlantic Ocean: hindcast and validation against altimeter data," *Weather and Forecasting*, vol. 19, no. 2, pp. 398–410, 2004.

[61] S. Guinehut, C. Coatanoan, A.-L. Dhomps, P. Y. Le Traon, and G. Larnicol, "On the use of satellite altimeter data in argo quality control," *Journal of Atmospheric and Oceanic Technology*, vol. 26, no. 2, pp. 395–402, 2009.

[62] V. G. Polnikov, F. A. Pogarskii, N. S. Zilitinkevich, and A. A. Kubryakov, "Use of along-track altimeter data to verify numerical wave models," *Izvestiya - Atmospheric and Oceanic Physics*, vol. 55, no. 9, pp. 1089–1097, 2019.

[63] B. Oztunali Ozbahceci, A. R. Turgut, A. Bozoklu, and S. Abdalla, "Calibration and verification of century based wave climate data record along the Turkish coasts using satellite altimeter data," *Advances in Space Research*, vol. 66, no. 10, pp. 2319–2337, 2020.

[64] V. Nourani, A. Davanlou Tajbakhsh, A. Molajou, and H. Gokcekus, "Hybrid wavelet-M5 model tree for rainfall-runoff modeling," *Journal of Hydrologic Engineering*, vol. 24, no. 5, Article ID 04019012, 2019.

[65] S. Emmanouil, S. G. Aguilar, G. F. Nane, and J. J. Schouten, "Statistical models for improving significant wave height predictions in offshore operations," *Ocean Engineering*, vol. 206, Article ID 107249, 2020.