IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 2 May 2024, accepted 12 May 2024, date of publication 14 May 2024, date of current version 21 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3401038

RESEARCH ARTICLE

An Adaptive Strategy for Wind Speed Forecasting Under Functional Data Horizon: A Way Toward Enhancing Clean Energy

MUHAMMAD UZAIR[1], ISMAIL SHAH[2,1], AND SAJID ALI[1]
[1]Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan
[2]Department of Statistical Sciences, University of Padua, 35121 Padua, Italy

Corresponding author: Ismail Shah (ismail.shah@unipd.it)

This work was supported by Università degli Studi di Padova/University of Padua, Open Science Committee.

ABSTRACT An important issue in competitive energy markets is the accurate and efficient wind speed forecasting for wind power production. However, wind speed forecasting models developed for one location usually do not match the other site for various reasons like changes in terrain, different wind speed patterns, and atmospheric factors such as temperature, pressure, humidity, etc. Thus, introducing a flexible model that captures all the features is a challenging task. This paper proposes a functional data analysis (FDA) approach to forecast the site variant wind daily profiles with higher accuracy. Unlike the traditional methods, the FDA is more attractive as it forecasts a complete daily profile, and thus, forecasts can be obtained in the ultra-short period. To this end, the wind speed data is first filtered for extreme values. The filtered series is then divided into deterministic (Component-I) and stochastic (Component-II) components. Component-I is modeled and forecasted based on the generalized additive modeling technique. On the other hand, Component-II is modeled and forecasted using functional models such as functional autoregressive (FAR) and FAR with explanatory variables (FARX). For comparison purposes, forecasts from the traditional univariate autoregressive integrated moving average (ARIMA), seasonal ARIMA (SARIMA), SARIMA with exogenous information (SARIMAX), and neural network autoregressive (NNAR) models are also obtained. For empirical analysis, the wind speed data are obtained from the NASA power project for the site Canada located in Durham, England, and one-day-ahead out-of-sample forecasts are obtained for a complete year. The forecasting performance of different models is assessed through different accuracy measures, namely mean error, root mean squared error, mean absolute error, and mean absolute standard error. The results indicate that the functional models outperform the classical ARIMA, SARIMA, SARIMAX, and a deep learning model, NNAR. Within the functional models, the forecasting ability of the FARX is superior to FAR.

INDEX TERMS Wind speed, forecasting, time series models, machine learning models, functional models.

I. INTRODUCTION

Wind energy is one of the sustainable energy resources with significant growth potential and is considered a highly effective energy resource for improving the environment and atmosphere. Wind speed (WS) is directly related to wind power and can help meet the increasing electricity demand in developing and developed countries [1]. Improved WS forecasting is an efficient tool for expanding clean energy sources, which is a dire need in the current era. It directly affects agriculture, engineering, climate change, weather prediction, and the power sector. The prediction of wind flow and weather patterns is important for determining the location of wind turbines and constructing wind projects in the engineering field. Weather predictions such as rainfall, air pressure, temperature, precipitation, humidity, and tornado warnings can be easily made by studying WS profiles. Moreover, WS is a natural source of clean energy that

The associate editor coordinating the review of this manuscript and approving it for publication was Akin Tascikaraoglu.

is inevitable due to today's climate changes in terms of the concentration of carbon dioxide in the atmosphere. Figure 1 reflects a global map of the sporadic nature of WS profiles for different sites that indicates that WS profiles are substantially different for different sites.[1] Generally, the forecasting model of WS developed for one site does not suit the other for various reasons, like changes in terrain, different wind speed patterns, and atmospheric factors. However, the generalization capability of the forecasting model can be improved [2].

For efficient management, accurate wind speed forecasting is essential to integrate it into the scheduling and dispatch decisions of the power system operators. Moreover, the restructuring of the electricity industry worldwide automatically enhances the significance of wind power prediction to system operators and traders. For proper and balanced transfer of electricity under load management constraints, a precise forecast of power generation is important. The transmission system may not be able to transport all of the wind electricity as wind farms expand in size, and thus, the strain they exert on the grids likewise expands. Experience around the globe has shown that precise and trustworthy forecasting systems for wind energy are universally acknowledged for promoting clean energy. Short-term WS forecasting helps in planning the economic load dispatching decisions and is considered to be the most crucial forecasting horizon. The literature is very rich in the context of WS forecasting based on a variety of models. However, these models are generally site-dependent, which provides a vast space for revising and studying better and consistent models [3].

Due to the variability and unpredictability of wind, wind forecasting is an essential element in the operation and management of wind farms. Accurate wind forecasting can improve the reliability and efficiency of wind energy production, reduce operational costs, and avoid curtailment. To this end, various methods have been proposed for wind forecasting, including statistical, machine learning, hybrid, and numerical weather prediction models. Statistical models, such as time series models and artificial neural networks (ANNs), use historical wind data and weather forecasts to generate a statistical model that can predict future wind conditions and, thus, are extensively used for wind speed forecasting. For example, [4] and [5] suggested different models based on ANNs for hourly mean WS data collected from North Dakota. The forecasting performance of the ANN models, such as radial basis, adaptive linear element, and backpropagation, was superior to that of the classical models. For wind speed prediction, some studies used backpropagation neural networks by eliminating seasonal effects from actual WS datasets based on seasonal exponential adjustment [6], [7], [8]. To better understand the wind profiles, [9] used eight different strategies for multi-step forecasting based on machine learning methods and suggested three methods of merging combined data-driven input rectification and model output rectification (COMB-DIRMO) models whose results were better than the classical methods in the literature. One and two-days-ahead hourly mean WS forecasts obtained by [10] using a fractional ARIMA model provide better results than the persistent models. Machine learning techniques such as support vector regression (SVR), decision trees (DT), and random forests (RF) can handle non-linear and complex relationships in data and be trained using historical wind and weather data to provide short-term wind forecasts. For example, based on univariate time series data, [11] applied an SVR model and used the phase space reconstruction procedure for feature selection. The accuracy of the suggested SVR model is compared with autoregressive (AR), autoregressive moving average (ARMA), and ARIMA models by using Akaike's information criterion. Reference [12] suggested an SVR model for short-term wind speed modeling and forecasting based on rolling origin re-calibration. Conventional time series models such as ARIMA, seasonal ARIMA (SARIMA), and ARIMA-generalized autoregressive conditional heteroskedasticity (ARIMA-GARCH) and with weather information such as MLR, ARIMA, and SARIMA with exogenous variable (ARIMAX and SARIMAX) are also used for comparison purposes. The practical application of the said models is evaluated from a data set obtained from the Korea Meteorological Administration. The results suggest that SVR significantly outperforms the other models.

Current research trend in wind forecasting aims to improve the accuracy and reliability of wind forecasting models by integrating multiple forecasting methods and data sources. Hybrid forecasting methods that combine numerical and statistical models have been proposed to improve wind forecasting accuracy. For example, hybrid models used least squared support vector machine (LSSVM) and modified multi-objective optimization algorithms to model and forecast wind speed [13], [14], [15], [16], [17]. The suggested model forecasts the wind profiles more accurately as the parameters are estimated through particle swarm optimization based on simulated annealing. Reference [18] proposed a hybrid forecasting model for wind speed using weather research and forecasting (WRF), ARIMA, and wavelet neural networks. Reference [19] developed a hybrid forecasting model based on WRF and a long short-term memory (LSTM) neural network. For example, [20] proposed a real-time updating method for short-term wind power forecasting based on the WRF model and variational data assimilation techniques, aiming to improve the accuracy and timeliness of wind power forecasts. The volatility of patterns in WS is also analyzed by [21] and [22] by using a model comprising of a combination based on deep learning and long-short term memory networks which are useful to remember the long-term patterns and are appropriate for the turbulent nature of WS profiles. A combination of linear models to detect linear trends while applying machine learning models to estimate and predict non-linear trends is generally effective in forecasting problems. The literature is

[1] source: https://earth.nullschool.net

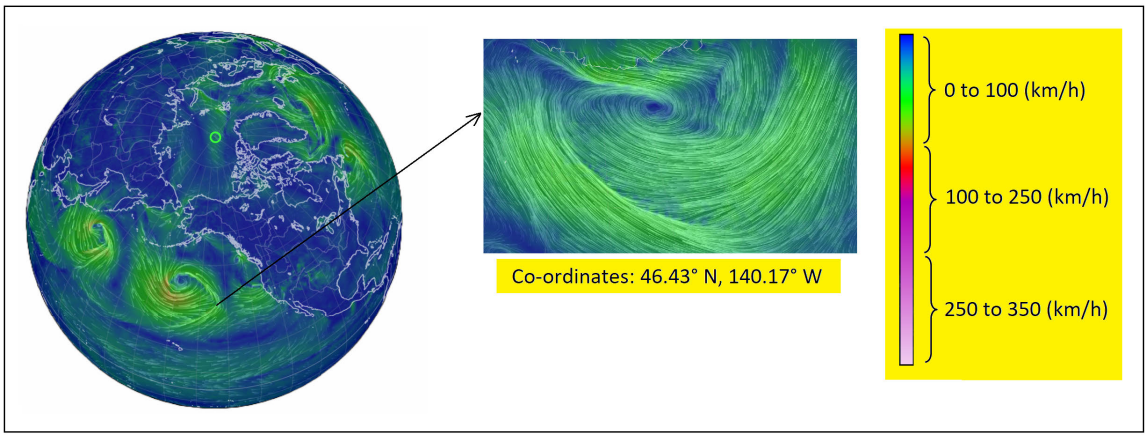

FIGURE 1. The site variant wind speed profiles around the globe.

very rich in the context of building such hybrid models [23], [24], [25], [26], [27]. The parameter estimation process becomes unreliable in the presence of extreme observation in time series. Different techniques are available in the literature to effectively handle outliers [28]. For example, [29] suggested the use of strong loss functions and regularization strategies, like ridge penalties and lasso penalties inside the Extreme Learning Machine (ELM) framework to lessen the impact of outliers in the training data, which improves the effectiveness and precision of ELM training, especially for wind speed forecasting. To deal with extremes and spurious noise in time series forecasting, [30] suggested approach uses a dynamic re-scaled lncosh loss function inside a neural network framework. Moreover, a flexible scaled Huber's loss function in an adaptive robust extreme learning machine (ARELM) model, which is a outliers handling technique, is used for wind speed forecasting to reduce the impact of outliers and data contamination [31].

Furthermore, a systematic review of some current studies is also discussed in Table 2 [32]. Reference [33] suggested an ensemble interval-valued method to forecast hourly WS data based on 1185 WS measurements from the site Shandong Penglai. For the same site, another short-term WS forecasting strategy, named the dynamic non-stationary fuzzy time series model, is developed by [34] to model and forecast a 10-minute interval WS data consisting of 3006 measurements, and the performance of the model is evaluated by error standard deviation (ESD), RMSE and MAPE. A comprehensive comparison among models including ARIMA, grey model (GM), and machine learning models including linear regression (LR), RF, and SVR, deep learning models such as ANN, LSTM, and convolutional neural network (CNN) is done by [35], which deals with a 15-minute interval data (1761 WS measurements from Gansu Province, China), is modeled to forecast WS. The study concludes that the CNN model outperforms the competitors based on MSE, RMSE, R^2, MAE, and MAPE. The generalized Gaussian process is adopted by [36] to model 13140 WS 6-hour interval measurements from Central China. An hour ahead WS forecast is suggested by [37], modeled by 730 hourly WS measurements from the site Synoptic Stations in Romania for 2016-2017 based on calibration models within the framework of generalized additive modeling (GAM). Another modified multi-objective optimization algorithm is developed by [38], in which short-term WS forecasting is done by modeling a 10-minute interval data including 2880 observations collected from the site Shandong Peninsula.

Due to its intermittent nature, WS data patterns reflect much more complex behavior. Precise forecasting can help to promote renewable energy which is environment friendly and can assist in operating the energy sector with an optimized decision regarding the power supply. Thus, this study makes an effort to track the site variant and intermittent nature of WS profiles with non-linear and non-stationary trends using the FDA approach. An appealing feature of the FDA approach is that, unlike the traditional methods, it provides forecasts for an entire day; hence, predictions can be obtained for any time within a day.

The rest of the article is structured as follows. Introduction to functional data analysis along with extreme values treatment is described in section II. Section III provides details about the general modeling framework. Within this section, the estimation and forecasting of functional models, along with their competitors, are also given. An empirical application of the proposed methodology is provided in section IV-C, where the results are discussed in detail. Finally, section V concludes the study.

II. FUNCTIONAL DATA ANALYSIS

In recent years, the FDA has become very popular due to its significant advantages based on data reduction and maintaining the identity of the data set. As data collection methods are much more refined nowadays, this area gets more attention due to its capability in the context of big data analysis. A functional datum is a curve or a

TABLE 1. List of all notations used in the manuscript.

Notation	Description	Notation	Description
P	Number of segments in which a time series is divided	N	Total number of observations
r	Width of window	t	Number of days
k	Number of hours in a day	w	Wind Speed (WS)
W_t	WS value at time t	$\hat{\lambda}$	Sample mean of WS
SD_W	Standard deviation of WS	$W_{t,k}$	WS at k^{th} day and t^{th}hour
$D_{t,k}$	Component-I	$S_{t,k}$	Component-II
$T_{t,k}$	Long trend	$Y_{t,k}$	Yearly seasonality
$U_{t,k}$	Indicator variable for seasonal effect	$\hat{D}_{t+1,k}$	Day-ahead forecast of Component-I
$\hat{S}_{t+1,k}$	Day-ahead forecast of Component-II	$\hat{W}_{t+1,k}$	Day-ahead forecast of WS
$\hat{T}_{t,k}$	Forecasted value of $T_{t,k}$	$\hat{Y}_{t,k}$	Forecasted value of $Y_{t,k}$
$\hat{U}_{t,k}$	Forecasted value of $U_{t,k}$	n	Number of seasons
η_i	Parameter of indicator variable $U_{t,k}$	υ	Number of Fourier Basis Functions
$\xi_{t,j}(\upsilon)$	Fourier Basis Functions	b_j	Parameters of $\xi_{t,j}(\upsilon)$
$S_t^{\star}(\upsilon)$	Functional curves	$S_t(\upsilon)$	Centralized functional curves
$\lambda(\upsilon)$	Functional mean curve	$\mathcal{H}$	Hilbert Space
$C(\upsilon,\gamma)$	Covariance operator	$\alpha_j(t)$	Functional principal component (FPC)
d_{k1}	First FPC score	$\alpha_j(\upsilon)$	j^{th} FPC
δ_j	j^{th} Eigenvalue	$\theta_{j,t}$	j^{th} FPC score
δ_j	Variance of j^{th} FPC score	R	Number of retained FPC
$\eta(\upsilon)$	Error function	$\hat{\delta}_j$	Sample eigenvalues
$\hat{\alpha}_j$	Orthogonal sample eigen functions	$\Upsilon_t(\upsilon)$	Error term of FAR(p) model
$\oint_f^{(\rho)}$	Scalar exogenous variable	$\lambda_f^{(\rho)}(t)$	Time invariant mean functions
$\oint_f^{(\rho)}(t)$	Functional Exogenous variable	$\eta_t(\upsilon)$	Error term of FARX(p) model
m	Number of dimensions	p	Number of lag
q	order of Moving Average (MA) process	d	Number of differences
$\alpha^{'}$	Intercept term of ARIMA model	β_j	Parameter of AR process
ϕ_j	Parameter of MA process	w_0 and w_{0j}	Biases on the nodes
w_j and w_{ij}	Connection weights between the layers of NNAR model	$f(\cdot)$	Activation function of the hidden layer
$y_{t-p^{\star}}$	Input nodes of NNAR model	y_t	Output nodes of NNAR model
$p^{\star}$	Number of input nodes	$q^{'}$	Number of hidden nodes
ϵ_t	White noise process of NNAR model	σ^2	Variance of ϵ_t

TABLE 2. Systematic review of recent studies.

No.	Authors	Location	Data Type	No. of Observations/Years	Methods	Accuracy Measures
1	Yan Hao et al. (2024)	Shandong Penglai	Hourly	1185 (Confidential)	Ensemble interval-valued	MAE, RMSE, MAPE
2	Shi et al. (2024)	Shandong Penglai	10 minute	3006 (2021.01.02 to 2021.01.21)	Dynamic Non-stationary Fuzzy Time Series Model	ESD, RMSE, MAPE
3	Li et al. (2023)	Gansu province	15 minute	1761 (2016.01.01 to 2016.01.19)	ARIMA, GM, LR, RF, SVR, ANN, LSTM, CNN	MSE, RMSE, R^2, MAE, MAPE
4	Heng et al. (2022)	Central China	6 hours	13140 (Confidential)	Generalized Gaussian Process	MAE, RMSE
5	Brabec et al. (2021)	Synoptic stations in Romania	Hourly	730 (2016-2017)	Calibration models within the framework of GAM	RMSE
6	Liu et al. (2020)	Shandong Peninsula	10 minute	2880 (Confidential)	Modified multi-objective optimization algorithm	MAE, RMSE, MAPE

surface varying over a continuum. Since a function has infinite points, the functional datum also lives in an infinite dimensional space. However, in practice, these data are observed on discrete points. Functional data may or may not be independent of each other and are useful because the derivatives are available for further analysis. Since it's a curve and not like a scalar quantity, being a single datum, the problem of multicollinearity is automatically resolved. It also solves the problem of high dimensionality and removes the noise from the data. Furthermore, it utilizes the inherent smoothness of the data. This technique was first introduced by [39]. The applications of the FDA are available in every field, including energy, medical, environmental, engineering, etc. For example, FDA is extensively used for modeling and forecasting electricity demand and prices [40], [41], [42], [43], air temperature and precipitation [44], [45], transportation [46] and so on.

A. WIND SPEED FORECASTING BASED ON FDA

Since the WS profiles contain non-stochastic and non-linear patterns like extreme measurements, long-term patterns, site

variant seasonality, and irregular behavior in any short span, modeling and forecasting WS profiles are challenging. When a model is trained using a small data set, the results tend to vary, and forecasting is generally unreliable. However, the usage of big data requires a suitable reduction technique which refers to the loss of important features of the data in terms of information. Thus, such a deficiency can be controlled using the FDA. This adaptive technique is helpful for forecasting long-period WS profiles and focuses on short-term forecasting accuracy. Wind speed profiles are intermittent and have more complex non-linear patterns, such as extreme wind speed at different time horizons, as depicted in Figure 2. The figure indicates multiple traits, such as long-term trends, seasonal periodicity, and changes in daily profiles. Incorporating such characteristics into the model substantially refines forecasting accuracy. To this end, the original WS time series is prepossessed for extreme values before modeling.

B. HANDLING EXTREME WS MEASUREMENTS

Extreme measurements are commonly present in the time series of WS, which directly makes the estimation procedure challenging and significantly affects the forecasting results. Thus, identifying and replacing extreme measurements with optimal values improves forecasting accuracy in such cases. Figure 3 presents the hourly box plots of WS plotted against each hour, indicating that there are many extreme measurements in the WS series. Identifying and replacing such extreme values can be made using several techniques [28]. In our work, extreme observations are replaced by the shifting filter on the WS (SFWS) strategy suggested by [47]. The SFWS is the generalization of the standard deviation filter on the WS (SDFWS), which operates on a rolling window of a fixed width rather than the entire time series. Using the SFWS, the first step divides the time series into $P = (N/r)$ segments, where N is the total number of observations in the series, and R denotes the window's width. Within each segment, the WS values whose absolute deviation taken from their sample mean $\hat{\lambda}$ is greater than some multiple (e.g., 1.96 in case of using normal distribution 95% confidence interval) of the sample standard deviation (SD) are classified as extreme values. The window then advances r observations, and the filtering procedure is repeated to identify the extreme values in the following window. The process is repeated until all P portions have been addressed using this approach. The subset, W_t^0, containing WS extreme values, with the rolling window of width r, is obtained as

$$W_t^0 = \bigcup_{i=1}^{p'} \{W_t : |W_t - \hat{\lambda}| \geq 1.96 * SD_W\}. \tag{1}$$

Here, W_t denotes the wind speed value at time t, $\hat{\lambda}$ represents the sample mean, and SD_W is the standard deviation of wind speed for each segment. Finally, the subset W_t^0 are the identified extreme values in the specific segment.

Once the extreme values are identified, they are replaced by average values using different methods. This work, in particular, replaces the extreme values with the median value of the associated window.

III. MODELING FRAMEWORK

Consider the WS profile time series is $W_{t,k}$ where $t \in N$ and $k = 1, 2, 3, \ldots, 24$. The WS profile is split into two components, namely Component-I and Component-II, denoted by $D_{t,k}$ and $S_{t,k}$ respectively, i.e.,

$$W_{t,k} = D_{t,k} + S_{t,k} \tag{2}$$

where Component-I refers to long pattern $T_{t,k}$ and yearly seasonality $Y_{t,k}$. This can be decomposed in the following equation:

$$D_{t,k} = T_{t,k} + Y_{t,k} \tag{3}$$

As the seasons vary region-wise, building an adaptive and portable model is the key to this study, which is achieved by capturing the seasonal effect in terms of an indicator variable. As narrated earlier, the WS profiles are site variants, so this factor is accommodated to build the suggested model as an adaptive measure for all sites. The seasons are studied for the corresponding sites and a dummy variable $U_{t,k}$ is included to incorporate such variation. Thus, equation (3) is reformed as under:

$$D_{t,k} = T_{t,k} + Y_{t,k} + U_{t,k} \tag{4}$$

The component $D_{t,k}$ in equation (4) is estimated and forecasted using GAM techniques while $S_{t,k}$ is based on the FDA, namely FAR(p) and FARX(p) models. Furthermore, classical ARIMA and NNAR models are also used as benchmark models. After modeling both components, the final forecast is obtained as follows

$$\hat{W}_{t+1,k} = \hat{D}_{t+1,k} + \hat{S}_{t+1,k} \tag{5}$$

A. AN ADAPTIVE STRATEGY FOR MODELING COMPONENT-I

The GAM technique is used to model Component-I, which comprises long-term patterns $T_{t,k}$, yearly $Y_{t,k}$, and seasonal $U_{t,k}$ periodicities. The time series $T_{t,k}$ $(1, 2, 3, 4, \cdots)$ which is a function of time t, and $Y_{t,k}$ (1, 2, 3, 4, …., 365, 1, 2, 3, 4, …., 365, 1, 2, 3, 4, …., 366), 365 for a normal year while 366 for a leap year (if any), are modeled through smooth functions. Furthermore, the seasonal site variant accommodator (SSVA), $U_{t,k}$ such that $U_{t,k} = \sum_{i=1}^{n} \eta_i \, T_{t,k}$ where n is a number of the seasons of any respective site that might be studied. Also, $T_{t,k} = 1$ where k refers to the *ith* season of the year, or zero otherwise, and the parameter η_i is estimated by the ordinary least squares (OLS). When the Component-I is modeled, a next-day forecast is obtained for $\hat{T}_{t+1,k} = \hat{T}_{t,k}$, $\hat{Y}_{t+1,k} = \hat{Y}_{t,k}$ and $\hat{U}_{t+1,k} = \hat{U}_{t,k}$. Thus, a day-ahead forecast for Component-I is achieved

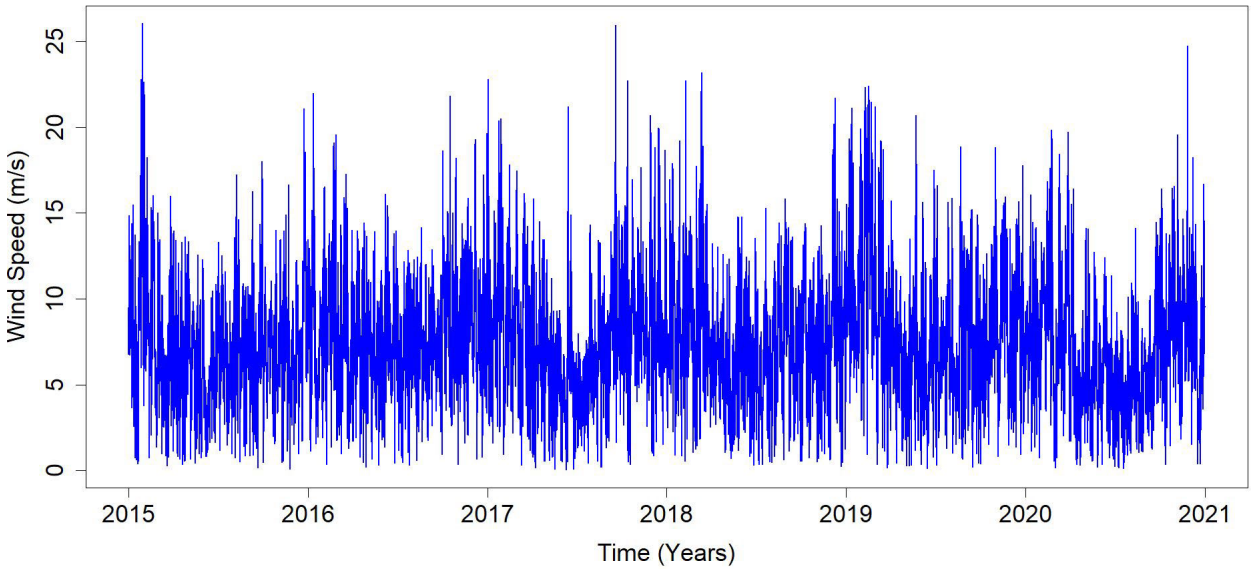

FIGURE 2. Time series plot of WS for the years 2015-2021.

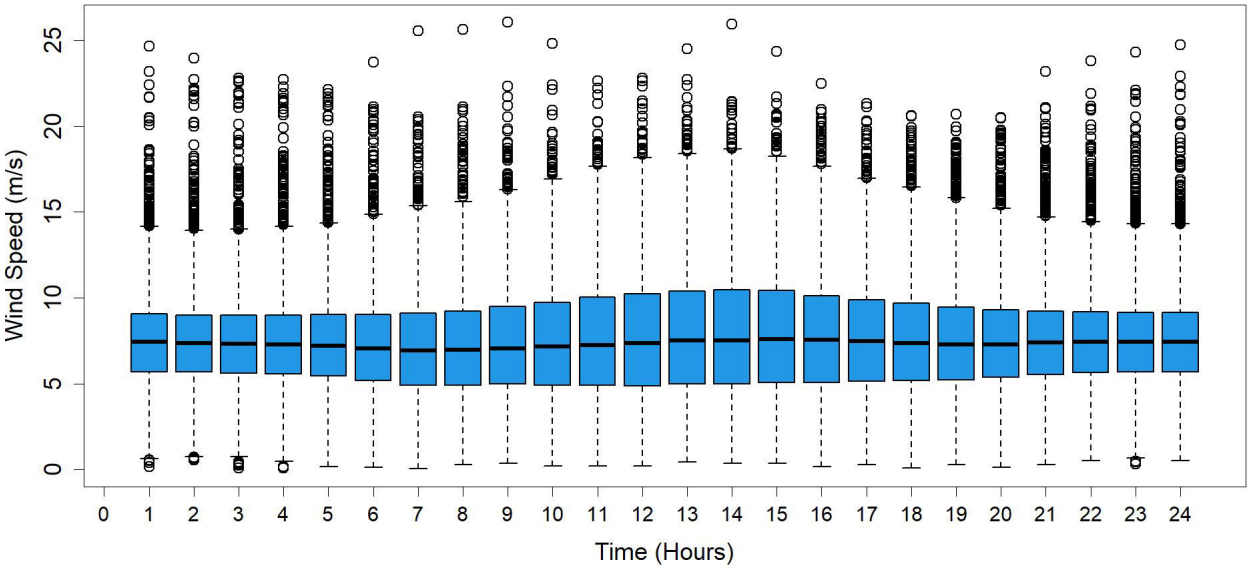

FIGURE 3. Hourly box-plots of canadian WS.

as follows:

$$\hat{D}_{t+1,k} = \hat{T}_{t,k} + \hat{Y}_{t,k} + \hat{U}_{t,k}. \tag{6}$$

The deduction of forecast Component-I leads to obtaining the intermittent part of WS profiles called Component-II $S_{t,k}$ as

$$S_{t,k} = W_{t,k} - \hat{D}_{t,k}. \tag{7}$$

B. A ROBUST STRATEGY FOR MODELING COMPONENT-II

This section provides details on the estimation of Component-II, i.e., $S_{t,k}$, using the models FAR(p) and FARX(p). In addition, this section gives insight into the functional methodology by explaining the functional principle component analysis technique.

1) THE FDA PARADIGM

FDA refers to the process where the discretized measurements (with or without equal intervals) are converted into curves based on certain period profiles (e.g., daily, monthly, or yearly). Thus, to work within the FDA framework, we must convert the profile into a functional object using a set of building blocks known as a basis function system (BFS) [42]. Using the FDA, big data is transformed into small data without losing any information. The BFS is crucial

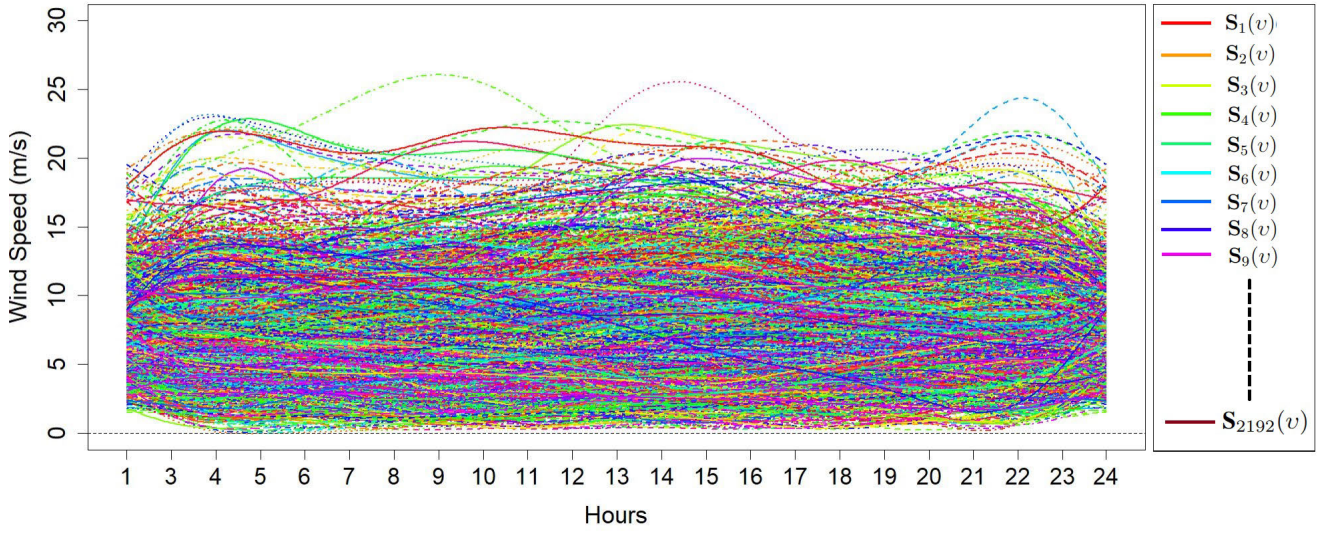

FIGURE 4. Daily functional trajectories of wind speed data measured for the period 2015-2021.

in converting WS profiles into functional data. Hence, $S_{t,k}$ is converted to a functional form referring to the daily WS profile for the k^{th} day, using Fourier basis functions (FBF), which is a linear combination of sine and cosine waves with increasing frequency and generally used for periodic data. The FBF captures the curves very well and is also useful in estimating their derivative, which is another tool for measuring the change in the underlying curves. Mathematically, the functional data is obtained as

$$S_t(\upsilon) = \sum_{j=1}^{\kappa} b_j \xi_{t,j}(\upsilon),\ \upsilon \in \aleph \tag{8}$$

where b_j, $(j = 1, 2, 3, \ldots, \kappa)$ are the parameters and $\xi_{t,j}(\upsilon)$ are the FBF. An example of functional data of 2192 daily WS profiles is plotted in Figure 4, where each curve represents a functional wind speed profile for a single day. This time series is known as the functional time series (FTS) where one can investigate the major source of variation among these curves.

Now, assume that $S_t(\upsilon)$ given in equation (8) is an arbitrary FTS defined on a common probability space (ψ, ω, π), where ψ, ω, and π denote the sample space, σ algebra on ψ, and the probability measure on ω, respectively. In general, it is assumed that the functions $S_t(\upsilon)$ are elements of the semi-metric and Hilbert space. Furthermore, it is assumed that functions are the elements of the square-integrable functions $S \in \wp^2[0, 1]$ residing in the Hilbert space $\mathcal{H}$ satisfying $||S_t||^2 = \int S^2(\upsilon)d\upsilon < \infty$ with an inner product $(u, v) = \int u(\upsilon)v(\upsilon)d\upsilon \ \ \forall\ u, v \in \wp^2[0, 1]$. The notation $S \in \wp^{\iota}_{\mathcal{H}}(\psi, \omega, \pi)$ is used to indicate for some $\iota > 0$, $E(||S_t||^{\iota}) < \infty$. Note that $\iota = 1$ results in the population mean curve $\lambda(\upsilon) = E(S(\upsilon))$, and $\iota = 2$ refers to the non-negative definite covariance operator given as

$$\begin{aligned} C(\upsilon, \gamma) &= Cov[S(\upsilon), S(\gamma)] \\ C(\upsilon, \gamma) &= E[\{S(\upsilon) - \lambda(\upsilon)\}\{S(\gamma) - \lambda(\gamma)\}]. \end{aligned} \tag{9}$$

The covariance operator $C(\upsilon, \gamma)$ in equation (9) allows the covariance operator of S, denoted by ϱ that can be written mathematically as

$$\varrho(\upsilon, \gamma) = \int_1^0 C(\upsilon, \gamma)\alpha(\upsilon)d\upsilon \tag{10}$$

2) FUNCTIONAL PRINCIPAL COMPONENT ANALYSIS

Suppose we have an FTS with k curves denoted by $S_1^{\star}(t), S_2^{\star}(t), \ldots, S_k^{\star}(t)$. To study the variation among the functional curves, we subtract the mean curve first, say $\lambda(t)$, i.e., $S_i(t) = S_i^{\star}(t) - \lambda(t)$ where $\lambda(t) = \frac{1}{\nu}\sum_{i=1}^{k} S_i^{\star}(t)$ to get centralized curves $S_i(t)$. Suppose the first FPC is $\alpha_j(t)$, we project the curves $S_i(t)$ to the first FPC $\alpha_j(t)$ so that the projection is the inner product of two functions $S_i(t)$ and $\alpha_j(t)$ where the inner product refers to the integral of the product of $S_i(t)$ and $\alpha_j(t)$ in functional space. This projection is called the first FPC score denoted by d_{k1} mathematically expressed as $d_{k1} = \int S_i(t)\alpha_j(t)dt$. Our objective is to maximize the variation for the first FPC score. As $S_i(t)$ is a centralized curve, the sample variance of d_{k1} will be the sum of the square of d_{k1}, i.e., we want to maximize $\sum d_{k1}^2$. To make d_{k1} identifiable, we added a constraint that $\int \alpha_j^2(t) = 1$. Hence, d_{k1} represents the strongest and most important mode of variations from the k curves. In this manner, we obtained the rest of the j FPCs such that all FPCs are orthogonal to each other.

To explore the major source of variation among WS profiles, we need to find some top j number of FPCs which is generally a small number. Hence, these top FPCs describe the major source of variation among multiple WS profiles $S_k(t)$. Thus, the curve $S_k(t)$ is projected to FPC scores for dimension reduction that projects the infinite function $S_k(t)$ to finite FPC scores. Using Mercer's lemma, there is an orthogonal sequence α_j of continuous functions in $\wp^2[0, 1]$ and a non-increasing sequence δ_j of positive numbers,

such that

$$C(\upsilon,\gamma)=\sum_{j=1}^{\infty}\delta_j\alpha_j(\upsilon)\alpha_j(\gamma) \tag{11}$$

where $\alpha_j(\upsilon)$ denotes the j^{th} FPC and δ_j denotes the j^{th} eigenvalue in the decreasing order [48]. By the separability of Hilbert spaces, the Karhunen–Lo'eve (KL) expansion [49] of a random process $S_t(\upsilon)$ is given as

$$S_t(\upsilon)=\lambda(t)+\sum_{j=1}^{\infty}\theta_{j,t}\alpha_j(\upsilon)\alpha_j(\gamma) \tag{12}$$

where $\theta_{j,t}$ is the j^{th} FPC score defined as $\theta_{j,t}=\int\chi(\upsilon)\alpha_j(\upsilon)d\upsilon_t$. The principal component scores (PCSs) comprise an uncorrelated sequence of random variables with zero mean and variance δ_j. The expansion facilitates dimension reduction as the first R terms often provide a good approximation to the infinite sums. Thus, the information contained in $S_t(\upsilon)$ can be adequately summarized by the R-dimensional vector $(\delta_1,\delta_2,\ldots,\delta_R)$. The approximated processes can be defined as follows:

$$S_t(\upsilon)=\sum_{j=1}^{R}\theta_{j,t}\alpha_j(\upsilon)+\eta(\upsilon) \tag{13}$$

where R denotes the number of retained principal components, and $\eta(\upsilon)$ denotes the error function with zero mean and finite variance. More details about the FPCA can be obtained by consulting [50], [51]. In practice, the mean curve $\lambda(\upsilon)$, FPCs $\alpha(\upsilon)$, and PCSs θ_j can only be estimated through realizations of a random process. Suppose $S_1(\upsilon),S_2(\upsilon),S_3(\upsilon),\ldots,S_t(\upsilon)$ are fully observed FTS, then the estimators for the mean function $\lambda(t)$ and covariance operator $C(\upsilon,\gamma)$ are given as

$$\hat{\lambda}(\upsilon)=\frac{1}{\nu}\sum_{t=1}^{\nu}S_t(\upsilon) \tag{14}$$

$$\hat{C}(\upsilon,\gamma)=\sum_{j=1}^{\infty}\hat{\delta}_j\hat{\alpha}_j(\upsilon)\hat{\alpha}_j(\gamma) \tag{15}$$

where $\hat{\delta}_1,\hat{\delta}_2,\hat{\delta}_3,\ldots.\geq 0$ are the sample eigenvalues of $C(\upsilon,\gamma)$ and $[\hat{\alpha}_1,\hat{\alpha}_2,\hat{\alpha}_3,\ldots.]$ are the corresponding orthogonal sample eigen functions. Reference [52] studied the estimators mentioned above and concluded that they are consistent estimators for the weekly dependent processes. Using the KL expansion, the realizations of the random process $S_t(\upsilon)$ can be written as follows:

$$S_t(\upsilon)=\hat{\lambda}(\upsilon)+\sum_{j=1}^{R}\hat{\theta}_{j,t}\hat{\alpha}_j(\upsilon)+\eta(\upsilon), \tag{16}$$

3) GENERALIZED FUNCTIONAL AUTOREGRESSIVE MODELS

The autoregressive (AR) models are well-known linear models where the response variable is regressed over its lagged values plus a noise term. If the response variable is functional, the AR model is called functional autoregressive (FAR) which is one of the best models for modeling the complex nature of FTS. The optimum value of lag order denoted by p is crucial to generalize the FAR model as FAR(p). Mathematically, a FAR(p) model can be written as

$$S_t(\upsilon)=\lambda(\upsilon)+\sum_{i=1}^{p}\phi_i S_{t-i}(\upsilon)+\Upsilon(\upsilon) \tag{17}$$

where $\Upsilon_t(\upsilon)\in\nu$ is i.i.d. in $\wp^2_{\mathcal{H}}$. In addition, $E(\Upsilon_t)=0$ and ϕ_i is linear operator bounded over mapping $\wp^2_{\mathcal{H}}\rightarrow\wp^2_{\mathcal{H}}$ such that equation (17) has a unique solution. To increase the forecasting accuracy, the FAR model can include other exogenous variables that may be scalars, vector-valued, or functional, and thus, denoted by FARX(p). For given trajectories $(S_1,S_2,\ldots,S_t)$ and exogenous variables $\oint_f^{(1)},\oint_f^{(2)},\ldots,\oint_f^{(\iota)}$, the goal is to then derive an empirical $\hat{S}_{t+1}$. The time-invariant mean functions $\lambda_f^{(1)}(t),\lambda_f^{(2)}(t),\ldots,\lambda_f^{(\rho)}(t)$ are assumed to be stationary for the functional exogenous variables $\oint_f^{(1)}(t),\oint_f^{(2)}(t),\ldots,\oint_f^{(\rho)}(t)$. Mathematically, a FARX(p) model is defined as

$$S_t(\upsilon)-\lambda(t)=\sum_{l=1}^{P}\psi_l(\upsilon)[S_{t-l}(\upsilon)-\lambda(\upsilon)]+\sum_{m=1}^{\rho}\phi_m(\upsilon)[\oint_f^{(m)}(\upsilon)-\lambda_f^{(m)}(t)]+\eta_t(\upsilon) \tag{18}$$

where $\eta_t(\upsilon)$ is the functional error term and $\psi_l(\upsilon)$, $\phi_m(\upsilon)$ are the functional operators of the model. It is important to note that scalar exogenous variables may also have some causal influence on the dynamics of the functional response, in addition to the lagged functional exogenous variables. Assume that we have ρ scalar exogenous variables $\oint_f^{(1)},\oint_f^{(2)},\ldots,\oint_f^{(\rho)}$ with mean $\lambda_f^{(1)},\lambda_f^{(2)},\ldots,\lambda_f^{(\rho)}$, then the FARX($p$) model can be expressed as follows:

$$\chi_t(\upsilon)-\lambda(t)=\sum_{l=1}^{P}\psi_l(\upsilon)[S_{t-l}(\upsilon)-\lambda(\upsilon)]+\sum_{m=1}^{\rho}\phi_m[\oint_f^{(m)}-\lambda_f^{(m)}]+\eta_t(\upsilon)\,. \tag{19}$$

The FARX(p) models in equations (18) and (19) are the generalizations of the model presented in equation (17) with various numbers of functional or scalar exogenous variables, respectively. The hierarchy of modeling and forecasting of equation (17) based on equations (14)-(16) and equation (18) is adopted which can be visualized in Figure 5. The adopted process is based on obtaining the FPCA score and FPCA values by fixing the dimension m. Based on Durbin–evinson and the innovations algorithm, the vector autoregressive (VAR) model is applied by fixing the lag p on the FPCA scores. The unexplained variation from the FPCA values and the residual from the VAR model is used to get optimum

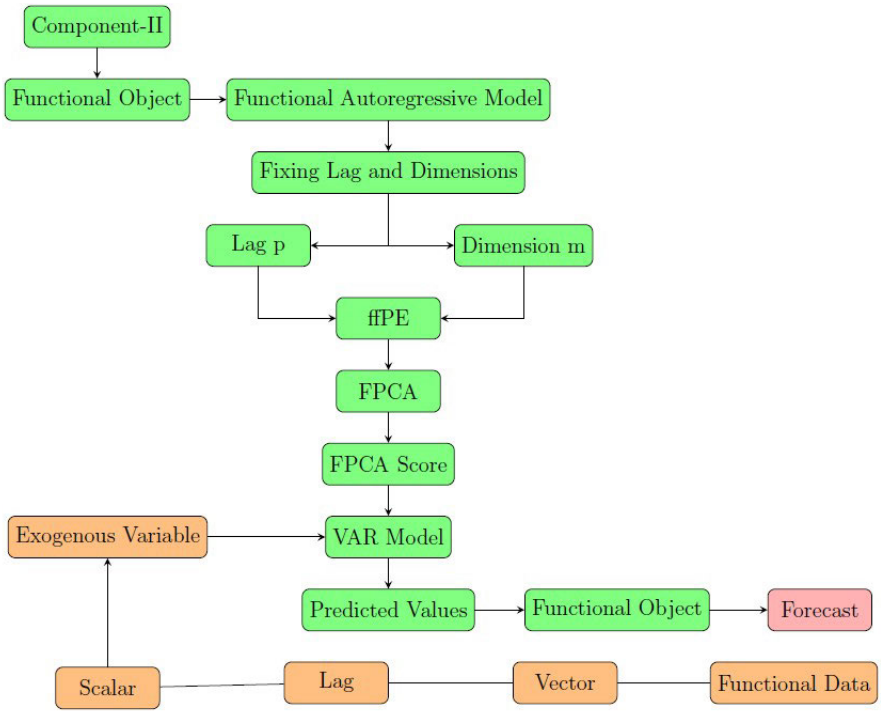

FIGURE 5. Flowchart for modeling and forecasting FAR(p) and FARX(p).

values of m and p, which minimizes the ffPE in equation (20) as explained in the following.

4) AN OPTIMUM VALUES OF p AND m BASED OF ffPE

Choosing an appropriate value of lag p and dimension m is a crucial part of forecasting FTS using the adaptive technique. The literature is rich in recommending these values for the FAR model. For example, the value of m for the FAR was suggested by [53], which is based on multistage hypothesis testing. Selecting an optimum number of FPCs for which the proportion of variance explained by them exceeds a certain threshold is a usual approach. In this context, a mechanical and automated technique is suggested for getting an optimum p and m for the FAR and FARX model, which is based on minimizing the MSE as studied by [54], called functional final prediction error (ffPE). Mathematically, the ffPE is given as

$$ffPE(p,m) = \frac{t + p \times m}{t - p \times m} Trace(\hat{\vartheta_\omega}) + \sum_{j>m} \hat{\delta}_j \tag{20}$$

where $\hat{\delta}_j$ is the j^{th} eigenvalue of $C(\upsilon, \gamma)$, $\hat{\vartheta_\omega}$ in (18) is the covariance matrix of the random vector $(\omega_1, \omega_2, \ldots, \omega_m)$ and is an unbiased estimator of ϑ_ω. The values of m and p that minimize the ffPE are the optimal values that can be used in the analysis.

C. COMPETITIVE MODELS

For a comprehensive comparison, the proposed models are compared with the traditional ARIMA, SARIMA, SARIMAX, and a deep learning method, the neural network autoregressive model (NNAR), to assess their forecasting performance [12]. The details for each model are as follows.

1) AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

The ARIMA is one of the most commonly applied time series models for forecasting a univariate time series. ARIMA is the extended form of the simple ARMA model being a tool that extrapolates the signal into the future to get forecasts based on separating signal from noise. It comprises three parametric parts known as the order of AR, MA, and the number of differences needed for time series to be stationary, i.e., p, q, and d, respectively. Mathematically, it can be written as

$$\begin{aligned} W_{t,k} = \alpha' + \beta_1 W_{t-1,k} + \beta_2 W_{t-2,k} + \ldots. + \beta_p W_{t-p,k} + \upsilon_t \\ + \phi_1 \upsilon_{t-1} + \phi_2 \upsilon_{t-2} + \ldots. + \phi_q \upsilon_{t-q} \end{aligned} \tag{21}$$

where $W_{t,k} = S_{t,k} - S_{t-d,k}$ is a d times differences series, α' is an intercept term, and β_i $(i = 1, \cdots, p)$ and ϕ_j $(j = 1, \cdots, q)$ are the parameters of AR and MA part, respectively, that are estimated using the maximum likelihood estimation (MLE) method.

2) SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

The SARIMA model is an extended form of ARIMA, a powerful technique designed for capturing seasonal patterns in time series data. It can identify both short- and long-term connections in the data, integrating the ideas of the ARIMA

model with seasonal elements. Since WS profiles have seasonal features that change over time, the seasonal and non-seasonal patterns are simultaneously included in the series, which can lower the performance of ARIMA in the wind power forecasting model. Mathematically, the SARIMA can be written as

$$\begin{aligned} W_{t,k} = \alpha' &+ \beta_1 W_{t-1,k} + \beta_2 W_{t-2,k} + \ldots + \beta_p W_{t-p,k} + \upsilon_t \\ &+ \phi_1 \upsilon_{t-1} + \phi_2 \upsilon_{t-2} + \ldots + \phi_q \upsilon_{t-q} \\ &+ \Theta_1 W_{t-s,k} + \Theta_2 W_{t-2s,k} + \ldots + \Theta_P W_{t-Ps,k} \\ &+ \Phi_1 \upsilon_{t-s} + \Phi_2 \upsilon_{t-2s} + \ldots + \Phi_Q \upsilon_{t-Qs} \end{aligned}$$

where $\Theta_i's$ and $\Phi_i's$ are the parameters of seasonal AR and MA processes, respectively. The model is denoted by SARIMA(p, d, q)(P, D, Q)s, where (P, D, Q) stands for the seasonal order, (p, d, q) for the non-seasonal order (the same as the ARIMA model), and s for the seasonal cycle length. The parameters of the model are generally estimated using the MLE method.

3) SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL WITH EXOGENOUS VARIABLE

The SARIMAX model is an expanded version of the SARIMA with the ability to capture the impact of exogenous variables in the time series data, leading to improved forecasting accuracy and providing a better understanding of the underlying dynamics of the process in many cases. Mathematically, it can be written as

$$\begin{aligned} W_{t,k} = \alpha' &+ \beta_1 W_{t-1,k} + \beta_2 W_{t-2,k} + \ldots + \beta_p W_{t-p,k} + \upsilon_t \\ &+ \phi_1 \upsilon_{t-1} + \phi_2 \upsilon_{t-2} + \ldots + \phi_q \upsilon_{t-q} \\ &+ \Theta_1 W_{t-s,k} + \Theta_2 W_{t-2s,k} + \ldots + \Theta_P W_{t-Ps,k} \\ &+ \Phi_1 \upsilon_{t-s} + \Phi_2 \upsilon_{t-2s} + \ldots + \Phi_Q \upsilon_{t-Qs} \\ &+ \beta' X_t \end{aligned}$$

where X_t denotes the exogenous variable at time 't' and β' represents the coefficient of X_t. The parameters estimation is done using the MLE approach.

4) NEURAL NETWORK AUTOREGRESSIVE MODEL

Traditional models such as ARIMA become less efficient in accommodating a large set of input for training purposes, especially in cases where there are complicated interactions or long-term reliance. Classical theories exploit the internment nature of complex structured profiles as they are directly concerned with the strong assumption about data. Artificial neural networks (ANN) allow complex nonlinear relationships between the response variable and its predictors. The strength of the Neural Network Autoregressive (NNAR) Model comes from the parallel processing of data, which eliminates the need for classical assumptions. As a result, the network model may be easily chosen based on the characteristics of the data. ANN includes three types of layers, namely input, output, and one or more hidden layer(s) with an activation function that determines the relationship (represented by sigmoid) between input $(y_{t-1}, y_{t-2}, \ldots, y_{t-p^\star})$ and output (y_t) of a node and network. The mathematical form of the NNAR model is:

$$y_t = \omega_0 + \sum_{j=1}^{q'} \omega_j f\left(\omega_{0j} + \sum_{i=1}^{p^\star} \omega_{ij} y_{t-i}\right) + \epsilon_t \qquad (22)$$

where w_0 and w_{0j} are the biases on the nodes, w_j and w_{ij} $(j = 1, 2, \cdots q', i = 1, 2, \cdots, p^\star,)$ are the connection weights between the layers of the model, $f(\cdot)$ is the activation function of the hidden layer, $p^\star$ is the number of input nodes, q' is the number of hidden nodes, and ϵ_t is a white noise process, i.e., $\epsilon_t \overset{\text{iid}}{\sim} N\left(0, \sigma^2\right)$. Furthermore, the sigmoid is based on the logistic function given as $f(x) = \left(1 + e^{-x}\right)^{-1}$.

IV. OUT-OF-SAMPLE FORECASTING

This section provides details about the data used in this study. It also discusses different accuracy measures and the results related to out-of-sample forecasting obtained through different models.

A. CANADIAN WIND SPEED DAILY PROFILES

To study the intermittent nature of WS profiles, the hourly wind speed data are collected for the site Canada, Durham in England from the website of NASA Power[2] The forecasting capability of the adaptive strategy is evaluated by the WS profiles measured for the period 2015-2021, consisting of 52,608 equidistant WS hourly measurements at 50 meters (m/s). A sample of ten daily profiles (24 hourly measurements per day) are plotted in Figure 6. The figure shows considerable variation among the daily profiles at different hours of the day.

For estimation purposes, the hourly WS measurements are divided into two parts, i.e., the training set contains 43,848 hourly measurements ranging from January 1, 2015, to December 31, 2020, while the test set has 8760 hourly measurements ranging from January 1, 2021, to December 31, 2021. The training set is used for model estimation, while the testing set is used for one-day-ahead out-of-sample forecasts. The modeling procedure is based on expanding window methodology and the twenty-four hours ahead forecasts are calculated for the entire year 2021. The data also show seasonal variation throughout the year. A sample plot of the WS profiles for the year 2015 is plotted for different seasons in Figure 7. The figure indicates that the WS profile has more variation in winter than in other seasons.

B. ACCURACY MEASURES FOR VALIDATION PURPOSE

To validate the accuracy of the forecasting models, different error measures are used. In particular, the forecast accuracy of a model is determined by the mean error (ME), mean absolute error (MAE), root-mean-squared error (RMSE), and mean absolute standard error (MASE) which are described in the following.

[2]https://power.larc.nasa.gov/data-access-viewer/

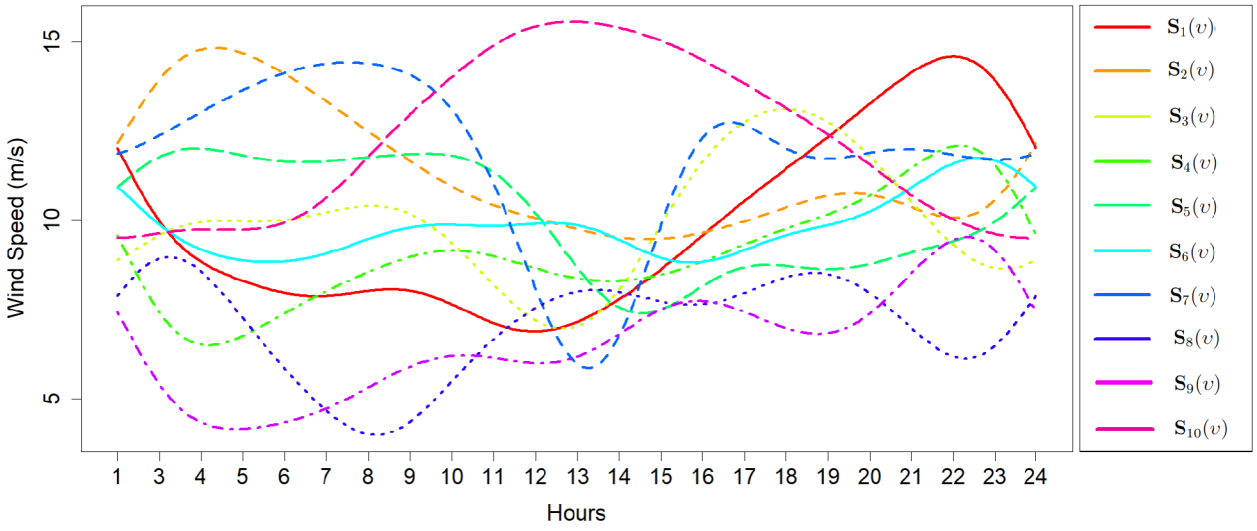

FIGURE 6. A sample of ten daily WS profiles.

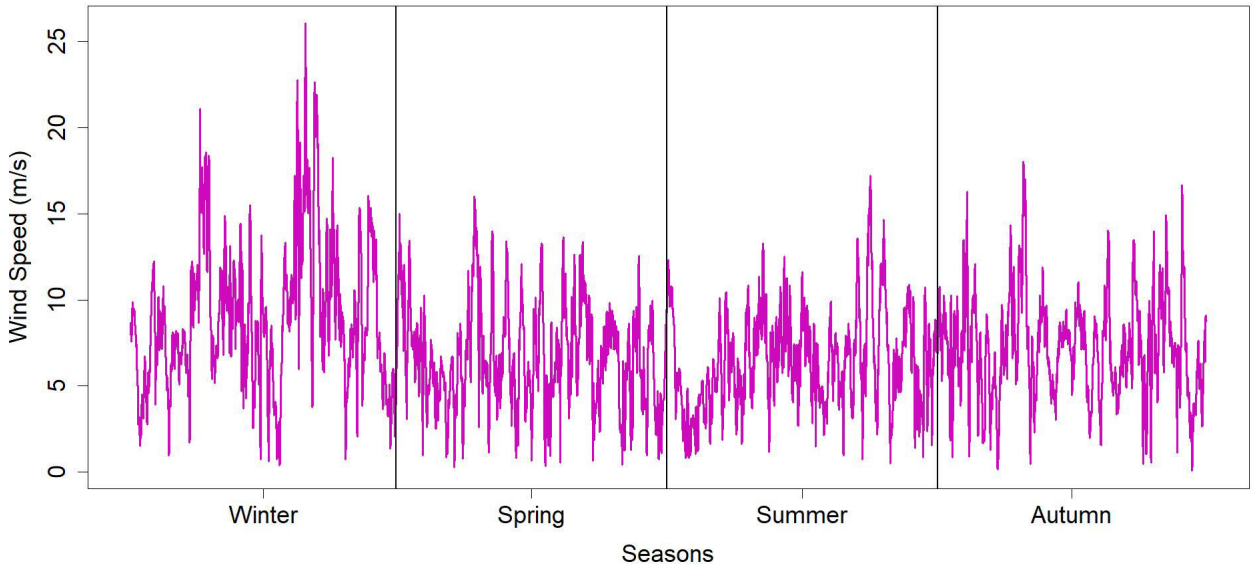

FIGURE 7. Season-wise division of WS profiles for the year 2015.

TABLE 3. One-day-ahead out-of-sample WS forecasting performance for different models.

Models	ME	RMSE	MAE	MASE
ARIMA	0.1099	2.8598	2.2260	14.8969
SARIMA	-0.1369	2.8717	2.2515	14.5241
SARIMAX	-0.1255	2.8732	2.2531	14.5675
NNAR	0.0950	2.8587	2.2237	13.8128
FAR(p)	0.0910	2.4550	1.8309	9.0914
FARX(p)	0.0135	2.1602	1.6626	8.1475

1) MEAN ERROR

The ME is obtained as the average difference of forecasted $\hat{W}_{t,k}$ and original $W_{t,k}$ series. Mathematically, it is defined as

$$\text{ME} = \text{mean}\,(W_{t,k} - \hat{W}_{t,k})$$

2) MEAN ABSOLUTE ERROR

The MAE is obtained as the average of the absolute differences between the forecasted $\hat{W}_{t,k}$ and observed $W_{t,k}$ values. The MAE can be written as

$$\text{MAE} = \text{mean}\,(|W_{t,k} - \hat{W}_{t,k}|)$$

3) ROOT MEAN SQUARED ERROR

The RMSE depends on the scale of the response variable and is a helpful tool for comparing forecasts of a time series across various models. It is calculated as the square root of the mean squared differences between forecasted $\hat{W}_{t,k}$ and observed $W_{t,k}$ data that can be written as

$$\text{RMSE} = \sqrt{\text{mean}\,(W_{t,k} - \hat{W}_{t,k})^2}$$

TABLE 4. p-values of DM test. H_0: the forecasting accuracy for the model in the row and the model in the column is the same; H_1: the forecasting accuracy for the model in the column is greater than the model in the row.

Models	ARIMA	SARIMA	SARIMAX	NNAR	FAR(p)	FARX(p)
ARIMA	–	0.207	0.178	0.641	<0.01	<0.01
SARIMA	0.793	–	0.441	0.196	<0.01	<0.01
SARIMAX	0.822	0.559	–	0.17	<0.01	<0.01
NNAR	0.359	0.804	0.83	–	<0.01	<0.01
FAR(p)	> 0.99	> 0.99	> 0.99	> 0.99	–	<0.01
FARX(p)	> 0.99	> 0.99	> 0.99	> 0.99	> 0.99	–

TABLE 5. Forecasting accuracy for one-day-ahead out-of-sample WS forecasts for different seasons.

Seasons	Errors	ARIMA	SARIMA	SARIMAX	NNAR	FAR(p)	FARX(p)
Winter	ME	-0.1476	0.1096	0.0896	-0.1470	-0.0375	-0.0370
	RMSE	3.1294	3.1539	3.1512	3.1180	2.6563	2.6183
	MAE	2.5457	2.5606	2.5577	2.5302	2.0205	1.9787
	MASE	17.0365	16.5178	16.5370	15.7171	9.7008	8.9915
Spring	ME	0.1736	-0.5270	-0.5547	0.0733	0.0792	0.0565
	RMSE	3.0297	3.0731	3.0786	3.0417	2.7106	2.5947
	MAE	2.3191	2.4000	2.4074	2.3108	2.0207	1.9182
	MASE	15.4644	15.4817	15.5652	14.4058	9.9065	8.5363
Summer	ME	-0.2575	-0.5742	-0.5321	-0.2404	-0.1723	-0.1328
	RMSE	2.2355	2.3010	2.2936	2.2288	1.9802	1.9777
	MAE	1.7908	1.8488	1.8429	1.7889	1.5100	1.3100
	MASE	11.9847	11.9260	11.9156	11.1120	7.4027	6.6749
Autumn	ME	0.8038	0.4559	0.5068	0.7887	0.6088	0.4563
	RMSE	2.9644	2.8877	2.8968	2.9538	2.5671	2.5345
	MAE	2.2640	2.2029	2.2105	2.2635	1.9117	1.8896
	MASE	15.1515	14.2107	14.2922	14.0602	9.3721	8.4086

4) MEAN ABSOLUTE STANDARD ERROR

To verify the accuracy of the forecast, we additionally apply the MASE [55]. Mathematically, it can be written as

$$\text{MASE} = \text{mean}\left|\frac{W_{t,k} - \hat{W}_{t,k}}{\frac{1}{n-1}\sum_{k=2}^{n}\left|W_{t,k} - W_{t,k-1}\right|}\right|$$

C. APPLICATION TO CANADIAN WS PROFILES

This section discusses the forecasting performance of different models described in section II-A when applied to the Canadian WS hourly data. For our functional models, the whole daily profile is used as a single functional datum, whereas in the case of ARIMA, SARIMA, SARIMAX, and NNAR, the univariate series of hourly data are used. The results for the functional models and the four benchmark models are listed in Table 3 in terms of different accuracy measures. This table shows that our proposed modeling strategy produces considerably lower forecasting errors for all models, especially for functional models. Note that the ME is 0.1099 for the traditional ARIMA, which is higher when compared to the ME value of 0.0910 obtained by the FAR(p). Comparing the competing models, the NNAR performs slightly better than ARIMA, SARIMA, and SARIMAX by producing relatively smaller errors. The SARIMA performs better than ARIMA and SARIMAX as the MASE for SARIMA is 14.5241, which is lower than ARIMA and SARIMAX. Note that both functional models perform better than the benchmarks as they produce smaller forecasting errors. For example, looking at the MAE in the table, the ARIMA and NNAR produced MAEs of 2.226 and 2.2237, respectively, which are considerably higher than the MAE values of 1.8309 and 1.6626 obtained by FAR(p) and FARX(p), respectively. Furthermore, Table 3 indicates that the functional model FARX(p) performs much better not only than the benchmarks but also outperforms the FAR(p). Hence, the FARX(p) model is overall best for studying and capturing the intermittent nature of WS profiles. It is worth mentioning that different naïve methods are also used for forecasting WS; however, as expected, their forecasting performance was too low. Hence, their results are not reported in this work.

To further validate the significance of the results listed in Table 3, The Diebold and Mariano (DM) test [56] is applied to each couple of predictors given in Table 3 and the results in the form of P-values are listed in Table 4. From this table, one can see that both functional models produced accurate forecasts compared to their competitors. Within the

functional model, the FARX is statistically more better at forecasting WS profiles than the FAR.

The WS profiles can be very different throughout the year as well as in different seasons in a year. To evaluate the performance of our proposed approaches, we compute the accuracy measures in terms of ME, RMSE, MAE, and MASE for different seasons of the year as well as for each month and the results are listed in Tables 5 and 6, respectively. The understudy site goes through four seasons, and the WS profiles vary considerably across seasons, resulting in different forecasting accuracy errors. For example, looking at Table 5, which lists the one-day-ahead out-of-sample forecasting errors for different seasons, one can notice that the errors vary through all seasons. More precisely, looking at the MAE values, one can see that the MAE values are lower in summer than in other seasons. Wind speed data in summer tends to show more consistent and stable patterns, which lowers forecasting errors. Although the forecasting accuracy for all the models is high, the functional models perform better, especially the FARX(p) model. The FARX(p) models produce an MAE value of 1.9787 in winter, whereas 1.3100 in summer. Note that functional models provide better forecasting results in steady weather circumstances and are adequate to capture these seasonal variations. On the other hand, both benchmark models produce similar results for all seasons; however, they are comparatively higher than the proposed functional models.

Table 6 lists the monthly forecasting accuracy for different models used in this study. The results suggest that the forecasting errors are lower from May to September compared to the other months of the year. The lowest errors are observed in June, whereas the higher errors correspond to March. Comparing the forecasting accuracy of different models, it is evident that the functional models perform relatively better than ARIMA, SARIMA, SARIMAX, and NNAR. Within the functional models, FARX(p) produces better results than FAR(p). The lowest MAE produced by FARX(p) is 1.4526 for June, whereas 2.0642 produced for October is the highest MAE value. Overall, the proposed forecasting procedure performs better by producing considerably lower errors.

Finally, Table 7 compares the models by listing hour-specific one-day-ahead out-of-sample forecasting errors for a whole year. In this table, the accuracy measures obtained for a specific hour are based on all the same specific hours during a complete year, with 365 forecasts for each hour. This table indicates that WS profiles are more stable in the first half of the day than compared to the second half. The forecasting errors for the initial hours are comparatively lower than those in the middle or later day part. The results suggest that our proposed functional models perform relatively well compared to classical ARIMA, SARIMA, SARIMAX, and NNAR. Furthermore, including an exogenous variable enhances the capability of FAR(p) in terms of FARX(p). Concerning the minimum forecasting errors, an MAE for the 1^{st} hour is 0.3867 obtained through FARX(p), whereas FAR(p) resulted in 0.4263. The errors vary throughout the day, with a maximum MAE obtained for the 4th hour. Comparing the results for the classical ARIMA, SARIMA, SARIMAX, and NNAR, one can see that the NNAR model performs well for every hour compared to other models. Moreover, models based on a functional approach outperform all four benchmarks for each hour of the day in terms of ME, RMSE, MAE, and MASE.

To summarize the work, one can see from Tables 3 to 7 that the proposed approach is efficient for WS forecasting and produces considerably lower errors that demonstrate the usefulness of our modeling framework. Moreover, the forecasting results show that the functional forecasting approach is superior to the classical ARIMA and NNAR methods. Within the functional models, the FARX(p) performs well compared to its counterpart when applied without any exogenous information.

Finally, we tried to compare our results with comparable studies in the literature. Reference [57] used an ANN technique to forecast the WS for the next hour for Newfoundland, Canada. This work merges deep learning models with various hyper-parameters, which are selected using a Bayesian optimization algorithm. This suggested ANN technique was compared with support vector machine (SVM), RF, and DT. The reported results indicate that the ANN technique works better with MAE 1.09 and RMSE 1.45, which are 60.88% and 70.59% higher than the accuracy measures of our suggested functional model FAR(p), respectively. Furthermore, based on single and multiple features selection, [58] proposed five novel hybrid neural network models, including radial basis function (RBF), extreme learning machine (ELM), multilayer perceptron (MLP), two walvet neural networks (WNN) trained by improved clonal selection algorithm (ICSA) and particle swarm optimization (PSO), to forecast the WS for 24 hours ahead using the data for Saskatchewan, Canada. Among the five developed neural network models based on single feature selection, the WNN trained by ICSA works better with MAE 3.3986 and RMSE 4.1489 for March, which are 40.65% and 36.39%, are higher than the accuracy measures of our suggested functional model FAR(p) for the same month. Similarly, among the five developed neural network models based on multiple feature selection, the ELM performs better with MAE 6.0130 and RMSE 8.3814 for March, which are 66.46% and 68.46%, more than the accuracy measures of our suggested functional model FAR(p) for the same month.

D. COMPUTATIONAL COMPLEXITY

Finally, we compare the computational cost for each model used in our study. For analysis, the programming environment R, a statistical computing language, is used to implement the models [59]. The deterministic components, Component-I, generally required 0.07 seconds to model and compute one step ahead forecast. This component is common in the case of FAR, FARX, ARIMA, and NNAR.

TABLE 6. One-day-ahead out-of-sample WS forecasts: month-specific forecasting errors.

Months	Errors	ARIMA	SARIMA	SARIMAX	NNAR	FAR(p)	FARX(p)	Months	Errors	ARIMA	SARIMA	SARIMAX	NNAR	FAR(p)	FARX(p)
	ME	-0.5384	-0.1312	-0.1700	-0.5272	-0.3231	-0.2852		ME	-0.7630	-0.9940	-0.9431	-0.7370	-0.5697	-0.4183
January	RMSE	2.9776	2.9425	2.9459	2.9660	2.2415	2.0133	July	RMSE	2.1417	2.2564	2.2368	2.1510	1.9472	1.6166
	MAE	2.3120	2.2741	2.2770	2.3018	1.7123	1.5626		MAE	1.7253	1.8061	1.7895	1.7289	1.4531	1.2562
	MASE	15.4726	14.6697	14.7220	14.2980	8.1537	7.6197		MASE	11.5464	11.6510	11.5704	10.7397	7.1975	6.4664
	ME	0.5814	0.8300	0.7860	0.5847	0.5347	0.5103		ME	0.1609	-0.1643	-0.0954	0.1470	0.0854	0.0746
February	RMSE	3.2233	3.3502	3.3378	3.1955	2.7401	2.1460	August	RMSE	2.4002	2.4153	2.4128	2.3852	2.0143	1.9032
	MAE	2.6723	2.7968	2.7853	2.6307	2.1271	1.7539		MAE	1.9349	1.9425	1.9429	1.9189	1.5527	1.4504
	MASE	17.8834	18.0418	18.0085	16.3414	10.4378	9.4657		MASE	12.9485	12.5305	12.5560	11.9196	7.6073	6.9094
	ME	0.6559	0.2432	0.2113	0.6323	0.4556	0.4402		ME	0.3126	0.2051	0.2818	0.3348	0.0166	0.0761
March	RMSE	3.5223	3.4012	3.4002	3.5523	2.6431	2.3113	September	RMSE	2.4084	2.3789	2.3865	2.3865	1.9885	1.9124
	MAE	2.8052	2.6773	2.6772	2.7981	2.0170	1.8375		MAE	1.8391	1.8200	1.8240	1.8270	1.5339	1.4878
	MASE	18.7730	17.2710	17.3101	17.6293	11.0991	8.9759		MASE	12.3078	11.7405	11.7930	11.3485	7.5362	6.8260
	ME	-0.6607	-1.3709	-1.4008	-0.6193	-0.4282	-0.4109		ME	1.2938	0.8403	0.8911	1.3019	0.8021	0.7429
April	RMSE	2.8503	3.0860	3.0987	2.8389	2.5021	2.1815	October	RMSE	3.2646	3.1237	3.1386	3.2712	3.0836	2.6461
	MAE	2.1624	2.4478	2.4628	2.1526	1.8562	1.7013		MAE	2.5271	2.3903	2.4047	2.5253	2.3207	2.0642
	MASE	14.4714	15.7901	15.9236	13.3715	9.3940	8.2602		MASE	16.8992	15.4192	15.5477	15.6968	11.2081	10.3272
	ME	0.2017	-0.4804	-0.5019	0.1937	0.1844	0.1511		ME	0.7581	0.3093	0.3347	0.7429	0.5997	0.4572
May	RMSE	2.6404	2.6911	2.6969	2.6530	2.6337	2.4864	November	RMSE	3.1468	3.0907	3.0951	3.1369	2.3888	2.0124
	MAE	1.9613	2.0763	2.0839	1.9600	1.8795	1.7864		MAE	2.4291	2.3923	2.3964	2.4178	1.7997	1.6068
	MASE	13.1166	13.3940	13.4735	12.1832	9.2097	8.3638		MASE	16.1892	15.4320	15.4940	15.0806	9.3108	8.0088
	ME	-0.0784	-0.5639	-0.5587	-0.0638	0.0294	0.0109		ME	-0.4107	-0.3003	-0.2797	-0.4292	-0.2696	-0.2687
June	RMSE	2.1526	2.2244	2.2243	2.1388	1.9786	1.9012	December	RMSE	3.1944	3.1754	3.1761	3.1910	2.9461	2.5312
	MAE	1.7097	1.7960	1.7957	1.7165	1.4940	1.4526		MAE	2.6678	2.6337	2.6328	2.6646	2.2326	1.9875
	MASE	11.4418	11.5855	11.6104	10.6622	7.4035	6.6482		MASE	17.8354	16.9894	17.0229	16.5725	10.5821	9.9350

TABLE 7. Hour-specific one-day-ahead out-of-sample forecasting errors.

Hours	Errors	ARIMA	SARIMA	SARIMAX	NNAR	FAR(p)	FARX(p)	Hours	Errors	ARIMA	SARIMA	SARIMAX	NNAR	FAR(p)	FARX(p)
	ME	0.6602	1.6807	1.6389	0.4904	-0.4264	-0.3867		ME	1.1732	1.5851	1.5511	0.6051	-0.4188	-0.4175
1	RMSE	0.6602	1.6807	1.6389	0.4904	0.4264	0.3867	13	RMSE	1.3677	1.6106	1.5784	0.7686	0.7528	0.7054
	MAE	0.6602	1.6807	1.6389	0.4904	0.4264	0.3867		MAE	1.1732	1.5851	1.5511	0.6632	0.5449	0.5118
	MASE	4.4181	10.8418	10.5966	4.5492	1.8974	1.8958		MASE	7.8515	10.2253	10.0288	7.3449	2.5090	2.4247
	ME	0.4505	1.4647	0.8470	0.1937	-0.8285	-0.7686		ME	1.2128	1.5601	1.5225	0.5875	-0.4014	-0.3977
2	RMSE	0.9869	1.4806	1.3468	0.9543	0.9209	0.8583	14	RMSE	1.3965	1.5867	1.5518	0.7468	0.7269	0.6808
	MAE	0.9505	1.4647	1.2649	0.2967	0.8285	0.7686		MAE	1.2128	1.5601	1.5225	0.6414	0.5184	0.4853
	MASE	3.0149	9.4486	9.2319	3.1316	3.7681	3.6867		MASE	8.1163	10.0638	9.8441	7.6155	2.3793	2.3070
	ME	0.3500	1.3715	1.3399	0.0728	-1.0597	-0.9939		ME	1.2293	1.5334	1.4934	0.5717	-0.3971	-0.3922
3	RMSE	1.4147	1.3890	1.3566	0.3053	1.1565	1.0893	15	RMSE	1.4008	1.5619	1.5251	0.7271	0.7077	0.6627
	MAE	1.2050	1.3715	1.3399	0.2541	1.0597	0.9940		MAE	1.2293	1.5334	1.4934	0.6220	0.5063	0.4739
	MASE	4.8727	8.8470	8.6632	4.7156	2.3596	2.3422		MASE	8.2265	9.89164	9.6558	7.7234	2.3234	2.2532
	ME	-1.1503	1.3345	1.3062	-1.0766	0.3157	0.0283		ME	1.2385	1.5143	1.4730	0.5666	-0.3868	-0.3817
4	RMSE	1.3746	1.3496	1.3204	0.2696	1.22834	1.1527	16	RMSE	1.3993	1.5432	1.5053	0.7146	0.6877	0.6442
	MAE	1.3157	1.3345	1.3062	0.2169	1.1503	1.0767		MAE	1.2385	1.5143	1.4730	0.6138	0.4893	0.4584
	MASE	5.2783	8.6085	8.4457	5.1191	2.1285	2.1124		MASE	8.2884	9.7686	9.5238	7.7980	2.2472	2.1773
	ME	0.3721	1.3562	1.3302	0.0973	-1.1115	-1.0381		ME	1.2242	1.4909	1.4490	0.5535	-0.3830	-0.3762
5	RMSE	1.2543	1.3687	1.3422	0.2933	1.1790	1.104176	17	RMSE	1.3788	1.5213	1.4830	0.6983	0.6717	0.6288
	MAE	1.1537	1.3562	1.3302	0.2482	1.1115	1.0381		MAE	1.2242	1.4909	1.4490	0.5979	0.4794	0.4484
	MASE	5.0891	8.7482	8.6003	4.9462	2.4904	2.4260		MASE	8.1925	9.6173	9.3689	7.7624	2.1982	2.1333
	ME	0.5190	1.4603	1.4368	0.2425	-0.9667	-0.9005		ME	1.1719	1.4457	1.4036	0.5091	-0.4037	-0.3952
6	RMSE	1.1442	1.4884	1.4656	0.4774	1.0811	1.0117	18	RMSE	1.3416	1.4871	1.4489	0.6811	0.6767	0.6341
	MAE	1.0519	1.4603	1.4368	0.3682	0.9667	0.9005		MAE	1.1719	1.4457	1.4036	0.5783	0.4948	0.4634
	MASE	4.4148	9.4202	9.2896	4.3017	3.4734	3.4599		MASE	7.8424	9.3260	9.0753	7.4920	2.2717	2.2019
	ME	0.6595	1.5544	1.5333	0.3713	-0.8292	-0.7757		ME	1.0892	1.3765	1.3338	0.4349	-0.4431	-0.4331
7	RMSE	1.0006	1.5937	1.5743	0.6185	1.0006	0.9367	19	RMSE	1.3090	1.4477	1.4104	0.6943	0.7097	0.66801
	MAE	0.8292	1.5544	1.5333	0.4791	0.8292	0.7757		MAE	1.1312	1.3765	1.3338	0.5952	0.5294	0.4976
	MASE	4.4135	10.0268	9.9137	4.2479	3.8031	3.6901		MASE	7.5701	8.8793	8.6241	7.1636	2.4395	2.3558
	ME	0.7298	1.5762	1.5563	0.4360	-0.7615	-0.7216		ME	0.9892	1.2879	1.2445	0.3448	-0.4924	-0.4819
8	RMSE	0.8846	1.6113	1.5930	0.6584	0.9415	0.8845	20	RMSE	1.2920	1.4138	1.3784	0.7426	0.7620	0.7234
	MAE	0.7298	1.5762	1.5563	0.5303	0.7615	0.7216		MAE	1.1201	1.3274	1.2898	0.6338	0.5744	0.5432
	MASE	4.88380	10.1679	10.0627	4.6397	3.5374	3.3887		MASE	7.4961	8.5629	8.3391	7.0271	2.6631	2.5561
	ME	0.8045	1.5861	1.5652	0.4831	-0.6915	-0.6687		ME	0.8793	1.1857	1.1494	0.2502	-0.5478	-0.5396
9	RMSE	0.9549	1.6174	1.5978	0.6837	0.8887	0.8380	21	RMSE	1.2933	1.3924	1.3551	0.8085	0.8266	0.7968
	MAE	0.8045	1.5861	1.5652	0.5670	0.6915	0.6687		MAE	1.1296	1.3051	1.2642	0.6818	0.6258	0.5980
	MASE	5.3841	10.2314	10.1200	5.0725	3.2785	3.0773		MASE	7.5595	8.4189	8.1736	7.0249	2.9316	2.7849
	ME	0.9129	1.6047	1.5813	0.5512	-0.5951	-0.5849		ME	0.7610	1.0877	1.0502	0.1462	-0.6056	-0.6009
10	RMSE	1.0850	1.6335	1.6111	0.7458	0.8475	0.7968	22	RMSE	1.3159	1.3760	1.3422	0.9015	0.8961	0.8764
	MAE	0.9129	1.6047	1.5813	0.6267	0.6496	0.6188		MAE	1.1565	1.2898	1.2537	0.7435	0.6801	0.6566
	MASE	6.1095	10.3514	10.2243	5.6989	3.0338	2.8909		MASE	7.7399	8.3203	8.1061	7.1822	3.2191	3.0266
	ME	1.0295	1.6195	1.5929	0.6073	-0.5044	-0.5020		ME	-0.6734	0.9891	0.9507	-0.6693	0.6394	0.0362
11	RMSE	1.2281	1.6461	1.6202	0.7935	0.8172	0.7661	23	RMSE	1.3552	1.3681	1.3378	1.0122	0.98580	0.9696
	MAE	1.0295	1.6195	1.5929	0.6759	0.6272	0.5923		MAE	1.1948	1.2851	1.2530	0.8148	0.7447	0.7226
	MASE	6.8894	10.4470	10.2992	6.4214	2.9037	2.7910		MASE	7.9960	8.2898	8.1014	7.4230	3.5424	3.3139
	ME	1.1143	1.6088	1.5785	0.6192	-0.4504	-0.4505		ME	0.5141	0.8858	0.8470	-0.0821	-0.7561	-0.7519
12	RMSE	1.3160	1.6337	1.6045	0.7900	0.7835	0.7342	24	RMSE	1.4119	1.3734	1.3468	1.1441	1.1071	1.0928
	MAE	1.1143	1.6088	1.5785	0.6821	0.5869	0.5526		MAE	1.2437	1.2937	1.2649	0.8976	0.8244	0.8030
	MASE	7.4570	10.3777	10.2057	3.1316	2.7094	2.6119		MASE	8.3230	8.3450	8.1783	7.7351	3.9368	3.6687

On the other hand, Table 8 lists the average time (in seconds) needed for model estimation and a one-step-ahead forecast for Component-II, which varies among models. Using ARIMA as a baseline model, the average time required for a forecast based on SARIMA and SARIMAX models is 6.75 and 9.03 times longer than that of the ARIMA

TABLE 8. Average time for a single one-step-ahead forecast.

Average Time	ARIMA	SARIMA	SARIMAX	NNAR	FAR(p)	FARX(p)
Time (s)	0.11	0.71	0.96	0.27	0.79	0.81
Relative time	1	6.75	9.03	2.55	7.48	7.65

model. Moreover, the average time required for a forecast based on NNAR is only 2.55 times longer than that of the ARIMA model. Lastly, the average time needed for a forecast based on functional models is 7.48 and 7.65 times longer than the ARIMA model. Note that the average time is less than one second for each model, indicating the model's computation efficiency. In addition, the GAM library in R is used to model and forecast Component-I [60]. Furthermore, the forecast library in R is used for ARIMA, SARIMA, SARIMAX, and NNAR [61]. The documentation provided by the packages provides in-depth details on the particular algorithms utilized in the estimation. Computation times are based on an Intel(R)-Core(TM) i7-4770 CPU running at 3.40 GHz, which has been used for all computations.

V. CONCLUSION

Modeling and forecasting WS is crucial as it is directly linked to various fields like agriculture, industry, the power sector, etc. However, due to the irregular and intermittent nature of the WS, the forecasting problem is challenging and gained the researcher's attention. The WS series generally contain complex features, like long trends and seasonal behavior, and they are also considered site variants as their characteristics vary with different regions. In this context, models developed for one region are generally unstable for any other site. To address these issues and to efficiently forecast the WS daily profiles, this research work proposes modeling the WS profiles under the paradigm of the FDA. The general modeling framework is based on extreme values treatment followed by decomposing WS data into two components. The first component models and forecasts the deterministic feature of the WS series, while the second component addresses the issue of forecasting stochastic variability in the WS series. For the estimation and forecasting of the second component, this study proposes the use of functional models that consider the WS data for a full day as a single functional datum. This methodology not only bypasses the issue of high-frequency data but also provides forecasts for any time within a day. Once both components are estimated and forecasted, they are aggregated to obtain the final forecast. To evaluate the performance of the proposed methodology, wind speed data from the NASA power project for the site Canada, located in Durham, England, are used. Using the data, one-day-ahead out-of-sample forecasts are obtained for a complete year, and the forecasting performance of different models is assessed through different accuracy measures, including the ME, RMSE, MAE, and MASE. For comparison purposes, the forecasts are also obtained through the classical ARIMA, SARIMA, SARIMAX, and NNAR models.

The study findings suggest that the proposed forecasting procedure is efficient in forecasting WS daily profiles by producing considerably lower forecasting errors. Both functional models perform relatively better than the ARIMA, SARIMA, SARIMAX, and NNAR. The functional model considering the exogenous information (FARX) produces better results than its counterpart (FAR). On the other hand, NNAR produces comparatively better results than ARIMA, SARIMA, and SARIMAX. In conclusion, the models under the functional paradigm perform better than the competitors.

Apart from excellent forecasting performance, the current study has some limitations. Additional meteorological features such as humidity, pressure, and temperature are not used within the models, which could enhance their forecasting ability. This study focuses on the point forecast, whereas interval forecast may also be important in many situations. In the future, the accuracy of functional models can be enhanced by utilizing more explanatory variables such as air temperature, relative humidity, sea level, and precipitation amount. Moreover, other machine learning techniques can also be compared with the proposed functional models to assess their performance for WS forecasting.

ACKNOWLEDGMENT

The authors would like to thank the six anonymous referees for their constructive comments to improve the quality and presentation of the work.

DATA AVAILABILITY

Data can be requested from the corresponding author.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

COMPUTATIONAL CODE AVAILABILITY

The R codes used in this study are available from the corresponding author on request.

REFERENCES

[1] L. S. Paraschiv and S. Paraschiv, "Contribution of renewable energy (hydro, wind, solar and biomass) to decarbonization and transformation of the electricity generation sector for sustainable development," *Energy Rep.*, vol. 9, pp. 535–544, Sep. 2023.

[2] A. Altan, S. Karasu, and E. Zio, "A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer," *Appl. Soft Comput.*, vol. 100, Mar. 2021, Art. no. 106996.

[3] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *Proc. North Amer. Power Symp.*, Sep. 2010, pp. 1–8.

[4] G. Li and J. Shi, "On comparing three artificial neural networks for wind speed forecasting," *Appl. Energy*, vol. 87, no. 7, pp. 2313–2320, Jul. 2010.

[5] M. Monfared, H. Rastegar, and H. M. Kojabadi, "A new strategy for wind speed forecasting using artificial intelligent methods," *Renew. Energy*, vol. 34, no. 3, pp. 845–848, Mar. 2009.

[6] Z.-H. Guo, J. Wu, H.-Y. Lu, and J.-Z. Wang, "A case study on a hybrid wind speed forecasting method using BP neural network," *Knowl.-Based Syst.*, vol. 24, no. 7, pp. 1048–1056, Oct. 2011.

[7] X. Huang, J. Wang, and B. Huang, "Two novel hybrid linear and nonlinear models for wind speed forecasting," *Energy Convers. Manage.*, vol. 238, Jun. 2021, Art. no. 114162.

[8] S. Wang, N. Zhang, L. Wu, and Y. Wang, "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method," *Renew. Energy*, vol. 94, pp. 629–636, Aug. 2016.

[9] J. Wang, Y. Song, F. Liu, and R. Hou, "Analysis and application of forecasting models in wind power integration: A review of multi-step-ahead wind speed forecasting models," *Renew. Sustain. Energy Rev.*, vol. 60, pp. 960–981, Jul. 2016.

[10] R. G. Kavasseri and K. Seetharaman, "Day-ahead wind speed forecasting using f-ARIMA models," *Renew. Energy*, vol. 34, no. 5, pp. 1388–1393, May 2009.

[11] G. Santamaría-Bonfil, A. Reyes-Ballesteros, and C. Gershenson, "Wind speed forecasting for wind farms: A method based on support vector regression," *Renew. Energy*, vol. 85, pp. 790–809, Jan. 2016.

[12] J.-Y. Ryu, B. Lee, S. Park, S. Hwang, H. Park, C. Lee, and D. Kwon, "Evaluation of weather information for short-term wind power forecasting with various types of models," *Energies*, vol. 15, no. 24, p. 9403, Dec. 2022.

[13] Z. Yang and J. Wang, "A hybrid forecasting approach applied in wind speed forecasting based on a data processing strategy and an optimized artificial intelligence algorithm," *Energy*, vol. 160, pp. 87–100, Oct. 2018.

[14] Q. Zhou, C. Wang, and G. Zhang, "Hybrid forecasting system based on an optimal model selection strategy for different wind speed forecasting problems," *Appl. Energy*, vol. 250, pp. 1559–1580, Sep. 2019.

[15] Y. Ren, P. N. Suganthan, and N. Srikanth, "A novel empirical mode decomposition with support vector regression for wind speed forecasting," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1793–1798, Aug. 2016.

[16] G.-R. Ji, P. Han, and Y.-J. Zhai, "Wind speed forecasting based on support vector machine with forecasting error estimation," in *Proc. Int. Conf. Mach. Learn. Cybern.*, vol. 5, Aug. 2007, pp. 2735–2739.

[17] A. Xu, M.-W. Tian, B. Firouzi, K. A. Alattas, A. Mohammadzadeh, and E. Ghaderpour, "A new deep learning restricted Boltzmann machine for energy consumption forecasting," *Sustainability*, vol. 14, no. 16, p. 10081, Aug. 2022.

[18] H. Zhang, J. Liu, and J. Zhao, "A hybrid wind speed forecasting model using wrf, arima, and wavelet neural networks," *Renew. Energy*, vol. 156, pp. 456–464, Mar. 2020.

[19] C. Zhu, Y. Wei, and J. Li, "A hybrid model based on WRF and LSTM neural network for wind speed forecasting," *Renew. Energy*, vol. 131, pp. 207–216, Jan. 2019.

[20] J. Guo, C. Feng, X. Zhang, W. Liu, and L. Wang, "Real-time updating method for short-term wind power forecasting based on WRF and variational data assimilation," *Energy*, vol. 215, Jun. 2021, Art. no. 119155.

[21] V. Bali, A. Kumar, and S. Gangwar, "Deep learning based wind speed forecasting—A review," in *Proc. 9th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Jan. 2019, pp. 426–431.

[22] M. Neshat, M. M. Nezhad, E. Abbasnejad, S. Mirjalili, L. B. Tjernberg, D. Astiaso Garcia, B. Alexander, and M. Wagner, "A deep learning-based evolutionary model for short-term wind speed forecasting: A case study of the lillgrund offshore wind farm," *Energy Convers. Manage.*, vol. 236, May 2021, Art. no. 114002.

[23] E. Cadenas and W. Rivera, "Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA–ANN model," *Renew. Energy*, vol. 35, no. 12, pp. 2732–2738, Dec. 2010.

[24] E. Cadenas and W. Rivera, "Wind speed forecasting in the south coast of Oaxaca, México," *Renew. Energy*, vol. 32, no. 12, pp. 2116–2128, Oct. 2007.

[25] J. C. Palomares-Salas, J. J. G. de la Rosa, J. G. Ramiro, J. Melgar, A. Aguera, and A. Moreno, "ARIMA vs. neural networks for wind speed forecasting," in *Proc. IEEE Int. Conf. Comput. Intell. Meas. Syst. Appl.*, May 2009, pp. 129–133.

[26] O. B. Shukur and M. H. Lee, "Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA," *Renew. Energy*, vol. 76, pp. 637–647, Apr. 2015.

[27] W. Zhang, J. Wang, J. Wang, Z. Zhao, and M. Tian, "Short-term wind speed forecasting based on a hybrid model," *Appl. Soft Comput.*, vol. 13, no. 7, pp. 3225–3233, 2013.

[28] I. Shah, S. Akbar, T. Saba, S. Ali, and A. Rehman, "Short-term forecasting for the electricity spot prices with extreme values treatment," *IEEE Access*, vol. 9, pp. 105451–105462, 2021.

[29] Y. Yang, H. Zhou, Y. Gao, J. Wu, Y.-G. Wang, and L. Fu, "Robust penalized extreme learning machine regression with applications in wind speed forecasting," *Neural Comput. Appl.*, vol. 34, no. 1, pp. 391–407, Jan. 2022.

[30] Y. Yang, H. Zhou, J. Wu, Z. Ding, Y.-C. Tian, D. Yue, and Y.-G. Wang, "Robust adaptive rescaled lncosh neural network regression toward time-series forecasting," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 9, pp. 5658–5669, Sep. 2023.

[31] Y. Yang, H. Zhou, J. Wu, Z. Ding, and Y.-G. Wang, "Robustified extreme learning machine regression with applications in outlier-blended wind-speed forecasting," *Appl. Soft Comput.*, vol. 122, Jun. 2022, Art. no. 108814.

[32] Z. H. Munim, "State-space TBATS model for container freight rate forecasting with improved accuracy," *Maritime Transp. Res.*, vol. 3, 2022, Art. no. 100057.

[33] Y. Hao, X. Wang, J. Wang, and W. Yang, "A new perspective of wind speed forecasting: Multi-objective and model selection-based ensemble interval-valued wind speed forecasting system," *Energy Convers. Manage.*, vol. 299, Jan. 2024, Art. no. 117868.

[34] X. Shi, J. Wang, and B. Zhang, "A fuzzy time series forecasting model with both accuracy and interpretability is used to forecast wind power," *Appl. Energy*, vol. 353, Jan. 2024, Art. no. 122015.

[35] X. Li, K. Li, S. Shen, and Y. Tian, "Exploring time series models for wind speed forecasting: A comparative analysis," *Energies*, vol. 16, no. 23, p. 7785, Nov. 2023.

[36] J. Heng, Y. Hong, J. Hu, and S. Wang, "Probabilistic and deterministic wind speed forecasting based on non-parametric approaches and wind characteristics information," *Appl. Energy*, vol. 306, Jan. 2022, Art. no. 118029.

[37] M. Brabec, A. Craciun, and A. Dumitrescu, "Hybrid numerical models for wind speed forecasting," *J. Atmos. Solar-Terr. Phys.*, vol. 220, Sep. 2021, Art. no. 105669.

[38] Z. Liu, P. Jiang, L. Zhang, and X. Niu, "A combined forecasting model for time series: Application to short-term wind speed forecasting," *Appl. Energy*, vol. 259, Feb. 2020, Art. no. 114137.

[39] J. O. Ramsay and C. J. Dalzell, "Some tools for functional data analysis," *J. Roy. Stat. Soc. B, Stat. Methodol.*, vol. 53, no. 3, pp. 539–561, Jul. 1991.

[40] I. Shah and F. Lisi, "Forecasting of electricity price through a functional prediction of sale and purchase curves," *J. Forecasting*, vol. 39, no. 2, pp. 242–259, Mar. 2020.

[41] J. Andersson and J. Lillestøl, "Modeling and forecasting electricity consumption by functional data analysis," *J. Energy Markets*, vol. 3, no. 1, pp. 3–15, Mar. 2010.

[42] F. Jan, I. Shah, and S. Ali, "Short-term electricity prices forecasting using functional time series analysis," *Energies*, vol. 15, no. 9, p. 3423, May 2022.

[43] Y. Chen, T. Koch, K. G. Lim, X. Xu, and N. Zakiyeva, "A review study of functional autoregressive models with application to energy forecasting," *WIREs Comput. Statist.*, vol. 13, no. 3, p. e1525, May 2021.

[44] A. R. Ghumman, H. Haider, and M. Shafiquzamman, "Functional data analysis of models for predicting temperature and precipitation under climate change scenarios," *J. Water Climate Change*, vol. 11, no. 4, pp. 1748–1765, Dec. 2020.

[45] K.-H. Kang and H.-S. Ahn, "Functional data analysis of temperature and precipitation data," *Korean J. Appl. Statist.*, vol. 19, no. 3, pp. 431–445, 2006.

[46] I. Shah, I. Muhammad, S. Ali, S. Ahmed, M. M. A. Almazah, and A. Y. Al-Rezami, "Forecasting day-ahead traffic flow using functional time series approach," *Mathematics*, vol. 10, no. 22, p. 4279, Nov. 2022.

[47] S. Borovkova and F. J. Permana, "Modelling electricity prices by the potential jump-diffusion," in *Stochastic Finance*. Boston, MA, USA: Springer, 2006, pp. 239–263.

[48] M. Loève, "Random functions à of exponential orthogonal composition," *Sci. J.*, vol. 84, pp. 159–162, Jan. 1946.

[49] K. Karhunen, "On linear methods in probability theory," *Annales Academiae Scientiarun Fennicae Ser. A1, Mathematica Phys.*, vol. 47, pp. 1–101, Mar. 1947.

[50] B. W. Silverman and J. O. Ramsay, *Applied Functional Data Analysis: Methods and Case Studies*. New York, NY, USA: Springer, 2002.

[51] H. L. Shang, "A survey of functional principal component analysis," *Adv. Stat. Anal.*, vol. 98, no. 2, pp. 121–142, Apr. 2014.
[52] S. Hörmann and P. Kokoszka, "Weakly dependent functional data," *Ann. Statist.*, vol. 38, no. 3, pp. 1845–1884, Jun. 2010.
[53] P. Kokoszka and M. Reimherr, "Determining the order of the functional autoregressive model," *J. Time Ser. Anal.*, vol. 34, no. 1, pp. 116–129, Jan. 2013.
[54] A. Aue, D. D. Norinho, and S. Hörmann, "On the prediction of stationary functional time series," *J. Amer. Stat. Assoc.*, vol. 110, no. 509, pp. 378–392, Jan. 2015.
[55] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *Int. J. Forecasting*, vol. 22, no. 4, pp. 679–688, Oct. 2006.
[56] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Statist.*, vol. 13, no. 3, p. 253, Jul. 1995.
[57] H. Rahaman, T. M. R. Bashar, M. Munem, M. H. H. Hasib, H. Mahmud, and A. N. Alif, "Bayesian optimization based ANN model for short term wind speed forecasting," in *Proc. IEEE Electr. Power Energy Conf. (EPEC)*, Nov. 2020, pp. 1–5.
[58] M. Abbasipour, M. A. Igder, and X. Liang, "A novel hybrid neural network-based day-ahead wind speed forecasting technique," *IEEE Access*, vol. 9, pp. 151142–151154, 2021.
[59] *R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing*, R Core Team, Vienna, Austria, 2022.
[60] T. Hastie, "GAM: Generalized additive models," R Package Version 1.22, 2022. [Online]. Available: https://CRAN.R-project.org/package=gam
[61] R. J. Hyndman and Y. Khandakar, "Automatic time series forecasting: TheforecastPackage forR," *J. Stat. Softw.*, vol. 27, no. 3, pp. 1–22, 2008.

MUHAMMAD UZAIR is currently pursuing the Ph.D. degree with the Department of Statistics, Quaid-i-Azam University (QAU), Islamabad, Pakistan. His research interests include functional data analysis, time series forecasting, and environmental studies.

ISMAIL SHAH received the master's degree from Lund University, Sweden, and the Ph.D. degree in statistics from the University of Padua, Italy. He is currently a Researcher with the Department of Statistical Sciences, University of Padua, and an Associate Professor with the Department of Statistics, Quaid-i-Azam University (QAU), Islamabad, Pakistan. His research interests include functional data analysis, time series analysis, energy economics, and applied and industrial statistics.

SAJID ALI received the Ph.D. degree in statistics from Bocconi University, Milan, Italy. He is currently an Associate Professor with the Department of Statistics, Quaid-i-Azam University (QAU), Islamabad, Pakistan. His research interests include Bayesian inference, construction of new flexible probability distributions, and process monitoring.