

Article

Uncertainty Quantification of Complex Weather Dynamics Using a Novel Functional Autoregressive Model

Ismail Shah ^{1,2} , Muhammad Uzair ^{1,*} , Sajid Ali ¹  and Sadiya M. Aljeddani ^{3,*} ¹ Department of Statistics, Quaid-i-Azam University, Islamabad 45320, Pakistan; ismail.shah@unipd.it (I.S.); sajidali@qau.edu.pk (S.A.)² Department of Statistical Sciences, University of Padova, 35121 Padova, Italy³ Mathematics Department, Al-Lith University College, Umm Al-Qura University, Al-Lith 21961, Saudi Arabia

* Correspondence: muzair@stat.qau.edu.pk (M.U.); smgadany@uqu.edu.sa (S.M.A.)

Abstract

Functional time series (FTS) modeling has emerged as a powerful framework for capturing complex temporal dependencies using the functional autoregressive models $FAR(p, m)$ and $FARX(p, m, \tau)$. These functional models characterize the evolution of functional observations by incorporating ‘ p ’ lagged functional responses, ‘ m ’ truncated dimensions from functional principal component analysis (FPCA), and τ number of scalar covariates with optimal parameter selection guided by the minimization of the functional final prediction error $ffPE(p, m)$. The aim of this study is to propose a computationally efficient FAR model that can integrate a number of functional covariates to achieve a high predictive accuracy in terms of standard out-of-sample accuracy measures. To this end, an integrated functional autoregressive model $FARX(p, m, \mathbf{g}, \tau)$ is developed, where $\mathbf{X}$ denotes the exogenous information, this being a lagged or modeled functional profile within the $FAR(p, m)$ framework, and ‘ $\mathbf{g}$ ’ represents a vector of optimal dimensions for a number of functional covariates. The theoretical contributions are twofold: first, deriving the distribution of the modified functional final prediction error, denoted as $ffPEX(p, m, \mathbf{g}, \tau)$; second, using this derivation to establish formal criteria for optimal model selection. To empirically investigate the predictive performance of the proposed model, hourly temperature data from the NASA POWER project are considered, and day-ahead out-of-sample forecasts over a full annual cycle are computed. The forecasting performance of the proposed model is assessed against state-of-the-art models using different error summary metrics. The results show that functional models consistently outperform traditional time series and neural network-based approaches, with $FARX(p, m, \mathbf{g}, \tau)$ achieving superior predictive accuracy compared to $FAR(p, m)$ and $FARX(p, m, \tau)$, thereby underscoring the efficacy of incorporating functional exogenous information in FTS modeling.



Academic Editors: MinJae Lee,
Carmen Patino-Alonso and
Jong-Min Kim

Received: 18 January 2026

Revised: 13 February 2026

Accepted: 25 February 2026

Published: 1 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: functional autoregressive models; forecasting; functional final prediction error; functional principal component analysis; ARIMA; temperature

MSC: 62M10; 62R10; 46-11; 62J05

1. Introduction

Climate patterns exhibit intricate variability, driven by complex interactions between atmospheric, oceanic, and terrestrial systems. Understanding these patterns is crucial for predicting extreme weather events, assessing climate change impacts, and developing effective mitigation strategies. Among the many climatic variables, temperature plays a pivotal

role, serving as a primary driver of atmospheric circulation, precipitation dynamics, and weather extremes [1]. It is a fundamental variable in meteorology, shaping the intricate patterns of weather systems and influencing atmospheric circulation. As key drivers of energy exchange within the Earth's climate system, temperature variations dictate the formation of storms, heatwaves, and cold spells, making them crucial parameters for understanding and predicting complex weather structures. Moreover, long-term temperature trends serve as vital indicators of environmental changes, including global warming and shifts in regional climate regimes. Accurate modeling and forecasting of temperature are essential for climate resilience, disaster preparedness, and sustainable environmental policies [2,3].

To accurately forecast air temperature, researchers used different methods that vary in complexity, methodology, and performance. For example, ref. [4] used linear regression to model the daily maximum temperature data with geopotential thickness forecasts as a covariate for the site in Nashville, Tennessee. The study reveals that forecasting accuracy of daily maximum temperature is high for warmer air, whereas in the winter season, the forecasting accuracy tends to decrease. Although this method improves daily maximum temperature modeling, it has limitations, including dependence on geopotential thickness forecasts, seasonal effects, and other environmental factors. In the context of producing long-horizon predictive densities, which are crucial for pricing weather derivatives, the daily average temperature data are modeled by various authors [5,6]. For pricing temperature derivatives, the Chicago Mercantile Exchange (CME) defines the daily average temperature underlying its contracts as the mean of the daily high and low temperatures. In an evaluation using daily average temperature data from the United States of America, ref. [7] proposed an Extreme Value Theory (EVT) approach that outperformed classical models such as AR, generalized autoregressive heteroskedasticity (GARCH), and standard regression.

Based on 62 years of historical time series (TS) data from Zhengzhou, China, ref. [8] modeled the daily average temperature using a mean-reverting Ornstein–Uhlenbeck process to provide accurate pricing for weather derivatives. Recent advances in nonlinear and complex time series modeling include Markov switching bilinear models for higher-order spectral analysis and threshold-based stochastic volatility frameworks which are capable of capturing asymmetric dynamics and regime changes [9,10]. Heatwaves significantly impact society and hence are referred to as highlighting the importance of modeling temperature. Ref. [11] evaluated the Regional Atmospheric Modelling System (RAMS) for predicting maximum and minimum summer temperatures in the Valencia Region from 2007 to 2010. The results show that the model predicts maximum temperatures well, with small errors of about 2 °C. The impact of global warming based on daily minimum and maximum temperature data from Coimbra in Portugal is assessed by [12] using the Hadley Centre model and a neural network. The results suggest that the two-layer neural network outperforms the competitors. Ref. [13] used the Abductive Induction Mechanism (AIM) as a modeling tool for temperature forecasting within a machine learning framework. Using daily maximum temperature data from the city of Dhahran in Saudi Arabia, the model was validated, achieving 97% accuracy within a range of ± 3 °C, and hence provides better predictions than other traditional models. A novel abductive network model is developed by [14] to forecast a day-ahead and hour-ahead temperature data. The aforementioned model was trained using 5 years of hourly data and tested for the next whole year, for which the mean absolute errors (MAEs) of day-ahead and hour-ahead forecasts were 1.68°F and 1.05°F, respectively. Multiple linear regression and three artificial neural networks—feed-forward backpropagation (FFBP), radial basis function (RBF), and generalized regression neural network (GRNN)—have also been applied to model daily minimum, average, and maximum temperatures [15]. These models are evaluated against the geographical location

of Turkey’s most important areas for agricultural production—the Geyve and Sakarya basins—which are located in the southeast of the Marmara region. The study reveals that the three ANN models outperform the classical benchmark.

A comparative study of data-intelligent models, namely a generalized regression neural network (GRNN), multivariate adaptive regression splines (MARSs), random forest (RF), and extreme learning machines (ELMs), is presented by [16], in which longitude, latitude, altitude, and periodicity are used as covariates rather than traditional atmospheric features such as humidity, precipitation, etc. These models were implemented at eleven different sites in Madhya Pradesh, central India, and the results show that the GRNN works more effectively for modeling air temperature, particularly in data-sparse regions where only geographic and topographic factors are utilized for temperature forecasting. One can also read [17] for an in-depth review of air temperature modeling using machine learning and neural network models, respectively. A systematic literature review (from 2019 to 2024) in connection with modeling and forecasting daily temperature is also given in Table 1.

Table 1. Summary of literature review.

No.	Authors	Location	Forecasting	Data Length (Years)	Methods	Error Metrics
1	Shah et al. (2024) [18]	Islamabad, Pakistan	Day-ahead	5	FAR, ARIMA, VAR	MAE, RMSE, MAPE
2	Selmy et al. (2024) [19]	Delhi	Day-ahead	5	ARIMA, SARIMA, LSTM, CNN-LSTM,	MAE, RMSE, MAPE
3	Alexander et al. (2024) [20]	Jena, Germany	Day-ahead	8	ARIMA, WaveNet, LSTM	MAE, RMSE, MSE
4	Uluocak et al. (2024) [21]	Adana and Ankara, Türkiye	Day-ahead	8	GRU-CNN, LSTM-CNN, FNN, ANFIS, ARMA, GRU, LSTM, CNN	MAE, RMSE, R ² , NSE
5	An et al. (2023) [22]	China	Day-ahead	10	MLR, SVR, GBRT, LSTM, MLP, GFS	MAE
6	Sen et al. (2023) [23]	Ottawa, Canada	Hour-ahead	11	ARIMA, ANN, LSTM, GRU, GA, DE, PSO	MAPE, MSE
7	Toharudin et al. (2023) [24]	Bandung, Indonesia	Day-ahead	5.5	LSTM, Facebook Prophet	RMSE
8	Elshewey et al. (2022) [25]	Delhi, India	Day-ahead	5	WD-SARIMAX	MAE, RMSE, R ² , MAPE, MSE, MedAE,
9	Gong et al. (2022) [26]	Europe	Hour-ahead	13	ConvLSTM, SAVP	MSC, ACC, SSIM, rG
10	Shin et al. (2022) [27]	South Korea	Day-ahead	6	ANN, DNN, ELM, LSTM, LSTM-PC	MAE, RMSE, R, Thiel’s U-statistic
11	Alomar et al. (2022) [28]	North America	Day-ahead	22	SVR, RT, QRT	MAE, RMSE, R, Thiel’s U-statistic
12	Haque et al. (2021) [29]	Beijing, China, Toronto, Las Vegas, Seattle, Dallas	Hour-ahead	4	SRN, GRU, LSTM, CNN, CNN-LSTM, GRU-LSTM	MAE, RMSE, R ²
13	Wang et al. (2021) [30]	Michigan, United States	3 days ahead	6	Exponential Smoothing	RMSE
14	Lee et al. (2020) [31]	South Korea	Day- and hour-ahead	10	MLP, RNN, CNN	MAE
15	Zhang et al. (2020) [32]	Mainland China	Day-ahead	67	CRNN	MAE, RMSE
16	Shin et al. (2020) [33]	South Korea	Day-ahead	3	Hybrid (GloSea5GC2, RELM), Climatology Model	RMSE
17	Wanishsakpong et al. (2020) [34]	Ranong and Phuket, Thailand	Month-ahead	10	ARIMA, ARIMAX	RMSE, RRMSE
18	Zahroh et al. (2019) [35]	Bandung, Indonesia	Day-ahead	5.5	LSTM	MAPE

Vector Autoregressive (VAR), Gated Recurrent Unit (GRU), Wavelet Decomposition (WD), Stochastic Adversarial Video Prediction (SAVP), Anomaly Correlation Coefficient (ACC), Structural Similarity Index (SSIM), Gradient Ratio (rG), Support Vector Regression (SVR), Regression Tree (RT), Quantile Regression Tree (QRT), Multilayer Perceptron (MLP), Convolutional Neural Network (CNN), (a Global Climate model (GloSea5GC2), Regularized Extreme Learning Machine (RELM), Genetic Algorithm (GA), Differential Evolution (DE), and Particle Swarm Optimization (PSO).

Accurate day-ahead temperature forecasts are essential for energy, agriculture, and public health, as they help optimize resource management and reduce climate-related risks. Traditional methods often struggle to capture the continuous and complex nature of temperature variations. To overcome this deficiency, this research work contributes an effort to model and forecast the daily temperature curves using functional data analysis (FDA). To this end, a novel functional autoregressive model $\text{FARX}(p, m, \mathbf{g}, \tau)$ is developed and the distribution of the modified functional final prediction error $\text{fFPEX}(p, m, \mathbf{g}, \tau)$ is derived for optimal model selection.

The rest of the article is structured as follows. Section 2 provides an introduction to functional data analysis along with the fundamental tools necessary for its implementation. Section 3 presents the development of a new functional model, including discussions on classical and functional competitors. An empirical application of the proposed methodology is provided in Section 4, where the results are analyzed in detail. Finally, Section 5 concludes the study.

2. Functional Data Analysis

FDA offers a strong and promising framework for treating complex and high-dimensional datasets. Within FDA, data are represented as smooth curves or functions rather than discrete points or vectors [36]. This technique contrasts with traditional approaches that often concentrate on summary statistics or discrete observations. There are diverse applications of FDA in various disciplines such as medicine, finance, neuroscience, economics, environmental studies, and quality control [37,38]. A systematic review about the applications of the FDA in various fields for the period from 1995 to 2010 is given by [39]. A recent study [40] involves forecasting crude oil prices using derivative information in a multivariate empirical mode decomposition (MEMD) model within the FDA framework, which results in better performance as compared with traditional time series models. An application of the FDA to analyze the behavioral patterns of the Colombian stock market, specifically the effects of COVID-19 and the aftermath of the Ukraine war, is conducted by [41]. Ref. [42] employed the FDA approach to predict age-specific brain cancer mortality trends for planning public health policies and resource allocation. FDA has also significantly contributed to numerical weather prediction. For example, it is utilized to create point and interval forecasts for drought time series patterns [43]. They developed a reliable predictive method for predicting drought intervals. Based on the FDA technique, ref. [44] illustrated the temporal changes of rainfall to assist in forecasting the future and to achieve a clear understanding of rainfall patterns. Ref. [45] implemented the FDA framework to forecast temperature and precipitation under various climate change scenarios. Ref. [46] performed FPCA to model the structure of rainfall. A functional autoregressive model of the first order, FAR(1), is explored to forecast hourly air temperature up to 24 h ahead, where the traditional time series model ARIMA served as the competitor, and the results suggest that FAR(1) is superior. Online auction prices and traffic volume are also predicted under the FDA framework by [47–49], respectively. Due to substantial fluctuations, ref. [50] indicates a non-linear and complex structure in electricity consumption, demand, and prices. Ref. [51–53] applied the FDA to estimate and forecast electricity consumption and demand, respectively. Hour-ahead electricity demand forecasting strategies are implemented using non-parametric FDA [54–57], and the results significantly outperform those obtained from the classical seasonal ARIMA (SARIMA) models.

2.1. Preliminaries

Let $\{S_t(v) : t \in \mathbb{Z}, v \in [0, 1]\}$ denote a stationary functional time series of observations, where t represents the time index and v the argument in the continuous domain of each function. Each S_t is assumed to belong to the Hilbert space $\mathcal{H} = L^2([0, 1])$ with the inner product $\langle x, y \rangle = \int_0^1 x(v)y(v) dv$. Thus, each S_t is square-integrable, satisfying $\|S_t\|^2 = \int_0^1 S_t^2(v) dv < \infty$, and all random functions are defined on a common probability space $(\Omega, \mathcal{A}, P)$. For $p > 0$, we write $S_t \in L^p_{\mathcal{H}}$ if $E[\|S_t\|^p] < \infty$. Since the series is not necessarily centered, we define the mean function $\lambda(v) = E[S_t(v)]$, $v \in [0, 1]$, and the covariance operator of the centered functions $S_t - \lambda$ as:

$$C(x)(v) = E[\langle S_t - \lambda, x \rangle (S_t(v) - \lambda(v))], \quad x \in \mathcal{H}$$

or equivalently via its kernel representation

$$C(x)(v) = \int_0^1 c(v, \gamma) x(\gamma) d\gamma, \quad x \in \mathcal{H} \quad (1)$$

with

$$c(v, \gamma) = \text{Cov}(S_t(v), S_t(\gamma)). \quad (2)$$

Equation (1) expresses the covariance operator C as an integral operator with kernel $c(v, \gamma)$, while Equation (2) defines this kernel as the covariance between the functional observations at points v and γ . For a more detailed discussion of covariance operators in functional time series, the reader is referred to [58].

2.2. Basis Function System

In the FDA framework, a basis function system (BFS) plays a crucial role in converting discrete data into a functional object, enabling a more compact representation without losing essential information. Various types of basis function systems exist, including polynomial, B-spline, and Fourier basis functions (FBFs), each suitable for different data structures. Among these, FBF, which consist of sine and cosine waves with increasing frequency, are particularly effective for modeling periodic data. FBFs are especially useful for cyclical environmental data, such as temperature, wind speed, and atmospheric pressure, where seasonal and diurnal variations play a crucial role. The sinusoidal components of the FBF effectively capture periodic fluctuations, providing a smooth and continuous representation of temperature profiles. A Fourier basis function is defined as

$$\xi_j(v) = \cos(j\varphi(v)) + \sin(j\varphi(v)), \quad (3)$$

where j denotes the Fourier frequency index and $\varphi(v)$ is the fundamental frequency function, typically defined as $\varphi(v) = 2\pi v$, $v \in [0, 1]$. Using M Fourier basis functions, the functional observation at time t can be approximated as

$$S_t(v) = \sum_{j=1}^M b_{t,j} \xi_j(v), \quad v \in [0, 1] \quad (4)$$

where $b_{t,j}$ are time-varying Fourier coefficients and $\{\xi_j(v)\}_{j=1}^M$ form the Fourier basis function system. An example of functional data, based on $M = 20$ FBFs represented in Figure 1, for the years 2019 to 2024 reflects 2192 daily temperature curves in Figure 2. The $S_t(v)$ is known as the functional time series (FTS) where one can investigate the major source of variation among different temperature curves.

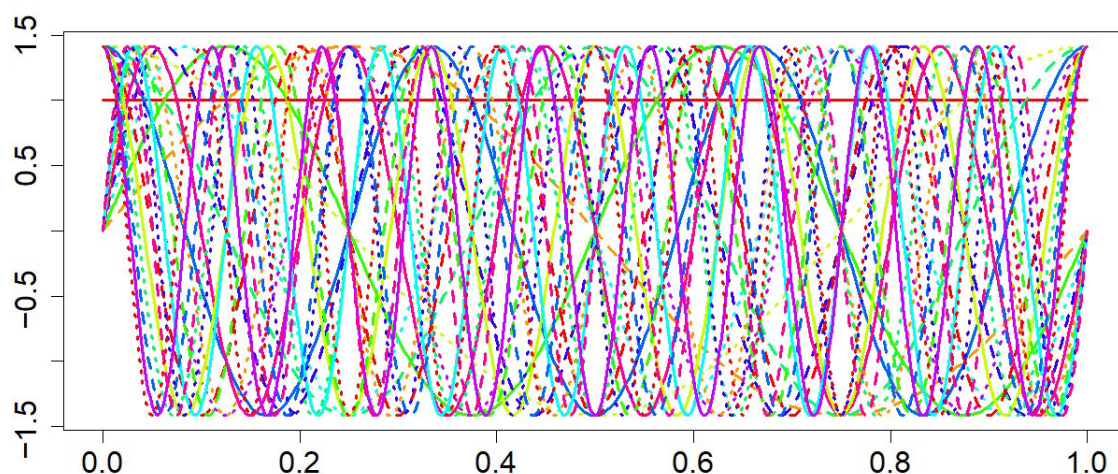


Figure 1. The Fourier basis functions using $M = 20$.

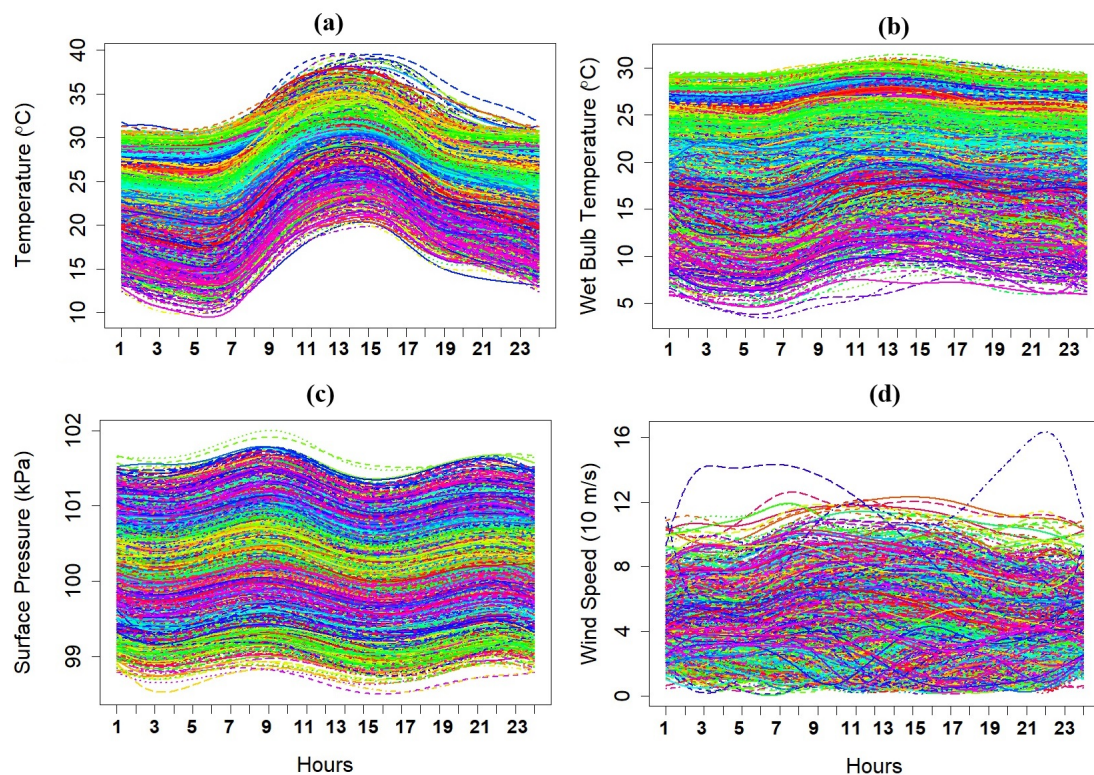


Figure 2. Functional data representation of (a) response variable; (b) first exogenous variable; (c) second exogenous variable; (d) third exogenous variable, under study using Fourier basis functions.

2.3. Functional Principal Component Analysis (FPCA)

Suppose we observe an FTS, $S_t(v)$, consisting of n curves, where $v \in [0, 1]$. The mean function is defined as $\lambda(v) = \frac{1}{n} \sum_{t=1}^n S_t(v)$. The mean-centered curves are then given by

$$S_t^*(v) = S_t(v) - \lambda(v). \quad (5)$$

Let $\alpha_j(v)$ denote the j th FPC. Each mean-centered curve $S_t^*(v)$ can be projected onto $\alpha_j(v)$ in the Hilbert space $\mathcal{H} = L^2([0, 1])$. The corresponding FPC score is defined as

$$\theta_{j,t} = \langle S_t^*, \alpha_j \rangle = \int_0^1 S_t^*(v) \alpha_j(v) dv. \quad (6)$$

The FPCs satisfy the normalization condition $\int_0^1 \alpha_j^2(v) dv = 1$ ensuring the identifiability of the scores. The covariance operator $C : \mathcal{H} \rightarrow \mathcal{H}$ of the mean-centered process is defined as $C(x)(v) = \mathbb{E}[\langle S_t^*, x \rangle S_t^*(v)]$, $x \in \mathcal{H}$. According to Mercer's theorem, C admits the spectral decomposition

$$C(x)(v) = \sum_{j=1}^{\infty} \delta_j \langle x, \alpha_j \rangle \alpha_j(v), \quad x \in \mathcal{H} \quad (7)$$

where $\{\delta_j\}_{j \geq 1}$ are the eigenvalues in decreasing order and $\{\alpha_j\}_{j \geq 1}$ are the corresponding orthonormal eigenfunctions [59]. Using the Karhunen–Loève expansion, each functional observation can be expressed as

$$S_t(v) = \lambda(v) + \sum_{j=1}^{\infty} \theta_{j,t} \alpha_j(v), \quad (8)$$

where $\mathbb{E}[\theta_{j,t}] = 0$ and $\text{Var}(\theta_{j,t}) = \delta_j$. For dimension reduction, we retain the first m FPCs and obtain the approximation [60,61]

$$S_t(v) = \lambda(v) + \sum_{j=1}^m \theta_{j,t} \alpha_j(v) + \sum_{j=m+1}^{\infty} \theta_{j,t} \alpha_j(v), \quad (9)$$

where the third term in the above expression accounts for the variability not explained by the first m functional principal components and has zero mean and finite variance [62]. In practice, the mean function and covariance kernel are estimated as

$$\hat{\lambda}(v) = \frac{1}{n} \sum_{t=1}^n S_t(v), \quad (10)$$

$$\hat{C}(v, \gamma) = \frac{1}{n-1} \sum_{t=1}^n [S_t(v) - \hat{\lambda}(v)] [S_t(\gamma) - \hat{\lambda}(\gamma)], \quad (11)$$

Let $\hat{\delta}_j$ and $\hat{\alpha}_j(v)$ denote the eigenvalues and eigenfunctions of $\hat{C}$. The reconstructed FTS is then given by

$$\hat{S}_t(v) = \hat{\lambda}(v) + \sum_{j=1}^m \hat{\theta}_{j,t} \hat{\alpha}_j(v), \quad (12)$$

where $\hat{\theta}_{j,t} = \int_0^1 [S_t(v) - \hat{\lambda}(v)] \hat{\alpha}_j(v) dv$.

3. Functional Time Series Modeling

An FTS captures nonlinear dependencies while preserving smooth temporal structures, making it particularly suitable for modeling complex dynamic processes. This section discusses the existing methodologies and the new developed approach which can effectively handle such dependencies. Despite its statistical relevance and mathematical appeal, functional time series modeling has practical limitations. Few user-friendly software packages exist, with notable exceptions being the *far* [63] and *fts* [64] packages in R version 4.5.0 (2025-04-11 ucrt). The lack of dedicated tools often requires manual implementation, limiting its use to academicians. This article aims to bridge this gap by introducing a prediction algorithm.

3.1. Functional Autoregressive Models

An autoregressive (AR) model is a well-known linear model where the response variable is regressed over its lagged values plus a noise term [65,66]. If the response observations are functions defined over a continuous domain (rather than scalar or vector values), the autoregressive model is referred to as an FAR model. An FAR process and its theoretical properties are described in a Hilbert space framework, which is one of the best techniques for modeling the complex nature of FTS [67]. In traditional univariate and multivariate time series analysis, reliable predictions rely on recursive methods like the Durbin–Levinson algorithm and the innovations algorithm [68], which systematically update predictions based on past observations. Prediction equations for general stationary FTS can be derived explicitly. However, their practical implementation is challenging due to the infinite-dimensional nature of functional data. As a result, most of the research in this area is limited to FAR(1) models primarily due to the computational and theoretical challenges of estimating higher-order FAR(p) models, as functional time series are infinite-dimensional and more complex models require substantial data and computation for reliable parameter estimation. Mathematically, the FAR(1) model [69] is given in Equation (13) as

$$S_t(v) = \lambda(v) + \phi S_{t-1}(v) + \mathcal{N}_t(v), \quad (13)$$

where $\lambda(v)$ is the functional mean curve, $S_{t-1}(v)$ is the first lagged functional variable, $\mathcal{N}_t(v)$ is i.i.d. in $L^2_{\mathcal{H}}$ such that $E[\mathcal{N}_t(v)] = 0$, and ϕ is the linear operator bounded over mapping $\mathcal{H} \rightarrow \mathcal{H}$ such that Equation (13) has a unique solution. If the FAR(1) approach is infeasible, the multiple testing procedure ensures a better fit to the data and can be used to determine the optimal order p [70]. A more general FAR(p) process is given by Equation (14), in which the kernel function yields a test statistic following an approximate chi-square distribution, with the degrees of freedom based on the number of FPCs being

$$S_t(v) = \lambda(v) + \sum_{i=1}^p \phi_i S_{t-i}(v) + \mathcal{N}_t(v), \quad (14)$$

where ϕ_i are the linear operators of corresponding p -lagged functional variables $S_{t-i}(v)$. Another approach for FAR(p, m) based on multivariate time series and FPCA, in which a VAR model is used with FPC scores. To implement the VAR model, an automatic procedure is suggested for selecting the optimal lag order (p) and dimension (m) by minimizing the final functional prediction error (ffPE) given in Equation (15).

$$ffPE(p, m) = \left[\frac{n + p \times m}{n - p \times m} \right] \text{trace}(\widehat{\Sigma}_{\mathcal{N}}) + \sum_{j=m+1}^{\infty} \hat{\delta}_j, \quad (15)$$

where the first part is the product of a penalty term and the trace of an estimated covariance matrix of the residuals from a VAR model fitted to the FPC scores. n is the sample size in the penalty term. The second term represents the sum of eigenvalues associated with ignored principal components. The values of m and p , which minimize the ffPE, are the optimal values that are used in FAR(p, m). The predictive performance of FAR(p, m) is also increased by incorporating the τ number scalar covariates, namely the model FARX(p, m, τ) given as in Equation (16)

$$S_t(v) = \lambda(v) + \sum_{i=1}^p \phi_i S_{t-i}(v) + \sum_{k=1}^{\tau} \zeta_k \mathbb{Z}_k + \mathcal{N}_t(v), \quad (16)$$

where ζ_k are coefficients of $\mathbb{Z}_k$ scalar covariates. Hence, the optimal orders p and m for Equation (16) are obtained by minimizing the ffPE, adjusted for degrees of freedom for scalar covariates as

$$ffPE(p, m) = \left[\frac{n + p \times m + \tau}{n - p \times m - \tau} \right] \text{trace}(\widehat{\Sigma}_{\mathcal{N}}) + \sum_{j=m+1}^{\infty} \hat{\delta}_j. \quad (17)$$

For detailed discussions about Equations (13)–(17) and their applications to real data, one can consult [18,52,58] and the references therein.

3.2. Building FARX($p, m, \underline{\mathbf{g}}, \tau$) Model

To enhance the forecasting accuracy of the FARX(p, m, τ) model, this work introduces a novel functional model, denoted as FARX($p, m, \underline{\mathbf{g}}, \tau$). This model incorporates $\underline{\mathbf{g}} = [g_1 \ g_2 \ \dots \ g_p]'$, where each g represents the optimum dimensions of corresponding functional exogenous variables, extending the standard framework. Additionally, a modified version of ffPE(p, m), is derived, leading to the formulation of ffPEX($p, m, \underline{\mathbf{g}}, \tau$) for improved model selection.

Suppose a functional stationary response $S_t(v)$ and functional exogenous variables $\mathbb{Y}_1(v), \mathbb{Y}_2(v), \dots, \mathbb{Y}_\rho(v)$ are given and the goal is to derive an empirical $\hat{S}_{t+1}(v)$. Then Equation (17) can be written as a FARX($p, m, \underline{g}, \tau$) given as

$$S_t(v) = \lambda(v) + \sum_{i=1}^p \phi_i S_{t-i}(v) + \sum_{l=1}^{\rho} \psi_l \mathbb{Y}_l(v) + \sum_{k=1}^{\tau} \zeta_k \mathbb{Z}_k + \mathcal{N}_t(v), \quad (18)$$

where ψ_l are the functional operators of the ρ functional covariates $\mathbb{Y}_l(v)$ of the above model. The rest of the notations are same as discussed in Section 3.1. In this work, we estimate the Equation (18) model using the FPCA approach. This method ensures both accurate model estimation and computational efficiency, as the first few FPCs capture most of the variation in the endogenous and exogenous functional variables. In this context, the identification step involves selecting the optimal dimensions m and lag order p for an endogenous functional variable and a set of optimal dimensions $\underline{g} = [g_1 \ g_2 \ \dots \ g_\rho]'$ for a ρ number of exogenous functional variables. To achieve this, the following steps are employed for selecting $p, m, \underline{g}$ and τ while a flowchart is given in Figure 3.

1. Based on the algorithm in Section 3.3, fix the dimension m of the functional endogenous variable. Using FPCA based on m dimensions, the estimated FPC scores are obtained $\hat{s}_{j,t} = \int \hat{S}_t(v) \hat{\alpha}_j(v) d(v)$ for each of $\hat{S}_t(v)$ such that we have vectors of j -variate-estimated FPC scores $\hat{\mathbf{S}}_t = [\hat{s}_{1,t} \ \hat{s}_{2,t} \ \dots \ \hat{s}_{m,t}]'$ where $t = 1, 2, 3, \dots, k$ and $j = 1, 2, 3, \dots, m$.
2. Fix the dimensions $\underline{g} = [g_1 \ g_2 \ \dots \ g_\rho]'$ for a ρ number of functional exogenous variables and obtain the estimated FPC scores for each of the $\mathbb{Y}_l(v)$.

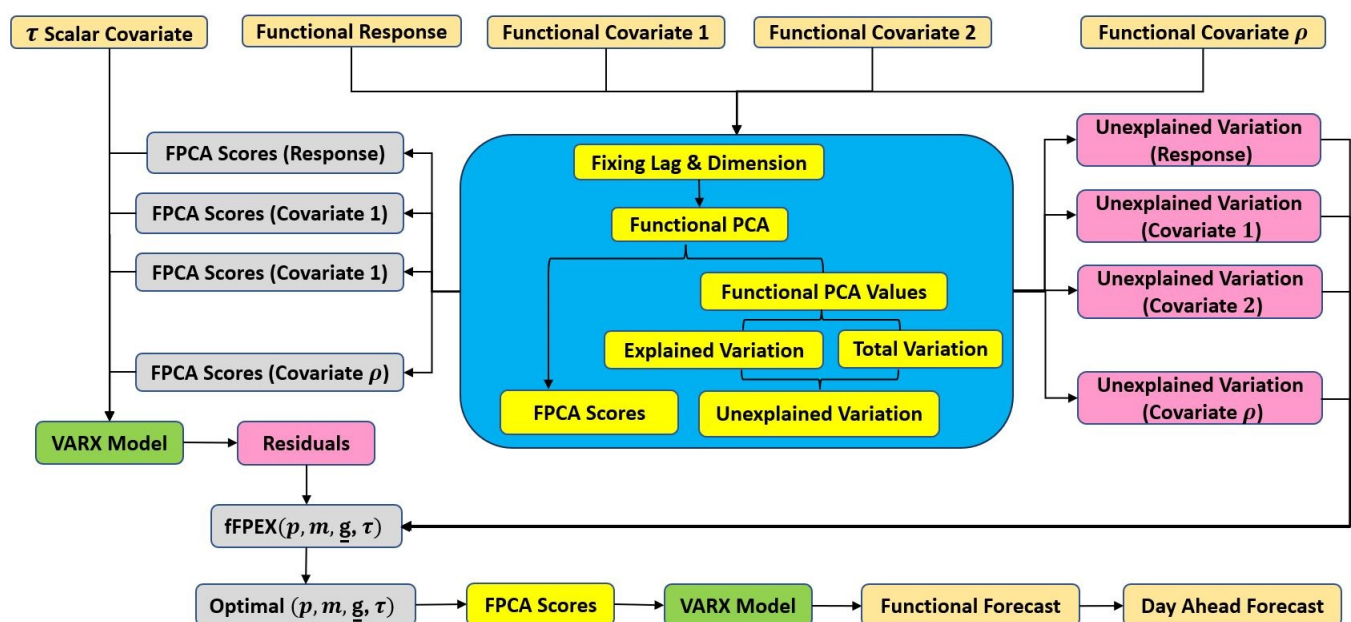


Figure 3. Framework for the proposed functional model.

$$\hat{y}_{j_1,t} = \int \hat{\mathbb{Y}}_1(v) \hat{\alpha}_{j_1}(v) d(v), j_1 = 1, 2, 3, \dots, g_1$$

$$\hat{y}_{j_2,t} = \int \hat{\mathbb{Y}}_2(v) \hat{\alpha}_{j_2}(v) d(v), j_2 = 1, 2, 3, \dots, g_2$$

and so on,

$$\hat{y}_{j_\rho,t} = \int \hat{\mathbb{Y}}_\rho(v) \hat{\alpha}_{j_\rho}(v) d(v), j_\rho = 1, 2, 3, \dots, g_\rho$$

Such that we have vectors of $j_1, j_1, \dots, j_\rho$ -variate-estimated FPC scores, respectively.

$$\hat{\mathcal{Y}}_{t_1} = [\hat{y}_{1,t_1} \quad \hat{y}_{2,t_1} \quad \dots \quad \hat{y}_{g_1,t_1}]'$$

$$\hat{\mathcal{Y}}_{t_2} = [\hat{y}_{1,t_2} \quad \hat{y}_{2,t_2} \quad \dots \quad \hat{y}_{g_2,t_2}]'$$

and so on,

$$\hat{\mathcal{Y}}_{t_\rho} = [\hat{y}_{1,t_\rho} \quad \hat{y}_{2,t_\rho} \quad \dots \quad \hat{y}_{g_\rho,t_\rho}]'$$

3. If the number of scalar exogenous variables τ is sufficiently large, then it is better to use their functional form instead of their scalar form. That is, fix the dimensions τ using the same algorithm as described in Section 3.3. Using FPCA based on τ dimensions, the estimated FPC scores are obtained $\hat{z}_{j,t} = \int \hat{\mathbb{Z}}_{\mathbf{k}}(v) \hat{\alpha}_j(v) d(v)$ for $\hat{\mathbb{Z}}_{\mathbf{k}}(v)$, such that we have vectors of j -variate-estimated FPC scores $\hat{\mathbf{Z}}_t = [\hat{z}_{1,t} \quad \hat{z}_{2,t} \quad \dots \quad \hat{z}_{\tau,t}]'$ where $t = 1, 2, 3, \dots, k$ and $j = 1, 2, 3, \dots, \tau$.
4. The first and second derivatives of functional data capture essential dynamic features, such as trends and curvature, making them valuable predictors for improving the accuracy and interpretability of functional response models. Therefore, using the same idea as discussed earlier, one can also use the derivatives of functional variables as predictors.
5. Once all vectors of estimated FPC scores from all functional and scalar exogenous variables are obtained, combine them into a single vector as follows: $\mathbf{C}_t = [\hat{\mathcal{Y}}_{t_1} \quad \hat{\mathcal{Y}}_{t_2} \quad \dots \quad \hat{\mathcal{Y}}_{t_\rho} \quad \hat{\mathbf{Z}}_t]'$.
6. Next, fix the lag order p and using the estimated vector of FPC scores of endogenous variable $\hat{\mathbf{S}}_t = [\hat{s}_{1,t} \quad \hat{s}_{2,t} \quad \dots \quad \hat{s}_{m,t}]'$ obtained in Step (1) and the $\mathbf{C}_t$ from Step (5), fit an appropriate multivariate model, for example the VAR model with an exogenous variable (VARX), given as

$$\mathbf{S}_t = \sum_{i=1}^p \Psi_i \mathbf{S}_{t-i} + \Gamma \mathbf{C}_t + \mathcal{N}_t$$

where Γ is a matrix of coefficients of $\mathbf{C}_t$ and $\mathcal{N}_t$ is a white noise process. Then obtain a one-step-ahead forecast for $\hat{\mathbf{S}}_{t+1}$ as

$$\hat{\mathbf{S}}_{t+1} = [\hat{s}_{1,t+1} \quad \hat{s}_{2,t+1} \quad \dots \quad \hat{s}_{m,t+1}]'$$

7. In the last step, the $\hat{\mathbf{S}}_{t+1}$ is reverted to a functional object using the KL theorem and a one-step-ahead forecast in a functional form $\hat{S}_{t+1}(v)$ is obtained as

$$\hat{S}_{t+1}(v) = \hat{\lambda}(v) + \sum_{j=1}^m \hat{s}_{t+1,j} \hat{\alpha}_j(v).$$

3.3. Selection of Optimal Orders $p, m, \underline{\mathbf{g}}, \tau$ by ffPE $\mathbf{X}(p, m, \underline{\mathbf{g}}, \tau)$

The performance of $\text{FAR}(p, m)$ clearly depends on the optimal selection of the order parameters p and m representing the lag and appropriate number of FPCs, respectively. Past studies suggest some techniques to get these optimal values for the FAR models. For example, the value of m was suggested by [70], based on multistage hypothesis testing. Also, a mechanical and automated technique is suggested for getting an optimum p and m based on minimizing the mean squared error (MSE) as proposed by [58], and called $\text{ffPE}(p, m)$ in Equations (15) and (17), as discussed in Section 3.1. In this section, we demonstrate a modified $\text{ffPE}(p, m)$ called $\text{ffPEX}(p, m, \underline{\mathbf{g}}, \tau)$, which is minimized for the

selection of optimal parameter values $p, m, \mathbf{g}, \tau$ where $\mathbf{g} = [g_1 \ g_2 \ \dots \ g_\rho]'$. We begin by analyzing the MSE. Since the eigenfunctions α_j are orthogonal and the FPCs $s_{j,t}$ are uncorrelated, we can partition the MSE as follows:

$$\begin{aligned} E\left\{\left\|S_{t+1}(v) - \hat{S}_{t+1}(v)\right\|^2\right\} &= E\left\{\left\|\sum_{j=1}^{\infty} s_{t+1,j}\alpha_j - \sum_{j=1}^m \hat{s}_{t+1,j}\alpha_j\right\|^2\right\} \\ &= E\left\{\left\|\mathbf{S}_{t+1} - \hat{\mathbf{S}}_{t+1}\right\|^2\right\} + \sum_{j=m+1}^{\infty} \delta_j + \sum_{l=1}^{\rho} \sum_{r=g_l+1}^{\infty} \mathcal{Z}_{l,r}. \end{aligned} \quad (19)$$

Here, the usual L^2 Euclidean norm of vectors is represented by $\|\cdot\|^2$. In the above equation, the first term $E\left\{\left\|\mathbf{S}_{t+1} - \hat{\mathbf{S}}_{t+1}\right\|^2\right\}$ represents the finite-dimensional MSE for the approximation based on the truncated score vectors. Essentially, it quantifies the error within the selected subspace of dimension m . This finite-dimensional space contains the most significant components (typically derived from the eigenfunctions of the functional data), allowing us to approximate $\mathbf{S}_{t+1}$ using a lower-dimensional projection. The second term $\sum_{j=m+1}^{\infty} \delta_j$ captures the information loss due to truncation at dimension m . Here, δ_j represents the eigenvalues associated with the components beyond m . By truncating at m , we exclude components with eigenvalues $\delta_{m+1}, \delta_{m+2}, \dots$ which means losing some functional information, as these components are no longer part of the approximation. The third term $\sum_{l=1}^{\rho} \sum_{r=g_l+1}^{\infty} \mathcal{Z}_{l,r}$ captures the information loss due to truncation at dimension $\mathbf{g} = [g_1 \ g_2 \ \dots \ g_\rho]'$ based on ρ number of functional covariates. Here l is the index of the functional covariate ranging from 1 to ρ , while r indexes the higher-order terms beyond the optimal dimension g_l . $\mathcal{Z}_{l,r}$ captures the information loss due to truncating the l^{th} functional covariate at its g_l^{th} dimension and corresponds to the eigenvalue of the r^{th} principal component of that covariate, quantifying the variance it explains. Summing over all components beyond g_l accounts for the total information lost due to dimension reduction, analogous to $\sum_{j=m+1}^{\infty} \delta_j$ for the response. Hence, the third term represents the cumulative contribution of the truncated higher-order components of the functional covariates beyond the selected optimal dimensions $\mathbf{g}$. Assuming the stationarity of the process $\mathbf{S}_t$ is a $(m + \mathbb{K} + \tau)$ -variate VARX(p), where $\mathbb{K} = \sum_{l=1}^{\rho} g_l$, has the following form:

$$\mathbf{S}_{t+1} = \mathbf{\Psi}_1 \mathbf{S}_t + \mathbf{\Psi}_2 \mathbf{S}_{t-1} + \dots + \mathbf{\Psi}_p \mathbf{S}_{t-p+1} + \mathbf{\Gamma} \mathbf{C}_t + \mathcal{N}_{t+1},$$

where $\mathbf{\Gamma}$ is a matrix of coefficients of $\mathbf{C}_t = [\hat{\mathbf{y}}_{t_1} \ \hat{\mathbf{y}}_{t_2} \ \dots \ \hat{\mathbf{y}}_{t_\rho} \ \hat{\mathbf{z}}_t]'$ and $\mathcal{N}_t$ is a white noise process such that

$$\sqrt{n}(\widehat{\mathcal{W}} - \mathcal{W}) \xrightarrow{D} N_{(p \times m^2) + \mathbb{K} + \tau} \left(\mathbf{0}, \Sigma_{\mathcal{N}} \otimes \Delta_{(p \times m + \mathbb{K} + \tau)}^{-1} \right), \quad (20)$$

where

$$\widehat{\mathcal{W}} = \text{vec} \left([\hat{\mathbf{\Psi}}_1, \dots, \hat{\mathbf{\Psi}}_p]', [\hat{\mathbf{\Gamma}}_1, \dots, \hat{\mathbf{\Gamma}}_{\mathbb{K} + \tau}]' \right)$$

is the least squares estimator of $\mathcal{W} = \text{vec} \left([\mathbf{\Psi}_1, \dots, \mathbf{\Psi}_p]', [\mathbf{\Gamma}_1, \dots, \mathbf{\Gamma}_{\mathbb{K} + \tau}]' \right)$, $\sqrt{n}(\widehat{\mathcal{W}} - \mathcal{W})$ is the difference between the estimator $\widehat{\mathcal{W}}$ and the true parameter vector $\mathcal{W}$ scaled by $\sqrt{n}$ and $\Delta_{(p \times m + \mathbb{K} + \tau)} = \text{Var}[\text{vec}(\mathbf{S}_1, \dots, \mathbf{S}_p)]$, and $\Sigma_{\mathcal{N}} = E[\mathcal{N}_1, \mathcal{N}_1']$. The asymptotic distribution suggests that as $n \rightarrow \infty$, this scaled difference converges to a normal distribution with a zero mean and a covariance structure-determined Kronecker product ($\otimes$) of $\Sigma_{\mathcal{N}}$ and $\Delta_{(p \times m + \mathbb{K} + \tau)}^{-1}$. Suppose that the $\widehat{\mathcal{W}}$ are estimated from an independent training sample $(\mathbf{R}_1, \dots, \mathbf{R}_t) \stackrel{D}{=} (\mathbf{S}_1, \dots, \mathbf{S}_t)$. It follows then that

$$\begin{aligned}
& \mathbb{E} \left\{ \left\| \mathbf{S}_{t+1} - \widehat{\mathbf{S}}_{t+1} \right\|^2 \right\} = \mathbb{E} \left\{ \left\| \mathbf{S}_{t+1} - \left(\widehat{\Psi}_1 \mathbf{S}_t + \widehat{\Psi}_2 \mathbf{S}_{t-1} + \dots + \widehat{\Psi}_p \mathbf{S}_{t-p+1} + \widehat{\Gamma} \mathbf{C}_t \right) \right\|^2 \right\} \\
& \mathbb{E} \left\{ \left\| \mathbf{S}_{t+1} - \widehat{\mathbf{S}}_{t+1} \right\|^2 \right\} \\
& = \mathbb{E} \left\{ \left\| \mathcal{N}_{t+1} \right\|^2 \right\} + \mathbb{E} \left\{ \left\| (\Psi_1 - \widehat{\Psi}_1) \mathbf{S}_t + \dots + (\Psi_p - \widehat{\Psi}_p) \mathbf{S}_{t-p+1} + \sum_{t=1}^{\mathbb{K}+\tau} (\Gamma_t - \widehat{\Gamma}_t) \mathbf{C}_t \right\|^2 \right\} \\
& = \text{trace}\{\Sigma_{\mathcal{N}}\} + \mathbb{E} \left\{ \left\| \left[I_{(p \times m + \mathbb{K} + \tau)} \otimes (\mathbf{S}'_t, \dots, \mathbf{S}'_{t-p+1}, \mathbf{C}'_1, \dots, \mathbf{C}'_{\mathbb{K}+\tau}) \right] (\mathcal{W} - \widehat{\mathcal{W}}) \right\|^2 \right\}
\end{aligned} \tag{21}$$

The first term $\text{trace}\{\Sigma_{\mathcal{N}}\}$ is the intrinsic noise variance from the white noise process $\mathcal{N}_{t+1}$ and second term is obtained from the estimation error, which is expressed compactly using the Kronecker product ($\otimes$) and vector notation $(\mathbf{S}'_t, \dots, \mathbf{S}'_{t-p+1}, \mathbf{C}'_1, \dots, \mathbf{C}'_{\mathbb{K}+\tau})$ represents the past values in the autoregressive component, and combined from a reduced dimension from functional covariates and scalar exogenous variables. The independence of $\widehat{\mathcal{W}}$ and $(\mathbf{S}_1, \dots, \mathbf{S}_t, \mathbf{C}_t)$ yields that

$$\begin{aligned}
& \mathbb{E} \left\{ \left\| \left[I_{(p \times m + \mathbb{K} + \tau)} \otimes (\mathbf{S}'_t, \dots, \mathbf{S}'_{t-p+1}, \mathbf{C}'_1, \dots, \mathbf{C}'_{\mathbb{K}+\tau}) \right] (\mathcal{W} - \widehat{\mathcal{W}}) \right\|^2 \right\} \\
& = \mathbb{E} \left[\text{trace} \left\{ (\mathcal{W} - \widehat{\mathcal{W}})' \left[I_{(p \times m + \mathbb{K} + \tau)} \otimes \Delta_{(p \times m + \mathbb{K} + \tau)} \right] (\mathcal{W} - \widehat{\mathcal{W}}) \right\} \right] \\
& = \text{trace} \left[\left(I_{(p \times m + \mathbb{K} + \tau)} \otimes \Delta_{(p \times m + \mathbb{K} + \tau)} \right) \right] \mathbb{E} \left[(\mathcal{W} - \widehat{\mathcal{W}})(\mathcal{W} - \widehat{\mathcal{W}})' \right]
\end{aligned}$$

$I_{(p \times m + \mathbb{K} + \tau)} \otimes \Delta_{(p \times m + \mathbb{K} + \tau)}$ represents the overall variance–covariance structure for the autoregressive, functional exogenous, and scalar terms. By using Equation (20), the last term can be approximated as

$$\frac{1}{n} \left(\text{trace} \left[\Sigma_{\mathcal{N}} \otimes I_{(p \times m + \mathbb{K} + \tau)} \right] + o(1) \right) \sim \frac{(p \times m + \mathbb{K} + \tau)}{n} \text{trace}\{\Sigma_{\mathcal{N}}\}.$$

Using the above results, Equation (21) can be written as

$$\mathbb{E} \left\{ \left\| \mathbf{S}_{t+1} - \widehat{\mathbf{S}}_{t+1} \right\|^2 \right\} = \text{trace}\{\Sigma_{\mathcal{N}}\} + \frac{(p \times m + \mathbb{K} + \tau)}{n} \text{trace}\{\Sigma_{\mathcal{N}}\}.$$

Replacing $\text{trace}\{\Sigma_{\mathcal{N}}\}$ by $\frac{n}{n - (p \times m + \mathbb{K} + \tau)} \text{trace}\{\widehat{\Sigma}_{\mathcal{N}}\}$,

$$\begin{aligned}
& \mathbb{E} \left\{ \left\| \mathbf{S}^*_{t+1} - \widehat{\mathbf{S}}^*_{t+1} \right\|^2 \right\} \\
& = \frac{n}{n - (p \times m + \mathbb{K} + \tau)} \text{trace}\{\widehat{\Sigma}_{\mathcal{N}}\} + \frac{(p \times m + \mathbb{K} + \tau)}{n} \times \frac{n}{n - (p \times m + \mathbb{K} + \tau)} \text{trace}\{\widehat{\Sigma}_{\mathcal{N}}\} \\
& = \left[\frac{n}{n - (p \times m + \mathbb{K} + \tau)} + \frac{(p \times m + \mathbb{K} + \tau)}{n} \times \frac{n}{n - (p \times m + \mathbb{K} + \tau)} \right] \text{trace}\{\widehat{\Sigma}_{\mathcal{N}}\} \\
& = \left[\frac{n + (p \times m + \mathbb{K} + \tau)}{n - (p \times m + \mathbb{K} + \tau)} \right] \text{trace}\{\widehat{\Sigma}_{\mathcal{N}}\}.
\end{aligned}$$

Using the above results, Equation (19) can be written as

$$\mathbb{E} \left\| S_{t+1}(v) - \widehat{S}_{t+1}(v) \right\|^2 \approx \left[\frac{n + (p \times m + \mathbb{K} + \tau)}{n - (p \times m + \mathbb{K} + \tau)} \right] \text{trace}\{\widehat{\Sigma}_{\mathcal{N}}\} + \sum_{j=m+1}^{\infty} \delta_j + \sum_{l=1}^{\rho} \sum_{r=g_l+1}^{\infty} \mathcal{Z}_{l,r}$$

In practice, the population-level quantities in the second and third terms are not observable and are therefore replaced by their corresponding empirical estimates. Thus, the

orders $p, m, \mathbb{K} = [g_1 \ g_2 \ \dots \ g_p]'$ and τ are simultaneously selected by minimizing the functional final perdition error criterion generalized for exogenous covariates as

$$fFPEX(p, m, \underline{\mathbf{g}}, \tau) = \left[\frac{n + (p \times m + \mathbb{K} + \tau)}{n - (p \times m + \mathbb{K} + \tau)} \right] \text{trace}\{\hat{\Sigma}_{\mathcal{N}}\} + \sum_{j=m+1}^{\infty} \hat{\delta}_j + \sum_{l=1}^{\rho} \sum_{r=g_l+1}^{\infty} \hat{Z}_{l,r} \quad (22)$$

The implementation of the $fFPEX(p, m, \underline{\mathbf{g}}, \tau)$ criterion renders the proposed forecasting framework fully automated, eliminating reliance on arbitrary tuning parameter specification. Notably, the determination of the optima dimension are now sample-adaptive and explicitly dependent on the observed sample size n . The empirical results presented in Section 4 confirm the strong practical efficacy of this approach.

3.4. Competitive Models

The traditional time series model, like the Autoregressive Integrated Moving Average (ARIMA), the neural network autoregressive model (NNAR), and two functional autoregressive models FAR(p, m) and FARX(p, m, τ) in Equations (14) and (16), as described in Section 3.1, are studied as benchmarks to compare the forecasting accuracy of the proposed functional models.

3.4.1. Autoregressive Integrated Moving Average Model

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, which is a tool that extrapolates the signal into the future to generate forecasts by separating signal from noise. It comprises three parametric parts known as the order of AR, MA, and the number of differences needed for a time series to be stationary; i.e., $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{d}$, respectively. Mathematically, it can be written as

$$(1 - B)^{\mathbf{d}} S_{t,k} = \alpha + \sum_{h=1}^{\mathbf{p}} \beta_h (1 - B)^{\mathbf{d}} S_{t-h,k} + \mathcal{N}_t + \sum_{h'=1}^{\mathbf{q}} \phi_{h'} \mathcal{N}_{t-h'} \quad (23)$$

where $S_{t,k}$ is the stochastic part obtained from partitioning the $S_{t,k}$, $\mathbf{q}$ is an integrated difference to achieve a stationarity, B is the backward shift operator, $\mathcal{N}_t$ is the white noise term, α is an intercept term, and β_h ($h = 1, \dots, \mathbf{p}$) and $\phi_{h'}$ ($h' = 1, \dots, \mathbf{q}$) are the parameters of the AR and MA parts, respectively, which are estimated using the maximum likelihood estimation method.

3.4.2. Neural Network Autoregressive Model

Artificial neural networks (ANNs) allow complex nonlinear relationships between the response variable and its predictors. The strength of the 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 a sigmoid) between the input ($S_{t-1,k}, S_{t-2,k}, \dots, S_{t-d,k}$) and output ($S_{t,k}$) of a node and network. The mathematical form of the NNAR model is

$$S_{t,k} = w_0 + \sum_{a=1}^q \omega_a f \left(w_{0a} + \sum_{z=1}^d \omega_{za} S_{t-z,k} \right) + \mathcal{N}_t \quad (24)$$

where w_0 and w_{0a} are the biases on the nodes, w_a and w_{za} ($a = 1, 2, \dots, q, z = 1, 2, \dots, d$) are the connection weights between the layers of the model, $f(\cdot)$ is the activation function of the hidden layer, d is the number of input nodes, q is the number of hidden nodes,

and $\mathcal{N}_t \stackrel{iid}{\sim} N(0, \sigma^2)$. Furthermore, the sigmoid is based on the logistic function given as $f(x) = (1 + e^{-x})^{-1}$.

4. Application to Real Data

4.1. Study Area

For the empirical investigation of the proposed FAR model, an hourly temperature time series (in °C) for the site Karachi, Pakistan (Latitude: 24.8607° N, Longitude: 67.0011° E), is considered (Figure 4). The data used in this study span the period from 1 January 2019 to 31 December 2024 and were collected from the NASA POWER data access portal (<https://power.larc.nasa.gov/data-access-viewer/> (accessed on 14 January 2025)).

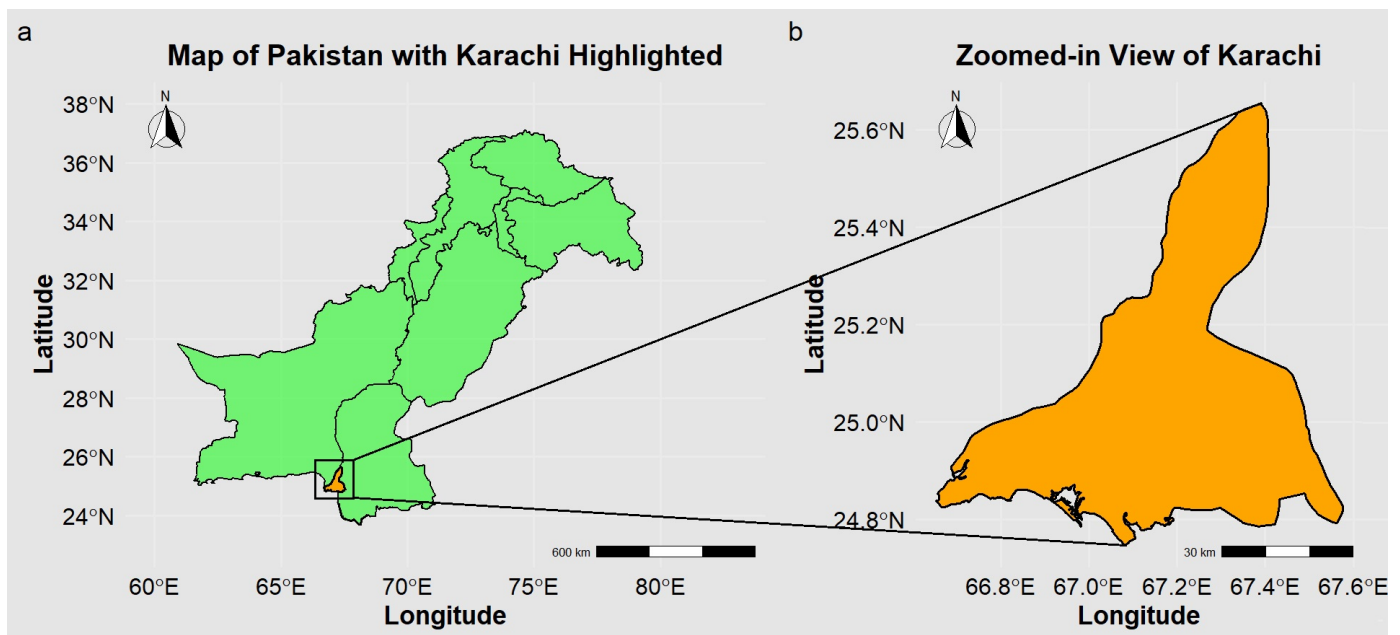


Figure 4. Spatial location of the study area, (a) map of Pakistan with Karachi highlighted; (b) zoomed-in view of Karachi.

In addition, exogenous variables such as wet-bulb temperature (°C), wind speed (10 m/s), and surface pressure (kPa) are also collected for the same location. The time series data for each variable consists of 52,632 hourly observations for 2192 days. The descriptive summaries for these variables are given in Table 2.

Table 2. Descriptive Summaries of Endogenous and Exogenous Variables.

Variables	Min.	Q1	Median	Q3	Max.	Mean	SD
Temperature (°C)	9.48	22.98	27.03	29.7	39.57	26.2	4.95
Wet-bulb temperature (°C)	3.36	18.44	24.43	26.98	31.43	22.35	5.73
Surface Pressure (kPa)	98.51	99.7	100.29	100.82	102	100.26	0.67
Wind Speed (10 m/s)	0.03	2.93	4.09	5.59	16.52	4.36	2.01

Figure 5 depicts the nature of variations for response and the three exogenous variables. The time series plot of the temperature data over six years clearly distinguishes the training set (43,848) used for modeling and estimation from the testing set (8784), which corresponds to out-of-sample forecasts. Since 2024 is a leap year, a total of 366 day-ahead forecasts are generated using the expanding window modeling technique. The hourly temperature data also exhibits seasonal variation. Figure 6a illustrates seasonal variations across different

seasons in 2024, showing that winter and autumn experience significant fluctuations in temperature as compared to the other three seasons.

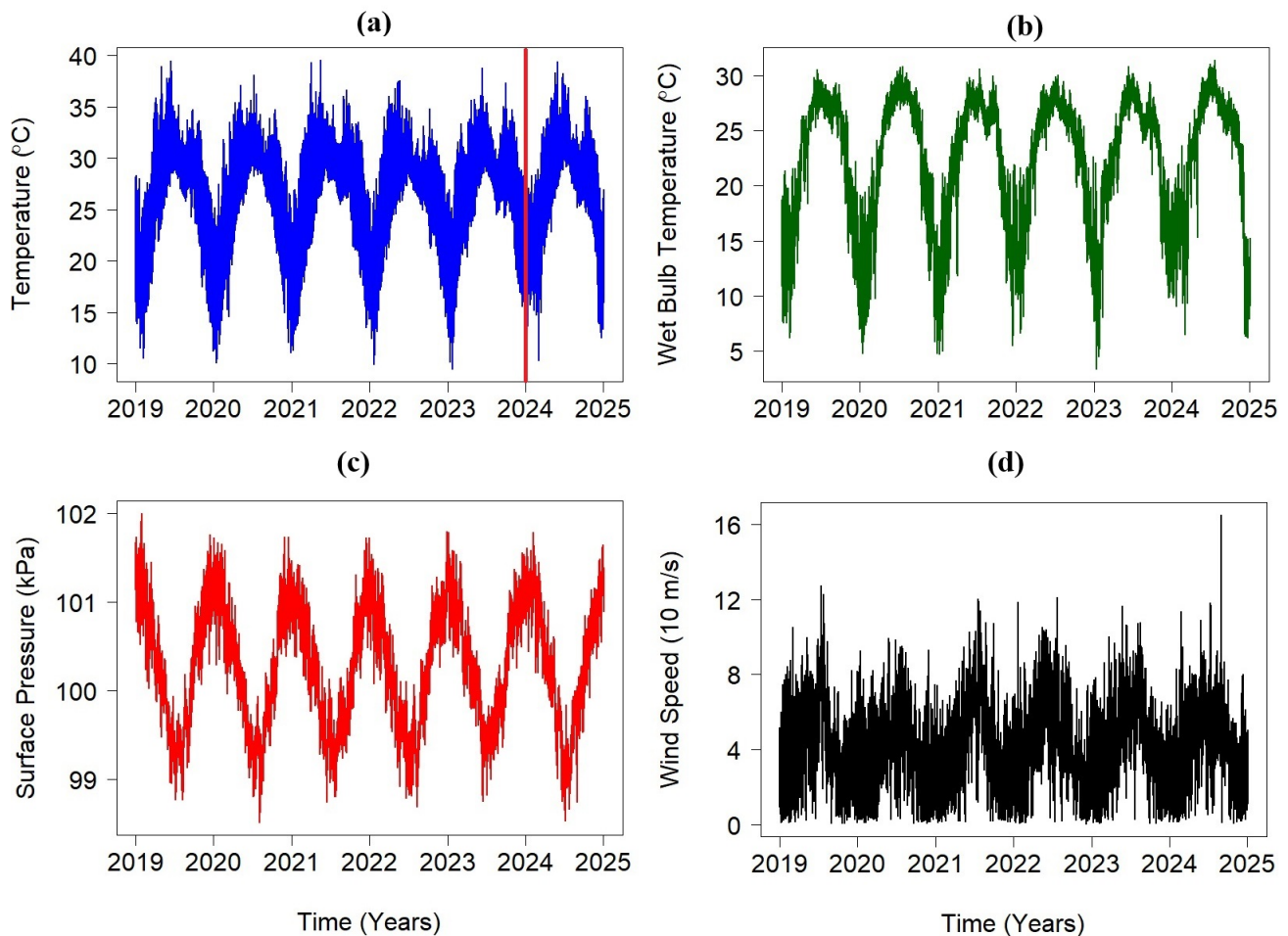


Figure 5. Time series plots of (a) endogenous variable indicating training and testing data partition; (b) first exogenous variable; (c) second exogenous variable; (d) third exogenous variable.

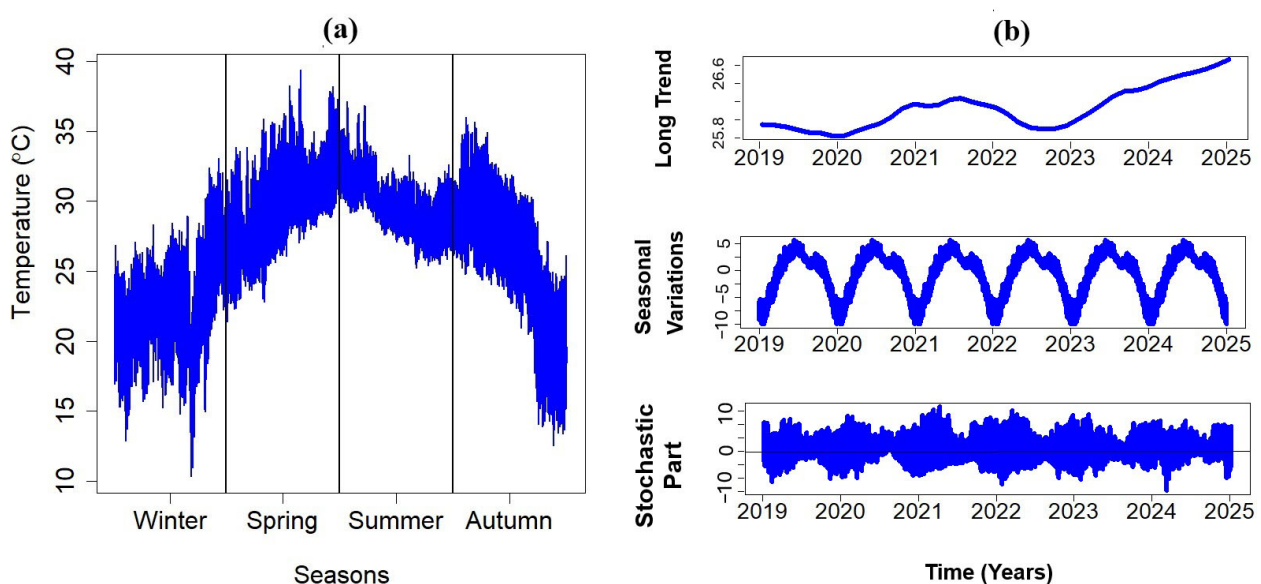


Figure 6. (a) Seasonal variations within an year; (b) long-trend, yearly, and seasonal variations for the whole dataset.

4.2. Data-Driven Modeling Procedure

Figure 7 clearly illustrates the step-by-step data-driven modeling procedure that is used to implement the proposed and competitive models, beginning with data preprocessing.

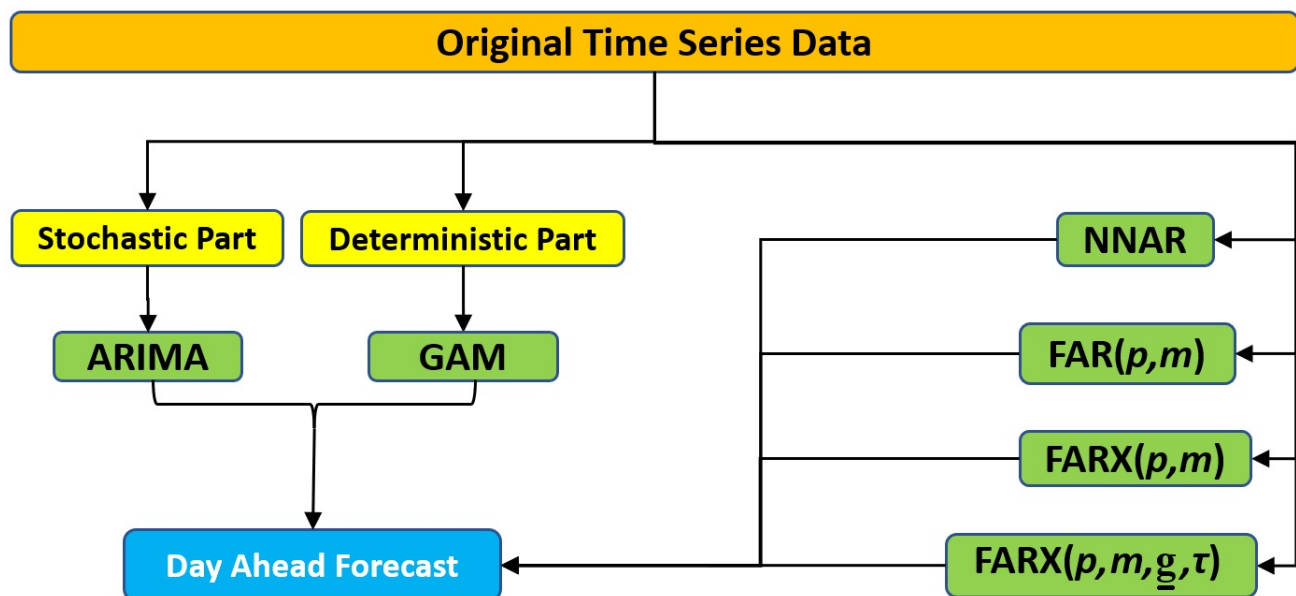


Figure 7. Flowchart of data-driven modeling procedure.

Extreme values in the times series of responses, as well as in the explanatory variables, pose challenges for estimation and forecasting. Identifying and adjusting these values improves prediction accuracy. This study applies the Shifting Filter on time series (SFT), an extension of [71], which operates on a rolling window instead of the entire series. Each of the original time series is divided into $P = (N/r)$ segments, where N is the total number of observations and r is the window width. Values deviating beyond 2.58 (based on a 99% confidence interval from a normal distribution) times the sample standard deviation σ from the mean $\hat{\lambda}$ are flagged as extreme [72]. The process iterates across all segments for all series as follows:

$$S_t^\circ = \bigcup_{i=1}^P S_t : |S_t - \hat{\lambda}| \geq 2.58 * SD_S,$$

where S_t is the temperature ($^{\circ}\text{C}$) at time t , $\hat{\lambda}$ is the sample mean, and SD_S is the standard deviation. The replacement rule can be mathematically defined as

$$\tilde{S}_t = \begin{cases} S_t, & \text{if } |S_t - \hat{\lambda}| < 2.58 \cdot SD_S \\ \tilde{\lambda}, & \text{otherwise} \end{cases}$$

Extreme values are replaced with the median ($\tilde{\lambda}$) of the respective window for improved forecasting stability.

Once $\tilde{S}_t$ are obtained for each TS, we move towards the data-driven framework as indicated in Figure 7, which includes the application of the traditional time series model ARIMA, a machine learning model NNAR, and three FAR models. Consider $S_{t,k}$ to be a filtered temperature time series where $t \in N$ and $k = 1, 2, 3, \dots, 24$. Due to significant seasonal variations ($U_{t,k}$), long trends ($T_{t,k}$), and year effects ($Y_{t,k}$), as represented in Figure 6b, we partitioned $S_{t,k}$ into deterministic ($D_{t,k}$) and stochastic ($S_{t,k}$) parts, such that $S_{t,k} = D_{t,k} + S_{t,k}$, the first term $D_{t,k} = U_{t,k} + T_{t,k} + Y_{t,k}$ is estimated via a generalized

additive model (GAM) and the second term $\mathcal{S}_{t,k} = S_{t,k} - \hat{D}_{t,k}$ is modeled by ARIMA. Finally, the results are combined ($\hat{D}_{t,k} + \hat{\mathcal{S}}_{t,k}$) for the day-ahead final forecast; that is, $\hat{S}_{t+1,k} = \hat{D}_{t+1,k} + \hat{\mathcal{S}}_{t+1,k}$. Additionally, models like NNAR and FAR can capture non-linearity directly without decomposing $S_{t,k}$, as they performed even better in handling complex patterns.

Order Identification and Estimation

For different variables used in the study, order identification is crucial for ensuring accurate time series modeling and forecasting. The optimal orders required for the implementation of the models in Figure 7 are estimated and given in Table 3.

Table 3. Optimal orders for various models.

ARIMA	NNAR	FAR(p, m)	FARX(p, m, τ)	FARX($p, m, \underline{g}, \tau$)
$p = 3$	$d = 32$	$p = 6$	$p = 6$	$p = 2$
$d = 0$	$q = 16$	$m = 14$	$m = 14$	$m = 15$
$q = 2$			$\tau = 3$	$g_1 = 15$
				$g_2 = 15$
				$g_3 = 9$
				$\tau = 6$

p : Autoregressive order, d : Integration order, q : Moving average order, d : Input neurons, q : Hidden layers, p : Lagged value, m : Dimensions of functional response variable, τ : number of scalar covariates, g_1, g_2 and g_3 : Dimensions of first, second, and third functional covariates.

The optimal orders for all models are selected based on the model selection criteria: for ARIMA, using AIC/BIC; for neural networks via cross-validation; and for FAR(p, m), FARX(p, m, τ) and FARX($p, m, \underline{g}, \tau$), the respective fFPE's in Equations (15), (17) and (22) are minimized to determine the right functional orders, which optimizes the dependency structure in FTS, enhancing predictive performance. Thus, based on Table 3, the estimated time series models are ARIMA(3, 0, 2), NNAR(32, 16), and the three FAR models FAR(6, 14), FARX(6, 14, 3), FARX(2, 15, $[15 \ 15 \ 9]'$, 6), respectively.

To evaluate the residual behavior of the selected models, the autocorrelation function (ACF) and partial ACF (PACF) plots are depicted in Figure 8. Overall, the final estimated residuals appear to be largely whitened, indicating that the fitted models can be regarded as adequate. Since the fitted models exhibit satisfactory diagnostic behavior, we now proceed to the out-of-sample forecasting analysis.

4.3. Out-of-Sample Forecasting

Out-of-sample forecasting evaluates a model's predictive performance on unobserved data beyond the estimation period. This approach helps assess the generalization and robustness of the model in real-world applications. To validate the accuracy of the forecasting models, different performance measures are used. In particular, the forecasting accuracy of a model is determined by the mean absolute error (MAE), root mean squared error (RMSE), mean absolute percentage error (MAPE) and mean absolute scaled error (MASE) for $\mathbb{T}$ number of out-of-sample forecasts used for accuracy evaluation, which are given as

$$\text{MAE} = \frac{1}{\mathbb{T}} \sum_{t=1}^{\mathbb{T}} |S_{t,k} - \hat{S}_{t,k}|$$

$$\text{RMSE} = \sqrt{\frac{1}{\mathbb{T}} \sum_{t=1}^{\mathbb{T}} (S_{t,k} - \hat{S}_{t,k})^2}$$

$$\text{MAPE} = \frac{1}{\mathbb{T}} \sum_{t=1}^{\mathbb{T}} \left| \frac{S_{t,k} - \hat{S}_{t,k}}{S_{t,k}} \right| \times 100$$

$$\text{MASE} = \frac{\sum_{t=1}^{\mathbb{T}} |S_{t,k} - \hat{S}_{t,k}|}{\frac{\mathbb{T}}{\mathbb{T}-1} \sum_{t=2}^{\mathbb{T}} |S_{t,k} - S_{t-1,k}|}$$

Using the proposed and competitor models, day-ahead out-of-sample forecasts are obtained for the complete leap year 2024. A univariate hourly time series dataset of temperature is used to get ARIMA(3, 0, 2) and NNAR(32, 16) while the daily temperature curves are used to estimate the three FAR models.

The results for various models are summarized in Table 4 by means of RMSE, MAE, MAPE and MASE. From the results we can see that the NNAR outperforms the classical ARIMA by producing relatively small RMSE (0.9720) and MAPE (2.7928). While the FAR and FARX models provide considerable improvement in forecasting in comparison with ARIMA and NNAR, the proposed FARX(2, 15, $[15 \ 15 \ 9]'$, 6) ensures superior performance within the class of functional models. It achieves the lowest RMSE (0.2942) and MAE (0.2004), indicating improved forecasting accuracy. Additionally, the significant reduction in MAPE (0.8213) highlights lower prediction errors, while the MASE (0.3188) is also the lowest among all models, showcasing its robustness. These findings confirm that FARX(2, 15, $[15 \ 15 \ 9]'$, 6) provides more accurate day-ahead temperature forecasts while minimizing overall errors, which is a major advancement in climate data analysis.

Table 4. One-day-ahead out-of-sample temperature forecasting performance for various models.

Models	RMSE	MAE	MAPE	MASE
ARIMA	1.1021	0.6924	2.8604	1.1013
NNAR	0.9720	0.6725	2.7928	1.0697
FAR(p, m)	0.7887	0.5370	2.1412	0.8542
FARX(p, m, τ)	0.7873	0.5360	2.1378	0.8526
FARX($p, m, \underline{\mathbf{g}}, \tau$)	0.2942	0.2004	0.8213	0.3188

To assess the significance of the differences between the accuracy metrics given in Table 4, we applied the Diebold and Mariano (DM) [73] test, which is a standard statistical method for comparing forecast accuracy across models. It evaluates whether differences in accuracy measures between forecast pairs are statistically significant and the null hypothesis states no difference in accuracy. Table 5 consists of findings of the DM test, and contains the p -values, using which we test the null hypothesis that the forecasting accuracies of the model in the row and the model in the column are the same against the alternative hypothesis that the forecasting accuracy of the model in the column is greater than the model in the row.

Table 5. p -values from the Diebold–Mariano test, where the null hypothesis assumes equal predictive accuracy between models, while the alternative suggests the column model outperforms the row model.

Models	ARIMA	NNAR	FAR(p, m)	FARX(p, m, τ)	FARX($p, m, \underline{\mathbf{g}}, \tau$)
ARIMA	–	<0.01	<0.01	<0.01	<0.01
NNAR	>0.99	–	<0.01	<0.01	<0.01
FAR(p, m)	>0.99	>0.99	–	<0.01	<0.01
FARX(p, m, τ)	>0.99	>0.99	>0.99	–	<0.01
FARX($p, m, \underline{\mathbf{g}}, \tau$)	>0.99	>0.99	>0.99	>0.99	–

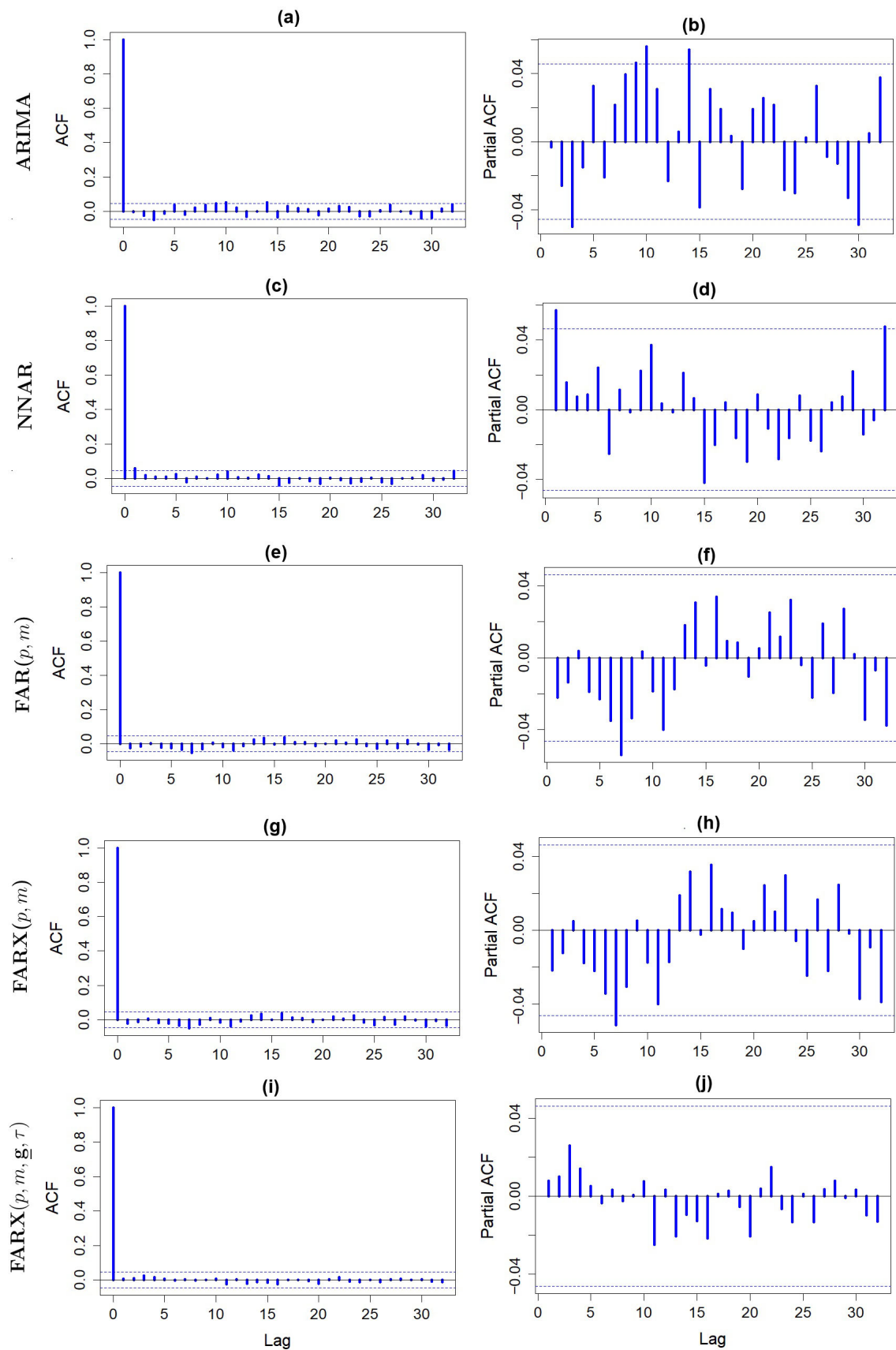


Figure 8. ACF and PACF plots for all models under study, (a) ACF for ARIMA; (b) PACF for ARIMA; (c) ACF for NNAR; (d) PACF for NNAR; (e) ACF for $FAR(p, m)$; (f) PACF for $FAR(p, m)$; (g) ACF for $FARX(p, m, \underline{g}, \tau)$; (h) PACF for $FARX(p, m, \underline{g}, \tau)$; (i) ACF for $FARX(p, m, \underline{g}, \tau)$; (j) PACF for $FARX(p, m, \underline{g}, \tau)$.

This table indicates that the functional models produce statistically more significant results than ARIMA and NNAR. Among the functional models, the proposed $\text{FARX}(p, m, \mathbf{g}, \tau)$ outperforms the other functional models, demonstrating its statistical significance for temperature forecasting.

To evaluate the performance of our proposed functional model, we compute the same accuracy measures for different seasons of the year as well as for each month, and the results are listed in Tables 6 and 7, respectively. The study site goes through four different seasons and the daily temperature curves are considerably different resulting in different forecasting accuracy errors. For example, looking at Table 6, 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 in summer, the value is 0.1550 for the $\text{FARX}(2, 15, [15 \ 15 \ 9]', 6)$ model, which is the lowest value as compared to other seasons and compared to the other models. Although the forecasting accuracies of all the models are high, the three FAR approaches have relatively better performance, especially the proposed $\text{FARX}(2, 15, [15 \ 15 \ 9]', 6)$ model. The error metrics of all models are relatively small in autumn, but large in winter. For instance, the $\text{FAR}(6, 14)$ and $\text{FARX}(6, 14, 3)$ produce MAPE values of 1.2031 and 1.1875 in autumn, but 3.4099 and 3.4241 in winter, respectively. On the other hand, the benchmark models $\text{ARIMA}(3, 0, 2)$ and $\text{NNAR}(32, 16)$ produces fairly large MAPE values. The results indicate that the proposed model performs better in summer and autumn compared to winter and spring. This can be attributed to the relatively smoother and more stable temperature patterns during these seasons, whereas winter and spring exhibit rapid fluctuations due to complex atmospheric interactions.

Table 6. Accuracy measures for one-day-ahead out-of-sample temperature forecasts for different seasons based on various models.

Seasons	Errors	ARIMA	NNAR	$\text{FAR}(p, m)$	$\text{FARX}(p, m, \tau)$	$\text{FARX}(p, m, \mathbf{g}, \tau)$
Winter	RMSE	1.2935	1.2562	0.9180	0.9204	0.3584
	MAE	0.9444	0.9645	0.6977	0.7008	0.2553
	MAPE	4.6811	4.8695	3.4099	3.4241	1.2796
	MASE	0.9805	1.0014	0.7244	0.7276	0.2651
Spring	RMSE	1.4741	1.1767	0.9757	0.9791	0.3604
	MAE	0.8976	0.8576	0.6904	0.6904	0.2509
	MAPE	3.6264	3.3937	2.6307	2.6418	0.9884
	MASE	0.9770	0.9334	0.7514	0.7554	0.2730
Summer	RMSE	0.7738	0.7140	0.6650	0.6640	0.2237
	MAE	0.4985	0.4667	0.4146	0.4159	0.1550
	MAPE	1.6006	1.4836	1.3111	1.3150	0.5073
	MASE	0.9778	0.9154	0.8133	0.8157	0.3041
Autumn	RMSE	0.6402	0.5542	0.4905	0.4898	0.1941
	MAE	0.4289	0.4015	0.3409	0.3370	0.1404
	MAPE	1.5389	1.4323	1.2031	1.1875	0.5114
	MASE	1.0445	0.9778	0.8302	0.8206	0.3419

Table 7 lists the monthly forecasting accuracy for the models used in this study. The results suggest that the forecasting errors are lower in June, September and November compared to the other months of the year. The lowest errors are observed in the month of June, whereas the higher errors correspond to the month of March. Comparing the forecasting accuracies of different models, it is evident that the functional models perform relatively better than ARIMA and NNAR. Among the functional models, the $\text{FARX}(2, 15, [15 \ 15 \ 9]', 6)$ gives better results than $\text{FAR}(6, 14)$ and $\text{FARX}(6, 14, 3)$. The lowest MAPE produced by our proposed functional model is 0.4081 for June, whereas the value of 1.4048, produced for February, is the highest MAPE value.

Table 7. One-day-ahead out-of-sample temperature forecasts: month-specific forecasting errors.

Months	Errors	ARIMA	NNAR	$FAR(p, m)$	$FARX(p, m, \tau)$	$FARX(p, m, \underline{g}, \tau)$	Months	Errors	ARIMA	NNAR	$FAR(p, m)$	$FARX(p, m, \tau)$	$FARX(p, m, \underline{g}, \tau)$
January	RMSE	1.0441	1.1219	0.8484	0.8507	0.2912	July	RMSE	0.8731	0.8616	0.8178	0.8156	0.2552
	MAE	0.7858	0.8907	0.6552	0.6560	0.2120		MAE	0.5706	0.5588	0.5031	0.5042	0.1814
	MAPE	3.9065	4.4024	3.1678	3.1675	1.0754		MAPE	1.7669	1.7291	1.5459	1.5497	0.5746
	MASE	0.9319	1.0563	0.7770	0.7780	0.2514		MASE	0.8855	0.8671	0.7807	0.7823	0.2815
February	RMSE	1.2813	1.2888	0.9585	0.9609	0.4120	August	RMSE	0.6580	0.6037	0.5751	0.5747	0.2318
	MAE	0.9975	0.9998	0.7142	0.7171	0.2960		MAE	0.4483	0.4011	0.3692	0.3699	0.1558
	MAPE	4.7546	4.7935	3.3520	3.3663	1.4048		MAPE	1.5386	1.3706	1.2467	1.2490	0.5359
	MASE	0.9573	0.9595	0.6855	0.6882	0.2841		MASE	0.9449	0.8453	0.7781	0.7797	0.3284
March	RMSE	2.1264	1.5213	1.1283	1.1315	0.4168	September	RMSE	0.6907	0.4020	0.3780	0.3583	0.1746
	MAE	1.1900	1.1580	0.8042	0.8086	0.2863		MAE	0.3753	0.3094	0.2684	0.2551	0.1270
	MAPE	5.7452	5.3436	3.5958	3.6115	1.3095		MAPE	1.3450	1.0949	0.9423	0.8945	0.4570
	MASE	1.0631	1.0345	0.7184	0.7224	0.2558		MASE	1.2614	1.0398	0.9021	0.8573	0.4269
April	RMSE	0.9866	0.9736	0.9405	0.9448	0.3628	October	RMSE	0.6771	0.7008	0.6261	0.6314	0.2133
	MAE	0.7717	0.7549	0.7040	0.7061	0.2664		MAE	0.4800	0.5043	0.4329	0.4337	0.1550
	MAPE	2.7698	2.7160	2.4912	2.4962	0.9867		MAPE	1.6210	1.7116	1.4546	1.4556	0.5416
	MASE	0.8584	0.8397	0.7831	0.7854	0.2963		MASE	0.9315	0.9786	0.8401	0.8417	0.3007
May	RMSE	0.9928	0.9370	0.8346	0.8373	0.2904	November	RMSE	0.5408	0.5123	0.4266	0.4330	0.1918
	MAE	0.7271	0.6565	0.5634	0.5678	0.2004		MAE	0.4297	0.3875	0.3184	0.3189	0.1388
	MAPE	2.3367	2.0995	1.8007	1.8128	0.6689		MAPE	1.6481	1.4811	1.2039	1.2034	0.5346
	MASE	1.0272	0.9274	0.7960	0.8022	0.2831		MASE	0.9343	0.8425	0.6923	0.6932	0.3017
June	RMSE	0.7753	0.6476	0.5685	0.5689	0.1752	December	RMSE	1.5119	1.3489	0.9460	0.9484	0.3653
	MAE	0.4757	0.4392	0.3702	0.3721	0.1269		MAE	1.0534	1.0052	0.7248	0.7303	0.2605
	MAPE	1.4929	1.3466	1.1349	1.1408	0.4081		MAPE	5.3870	5.4079	3.7061	3.7348	1.3666
	MASE	1.0418	0.9619	0.8107	0.8150	0.2779		MASE	0.9921	0.9467	0.6826	0.6878	0.2454

Finally, Table 8 provides a comparison of the models for each hour of the day. The results are attributed to the one-day-ahead out-of-sample forecasts for a whole year. That is, the accuracy measures are obtained for the first hour based on all first hours for a complete cycle in 2024 (366 days), and so forth. This table indicates that daily temperature curves are more stable in the second-to-sixth and eighteenth-to-twenty-first hours of the day than they are in the first, seventh, thirteenth, and the twenty-fourth hour. The forecasting errors for the thirteenth and fourteenth hours are comparatively lower than in the first and last hour of the day. Additionally, higher prediction errors around dawn are observed, likely caused by abrupt temperature transitions and lower signal-to-noise ratios at these times, which increase the difficulty of accurate short-term forecasting. The results suggest that our proposed functional model $FARX(2, 15, [15 \ 15 \ 9]', 6)$ is more efficient at capturing the non-linear trends and performs significantly better compared to the traditional ARIMA and NNAR within the class of functional models. Hence, the inclusion of a number of exogenous variables enhances the capability of $FAR(6, 14)$ and $FARX(6, 14, 3)$.

Table 8. Hour-specific one-day-ahead out-of-sample forecasting errors for temperature.

Hours	Errors	ARIMA	NNAR	$FAR(p, m)$	$FARX(p, m, \tau)$	$FARX(p, m, \underline{g}, \tau)$	Hours	Errors	ARIMA	NNAR	$FAR(p, m)$	$FARX(p, m, \tau)$	$FARX(p, m, \underline{g}, \tau)$
1	RMSE	0.9675	0.8736	0.2658	0.2674	0.4146	13	RMSE	1.2435	1.0984	1.0217	1.0236	0.1243
	MAE	0.5696	0.5823	0.1888	0.1880	0.2831		MAE	0.8210	0.7966	0.7409	0.7421	0.0799
	MAPE	2.7269	2.8245	0.8946	0.8962	1.3748		MAPE	2.7335	2.6322	2.4490	2.4526	0.2726
	MASE	1.0825	1.1065	0.3588	0.3572	0.5380		MASE	1.0176	0.9874	0.9183	0.9198	0.0990
2	RMSE	0.9744	0.9191	0.2568	0.2579	0.1878	14	RMSE	1.2470	1.1036	1.0463	1.0478	0.1691
	MAE	0.5856	0.6094	0.1754	0.1758	0.1331		MAE	0.8251	0.8044	0.7542	0.7552	0.0968
	MAPE	2.8407	3.0073	0.8257	0.8282	0.6338		MAPE	2.7504	2.6557	2.4991	2.5029	0.3244
	MASE	1.0766	1.1204	0.3225	0.3232	0.2447		MASE	1.0297	1.0038	0.9412	0.9424	0.1208
3	RMSE	0.9922	0.9248	0.3398	0.3397	0.1589	15	RMSE	1.2243	1.0939	1.0443	1.0465	0.2320
	MAE	0.6025	0.6073	0.2371	0.2371	0.1164		MAE	0.8207	0.8094	0.7600	0.7603	0.1475
	MAPE	2.9540	3.0337	1.1291	1.1297	0.5349		MAPE	2.7664	2.7094	2.5471	2.5487	0.4940
	MASE	1.0776	1.0863	0.4240	0.4242	0.2082		MASE	1.0366	1.0223	0.9598	0.9603	0.1863
4	RMSE	1.0209	0.9145	0.4320	0.4329	0.1756	16	RMSE	1.1947	1.0533	1.0437	1.0482	0.3303
	MAE	0.6276	0.6153	0.3169	0.3185	0.1300		MAE	0.8025	0.7845	0.7741	0.7802	0.2495
	MAPE	3.1106	3.0861	1.5174	1.5240	0.5956		MAPE	2.7748	2.6959	2.6526	2.6740	0.8707
	MASE	1.0829	1.0618	0.5468	0.5496	0.2244		MASE	1.0266	1.0035	0.9902	0.9981	0.3192

Table 8. Cont.

Hours	Errors	ARIMA	NNAR	FAR(p, m)	FARX(p, m, τ)	FARX($p, m, \underline{g}, \tau$)	Hours	Errors	ARIMA	NNAR	FAR(p, m)	FARX(p, m, τ)	FARX($p, m, \underline{g}, \tau$)
5	RMSE	1.0489	0.9513	0.5264	0.5280	0.1861	17	RMSE	1.1610	1.0083	0.9853	0.9893	0.3261
	MAE	0.6481	0.6405	0.3779	0.3808	0.1219		MAE	0.7688	0.7536	0.7234	0.7273	0.2512
	MAPE	3.2162	3.2141	1.8461	1.8565	0.5786		MAPE	2.8123	2.7290	2.6079	2.6220	0.9071
	MASE	1.0642	1.0517	0.6206	0.6253	0.2002		MASE	1.0285	1.0082	0.9678	0.9730	0.3360
6	RMSE	1.0852	0.9843	0.6217	0.6200	0.2210	18	RMSE	1.1184	0.9714	0.9231	0.9235	0.3684
	MAE	0.6764	0.6661	0.4279	0.4273	0.1560		MAE	0.7344	0.6892	0.6854	0.6830	0.2882
	MAPE	3.3638	3.3555	2.1425	2.1370	0.7077		MAPE	2.9013	2.7089	2.6769	2.6703	1.1229
	MASE	1.0962	1.0794	0.6935	0.6925	0.2528		MASE	1.0239	0.9608	0.9555	0.9521	0.4018
7	RMSE	1.1272	1.0036	0.6943	0.6912	0.3283	19	RMSE	1.0298	0.8876	0.8178	0.8187	0.3043
	MAE	0.6861	0.6715	0.4890	0.4875	0.2428		MAE	0.6512	0.6000	0.5840	0.5814	0.2206
	MAPE	3.3010	3.2858	2.3783	2.3639	1.1503		MAPE	2.6950	2.4874	2.3813	2.3740	0.8757
	MASE	1.1042	1.0806	0.7869	0.7845	0.3907		MASE	1.0187	0.9386	0.9135	0.9095	0.3451
8	RMSE	1.1043	0.9855	0.6914	0.6903	0.3360	20	RMSE	0.9974	0.8428	0.7720	0.7759	0.2804
	MAE	0.6846	0.6732	0.5061	0.5063	0.2499		MAE	0.6129	0.5645	0.5307	0.5320	0.1947
	MAPE	3.0044	2.9736	2.1990	2.1960	1.0419		MAPE	2.6254	2.4273	2.2383	2.2460	0.7976
	MASE	1.0515	1.0340	0.7773	0.7776	0.3838		MASE	1.0188	0.9383	0.8822	0.8844	0.3236
9	RMSE	1.1115	0.9322	0.7256	0.7238	0.3447	21	RMSE	0.9972	0.8712	0.7481	0.7525	0.2687
	MAE	0.7109	0.6463	0.5318	0.5309	0.2600		MAE	0.6028	0.5704	0.5149	0.5166	0.1972
	MAPE	2.8150	2.5586	2.0833	2.0763	0.9898		MAPE	2.6531	2.5108	2.2339	2.2426	0.8209
	MASE	1.0396	0.9451	0.7777	0.7765	0.3802		MASE	1.0454	0.9892	0.8928	0.8959	0.3420
10	RMSE	1.1806	1.0114	0.8336	0.8343	0.3465	22	RMSE	1.0114	0.8914	0.7448	0.7491	0.2754
	MAE	0.7737	0.7318	0.6075	0.6097	0.2575		MAE	0.5939	0.5802	0.5112	0.5136	0.2034
	MAPE	2.8290	2.6613	2.1932	2.1996	0.9148		MAPE	2.6777	2.6370	2.2792	2.2900	0.8583
	MASE	1.0249	0.9694	0.8048	0.8077	0.3411		MASE	1.0446	1.0205	0.8993	0.9035	0.3577
11	RMSE	1.2384	1.0871	0.9507	0.9506	0.3125	23	RMSE	1.0097	0.8585	0.7501	0.7538	0.3214
	MAE	0.8267	0.7905	0.6877	0.6876	0.2155		MAE	0.5813	0.5708	0.5049	0.5086	0.2268
	MAPE	2.8801	2.7358	2.3755	2.3735	0.7479		MAPE	2.6842	2.6486	2.3189	2.3334	1.0017
	MASE	1.0112	0.9669	0.8412	0.8410	0.2636		MASE	1.0545	1.0355	0.9160	0.9227	0.4114
12	RMSE	1.2481	1.0956	0.9987	0.9993	0.2212	24	RMSE	1.0165	0.8786	0.7799	0.7820	0.5023
	MAE	0.8379	0.7981	0.7240	0.7250	0.1491		MAE	0.5731	0.5845	0.5103	0.5134	0.3389
	MAPE	2.8288	2.6782	2.4242	2.4265	0.5102		MAPE	2.7055	2.7713	2.4133	2.4251	1.5841
	MASE	1.0167	0.9684	0.8785	0.8797	0.1809		MASE	1.0745	1.0958	0.9568	0.9626	0.6354

To summarize the work, one can see from Tables 4–8 that the proposed approach is efficient, producing significantly lower error metrics than the other competitors, which indicates better predictive power for modeling and forecasting the daily temperature curves and hence demonstrates the usefulness of our estimation framework. Moreover, the forecasting results show that the functional forecasting approach is superior to the classical ARIMA and NNAR methods.

4.4. Computational Complexity

Finally, we compare the average computational time in seconds for the five models as given in Table 9. For the functional models, authors wrote their own code and the necessary programming is carried out in R which is a statistical computing language and all the computations have been performed using an Intel(R)-Core(TM) i7-4770 CPU running at 3.40 GHz. For deterministic part, the library gam is used while for ARIMA and NNAR, the library forecast is used. For the applications of functional models the libraries fda and vars. The documentation provided by the packages includes in-depth details on the particular algorithms utilized in the estimation process.

Table 9. Average time in seconds for one-day-ahead forecast of temperature data.

Average Time	ARIMA	NNAR	FAR(p, m)	FARX(p, m, τ)	FARX($p, m, \underline{g}, \tau$)
Time (s)	0.33	1.05	1.14	1.17	2.17
Relative Time	1	3.18	3.45	3.55	6.58

Table 9 presents the average computation time in seconds for one-day-ahead temperature forecasting across various models. ARIMA exhibits the lowest execution time (0.33 s), serving as the baseline. The functional models, particularly FARX($p, m, \underline{g}, \tau$), show increased computational cost, with a maximum time of 2.17 s. This rise reflects the added

complexity due to functional and exogenous components. However, such models provide improved predictive performance, justifying the additional time.

5. Conclusions

Accurate modeling and forecasting of daily temperature curves are essential for many practical applications, such as weather forecasting, agriculture, energy management, coastal activities, engineering, and climate analysis. However, due to the complex and non-linear patterns in weather data, the forecasting problem is a challenging task. To address this issue, a novel functional autoregressive model, namely $\text{FARX}(p, m, \mathbf{g}, \tau)$, is proposed to accommodate any number of functional and scalar covariates. The optimum orders used for different parts of the proposed model are estimated by minimizing a modified functional final prediction error called $\text{ffPEX}(p, m, \mathbf{g}, \tau)$. To evaluate the performance of the proposed methodology, the daily temperature curve data obtained from NASA POWER for Karachi in Pakistan is used, and one-day-ahead out-of-sample forecasts are obtained for a complete year. For comparison purposes, forecasts are also obtained using the classical ARIMA and NNAR models, as well as two functional models, namely $\text{FAR}(p, m)$ and $\text{FARX}(p, m)$. The forecasting performance of the different models is assessed through different accuracy measures, including the RMSE, MAE, MAPE, and MASE.

The study findings suggest that the proposed functional model $\text{FARX}(p, m, \mathbf{g}, \tau)$ is efficient at forecasting daily temperature curves by producing considerably lower forecasting errors compared to the competitors. Both of the functional models, $\text{FAR}(p, m)$ and $\text{FARX}(p, m)$, perform relatively better than ARIMA and NNAR. Among the functional models, our proposed model, taking into consideration several functional and scalar exogenous variables, produces the best forecasting results. The proposed model effectively captures functional temperature dynamics, particularly for short-term horizons; however, it may be less accurate under extreme or abrupt climatic events. Future work could incorporate additional high-frequency exogenous variables and explore hybrid approaches with deep learning to further enhance predictive performance. Additionally, the model should be evaluated across regions with diverse climatic patterns, such as continental or tropical climates, to assess its robustness and generalizability beyond Karachi.

Author Contributions: Conceptualization, I.S. and M.U.; methodology, I.S. and M.U.; software, M.U. and S.A.; validation, S.A. and S.M.A.; formal analysis, I.S. and M.U.; investigation, S.A. and S.M.A.; resources, S.A. and S.M.A.; data curation, M.U. and S.M.A.; writing—original draft preparation, I.S.; writing—review and editing, S.A. and S.M.A.; visualization, M.U. and S.M.A.; supervision, I.S.; project administration, I.S. and S.A.; funding acquisition, S.M.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by Umm Al-Qura University, Saudi Arabia, under grant number: 26UQU4310037GSSR03.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study were collected from the NASA POWER data access portal (<https://power.larc.nasa.gov/data-access-viewer/>), accessed on 14 January 2025.

Acknowledgments: The authors extend their appreciation to Umm Al-Qura University, Saudi Arabia, for funding this research work through grant number: 26UQU4310037GSSR03.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tajfar, E.; Bateni, S.M.; Margulis, S.A.; Gentine, P.; Auligne, T. Estimation of turbulent heat fluxes via assimilation of air temperature and specific humidity into an atmospheric boundary layer model. *J. Hydrometeorol.* **2020**, *21*, 205–225. <https://doi.org/10.1175/JHM-D-19-0104.1>.
2. Valipour, M.; Bateni, S.M.; Gholami Sefidkouhi, M.A.; Raeini-Sarjaz, M.; Singh, V.P. Complexity of forces driving trend of reference evapotranspiration and signals of climate change. *Atmosphere* **2020**, *11*, 1081. <https://doi.org/10.3390/atmos11101081>.
3. Cifuentes, J.; Marulanda, G.; Bello, A.; Reneses, J. Air temperature forecasting using machine learning techniques: A review. *Energies* **2020**, *13*, 4215. <https://doi.org/10.3390/en13164215>.
4. Massie, D.R.; Rose, M.A. Predicting daily maximum temperatures using linear regression and Eta geopotential thickness forecasts. *Weather. Forecast.* **1997**, *12*, 799–807. [https://doi.org/10.1175/1520-0434\(1997\)012<0799:PDMTUL>2.0.CO;2](https://doi.org/10.1175/1520-0434(1997)012<0799:PDMTUL>2.0.CO;2).
5. Campbell, S.D.; Diebold, F.X. Weather forecasting for weather derivatives. *J. Am. Stat. Assoc.* **2005**, *100*, 6–16. <https://doi.org/10.1198/016214504000001051>.
6. Svec, J.; Stevenson, M. Modelling and forecasting temperature based weather derivatives. *Glob. Financ. J.* **2007**, *18*, 185–204. <https://doi.org/10.1016/j.gfj.2006.04.006>.
7. Dupuis, D.J. Forecasting temperature to price CME temperature derivatives. *Int. J. Forecast.* **2011**, *27*, 602–618. <https://doi.org/10.1016/j.ijforecast.2010.03.004>.
8. Wang, Z.; Li, P.; Li, L.; Huang, C.; Liu, M. Modeling and forecasting average temperature for weather derivative pricing. *Adv. Meteorol.* **2015**, *2015*, 837293. <https://doi.org/10.1155/2015/837293>.
9. Cavicchioli, M.; Ghezel, A.; Zemmouri, I. (Bi)spectral analysis of Markov switching bilinear time series. *Stat. Methods Appl.* **2025**, 1–30. <https://doi.org/10.1007/s10260-025-00826-9>.
10. Alraddadi, R. The logTG-SV model: A threshold-based volatility framework with logarithmic shocks for exchange rate dynamics. *AIMS Math.* **2025**, *10*, 19495–19511. <https://doi.org/10.3934/math.2025870>.
11. Gomez, I.; Estrela, M.J.; Caselles, V. Operational forecasting of daily summer maximum and minimum temperatures in the Valencia Region. *Nat. Hazards* **2014**, *70*, 1055–1076. <https://doi.org/10.1007/s11069-013-0861-1>.
12. Trigo, R.M.; Palutikof, J.P. Simulation of daily temperatures for climate change scenarios over Portugal: A neural network model approach. *Clim. Res.* **1999**, *13*, 45–59. <https://doi.org/10.3354/cr013045>.
13. Abdel-Aal, R.E.; Elhadidy, M.A. Modeling and forecasting the daily maximum temperature using abductive machine learning. *Weather. Forecast.* **1995**, *10*, 310–325. [https://doi.org/10.1175/1520-0434\(1995\)010<0310:MAFTDM>2.0.CO;2](https://doi.org/10.1175/1520-0434(1995)010<0310:MAFTDM>2.0.CO;2).
14. Abdel-Aal, R.E. Hourly temperature forecasting using abductive networks. *Eng. Appl. Artif. Intell.* **2004**, *17*, 543–556. <https://doi.org/10.1016/j.engappai.2004.04.002>.
15. Ustaoglu, B.; Cigizoglu, H.K.; Karaca, M. Forecast of daily mean, maximum and minimum temperature time series by three artificial neural network methods. *Meteorol. Appl.* **2008**, *15*, 431–445. <https://doi.org/10.1002/met.83>.
16. Sanikhani, H.; Deo, R.C.; Samui, P.; Kisi, O.; Mert, C.; Mirabbasi, R.; Gavili, S.; Yaseen, Z.M. Survey of different data-intelligent modeling strategies for forecasting air temperature using geographic information as model predictors. *Comput. Electron. Agric.* **2018**, *152*, 242–260. <https://doi.org/10.1016/j.compag.2018.07.008>.
17. Tran, T.T.K.; Bateni, S.M.; Ki, S.J.; Vosoughifar, H. A review of neural networks for air temperature forecasting. *Water* **2021**, *13*, 1294. <https://doi.org/10.3390/w13091294>.
18. Shah, I.; Mubassir, P.; Ali, S.; Albalawi, O. A functional autoregressive approach for modeling and forecasting short-term air temperature. *Front. Environ. Sci.* **2024**, *12*, 1411237. <https://doi.org/10.3389/fenvs.2024.1411237>.
19. Selmy, H.A.; Mohamed, H.K.; Medhat, W. A predictive analytics framework for sensor data using time series and deep learning techniques. *Neural Comput. Appl.* **2024**, *36*, 6119–6132. <https://doi.org/10.1007/s00521-023-09398-9>.
20. Alexander, V.O.; Ali, M.I. Enhancing Time Series Data Predictions: A Survey of Augmentation Techniques and Model Performance. In *Proceedings of the 2024 Australasian Computer Science Week (ACSW 2024)*; Association for Computing Machinery: New York, NY, USA, 2024. <https://doi.org/10.1145/3641142.3641143>.
21. Uluocak, I.; Bilgili, M. Daily air temperature forecasting using LSTM-CNN and GRU-CNN models. *Acta Geophys.* **2024**, *72*, 2107–2126. <https://doi.org/10.1007/s11600-023-01241-y>.
22. An, H.; Li, Q.; Lv, X.; Li, G.; Qian, Q.; Zhou, G.; Nie, G.; Zhang, L.; Zhu, L. Forecasting daily extreme temperatures in Chinese representative cities using artificial intelligence models. *Weather. Clim. Extrem.* **2023**, *42*, 100621. <https://doi.org/10.1016/j.wace.2023.100621>.
23. Sen, A.; Mazumder, A.R.; Dutta, D.; Sen, U.; Syam, P.; Dhar, S. Comparative evaluation of metaheuristic algorithms for hyperparameter selection in short-term weather forecasting. *arXiv* **2023**, arXiv:2309.02600. <https://doi.org/10.48550/arXiv.2309.02600>.

24. Toharudin, T.; Pontoh, R.S.; Caraka, R.E.; Zahroh, S.; Lee, Y.; Chen, R.C. Employing long short-term memory and Facebook prophet model in air temperature forecasting. *Commun. Stat.-Simul. Comput.* **2023**, *52*, 279–290. <https://doi.org/10.1080/03610918.2020.1854302>.
25. Elshewey, A.M.; Shams, M.Y.; Elhady, A.M.; Shohieb, S.M.; Abdelhamid, A.A.; Ibrahim, A.; Tarek, Z. A Novel WD-SARIMAX model for temperature forecasting using daily delhi climate dataset. *Sustainability* **2022**, *15*, 757. <https://doi.org/10.3390/su15010757>.
26. Gong, B.; Langguth, M.; Ji, Y.; Mozaffari, A.; Stadtler, S.; Mache, K.; Schultz, M.G. Temperature forecasting by deep learning methods. *Geosci. Model Dev.* **2022**, *15*, 8931–8956. <https://doi.org/10.5194/gmd-15-8931-2022>.
27. Shin, K.-H.; Jung, J.-W.; Chang, K.-H.; Kim, K.; Jung, W.-S.; Lee, D.-I.; You, C.-H. Dynamical prediction of two meteorological factors using the deep neural network and the long short-term memory (II). *J. Korean Phys. Soc.* **2022**, *80*, 1081–1097. <https://doi.org/10.1007/s40042-022-00472-4>.
28. Alomar, M.K.; Khaleel, F.; Aljumaily, M.M.; Masood, A.; Razali, S.F.M.; AlSaadi, M.A.; Al-Ansari, N.; Hameed, M.M. Data-driven models for atmospheric air temperature forecasting at a continental climate region. *PLoS ONE* **2022**, *17*, e0277079. <https://doi.org/10.1371/journal.pone.0277079>.
29. Haque, E.; Tabassum, S.; Hossain, E. A comparative analysis of deep neural networks for hourly temperature forecasting. *IEEE Access* **2021**, *9*, 160646–160660. <https://doi.org/10.1109/ACCESS.2021.3131533>.
30. Wang, H.; Pathan, M.S.; Lee, Y.H.; Dev, S. Day-ahead forecasts of air temperature. In *2021 IEEE USNC-URSI Radio Science Meeting (Joint with AP-S Symposium)*; IEEE: New York, NY, USA, 2021; pp. 94–95. <https://doi.org/10.23919/USNC-URSI51813.2021.9703606>.
31. Lee, S.; Lee, Y.-S.; Son, Y. Forecasting daily temperatures with different time interval data using deep neural networks. *Appl. Sci.* **2020**, *10*, 1609. <https://doi.org/10.3390/app10051609>.
32. Zhang, Z.; Dong, Y. Temperature Forecasting via Convolutional Recurrent Neural Networks Based on Time-Series Data. *Complexity* **2020**, *2020*, 3536572. <https://doi.org/10.1155/2020/3536572>.
33. Shin, J.-Y.; Kim, K.R.; Ha, J.-C. Seasonal forecasting of daily mean air temperatures using a coupled global climate model and machine learning algorithm for field-scale agricultural management. *Agric. For. Meteorol.* **2020**, *281*, 107858. <https://doi.org/10.1016/j.agrformet.2019.107858>.
34. Wanishsakpong, W.; Owusu, B.E. Optimal time series model for forecasting monthly temperature in the southwestern region of Thailand. *Model. Earth Syst. Environ.* **2020**, *6*, 525–532. <https://doi.org/10.1007/s40808-019-00698-5>.
35. Zahroh, S.; Hidayat, Y.; Pontoh, R.S.; Santoso, A.; Sukono, F.; Bon, A.T. Modeling and forecasting daily temperature in Bandung. In *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Riyadh, Saudi Arabia, 26–28 November 2019; pp. 406–412.
36. Ramsay, J.O.; Silverman, B.W. *Applied Functional Data Analysis: Methods and Case Studies*; Springer: Berlin/Heidelberg, Germany, 2002. https://doi.org/10.1007/978-0-387-22465-7_2.
37. Uzair, M.; Shah, I.; Ali, S. An adaptive strategy for wind speed forecasting under functional data horizon: A way towards enhancing clean energy. *IEEE Access* **2024**, *12*, 68730–68746. <https://doi.org/10.1109/ACCESS.2024.3401038>.
38. Naeem, N.; Ali, S.; Shah, I. Functional EWMA control chart for phase II profile monitoring. *J. Stat. Comput. Simul.* **2025**, *95*, 96–116. <https://doi.org/10.1080/00949655.2024.2419864>.
39. Ullah, S.; Finch, C.F. Applications of functional data analysis: A systematic review. *BMC Med. Res. Methodol.* **2013**, *13*, 43. <https://doi.org/10.1186/1471-2288-13-43>.
40. Tao, Z.; Wang, M.; Liu, J.; Wang, P. A Functional Data Analysis Framework Incorporating Derivative Information and Mixed-Frequency Data for Predictive Modeling of Crude Oil Price. *IEEE Trans. Ind. Inform.* **2025**, *21*, 3226–3235. <https://doi.org/10.1109/TII.2024.3523571>.
41. Rodríguez Cuadro, D.; Pérez-Plaza, S.; Castaño-Martínez, A.; Fernández-Palacín, F. A Study of the Colombian Stock Market with Multivariate Functional Data Analysis (FDA). *Mathematics* **2025**, *13*, 858. <https://doi.org/10.3390/math13050858>.
42. Pokhrel, K.P.; Tsokos, C.P. Forecasting age-specific brain cancer mortality rates using functional data analysis models. *Adv. Epidemiol.* **2015**, *2015*, 721592. <https://doi.org/10.1155/2015/721592>.
43. Beyaztas, U.; Yaseen, Z.M. Drought interval simulation using functional data analysis. *J. Hydrol.* **2019**, *579*, 124141. <https://doi.org/10.1016/j.jhydrol.2019.124141>.
44. Hael, M.A.; Yongsheng, Y.; Saleh, B.I. Visualization of rainfall data using functional data analysis. *SN Appl. Sci.* **2020**, *2*, 461. <https://doi.org/10.1007/s42452-020-2238-x>.
45. Ghumman, A.R.; Haider, H.; Shafiqzamman, M. Functional data analysis of models for predicting temperature and precipitation under climate change scenarios. *J. Water Clim. Change* **2020**, *11*, 1748–1765. <https://doi.org/10.2166/wcc.2019.172>.

46. Hael, M.A. Modeling of rainfall variability using functional principal component method: A case study of Taiz region, Yemen. *Model. Earth Syst. Environ.* **2021**, *7*, 17–27. <https://doi.org/10.1007/s40808-020-00876-w>.
47. Wang, S.; Jank, W.; Shmueli, G. Explaining and forecasting online auction prices and their dynamics using functional data analysis. *J. Bus. Econ. Stat.* **2008**, *26*, 144–160. <https://doi.org/10.1198/073500106000000477>.
48. Wagner-Muns, I.M.; Guardiola, I.G.; Samaranayake, V.A.; Kayani, W.I. A functional data analysis approach to traffic volume forecasting. *IEEE Trans. Intell. Transp. Syst.* **2017**, *19*, 878–888. <https://doi.org/10.1109/TITS.2017.2706143>.
49. Shah, I.; Muhammad, I.; Ali, S.; Ahmed, S.; Almazah, M.M.A.; Al-Rezami, A.Y. Forecasting day-ahead traffic flow using functional time series approach. *Mathematics* **2022**, *10*, 4279. <https://doi.org/10.3390/math10224279>.
50. Liebl, D. Modeling and forecasting electricity spot prices: A functional data perspective. *Ann. Appl. Stat.* **2013**, *12*, 1562–1592. <https://doi.org/10.1214/13-AOAS652>.
51. Lisi, F.; Shah, I. Forecasting next-day electricity demand and prices based on functional models. *Energy Syst.* **2020**, *11*, 947–979. <https://doi.org/10.1007/s12667-019-00356-w>.
52. Jan, F.; Shah, I.; Ali, S. Short-Term Electricity Prices Forecasting Using Functional Time Series Analysis. *Energies* **2022**, *15*, 3423. <https://doi.org/10.3390/en15093423>.
53. Shah, I.; Jan, F.; Ali, S. Functional data approach for short-term electricity demand forecasting. *Math. Probl. Eng.* **2022**, *2022*, 6709779. <https://doi.org/10.1155/2022/6709779>.
54. Chen, Y.; Koch, T.; Lim, K.G.; Xu, X.; Zakiyeva, N. A review study of functional autoregressive models with application to energy forecasting. *Wiley Interdiscip. Rev. Comput. Stat.* **2021**, *13*, e1525. <https://doi.org/10.1002/wics.1525>.
55. Frois Caldeira, J.; Gupta, R.; Suleman, M.T.; Torrent, H.S. Forecasting the term structure of interest rates of the BRICS: Evidence from a nonparametric functional data analysis. *Emerg. Mark. Financ. Trade* **2021**, *57*, 4312–4329. <https://doi.org/10.1080/1540496X.2020.1808458>.
56. Vilar, J.M.; Cao, R.; Aneiros, G. Forecasting next-day electricity demand and price using nonparametric functional methods. *Int. J. Electr. Power Energy Syst.* **2012**, *39*, 48–55. <https://doi.org/10.1016/j.ijepes.2012.01.004>.
57. Curceac, S.; Ternynck, C.; Ouarda, T.B.M.J.; Chebana, F.; Niang, S.D. Short-term air temperature forecasting using nonparametric functional data analysis and SARMA models. *Environ. Model. Softw.* **2019**, *111*, 394–408. <https://doi.org/10.1016/j.envsoft.2018.09.017>.
58. Aue, A.; Norinho, D.D.; Hörmann, S. On the prediction of stationary functional time series. *J. Am. Stat. Assoc.* **2015**, *110*, 378–392. <https://doi.org/10.1080/01621459.2014.909317>.
59. Karhunen, K. Under linear methods in probability theory. *Ann. Acad. Sci. Fenn. Ser. A. I. Math.-Phys.* **1947**, *37*, 3–79.
60. Shang, H.L. A survey of functional principal component analysis. *ASTA Adv. Stat. Anal.* **2014**, *98*, 121–142. <https://doi.org/10.1007/s10182-013-0213-1>.
61. Hall, P.; Hosseini-Nasab, M. On properties of functional principal components analysis. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2006**, *68*, 109–126. <https://doi.org/10.1111/j.1467-9868.2005.00535.x>.
62. Hörmann, S.; Kokoszka, P. Weakly dependent functional data. *Ann. Stat.* **2010**, *38*, 1845–1884. <https://doi.org/10.1214/09-AOS768>.
63. Damon, J.; Guillas, S. far: Modelization for Functional AutoRegressive Processes. 2024. Available online: <https://CRAN.R-project.org/package=far> (14 January 2025).
64. Hyndman, R.J.; Shang, H.L. ftsa: Functional Time Series Analysis. 2025. Available online: https://CRAN.R-project.org/package=fts_a (14 January 2025).
65. Dickey, D.A.; Fuller, W.A. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.* **1979**, *74*, 427–431. <https://doi.org/10.1080/01621459.1979.10482531>.
66. Dickey, D.A.; Fuller, W.A. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* **1981**, 1057–1072. <https://doi.org/10.2307/1912517>.
67. Hörmann, S.; Kokoszka, P. Functional time series. In *Handbook of Statistics*; Elsevier: Amsterdam, The Netherlands, 2012; Volume 30, pp. 157–186. <https://doi.org/10.1016/B978-0-444-53858-1.00007-7>.
68. Brockwell, P.J.; Davis, R.A. *Time Series: Theory and Methods*; Springer: Berlin/Heidelberg, Germany, 1991. <https://doi.org/10.1007/978-1-4419-0320-4>.
69. Bosq, D. *Linear Processes in Function Spaces: Theory and Applications*; Springer: Berlin/Heidelberg, Germany, 2000. <https://doi.org/10.1007/978-1-4612-1154-9>.
70. Kokoszka, P.; Reimherr, M. Determining the order of the functional autoregressive model. *J. Time Ser. Anal.* **2013**, *34*, 116–129. <https://doi.org/10.1111/j.1467-9892.2012.00816.x>.
71. Borovkova, S.; Permana, F.J. Modelling electricity prices by the potential jump-diffusion. In *Stochastic Finance*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 239–263. https://doi.org/10.1007/0-387-28359-5_9.

72. Shah, I.; Akbar, S.; Saba, T.; Ali, S.; Rehman, A. Short-term forecasting for the electricity spot prices with extreme values treatment. *IEEE Access* **2021**, *9*, 105451–105462. <https://doi.org/10.1109/ACCESS.2021.3100076>.
73. Diebold, F.X.; Mariano, R.S. Comparing predictive accuracy. *J. Bus. Econ. Stat.* **2002**, *20*, 134–144. <https://doi.org/10.1198/073500102753410444>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.