



## OPEN High resolution temperature forecasting using functional time series decomposition and advanced predictive models

Huda M. Alshanbari<sup>1</sup>, Musaad S. Aldhabani<sup>2</sup>, Naveed Iqbal<sup>3</sup>, Abdulkafi Mohammed Saeed<sup>4</sup> & Noorullah Noori<sup>5</sup>✉

The air temperature (AT) is a significant factor in the environmental processes, human health, agriculture, and energy systems, so correct forecasting is very important. This paper explores the prediction of AT in a functional time series (FTS) model that models the high-frequency temperature data as smooth curves over the day that can model both deterministic seasonal cycles and other stochastic dynamics. Smoothing splines and Fourier-based functional representations are used to model and predict short-term variations in temperature by a functional autoregressive model, FAR(1). The performance of FAR(1) is systematically contrasted to those of classical statistical models such as ARIMA and VAR and machine-learning-based ones such as artificial neural networks (ANN) and autoregressive neural networks (ARNN). Accuracy of an error forecast is measured in terms of mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE). Empirical findings show the FAR(1) model records reduced forecasting errors and increased predictive stability both in months and in hours than the competition methods. Such results outline the utility of functional data analysis to utilize the smooth and periodic organization of temperature data. The research offers a practical and holistic model of predicting high-frequency AT, and it can be employed in the better decision-making in agricultural, energy and safety management among the population that operate during the climate variability.

**Keywords** Functional autoregressive, Functional data analysis, ARIMA, Vector autoregressive

AT is one of the most basic climatic variables and determines much of the environmental, economic and social processes, such as agriculture, energy demand, ecosystem dynamics and human health. Proper predictions of the AT are thus critical to successful planning and risk management, especially as the climate variability and extreme weather are becoming more frequent. Short term temperature projections are used in operations like balancing energy loads, frost and heatwave alerts, urban climate adaptation, as well as important inputs to hydrological and environmental models. But the AT data have complicated temporal variations due to diurnal nature, seasonal variations, and stochastic variations, which are not easy to represent using the classical time series models. This is what drives the utilization of high order modeling architectures that are capable of modelling smooth periodic behaviour and dynamic time dependence. Therefore, the creation of effective and understandable forecasting strategies of AT is a significant research topic that has direct practical implications to climate-sensitive decision-making.

Such activities that are affected include social and economic activities due to the fact that the weather patterns should be properly modeled and predictable. The future of an extreme smell provides us with valuable information that determines the efficiency of many tasks and helps prevent harm, fatalities, and property loss, as well as damage to property and loss of life and property, in the future, as well as reduce their risks and threats, when properly managed<sup>1</sup>. AT is a vital aspect in the case of growing purposes because it assists in establishing the moisture in the soil and the growth of the crops too<sup>33</sup>. The crops grow faster due to high temperatures, and it

<sup>1</sup>Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O.Box 84428, Riyadh 11671, Saudi Arabia. <sup>2</sup>Department of Mathematics, Faculty of Science, University of Tabuk, P.O.Box741, Tabuk 71491, Saudi Arabia. <sup>3</sup>Department of Mathematics, College of Science, University of Ha'il, Ha'il, Saudi Arabia. <sup>4</sup>Department of Mathematics, College of Science, Qassim University, Buraydah 51452, Saudi Arabia. <sup>5</sup>Department of Mathematics, Kabul University, Kabul, Afghanistan. ✉email: noorullah.noori@ku.edu.af

lowers the yield potential potential<sup>15</sup>. Temperature variations, however, do not play out equally across different industries, and ignoring this difference can result in more effective long-term predictions of power and energy planning as well as power consumption planning<sup>26</sup>. Weather-related variables in the fast-growing weather derivatives markets are also often used to speculate on the forecasts of the variable reviewed by the weather market operators<sup>6</sup>. The next-generation Weather Research and Forecast model has improved the prediction of AT and severe weather over the Ararat Valley mostly because it was used to predict summertime and severe weather<sup>4,9,32</sup>. Recent machine learning research also shows that environmental processes over and above the atmosphere with temperature drivers can be highly predicted, e.g.<sup>24</sup>. proposed a deep learning model to predict soil temperature, and the possibility of AI models to be used in the same manner as our functional data-driven models of AT.

The modeling and forecasting of AT have been the focus of a large number of studies since it is a fundamental field of study in most fields of study<sup>3,23,34</sup>. As an example, the authors of 2018, time-specific seasonal ARIMA (SARIMA) were used to predict the monthly mean temperatures. They utilized one-hour temperature records of a weather station in Nanjing, China, between January 1951 and December 2017<sup>14</sup>. studied the global surface temperature and concluded that dangerous climate change, as judged from likely effects on sea level and extinction of species<sup>25</sup>. studied study proposes novel algorithmic approaches utilizing machine learning techniques to predict wind turbine power. Applied algorithms include extremely randomized trees, light gradient boosting machine, ensemble methods, and the CNN-LSTM method. Based on the provided results, the lowest mean square error value is related to the CNN-LSTM method, indicating that this method is more accurate. Also, the ensemble method provides admissible results despite the high speed of the algorithm. Some of the recent studies are<sup>2,10–13,16,17,19</sup>.

This paper builds on the current literature on a topic of AT prediction through the use of functional data analysis (FDA) methods, which are effectively adapted to verify the complicated, smooth, and dynamic temporal features of temperature sequences. Even though FDA has found application in several fields, including environmental science, econometrics, and biostatistics, it has not been extensively used in AT forecasting yet<sup>20,29</sup>. Continued progress in the field of statistical time series modeling and data collection technologies has further enhanced the accuracy of the forecasting process, which is why further work on the FTS methods is considered in this area of knowledge<sup>18,21,22,27,30</sup>. In that regard, this study examines the application of FTS in improving short-term prediction of AT, where a FAR model of one with a history of AT data between January 2018 and December 2024 was trained. MAE, MAPE and RMSE are used to measure the predictive performance of the proposed framework, which gives a quantitative measure of the forecasting accuracy of the proposed framework.

The research is designed in the following way: - The “General model” section presents a general introduction about the functional data analysis, introduces the proposed model and reviews of other competing models. - “Data, and accuracy measures” section provides the data, and accuracy measures. - The results and the discussion in detail are provided in the section of the “Result and Discussion”. - Lastly, the “conclusion” section has talks and conclusion on the findings.

## General model

The components estimation procedure is used to capture different features of the AT series. Suppose the AT series  $y_{k,l}$  where  $l \in \times$  and  $k = 1, \dots, 24$ . The AT series is decomposed into a deterministic component denoted by  $D_{k,l}$  and a stochastic component denoted by  $x_{k,l}$  as

$$y_{k,l} = D_{k,l} + \mathbb{X}_{k,l}. \quad (2.1)$$

Here, the deterministic component in 2.1 comprised of the complex structure of the seasonal component, such as annual seasonality ( $s_{k,l}$ ), weekly seasonality ( $w_{k,l}$ ), and long-trend component ( $c_{k,l}$ ). Mathematically, it can be written as

$$D_{k,l} = c_{k,l} + s_{k,l} + w_{k,l} \quad (2.2)$$

The deterministic component  $D_{k,l}$  in 2.2 is estimated through a smoothing spline. In the case of stochastic component  $\mathbb{X}_{k,l}$ , three different models, namely the ARMIA, VAR, ANN, ARNN, and FAR(1) are used. Once both the components are modeled and forecasted, the one-day-ahead out-of-sample forecast is obtained. The stochastic component, which is modeled and forecasted, is discussed in detail in the next section.

One transforms the observed hourly price series to functional observations by first expanding into Fourier basis with smoothing parameters chosen through generalized cross-validation. The deterministic element is specified by the estimated smooth functional trend whereas the stochastic element is derived as the functional process of residue. The autoregressive structure is then modeled by a functional residue. The functional final prediction error criteria is used to select model order and truncation dimensions, and insures a fully data-driven and reproducible decomposition.

## Functional data analysis

The methodology was first proposed by FDA in<sup>28</sup>. Many of the classical statistical approaches have been extended and modified to the FDA model<sup>8</sup>. It is also possible to use data presented in curves, shapes or patterns - and not only discrete points - in FDA. This method is particularly suitable for information that fluctuates slowly by time or space citepramsay2005principal. In practice most functional data is measured discretely and in a high frequency and can easily be converted into functions. Functional data generally is curved and a suitable set of basis functions is used. The coefficients  $\varrho_t$  and basis functions  $\vartheta_t(i)$  may be expressed to represent a set of functions, such as a basis function system, or  $\mathbb{X}(i)$ .

$$\mathbb{X}(i) = \sum_{t=1}^T \varrho_t \vartheta_t(i), \quad , \quad 1 \in I \quad (3.1)$$

We will simplify the expression by omitting (i) in the function where the meaning is evident. In order to deal with the various types of capabilities of the data, we employ certain trendy basis features, such as the exponential basis, the B-spline basis, the Fourier basis, and the polynomial main components. An example here is that B-spline basis may be applied to non-periodic data, whereas Fourier basis may be applied well to periodic data. In this paper we have used the approach based on Fourier basis since our data is periodic. The 24 hourly temperature observations on each day were considered as one functional observation. The hourly index was linearly mapped on to the unit interval. A Fourier basis was used to represent  $[0,1]$ , and daily temperature curves and it is highly appropriate to express the high diurnal periodicity of AT. A variance-explained criterion was used to select the number of basis functions and keep the minimum number of basis functions that account to at least 95% of the overall variability, which provides a parsimonious and smooth functional representation.

### Functional autoregressive model

In high frequency environmental data, functional time series may arise when observations are functions, not a scalar or a vector, and in such applications this is a natural state of affairs, like hourly air temperature curves. In order to model time dependence between one functional observation and the next the Functional Autoregressive model of order one, FAR(1), generalizes the classical autoregressive model to an infinite dimensional function space.

Suppose that  $(\mathbb{X}_t(i) * i \in \mathbb{Z})$  is a functional time series, the values of which lie in a separable Hilbert space  $(\mathcal{H})$ , ( $L^2([0, 1])$  in general). The FAR(1) model is defined as  $[\mathbb{X}_t = \Psi(\mathbb{X}_{t-1}) + \varepsilon_t, \quad t \in \mathbb{Z},]$  In which, the bounded linear operator, which describes the response to time-dependent variations between successive curves, is denoted by  $(\Psi : \mathcal{H} \rightarrow \mathcal{H})$  and the sequence of independent and identically distributed functions with zero mean and finite second moment.

The FAR(1) process has a unique strictly stationary solution under conducive conditions on the operator norm of  $(\Psi)$ . Practically, the operator  $(\Psi)$  is approximated by functional principal component analysis (FPCA) and the large-dimensional problem is approximated in a finite-dimensional subspace. This model retains the smooth form of functional observations and defining the smooth, curve-based processes and dynamics of functional observations, FAR(1) is specifically well suited to forecasting smooth processes, including daily AT profiles.

#### Operators in the Hilbert space $(L^2[0, 1])$

Linear operators in FTS modeling operate on the elements of the Hilbert space of  $L^2[0, 1]$  denoted by  $\mathcal{H}$ . The operator norm  $\Psi$  is used to measure the size of a linear operator (bounded or not) titled as.

$$\|\Psi\|_L = \sup_{\|x\| \leq 1} \|\Psi(x)\|.$$

It is a norm that makes the operator stable and is significant in determining stationarity conditions of functional autoregressive models.

A particular operator, denoted by the symbol  $\Psi$  is called compact when it has a representation, in the form of a sequence of decreasing coefficients, denoted by the  $\{\lambda_t\}$ , and orthonormal basis functions, i.e.

$$\Psi(x) = \sum_{t=1}^{\infty} \lambda_t \langle x, \nu_t \rangle f_t,$$

where  $\lambda_t \rightarrow 0$  as  $t \rightarrow \infty$ . Compact operators have very convenient applications in practice, since they can be very well approximated by a finite number of basis functions.

Another most commonly employed type of compact operators are the Hilbert Schmidt operators, where the squared coefficients are summable,

$$\sum_{t=1}^{\infty} \lambda_t^2 < \infty.$$

A natural norm is the one that is confessed by such operators.

$$\|\Psi\|_S = \left( \sum_{t=1}^{\infty} \lambda_t^2 \right)^{1/2},$$

that is more powerful than the operator norm. Assuming, further, that the operator is symmetric, positive, it has a spectral decomposition of the form.

$$\Psi(x) = \sum_{t=1}^{\infty} \lambda_t \langle x, \nu_t \rangle \nu_t.$$

In case the absolute value of the sum of the coefficients is finite, in other words,

$$\sum_{t=1}^{\infty} |\lambda_t| < \infty,$$

the operator is known as nuclear, which is a stronger regularity requirement. These norms of operators fulfill the ordering.

$$\|\cdot\|_n \geq \|\cdot\|_S \geq \|\cdot\|_L.$$

Although the FAR(1) model is theoretically formulated as a set of bounded linear operators acting on a Hilbert space, the operationalization of the model in the present study uses basis expansion, truncation and finite dimensional approximation which are clearly explained in the following sections.

#### Estimation of the operator $\Psi$

In order to ensure a stationary solution, some assumptions have to be made when estimating the autoregressive operator sigma in the model. In particular, the operator norm of  $\Psi$  can be used to assure stationarity. The sufficient condition is that an integer  $t_0$  exists such that  $\|\Psi^{t_0}\|_L < 1$ . Similarly, this assumes that constants  $a > 0$  and  $0 < b < 1$ , exist for which  $\|\Psi^t\|_L \leq ab^t$  is also true for all constants  $t \geq 0$ . Under these assumptions, it is shown that there is a unique strictly stationary solution in the functional autoregressive process in certain contexts<sup>5</sup>.

Since the functional data are defined in infinite-dimensional space that is not locally compact, Lebesgue measure is not present, i.e., probability densities cannot be defined so. Therefore, likelihood-based estimation methods cannot be employed in the estimation of the operator  $\Psi$ . Rather, estimation is conducted by the moment method. In this case, the autoregressive operator is  $\Psi = C\Gamma^{-1}$ , where  $\Gamma = E(x_t \otimes x_t)$  is the covariance operator, and  $C = E(x_t \otimes x_{t+1})$  is the lag-one cross-covariance operator, and  $\otimes$  the Kronecker product  $\hat{\Gamma}_i$ . The numerical equivalents are noted with  $\hat{\Gamma}_i$  and  $\hat{C}_i$ , respectively, and are used in practice to estimate the  $\Psi$ .

For simplicity, the functional time series must be mean centered, i.e.,  $E(x_i) = 0$ . This is done in practice by subtracting the sample mean curve for each observation. Finally, we estimate the sample covariance operators  $\hat{\Gamma}_i = \frac{1}{n} \sum_{i=0}^{n-1} x_i \otimes x_i$  and  $\hat{C}_i = \frac{1}{n} \sum_{i=0}^{n-1} x_i \otimes x_{i+1}$ , which report the temporal variation of daily curves of temperature.

Practically, covariance operator, denoted as  $\Gamma$ , is estimated using discretely observed functional data and inverting it is numerically unstable. As a solution to this problem, a dimensionality reduction is carried out with the help of FPCA. In particular, decomposition of the set of eigenvalues and eigenfunctions is made up of a finite set of eigenvalues and eigenfunctions and only the first  $w$  components that capture most of the overall variability are retained. The truncation provides a stable and well-conditioned approximation of the covariance operator inverse.

The empirical covariance and cross covariance operators are calculated as

$$\hat{\Gamma} = \frac{1}{n} \sum_{i=1}^n x_i \otimes x_i, \quad \hat{C} = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i \otimes x_{i+1},$$

then the FAR(1) dependence operator is estimated using them. The approximate inverse of using the truncated inverse of the  $\Gamma$  is in the form of:

$$\hat{\Gamma}^{-1}(x) = \sum_{t=1}^w \hat{\lambda}_t^{-1} \langle x, \hat{\nu}_t \rangle \hat{\nu}_t,$$

and the empirical eigenvalues and eigenfunctions are denoted by  $\hat{\lambda}_t$  and  $\hat{\nu}_t$  respectively, as: In FAR(1) structure, the dependence operator is approximated by the expression, which is given by,  $\Psi = C\Gamma^{-1}$  and the estimator is given by,  $\hat{\Psi} = \hat{C}\hat{\Gamma}^{-1}$ . In its application, the observed curves are smoothed before FPCA to minimize measurement noise and assure stable estimation. Through the choice of truncation level, where the retained components capture between 90% to 95% of overall variation, a compromise was made between the performance of the model and computational efficiency.

This is because the FAR(1) model can estimate the strong temporal variation in AT curves in this FPCA-based estimation strategy that is numerically stable and can be applied in large scale AT forecasting models.

#### Competing models

In order to compare the suggested FAR(1) model, a number of rival methods were taken into account. The information criteria (AIC/BIC) was used to select ARIMA models amongst different candidate orders and residual diagnostics validated that there was no serial correlation. The VAR model was also built on hourly temperature vectors that were lagged to reflect the presence of cross-hour dependence, the lag order was chosen

using AIC. ANN and ARNN models were used as machine learning baselines. Validation-based tuning was used to select network architectures, which are denoted by ARNN (p, k), where p is the lagged number of inputs, and k is the number of hidden neurons. Each of the neural networks followed a logistic sigmoid activation function in the hidden layer and a linear output layer, was trained with the Adam optimizer and a learning rate of 0.001, and early stopping to limit overfitting.

#### Autoregressive integrated moving average (ARIMA) model

The ARIMA model is a classical linear model in TS that is extensive in univariate forecasting because it has a strong statistical basis and can be interpreted. It is a combination of AR and moving average applied on the differenced series to deal with non-stationarity in the mean. AR terms embody the temporal persistence and the moving average component is used to describe the impact of historical shocks. The information criteria (e.g. AIC) are usually used to select model orders; residual diagnostics are also used to verify that the model fits well. Though ARIMA can do well with linear and moderately structured data, it represents observations as discrete points and can not well represent smooth curve-based dynamics that are found in high-frequency series of AT, which prompts the adoption of functional alternatives like FAR models.

Suppose that  $y_t$  is a univariate TS. ARIMA (p,d,q) model is described such that (d) order differencing is applied to ( $y_t$ ) to make the series stationary.

$$x_t = \nabla^d y_t = (1 - B)^d y_t,$$

where (B) is the backshift operator, ( $B y_t = y_{t-1}$ ).

The differenced series ( $x_t$ ) follows an autoregressive moving average process ARMA((p,q)),

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q},$$

where:

( $\phi_1, \dots, \phi_p$ ) are autoregressive coefficients, ( $\theta_1, \dots, \theta_q$ ) are moving average coefficients,  $\varepsilon_t$  is a white-noise error term with zero mean and variance ( $\sigma^2$ ).

Equivalently, using lag-polynomial notation, the ARIMA((p,d,q)) model can be written as

$$\Phi(B)(1 - B)^d y_t = \Theta(B)\varepsilon_t,$$

where

$$\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p, \quad \Theta(B) = 1 + \theta_1 B + \dots + \theta_q B^q.$$

#### Vector autoregressive

In TS analysis, the VAR model has been a widely used and useful tool, especially to describe the dynamic relationship of multiple variables over time. It generalises the univariate AR model to multivariate ones. The prediction and forecasting capability and flexibility of VAR models make it very effective in interpreting and predicting actual world behavior since they are able to capture numerous interdependences between variables. A VAR model is a collection of equations, which explain the dynamic of various TS variables. In this model, the variables are linear functions of their past, and the previous value of all other variables. The overall structure of a VAR model when the TS is represented by a vector of univariates, i.e. by vectors of the variables, is:

$$\mathbf{x}_t = C + \Delta_1 \mathbf{x}_{t-1} + \Delta_2 \mathbf{x}_{t-2} + \dots + \Delta_p \mathbf{x}_{t-p} + \varepsilon_t$$

The present values of the response TS variables with  $\times$  unique variables of response are represented as the vector  $\mathbf{x}_t$ . The fixed vector  $C$  characterizes the base levels or intercepts of such variables. The AR coefficient matrices of the variables are  $\Delta_t$  are  $\times \times \times$  matrices of dimensions of size, for example, at a given time lag  $t$ , say at lag 1 to 10, denoted as the matrices of size of p, they are the relationship between the variables at various time lags more so how the former values up to lag p have an influence on the present values.

There are common AIC, BIC and HQ information criteria used together with the cross-validation methods to estimate the best model order. Typical estimated models are those that are estimated during a 365-day period and that the value adopted of p might be different. Nevertheless, VAR(5) model is usually regarded as suitable, because most of the time residuals are white noise, the value of AIC is lower than other models of higher order. Ordinary least squares (OLS) and maximum likelihood estimation (MLE) are widely used in estimating and inferring parameters of VAR models.

#### Artificial neural networks

Artificial neural networks (ANNs) refer to a machine learning method that is based on the human brain process of information processing. They are a number of layers of interconnected nodes resembling biological neurons. Such networks are desired to resemble the learning and the decision-making processes of the brain.

The input layer provides the system with information. It is then passed through one or more intermediate layers to the output layer. All nodes in the network get signals as sent by other nodes, process the inputs and produce a resultant output. The weight of the connections between nodes is known to have an impact on the influences the node has on the other. The feedforward neural network is one of the commonly used types of ANN. This arrangement is unidirectional in the flow of data, i.e. the input layer to the output layer without feedback loops. In the case of the neural network-based models, nonlinear temperature dynamics were represented by the

use of a logistic sigmoid activation function in the hidden layer, and a linear activation function in the output layer to make sure that forecasts were in the original temperature scale.

### Autoregressive neural networks

The Autoregressive neural networks (ARNN) is a hybrid model that blends traditional autoregressive concepts with the capability of neural networks to handle nonlinear relationships. It takes previous values of time series as inputs in order to make future observations. Its hidden layers have the ability to determine complicated nonlinear relationships over time (unlike linear AR models). This type of structure is referred to as ARNN( $p$ ,  $k$ ), which has  $p$  lags and  $k$  hidden neurons. It is not restrictive of strict stationarity, which affords it greater applicability to a variety of datasets. With the inputs of the recent observations, the ARNN makes good one-step-ahead predictions that fit multi-dimensional time series patterns. The equation of the model will be:

$$x_{jt} = g \left( \sum_{l=1}^{\infty} w_{jl} x_{l(t-1)} \right) + \kappa y_{j(t-1)}^T \epsilon_t$$

In this case,  $x_{jt}$  is the response at node  $j$ ,  $t$ . The scoring pattern denoted by  $w_{ji}$  denotes the relationship strength between node  $j$  and node  $i$ . The expression denotes external effects, and so the cost is captured by the term of the form of  $\kappa y_{j(t-1)}^T$ ,  $g(\cdot)$  is a nonlinear activation function and the error term during time  $t$  is denoted by the term of the form of the error term denoted by the term  $\epsilon_t$ . The ARNN model updates its weights with the help of estimation techniques, which include profile least squares estimation and local linear approximation. These techniques aid in estimation of the unknown relationship and provide quality estimates of parameters. The model is based on a training procedure over historical data used to perform one-step-ahead and multi-step predictions. The ARNN(3,7) model is adopted in this work, that is, three past observations were taken as input and the hidden layer consisted of seven neurons, the results of which were determined after trial and error.

### Data and accuracy measures

#### Data

The paper is based on hourly AT measurements of Tabuk, in Saudi Arabia, which were provided through official weather records. Our data is the period between 1 January 2018 and 31 December 2024. To prevent information leakage and have a realistic forecasting setup, the data were divided chronologically: the data between 1 January 2018 and 31 December 2023 was used to train and tune the model, whereas the data between 1 January 2024 and 31 December 2024 were used in out-of-sample testing.

#### Forecasting protocol

The prediction horizon used to forecast was one day ahead. A strategy of expanding window was used, in which new models were re-estimated with the arrival of new observations and they imitated real-time forecasting conditions. Performance was assessed over model performance on MAE, MAPE and RMSE valued on the original temperature scale across all the methods in order to provide fair comparison. Even though the last year (2024) is marked as a test period, the forecasting is made using an expanding-window strategy whereby, model parameters are re-estimated one by one as new data is supplied. This rolling-origin assessment produces day-by-day one-step-ahead predictions across the whole test year, which offers a variety of out-of-sample validation points as opposed to a single stationary split. Also, robustness is evaluated at hourly, monthly, and annual scales so that the comparisons of performance remain consistent at varying time scales and seasonal factors.

#### Forecasting accuracy metrics

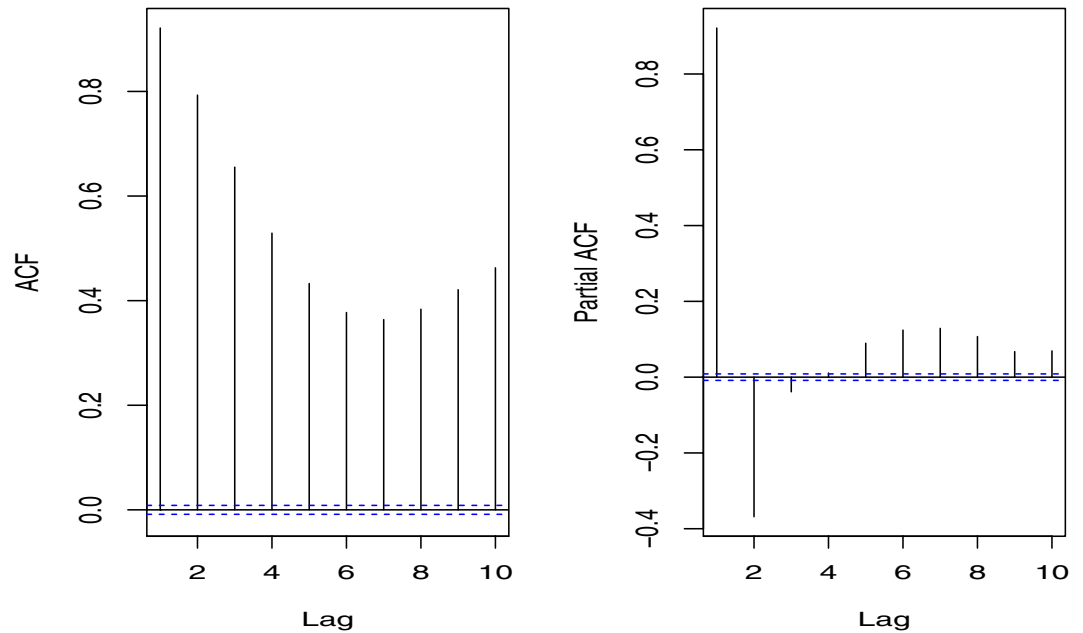
We apply three measures of standard error MAE, MAPE, and RMSE to assess the accuracy of the forecasting models. These error measures which are used to measure the degree of the closeness of the forecasts and actual values are frequently descriptive statistics. In mathematics, they are as follows:

$$\begin{aligned} \text{MAE} &= \frac{1}{n} \sum_{k=1}^n |y_{k,l} - \hat{y}_{k,l}| \\ \text{MAPE} &= \frac{1}{n} \sum_{k=1}^n \left( \left| \frac{y_{k,l} - \hat{y}_{k,l}}{y_{k,l}} \right| \right) \times 100 \\ \text{RMSE} &= \sqrt{\text{mean}(y_{k,l} - \hat{y}_{k,l})^2} \end{aligned}$$

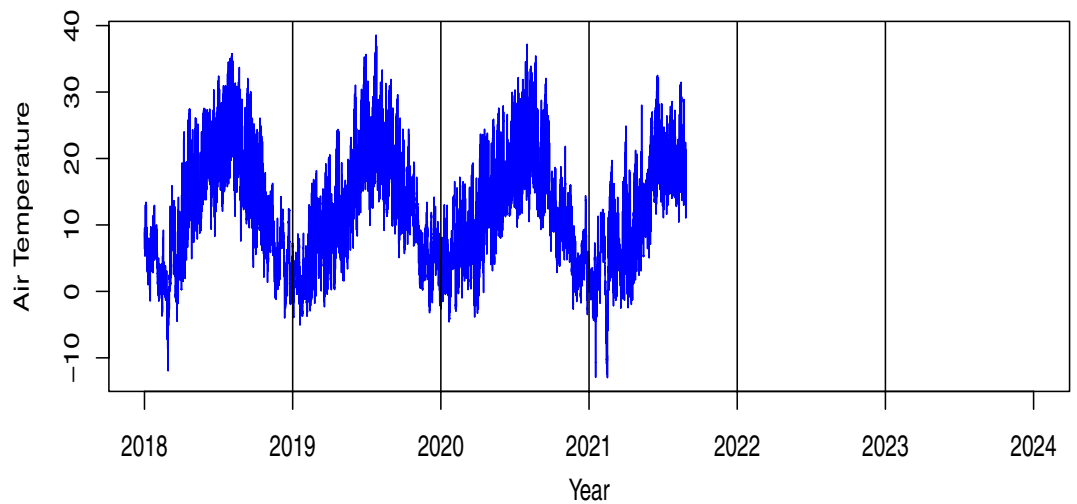
Where:  $n$  is the observations of the testing forecast period.  $y_{k,l}$  and  $\hat{y}_{k,l}$  are the actual and predicted values of the AT on the  $k$ th day ( $k = 1, 2, \dots, 365$ ) and  $l^{\text{th}}$  hour ( $l = 1, 2, \dots, 24$ ).

### Results and discussion

The Fig. 1 will have several plots illustrating both the autocorrelation function (ACF) and partial autocorrelation function (PACF) at different lags that will be used to analyze the TS data. The ACF plots show the relationship between observations at various lags, which can be used to determine the presence of seasonality or trend items. The direct correlation between observations at particular lags is isolated in the PACF plots and intermediate lags



**Fig. 1.** Autocorrelation (ACF) and partial autocorrelation (PACF) diagnostics for the AT series, illustrating temporal dependence across lags.

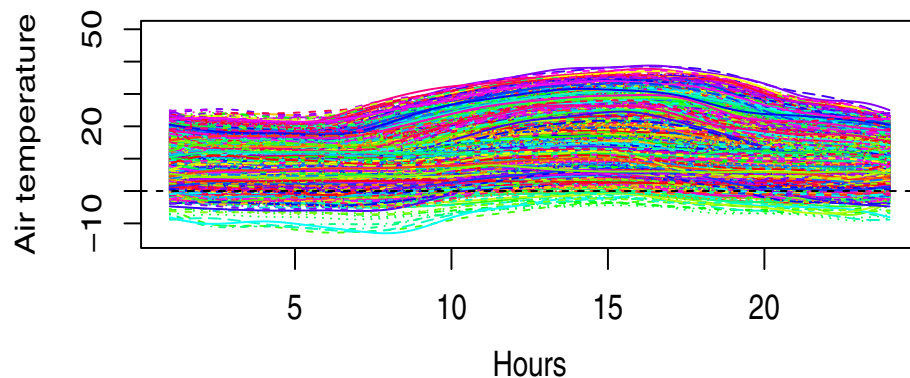


**Fig. 2.** AT data used for FTS modeling (2018–2024).

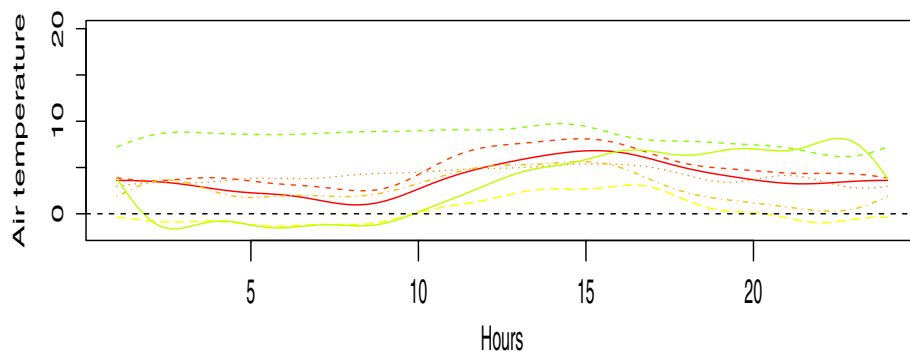
are controlled, which helps in the selection of order of AR models. These plots in combination aid in diagnostic checking and model identification of TS analysis.

The Fig. 2 demonstrates that the AT changes over the years, with distinct seasonal and interannual changes. The pattern is strong in terms of period and could be observed as temperatures were higher in the season of warmth and when it was colder they were lower which is a natural annual cycle of temperature. Although consumers experience temporary fluctuations, the general arrangement of the series is consistent over the years, suggesting that there is a consistent pattern of climatic behavior in the period of study. The large variation in temperature indicates the existence of extreme low and extreme high values and hence the significance of having strong forecasting models that can be used to captivate such a variation.

The Fig. 3 shows the transformation of the discrete TS data into a continuous functional form by implementing the Fourier basis function. It shows hourly AT measurements taken between January 1, 2018, and December 31, 2024, in which the vertical axis shows the temperature in degrees Celsius and the horizontal axis shows time of day in hours. The Fourier basis functions are used to smooth the discrete observations to functional curves to capture daily and seasonal repetitive patterns. This method enables to be more flexible in modeling and analysis of the underlying dynamics of the temperature over multi-year time range.



**Fig. 3.** The transformation of the discrete AT data into functional data.



**Fig. 4.** shows the functional data of one week.

Model	MAE	MAPE	RMSE
ARIMA	30.73	28.99	29.46
VAR	24.65	22.83	23.73
ANN	28.36	26.75	27.88
ARNN	25.23	23.46	24.30
FAR(1)	21.44	19.52	20.92

**Table 1.** Overall forecast results for AT.

The Fig. 4 indicates one week of AT converted into discrete measurements into functional curves with the use of Fourier basis. The time of the day is indicated by the horizontal axis (in hours) and temperature by the vertical axis (in degrees Celsius). Every straight line symbolizes the daily pattern of the temperature that is continuously changing, and that emphasizes the diurnal cycle and daily changes throughout the week.

Table 1 gives a summary of the overall forecasting of all the competing AT prediction models. Compared to ARIMA and ANN, VAR and ARNN are better in terms of multivariate and nonlinear dependencies, which makes them superior to classical and machine learning methods. Nevertheless, the proposed FAR(1) model presents the least MAE and MAPE and RMSE values, which means that the accuracy of forecasting is improved with clear and consistent results. These findings affirm that the approach of modeling AT as a functional process is a more powerful and effective framework compared to the TS and neural network based AT modeling.

The comprehensive comparison of AT forecasting errors in all competing models was reported in Table 2. The proposed FAR(1) model shows the smallest values of MAE, MAPE, and RMSE throughout the full 24-hour cycle, which means that it is more accurate and the model is stable during the day and night. The errors in classical models like ARIMA and VAR are relatively higher especially in early morning and late evening hours when the temperature dynamics are more prone to be volatile. Even though ANN and ARNN are able to pick up certain nonlinear patterns, their prediction errors are higher and more volatile than those of the functional model in most of the hours. On the whole, these findings indicate that FAR(1) as a smooth functional process modeling of temperature better reflects intraday dynamics and shows strong improvements over conventional TS and machine-learning algorithms.

Hour	Errors	ARIMA	VAR	ANN	ARNN	FAR(1)
1	MAE	29.95	23.30	27.37	24.56	20.99
	MAPE	25.28	21.38	23.98	21.88	17.00
	RMSE	27.36	22.56	25.36	23.56	19.22
2	MAE	29.39	25.55	27.16	26.67	21.39
	MAPE	26.24	23.81	25.79	24.13	19.60
	RMSE	27.33	24.16	26.99	25.89	20.03
3	MAE	32.94	26.95	29.88	28.88	23.09
	MAPE	29.98	23.77	26.66	25.57	21.21
	RMSE	30.47	25.11	27.99	26.85	22.45
4	MAE	29.75	25.10	28.33	26.56	22.67
	MAPE	27.57	23.46	26.03	24.45	20.93
	RMSE	28.49	24.81	27.19	25.04	21.28
5	MAE	31.84	26.93	29.25	27.77	19.13
	MAPE	29.73	24.85	27.88	25.92	17.18
	RMSE	30.94	25.16	28.95	26.97	18.53
6	MAE	33.44	27.34	29.05	28.80	22.71
	MAPE	32.18	25.22	27.39	26.70	20.09
	RMSE	32.91	26.51	28.77	27.99	21.77
7	MAE	31.23	26.83	28.27	27.77	21.99
	MAPE	29.78	24.55	26.05	25.15	19.99
	RMSE	30.39	25.33	27.08	26.99	20.19
8	MAE	30.66	24.59	27.32	26.88	22.00
	MAPE	28.09	22.06	25.89	24.91	20.76
	RMSE	29.98	23.75	26.05	25.17	21.10
9	MAE	30.55	25.53	28.72	27.03	20.78
	MAPE	28.68	23.89	26.67	25.33	18.07
	RMSE	29.22	24.85	27.95	26.89	19.75
10	MAE	29.27	25.08	27.87	26.66	21.78
	MAPE	27.33	22.29	25.07	24.15	19.19
	RMSE	28.38	24.87	26.65	25.57	20.79
11	MAE	29.34	25.77	27.09	26.11	19.48
	MAPE	27.44	23.35	25.33	24.61	17.17
	RMSE	28.81	24.25	26.95	25.49	18.97
12	MAE	31.82	26.22	28.97	27.73	20.08
	MAPE	29.88	24.65	26.07	25.34	18.38
	RMSE	30.18	25.85	27.63	26.54	19.96
13	MAE	29.19	25.85	27.08	26.93	20.56
	MAPE	27.16	23.73	25.23	24.86	18.58
	RMSE	28.78	24.05	26.88	25.05	19.15
14	MAE	33.49	25.82	28.05	27.05	22.45
	MAPE	31.66	23.67	26.28	25.56	20.61
	RMSE	32.98	24.01	27.99	26.06	21.18
15	MAE	30.83	25.19	27.63	26.65	19.88
	MAPE	28.94	23.63	25.00	24.04	17.98
	RMSE	29.32	24.85	26.35	25.26	18.35
16	MAE	28.56	22.77	25.36	24.76	19.56
	MAPE	26.23	20.98	23.83	22.05	17.63
	RMSE	27.08	21.15	24.05	23.00	19.95
17	MAE	32.94	24.38	29.45	27.86	21.67
	MAPE	30.55	22.62	27.94	25.11	19.04
	RMSE	31.89	23.85	28.17	26.90	20.07
18	MAE	29.99	25.11	27.83	26.88	20.92
	MAPE	27.17	23.39	25.68	24.11	18.58
	RMSE	28.10	24.95	26.95	25.58	19.26
Continued						

Hour	Errors	ARIMA	VAR	ANN	ARNN	FAR(1)
19	MAE	31.09	23.18	27.14	25.95	18.66
	MAPE	29.46	21.75	25.90	23.32	16.00
	RMSE	30.78	22.05	26.05	24.47	17.25
20	MAE	29.93	22.99	27.11	26.66	18.99
	MAPE	27.46	20.78	25.08	24.34	16.23
	RMSE	28.78	21.25	26.32	25.66	17.07
21	MAE	31.34	25.93	29.17	27.88	20.14
	MAPE	29.66	23.03	27.88	25.13	18.85
	RMSE	30.98	24.55	28.05	26.77	19.03
22	MAE	31.77	25.16	28.91	26.45	22.35
	MAPE	29.88	23.28	26.35	24.11	20.03
	RMSE	30.18	24.85	27.95	25.67	21.56
23	MAE	32.07	26.49	30.01	29.05	21.03
	MAPE	30.16	24.08	28.16	27.67	19.65
	RMSE	31.84	25.75	29.51	28.96	20.89
24	MAE	28.99	25.19	28.92	26.23	20.88
	MAPE	26.04	23.38	26.56	24.86	18.33
	RMSE	27.98	24.75	27.85	25.06	19.96

**Table 2.** Hour-specific AT forecasting errors.

Table 3 is a table that gives a complete monthly comparison of AT forecasting errors in all the competing models. The proposed FAR(1) model is the lowest in the value of MAE and MAPE and RMSE in all months, which implies its high accuracy and strength under various seasonal conditions. The classical models of ARIMA and VAR have moderate performance but have significantly higher error especially when there are higher variability of the temperature in the month. The ANN and ARNN models identify a few nonlinear trends and yet perform poorly as compared to the functional approach with less predictive error dynamics between months. On the whole, the findings confirm that the trend of AT modeling as a smooth functional process on the basis of FAR(1) shows definite and consistent improvements compared to the classical TS and machine-learning approaches.

The better functioning of FAR(1) is especially seen in cases when it is winter and temperature dynamics are more hostile and nonlinear effects are sharp. During such timeframes, FAR(1) has reached up to 20% of MAE and RMSE gains over ARNN, which demonstrates that it is more robust when faced with difficult conditions. Hourly analysis also demonstrates that FAR(1) has a steady accuracy in the whole 24 hours cycle with significant increases in early morning and late evening hours. Although it is linear, FAR(1) has an advantage over ANN-based models because it is capable of making good use of the smooth, periodic, functional characteristics of temperature curves. However, the paper is restricted to one city, one FAR order, and it does not involve deep learning benchmarks, in the future research works it will be eliminated by applying LSTM and Transformer models into the same functional framework.

## Conclusion

AT is a fundamental meteorological parameter that has vast impacts on numerous processes of the surrounding environment, economy, and society, and it is necessary to make effective predictions. Nevertheless, complicated deterministic structures are present along with stochastic variability of time series data, and prediction of all these components continues to be a major difficulty in AT prediction. In order to solve this problem, a functional data analysis framework of one-day-ahead AT forecasting has been used in this study where a component estimation approach is implemented which explicitly decomposes the series into deterministic and stochastic parts. The deterministic component is determined with the help of smoothing splines to identify long-term trends and seasonal dynamics whereas the stochastic component is modeled with the help of a set of rival models such as FAR, ARIMA, VAR, ANN, and ARNN.

Empirical analysis is performed based on AT data of Tabuk, Saudi Arabia, and forecasting performance of a full year of out-of-sample predictions is assessed with the measures of MAE, MAPE and RMSE. The findings confirm that the presented component-based functional approach provides a systematically higher predictive accuracy, to a significant degree of minimizing forecast errors as compared to classical time series and neural network benchmarking. Specifically, the multivariate and functional models are superior to the univariate statistical and machine-learning approaches and this pointed out the benefit of considering the underlying functional structure of AT dynamics.

In spite of these positive outcomes, some limitations are to be mentioned. The discussion is limited to one geographical location as well as only linear, parametric models are analyzed. Besides, the possible effect of exogenous variables is not taken into account, and it is possible that the further research will allow covering only significant gaps by including more covariates and nonlinear models.

Month	Errors	ARIMA	VAR	ANN	ARNN	FAR(1)
January	MAE	29.29	25.88	27.85	26.56	19.64
	MAPE	27.93	23.38	25.23	24.84	17.11
	RMSE	28.04	24.68	26.75	26.12	18.72
February	MAE	29.73	25.56	28.19	27.10	21.13
	MAPE	26.44	23.16	26.15	25.26	19.55
	RMSE	27.74	24.55	27.55	26.70	20.94
March	MAE	31.90	25.17	28.80	26.86	20.78
	MAPE	29.02	23.75	26.06	24.02	18.24
	RMSE	30.44	24.06	27.85	25.00	19.88
April	MAE	31.75	25.63	29.39	27.34	21.76
	MAPE	29.53	23.45	27.38	25.17	19.89
	RMSE	30.84	24.86	28.95	26.09	20.18
May	MAE	31.59	25.28	28.59	26.88	19.78
	MAPE	29.90	23.88	26.80	24.00	17.12
	RMSE	30.14	24.88	27.35	25.13	18.76
June	MAE	30.77	25.33	27.96	26.94	19.89
	MAPE	28.93	23.77	25.89	24.00	17.29
	RMSE	29.24	24.96	26.25	25.38	18.73
July	MAE	30.44	24.83	28.77	26.13	20.23
	MAPE	28.82	22.02	26.35	24.67	18.23
	RMSE	29.14	23.76	27.65	25.99	19.78
August	MAE	30.04	23.97	27.08	25.02	19.45
	MAPE	28.68	21.12	25.49	23.56	17.73
	RMSE	29.84	22.87	26.85	24.88	18.99
September	MAE	30.84	24.98	27.54	26.77	18.56
	MAPE	28.49	22.28	25.99	24.25	16.13
	RMSE	29.98	23.16	26.25	25.62	17.76
October	MAE	29.48	24.86	27.08	26.02	20.66
	MAPE	27.75	22.16	25.59	24.34	18.13
	RMSE	28.14	23.56	26.95	25.83	19.67
November	MAE	30.48	25.87	27.56	26.66	20.24
	MAPE	28.88	23.18	25.98	24.04	18.19
	RMSE	29.14	24.96	26.55	25.67	19.58
December	MAE	32.11	26.89	29.63	27.88	21.76
	MAPE	30.25	24.94	27.19	25.99	19.07
	RMSE	31.84	25.16	28.75	26.45	20.65

**Table 3.** Month-specific AT forecasting errors.

## Data availability

Data of this research and the code used in the analyses described in this paper are available from the corresponding author upon reasonable request.

Received: 15 November 2025; Accepted: 16 February 2026

Published online: 25 February 2026

## References

- Al-Matarneh, L., Sheta, A., Bani-Ahmad, S., Alshaer, J. & Al-Oqily, I. Development of temperature-based weather forecasting models using neural networks and fuzzy logic. *Int. J. Multimed. Ubiquitous Eng.* **9**(12), 343–366 (2014).
- Arshad, I. et al. Performance of classification algorithms under class imbalance: Simulation and real-world evidence. *IEEE Access* <https://doi.org/10.1109/access.2025.3620264> (2025).
- Asha, J., Rishidas, S. et al., “Forecasting performance comparison of daily maximum temperature using ARMA based methods,” in *Journal of Physics: Conference Series*, **1921**(1), 012041 (2021).
- Atsatryan, H., et al. “WRF-ARW model for prediction of high temperatures in south and South East Regions of Armenia,” in *2015 IEEE 11th International Conference on e-Science*, 207–213 (2015).
- Bosq, D. *Linear Processes in Function Spaces: Theory and Applications*, **149**, (Springer Science & Business Media, 2000).
- Campbell, S. D. & Diebold, F. X. Weather forecasting for weather derivatives. *J. Am. Stat. Assoc.* **100**(469), 6–16 (2005).
- Curceac, S., Ternynck, C., Ouada, 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.* **111**, 394–408 (2019).
- Ferraty, F., & Vieu, P. *Nonparametric Functional Data Analysis: Theory and Practice*. (Springer, 2006).
- Gevorgyan, A. Summertime wind climate in Yerevan: Valley wind systems. *Clim. Dyn.* **48**(5), 1827–1840 (2017).

10. Guo, Q. et al. Air pollution forecasting using artificial and wavelet neural networks with meteorological conditions. *Aerosol Air Qual. Res.* **20**(6), 1429–1439 (2020).
11. Guo, Q., He, Z. & Wang, Z. Long-term projection of future climate change over the twenty-first century in the Sahara region in Africa under four Shared Socio-Economic Pathways scenarios. *Environ. Sci. Pollut. Res.* **30**(9), 22319–22329 (2023).
12. Guo, Q., He, Z. & Wang, Z. Predicting of daily PM<sub>2.5</sub> concentration employing wavelet artificial neural networks based on meteorological elements in Shanghai, China. *Toxics* **11**(1), 51 (2023).
13. Guo, Q., He, Z. & Wang, Z. Monthly climate prediction using deep convolutional neural network and long short-term memory. *Sci. Rep.* **14**(1), 17748 (2024).
14. Hansen, J. et al. Global temperature change. *Proc. Natl. Acad. Sci. U. S. A.* **103**(39), 14288–14293 (2006).
15. Hatfield, J. L. et al. Temperature extremes: Effect on plant growth and development. *Weather Clim. Extrem.* **10**, 4–10 (2015).
16. He, Z. & Guo, Q. Comparative analysis of multiple deep learning models for forecasting monthly ambient PM<sub>2.5</sub> concentrations: A case study in Dezhou City, China. *Atmosphere* **15**(12), 1432 (2024).
17. He, Z., Guo, Q., Wang, Z. & Li, X. Prediction of monthly PM<sub>2.5</sub> concentration in Liaocheng in China employing artificial neural network. *Atmosphere* **13**(8), 1221 (2022).
18. Jan, F. et al. Advancing economic forecasting with hybrid time series and functional models: Evidence from electricity demand data. *IEEE Access* **13**, 197864–197875 (2025).
19. Jan, F., Iftikhar, H., Tahir, M. & Khan, M. Forecasting day-ahead electric power prices with functional data analysis. *Front. Energy Res.* **13**, 147–168 (2025).
20. Jan, F., Shah, I. & Ali, S. Short-term electricity prices forecasting using functional time series analysis. *Energies* **15**(9), 3423 (2022).
21. Kowal, D. R. Integer-valued functional data analysis for measles forecasting. *Biometrics* **75**(4), 1321–1333 (2019).
22. Shang, H. L. & Haberman, S. Weighted compositional functional data analysis for modeling and forecasting life-table death counts. *J. Forecast.* **43**(8), 3051–3071 (2024).
23. Liu, Y., Ortega-Farías, S., Tian, F., Wang, S. & Li, S. Estimation of surface and near-surface air temperatures in arid northwest China using Landsat satellite images. *Front. Environ. Sci.* **9**, 791336 (2021).
24. Makumbura, R. K. et al. Advancing water quality assessment and prediction using machine learning models, coupled with explainable artificial intelligence (XAI) techniques like shapley additive explanations (SHAP) for interpreting the black-box nature. *Results Eng.* **23**, 102831 (2024).
25. Malakouti, S. M., Ghiasi, A. R., Ghavifekr, A. A. & Emami, P. Predicting wind power generation using machine learning and CNN-LSTM approaches. *Wind Eng.* **46**(6), 1853–1869 (2022).
26. Moral-Carcedo, J. & Pérez-García, J. Temperature effects on firms' electricity demand: An analysis of sectorial differences in Spain. *Appl. Energy* **142**, 407–425 (2015).
27. Poldrack, R.A., Mumford, J.A., & Nichols, T.E. *Handbook of Functional MRI Data Analysis*. (Cambridge University Press, 2024).
28. Ramsay, J. O. When the data are functions. *Psychometrika* **47**, 379–396 (1982).
29. Shah, I. et al. Forecasting day-ahead traffic flow using functional time series approach. *Mathematics* **10**(22), 4279 (2022).
30. Shukla, S., Hong, T., & Hyndman, R.J. "Functional data analysis for peak shape forecasting," (2024).
31. Shumway, R. H., Stoffer, D. S. & Stoffer, D. S. *Time Series Analysis and Its Applications* Vol. 3 (Springer, 2000).
32. Skamarock, W. C. et al. A description of the advanced research WRF version 3. *NCAR Technical Note* **475**(125), 10–5065 (2008).
33. Webber, H. et al. Simulating canopy temperature for modelling heat stress in cereals. *Environ. Model. Softw.* **77**, 143–155 (2016).
34. Zhu, Y. et al. Forecast calibrations of surface air temperature over Xinjiang based on U-net neural network. *Front. Environ. Sci.* **10**, 1011321 (2022).

## Acknowledgements

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R 299), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

## Author contributions

Huda M Alshabari, Musaad S. Aldhabani, Naveed Iqbal, Abdulkafi Mohammed Saeed, Noor Ullah Noori; Conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing-original draft preparation, writingreview and editing, visualization, supervision, project administration. All authors have read and agreed to the published version of the manuscript.

## Funding

There was no financial support funding for this study.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to N.N.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026