Modeling Earth Systems and Environment (2024) 10:1855–1867
https://doi.org/10.1007/s40808-023-01874-4

ORIGINAL ARTICLE

A hybrid Facebook Prophet-ARIMA framework for forecasting high-frequency temperature data

Mohammed Elseidi[1]

Received: 18 August 2023 / Accepted: 14 September 2023 / Published online: 10 October 2023

Abstract

High-frequency temperature data, such as hourly or daily measurements, show complex seasonal patterns and short-term dynamics that affect various environmental systems and processes. Forecasting temperatures accurately can help environmental scientists manage these systems and processes better. However, most forecasting methods cannot capture both the seasonal and non-seasonal factors of high-frequency temperature data. In this study, we propose a hybrid framework that combines the ARIMA and the Prophet models to improve forecasting accuracy. The Prophet model is a popular method that can model complex seasonality and general trends in time series data, but it often misses short-term dynamics. The ARIMA model can capture short-term dynamics, but it cannot handle complex seasonality. By combining these two models, we can take advantage of their strengths and overcome their weaknesses. The hybrid approach uses Prophet to model the complex seasonality and any general trends in the data, while ARIMA models the short-term dynamics in the residuals of Prophet. We applied our hybrid model to three real-world time series: the average, maximum, and minimum daily temperatures in Ross, USA, from 2017 to 2021. We evaluated our model using graphical measures, statistical tests, and loss function criteria on both the training and the test data. Our results show that our hybrid model outperformed the Prophet model in capturing short-term dynamics and forecasting high-frequency temperature data. This method is particularly useful for time series problems that involve complex seasonality and short-term dynamics.

Keywords Temperature · Forecasting · Facebook Prophet · ARIMA · Complex seasonality · Short-term dynamics

Introduction

Time series analysis is a powerful tool for studying complex systems and forecasting their future behavior. It involves organizing data into a chronological sequence and applying statistical methods to discover patterns and trends. These insights can then be used to make predictions about the system of interest. Time series analysis has applications in various domains, such as energy, business, the economy, health, and the environment. One of the most important applications of time series analysis is understanding the dynamics and trends of ecological and environmental systems, which affect many aspects of biodiversity, conservation, and human well-being. For instance, time series analysis can help identify the drivers and patterns of population fluctuations, the effects of climate change and variability on ecosystems, the relationships between environmental variables and health outcomes, and the challenges and opportunities for sustainable management of natural resources (Zuur et al. 2007). Temperature is also a crucial factor in the weather, which has a significant impact on human life and the planet. Weather monitoring and forecasting are essential for many purposes, such as avoiding flight disruptions (Gultepe et al. 2019), optimizing agricultural productivity (Cogato et al. 2019), and managing energy supply and demand (Liu et al. 2018; Kim et al. 2020). Moreover, climate change and global warming have increased the occurrence of extreme weather events, which require more quantitative approaches to understand and mitigate their effects (Wanishsakpong and Owusu 2020). Therefore, understanding how temperature affects the environment and developing better methods for predicting future temperatures are important goals for time series analysis. Quantitative methods for modeling and predicting temperature have been developed using various approaches, such as statistical time series analysis (Papacharalampous et al. 2018;

✉ Mohammed Elseidi
dr_mohamedelsiedy@mans.edu.eg

[1] Department of Applied Statistics and Insurance, Faculty of Commerce, Mansoura University, Mansoura, Egypt

Tektaş 2010), artificial intelligence (Lee et al. 2020; Azari et al. 2022; Anjali et al. 2019; Molavi Nojumi et al. 2022; Smith et al. 2006) and hybrid methods that combine both (Arslan 2022; Faruk 2010). Each approach has its strengths and weaknesses, which should be taken into account when selecting one for a specific task. Machine learning (ML) techniques have been successfully applied to many domains, such as computer vision, natural language processing, etc. Due to their versatility and performance, ML techniques have also been used for air temperature forecasting. Several ML-based forecasting methods have been explored in the literature, such as artificial neural networks (ANN) and support vector machines (SVM), which are popular methods for analyzing air temperature time series. Cifuentes et al. (2020) provides a comprehensive review of the different ML techniques for air temperature forecasting, their pros and cons, and their potential for further improvement. However, ML techniques face some challenges when dealing with seasonal patterns, which require careful data preprocessing and trend removal to build an effective neural forecaster (Zhang and Qi 2005). Statistical time series methods are another widely used approach for forecasting temperatures. Among them, the Autoregressive Integrated Moving Average (ARIMA) model is a well-known and respected statistical forecasting tool for air temperature (Box George et al. 1976). The ARIMA model follows the Box–Jenkins methodology, which has a remarkable track record for accuracy and efficiency when modeling various time series. However, the ARIMA model requires some assumptions about the linearity and distribution of the time series, which may not always hold. Moreover, the ARIMA model has a limitation when dealing with seasonal time series, which led to the development of the Seasonal ARIMA (SARIMA) model (Chen et al. 2018). However, the SARIMA model is not suitable for modeling complex or high-frequency seasonal fluctuations, such as those found in daily data, which may be affected by multiple seasonal patterns, calendar effects, and non-integer seasonal frequencies (Elseidi 2023). Complex seasonal patterns in time series analysis pose a challenge for forecasting methods. Some of the methods that have been proposed to address this challenge are dynamic harmonic regression (DHR), TBATS, and Prophet. The DHR method models the temperature signal as a combination of a deterministic global trend, cyclic oscillations using polynomial functions and Fourier terms, and stochastic short-term dynamics using ARIMA models (Ye et al. 2013; Hyndman and Athanasopoulos 2018; Young et al. 1999; Gonçalves et al. 2021). A general framework that integrates various techniques for time series forecasting is the TBATS method, which includes Box–Cox transformations, state-space models, exponential smoothing, Fourier terms, and ARIMA models (De Livera et al. 2011). This method can handle time series with nonlinear and changing seasonality, but it may be computationally intensive for long time series. Another novel approach for time series forecasting with complex seasonality is the Prophet method, developed by Facebook (Taylor and Letham 2018). This method uses an additive model that incorporates yearly, monthly, and daily seasonality and holiday effects. The Prophet method is suitable for time series with complex seasonal effects and multiple seasons of data. It is also robust to missing data, trend changes, and outliers. Several studies have applied the Prophet method to temperature data, such as: Toharudin et al. (2023), Thiyagarajan et al. (2020), Asha et al. (2020), Oo and Sabai (2020), Haris et al. (2022), Papacharalampous et al. (2018). However, the Prophet method has some limitations when modeling the short-term dynamics of high-frequency temperature data. According to Elseidi (2023), the Prophet method does not capture the short-term correlations and fluctuations in the residuals well. To address the challenge of forecasting high-frequency data with complex seasonal patterns, changepoints, and short-term dynamics, we propose a hybrid model that leverages the strengths of Facebook Prophet and ARIMA. Our hybrid model consists of two steps: first, we apply the Prophet model to capture the seasonal and trend components of the data and handle any changepoints. Second, we fit an ARIMA model to the residuals of the Prophet model to account for the stochastic dynamics. By combining these two methods, we can achieve more accurate and robust forecasting performance for high-frequency data. The paper proceeds as follows: Forecasting methods are discussed in Sect. 2. Section 3 is dedicated to applying the models to real-world data. This paper concludes in Sect. 4.

Methods

Facebook Prophet

Facebook Prophet is a cutting-edge technique designed to handle complex seasonality, developed by Facebook's data science team (Taylor and Letham 2018). It is based on Harvey's (Harvey and Peters 1990) approach of decomposing the time series into three components: trend, seasonality, and holidays. This method can be expressed as a nonlinear regression model:

$$y_t = g(t) + s(t) + h(t) + \varepsilon_t, \tag{1}$$

The Facebook Prophet method utilizes a piecewise-linear trend ($g(t)$), multiple seasonal patterns ($s(t)$), holiday effects ($h(t)$), and a white noise error term (ε_t) to model time series data. It automatically detects changes in trends by selecting changepoints from data and models both linear and logistic growth curves. Additionally, it has a yearly seasonal component modeled with the Fourier series and a weekly seasonal component modelled with dummy variables, as well as the ability to include user-defined holidays. Facebook Prophet provides two models for growth forecasting: the nonlinear Saturating Growth model and the piecewise linear model. The logistic growth model, in its most basic form, is often used for nonlinear modelling, expressed as follows:

$$g(t) = \frac{C}{1 + \exp(-k(t - m))} \tag{2}$$

where C is the carrying capacity, k is the growth rate, and m is an offset. To account for changes in the trend, the Prophet growth model includes turning points at which the rate of growth can fluctuate. Equation 3 then represents the piecewise logistic growth model:

$$g(t) = \frac{C(t)}{1 + \exp\left(-\left(k + \boldsymbol{a}(t)^T \boldsymbol{\delta}\right)\left(t - \left(m + \boldsymbol{a}(t)^T \gamma\right)\right)\right)} \tag{3}$$

Here, $\boldsymbol{\delta}$ is a vector of rate adjustments, and γ is the correct adjustment at each changepoint j. An efficient and frequently used model is based on a piecewise constant rate of growth to predict situations that do not show signs of saturation growth. This trend-predicting model can be found in Eq. 4:

$$g(t) = \left(k + \boldsymbol{a}(t)^T \boldsymbol{\delta}\right) t \left(m + \boldsymbol{a}(t)^T \gamma\right) \tag{4}$$

For a given period (e.g. 365.25 days for yearly data or 7 days for weekly data), the Fourier series can be used by Facebook Prophet to create a flexible model of periodic effects, as expressed in Eq. 5.

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right)\right) \tag{5}$$

To incorporate a list of holidays into the model, we can assume that the effects of each holiday are independent. Let D_i be the collection of past and future dates for each holiday i. To indicate if time t occurs during holiday i, we introduce an indicator function and assign each holiday a parameter κ_i to represent the change in the prediction. Similar to seasonality, this is implemented by creating a matrix of regressors, as expressed in Eq. (6):

$$Z(t) = [\mathbf{1}(t \in D1), \ldots, \mathbf{1}(t \in DL)] \tag{6}$$

and taking

$$h(t) = Z(t)\boldsymbol{\kappa} \tag{7}$$

ARIMA

ARIMA stands for Autoregressive Integrated Moving Average, which is a class of models for univariate time series analysis and forecasting (Fattah et al. 2018; Torres et al. 2005). An ARIMA model can be expressed as:

$$ARIMA(p, d, q) = AR(p) + I(d) + MA(q) \tag{8}$$

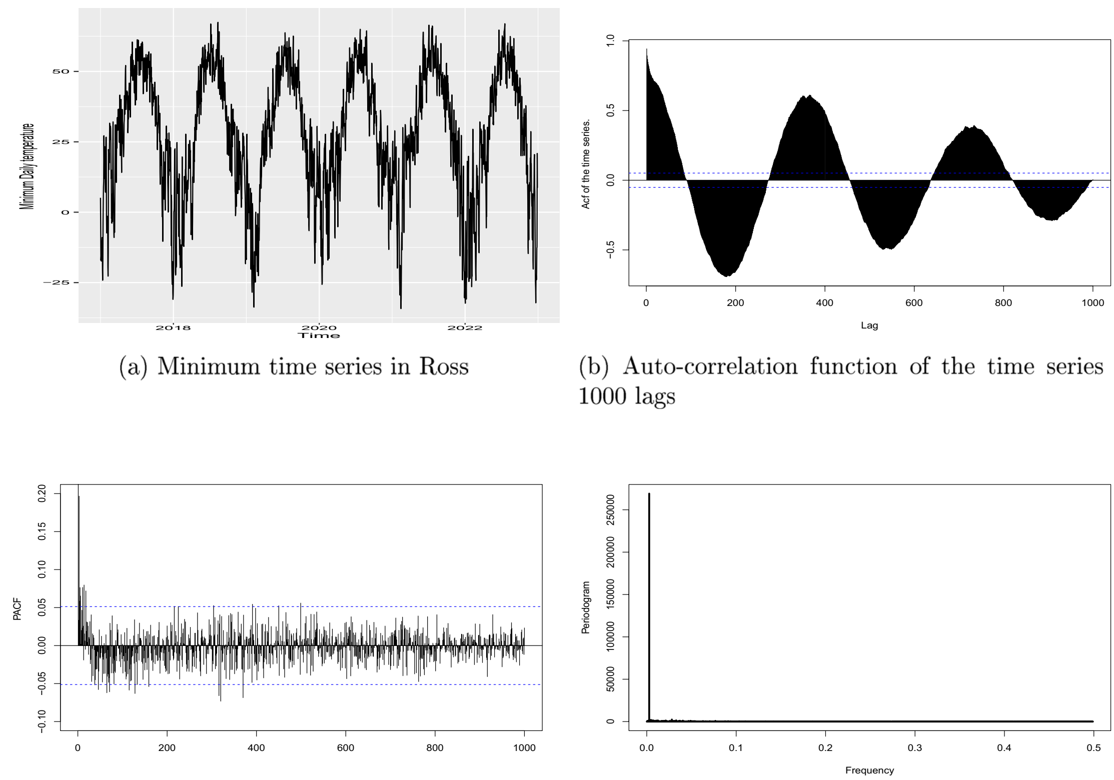
where p is the order of the autoregressive (AR) part, d is the degree of differencing, and q is the order of the moving average (MA) part. The AR part captures the correlation between the current value and the previous values of the time series, while the MA part captures the correlation between the current value and the previous errors. The differencing part transforms the original time series into a stationary one by subtracting the previous values from the current values. The general form of an ARIMA model is:

$$\phi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t \tag{9}$$

where y_t is the observed value at time t, ε_t is the error term at time t, B is the backward shift operator such that $By_t = y_{t-1}$, $\phi(B)$ is a polynomial of degree p such that $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$, and $\theta(B)$ is a polynomial of degree q such that $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$. To fit an ARIMA model to a given time series, we need to determine the values of p, d, and q using various methods such as autocorrelation function (ACF), partial autocorrelation function (PACF), information criteria (AIC, BIC), or cross-validation. Then, we need to estimate the parameters ϕ_i and θ_j using maximum likelihood estimation or other methods. Finally, we need to check the model fit and assumptions using diagnostic plots and tests. ARIMA models are widely used in various fields such as economics, finance, engineering, and medicine. They can capture various patterns and trends in time series data and provide accurate forecasts for future values.

The proposed method (Prophet-ARIMA)

Hybridization, which involves combining the strengths of two or more distinct models, is a popular approach for data analysis, particularly for time series data (Kong et al. 2021; Zhang 2003; Pan et al. 2009). This study focused on using such a technique to enhance the modeling and prediction accuracy of the Facebook Prophet model for high-frequency temperature data. By combining the advantages of multiple models, hybridization can improve prediction accuracy beyond what is achievable with individual models. The proposed model combines the strengths of the Prophet model



(a) Minimum time series in Ross

(b) Auto-correlation function of the time series for 1000 lags

(c) Partial auto-correlation function of the time series for 1000 lags

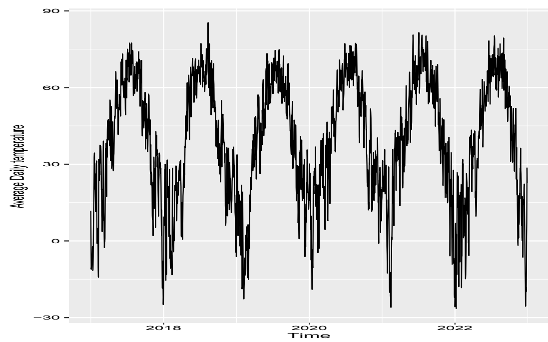
(d) Periodogram

Fig. 1 Minimum daily temperature time series and its characteristics for Ross Station in the USA. **a** The time series plot. **b** The autocorrelation function of 1000 lags. **c** The partial autocorrelation function of 1000 lags. **d** The periodogram of the time series

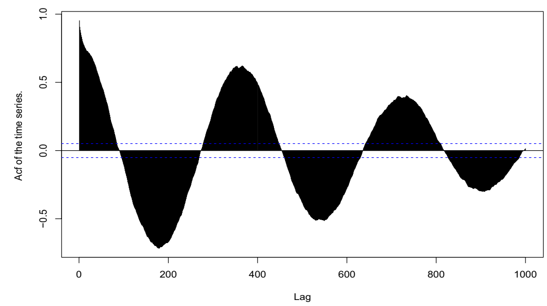
and the well-known ARIMA model. The Prophet model is beneficial for modeling complex seasonality, general trends, change points, and holidays, while the ARIMA model can model short-term random dynamics. By merging the two models, a comprehensive prediction model can be created that can more accurately forecast more than one problem that each model cannot address individually. The idea of merging different models has been proven successful in various studies, such as the DHR (Ye et al. 2013) and TBATS (Ye et al. 2013) models. Furthermore, Elseidi (2023) showed the usefulness of using the ARIMA model within the DHR and TBATS techniques for modeling the short-term behavior of temperature data and the drawbacks of the prophet method in dealing with these dynamics.

The steps for implementing the proposed hybrid approach combining the Prophet and ARIMA models are as follows:

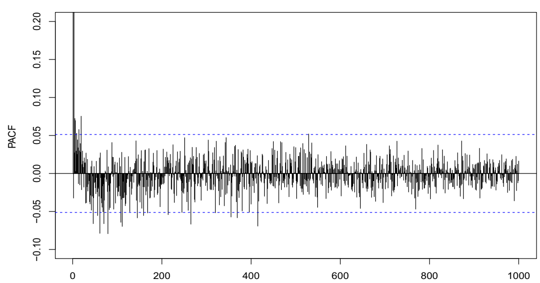
1. Prepare and preprocess the data for analysis. This includes ensuring that the temperature data is in a time series format with a date column, a temperature data column, and other relevant columns.
2. Conduct Prophet Model Analysis: Fit the Prophet algorithm to the data to identify trends, seasonality, and changepoints.
3. Evaluate the residuals of the Prophet model to assess the accuracy of the model's fit. Residuals should be normally distributed and show no clear patterns. If the residuals appear normally distributed and there are no clear patterns, then the Prophet model has provided a good fit for the data. If the residuals do not appear to be normally distributed or there are clear patterns in the residuals, then further optimization is necessary (Steps 4 and 5).
4. Use the residuals from the Prophet model as input to an ARIMA model to handle the remaining short-term dynamics in the residuals.
5. To obtain the final forecast, combine the forecasted values from the Prophet model with the forecasted values from the Autoregressive Integrated Moving Average (ARIMA) model.



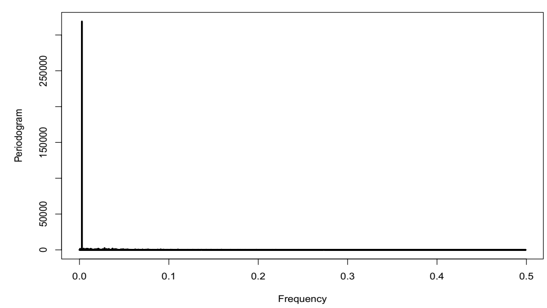
(a) Average time series in Ross



(b) Auto-correlation function of the time series for 1000 lags



(c) Partial auto-correlation function of the time series for 1000 lags



(d) Periodogram

Fig. 2 Average daily temperature time series and its characteristics for Ross Station in the USA. **a** The time series plot. **b** The autocorrelation function of 1000 lags. **c** The partial autocorrelation function of 1000 lags. **d** The periodogram of the time series

Data analysis

Data

We used three real-world time series to compare the performance of the proposed hybrid technique and the Prophet method: the average, maximum, and minimum daily temperatures in Ross, USA, from 2017 to 2021. The data source was the North Dakota Agricultural Weather Network (NDAWN). We split the data into two parts: the training data set (2017–2020) and the test data set (2021). We chose the best models based on their accuracy in modeling and forecasting.

Descriptive analysis

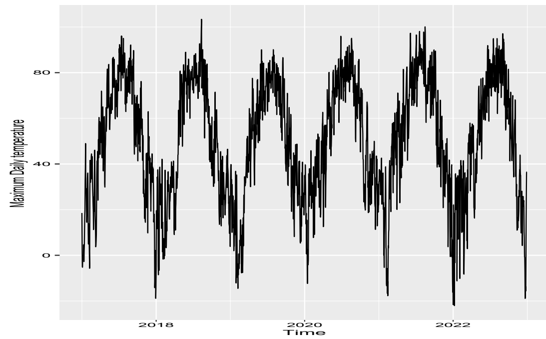
Descriptive statistics and graphical representations were used to analyze the data accurately. The minimum, average, and maximum daily time series are summarized in Table 1. Figures 1, 2, and 3 show the visual illustrations of these time series. Each figure has four panels: the daily time series (upper left), the autocorrelation function of 1000 lags (upper right), the partial autocorrelation function of 1000 lags (lower left), and the periodogram of the time series (lower right). The plots of the time series and their autocorrelation functions indicate that an annual seasonal pattern is the most dominant in each time series. The fundamental frequencies and amplitudes that affect the time series can be identified by spectral analysis and reference data attributes (Caporin and Elseidi 2023). The periodograms of all three-time series confirm this by having only one significant peak corresponding to the annual pattern.

Training data

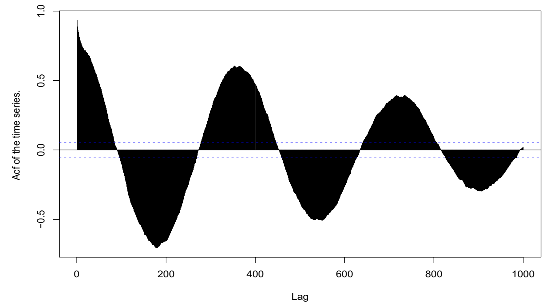
Evaluation criteria of the modeling accuracy

The modeling performance of the competing forecasting models was evaluated using three distinct approaches: visual inspection, statistical tests, and error metrics.

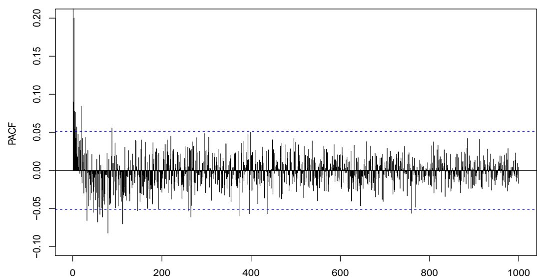
- For visual inspection, we plotted the residuals, their autocorrelation function (ACF), and their histogram



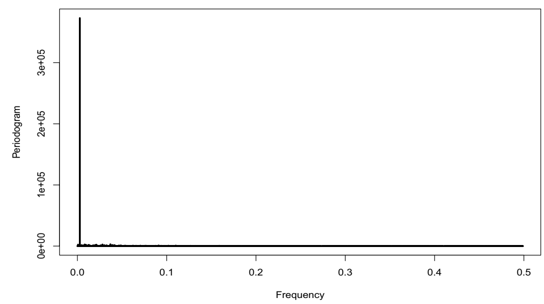
(a) Maximum time series in Ross



(b) Auto-correlation function of the time series for 1000 lags



(c) Partial auto-correlation function of the time series for 1000 lags



(d) Periodogram

Fig. 3 Maximum daily temperature time series and its characteristics for Ross Station in the USA. **a** The time series plot. **b** The autocorrelation function of 1000 lags. **c** The partial autocorrelation function of 1000 lags. **d** The periodogram of the time series

Table 1 Descriptive statistics of the Minimum, Average and Maximum air temperature time series

Measures	Minimum	Average	Maximum
Mean	27.70	38.90	50.11
Median	29.57	40.15	51.22
S.D	22.46	24.00	26.15
Skewness	– 0.50	– 0.41	– 0.32
Kurtosis	– 0.51	– 0.68	– 0.79
Minimum	– 34.24	– 26.45	– 22.04
Maximum	67.44	85.40	103.35

with a normal curve. This helped us check the distribution and randomness of the errors.

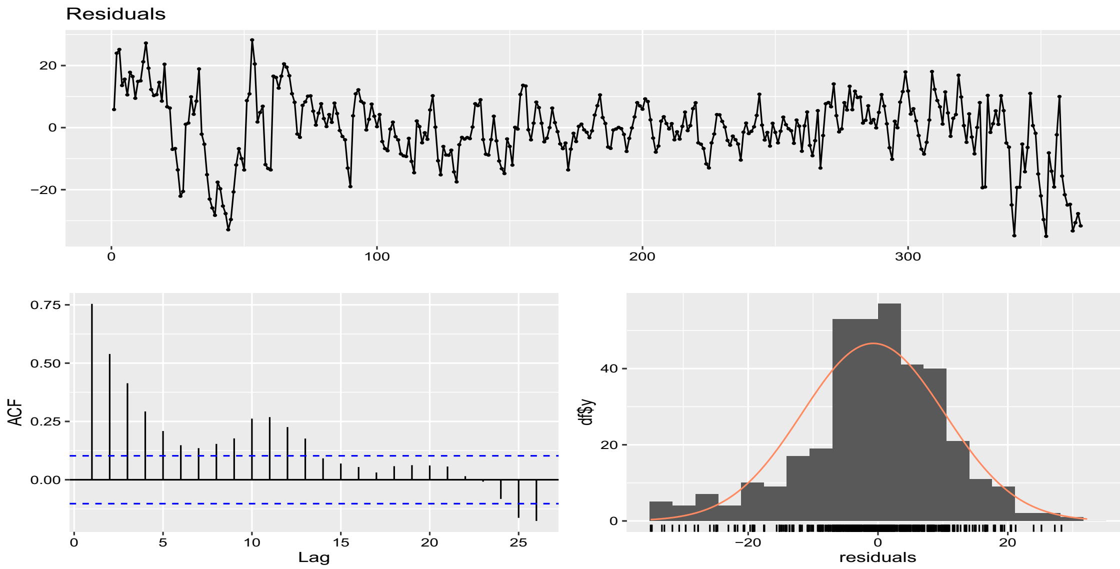
- For statistical tests, we used the Ljung–Box test to see if there was any autocorrelation in the residuals. Autocorrelation can affect the accuracy and interpretation of forecasts. The lag order for the test should be twice the annual seasonality of the series.
- For error metrics, we calculated the root mean squared error (RMSE) for each model. RMSE measures how close the predictions are to the actual values by taking the square root of the average squared difference between them. It is a common way to evaluate the performance of statistical models. The formula for RMSE is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(P_i - A_i\right)^2} \tag{10}$$

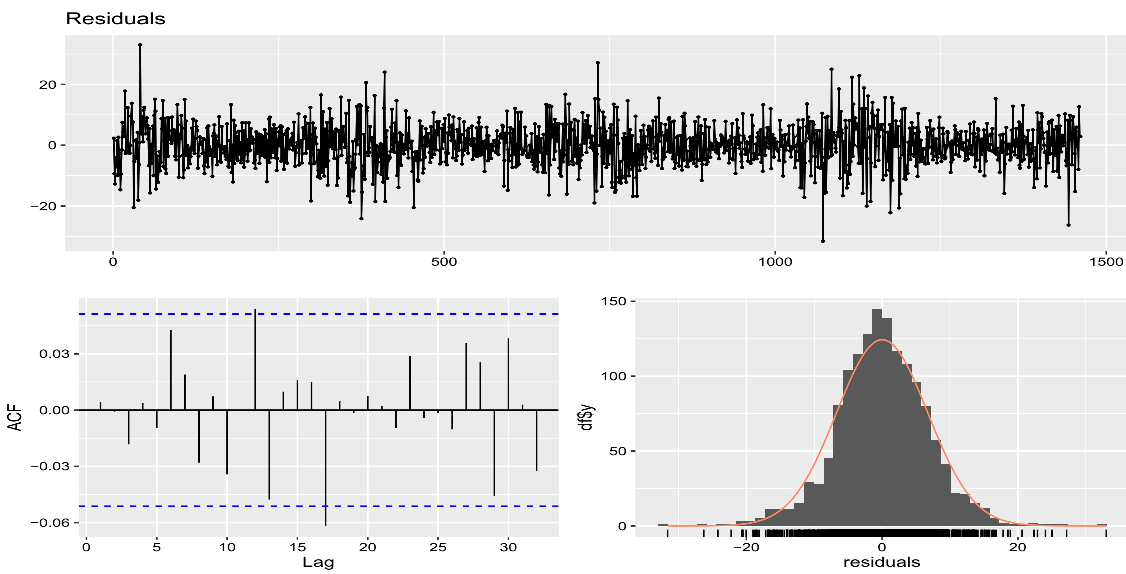
where n is the number of observations, P is the predicted value, and A is the actual value.

Modeling results

We compared the forecasting performance of the Prophet model and our proposed hybrid model using three methods: visual inspection, statistical tests, and error metrics. The results of the visual inspection are shown in Figs. 4, 5, and 6 for the minimum, average, and maximum temperature series, respectively. Each figure has two subplots: the upper one for the Prophet model and the lower one for our hybrid model. The results show that our hybrid model fits the data better than the Prophet model, as it captures both the seasonal



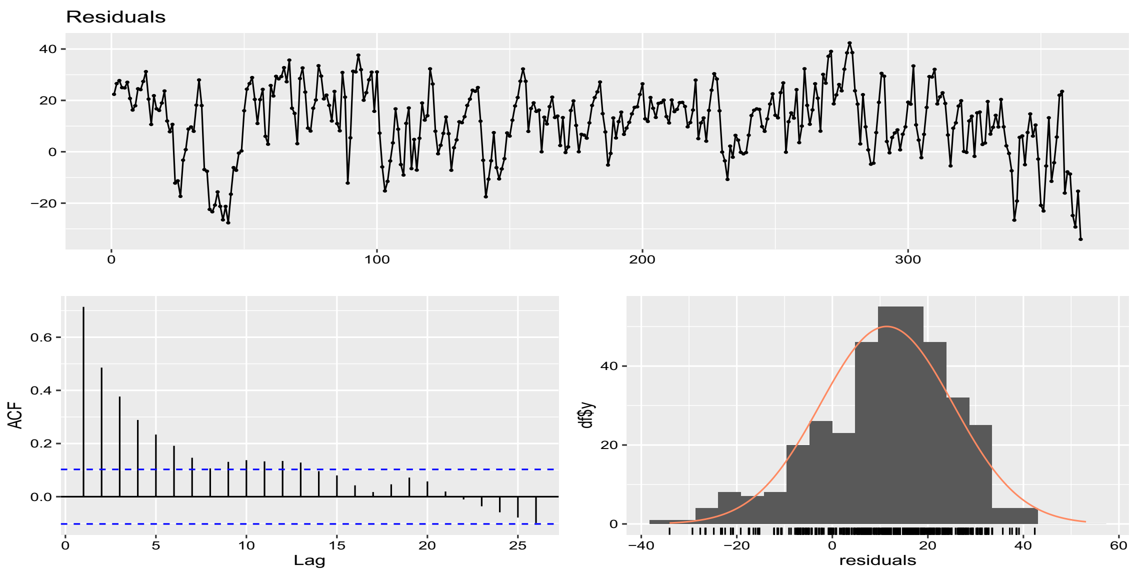
(a) Residuals diagnostics of Prophet



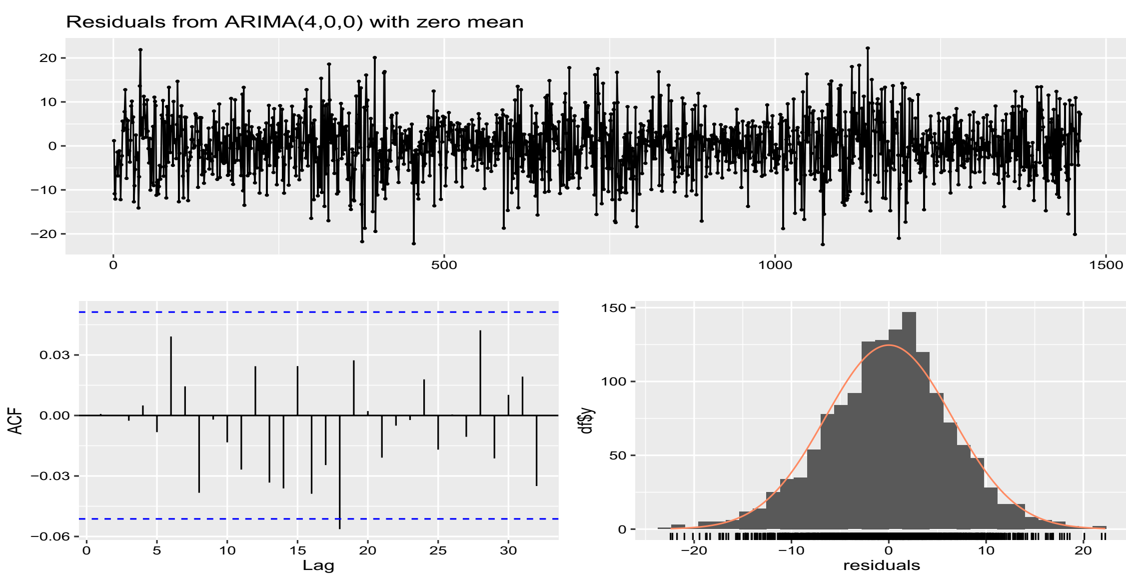
(b) Residuals diagnostics of Prophet-Arima $(5, 0, 5)$

Fig. 4 The graphical accuracy measures of the training data of Minimum daily temperature, we present the residuals time series, ACF of the residuals, and the histogram of the residuals in comparison with the normal curve

patterns and the short-term fluctuations more accurately. Moreover, the residuals of our hybrid model are closer to a normal distribution and have no significant autocorrelation, unlike the Prophet model's residuals. This indicates that our hybrid model does not leave any systematic errors in the data. The results of the statistical tests are given in Table 2, which shows the p-values of the Ljung–Box test for both models. The Ljung–Box test is used to test whether there is any autocorrelation in the residuals up to a certain lag order. We used a lag order of 730, which is twice the annual seasonality of the series. The results show that none of the p-values for our hybrid model are significant at the 0.05 level, meaning that we cannot reject the null hypothesis that there is no autocorrelation in the residuals. On the contrary, all of the p-values for the Prophet model are significant at the 0.05 level, meaning that we can reject the null hypothesis that there is no autocorrelation in the residuals. This



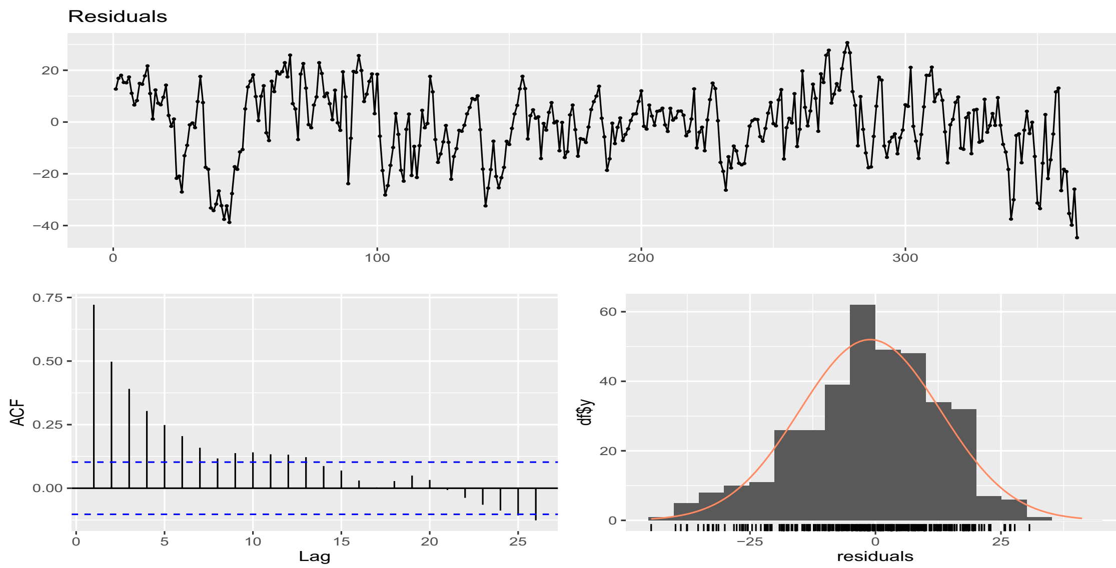
(a) Residuals diagnostics of Prophet



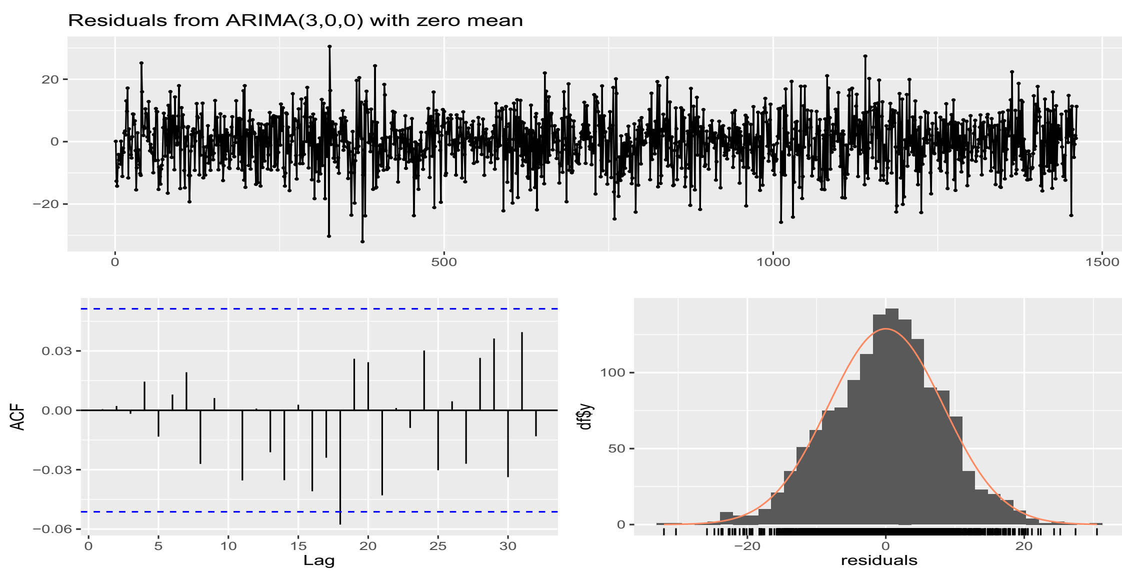
(b) Residuals diagnostics of Prophet-ARIMA$(3, 0, 4)$

Fig. 5 The graphical accuracy measures of the training data of Average daily temperature, we present the residuals time series, ACF of the residuals, and the histogram of the residuals in comparison with the normal curve

confirms that our hybrid model handles short-term dynamics better than the Prophet model. The results of the error metrics are presented in Table 3, which shows the root mean squared error (RMSE) values for both models. The RMSE is a measure of how close the predictions are to the actual values by taking the square root of the average squared difference between them. The results show that our hybrid model has lower RMSE values than the Prophet model for all the series, indicating that our hybrid model has smaller prediction errors. This demonstrates that our hybrid model is more accurate than the Prophet model in modeling the temperature series. The ARIMA model parameters for each series are as follows: (2,0,3) for the minimum series, (4,0,0) for the average series, and (3,0,0) for the maximum series. Our hybrid model leverages the advantages of both the Prophet and the ARIMA models to produce more accurate forecasts. The



(a) Residuals diagnostics of Prophet



(b) Residuals diagnostics of Prophet-ARIMA$(1, 0, 2)$

Fig. 6 The graphical accuracy measures of the training data of Maximum daily temperature, we present the residuals time series, ACF of the residuals, and the histogram of the residuals in comparison with the normal curve

Table 2 P-value of the Ljung–Box test in the training data-set

Methods	LB (Minimum)	LB (Average)	LB (Maximum)
Prophet	0.00	0.00	0.00
Prophet-ARIMA	**0.54**	**0.30**	**0.28**

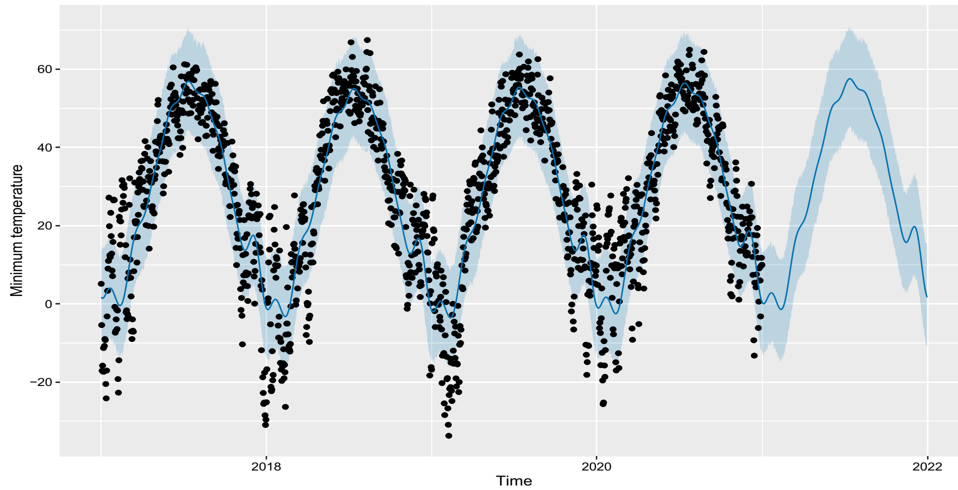
The bold p-values indicate that they are not statistically different from zero

Table 3 RMSE in the training data-set

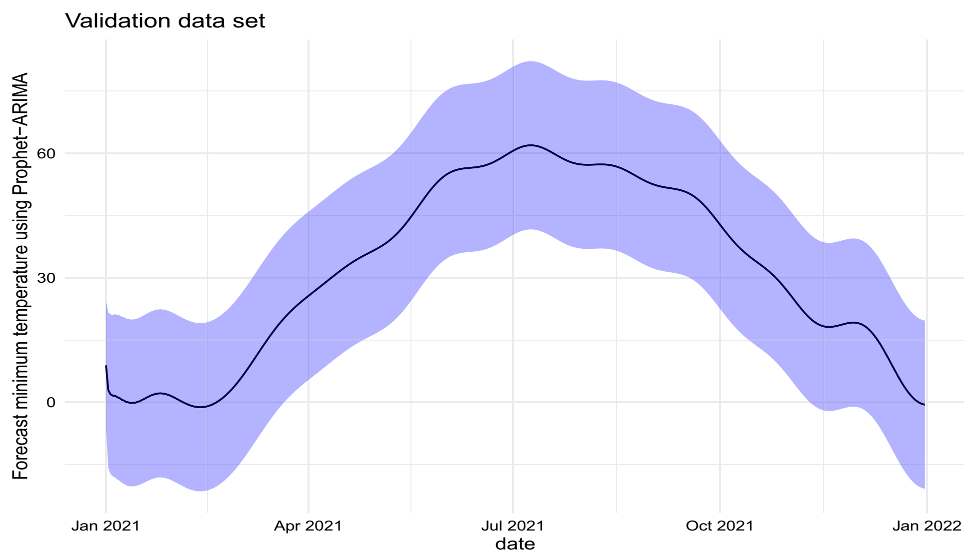
Methods	RMSE (Minimum)	RMSE (Average)	RMSE (Maximum)
Prophet	9.96	10.14	11.46
Prophet-ARIMA	**6.73**	**6.52**	**8.32**

The bold values in the tables highlight the best methods for each criterion

Fig. 7 The predictions of the Prophet and proposed models, respectively, for minimum daily temperatures



(a) Predictions of the prophet



(b) Predictions of the Proposed method

Prophet model handles complex seasonality, general trends, change points, and holidays, while the ARIMA model captures short-term random dynamics. By combining the two models, we can create a comprehensive prediction model that can address multiple problems that each model alone cannot solve.

Test data

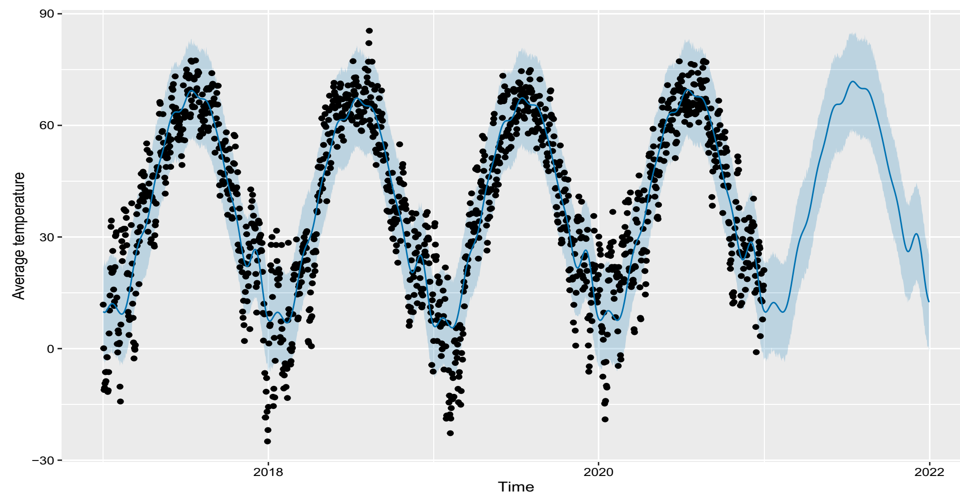
Evaluation criteria of the prediction accuracy

To compare the effectiveness of different forecasting models, we use the root mean square error (RMSE) as a loss function criterion. The RMSE is calculated by Eq. 10 and it provides a simple and reliable way to measure the accuracy of the predictions.

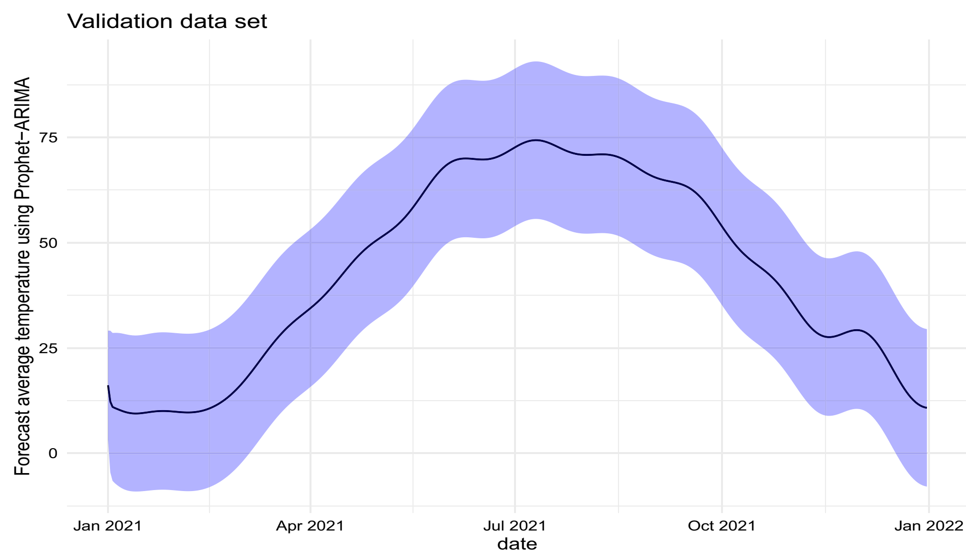
Forecasting results

The forecasting performance of the Prophet and the proposed hybrid models was evaluated using test data. Table 4 shows the root mean square error (RMSE) values for both models for each time series. The proposed hybrid model achieved lower RMSE values than the Prophet model for all time series, indicating higher accuracy. Figures 7, 8 and 9 compare the forecasts of the two models for daily minimum, average, and maximum temperatures, respectively. The figures clearly illustrate that the proposed hybrid model can capture the short-term dynamics and complex seasonality of the temperature data better than the Prophet model. The main contribution of the proposed hybrid model is that it can effectively model the non-seasonal elements in the residuals

Fig. 8 The predictions of the Prophet and proposed models, respectively, for average daily temperatures



(a) Predictions of the Prophet



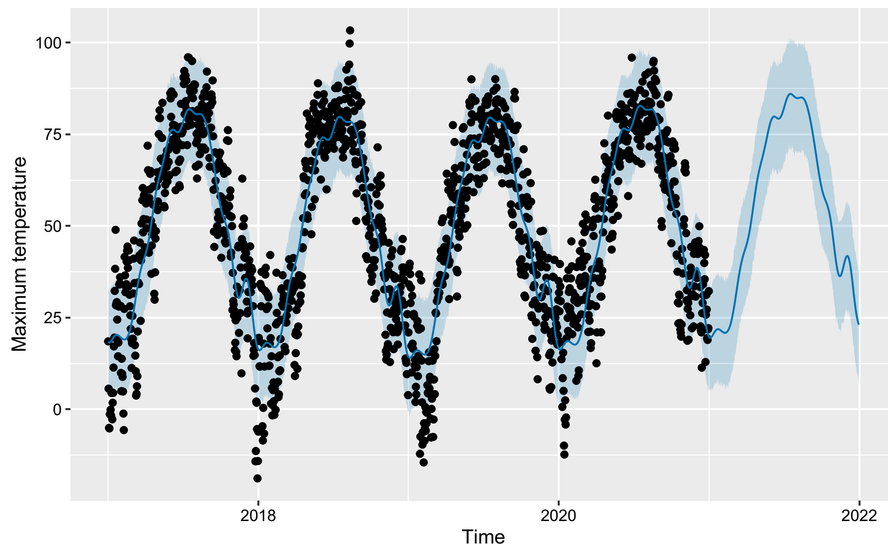
(b) Predictions of the of Proposed method

of the Prophet model using ARIMA, which improves forecasting accuracy and reduces error. The results of the residual diagnostics, which are presented in the modeling section, confirm this finding. Therefore, the proposed hybrid model is a powerful and robust method for forecasting high-frequency temperature data.

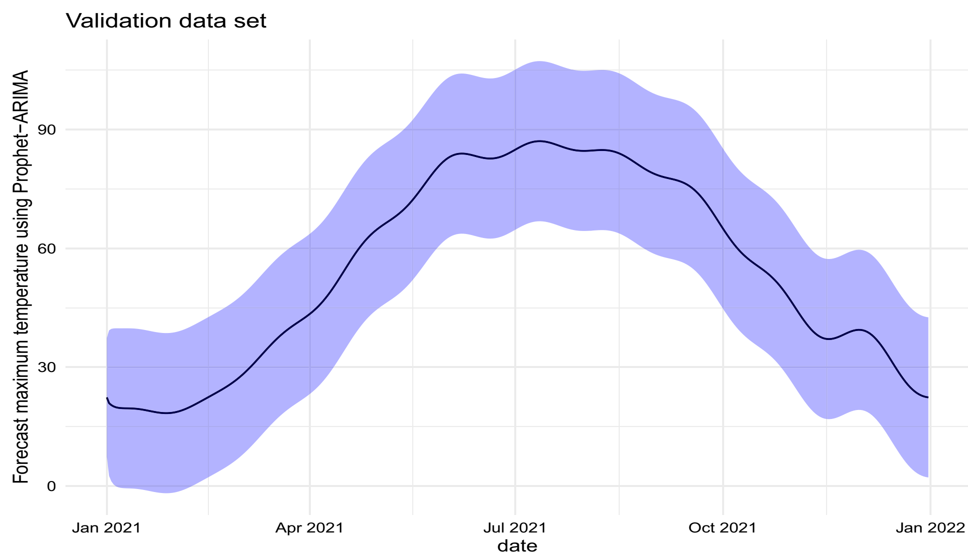
Conclusion

This paper presents a novel hybrid framework that combines the ARIMA and the Prophet models to forecast high-frequency temperature data. We tested our hybrid model on three real-world time series: the daily minimum, average, and maximum temperatures in Ross, USA, from 2017 to 2021. We evaluated the performance of our hybrid and the Prophet models using graphical measures, statistical tests, and error metrics on both the training and test data. Our results demonstrate that our hybrid model achieves higher modeling and forecasting accuracy than the Prophet model. Our hybrid model can handle both the complex seasonality and the short-term dynamics of the temperature data, while the Prophet model can only capture the complex seasonality. Our hybrid model integrates the strengths of the ARIMA and Prophet models to produce more accurate forecasts. The ARIMA model improves forecasting accuracy and reduces error by modeling the short-term random dynamics in the residuals of the Prophet model. The Prophet model captures the data's

Fig. 9 The predictions of the Prophet and proposed models, respectively, for Maximum daily temperatures



(a) Predictions of the Prophet



(b) Predictions of the Proposed method

Table 4 RMSE in the test data-set

Methods	RMSE (Minimum)	RMSE (Average)	RMSE (Maximum)
Prophet	10.99	11.75	14.09
Prophet-ARIMA	**10.92**	**11.65**	**14.02**

The bold values in the tables highlight the best methods for each criterion

complex seasonality, general trends, change points, and holidays. By combining the two models, we can create a comprehensive prediction model that can solve multiple problems that each model alone cannot address. Our hybrid model is a powerful and robust method for forecasting high-frequency temperature data.

Data availability The data that support the findings of this study are available on request from the corresponding author.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Anjali T, Chandini K, Anoop K, Lajish V (2019) Temperature prediction using machine learning approaches. In 2019 2nd International Conference on intelligent computing, instrumentation and control technologies (ICICICT), volume 1, pp 1264–1268. IEEE

Arslan S (2022) A hybrid forecasting model using lstm and prophet for energy consumption with decomposition of time series data. PeerJ Comput Sci 8:e1001

Asha J, Rishidas S, SanthoshKumar S, Reena P (2020) Analysis of temperature prediction using random forest and Facebook prophet algorithms. In: International Conference on innovative data communication technologies and application, Springer, pp 432–439

Azari B, Hassan K, Pierce J, Ebrahimi S (2022) Evaluation of machine learning methods application in temperature prediction. Environ Eng 8:1–12

Box George E, Jenkins Gwilym M, Reinsel Gregory C, Ljung Greta M (1976) Time series analysis: forecasting and control. Holden Bay, San Francisco

Caporin M, Elseidi M (2023) Quantile regression-based seasonal adjustment. Int J Comput Econ Economet 13(3):270–304

Chen P, Niu A, Liu D, Jiang W, Ma B (2018) Time series forecasting of temperatures using Sarima: an example from Nanjing. IOP Conf Ser Mater Sci Eng 394:052024

Cifuentes J, Marulanda G, Bello A, Reneses J (2020) Air temperature forecasting using machine learning techniques: a review. Energies 13(16):4215

Cogato A, Meggio F, De Antoni Migliorati M, Marinell F (2019) Extreme weather events in agriculture: a systematic review. Sustainability 11(9):2547

De Livera AM, Hyndman RJ, Snyder RD (2011) Forecasting time series with complex seasonal patterns using exponential smoothing. J Am Stat Assoc 106(496):1513–1527

Elseidi M (2023) Forecasting temperature data with complex seasonality using time series methods. Model Earth Syst Environ 9(2):2553–2567

Faruk DÖ (2010) A hybrid neural network and Arima model for water quality time series prediction. Eng Appl Artif Intell 23(4):586–594

Fattah J, Ezzine L, Aman Z, El Moussami H, Lachhab A (2018) Forecasting of demand using Arima model. Int J Eng Bus Manag 10:1847979018808673

Gonçalves AM, Costa C, Costa M, Lopes SO, Pereira R (2021) Temperature time series forecasting in the optimal challenges in irrigation (TO CHAIR). In: Gaspar-Cunha A, Periaux J, Giannakoglou KC, Gauger NR, Quagliarella D, Greiner D (eds) Advances in evolutionary and deterministic methods for design, optimization and control in engineering and sciences. Computational methods in applied sciences, vol 55. Springer, Cham, pp 423–435. https://doi.org/10.1007/978-3-030-57422-2_27

Gultepe I, Sharman R, Williams PD, Zhou B, Ellrod G, Minnis P, Trier S, Griffin S, Yum S, Gharabaghi B et al (2019) A review of high impact weather for aviation meteorology. Pure Appl Geophys 176(5):1869–1921

Haris MD, Adytia D, Ramadhan AW (2022) Air temperature forecasting with long short-term memory and prophet: a case study of Jakarta, Indonesia. In: 2022 International Conference on data science and its applications (ICoDSA), pp 251–256. IEEE

Harvey AC, Peters S (1990) Estimation procedures for structural time series models. J Forecast 9(2):89–108

Hyndman RJ, Athanasopoulos G (2018) Forecasting: principles and practice. OTexts

Kim MK, Kim Y-S, Srebric J (2020) Predictions of electricity consumption in a campus building using occupant rates and weather elements with sensitivity analysis: artificial neural network vs. linear regression. Sustain Cities Soc 62:102385

Kong YH, Lim KY, Chin WY (2021) Time series forecasting using a hybrid prophet and long short-term memory model. In: International Conference on soft computing in data science, pp 183–196. Springer

Lee S, Lee Y-S, Son Y (2020) Forecasting daily temperatures with different time interval data using deep neural networks. Appl Sci 10(5):1609

Liu Y, Roberts MC, Sioshansi R (2018) A vector autoregression weather model for electricity supply and demand modeling. J Mod Power Syst Clean Energy 6(4):763–776

Molavi Nojumi M, Huang Y, Hashemian L, Bayat A (2022) Application of machine learning for temperature prediction in a test road in Alberta. Int J Pavement Res Technol 15(2):303–319

Oo ZZ, Sabai P (2020) Time series prediction based on Facebook prophet: a case study, temperature forecasting in Myintkyina. Int J Appl Math Electron Comput 8(4):263–267

Pan F, Zhang H, Xia M (2009) A hybrid time-series forecasting model using extreme learning machines. In: 2009 Second International Conference on intelligent computation technology and automation, volume 1, pp 933–936. IEEE

Papacharalampous G, Tyralis H, Koutsoyiannis D (2018) Predictability of monthly temperature and precipitation using automatic time series forecasting methods. Acta Geophys 66(4):807–831

Smith BA, McClendon RW, Hoogenboom G (2006) Improving air temperature prediction with artificial neural networks. Int J Comput Intell 3(3):179–186

Taylor SJ, Letham B (2018) Forecasting at scale. Am Stat 72(1):37–45

Tektaş M (2010) Weather forecasting using Anfis and Arima models. Environ Res Eng Manag 51(1):5–10

Thiyagarajan K, Kodagoda S, Ulapane N, Prasad M (2020) A temporal forecasting driven approach using Facebook's prophet method for anomaly detection in sewer air temperature sensor system. In: 2020 15th IEEE Conference on industrial electronics and applications (ICIEA), pp 25–30. IEEE

Toharudin T, Pontoh RS, Caraka RE, Zahroh S, Lee Y, Chen RC (2023) Employing long short-term memory and Facebook prophet model in air temperature forecasting. Commun Stat-Simul Comput 52(2):279–290

Torres JL, Garcia A, De Blas M, De Francisco A (2005) Forecast of hourly average wind speed with Arma models in Navarre (Spain). Sol Energy 79(1):65–77

Wanishsakpong W, Owusu BE (2020) Optimal time series model for forecasting monthly temperature in the southwestern region of Thailand. Model Earth Syst Environ 6(1):525–532

Ye L, Yang G, Van Ranst E, Tang H (2013) Time-series modeling and prediction of global monthly absolute temperature for environmental decision making. Adv Atmos Sci 30(2):382–396

Young PC, Pedregal DJ, Tych W (1999) Dynamic harmonic regression. J Forecast 18(6):369–394

Zhang GP (2003) Time series forecasting using a hybrid Arima and neural network model. Neurocomputing 50:159–175

Zhang GP, Qi M (2005) Neural network forecasting for seasonal and trend time series. Eur J Oper Res 160(2):501–514

Zuur AF, Ieno EN, Smith GM et al (2007) Analysing ecological data, vol 680. Springer