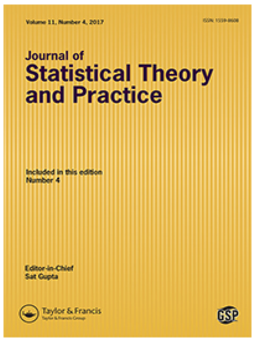

Taylor & Francis
Taylor & Francis Group

Journal of Statistical Theory and Practice

ISSN: 1559-8608 (Print) 1559-8616 (Online) Journal homepage: http://www.tandfonline.com/loi/ujsp20

Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq

Tara Ahmed Chawsheen & Mark Broom

JOURNAL OF STATISTICAL THEORY AND PRACTICE
2017, VOL. 11, NO. 4, 604–633
http://dx.doi.org/10.1080/15598608.2017.1292484

Seasonal time-series modeling and forecasting of monthly mean temperature for decision making in the Kurdistan Region of Iraq

Tara Ahmed Chawsheen and Mark Broom

Department of Mathematics, City, University of London, London, United Kingdom

ABSTRACT

A generalized structural time-series modeling framework was used to analyze the monthly records of mean temperature, one of the most important environmental parameters, using classical stochastic processes. In this article we are using the SARIMA Box–Jenkins model and obtain a medium-term (10 years) forecast of the mean temperature in Erbil. A prediction of the monthly mean temperature during the past 287 months ($\simeq$24 years) using the SARIMA$(0,1,2)(0,1,1)_{12}$ model predicts that the average temperature in the governorate of Erbil, Iraq, will be stable for the next 10 years. The evaluation of prediction accuracy shows that our model performs equally well when applying it to different periods of time for which data is available. The method used here could easily be applied by the decision makers responsible for providing water and electricity in the Kurdistan Region.

ARTICLE HISTORY
Received 8 November 2016
Accepted 27 January 2017

KEYWORDS
Climate change; forecasting; Fourier method; Kurdistan Region of Iraq; SARIMA model; stochastic process

AMS SUBJECT CLASSIFICATION
62M10

1. Introduction

In 2009, the United Nations Environment Programme (UNEP) reported a sudden surge in global temperature of approximately 0.5°C. However, to this day there is no consensus among scientists on how to gauge the magnitude of climate change and its effects on a regional level. It was widely assumed that its effects on surface and ocean temperature would appear gradually and slowly due to their weak and delayed response to the greenhouse gas levels (i.e., carbon dioxide, water vapor, ozone, methane, various nitrous oxides, and other industrial gases). Since the emission of these gases is coupled to world population growth and technological advances, it is hard to predict when (if at all) this phenomenon will reach a stable equilibrium again (UNEP 2009). Climate variation over any region has become a topic of interest all over the world, due to its immediate effect on the daily lives of humans (Ghahraman 2006). The Kurdistan Region of Iraq is affected by changes in climate conditions in the fields of agriculture, architecture, road construction etc. Situated in the north of Iraq, Kurdistan has been facing the consequences of severe drought in the Fertile Crescent since the 1970s (El-Kadi 2001).

In terms of the Koppen classification of climes, Kurdistan exhibits an arid to semi-arid climate. It is hot and dry in summer and cold and wet in winter, with short spring and autumn seasons (Turkes 1996c). In winter, the weather is shaped by Mediterranean

CONTACT Mark Broom Mark.Broom@city.ac.uk Department of Mathematics, City, University of London, Northampton Square, London EC1V 0HB, United Kingdom.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/ujsp.

cyclones passing Iraqi Kurdistan on their way to the northeast, and by Arabian Sea cyclones moving northward across the Persian Gulf, both of which typically carry a great amount of moisture leading to a large amount of precipitation. Other sources of rain and snow include occasional European winter low-pressure systems moving eastward to the southeast part of Turkey and the adjacent Kurdish territories (Turkes 1999). In summer, the region falls under the influence of subtropical high pressure belts and Mediterranean anticyclones, which carry sand and dust to the region. Temperatures may reach up to 50°C in summer and drop as low as −10°C in winter (Keller and Blodgett 2006). The statistical analysis of the climatological records contributes to the understanding of the underlying causes of drought and consequently facilitates taking measures to prepare for (if not prevent altogether) natural disasters such as crop failures or flooding or dust storms.

As pointed out by Al-Kubaisi and Gardi (2012), who compared mean air temperature, the number of dust storms, and precipitation figures over a period of 11 years from 1998 to 2009, there is an unmistakable interrelation among the three. While the mean temperature fell from 22.7°C in 1998 to 22.5°C in 2000 and 21.3°C in 2009, annual rainfall at first decreased from 310.3 mm to 268.5 mm, finally climbing back to 295.6 mm, and the number of dust storms went from 63 to 94 and back down to 64, respectively (Al-Kubaisi and Gardi 2012).

Kurdistan has been going through a period of drought over the past few years; as a result, many of the inhabitants of rural areas left their villages and migrated to the cities where the water scarcity issue is becoming more pressing due to increasing population (Zakaria et al. 2013). The situation is complicated further by the political circumstances, such as mass immigration from southern Iraq and Syria. Another study by Khalid (2014) and Eklund and Pilesjö (2012) showed that in the last two decades, Erbil, Iraq, had expanded significantly both in population size and in area, a process in which much of the natural soil was removed. As a result, the microclimate of this area has changed and now exhibits the urban heat island (UHI) phenomenon. The temperature records make this evident: While the average maximum temperature in Erbil city was 14.34°C in 1975–1990, it had reached 14.74°C in 1985–2000 and 15.70°C in 1995–2012 (see Saeed and Abas 2012).

In order to accurately determine the need for electricity and water and plan their provision accordingly, a precise quantitative understanding and monitoring of various climate parameters, such as temperature, precipitation, humidity, and wind, is indispensable. This study aims to examine the time evolution of the mean temperature in Erbil between January 1992 and November 2015 by separating seasonal effects from long-term trends and, using the seasonal autoregressive integrated moving average (SARIMA) method, create a model that accurately predicts the monthly mean temperature until December 2025, thus enabling the political authorities to make informed decisions on climate related matters.

To model a time-series event as a function of its past values, analysts identify patterns in past values and project them into the future. In particular, the Box–Jenkins methodology could be applied to any environmental parameters, for example, wind speed, precipitation, humidity, and evaporation. Box–Jenkins methodology has been used by many researchers starting with the studies by Intergovernmental Panel on Climate Change (IPCC 2013), Lee and Ko (2011), Ghil et al. (2002), Mann (2008), and Mann and Park

(1996), who predicted the variation in temperature in different places in the world by using different statistical approaches, including bivariate time-series models, and time-series smoothing both in the univariate and the multivariate setting. The most significant feature of the univariate time-series model is its ability to determine the trend and random residuals about the time-series data by using an autoregressive integrated moving average (ARIMA) (Romilly 2005).

Autoregressive integrated moving average (ARIMA) and seasonal integrated moving average (SARIMA) techniques have been broadly applied to forecast how variables change over time. These techniques typically use (seasonal) autoregressive terms and seasonal moving average terms to forecast the changes of time series. As generally reported, these forecasting techniques regard both the preceding values of a variable and the corresponding error terms as essential information in forecasting future values. Given a large time-series data set, ARIMA and SARIMA methods show high forecast accuracy. Forecasting analysis in a variety of fields such as air temperature, electricity demand, wheat prices, inflation, unemployment, reliability, and fishery landings has demonstrated the validity of ARIMA and/or SARIMA models (Choi, Roberts, and Lee 2015). In other instances, the deterministic stochastic combined technique has been successfully used by Ye et al. (2013) to predict global temperature as recorded by the National Climate Data Centre (NCDC), and a time-series approach has been implemented by Mraoua and Bari (2007) to accurately model weather derivative pricing in Morocco. SARIMA itself has been applied to local temperature forecasting in the Ashanti region of northern Ghana by Asamoah-Boaheng (2014).

Regional changes were observed in the mean temperatures in Turkey from 1950 to 1994 over the course of a study conducted by Can and Atimtay (2002) using time-series analysis of mean temperature data. Their study established a statistically significant cooling trend at 21 stations, as well as a warming trend at one station and no trend at 36 stations. Hansen et al. (2006) in their study focused on global temperature change, while Rahmstorf et al. (2007) compared recent climate observations to projections. Zakaria et al. (2012) applied ARIMA models for weekly rainfall data for the period 1990–2011 from four rainfall stations in the northwest of Iraq: Sinjar, Mosul, Rabeaa, and Talafar. Four SARIMA models were developed for these stations: $(3,0,2)(2,1,1)_{30}$, $(1,0,1)(1,1,3)_{30}$, $(1,1,2)(3,0,1)_{30}$, and $(1,1,1)(0,0,1)_{30}$, respectively.

In this article, Box–Jenkins methodology and in particular the method of the seasonal autoregressive integrated moving average (SARIMA) model are applied to temperature data from the Kurdistan Region of Iraq.

2. Methods

2.1. *Data*

The data covers the 287-month period from January 1992 to November 2015 and was compiled using measured results made available via the Kurdistan Regional Statistics Office (KRSO), the Ministry of Planning—Kurdistan Region 2015 Bulletin, and the Environmental Statistics Bulletin—Iraq (CSOI) 2014.

The maximum and minimum temperature data in degrees centigrade are recorded in different locations in the Kurdistan Region by the General Directorate of Meteorology and Seismology office in Erbil. They send the data to the Central Statistical Organization branches in Erbil and Baghdad; after that the data are ready for publication. The data can also be obtained from the United Nations Food and Agriculture Organization Coordination Office for North Iraq and from the General Directorate of Agriculture in Erbil at the Agro-Meteorological Sub-sector Department. The data were analyzed using the Statgraphics Centurion XVII software package.

2.2. *Box–Jenkins methods*

An integral notion in the Box–Jenkins framework is that of a stationary stochastic process. A stochastic process is called stationary if its probability distribution is independent of time. This immediately implies that the mean and variance functions of a stationary process are time-independent. In particular, stationary processes cannot exhibit any sort of trend. The basic idea of the Box–Jenkins method is to transform any given stochastic process into a stationary one by separating the trend from the noise. A stationary process is completely determined by its mean, variance, and autocorrelation function, that is, the correlation between two values separated by a lag of k time steps. Comparing the autocorrelation of a given model to the one obtained from a data set is a crucial step in identifying accurate and reliable models (Chatfield 2004).

2.2.1. *Autoregressive moving average process (ARMA) or mixed process*

In order to reproduce autocorrelation patterns, a more general approach is needed. One option is to use a combination of autoregressive and moving average methods, namely, the ARMA(p,q) model, which treats a variable as a linear function of the p preceding values and the statistical errors associated to the q previous values (Jeffrey 1990). The most general form that this model can take is

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q},$$

which can be more succinctly expressed as in Eq. (1):

$$\begin{aligned} & y_t - \varphi_1 y_{t-1} - \varphi_2 y_{t-2} - \dots - \varphi_p y_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} + \mu \Rightarrow \\ & (1 - \varphi_1 \beta - \varphi_2 \beta^2 - \dots - \varphi_p \beta^p) y_t = (1 - \theta_1 \beta - \theta_2 \beta^2 - \dots - \theta_q \beta^q) a_t + \mu \ \Rightarrow \\ & \varphi(\beta) y_t = \theta(\beta) a_t + \mu \end{aligned} \tag{1}$$

where β denotes the backward shift operator ($\beta y_t = y_{t-1}$), φ and θ are polynomials of degree p and q, respectively, a_t denotes a purely random process, and μ is a constant.

2.2.2. *ARIMA models*

The ARMA model can be further refined by passing to the autoregressive integrated moving average or ARIMA(p,d,q). “Integration” here refers to the process of differencing in order to turn a nonstationary time series into a stationary one. The parameter p stands for the number of autoregressive terms, q is the number of statistical errors taken into account, representing the moving-average approach, and d is the number of nonseasonal differences (Chatfield 2004). A model of this type can be expressed as in Eq. (2):

$$(1-\beta)^d(1-\varphi_1\beta-\varphi_2\beta^2-\ldots-\varphi_p\beta^p)y_t=(1-\theta_1\beta-\theta_2\beta^2-\ldots-\theta_q\beta^q)a_t \Rightarrow$$
$$(1-\beta)^d\varphi(\beta)y_t=\theta(\beta)a_t \tag{2}$$

2.2.3. *SARIMA models*

Lastly, a problem arising in many applications is that of periodicity. In our case, the mean temperature clearly follows, to some degree, annual cycles, and these patterns need to be taken into account separately. Box and Jenkins (1970) incorporated seasonality into existing ARIMA approaches, arriving at the seasonal autoregressive moving average model or SARIMA(p,d,q)(P,D,Q)$_{S,}$ which can be written as

$$\varphi_p(\beta)\Phi_P(\beta^S)\,W_t=\theta_q\,(\beta)\Theta_Q(\beta^S)\,a_t \tag{3}$$

where β again denotes the backward shift operator, and S denotes the number of data points in a season, so that $\beta^S y_t=y_{t-S}$ is a shift of a full season (in our case S = 12 representing 12 months in a year). P and Q are terms equivalent to p and q, except they are applied to the series in steps of size S to remove seasonality, before the ARMA transformation with p and q is carried out; $\varphi_p, \Phi_P, \theta_q, \Theta_Q$ are polynomials of degree p, P, q, Q, respectively; a_t denotes a purely random process; and W_t is a differenced series used if the original process y_t is not stationary. The differencing refers to subtracting the earlier value of the time series observations from the present value and can be written as

$$W_t=\nabla^d\nabla_S^D y_t\,, \tag{4}$$

with $\nabla^d=(1-\beta)^d$ being the nonseasonal differencing and $\nabla_S^D=(1-\beta^s)^D$ the seasonal differencing. The superscripts d and D indicate the order of the nonseasonal and seasonal differencing, respectively (Chatfield 2004: Ye et al. 2013).

2.3. *Fitting Box–Jenkins models*

Following Box and Jenkins (1970), forecasts can be derived from the preceding model in four steps: (1) model identification, (2) estimation of model parameters, (3) diagnostic checking, and (4) application of the model forecasting (Box, Jenkins, and Reinsel 1994). The standard approach to model identification is to match both the autocorrelation function (ACF) and the partial autocorrelation function (PACF), which serves to isolate particularly strong self-correlations, for example, due to seasonality effects, to the ones exhibited by a given data set (Pankratz 1983); this procedure not only allows for model identification but also gives a first estimate of the model parameters. This estimate is then refined by other statistical methods, such as a mean-squares or maximum likelihood fit.

In the next step, the chosen model is checked against the time series by analyzing the series of residuals, sample correlations, and the residual histogram and performing a diagnosis test (Chatfield 2004). One such test is the Ljung–Box lack-of-fit test, which amounts to computing the following quantity:

$$Q=n(n+2)\sum_{k=1}^{h}\frac{r_k^2}{n-k} \tag{5}$$

where h is the maximum considered lag, n is the number of observations in the series, and r_k is the autocorrelation at lag k. Denote by m the number of model parameters fitted to the data, and under the null hypothesis the statistic Q is assumed to have a chi-squared distribution with (h – m) degrees of freedom. This hypothesis, and thus the model, is then rejected or accepted accordingly.

In general, among the models that pass this test, the ones with fewer parameters yield more accurate forecasts. Different models are compared by using either Akaike's information criterion (AIC) (Zakaria et al. 2012) or the Schwarz Bayesian information criterion (SBIC) (Schwarz 1978). When we add parameters to the fitted models, the value of the likelihood will rise and cause the problem of overfitting. AIC and SBIC are used to deal with this problem, by initiating a penalty term for the number of parameters in the model, with this value being greater in SBIC than in AIC (Schwarz 1978).

The AIC amounts to minimizing the following quantity:

$$AIC = -2\log_e(L) + 2(p + q + P + Q + C) \tag{6}$$

where L is the maximum likelihood, p the nonseasonal autoregressive order, q the nonseasonal moving average order, P the seasonal autoregressive order, Q the seasonal moving average order, and C the constant of the model.

The SBIC is computed as

$$SBIC = -2\log_e(L) + 2(p + q + P + Q + C)\log_e(n) \tag{7}$$

where n is the sample size.

2.3.1. *Fitting Box–Jenkins models for a seasonal model*

A seasonal model is identified using the following steps:

Step 1: Examine the time-series plots for seasonality and trend (i.e., check for stationarity).

Step 2: Carry out the necessary transformation of the data according to whether or not the data exhibit trend and seasonality effects, turn the data into a stationary series using both seasonal and nonseasonal differencing, and apply, for example, a natural log normal transformation.

Step 3: Examine the ACF and PACF of the new data, transformed data, and (if necessary) differenced data as they are the principal tools used to identify the AR and MA terms. Generally, to select nonseasonal terms we check the early lags of estimated ACF and PACF coefficients. Spikes in the ACF indicate nonseasonal MA terms, while spikes in the PACF are a sign of nonseasonal AR terms. For the seasonal terms, we study the patterns across lags that are multiples of S. For example, for monthly data, we look at lags 12, 24, 36, 48, and so on. The ACF and PACF may then be examined for spikes at the seasonal lags in the same way as for the earlier lags.

Step 4: When the model is selected, its parameters can be estimated using statistical techniques, such as the maximum likelihood, least-squares, or Yule–Walker method. The selected model or models should be those that might be reasonable on the basis of step 3, including the transformation and any differencing we made on the original data before looking at the ACF and PACF.

Step 5: Perform tests on the residuals in order to determine whether the model is adequate for the data. It is sensible to use a p-value threshold of 0.05 (and equivalently a confidence level of 95%), since this is the most widely used value and allows comparison to other studies. Test for the model's in-sample fitting performance, which is measured by the stationary R-squared and R-squared model fit, as well as AIC and SBIC. Test for the model's out-of-sample forecasting accuracy, the magnitude of error, which is measured by the root mean squared error (RMSE), the mean absolute error (MAE), and the mean absolute percentage error (MAPE). Also check for bias in the estimators, for instance the mean error (ME) and mean percentage error (MPE) are used as measures of biased estimators. It is necessary to check for the assumptions of normality and homoscedasticity, and also to check for autocorrelations (using the Ljung–Box test), in addition to plotting ACF. It is essential to compare AIC or SBIC values if several models have been tried (Ye et al. 2013). We recommend this procedure, with the full range of diagnostic tests, for SARIMA model selection for similar data in general.

If the results are unsatisfactory, we must go back to step 3 or maybe even step 2 (Chatfield 2004), and try a different set of parameters. At this stage we can select different models if any individual coefficients fall outside some specific interval around zero. In that case we can depend on the estimated ACF and PACF coefficient values to be more accurate and compare them with the appropriate confidence interval, which can be found by referring to the cumulative distribution function (cdf) for a normal distribution. For example, the 0.975 probability point of the standard normal is 1.96. The 95% confidence interval for ACF and PACF coefficients is therefore $\pm 1.96/\sqrt{n}$, where n is the number of observations in the series. Any coefficients outside this critical interval are evidence that the coefficients are significantly different from zero at the 95% confidence level and this interval is called the Bartlett range (Box, Jenkins, and Reinsel 1994).

3. Results

3.1. *Pattern of the Erbil mean temperature*

The most common patterns in time-series data are increasing or decreasing overall trend, cycles, seasonality, and irregular fluctuations. These are identified by plotting the original mean temperature data versus monthly recorded data over 32 years, as in Figure 1a. It appears from Figures 1a–1f, that there is a seasonality effect on the mean temperature data. The overall mean temperature during the studied period January 1992–November 2015 appears to exhibit a slight trend. In addition, there is a regular cycle with a period of 12 months, rising to a peak in July or August during the summer months and falling off in December (Figure 1d). Thus, the seasonal time-series decomposition method is suitable for our data.

Generally, it is difficult to detect any pure cycle and trend in Figure 1a, but the spectral density of the data in the periodogram (Figure 1b) shows a sharp spike at exactly the right frequency, thus indicating a hidden cycle. Although we are certain of a cyclical effect in the mean temperature data, we still used a periodogram based on Fourier decomposition; this fits the data to a sum of sine waves of different frequencies (Gottman 1981). For strongly seasonal data, for example, one cycle every 12 months, there will be a large spike at 1/12. The multiplicative

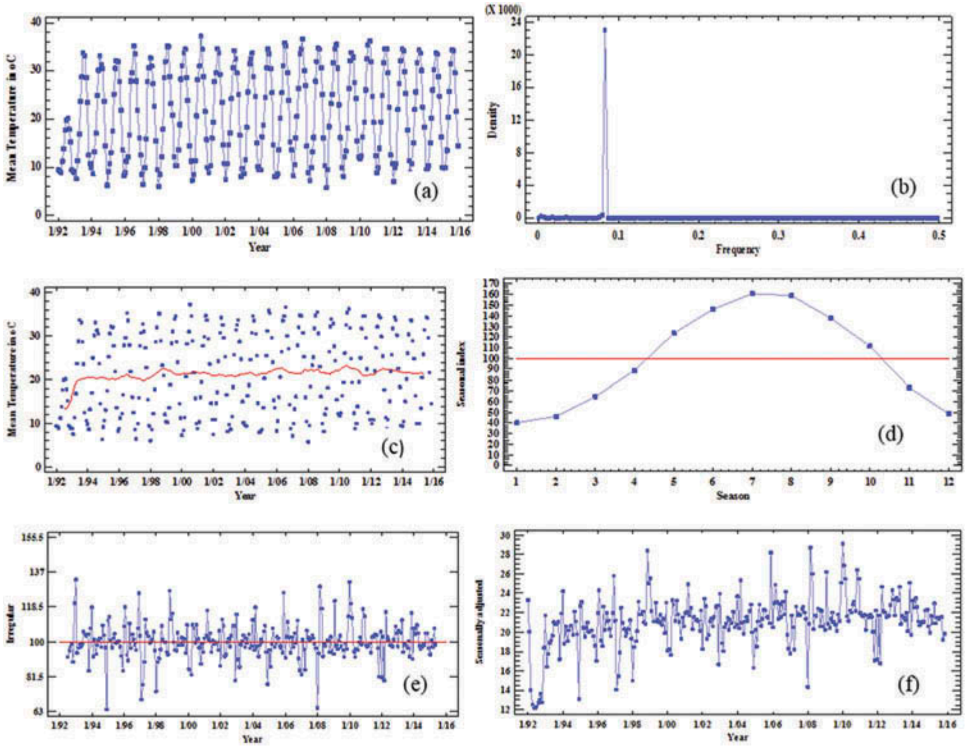

Figure 1. Time sequence and the decomposed plot of monthly Erbil mean temperature: (a) time-series plot of the original data; (b) temperature periodogram; (c) plot of trend cycle component; (d) plot of seasonal indices; (e) plot of irregular or residual component; (f) seasonally adjusted data.

seasonal decomposition has been applied on Kurdistan temperature data. Figure 1c displays the trend and cyclical effects in the original data set; a moving average of length equal to the seasonal order has been added. The moving average estimates the combined trend and cycle components, which are not usually separated (Dagum 2010; Grieser, Tromel, and Schonwiese 2002), and the seasonal indices estimate the seasonal component.

When using a multiplicative model, the indices are expressed as percentages. Figure 1d shows the seasonal indices for each season, scaled such that an average season corresponds to 100. The indices range from a low of 40.504 in January to a high of 160.589 in July. This indicates that there is a seasonal swing from 40.504% of the average to 160.589% of the average throughout 12 months. For example the index 0.91 in April indicates that the mean temperature is at 91% of the baseline. Note the strong seasonal effect for the temperature data, rising from a low in January to a peak in July or August and then falling off again. Finally the irregular component is displayed in Figure 1e. For the multiplicative model, this component is also expressed on a percentage basis, with the average value scaled to equal 100. In January 1993, the irregular component rose to approximately 133%, implying that temperature during that month was 33% more than expected, while in January 2008 the figure shows that the irregular component has declined to approximately 65% less than expected. The region has faced the same situation when in December 1994 the temperature was 64% below the average.

Once the decomposition has been performed, we can take the original data and divide it by the estimated seasonal indices to obtain the seasonally adjusted data Y'_t (Chatfield 2004), defined by

$$Y'_t = \frac{Y_t}{S_t}, \tag{8}$$

where Y_t is the seasonal component. The seasonally adjusted data is plotted in Figure 1f. Appendix A shows the mean temperature seasonal adjusted time series data in centigrade together with the other components. Table 5 in Appendix A explains each step of the seasonal decomposition. The trend-cycle column shows the results of a centered moving average of length 12 applied to temperature. The seasonality column shows the data divided by the moving average and multiplied by 100. Seasonal indices are then computed for each season by averaging the ratios across all observations in that season, and scaling the indices so that an average season equals 100. The data are then divided by the trend-cycle and seasonal estimates to give the irregular or residual component. This component is then multiplied by 100 (see Yi-Hui 2011; Theodosiou 2011).

3.2. *Fitting a SARIMA model*

The model development process begins by studying the original plot, autocorrelation function (ACF), partial autocorrelation function (PACF), and objective test of the raw data to ensure that the assumption of stationarity is met. In Figures 2a and 2b from the correlogram, most of the spikes in both the ACF and the PACF were found to be outside the confidence limits. Also, the ACF and PACF show a cyclic or seasonal variation of the correlations in the form of sinusoidal waves. Furthermore, both the ACF and the PACF

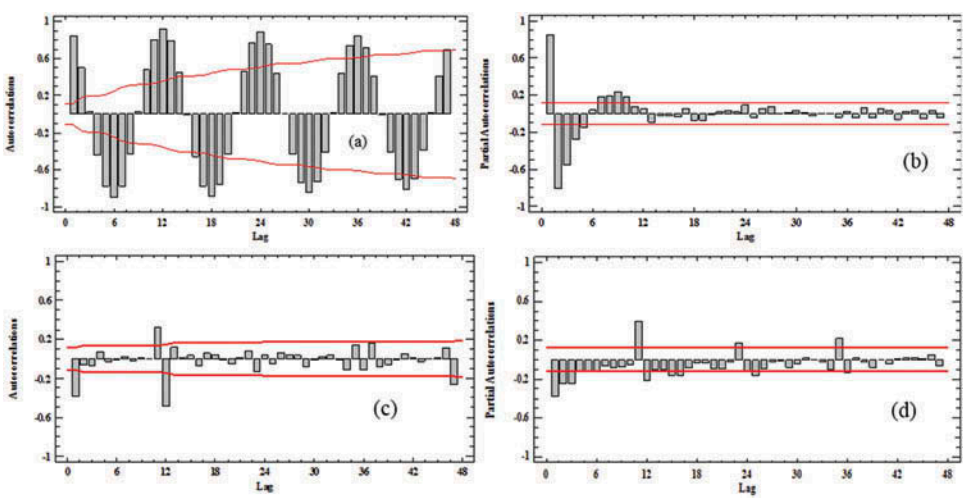

Figure 2. Correlogram plots: (a) Estimated autocorrelations for mean temperature (ACF), showing the correlogram for the original mean temperature data. Here the 12, 24, 36, and 48 autocorrelation coefficients are statistically significant at the 95.0% confidence level. (b) Estimated partial autocorrelation (PACF). (c) Estimated autocorrelation for adjusted mean temperature (ACF). (d) Estimated partial autocorrelation function for adjusted mean temperature (PACF).

show decay of the spikes indicating that the series has component problems. This is a clear indication of a seasonality of order 12.

The next step is to difference the series, by taking one regular difference to remove the seasonal trend in the data and then one seasonal differencing to take out a seasonal random walk type of nonstationarity. In order to make the series stationary around its variance, we applied a natural log transformation. Following the Box–Jenkins technique we depend on ACF or PACF plots to fit the order of the seasonal model (Chatfield 2004).

From Figures 2c and 2d we can choose our model, depending on the ACF and PACF spikes at low lags. To determine the nonseasonal AR terms, we look at the PACF, which shows clear spikes at lags 1, 2 and 3. Thus, the nonseasonal AR terms are determined to be of order 3. There are three spikes at lags 1, 11, and 12 in ACF, so we have three terms for nonseasonal MA. Now for the seasonal part of the model, in this case we look at lags 12, 24, 36, and 48 for both ACF and PACF. From the PACF we indicate that there are three significant spikes at lags 12, 24, and 36; thus, the order of the seasonal AR is three. In the ACF, there are two spikes at lags 12 and 48; this means that the order of the seasonal MA is two. Therefore, our base model is SARIMA$(3,1,3)(3,1,2)_{12}$. The model coefficient summary is given in Tables 1a–1i.

SARIMA model coefficient summary, (a)–(i): Starting from SARIMA$(3,1,3)(3,1,2)_{12}$, we arrive at our final model SARIMA$(0,1,2)(0,1,1)_{12}$ by, at each step, dropping the term with the highest p value associated to it and reestimating the remaining parameters until all p values for all estimated parameters are below 0.05.

Table 6 in Appendix B shows the estimated autocorrelations (partial autocorrelations) between values of adjusted mean temperature in degrees centigrade at various lags. We get two alternative models from it depending on the 95% confidence interval for ACF and PACF coefficients. For the Bartlett range, where n = 287, they are significantly different from zero at the 95% confidence level. The models are SARIMA$(3,1,3)(3,1,3)_{12}$ and SARIMA$(2,1,3)(3,1,3)_{12}$ when we select first and second spikes in the PACF instead of three (the PACF accounts for the correlations at all lower lags). Steps of estimating SARIMA$(3,1,3)(3,1,3)_{12}$ and SARIMA$(2,1,3)(3,1,3)_{12}$ parameters are shown in Tables 7A and 7B of Appendix C. The main conclusion between these models is that all three reduce to the same SARIMA$(0,1,2)(0,1,1)_{12}$ model, which is the model that we select. This model is appropriate for predicting future values from 2015(Dec) to 2025(Jan–Dec). It is stable when we delete from or add years to the original period from 1992(Jan–Dec) to 2015(Jan–Nov) and attempt to predict the given data, for example, selecting the 1993(Jan–Dec) to 2010(Jan–Dec) period to predict year 2011, and so on. This model is a final model that works under all conditions for various periods.

3.2.1. *Model estimation and evaluation*

Table 2 shows summarized results of seven tests run on the residuals to determine whether the model is adequate for predicting the mean temperature in Erbil and on the basis of historical data from 1992(Jan–Dec) to 2015(Jan–Nov).The magnitudes of error in the model are 1.696°C, 1.294°C, and 7.796% for RMSE, MAE, and MAPE, respectively, relative to the average of the predicted temperature at 21.028°C. The model shows no sign of biased estimations across the entire duration of the prediction period (10 years), based on the values of both ME and MPE as they are too close to zero. The fitted model is supported by the small value of AIC and SBIC. Since no tests are statistically significant at

Table 1. SARIMA model terms selection procedures.

$SARIMA(3,1,3)(3,1,2)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.359	1.291	0.781
AR(2)	−0.192	0.959	0.841
AR(3)	0.118	0.409	0.772
MA(1)	0.194	1.282	0.880
MA(2)	0.210	1.184	0.859
MA(3)	0.404	0.986	0.682
SAR(1)	−0.274	1.884	0.885
SAR(2)	0.014	0.065	0.827
SAR(3)	−0.025	0.073	0.734
SMA(1)	0.663	1.883	0.725
SMA(2)	0.243	1.754	0.890

(a)

$SARIMA(3,1,3)(3,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.535	1.446	0.712
AR(2)	−0.167	0.941	0.860
AR(3)	0.149	0.449	0.741
MA(1)	0.019	1.435	0.989
MA(2)	0.331	1.153	0.774
MA(3)	0.457	1.052	0.665
SAR(1)	−0.016	0.063	0.798
SAR(2)	0.015	0.060	0.805
SAR(3)	−0.008	0.060	0.895
SMA(1)	0.925	0.019	0.000

(b)

$SARIMA(3,1,3)(2,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.538	1.359	0.693
AR(2)	−0.164	0.900	0.856
AR(3)	0.144	0.427	0.736
MA(1)	0.016	1.349	0.991
MA(2)	0.335	1.089	0.759
MA(3)	0.451	1.001	0.653
SAR(1)	−0.014	0.063	0.820
SAR(2)	0.019	0.062	0.762
SMA(1)	0.927	0.017	0.000

(c)

$SARIMA(3,1,3)(1,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.439	1.532	0.775
AR(2)	−0.195	1.038	0.851
AR(3)	0.143	0.467	0.761
MA(1)	0.117	1.522	0.939
MA(2)	0.250	1.290	0.846
MA(3)	0.445	1.123	0.692
SAR(1)	−0.018	0.066	0.782
SMA(1)	0.923	0.018	0.000

(d)

$SARIMA(3,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.364	1.424	0.798
AR(2)	−0.189	1.012	0.852
AR(3)	0.123	0.430	0.775
MA(1)	0.192	1.416	0.892
MA(2)	0.214	1.273	0.867
MA(3)	0.406	1.049	0.699
SMA(1)	0.925	0.017	0.000

(e)

$SARIMA(2,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.397	0.759	0.601
AR(2)	0.099	0.336	0.769
MA(1)	0.162	0.757	0.831
MA(2)	0.511	0.718	0.477
MA(3)	0.137	0.164	0.404
SMA(1)	0.932	0.014	0.000

(f)

$SARIMA(1,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.242	0.521	0.642
MA(1)	0.319	0.516	0.537
MA(2)	0.336	0.321	0.298
MA(3)	0.171	0.139	0.218
SMA(1)	0.927	0.017	0.000

(g)

$SARIMA(0,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
MA(1)	0.560	0.060	0.000
MA(2)	0.194	0.068	0.005
MA(3)	0.115	0.058	0.055
SMA(1)	0.926	0.017	0.000

(h)

$SARIMA(0,1,2)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
MA(1)	0.588	0.059	0.000
MA(2)	0.237	0.058	0.000
SMA(1)	0.933	0.016	0.000

(i)

Table 2. Model testing.

Model	RMSE	MAE	MAPE	ME	MPE	AIC	SBIC
SARIMA(0,1,2)(0,1,1)$_{12}$	1.696	1.294	7.796	0.143	−0.294	1.078	1.116

the 95% or higher confidence level, the proposed model, SARIMA(0,1,2)(0,1,1)$_{12}$, passes all tests. Therefore, it is considered a good model for forecasting.

In model selection and validity model testing criteria for mean temperature forecasting, RMSE is the root mean squared error, MAE mean absolute error, MAPE mean absolute percentage error, ME mean error, MPE mean percentage error, AIC Akaike's information criterion, and SBIC Schwarz Bayesian information criterion.

The model parameters (autoregressive, moving average, seasonal autoregressive, and seasonal moving average) are estimated using maximum likelihood estimation. The estimates of the parameters are shown in Tables 1a–1i. Based on the 95% confidence level, we conclude that all the coefficients of the SARIMA(0,1,2)(0,1,1)$_{12}$ model are significantly different from zero. Furthermore, the model reproduces the data under study very well, as indicated by the stationary R-squared (0.514) and R-squared (0.96).

The mathematical equation for the SARIMA(0,1,2)(0,1,1)$_{12}$ model's estimated coefficients is

$$\begin{aligned}
&(1-\beta)(1-\beta^{12})y_t = (1+\theta_1\beta+\theta_2\beta^2)(1+\Theta_1\beta^{12})a_t \quad \Rightarrow \\
&(1-\beta-\beta^{12}+\beta^{13})y_t = (1+\theta_1\beta+\theta_2\beta^2+\Theta_1\beta^{12}+\Theta_1\theta_1\beta^{13}+\Theta_1\theta_2\beta^{14})a_t \quad \Rightarrow \\
&y_t = y_{t-1}+y_{t-12}-y_{t-13}+a_t+\theta_1 a_{t-1}+\theta_2 a_{t-2}+\Theta_1 a_{t-12}+\Theta_1\theta_1 a_{t-13}+\Theta_1\theta_2 a_{t-14}
\end{aligned} \tag{9}$$

Using the parameters we estimate from the data, this gives

$$\begin{aligned}
&y_t = y_{t-1}+y_{t-12}-y_{t-13}+a_t+0.588a_{t-1}+0.237a_{t-2}+0.933a_{t-12}+0.549a_{t-13} \\
&\quad +0.221a_{t-14} \quad \Rightarrow \\
&\hat{y}_{288} = y_{288-1}+y_{288-12}-y_{288-13}+\hat{a}_{288}+0.588a_{288-1}+0.237a_{288-2}+0.933a_{288-12} \\
&\quad +0.549a_{288-13}+0.221a_{288-14} \quad \Rightarrow \\
&\hat{y}_{288} = y_{287}+y_{276}-y_{275}+0+0.588a_{287}+0.237a_{286}+0.933a_{276}+0.549a_{275}+0.221a_{274}
\end{aligned} \tag{10}$$

Equation (10) may now be used to forecast the Erbil future mean temperature value for the coming 10 years (121 months) starting from December 2015, $\hat{y}_{288}$, where $\hat{y}$ is the predicted value and 288 is the number of months that have passed since January 1992 (see Appendix D).

3.2.2. *Model diagnostics*

In time-series modeling, the selection of a best model fit to the data is directly related to whether the residual analysis is performed well. One of the assumptions of the SARIMA (seasonal ARIMA) model is that, for a good model, the residuals must follow a white noise process; that is, the residuals have zero mean, have constant variance (homoscedasticity), and also are uncorrelated with past values. A special case of this process is when the residuals are normally distributed, when they are said to follow a Gaussian white noise process. It is such a process that we test for here.

For our selected SARIMA$(0,1,2)(0,1,1)_{12}$ model, normality is tested by a normal probability plot as shown in Figure 3a, the standardized residual in Figure 3b, the periodogram in Figure 3c, and in Figure 3d the histogram of residuals. The four figures of residuals for mean temperature data shows that the residuals of the model are consistent with a normal distribution assumption. Table 3 indicates that the SARIMA$(0,1,2)(0,1,1)_{12}$ model residuals are uncorrelated as well as independent for all three tests indicated.

In the residual autocorrelation and independence test for the selected model, RUNS is the test for excessive runs up and down, RUNM the test for excessive runs above and below median, and AUTO the Ljung–Box test for excessive autocorrelation.

In order to determine whether the residuals are randomly distributed, three tests have been performed. In the first step, we counted how often the sequence exceeded the median, finding 143 as opposed to 138, which is expected for a random sequence. In the second test, we determined how often the sequence increased, finding 187 steps as compared to the expected 182.3. Both of these tests result in a p value that is larger than 0.05, which indicates that there is no reason to reject the hypothesis of randomness at a 95% confidence level. Third, the p value (0.819) for the Ljung–Box statistic exceeds 5% as well, indicating that there is no significant departure from white noise for the residuals; that is, there is no indication of autocorrelation in residuals of the selected model. Thus, the selected model SARIMA $(0,1,2)(0,1,1)_{12}$ satisfies all the model assumptions.

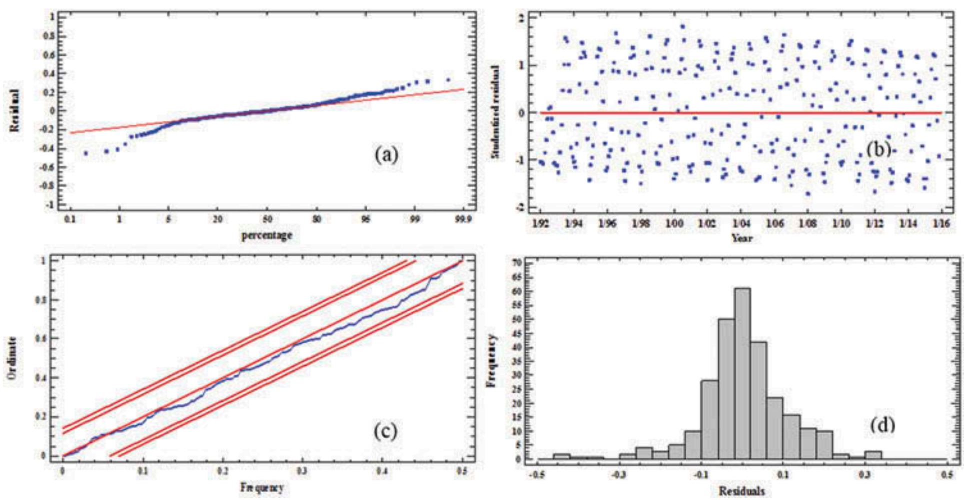

Figure 3. Residual plot for SARIMA$(0,1,2)(0,1,1)_{12}$ model: (a) normal probability plot for the residual, (b) the standardized plot for residuals, (c) periodogram for residuals, and (d) histogram for residuals.

Table 3. Test for autocorrelation and independence.

Tests	Test statistic value	p value
RUNS: Runs above and below median	0.545	0.586
RUNM: Runs up and down	0.599	0.549
AUTO: Ljung–Box test	15.067	0.819

The estimated white noise variance at 271 degrees of freedom was 0.014 and the estimated white noise standard deviation was 0.119; also, the difference in variance and difference in mean test were "OK," which indicate that our selected model residuals are homogeneous, that is, that there are no significant departures from white noise for the residuals at 95%. The current model is adequate for the data as the selected model SARIMA$(0,1,2)(0,1,1)_{12}$ satisfies all our model assumptions (normality, uncorrelated residuals, and homoscedasticity). Therefore, the selected model is considered a good model to forecast future values.

Looking at Figure 4, the autocorrelation checks of the residuals indicate that the model is good because they resemble a white noise process; that is, the residuals have zero mean, constant variance, and are also uncorrelated. Since the model diagnostic tests show that all the parameter estimates are significant and the residual series for the model are random, it can then be concluded that a SARIMA$(0,1,2)(0,1,1)_{12}$ model is adequate for the Erbil mean temperature series. Therefore, SARIMA$(0,1,2)(0,1,1)_{12}$ is used to forecast the future mean temperature series of the Kurdistan Region.

3.3. *Forecasting using SARIMA $(0,1,2)(0,1,1)_{12}$*

The performance of SARIMA$(0,1,2)(0,1,1)_{12}$ model for the Erbil mean temperature is now evaluated by forecasting the data one step prediction for years 2014(Jan–Dec)–2015(Jan–Nov) to indicate the model's adequacy and performance and for comparison purposes. Using the selected model, the 23 months of forecast are shown in Table 4 and Figure 5.

Checking the SARIMA$(0,1,2)(0,1,1)_{12}$ model by predicting the existing mean temperature data in January 2014 through November 2015, we find the following.

It appears from Figure 5 that the selected model is very well suited for predicting the future development of the Erbil mean temperature, as the differences between the actual data (solid line) and forecast data (dashed line) are very small; the lower line represents their residual values as tabulated in Table 4.

Figure 6 shows the forecast values for the mean temperatures for 121 months from December 2015 to December 2025. The forecast mean temperatures are represented by the solid line, and the dashed lines indicate the 95% confidence band. In fact, this figure does not give us a clear trend of future mean temperature, as it may both increase or decrease

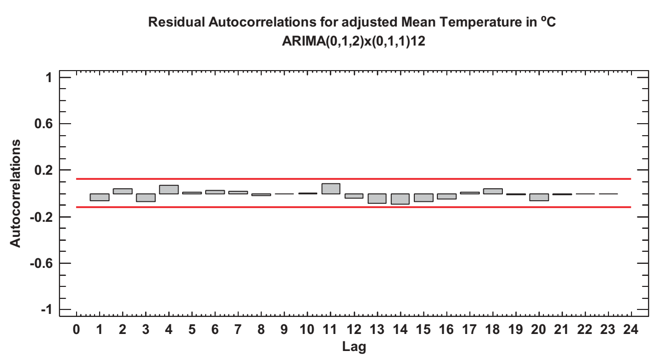

Figure 4. ACF of residuals for SARIMA$(0,1,2)(0,1,1)_{12}$ model.

Table 4. Forecast mean temperatures value (°C) for January 2014 to November 2015.

Period	Actual data	Forecast data	Residual	Period	Actual data	Forecast data	Residual
Jan-14	9.40	8.35	1.05	Jan-15	8.07	8.56	−0.49
Feb-14	10.40	10.71	−0.31	Feb-15	9.98	9.90	0.08
Mar-14	15.10	14.12	0.98	Mar-15	13.44	13.88	−0.44
Apr-14	20.10	20.42	−0.32	Apr-15	20.47	19.04	1.43
May-14	26.05	27.20	−1.15	May-15	25.78	26.83	−1.05
Jun-14	30.30	32.19	−1.89	Jun-15	31.42	31.07	0.35
Jul-14	33.70	34.83	−1.13	Jul-15	34.34	34.63	−0.29
Aug-14	33.50	34.51	−1.01	Aug-15	34.06	34.09	−0.03
Sep-14	28.90	29.78	−0.88	Sep-15	29.60	29.58	0.02
Oct-14	24.85	23.89	0.96	Oct-15	21.50	23.97	−2.47
Nov-14	16.30	16.11	0.19	Nov-15	14.50	15.10	−0.60
Dec-14	10.00	10.61	−0.61				

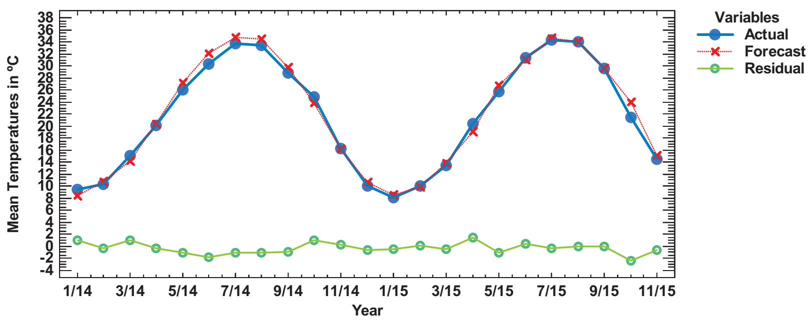

Figure 5. The forecast mean temperatures values (°C) for January 2014 to November 2015.

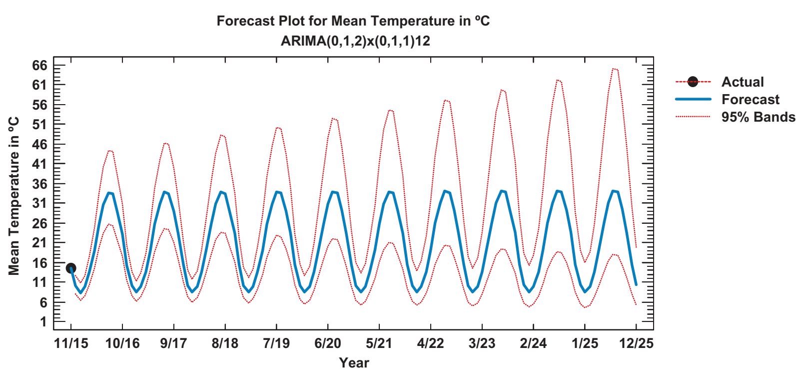

Figure 6. The Erbil mean temperature forecast from December 2015 to December 2025.

within the confidence limits. We therefore decided to plot the future value for the same period using the data from January 1992 to November 2015 as a base for our forecast instead of just using the last month of the existing data, November 2015, in order to show the more striking graph shown in Figure 7. Table 8 in Appendix D shows the 121 months of forecast. The forecast mean temperature in Erbil for the next 10 years looks flattened when compared to previous values in Figure 7, meaning that the temperature is predicted to be stable with the same pattern in the future.

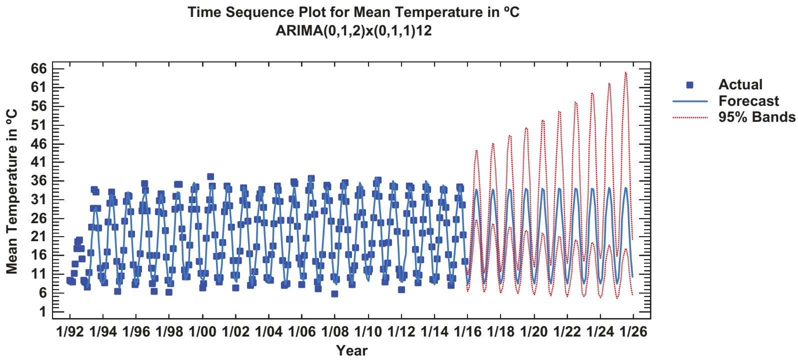

Figure 7. The Erbil mean temperature forecast from December 2015 to December 2025.

The change in temperature is clear in Erbil, where the lowest temperature recorded was 5.80°C in January 2008 while the highest temperature recorded was 37.25°C in July 2000. A change in temperature happened in August 1992, when 20.15°C degrees was recorded. The temperatures in January, February, March, April, November, and December were below the temperature mean of 21°C, while the other months were above the mean. From Figure 7 it appears that the predicted mean temperature in January has decreased from 9.4°C in 2014 to 8.1°C in 2015 but is projected to rise in 2016 until 2025, while the average temperatures in July, August, and December for the same period will be generally around 34°C, 33.4°C, and 10°C.

4. Conclusions

The temperature in the Kurdistan Region has changed similarly to many other areas in the world, due to climate change. Many researchers have studied these phenomena in different places by using various methods and statistical tools, among them the seasonal time-series method. In the Kurdistan Region, studies on rainfall and on electricity demand in both Sinjar district and Sulaymaniyah Governorate have been carried out using ARIMA and SARIMA, respectively. Although Erbil is the capital of Iraqi Kurdistan and it shows a significant shift in temperature over the last decades, until now no time-series-based studies in that direction have been undertaken.

In general, the pattern of mean temperatures in the Erbil Kurdistan Region from January 1992 to November 2015 was observed to be not stationary and increasing over time. The nonstationarity of the mean temperature series was verified by the plot of the sample ACF and PACFs. The data cover 287 time periods. Currently, a seasonal autoregressive integrated moving average (SARIMA) model has been selected by following the procedures of the Box–Jenkins SARIMA model building. The underlying assumption is that the best forecast for future data is given by a parametric model relating the most recent data value to previous data values and previous noise. Each value of mean temperature has been adjusted in the following way before the model was fitted: (1) Seasonal and nonseasonal differences are applied to remove the effect of trend and take out a seasonal random walk type of nonstationarity, that is, to make the series stationary around its mean, (2) a natural log transformation was applied to make the series stationary around its variance, and (3) a multiplicative seasonal adjustment was applied.

Using the ACF and PACF estimated coefficient plots in Figures 2c and 2d, as well as Tables 7A and 7B (Appendix C), three models were developed, namely, SARIMA(3,1,3)(3,1,2)$_{12}$, SARIMA(3,1,3)(3,1,3)$_{12}$, and SARIMA(2,1,3)(3,1,3)$_{12}$; each of them leads to the same model, SARIMA (0,1,2)(0,1,1)$_{12}$. We get this particular model based on the significance terms in the model. Terms with p values less than 0.05 are considered statistically significantly different from zero at the 95.0% confidence level. Starting with the base model, SARIMA(3,1,3)(3,1,2)$_{12}$, the p values for AR(3), MA(3), SAR(3), and SMA(2) terms in the model are greater than 0.05, so they are not statistically significant. We should therefore consider reducing the order of the terms depending on the maximum p values among them; this is illustrated in Tables 1a–1i. Here at each step, we are dropping the term with the highest p value associated to it and reestimating the remaining parameters until all p values for all estimated parameters are below 0.05.

The model diagnostics were performed through careful examination of the model residuals. The model residuals were found to be following a white-noise process with a mean of zero and a constant variance, hence uncorrelated. The comparison for choosing the best model to represent the data is based on the RMSE, MAE, and MAPE values of 1.696°C, 1.294°C, and 7.796%, respectively. No bias was detected in the model, based on the values of both ME and MPE (0.143 and −0.294, respectively), as they are close to zero. The fitted model is supported by the small values of AIC and SBIC of 1.07 and 1.116, respectively.

As no tests are statistically significant at the 95% or higher confidence level, the existing model is sufficient for the data. Furthermore, the model residuals satisfy our assumptions of normality, homoscedasticity, and being uncorrelated with past values, through a normal probability plot, standardized residual plot, periodogram, and histogram (see Figures 3 and 4) and the Ljung–Box test. The statistical analysis leads us to conclude that there is no reason to reject the hypothesis that the residuals follow a white-noise process at the 95% confidence level. In addition, the value of R-squared at the 95% confidence level was 0.96, which means that our model explains the data well. Perhaps good indicators that our SARIMA(0,1,2)(0,1,1)$_{12}$ model represents this region well are the small values of the estimated variance and standard deviation for the model input white noise 0.014 and 0.119.

The selected model is further validated by predicting the mean temperature of January 2014 to November 2015 and reproducing the known seasonal patterns in its forecasts. This shows that the estimated forecast mean temperature was identical or very close to the actual real data. The pattern of mean temperatures in Erbil from December 2015 to December 2025 was observed to be stationary, and hence does not follow any particular pattern (neither increasing nor decreasing).

Similar investigations have been carried out in the wider region by other researchers. Tektas (2010) used an ARIMA(2,1,1) model to predict the weather of the Göztepe Region in İstanbul, Turkey. This involved data from 2000 to 2008 collected on a daily basis. We note that the model fits their data less well than our SARIMA model fits ours (assessed through the standard criteria). A SARIMA(0,0,1)(0,1,1)$_{12}$ model was adopted by Sarraf, Vahdat, and Behbahaninia (2011) to forecast average monthly temperature at Ahwaz synoptic station in Iran, using average monthly temperature data from 1990 to 2010.

This model was a good fit to their data, and they used it to predict the average temperature for 2010–2011, with a particular applications for agriculture within the region that year.

Air temperature of the southern Caspian Sea region (Anzali, Ramsar, and Babolsar synoptic stations) was modeled by Khajavi et al. (2012) in Iran. A SARIMA$(1,0,0)(0,1,1)_{12}$ model was chosen to forecast future mean monthly temperature at the Anzali and Babolsar stations, while a SARIMA$(0,0,2)(0,1,1)_{12}$ model was used for the Ramsar mean monthly temperature. They compared forecast temperature at all stations with real data for the year 2005–2006, with good predictability.

The monthly mean temperature at the Shiraz Synoptic Station in the south of Iran was used in a study by Babazadeh and Shamsnia (2014). They used a SARIMA$(2,1,0)(2,1,0)_{12}$ model to forecast the future mean temperature in the region, using the historical mean temperature data in the region from 21 years. Their chosen model again produced reliable forecasts for future mean temperature in the Shiraz Region, and was also used to forecast crop productions in the years 2008–2009 and 2009–2010.

The fitted SARIMA models that we have discussed are all quite similar in character, and it appears that this is a good general model for fitting temperature data, providing a good fit in the cases considered. We note, however, that in fitting the preceding models, none of the named studies carried out the full range of tests and procedures that we outline in section 2.3.1, which we believe should be followed (see, e.g., Chatfield 2004). We would recommend the procedure carried out in our article for the selection of SARIMA models for equivalent data elsewhere. Based upon our results and the model diagnostics performed, the identified model was found to be a good model for predicting future mean temperatures in the Kurdistan Region. Potential applications include the forecasting of crop yields as in Babazadeh and Shamsnia (2014) or the prediction of power requirements for temperature-sensitive energy usage such as heating and refrigeration, or of adverse weather events.

References

Al-Kubaisi, Q. Y., and M. M. Gardi. 2012. Dust storm in Erbil City as a result of climatic change in Kurdistan Region Iraq. *Iraqi Journal of Science* 53 (3C):40–44.

Asamoah-Boaheng, M. 2014. Using SARIMA to forecast monthly mean surface air temperature in the Ashanti Region of Ghana. *International Journal of Statistics and Applications* 4 (6):292–98. doi:10.5923/j.statistics.20140406.06.

Babazadeh, H., and S. A. Shamsnia. 2014. Modeling climate variables using time series analysis in arid and semi arid regions. *African Journal of Agricultural Research* 9 (26):2018–27.

Box, G. E., and G. M. Jenkins. 1970. *Time series analysis: Forecasting and control*. San Francisco, CA: Holden-Day.

Box, G. E., G. M. Jenkins, and G. C. Reinsel. 1994. *Time series analysis: Forecasting and control*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall.

Can, A., and A. T. Atimtay. 2002. Time series analysis of mean temperature data in Turkey. *Applied Time Series* 4:20–23.

Chatfield, C. 2004. *The analysis of time series: An introduction*, 6th ed. London, UK: Chapman & Hall/CRC.

Choi, J., D. C. Roberts, and E. Lee. 2015. Forecasting oil production in North Dakota using the seasonal autoregressive integrated moving average (S-ARIMA). *Natural Resources* 6 (1):16–26.

Dagum, E. B. 2010. Time series and decomposition. *Statistica* 70 (4):433–57.

Eklund, L., and P. Pilesjö. 2012. Migration patterns in Duhok Governorate, Iraq, 2000–2010. *Open Geography Journal* 5:48–58.

El-Kadi, A. 2001. Variation of rainfall and drought conditions in Gaza–Palestine: On a regional and global context. *Journal of the Islamic University* 9 (2):41–66.

Ghahraman, B. 2006. Time trend in mean annual temperature of Iraq. *Turkish Journal of Agricultural and Forestry* 30:439–48.

Ghil, M., M. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou. 2002. Advanced spectral methods for climate time series. *Reviews of Geophysics* 40 (1):1–41.

Gottman, J. M. 1981. *Time-series analysis, A comprehensive introduction for social scientists*. New York, NY: Cambridge University Press.

Grieser, J., S. Tromel, and C. D. Schonwiese. 2002. Statistical time series decomposition into significant components and application to European temperature. *Theoretical and Applied Climatology* 71:171–83.

Hansen, J., M. Sato, R. Ruedy, K. Lo, D. W. Lea, and M. Medina-Elizade. 2006. Global temperature change. *Proceedings of the National Academy of Sciences* 103 (39):14288–93.

Intergovernmental Panel on Climate Change. 2013. Summary for policymakers. In *Climate change 2013: The physical science basis. Contribution of Working Group I to the fifth assessment report of the Intergovernmental Panel on Climate Change*, ed. T. F. Stocker, D. Qin, G.-K. Plattner, M. M. B. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, 1–28. New York, NY: Cambridge University Press.

Jeffrey, J. 1990. Improving forecasts by decomposing the error. *Journal of Business Forecasting Methods & Systems* 9 (1):12–15.

Keller, E. A., and R. H. Blodgett. 2006. *Natural hazards: Earth's processes as hazards, disasters, and catastrophes*. Upper Saddle River, NJ: Pearson-Prentice Hall.

Khajavi, E., J. Behzadi, M. T. Nezami, A. Ghodrati, and M. A. Dadashi. 2012. Modeling ARIMA of air temperature of the southern Caspian Sea coasts. *International Research Journal of Applied and Basic Sciences* 3 (6):1279–87.

Khalid, N. J. 2014. Urban heat island in Erbil City. Student thesis series, INES, NGEM01 20131, Report No. 308, 1–57. http://lup.lub.lu.se/student-papers/record/4449039 (accessed July 2, 2015).

Lee, C., and C. Ko. 2011. Short-term load forecasting using lifting scheme and ARIMA models. *Expert Systems with Applications* 38:5902–11.

Mann, M. E. 2008. Smoothing of climate time series revisited. *Geophysical Research Letters* 35 (L16708):1–5.

Mann, M. E., and J. Park. 1996. Joint spatio-temporal modes of surface temperature and sea level pressure variability in the Northern Hemisphere during the last century. *Journal of Climate* 9 (9):2137–62.

Mraoua, M., and D. Bari. 2007. Temperature stochastic modeling and weather derivatives pricing: Empirical study with Moroccan data. *Afrika Statistika* 2 (1):22–43.

Pankratz, A. 1983. *Forecasting with univariate Box–Jenkins models concepts and cases*. New York, NY: John Wiley & Sons.

Rahmstorf, S., A. Cazenave, J. Church, J. E. Ansen, R. F. Keeling, D. E. Parker, and R. C. J. Somerville. 2007. Recent climate observations compared to projections. *Science* 316 (5825):709.

Romilly, P. 2005. Time series modelling of global mean temperature for managerial decision-making. *Journal of Environmental Management* 76:61–70.

Saeed, M. A., and K. A. Abas. 2012. Analysis of climate and drought conditions in the Federal Region of Kurdistan. *International Scientific Journal, Environmental Science* 2:1–8.

Sarraf, A., S. F. Vahdat, and A. Behbahaninia. 2011. Relative humidity and mean monthly temperature forecasts in ahwaz station with ARIMA model in time series analysis. *International Conference on Environment and Industrial Innovation, IPCBEE*, Kuala Lumpur, Malaysia, vol. 12, 149–53.

Schwarz, G. E. 1978. Estimating the dimension of a model. *Annals of Statistics* 6 (2):461–64.

Tektaş, M. 2010. Weather forecasting using ANFIS and ARIMA models. A case study for Istanbul. *Environmental Research, Engineering and Management* 51 (1):5–10.
Theodosiou, M. 2011. Forecasting monthly and quarterly time series using stl decomposition. *International Journal of Forecasting* 27 (4):1178–95.
Turkes, M. 1996c. Metrological drought in Turkey: A historical perspective, 1930–93. *Drought Network News* [University of Nebraska] 8:17–21.
Turkes, M. 1999. Vulnerability of Turkey to desertification with respect to precipitation and aridity conditions. *Turkish Journal of Engineering and Environmental Science* 23 (5):363–80.
United Nation Environment Programme. 2009. *Climate change*. New York, NY: United Nations.
Ye, L. M., G. X. Yang, E. Van Ranst, and H. J. Tang. 2013. Time-series modeling and prediction of global monthly absolute temperature for environmental decision making. *Advances in Atmospheric Sciences* 30 (2):382–96.
Yi-Hui, L. 2011. Analysing and forecasting the reliability for repairable system using the time series decomposition method. *International Journal of Quality & Reliability Management* 28 (3):317–27.
Zakaria, S., N. Al-Ansari, S. Knutsson, and T. Al-Badrany. 2012. ARIMA models for weekly rainfall in the semi-arid Sinjar District at Iraq. *Journal of Earth Sciences and Geotechnical Engineering* 2 (3):25–55.
Zakaria, S., Y. Mustafa, D. Mohammed, S. Ali, N. Al-Ansari, and S. Knutsson. 2013. Estimation of annual harvested runoff at Sulaymaniyah Governorate, Kurdistan region of Iraq. *Natural Science* 5 (12):1272–83.

Appendix A

Table 5. Seasonal decomposition method: Multiplicative of Erbil mean temperature (°C) from January 1992 to November 2015.

Period	Data	Trend cycle	Seasonality	Irregular	Seasonally adjusted	Period	Data	Trendcycle	Seasonality	Irregular	Seasonally adjusted
Jan-92	9.43				23.28	Aug-97	31.06	19.72	157.51	99.27	19.58
Feb-92	9.25				20.05	Sep-97	27.35	19.90	137.43	99.99	19.90
Mar-92	9.03				14.03	Oct-97	22.7	20.08	113.05	101.15	20.31
Apr-92	11.20				12.56	Nov-97	15.54	20.20	76.95	105.20	21.25
May-92	13.75				11.14	Dec-97	9.4	20.28	46.35	95.29	19.33
Jun-92	17.80				12.18	Jan-98	6.09	20.44	29.80	73.57	15.04
Jul-92	19.85	13.46	147.51	91.85	12.36	Feb-98	8.52	20.71	41.15	89.18	18.47
Aug-92	20.15	13.36	150.83	95.07	12.70	Mar-98	12.27	20.94	58.60	91.06	19.07
Sep-92	17.89	13.39	133.64	97.23	13.02	Apr-98	17.95	21.03	85.37	95.78	20.14
Oct-92	15.18	13.72	110.63	98.99	13.58	May-98	28.55	21.27	134.24	108.74	23.13
Nov-92	9.35	14.37	65.07	88.96	12.78	Jun-98	31.95	21.60	147.91	101.19	21.86
Dec-92	8.93	15.24	58.61	120.50	18.36	Jul-98	35.10	21.89	160.33	99.84	21.86
Jan-93	8.78	16.27	53.97	133.24	21.68	Aug-98	35.06	22.14	158.35	99.80	22.10
Feb-93	7.55	17.39	43.42	94.11	16.36	Sep-98	28.88	22.27	129.70	94.36	21.01
Mar-93	11.40	18.38	62.04	96.40	17.71	Oct-98	23.28	22.36	104.13	93.17	20.83
Apr-93	16.85	19.17	87.89	98.60	18.90	Nov-98	20.78	22.39	92.82	126.90	28.41
May-93	23.65	19.65	120.35	97.49	19.16	Dec-98	12.15	22.34	54.38	111.80	24.98
Jun-93	28.70	19.86	144.54	98.88	19.63	Jan-99	10.35	22.25	46.51	114.84	25.55
Jul-93	33.75	19.97	169.00	105.24	21.02	Feb-99	10.23	22.17	46.13	99.99	22.17
Aug-93	33.10	20.06	165.02	104.01	20.86	Mar-99	13.58	22.14	61.35	95.32	21.10
Sep-93	28.65	20.18	141.96	103.29	20.84	Apr-99	18.81	22.15	84.92	95.27	21.10
Oct-93	23.50	20.26	115.97	103.76	21.03	May-99	28.40	21.92	129.59	104.97	23.01
Nov-93	12.55	20.32	61.75	84.42	17.16	Jun-99	31.05	21.62	143.64	98.27	21.24
Dec-93	10.65	20.43	52.12	107.16	21.90	Jul-99	33.80	21.45	157.56	98.12	21.05
Jan-94	9.80	20.46	47.89	118.23	24.20	Aug-99	34.50	21.25	162.35	102.33	21.74
Feb-94	8.65	20.36	42.48	92.08	18.75	Sep-99	28.55	21.08	135.43	98.53	20.77
Mar-94	13.25	20.36	65.09	101.14	20.59	Oct-99	23.90	21.07	113.44	101.50	21.38
Apr-94	17.00	20.43	83.20	93.33	19.07	Nov-99	14.55	21.14	68.82	94.09	19.89
May-94	24.90	20.52	121.37	98.31	20.17	Dec-99	11.20	21.13	53.01	108.98	23.03
Jun-94	30.10	20.42	147.43	100.86	20.59	Jan-00	7.35	21.27	34.56	85.32	18.15
Jul-94	33.10	20.21	163.76	101.98	20.61	Feb-00	8.40	21.42	39.22	85.01	18.21
Aug-94	31.25	20.27	154.16	97.17	19.70	Mar-00	11.35	21.44	52.95	82.27	17.64
Sep-94	30.40	20.35	149.39	108.69	22.12	Apr-00	20.75	21.35	97.18	109.02	23.28
Oct-94	23.60	20.36	115.90	103.70	21.12	May-00	28.20	21.30	132.37	107.22	22.84
Nov-94	14.45	20.39	70.86	96.88	19.76	Jun-00	30.95	21.30	145.28	99.39	21.17
Dec-94	6.35	20.43	31.09	63.92	13.06	Jul-00	37.25	21.32	174.73	108.81	23.20

Jan-95	9.20	20.41	45.08	111.31	22.71	Aug-00	34.61	21.46	161.29	101.66	21.81
Feb-95	10.65	20.40	52.22	113.18	23.08	Sep-00	28.90	21.73	133.00	96.77	21.03
Mar-95	13.15	20.31	64.73	100.58	20.43	Oct-00	21.55	21.87	98.54	88.17	19.28
Apr-95	17.4	20.13	86.43	96.96	19.52	Nov-00	15.74	21.80	72.22	98.73	21.52
May-95	25.2	20.02	125.86	101.95	20.41	Dec-00	10.00	21.73	46.03	94.62	20.56
Jun-95	30.6	20.07	152.48	104.31	20.93	Jan-01	8.90	21.57	41.27	101.89	21.97
Jul-95	32.15	20.14	159.65	99.42	20.02	Feb-01	10.20	21.42	47.62	103.21	22.11
Aug-95	31.95	20.15	158.58	99.95	20.14	Mar-01	16.05	21.40	75.00	116.54	24.94
Sep-95	27.75	20.14	137.82	100.27	20.19	Apr-01	19.40	21.50	90.24	101.23	21.76
Oct-95	21.85	20.06	108.90	97.44	19.55	May-01	27.80	21.54	129.07	104.55	22.52
Nov-95	13.6	20.14	67.51	92.30	18.59	Jun-01	29.70	21.53	137.95	94.38	20.32
Dec-95	8.3	20.22	41.06	84.41	17.06	Jul-01	34.65	21.50	161.16	100.36	21.58
Jan-96	8.9	20.30	43.83	108.22	21.97	Aug-01	33.70	21.46	157.02	98.97	21.24
Feb-96	11.2	20.50	54.63	118.40	24.28	Sep-01	29.30	21.43	136.74	99.49	21.32
Mar-96	12.3	20.58	59.77	92.88	19.11	Oct-01	23.55	21.28	110.69	99.04	21.07
Apr-96	16.55	20.59	80.40	90.19	18.57	Nov-01	14.70	21.04	69.88	95.53	20.10
May-96	27.95	20.69	135.10	109.43	22.64	Dec-01	10.80	20.94	51.58	106.05	22.20
Jun-96	29.6	20.97	141.17	96.57	20.25	Jan-02	7.40	20.97	35.29	87.13	18.27
Jul-96	35.25	21.13	166.82	103.88	21.95	Feb-02	10.80	20.91	51.65	111.96	23.41
Aug-96	33.6	20.92	160.61	101.23	21.18	Mar-02	14.60	20.88	69.94	108.68	22.69
Sep-96	27.92	20.63	135.36	98.49	20.31	Apr-02	17.20	20.96	82.08	92.08	19.30
Oct-96	21.86	20.50	106.62	95.40	19.56	May-02	24.30	21.10	115.14	93.27	19.68
Nov-96	16.07	20.47	78.50	107.32	21.97	Jun-02	30.80	21.07	146.17	100.00	21.07
Dec-96	12.54	20.51	61.14	125.71	25.78	Jul-02	34.30	21.05	162.91	101.45	21.36
Jan-97	8.56	20.44	41.87	103.37	21.13	Aug-02	32.60	21.06	154.78	97.55	20.55
Feb-97	6.48	20.23	32.03	69.43	14.04	Sep-02	29.60	20.85	141.97	103.29	21.54
Mar-97	9.96	20.10	49.55	77.00	15.48	Oct-02	25.20	20.77	121.32	108.56	22.55
Apr-97	15.96	20.11	79.36	89.03	17.90	Nov-02	16.60	20.88	79.52	108.72	22.69
May-97	27.78	20.12	138.04	111.82	22.50	Dec-02	8.10	20.96	38.65	79.46	16.65
Jun-97	30.66	19.97	153.52	105.03	20.98	Jan-03	9.70	20.91	46.39	114.54	23.95
Jul-97	32.64	19.74	165.37	102.98	20.33	Feb-03	8.70	20.88	41.67	90.31	18.86
Mar-03	11.60	20.89	55.52	86.28	18.02	Feb-09	12.10	21.53	56.20	121.82	26.23
Apr-03	18.30	20.84	87.82	98.52	20.53	Mar-09	13.40	21.29	62.94	97.79	20.82
May-03	25.70	20.75	123.86	100.33	20.82	Apr-09	18.20	21.23	85.75	96.20	20.42
Jun-03	31.40	20.77	151.20	103.44	21.48	May-09	26.40	21.25	124.21	100.62	21.39
Jul-03	32.50	20.85	155.91	97.08	20.24	Jun-09	32.20	21.32	151.06	103.34	22.03
Aug-03	33.70	20.87	161.47	101.77	21.24	Jul-09	33.80	21.56	156.78	97.63	21.05
Sep-03	28.80	21.10	136.52	99.33	20.95	Aug-09	32.90	21.72	151.47	95.47	20.74
Oct-03	24.70	21.31	115.92	103.72	22.10	Sep-09	27.60	21.86	126.27	91.87	20.08
Nov-03	15.00	21.28	70.51	96.39	20.51	Oct-09	24.40	22.03	110.76	99.11	21.83

(*Continued*)

Table 5. (Continued).

Period	Data	Trend cycle	Seasonality	Irregular	Seasonally adjusted	Period	Data	Trendcycle	Seasonality	Irregular	Seasonally adjusted
Dec-03	10.10	21.20	47.64	97.95	20.76	Nov-09	15.50	22.06	70.26	96.05	21.19
Jan-04	9.60	21.26	45.15	111.47	23.70	Dec-09	12.20	22.06	55.30	113.69	25.08
Feb-04	9.40	21.35	44.04	95.45	20.37	Jan-10	11.80	22.14	53.29	131.58	29.13
Mar-04	16.30	21.41	76.12	118.29	25.33	Feb-10	12.40	22.35	55.49	120.27	26.88
Apr-04	18.70	21.52	86.91	97.50	20.98	Mar-10	16.40	22.66	72.38	112.47	25.48
May-04	24.50	21.53	113.78	92.16	19.85	Apr-10	19.30	22.85	84.46	94.76	21.65
Jun-04	30.80	21.43	143.74	98.34	21.07	May-10	26.10	23.03	113.36	91.82	21.14
Jul-04	34.60	21.28	162.60	101.25	21.55	Jun-10	32.50	23.19	140.14	95.87	22.23
Aug-04	33.60	21.19	158.60	99.96	21.18	Jul-10	35.40	23.08	153.39	95.51	22.04
Sep-04	30.50	21.04	144.98	105.48	22.19	Aug-10	36.20	22.86	158.37	99.82	22.82
Oct-04	25.50	20.99	121.47	108.68	22.82	Sep-10	31.80	22.65	140.37	102.13	23.14
Nov-04	14.60	21.10	69.19	94.59	19.96	Oct-10	24.80	22.55	109.96	98.39	22.19
Dec-04	7.95	21.16	37.57	77.25	16.34	Nov-10	19.30	22.51	85.73	117.21	26.39
Jan-05	8.20	21.23	38.63	95.38	20.25	Dec-10	12.40	22.43	55.27	113.64	25.49
Feb-05	8.55	21.34	40.06	86.82	18.53	Jan-11	8.90	22.37	39.79	98.24	21.97
Mar-05	13.60	21.38	63.62	98.86	21.13	Feb-11	10.00	22.26	44.92	97.36	21.67
Apr-05	20.35	21.28	95.64	107.30	22.83	Mar-11	13.90	22.10	62.90	97.73	21.60
May-05	25.45	21.24	119.81	97.05	20.62	Apr-11	19.40	21.91	88.55	99.34	21.76
Jun-05	31.20	21.51	145.05	99.23	21.34	May-11	25.00	21.53	116.14	94.08	20.25
Jul-05	35.80	21.75	164.57	102.48	22.29	Jun-11	31.70	21.18	149.71	102.42	21.69
Aug-05	35.25	21.84	161.37	101.71	22.22	Jul-11	34.60	21.03	164.55	102.47	21.55
Sep-05	29.65	22.03	134.59	97.93	21.57	Aug-11	34.50	20.95	164.71	103.81	21.74
Oct-05	23.95	22.14	108.18	96.79	21.43	Sep-11	29.60	20.82	142.19	103.46	21.54
Nov-05	15.30	22.19	68.95	94.27	20.92	Oct-11	22.40	20.80	107.71	96.38	20.04
Dec-05	13.70	22.33	61.37	126.16	28.17	Nov-11	12.50	21.00	59.52	81.37	17.09
Jan-06	8.30	22.37	37.10	91.60	20.49	Dec-11	10.80	21.12	51.13	105.13	22.20
Feb-06	10.60	22.38	47.36	102.65	22.97	Jan-12	6.95	21.14	32.87	81.15	17.16
Mar-06	16.00	22.46	71.25	110.71	24.86	Feb-12	10.00	21.14	47.31	102.54	21.67
Apr-06	20.60	22.60	91.16	102.27	23.11	Mar-12	10.80	21.14	51.09	79.38	16.78
May-06	26.40	22.69	116.36	94.26	21.39	Apr-12	22.00	21.28	103.41	116.01	24.68
Jun-06	33.50	22.45	149.22	102.09	22.92	May-12	27.35	21.62	126.51	102.48	22.15
Jul-06	34.60	22.20	155.83	97.03	21.55	Jun-12	32.20	21.86	147.30	100.77	22.03
Aug-06	36.70	22.13	165.81	104.51	23.13	Jul-12	34.65	21.96	157.78	98.25	21.58
Sep-06	30.00	21.99	136.42	99.25	21.83	Aug-12	34.30	22.07	155.39	97.94	21.62
Oct-06	27.00	21.69	124.47	111.37	24.16	Sep-12	29.85	22.28	133.96	97.46	21.72
Nov-06	14.40	21.57	66.76	91.27	19.69	Oct-12	25.40	22.46	113.10	101.20	22.73
Dec-06	8.90	21.57	41.26	84.83	18.30	Nov-12	17.75	22.39	79.28	108.38	24.27
Jan-07	7.20	21.52	33.46	82.62	17.78	Dec-12	11.35	22.30	50.90	104.65	23.33
Feb-07	10.00	21.43	46.66	101.12	21.67	Jan-13	8.80	22.26	39.52	97.58	21.73

Mar-07	13.20	21.37	61.77	95.98	20.51	Feb-13	10.85	22.23	48.81	105.79	23.52
Apr-07	16.20	21.33	75.97	85.22	18.17	Mar-13	15.00	22.22	67.52	104.91	23.31
May-07	27.90	21.32	130.86	106.00	22.60	Apr-13	22.00	22.11	99.51	111.64	24.68
Jun-07	32.00	21.45	149.16	102.04	21.89	May-13	25.70	22.01	116.75	94.57	20.82
Jul-07	34.80	21.46	162.18	100.99	21.67	Jun-13	31.65	21.96	144.10	98.58	21.65
Aug-07	34.50	21.38	161.34	101.69	21.74	Jul-13	34.40	21.92	156.96	97.74	21.42
Sep-07	30.70	21.59	142.21	103.47	22.34	Aug-13	33.70	21.92	153.72	96.89	21.24
Oct-07	25.20	22.10	114.03	102.03	22.55	Sep-13	30.15	21.91	137.62	100.13	21.94
Nov-07	16.10	22.28	72.28	98.81	22.01	Oct-13	22.50	21.83	103.05	92.21	20.13
Dec-07	10.40	22.15	46.96	96.55	21.38	Nov-13	18.35	21.77	84.30	115.24	25.09
Jan-08	5.80	22.13	26.21	64.71	14.32	Dec-13	9.60	21.73	44.18	90.84	19.74
Feb-08	9.60	22.17	43.30	93.85	20.81	Jan-14	9.40	21.64	43.43	107.24	23.21
Mar-08	18.50	22.21	83.29	129.42	28.75	Feb-14	10.40	21.60	48.14	104.34	22.54
Apr-08	23.20	22.12	104.90	117.68	26.03	Mar-14	15.10	21.54	70.09	108.91	23.46
May-08	25.10	22.03	113.94	92.30	20.33	Apr-14	20.10	21.59	93.10	104.44	22.55
Jun-08	31.70	22.02	143.98	98.50	21.69	May-14	26.05	21.60	120.59	97.68	21.10
Jul-08	34.70	22.10	157.01	97.77	21.61	Jun-14	30.30	21.53	140.71	96.26	20.73
Aug-08	35.60	22.30	159.61	100.60	22.44	Jul-14	33.70	21.49	156.78	97.63	20.99
Sep-08	30.60	22.20	137.86	100.31	22.26	Aug-14	33.50	21.42	156.38	98.57	21.11
Oct-08	23.00	21.78	105.63	94.51	20.58	Sep-14	28.90	21.34	135.46	98.55	21.03
Nov-08	16.20	21.62	74.93	102.44	22.15	Oct-14	24.85	21.28	116.77	104.48	22.23
Dec-08	10.00	21.70	46.09	94.76	20.56	Nov-14	16.30	21.29	76.58	104.69	22.28
Jan-09	8.20	21.68	37.82	93.39	20.25	Dec-14	10.00	21.32	46.90	96.43	20.56

Period	Data	Trend cycle	Seasonality	Irregular	Seasonally adjusted
Jan-15	8.07	21.39	37.72	93.13	19.92
Feb-15	9.98	21.44	46.54	100.87	21.63
Mar-15	13.44	21.50	62.52	97.15	20.88
Apr-15	20.47	21.39	95.72	107.38	22.96
May-15	25.78	21.17	121.77	98.64	20.88
Jun-15	31.42				21.50
Jul-15	34.34				21.38
Aug-15	34.06				21.47
Sep-15	29.60				21.54
Oct-15	21.50				19.24
Nov-15	14.50				19.82

Appendix B

Table 6. Estimated autocorrelations function (ACF) and partial autocorrelations function (PACF) for Erbil adjusted mean temperature (°C)

Lag	ACF	Lag	PACF
1	−0.38*	1	−0.38*
2	−0.07	2	−0.25*
3	−0.07	3	−0.25
4	0.08	4	−0.11
5	−0.03	5	−0.11
6	−0.02	6	−0.11
7	0.02	7	−0.06
8	−0.02	8	−0.08
9	0.00	9	−0.07
10	0.00	10	−0.05
11	0.32*	11	0.40
12	−0.49***	12	−0.22**
13	0.12	13	−0.10
14	0.00	14	−0.10
15	0.03	15	−0.16
16	−0.07	16	−0.16
17	0.06	17	−0.09
18	0.04	18	−0.04
19	−0.02	19	−0.03
20	−0.05	20	−0.09
21	0.02	21	−0.09
22	0.08	22	−0.02
23	−0.13	23	0.17
24	0.04	24	−0.12**
25	−0.05	25	−0.16
26	0.07	26	−0.10
27	0.04	27	−0.02
28	0.04	28	−0.01
29	−0.08	29	−0.07
30	−0.01	30	−0.04
31	0.03	31	0.02
32	0.04	32	0.00
33	−0.01	33	−0.02
34	−0.11	34	−0.10
35	0.15	35	0.22
36	−0.12**	36	−0.13**
37	0.16	37	0.02
38	−0.08	38	−0.02
39	−0.06	39	−0.08
40	−0.01	40	−0.01
41	0.05	41	−0.04
42	0.01	42	0.01
43	−0.03	43	0.02
44	0.00	44	0.02
45	0.00	45	0.01
46	0.11	46	0.06
47	−0.26	47	−0.06
48	0.20**	48	−0.10

Note. Table 6 shows the estimated autocorrelations (partial autocorrelations) between values of adjusted mean temperature (°C) at various lags. The lag k autocorrelation (partial autocorrelations) coefficient measures the correlation between values of adjusted mean temperature at time t and time $t - k$ ($t + k$).
*Nonseasonal terms.
**Seasonal terms.
***Nonseasonal and seasonal terms.

Appendix C

Table 7A. SARIMA$(3,1,3)(3,1,3)_{12}$ model terms selection procedures, for SARIMA model coefficient summary, (a)–(j): Starting from SARIMA$(3,1,3)(3,1,3)_{12}$, we arrive at our final model SARIMA$(0,1,2)(0,1,1)_{12}$ by, at each step, dropping the term with the highest p value associated to it and reestimating the remaining parameters until all p values for all estimated parameters are below 0.05.

SARIMA$(3,1,3)(3,1,3)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.368	1.181	0.755
AR(2)	−0.192	0.889	0.830
AR(3)	0.112	0.379	0.768
MA(1)	0.183	1.173	0.876
MA(2)	0.214	1.081	0.843
MA(3)	0.399	0.911	0.662
SAR(1)	−0.282	2.049	0.891
SAR(2)	0.050	1.566	0.974
SAR(3)	−0.029	0.071	0.683
SMA(1)	0.658	2.046	0.748
SMA(2)	0.289	1.595	0.856
SMA(3)	−0.036	1.460	0.980

(a)

SARIMA$(3,1,3)(3,1,2)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.359	1.291	0.781
AR(2)	−0.192	0.959	0.841
AR(3)	0.118	0.409	0.772
MA(1)	0.194	1.282	0.880
MA(2)	0.210	1.184	0.859
MA(3)	0.404	0.986	0.682
SAR(1)	−0.274	1.884	0.885
SAR(2)	0.014	0.065	0.827
SAR(3)	−0.025	0.073	0.734
SMA(1)	0.663	1.883	0.725
SMA(2)	0.243	1.754	0.890

(b)

SARIMA$(3,1,3)(3,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.535	1.446	0.712
AR(2)	−0.167	0.941	0.860
AR(3)	0.149	0.449	0.741
MA(1)	0.019	1.435	0.989
MA(2)	0.331	1.153	0.774
MA(3)	0.457	1.052	0.665
SAR(1)	−0.016	0.063	0.798
SAR(2)	0.015	0.060	0.805
SAR(3)	−0.008	0.060	0.895
SMA(1)	0.925	0.019	0.000

(c)

SARIMA$(3,1,3)(2,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.538	1.359	0.693
AR(2)	−0.164	0.900	0.856
AR(3)	0.144	0.427	0.736
MA(1)	0.016	1.349	0.991
MA(2)	0.335	1.089	0.759
MA(3)	0.451	1.001	0.653
SAR(1)	−0.014	0.063	0.820
SAR(2)	0.019	0.062	0.762
SMA(1)	0.927	0.017	0.000

(d)

SARIMA$(3,1,3)(1,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.439	1.532	0.775
AR(2)	−0.195	1.038	0.851
AR(3)	0.143	0.467	0.761
MA(1)	0.117	1.522	0.939
MA(2)	0.250	1.290	0.846
MA(3)	0.445	1.123	0.692
SAR(1)	−0.018	0.066	0.782
SMA(1)	0.923	0.018	0.000

(e)

SARIMA$(3,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.364	1.424	0.798
AR(2)	−0.189	1.012	0.852
AR(3)	0.123	0.430	0.775
MA(1)	0.192	1.416	0.892
MA(2)	0.214	1.273	0.867
MA(3)	0.406	1.049	0.699
SMA(1)	0.925	0.017	0.000

(f)

SARIMA$(2,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value

SARIMA$(1,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value

SARIMA$(0,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value

(*Continued*)

Table 7A. (Continued).

SARIMA(3,1,3)(3,1,3)$_{12}$				SARIMA(3,1,3)(3,1,2)$_{12}$				SARIMA(3,1,3)(3,1,1)$_{12}$			
Parameter	Estimate	Standard error	p value	Parameter	Estimate	Standard error	p value	Parameter	Estimate	Standard error	p value
AR(1)	−0.397	0.759	0.601	AR(1)	−0.242	0.521	0.642	MA(1)	0.560	0.060	0.000
AR(2)	0.099	0.336	0.769	MA(1)	0.319	0.516	0.537	MA(2)	0.194	0.068	0.005
MA(1)	0.162	0.757	0.831	MA(2)	0.336	0.321	0.298	MA(3)	0.115	0.058	0.055
MA(2)	0.511	0.718	0.477	MA(3)	0.171	0.139	0.218	SMA(1)	0.926	0.017	0.000
MA(3)	0.137	0.164	0.404	SMA(1)	0.927	0.017	0.000				
SMA(1)	0.932	0.014	0.000								
	(g)				(h)				(i)		

SARIMA(0,1,2)(0,1,1)$_{12}$			
Parameter	Estimate	Standard error	p value
MA(1)	0.588	0.059	0.000
MA(2)	0.237	0.058	0.000
SMA(1)	0.933	0.016	0.000
	(j)		

Table 7B. $(2, 1, 3)(3, 1, 3)_{12}$ model terms selection procedures.

(a) SARIMA$(2,1,3)(3,1,3)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.335	0.748	0.655
AR(2)	0.012	0.313	0.970
MA(1)	0.220	0.745	0.768
MA(2)	0.397	0.672	0.555
MA(3)	0.176	0.163	0.281
SAR(1)	−0.281	2.146	0.896
SAR(2)	0.056	1.557	0.971
SAR(3)	−0.027	0.075	0.723
SMA(1)	0.662	2.142	0.757
SMA(2)	0.292	1.588	0.854
SMA(3)	−0.039	1.453	0.979

(b) SARIMA$(2,1,3)(3,1,2)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.401	0.739	0.588
AR(2)	0.077	0.326	0.814
MA(1)	0.154	0.737	0.835
MA(2)	0.496	0.689	0.472
MA(3)	0.154	0.163	0.343
SAR(1)	−0.260	1.692	0.878
SAR(2)	0.016	0.064	0.805
SAR(3)	−0.026	0.073	0.720
SMA(1)	0.682	1.690	0.687
SMA(2)	0.231	1.590	0.885

(c) SARIMA$(2,1,3)(3,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.420	0.707	0.553
AR(2)	0.145	0.335	0.666
MA(1)	0.139	0.705	0.844
MA(2)	0.570	0.691	0.410
MA(3)	0.117	0.167	0.484
SAR(1)	−0.013	0.061	0.828
SAR(2)	0.019	0.059	0.754
SAR(3)	−0.005	0.060	0.932
SMA(1)	0.929	0.018	0.000

(d) SARIMA$(2,1,3)(2,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.425	0.686	0.536
AR(2)	0.134	0.326	0.681
MA(1)	0.133	0.684	0.846
MA(2)	0.562	0.667	0.400
MA(3)	0.124	0.166	0.455
SAR(1)	−0.012	0.061	0.843
SAR(2)	0.022	0.061	0.715
SMA(1)	0.931	0.016	0.000

(e) SARIMA$(2,1,3)(1,1,1)_{12}$

Parameter	Estimate	Standard errorr	p value
AR(1)	−0.425	0.682	0.534
AR(2)	0.144	0.330	0.663
MA(1)	0.135	0.681	0.843
MA(2)	0.571	0.669	0.394
MA(3)	0.117	0.168	0.486
SAR(1)	−0.017	0.063	0.787
SMA(1)	0.929	0.016	0.000

(f) SARIMA$(2,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.397	0.759	0.601
AR(2)	0.099	0.336	0.769
MA(1)	0.162	0.757	0.831
MA(2)	0.511	0.718	0.477
MA(3)	0.137	0.164	0.404
SMA(1)	0.932	0.014	0.000

(g) SARIMA$(1,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
AR(1)	−0.242	0.521	0.642
MA(1)	0.319	0.516	0.537
MA(2)	0.336	0.321	0.298
MA(3)	0.171	0.139	0.218
SMA(1)	0.927	0.017	0.000

(h) SARIMA$(0,1,3)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
MA(1)	0.560	0.060	0.000
MA(2)	0.194	0.068	0.005
MA(3)	0.115	0.058	0.055
SMA(1)	0.926	0.017	0.000

(i) SARIMA$(0,1,2)(0,1,1)_{12}$

Parameter	Estimate	Standard error	p value
MA(1)	0.588	0.059	0.000
MA(2)	0.237	0.058	0.000
SMA(1)	0.933	0.016	0.000

Note. SARIMA model coefficient summary, (a)–(i): Starting from SARIMA$(2,1,3)(3,1,3)_{12}$, we arrive at our final model SARIMA$(0,1,2)(0,1,1)_{12}$ by, at each step, dropping the term with the highest p value associated to it and reestimating the remaining parameters until all p values for all estimated parameters are below 0.05.

Appendix D

Table 8. Forecast value of Erbil mean from December 2015 to December 2025.

Number of the month	Period	Forecast	Lower 95% limit	Upper 95% limit	Number of the month	Period	Forecast	Lower 95% limit	Upper 95% limit
288	Dec-15	10.03	7.93	12.69	333	Sep-19	29.06	19.47	43.39
289	Jan-16	8.38	6.50	10.80	334	Oct-19	23.37	15.60	34.99
290	Feb-16	9.83	7.60	12.72	335	Nov-19	15.42	10.26	23.16
291	Mar-16	13.52	10.41	17.55	336	Dec-19	10.18	6.74	15.38
292	Apr-16	18.94	14.54	24.67	337	Jan-20	8.43	5.56	12.78
293	May-16	25.58	19.58	33.43	338	Feb-20	9.89	6.50	15.04
294	Jun-16	30.67	23.41	40.20	339	Mar-20	13.60	8.92	20.75
295	Jul-16	33.70	25.63	44.30	340	Apr-20	19.06	12.45	29.17
296	Aug-16	33.38	25.31	44.02	341	May-20	25.75	16.77	39.53
297	Sep-16	28.93	21.87	38.26	342	Jun-20	30.87	20.04	47.55
298	Oct-16	23.26	17.53	30.85	343	Jul-20	33.91	21.95	52.40
299	Nov-16	15.35	11.53	20.42	344	Aug-20	33.59	21.67	52.07
300	Dec-16	10.14	7.57	13.56	345	Sep-20	29.11	18.72	45.26
301	Jan-17	8.39	6.25	11.27	346	Oct-20	23.40	15.00	36.50
302	Feb-17	9.84	7.30	13.27	347	Nov-20	15.44	9.87	24.16
303	Mar-17	13.54	10.01	18.31	348	Dec-20	10.20	6.49	16.04
304	Apr-17	18.97	13.98	25.73	349	Jan-21	8.44	5.35	13.33
305	May-17	25.62	18.83	34.87	350	Feb-21	9.91	6.25	15.69
306	Jun-17	30.72	22.51	41.93	351	Mar-21	13.62	8.57	21.65
307	Jul-17	33.75	24.65	46.21	352	Apr-21	19.09	11.97	30.43
308	Aug-17	33.43	24.35	45.92	353	May-21	25.79	16.12	41.24
309	Sep-17	28.97	21.03	39.91	354	Jun-21	30.92	19.27	49.60
310	Oct-17	23.29	16.86	32.18	355	Jul-21	33.96	21.10	54.66
311	Nov-17	15.37	11.09	21.30	356	Aug-21	33.64	20.84	54.32
312	Dec-17	10.15	7.29	14.14	357	Sep-21	29.15	18.00	47.22
313	Jan-18	8.40	6.01	11.75	358	Oct-21	23.44	14.43	38.08
314	Feb-18	9.86	7.03	13.83	359	Nov-21	15.47	9.49	25.21
315	Mar-18	13.56	9.63	19.09	360	Dec-21	10.22	6.24	16.73
316	Apr-18	19.00	13.45	26.83	361	Jan-22	8.46	5.14	13.91
317	May-18	25.66	18.12	36.36	362	Feb-22	9.92	6.01	16.37
318	Jun-18	30.77	21.65	43.73	363	Mar-22	13.65	8.24	22.59
319	Jul-18	33.80	23.72	48.19	364	Apr-22	19.12	11.51	31.75
320	Aug-18	33.49	23.42	47.88	365	May-22	25.83	15.50	43.03
321	Sep-18	29.02	20.23	41.62	366	Jun-22	30.96	18.52	51.77
322	Oct-18	23.33	16.22	33.56	367	Jul-22	34.02	20.28	57.05
323	Nov-18	15.39	10.67	22.21	368	Aug-22	33.70	20.03	56.70
324	Dec-18	10.17	7.01	14.75	369	Sep-22	29.20	17.30	49.29
325	Jan-19	8.42	5.78	12.25	370	Oct-22	23.47	13.86	39.75
326	Feb-19	9.87	6.76	14.42	371	Nov-22	15.49	9.12	26.31
327	Mar-19	13.58	9.27	19.90	372	Dec-22	10.23	5.99	17.47
328	Apr-19	19.03	12.94	27.97	373	Jan-23	8.47	4.94	14.52
329	May-19	25.70	17.43	37.91	374	Feb-23	9.94	5.78	17.09
330	Jun-19	30.82	20.83	45.59	375	Mar-23	13.67	7.92	23.58
331	Jul-19	33.86	22.82	50.24	376	Apr-23	19.15	11.06	33.15
332	Aug-19	33.54	22.53	49.92	376	May-23	25.87	14.89	44.93

Number of the month	Period	Forecast	Lower 95% limit	Upper 95% limit
378	Jun-23	31.01	17.79	54.05
379	Jul-23	34.07	19.48	59.58
380	Aug-23	33.75	19.24	59.21
381	Sep-23	29.25	16.62	51.48
382	Oct-23	23.51	13.31	41.52
383	Nov-23	15.51	8.76	27.48
384	Dec-23	10.25	5.76	18.25
385	Jan-24	8.48	4.75	15.16
386	Feb-24	9.95	5.55	17.85
387	Mar-24	13.69	7.61	24.64

(Continued)

Table 8. (Continued).

Number of the month	Period	Forecast	Lower 95% limit	Upper 95% limit
388	Apr-24	19.18	10.62	34.63
389	May-24	25.91	14.30	46.94
390	Jun-24	31.06	17.08	56.47
391	Jul-24	34.12	18.71	62.25
392	Aug-24	33.80	18.47	61.87
393	Sep-24	29.29	15.95	53.79
394	Oct-24	23.55	12.78	43.39
395	Nov-24	15.54	8.41	28.72
396	Dec-24	10.26	5.52	19.07
397	Jan-25	8.50	4.56	15.85
398	Feb-25	9.97	5.33	18.66
399	Mar-25	13.71	7.30	25.75
400	Apr-25	19.21	10.19	36.20
401	May-25	25.95	13.72	49.07
402	Jun-25	31.11	16.39	59.03
403	Jul-25	34.18	17.95	65.08
404	Aug-25	33.85	17.72	64.68
405	Sep-25	29.34	15.30	56.24
406	Oct-25	23.59	12.26	45.36
407	Nov-25	15.56	8.06	30.03
408	Dec-25	10.28	5.30	19.94

Note. Table 8 shows the forecast values for mean temperature (°C) from December 2015 to December 2025.