Comparison of SARFIMA and LSTM methods to model and to forecast Canadian temperature

Sami Khedhiri
School of Mathematical and Computational Sciences
University of Prince Edward Island, Canada
E-mail: skhedhiri@upei.ca

An empirical study is performed using seasonal autoregressive fractionally integrated moving average (SARFIMA) time series models and long short-term memory (LSTM) methods from deep learning to evaluate their statistical performance for the fit and prediction of temperature variables. These two methods are applied to monthly minimum and maximum air temperature data collected from four Canadian stations, covering the eastern, western, and central regions of Canada. It is a generally accepted fact that accurate temperature forecasting is important to everyone and stakeholders in several economic activities including the tourism, agriculture, and energy sectors rely on this information. Therefore, it is crucial to select the best methods which provide appropriate statistical modelling and an accurate prediction of temperature. The results of this study show that deep learning LSTM models have better fit and smaller root-mean-square errors compared to time series SARFIMA models.

Keywords:
SARFIMA,
long short-term memory,
root-mean-square error,
temperature prediction

Introduction

The importance of accurate temperature forecasting is crucial and unequivocal for decision-making and risk management at the regional level, as shown by Lai–Dzombak (2020). Several scholars have used statistical models and time series methods to forecast climate variables (Papacharalampous et al. 2018, Raible et al. 1999, Agbo 2021, Kirkwood et al. 2021). Furthermore, it is suggested that climate change is predominantly governed by changes in temperature and precipitation, and that research focusing on the assessment of climate change is often based on the statistical analysis of temperature and precipitation time series (Dimri et al. 2020). However, there are a few alternative statistical methods that can also be applied, and several studies have shown that they can produce good results for modelling and forecasting weather variables. These methods include regression analysis (Massie–Rose 1997), autoregressive integrated moving average models, genetic algorithms, adaptive splines threshold autoregressive, and the K-nearest neighbour method (Shankarnarayan–Ramakrishna 2021). In addition, methods based on deep neural networks have also been used to forecast climate variables (Datta et al. 2019, Park et al. 2019). One technique that has been particularly popular in deep learning neural networks is called the long short-term memory (LSTM) method and is based on an advanced deep learning model which was initially applied to tackle complex problems such as natural language processing (Yao–Guan 2018), Sutskever et al. (2014) introduced computational methods to analyse speech, captioning, and handwriting recognition. Khedhiri (2021) performed a statistical analysis based on the LSTM neural network method to forecast COVID-19 infections. In fact, the long short-term memory model was originally introduced by Hochreiter–Schmidhuber (1997) to overcome the vanishing gradient problems in recurrent neural networking (RNN). Recently, however, this method has been widely used in climate studies as well. For instance, Tran Anh et al. (2019) applied an LSTM network to fit rainfall series and evaluate the prediction accuracy. Lee et al. (2020) studied the LSTM model with rainfall data and conducted a rainfall-runoff analysis, and Gonzalez–Yu (2018) introduced a novel LSTM neural network that they used in nonlinear system modelling. Dong et al. (2018) and Xiaoyun et al. (2016) used recurrent neural networks to predict wind power. In a recent study, Kreuzer et al. (2020) applied a convolutional LSTM neural network and compared its forecast accuracy with the univariate LSTM neural network and the seasonal autoregressive integrated moving average (SARIMA) model for temperature data related to German stations. Their results show better performance of the multivariate LSTM and convolutional LSTM networks compared to the univariate LSTM network and the short-memory SARIMA model. In this study, we investigate the long-term memory dependence of climate data by considering the estimation of the model with possible fractional differencing. The data set in this study includes thousands of monthly observations and spans for over 100 years; thus, it would be natural to look beyond short-memory ARMA-type models in order to obtain a better identification of the statistical characteristics of the temperature series. This allows us to use methods which produce accurate long-term monthly forecasts.

Furthermore, several scholars have used SARFIMA models in their studies. This includes Qi et al. (2020), who compared the accuracy of SARIMA and SARFIMA models for predicting the incidence of haemorrhagic fever with renal syndrome, and Mostafaei–Sakhabakhsh (2012), who used the SARFIMA model to study and forecast Iran's oil supply.

In this paper, we are interested in the analysis of two alternative methods to fit and forecast monthly minimum and maximum temperature series with historical data collected from the late nineteenth century to the present time. The datasets are from four Canadian stations: Calgary, Montreal, Toronto, and Vancouver. The first method employs SARFIMA time series models, and the second method is based on LSTM modelling. Using programs with built-in functions in MATLAB and R coding, we performed an empirical statistical analysis of these methods and compared their fit and prediction accuracy for the air temperature data.

Material and methods

a) SARFIMA method

Autoregressive integrated moving average models, also called ARIMA models, represent some of the most widely used models and forecasting methods for time series data. Although the method can handle data with a trend, it does not support a time series with a seasonal component (Browniee 2018). Their extension to ARIMA models, which supports the direct modelling of the seasonal component of the series, is called seasonal ARIMA or SARIMA models (Bell et al. 2012). Time series ARIMA models are widely used (Wibowo et al. 2021). However, despite their popularity and their applications in numerous studies, these models are only limited to either stationary *I(0)* or integrated *I*(*d*) time series after an integer number of differencing, and they are not suited to handle the long-run dependence of the series which would require fractional differencing to obtain a stationary time series. These long-memory processes, called ARFIMA models, are stationary with autocorrelation functions that decay more slowly than the ARMA short-memory model. They were introduced by Granger–Joyeux (1980) to provide a parsimonious parameterisation of long-memory processes that nests the autoregressive moving average model which is widely used for short-memory time series. In these models, the fractional degree of integration provides a natural generalisation of ARIMA models and allows the modelling of long-run effects which only die out at long horizons. Therefore, using different parameters for different types of dependence would facilitate the estimation and would make the interpretation of the statistical results much clearer, as discussed in Sowell (1992).

However, fractional differencing (or integration) alone does not cover seasonal time series; thus, seasonal ARFIMA models (SARFIMA) were developed to deal with seasonality. These models can be represented by a seasonal time series (y_t) as follows:

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D(y_t-\mu)=\theta(B)\Theta(B^s)\varepsilon_t \qquad (1)$$

where μ is the mean of the time series, ε_t is a white noise process with zero mean, s denotes the seasonal period which is 12 for our monthly data, d is the annual differencing, and D is the seasonal differencing. B refers to the backshift operator and is used to define lagged observations, and ϕ, Φ, θ, and Θ are polynomials in B with degrees p, P, q, and Q, respectively. Furthermore, a time series y_t given by the SARFIMA $(p, d, q)(P, D, Q)_s$ model is stationary if $(d+D) < 0.5$ and $D < 0.5$. In addition, the series is said to have a long memory if $0 < (d+D) < 0.5$ with $0 < D < 0.5$. Negative values, that is, $-0.5 < d+D < 0$ and $-0.5 < D < 0$ the time series, are said to have intermediate memory (Bisognin–Lopes 2009). Note also that if, $d > 0.5$ or $D > 0.5$ then y_t is nonstationary. If, $d = D = 0$ then we have the traditional seasonal ARMA model. A review on the practical use of these models and their statistical analysis in R can be found in (Contreras-Reyes–Palma 2013).

b) Long short-term memory method

LSTM networks are an extension of recurrent neural networks which allow us to remember and account for past data in memory and to resolve the issue of vanishing gradient that characterizes RNN. An LSTM model would generally be well-suited to fit and predict most time series given time lags of unknown duration. It trains the model using back-propagation and can remember long-term dependence using memory cells and gates. The LSTM can read, write, and delete information from its memory which can be thought of as a gated cell that decides either to store or delete information based on its importance. The importance assessment occurs through the weights learned by the algorithm. There are three gates in the LSTM network: input, forget, and output, and they can perform back-propagation. LSTM has the advantages of short training time and high accuracy. Following DiPietro–Hager (2020), an LSTM model can be defined as follows:

$$\begin{aligned}
for_t &= \sigma(y_t W_1^{for} + s_{t-1} W_2^{for} + b_{for}) \\
inp_t &= \sigma(y_t W_1^{inp} + s_{t-1} W_2^{inp} + b_{inp}) \\
out_t &= \sigma(y_t W_1^{out} + s_{t-1} W_2^{out} + b_{out}) \\
M_t &= \sigma(for_t * M_{t-1} + inp_t * H_t) \\
s_t &= \tanh(M_t) * out_t \\
H_t &= \tanh(y_t W_1^{g} + s_{t-1} W_2^{g} + b_H)
\end{aligned} \tag{2}$$

where *inp*, *out*, and denote the input, output, and forget gates, respectively, and y and s refer to the number of input and hidden states, respectively. W_1 and W_2 are weight matrices which are adjusted during the network learning phase, and b is the bias. H is the candidate hidden state, and M is the unit internal memory. In addition, *tanh* and σ are the hyperbolic tangent activation function and the sigmoid layer, respectively, and „*" refers to the point-wise multiplication of two vectors. This type of deep learning network is popular, and several scholars have used it in several applications. Chandra (2021) applied this type of deep learning neural network to forecast COVID-19 infections in India, and Yu et al. (2019) provided a detailed review of recurrent neural networks and LSTM models. In addition, Park et al. (2019) introduced a temperature prediction method with a missing data refinement model on a LSTM network.

We use the adaptive moment estimation optimiser which is an extension of stochastic gradient descent. ADAM is based on adaptive moment estimation and has two parts: an adaptive gradient algorithm and root-mean-square propagation. An application of LSTM methods with an ADAM optimiser to predict wind power generation and temperature with data from Estonia can be found in Misha et al. (2019). The ADAM optimiser computes the individual adaptive learning rates for different parameters from estimates of the first and second moments of the gradient (Kingman–Ba 2015).

Monthly temperature records were collected from four major Canadian cities. As an example, Figure A1 in Appendix shows the minimum and maximum temperature records for Calgary. All the datasets are freely available online from Government of Canada website [1] and cover the time period from November 1881 to July 2012 for Calgary, from July 1871 to September 2016 for Montreal, from March 1840 to June 2013 for Toronto, and from October 1898 to June 2013 for Vancouver. These four cities are the largest in Canada and cover the eastern, western, and central regions of the country. In addition, it should be indicated that the collected temperature records were measured in degrees Celsius.

Results and discussion

We start the first part of our empirical analysis by fitting a seasonal ARIMA model to each temperature series, and the results are reported in column 2 of Table 1 which shows the estimates of model (1) parameters. The function *auto.arima* in R allows us to find the best model in terms of lag length according to the Bayesian information criterion (BIC).

Table 1

Selected SARIMA models based on BIC and seasonal unit root test

Temperature series	SARIMA models	Canova and Hansen seasonal unit root test
Maximum in Calgary	(2,0,0) (2,1,0) (12)	6.06 (*)
Minimum in Calgary	(1,0,0) (2,1,0) (12)	7.37 (*)
Maximum in Montreal	(1,0,2) (0,1,1) (12)	2.67
Minimum in Montreal	(1,0,4) (0,1,1) (12)	5.28 (*)
Maximum in Toronto	(1,0,0) (2,1,0) (12)	4.53 (*)
Minimum in Toronto	(1,0,1) (2,1,0) (12)	5.46 (*)
Maximum in Vancouver	(2,0,2) (1,1,1) (12)	6.91 (*)
Minimum in Vancouver	(3,0,3) (2,1,1) (12)	5.67 (*)

Figure A2 in Appendix shows the persistence of the autocorrelation function for the temperature series typical of seasonal unit root presence in the data. To further investigate the non-stationarity characteristics of the series, we computed the KPSS test [1](Kwiatkowski et al. 1992), and the results (test values are not reported) show no unit root. In addition, we computed the Canova and Hansen test for seasonal stability (Canova–Hansen 1995) and the results are reported in the third column of Table 1. The starred values imply rejection of the null hypothesis of no seasonal unit root at the 5% significance level. Therefore, except for the maximum temperature in Montreal, all other series exhibit seasonal unit roots (seasonal instability). This

[1] The abbreviation stands for the names of the four authors who developed the test: Kwiatkowski, Phillips, Schmidt, Shin.

finding also confirms our previous selection of SARIMA models with seasonal unit root specification for monthly temperature data.

Next, we decompose each temperature data into its three time-series components: seasonal, trend, and random components (S_t, TR_t, u_t). The idea is to check for fractional differencing in each component of the series. We analysed the memory of each by checking the differencing parameter. The reported values in Table 2 are estimated from the model with the lowest p-value of the Box-Ljung test for model residuals. Using R packages *arfima* (Veenstra–McLeod 2017) and *ArfimaMLM* (Kraft et al. 2017), we perform the computations of the fractional difference parameters for the trend and seasonal components of each temperature series, and we compute the forecasts and forecast errors from the selected models. The results in Table 2 show that most of the seasonal fractional differencing parameters are greater than 0.5 which indicates the non-stationarity of the series seasonal components.

Figure A3 in Appendix shows the time-series decomposition of temperature data in Calgary, as an illustration, and Figure A4 in Appendix displays the slow decay of the autocorrelation function (ACF) for the seasonal component, whereas the trend and the random components have a faster decaying ACF. Next, we apply the ARFIMA method to the trend and the random components to obtain a more accurate and detailed identification of the time series and to check for the source of non-stability in the series. Furthermore, we compute the fractional differencing parameters for the trend and random components, and the results are reported in Table 2.

Table 2

Fractional differencing parameters for trend, seasonal, and random components

Series	Calgary		Montreal		Toronto		Vancouver	
	maximum	minimum	maximum	minimum	maximum	minimum	maximum	minimum
	temperature							
TR_t	0.416	0.244	0.386	0.407	0.590	0.377	0.098(*)	0.391
u_t	0.047(*)	0.021(*)	0.003(*)	0.048(*)	0.016(*)	0.021(*)	0.004(*)	0.093(*)
S_t	0.659	0.810	0.897	0.754	0.561	0.487	0.782	0.512

(*) Non-significance at 5%.

It is shown that for most temperature series, the trend component exhibits a long memory because most values of the differencing parameter are significantly greater than zero (except for the maximum temperature in Vancouver). However, the random components are short-memory processes with non-statistically significant differencing parameters. Thus, our statistical analysis shows an important finding that the temperature series have seasonal unit roots and that their trend components have a long memory. This result means that we can expect more irregular and volatile seasonal temperatures in the future.

To estimate and validate the models and assess their predictions, each temperature data is divided into two parts, one for the estimation and the other for forecasting. Because the number of available observations differs between the weather stations, we chose the subsample for model prediction to be approximately ten percent of the total sample. Table 3 lists the exact dates of the data subsamples.

Table 3

Data ranges for modelling and forecasting

City	For modelling	For forecasting
Calgary	November 1881 – May 1999	June 1999 – July 2012
Montreal	July 1871 – February 2012	March 2012 – July 2016
Toronto	March 1840 – January 1996	February 1996 – June 2013
Vancouver	October 1898 – August 2001	September 2001 – June 2013

We used MATLAB programs to fit each dataset to LSTM models, and we used available R coding to fit these temperature data to SARFIMA models. Next, the root-mean-square errors were computed by comparing the observed and predicted values of the data. For Calgary, the total sample ranges from November 1881 to July 2021 for a total of 1586 monthly records. There were 1742 monthly observations for Montreal from July 1871 to July 2016, 2079 observations for Toronto from March 1840 to June 2013, and 1412 temperature records for Vancouver from October 1898 to June 2013.

Next, for each temperature data, the selected SARFIMA model was estimated, and its forecasts and forecast errors were computed for the maximum and minimum temperature data available from the four Canadian stations. The estimated models for the Calgary data are as follows:

$SARFIMA(2,0.416,0)(2,0.659,0)_{12}$ for maximum temperatures and

$SARFIMA(1,0.244,0)(2,0.81,0)_{12}$ for minimum temperatures.

Tables 1 and 2 report the estimated model parameters for the remaining series.

In addition, we investigate the fit of the LSTM model for each dataset and assess its prediction performance. Our methodology for the LSTM application can be described as follows: For each minimum and maximum temperature variable, we partition the training, and we test the data. The training is performed on the first 90% of the series, the test is on the last 10%, and the data need to be standardised to zero mean and unit variance to prevent the training from diverging. To produce one-step ahead forecasting, the training sequence with values shifted by one time step is specified as the response. Thus, at each time step of the input sequence, the LSTM network, whose layer has 200 hidden units, learns to predict the values of the following time step. Next, we used the ADAM optimiser and trained it for 250 epochs. To find multi-step-ahead forecasts, we use the previous prediction as a function in a MATLAB program. The following step consists of updating the network state with the observed values instead of the predicted values and computing the forecast errors for each series. As an illustration, Figure A4 in

Appendix shows the monthly maximum and minimum temperature data for the city of Calgary, and Figure A5 in Appendix displays the observed and forecasted values of Calgary temperatures with the LSTM network. Figure A6 in Appendix shows the forecast errors of the selected SARFIMA models for the Calgary data. The root-mean-square errors for each series are listed in Table 4. This shows that the fit and prediction performance of the LSTM model are superior to those of the SARIMA model with lower RMSE (root-mean-square error) values.

Table 4

Root-mean-square error for temperature predictions

Series	Calgary		Montreal		Toronto		Vancouver	
	max	min	max	min	max	min	max	min
SARFIMA	3.93	3.28	3.06	2.89	3.10	2.75	2.09	2.18
LSTM	2.76	2.13	2.18	1.98	2.02	1.83	1.74	1.62

It can be seen from the results in Table 4 that the LSTM neural network with its ability to account better for the dependence of the observations, without a potential misspecification regarding either component of the time series, fits better and outperforms the traditional time series SARIMA models because the reported root-mean-square forecast error values are notably lower for most temperature series. The higher forecast errors cast by the seasonal ARFIMA models can be explained by the difficulty in correctly specifying the model and identifying the lag lengths and the differencing parameters in each component of the time series.

Conclusion

Statistical model-based temperature forecasting has recently gained a lot of interest, and researchers have attempted to develop alternative methods which provide accurate temperature predictions in the short and long terms. This study contributes to this topic by conducting an empirical analysis in order to compare the performance and to investigate the prediction accuracy of two competing statistical methods which are popular and have been successfully used in several applied studies in meteorology and other areas. In the first method, we use seasonal time series autoregressive fractionally integrated moving average models to fit the minimum and maximum temperature series collected from historical data for four Canadian cities. In the second method, we apply a LSTM deep learning network to the same data. The results of this study show that the LSTM model has a better fit and has lower root-mean-square forecast errors and thus it outperforms the SARFIMA time series model in terms of fit and prediction accuracy. However, the time series SARFIMA model, with its ability to accurately identify the potential correlation between long-range observations, revealed important statistical properties of the temperature data, which could not be depicted by the LSTM model. This shows the non-stationarity of the seasonal components of most temperature data, which means that we can expect more variation in the seasonal temperatures.

Appendix

Figure A1

Monthly maximum and minimum temperature in Calgary
(November 1881 to July 2012)

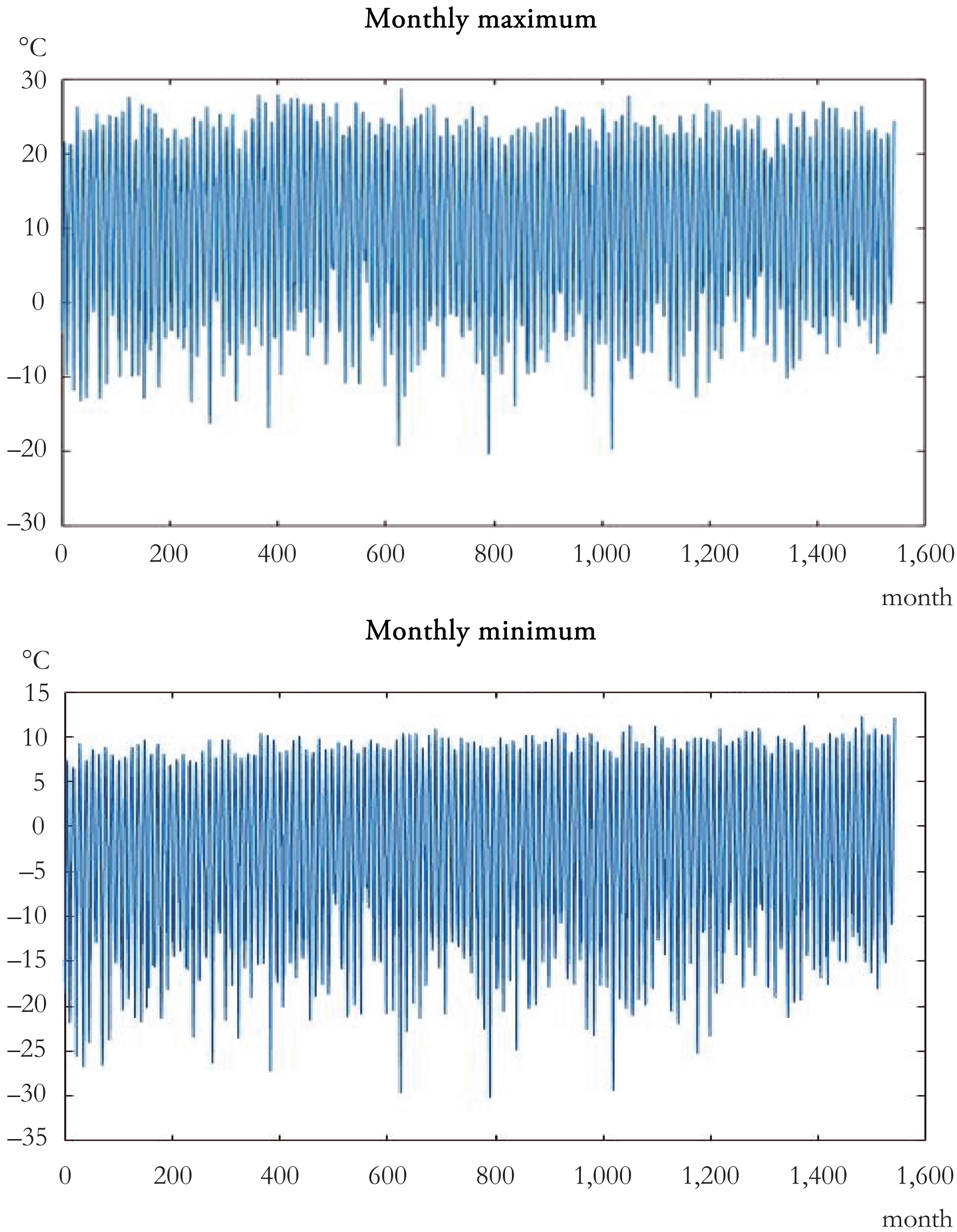


Figure A2

ACF and PACF for maximum temperature series

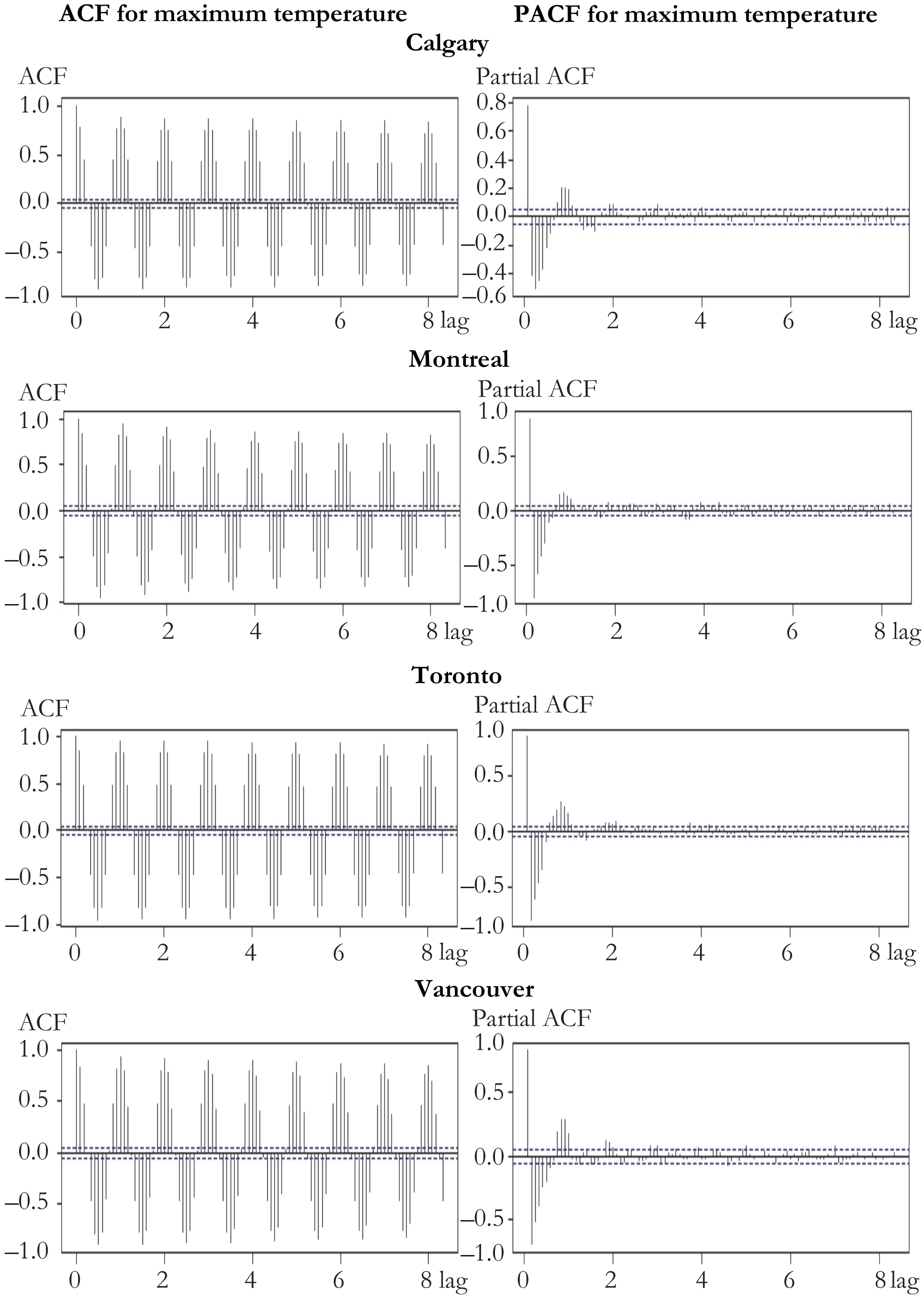


Figure A3

Time series decomposition of monthly maximum temperature in Calgary

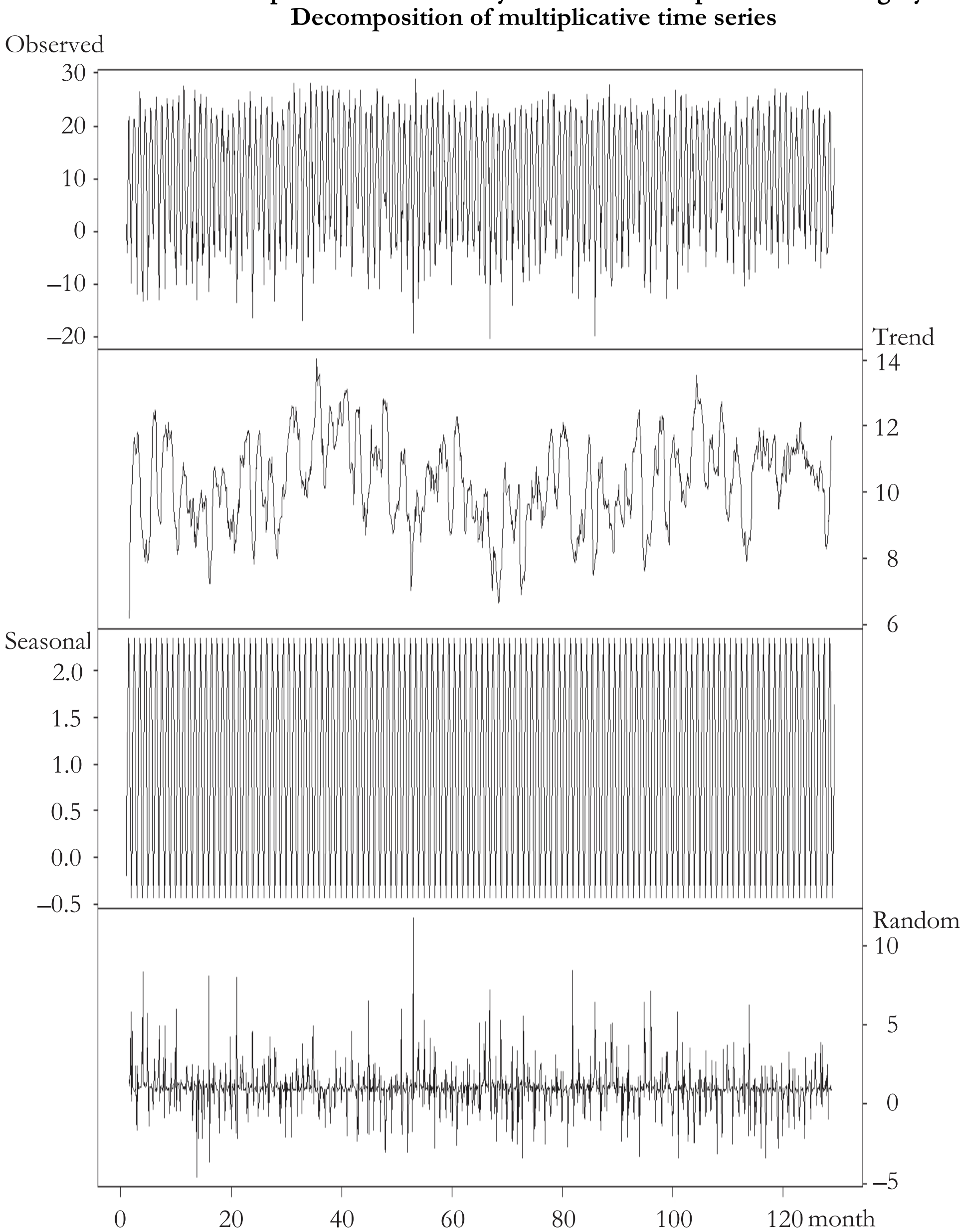


Figure A4

ACF for the components of monthly maximum temperature in Calgary

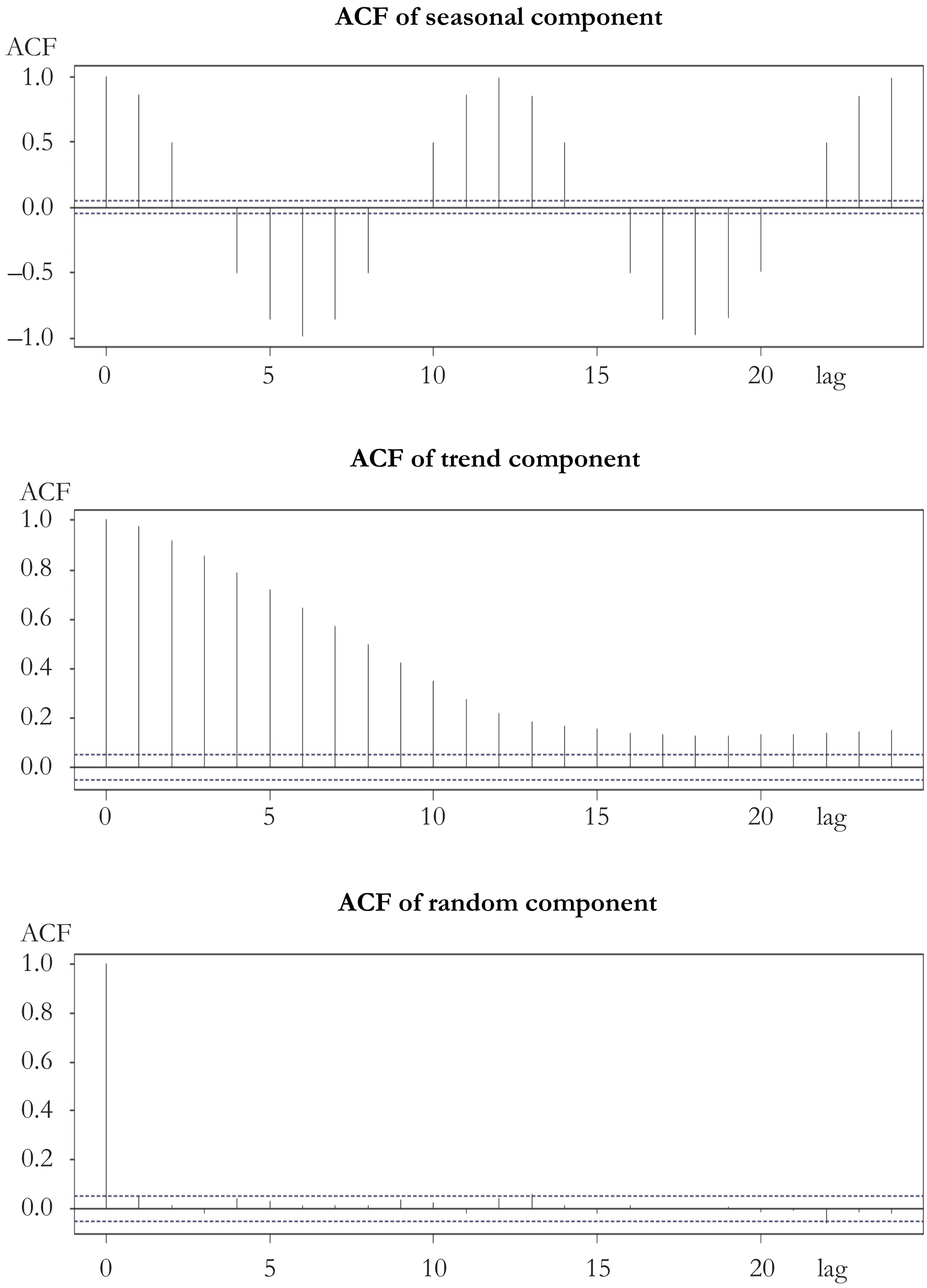


Figure A5

Calgary temperature updated forecast errors with LSTM network
(Forecast range: June 1999 to July 2012)

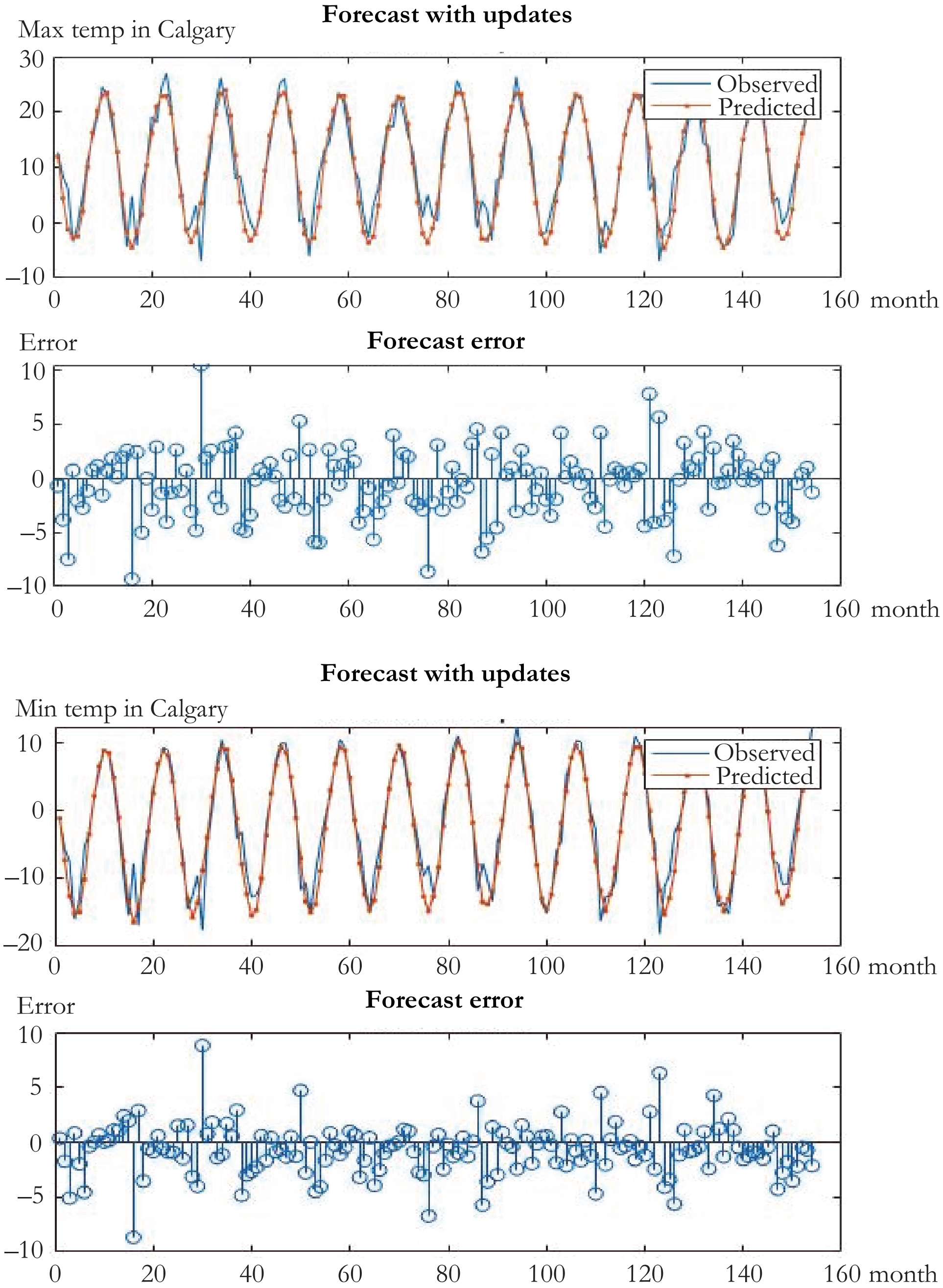
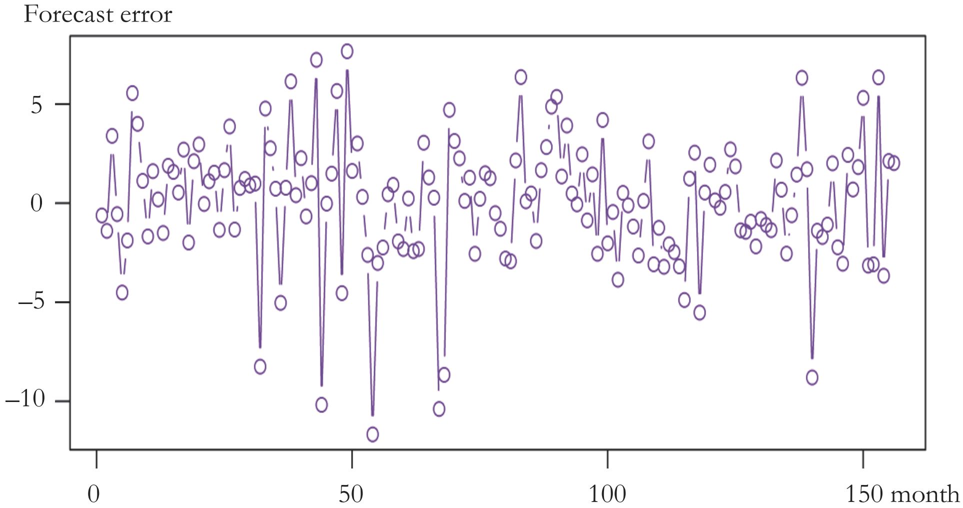


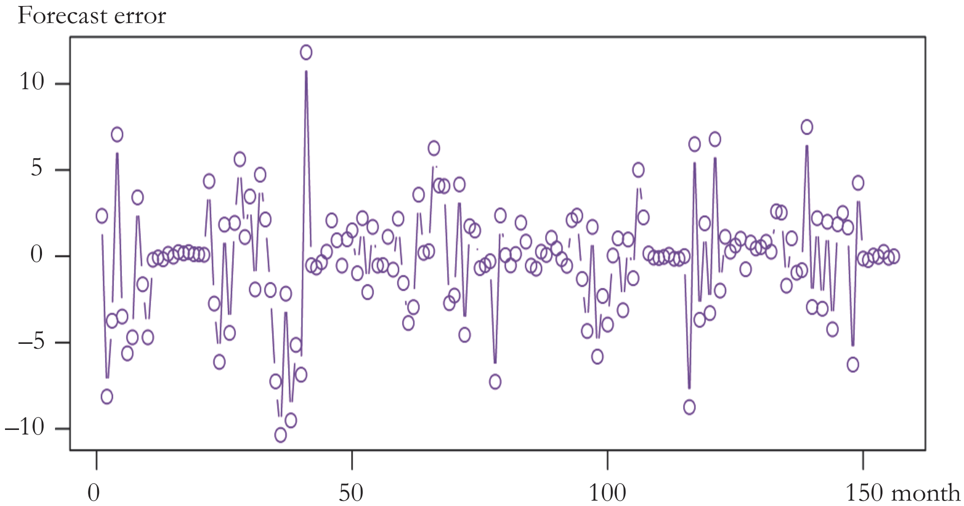
Figure A6

Calgary temperature forecast errors with SARFIMA models
(Forecast range: June 1999 to July 2012)

Maximum temperature



Minimum temperature



REFERENCES

AGBO, E. P. (2021): The role of statistical methods and tolls for weather forecasting and modelling. In: SAIFULLAH, M.: *Weather Forecasting* https://doi.org/10.5772/intechopen.96854

BELL, W. R.–HOLAN, S. H.–MCELROY, T. S. (2012): *Economic time series modeling and seasonality* Chapman and Hall/CRC, Boca Raton-London-New York.

BISOGNIN, C.–LOPES, S. R. C. (2009): Properties of seasonal long memory processes *Mathematical and Computer Modelling* 49 (9-10): 1837–1851. https://doi.org/10.1016/j.mcm.2008.12.003

CANOVA, F.–HANSEN, B. E. (1995): Are seasonal patterns constant over time? A test for seasonal stability *Journal of Business and Economic Statistics* 13 (3): 237–252. https://doi.org/10.1080/07350015.1995.10524598

CONTRERAS-REYES, J.–PALMA, W. (2013): Statistical analysis of autoregressive fractionally integrated moving average models in R *Computational Statistics* 28: 2309–2331. https://doi.org/10.1007/s00180-013-0408-7

DATTA, A.–BISWAS, S.–SI, S. (2019): Complete statistical analysis to weather forecasting *International Conference on Computational Intelligence in Pattern Recognition* (CIPR 2019), Indian Institute of Engineering Science and Technology (IIEST) Shibpur, Kolkata, India.

DIMRI, T.–AHMAD, S.–SHARIF, M. (2020): Time series analysis of climate variables using seasonal ARIMA approach *Journal of Earth System Science* 129: 149. https://doi.org/10.1007/s12040-020-01408-x

DIPIETRO, R.–HAGER, G. D. (2020): Deep learning: RNNs and LSTM. In: ZHOU, S. K.–RUECKERT, D.–FICHTINGER, G. *Handbook of Medical Image Computing and Computer Assisted Intervention* pp. 503–519., The Elsevier and MICCAI Society Book Series. https://doi.org/10.1016/B978-0-12-816176-0.00026-0

DONG, D.–SHENG, Z.–YANG, T. (2018): *Wind power prediction based on recurrent neural network with long short-term memory units* 2018 International Conference on Renewable Energy and Power Engineering (REPE), Ontario. https://doi.org/10.1109/REPE.2018.8657666

GONZALEZ, J.–YU, W. (2018): Non-linear system modeling using LSTM neural networks *IFAC-PapersOnLine* 52 (13): 485–489. https://doi.org/10.1016/j.ifacol.2018.07.326

GRANGER, C. W. J.–JOYEUX, R. (1980): An introduction to long memory time series models and fractional differencing *Journal of Time Series Analysis* 1 (1): 15–39. https://doi.org/10.1111/j.1467-9892.1980.tb00297.x

HOCHREITER, S.–SCHMIDHUBER, J. (1997): Long short-term memory *Neural Computation* 9 (8): 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

ISMAIL, L.–MATERWALA, H.–ZNATI, T.–TURAEV, S.–KHAN, M. (2020): Tailoring time series models for forecasting coronavirus spread: Case studies of 187 countries *Computational and Structural Biotechnology Journal* 18: 2972–3206. https://doi.org/10.1016/j.csbj.2020.09.015

KHEDHIRI, S. (2021): Forecasting COVID-19 infections in the Arabian Gulf region *Modeling Earth Systems and Environment* https://doi.ord/10.1007/s40808-021-01332-z

KIRKWOOD, C.–ECONOMOU, T.–ODBERT, H.–PUGEAULT, N. (2021): A framework for probabilistic weather forecast post-processing across models and lead times using machine learning *Philosophical Transactions of the Royal Society A*, 379 (2194): 1–15. https://doi.org/10.1098/rsta.2020.0099

KREUZER, D.–MUNZ, M.–SCHLUTER, S. (2020): Short-term temperature forecasts using a convolutional neural network – An application to different weather stations in Germany *Machine Learning with Applications* 2: 100007. https://doi.org/10.1016/j.mlwa.2020.100007.

KWIATKOWSKI, D.–PHILLIPS, P. C. B.–SCHMIDT, P.–SHIN, Y. (1992): Testing for the null hypothesis of stationarity against the alternative of a unit root *Journal of Econometrics* 54 (1-3): 159–178. https://doi.org/10.1016/0304-4076(92)90104-Y

LAI, Y.–DZOMBAK, D. D. (2020): Use of the autoregressive integrated moving average (ARIMA) model to forecast near-term regional temperature and precipitation *Weather and Forecasting* 35 (3): 959–976. https://doi.org/10.1175/WAF-D-19-0158.1

LEE, D.–LEE, G.–KIM, S.–JUNG, S. (2020): Future runoff analysis in the Mekong River Basin under climate change scenario using deep learning *Water* 12: 1556. https://doi.org/10.3390/w12061556

MASSIE, D.–ROSE, M. (1997): Predicting daily maximum temperature using linear regression and Eta geopotential thickness forecasts *Weather Forecasting* 12 (4): 799–807. https://doi.org/10.1175/1520-0434(1997)012%3C0799:PDMTUL%3E2.0.CO;2

MISHA, S.–BORDIN, C.–TAHARAGUCHI, K.–PALU, I. (2019): Comparison of deep learning models for multivariate prediction of time series wind power generation and temperature *Energy Reports* 6 (3): 273–286. https://doi.org/10.1016/j.egyr.2019.11.009

MOSTAFAEI, H.–SAKHABAKHSH, L. (2012): Using SARFIMA model to study and to predict the Iran's oil supply *International Journal of Energy Economics and Policy* 2 (1): 41–49.

PAPACHARALAMPOUS, G.–TYRALIS, H.–KOUTSOYIANNIS, D. (2018): Predictability of monthly temperature and precipitation using automatic time series forecasting methods *Acta Geophysica* 66: 807–831. https://doi.org/10.1007/s11600-018-0120-7

PARK, I.–KIM, H. S.–LEE, J.–KIM, J. H.–SONG, C. H.–KIM, H. K. (2019): Temperature prediction using the missing data refinement model based on a long short-term memory neural network *Atmosphere* 10 (11): 718. https://doi.org/10.3390/atmos10110718

QI, C.–ZHANG, D.–ZHU, Y.–LIU, L.–LI, C.–WANG, Z.–LI, X. (2020): SARFIMA model prediction for infectious diseases: Application to hemorrhagic fever with renal syndrome and comparing with SARIMA *BMC Medical Research Methodology* 20: 243. https://doi.org/10.1186/s12874-020-01130-8

Raible, C.–Bischof, G.–Fraedrich, K.–Kirk, E. (1999): Statistical single-station short-term forecasting of temperature and probability of precipitation: Area interpolation and NWP combination *Weather and Forecasting* 14 (2): 203–214. https://doi.org/10.1175/1520-0434(1999)014<0203:SSSSTF>2.0.CO;2

Shankarnarayan, V.K.–Ramakrishna, H. (2021): Comparative Study of Three Stochastic Future Weather Forecast Approaches: A Case Study *Data Science and Management* 3: 3–12. https://doi.org/10.1016/j.dsm.2021.07.002

Sowell, F. (1992): Maximum likelihood estimation of stationary fractionally integrated time series models *Journal of Econometrics* 53 (1-3): 165–188. https://doi.org/10.1016/0304-4076(92)90084-5

Sutskever, I.–Vinyals, O.–Le, Q. V. (2014): Sequence to sequence learning with neural networks In: Ghahramani, Z. W.–Max, C. *Advances in Neural Information Processing Systems* pp. 3104–3112., MIT, New York.

Tran Anh, D.–Van, S. P.–Dang, T. H.–Hoang, L. P. (2019): Downscaling rainfall using deep learning long short-term memory and feedforward neural network *International Journal of Climatology* 39 (10): 4170–4188. https://doi.org/10.1002/joc.6066

Wibowo, W.–Ulama, B. S. S.–Siagian, T. H–Purwa, T.–Wilantari, R. N. (2021): Impact of earthquakes on the number of airline passenger arrivals and departures: A case study of West Nusa Tenggara Province, Indonesia *Regional Statistics* 11 (3): 133–157. https://doi.org/10.15196/RS110302

Yao, L.–Guan, Y. (2018): An improved LSTM structure for natural language processing *2018 IEEE International Conference of Safety Produce Informatization (IICSPI)* pp. 565–569. https://doi.org/10.1109/IICSPI.2018.8690387

Yu, Y.–Si, X.–Hu, C.–Zhang, J. (2019) : A Review of recurrent neural networks: LSTM cells and network architectures *Neural Communication* 31 (7): 1235–1270. https://doi.org/10.1162/neco_a_01199

Xiaoyun, Q.–Xiaoning, K.–Chao, Z.–Shuai, J.–Xiuda, M. (2016): Short-term prediction of wind power based on deep long short-term memory *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)* pp. 1148–1152. https://doi.org/10.1109/APPEEC.2016.7779672

INTERNET SOURCES

Browniee, J. (2018): A gentle introduction to SARIMA for time series forecasting in Python *Machine Learning Mastery* https://machinelearningmastery.com/sarima-for-time-series-forecasting-in-python/ (downloaded: July 2021).

Chandra, R.–Jain, A.–Chauhan, D. (2021): *Deep learning via LSTM models for COVID-19 infection forecasting in India* https://arxiv.org/pdf/2101.11881.pdf (downloaded: July 2021).

Kingman, D. P.–Ba, L. J. (2015): Adam: A method for stochastic optimization *International Conference on Learning Representations* May 7-9, 2015, San Diego, USA. https://hdl.handle.net/11245/1.505367 (downloaded: July 2021).

KRAFT, P.–WEBER, C.–LEBO, M. (2017): *The ArfimaMLM package for R*
https://pwkraft.github.io/files/articles/ArfimaMLM-documentation.pdf
(downloaded: July 2021).
VEENSTRA, J.Q–MCLEOD, A. (2017): *The arfima package for R*
https://cran.r-project.org/web/packages/arfima/arfima.pdf
(downloaded: July 2021).

DATABASE/WEBSITE

[1] Government of Canada:
https://climate.weather.gc.ca/historical_data/search_historic_data_e.html
(downloaded: July 2021).