JOURNAL OF WATER AND LAND DEVELOPMENT

e-ISSN 2083-4535

Polish Academy of Sciences (PAN) Institute of Technology and Life Sciences – National Research Institute (ITP – PIB)

JOURNAL OF WATER AND LAND DEVELOPMENT
DOI: 10.24425/jwld.2022.140785
2022, No. 53 (IV–VI): 100–107

The SARIMA model-based monthly rainfall forecasting for the Turksvygbult Station at the Magoebaskloof Dam in South Africa

Kassahun Birhanu Tadesse, Megersa Olumana Dinka

University of Johannesburg, Faculty of Engineering and the Built Environment,
Department of Civil Engineering Sciences, PO Box 524, Auckland Park, 2006 Johannesburg, South Africa

RECEIVED 24.04.2019 ACCEPTED 16.10.2019 AVAILABLE ONLINE –.–.2022

Abstract: Rainfall forecast information is important for the planning and management of water resources and agricultural activities. Turksvygbult rainfall near the Magoebaskloof Dam (South Africa) has never been modelled and forecasted. Hence, the objective of this study was to forecast its monthly rainfall using the SARIMA model. GReTL and automatic XLSTAT software were used for forecasting. The trend of the long-term rainfall time series (TS) was tested by Mann–Kendall and its stationarity was proved by various unit root tests. The TS data from Oct 1976 to Sept 2015 were used for model training and the remaining data (Oct 2015 to Sept 2018) for validation. Then, all TS (Oct 1976 to Sept 2018) were used for out of sample forecasting. Several SARIMA models were identified using correlograms that were derived from seasonally differentiated TS. Model parameters were derived by the maximum likelihood method. Residual correlogram and Ljung–Box Q tests were used to check the forecast accuracy. Based on minimum Akaike information criteria (*AI*) value of 5642.69, SARIMA (2, 0, 3) $(3, 1, 3)_{12}$ model was developed using GReTL as the best of all models. SARIMA (1, 0, 1) $(3, 1, 3)_{12}$, with minimum *AI* value of 5647.79, was the second-best model among GReTl models. This second model was also the first best automatically selected model by XLSTAT. In conclusion, these two best models can be used by managers for rainfall forecasting and management of water resources and agriculture, and thereby it can contribute to economic growth in the study area. Hence, the developed SARIMA forecasting procedure can be used for forecasting of rainfall and other time series in different areas.

Keywords: automatic forecasting, correlogram, Mann–Kendall test, stationarity test, time series, unit root, XLSTAT

INTRODUCTION

Limited availability and unreliable rainfall have affected crop production and threatened food security especially in arid and semi-arid areas in the world. Crop production is sensitive to extreme climatic and weather conditions among several factors of production. Thus, the information on future occurrence and reliability of the agrometeorological characteristics is important for planning irrigated and rainfed agriculture. Agrometeorological conditions are modelled and predicted using their historical time series data as the past patterns in the time series data are believed to manifest themselves in the future [BOX, JENKINS 1976].

In areas where only one kind of meteorological record is available, a univariate BOX and JENKINS methodology [BOX, JENKINS 1976] are commonly used for analysis and forecasting of time series. Autoregressive integrated moving average (ARIMA) models have become a major tool in agrometeorological and hydrological applications to understand the characteristics of the precipitation, temperature, river flows, and so on. Seasonal ARIMA (SARIMA) models are important to model and forecast time series that show a periodic or non-stationary behaviour both within and across seasons [BOX *et al.* 2015; BOX, JENKINS 1976; BROCKWELL *et al.* 2002].

Several forecasting studies were conducted using SARIMA models in different areas. MURAT *et al.* [2018] used ARIMA and regression models for forecasting daily rainfall and temperature time series in Europe. SARIMA models were used to forecast rainfall [Arumugam, Saranya 2018; Papalaskaris *et al.* 2016; Takele, Gebretsidik 2015].

Sarima models have been used for different forecasting applications such as temperature [Chen *et al.* 2018], river flow in South Africa [Tadesse, Dinka 2017; Tadesse *et al.* 2017], malaria prevalence in Ghana [Anokye *et al.* 2018], reservoir inflow at the Koga Dam, Ethiopia [Tadesse, Dinka 2017], and annual rainfall and maximum temperature over the Tordzie watershed in Ghana [Nyatuame, Agodzo 2018]. A wavelet-SARIMA-ANN hybrid model has been used to forecast precipitation in Iran [Shafaei *et al.* 2016].

However, forecasting models developed for one area may not be used for the other areas due to a peculiar characteristic of rainfall and river flow time series. Although many rainfall forecasting studies are conducted in different areas of the world, the rainfall at the Turksvygbult area near the Magoebaskloof Dam has never been modelled and forecasted. The Magoebaskloof Dam in South Africa was designed to supply water for irrigation and domestic uses in the surrounding area. Hence, information on future rainfall characteristics is vital for the optimal operation of the dam and planning of irrigated and rainfed agriculture in the area.

Therefore, this study aimed to inspect statistical characteristics of the monthly rainfall time series and select the best forecasting model that could be used to plan reservoir water release policy, irrigation, and rainfed agricultural activities. Moreover, this study evaluated the rainfall forecast accuracy of automatic XLSTAT time series forecasting software and expert-based non-automatic GReTL software.

STUDY METHODS

STUDY AREA

South Africa is a relatively arid country with an even spatial rainfall distribution with two-thirds of the country receiving less than 500 mm of annual rainfall [King, Pienaar 2011]. Rainfall occurs mainly in summer months in most of the areas. The Turksvygbult rainfall station is located near the Magoebaskloof Dam in the Limpopo province, South Africa. The Magoebaskloof Dam is located on the Politsi River, a major tributary of the Groot Letaba River. According to Haumann [2006], the dam storage capacity is $5.5 \cdot 10^6$ m^3 and yield $4.92 \cdot 10^6$ m$^3 \cdot$y^{-1}. The Magoebaskloof Dam was commissioned in 1971. It is owned and operated by the Department of Water Affairs. The dam was originally intended to supply water for irrigation purposes only, however, the need arose later for domestic and industrial water in Politsi, Duiwelsskloof, and Ga-Kgapane. Water allocation from the dam totals $13.1 \cdot 10^6$ m$^3 \cdot$y^{-1} of which $11.044 \cdot 10^6$ m$^3 \cdot$y^{-1} is for agriculture and $2.034 \cdot 10^6$ m$^3 \cdot$y^{-1} for domestic uses.

DATA ACQUISITION AND ARRANGEMENT

The Turksvygbult rainfall data were collected from the Department of Water and Sanitation, South Africa. The rainfall was modelled and forecasted using the Gnu Regression, Econometrics, and Time-series Library (GRETL) software [Baiocchi, Distaso 2003], and the automatic forecasting Addinsoft XLSTAT version 2019.1 software. The rainfall has had 504 monthly observations from October 1976 to September 2018. The time-series data from October 1976 to September 2015 were used for model training. The rest observations from October 2015 to September 2018 were used for model validation. Then, all observations from 1976 to 2018 were used to forecast four years of rainfall from October 2018 to September 2022.

SARIMA MODELLING

SARIMA $(p, d, q)(P, D, Q)$ model with seasonal and non-seasonal components of (p, d, q) and (P, D, Q), respectively may be described as shown in Equation (1) [Box, Jenkins 1976].

$$\Phi_P(B^S)\varphi_p(B)(1-B^S)^D(1-B)^d Y_t = \Theta_Q(B^S)\theta_q(B)\varepsilon_t \qquad (1)$$

where: B is lag operator defined as $B^k Y_t = Y_{t-k}$,

$$\Phi_P(B^S) = 1 - \Phi_S B^S - \Phi_{2S} B^{2S} - \ldots \Phi_{PS} B^{PS},$$
$$\Theta_Q(B^S) = 1 - \Theta_S B^S - \Theta_{2S} B^{2S} - \ldots \Theta_{QS} B^{QS},$$
$$\varphi_p(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots \varphi_p B^p,$$
$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots \theta_q B^q$$

where: $\theta(B)$ and $\varphi(B)$ = polynomials of order q and p, respectively; $\Theta(B^S)$ and $\Phi(B^S)$ = polynomials in B of degrees Q and P, respectively; Y_t = rainfall at time t; ε_t = independently distributed random variable at time t; P and p = orders of the non-seasonal autoregression; Q and q = orders of seasonal autoregression; D and d = numbers of difference terms for seasonal and non-seasonal components. This methodology has the following steps [Box *et al.* 2008; Box, Jenkins 1976].

Identification of the model and parameter estimation

Before model identification, the stationarity of the rainfall time series was evaluated by Mann–Kendall trend test (MK) [Kendall 1975], Kwiatkowski–Phillips–Schmidt–Shin (KPSS) [Kwiatkowski *et al.* 1992] at 0.5 significance level ($\alpha = 0.05$), autocorrelation function (ACF) and partial autocorrelation functions (PACF) [Singh *et al.* 2011]. Additionally, the unit root test was performed according to Dickey and Fuller (DF) [1979], Phillips and Perron (PP) [1988]. Initially, few SARIMA (p, d, q) $(P, D, Q)_{12}$ models were identified from the ACF and PACF diagram of the seasonally differenced rainfall time series [Cryer, Chan 2008; Hipel *et al.* 1977; Salas *et al.* 1980]. Then their model parameters were estimated based on maximum likelihood [Hamilton 1994].

Diagnostic checking and forecasting

Autocorrelations of model forecast residuals were diagnosed by Ljung, Box Q tests [Ljung, Box 1978] in terms of ACF and PACF plots. The performance of the model forecasts were evaluated in terms of three criteria: Akaike information [Akaike 1974], Schwarz–Bayes [Schwarz 1978], and Hannan–Quinn [Hannan, Quinn 1979].

RESULTS AND DISCUSSION

DESCRIPTIVE STATISTICS OF THE TIME SERIES

The minimum and maximum Turksvybult monthly rainfall for 504 rainfall observations were 0.0 and 1555.7 mm, respectively. The mean monthly rainfall was 113.8 with a standard deviation of

150.7 mm, which shows the high variability of the monthly rainfall in the study area. However, its average annual rainfall was 1365.4 mm which was above 500 mm of annual rainfall received in most areas of South Africa as it was described in section "The study area".

The high variable rainfall characteristic is challenging for managers while planning water resources management or agricultural activities in the study area. Therefore, it necessitates modelling and forecasting of rainfall to provide managers with accurate information on future occurrence and amount of rainfall.

STATIONARITY OF THE TIME SERIES

Graphical observation

One of the techniques for determining whether a rainfall time series is stationary or not is a visual observation of the rainfall time series. The plot of the Turksvygbult rainfall time series is shown in Figure 1. The seasonality of the rainfall time series can be seen from the periodic (cyclic) peaks and lows for wet and dry seasons, respectively. Based on this graphical observation, the time series was not seasonally stationary. Hence, it is necessary to avoid the periodic component of the time series by seasonal differencing before using it for modelling and forecasting. Besides, from the visual observation of Figure 1, no increasing or decreasing trend of the rainfall time series was seen. In addition to graphical observations, statistical methods of stationarity and trend evaluation results were shown in Table 1.

Statistical methods

From Table 1, the absolute values of the observed (computed) for Dickey and Fuller (DF) and Phillips–Perron (PP) tests were greater than their critical values at 0.5 significance level ($\alpha = 0.05$). This proved the absence of a unit root. In another words, the time series is stationary and shows no upward or downward trends. Similarly, the computed/observed p-values for DF and PP tests are lower than the significance level. This means the time series has no unit root and therefore it was stationary at its level or at original (untransformed) time-series data. A different way of interpreting the stationarity is Kwiatkowski–Phillips–Schmidt–Shin (KPSS) trend test. The KPSS trend test proved the stationarity of the time series since its computed p-value of 0.493 was greater than $\alpha = 0.05$.

Like the KPSS test, the non-seasonal Mann–Kendall (MK) trend test showed the absence of a trend in the rainfall time series as its p-value of 0.074 (Tab. 1) was greater than $\alpha = 0.05$. Furthermore, the seasonal MK test for 12 months of seasonality ($S = 12$) showed the absence of trend due to seasonality as the

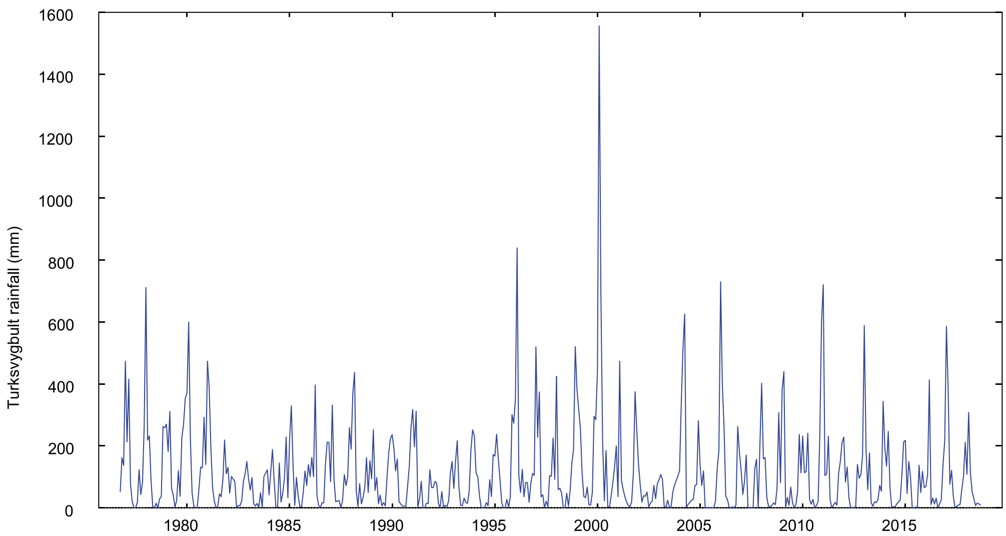

Fig. 1. Time series plot of the Turksvygbult rainfall from 1976 to 2018; source: own study

Table 1. The unit root and Mann–Kendall trend tests

Unit root and stationarity tests				Mann-Kendall trend test		
parameter	DF	PP	KPSS	parameter	Mann–Kendall	seasonal Mann–Kendall
(observed value)	−10.159	−9.869	0.056	Kendall's	−0.053	−0.071
(critical value)	−3.428	−1.941	0.148	Sen's slope	−0.027	−0.219
p-value	<0.0001	<0.0001	0.493	p-value	0.074	0.058
α	0.05	0.05	0.05	α	0.05	0.05
Remark	stationary	stationary	stationary	Remark	no trend	no trend

Explanations: DF = Dickey and Fuller, PP = Phillips–Perron, KPSS = Kwiatkowski–Phillips–Schmidt–Shin.
Source: own study.

p-value of 0.058 was greater than α = 0.05. These statistical MK tests proved the appropriateness of graphical evaluation made in section "Graphical observation" for the absence of a trend in the rainfall time series. Negative signs of the Sen's slope estimator at MK tests (Tab. 1) were an indicator of a slightly decreasing trend of rainfall time series due to climate change though it was not statistically significant. The decrease in trend was greater for the seasonal component of a time series with a Sen's slope of –0.219 as compared to that of a non-seasonal component with a Sen's slope of –0.027. All three statistical tests (DF, PP, KPPS, and MK) proved the stationarity and the absence of a significant trend in the rainfall time series.

IDENTIFICATION OF MODELS

The ACF and PACF plots of the measured rainfall time series (original data) have been shown in Figure 2. The seasonality of the time series is evidenced obviously from the sinusoidal graph of the ACF and the significant spikes detected at intervals of 12 months. This supports the evidence from the graphical method in section "Graphical observation" for the availability of seasonality in the rainfall time series. Therefore, seasonal differencing (D = 1) was made to make the time series seasonally stationary. As it was well explained in the section mentioned, the time series was stationary and had no significant trend, hence there was no need for differencing the non-seasonal component. Therefore, the non-seasonal differencing was made zero (d = 0).

Hence, SARIMA (p, 0, q) (P, 1, Q)$_{12}$ models were suggested for model identification. Initial, parameter p, q, P and Q of these models were determined based on the characteristics of ACF and PACF plots of seasonally differenced time series (Fig. 3).

From Figure 3, the ACF of the seasonally differenced time series has significant spikes at the 1st lag and cuts off after the 2nd lag for the non-seasonal component. For the seasonal component of the time series at S = 12, ACF has a significant spike at lag 12th and then cuts off after lag 24th. Similarly, PACF of the seasonally differenced series has significant spikes at 1st lag for the non-seasonal and significant spikes at 12th and 24th and 36th lags for seasonal components. Hence, one for each autoregressive (AR) and seasonal moving average (SMA) parameters for the non-seasonal component, and three for seasonal autoregressive (SAR) and two for seasonal moving average (SMA) parameters were suggested as SARIMA (1, 0, 1) (3, 1, 2) model. However, by adding one parameter for AR and SMA, SARIMA (2, 0, 1) (3, 1, 3)$_{12}$ was identified as an initial model.

However, as model identification involves much trial and error, several SARIMA models within ranges of three-parameter model components were evaluated using GReTl and automatic XLSTAT forecasting software. Finally, the top best three models from both software solutions are selected and presented in Table 2. The minimum value of Akaike information (AI) (bolded value of 5642.69), SARIMA (2, 0, 3) (3, 1, 3)$_{12}$ was found to perform best in forecasting mean monthly rainfall at the Turksvygbult station using the GReTL software.

The XLSTAT automatically selected SARIMA (1, 0, 1) (3, 1, 3)$_{12}$ as the best rainfall forecasting model among many SARIMA models based on Akaike information criteria (AIC). This automatically selected the SARIMA model which was the second-best model identified by GReTL. As shown in Table 3, the GReTL software forecasted this model at minimum AI and SB criteria as compared to that of XLSTAT. This implies that GReTL has the higher forecast accuracy than automatic forecasting by the XLSTAT software.

The GReTL forecasted the SARIMA (2, 0, 3) (3, 1, 3)$_{12}$ model which was also the best of all automatic XLSTAT forecasting models in terms of minimum AI value. Besides, the

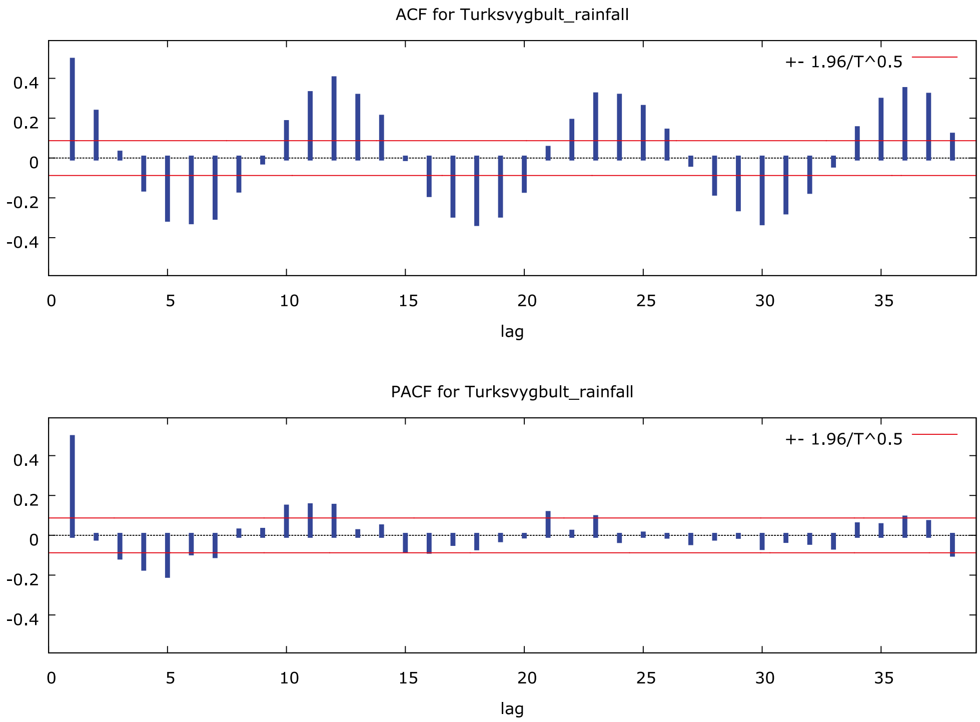

Fig. 2. Autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the rainfall time series for original data; source: own study

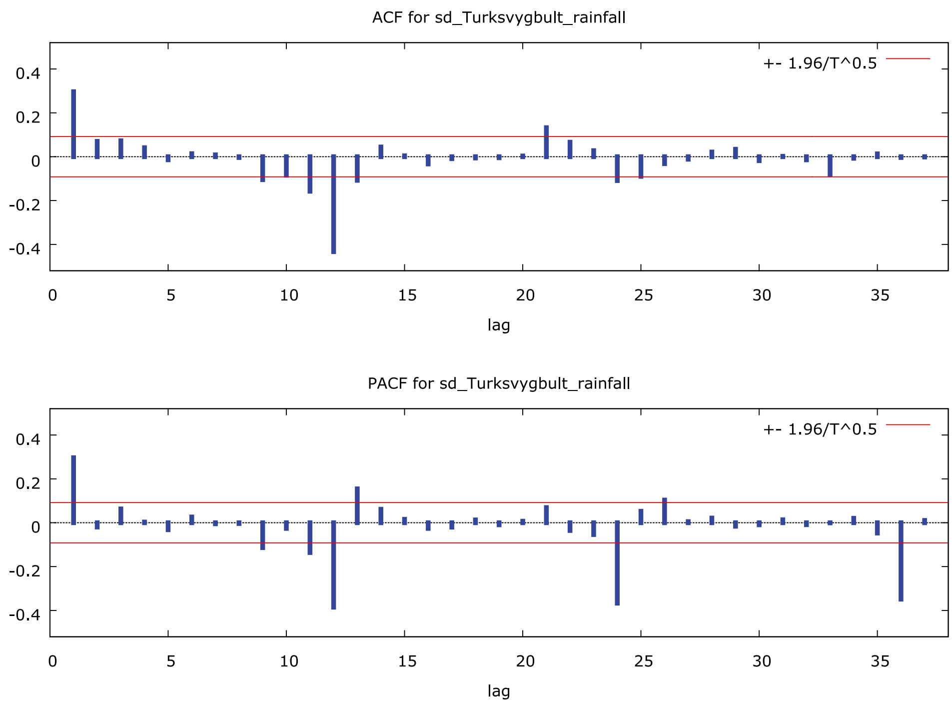

Fig. 3. Autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of seasonally differenced rainfall time series; source; own study

Table 2. Performance evaluations of the alternative models

ARIMA models						Automatic XLSTAT forecasting				GReTL forecasting			
p	*d*	*q*	*P*	*D*	*Q*	AI	AIc	SB	rank	AI	HB	SB	rank
1	0	1	3	1	3	5684.40	5685.00	5729.76	1	**5647.79**	5660.77	5258.68	2
3	0	2	1	1	2	5689.83	5691.28	5735.18	2	5950.69	5965.44	5988.20	3
2	0	3	3	1	3	5692.96	5693.91	5750.67	3	**5642.69**	5662.16	5692.11	1

Explanations: and p = orders of the non-seasonal autoregression; Q and q = orders of seasonal autoregression; D and d = numbers of difference terms for seasonal and non-seasonal components, AI = Akaike information, AI_c = Akaike information corrected, SB = Schwarz Bayes, and HQ = Hannan–Quinn criteria, and the bolded are the minimum AI values.
Source: own study.

Table 3. Significant parameters of SARIMA (2, 0, 3) (3, 1, 3)$_{12}$ model

Parameter	Coefficient	*SD*	*z*	*p*-value	Significance
φ_1	0.888457	0.0856824	10.37	3.42e–025	***
φ_2	−0.816589	0.0736300	−11.09	1.40e–028	***
Φ_1	−0.969379	0.0811488	−11.95	6.84e–033	***
Φ_2	−0.483461	0.0912276	−5.300	1.16e–07	***
Φ_3	−0.130332	0.0432574	−3.013	0.0026	***
θ_1	−0.573597	0.0893199	−6.422	1.35e–010	***
θ_2	0.580649	0.0756998	7.670	1.71e–014	***
θ_3	0.292143	0.0510679	5.721	1.06e–08	***
Θ_1	0.153259	0.0890021	1.722	0.0851	*
Θ_2	−0.492000	0.0659440	−7.461	8.60e–014	***
Θ_3	−0.457309	0.0992678	−4.607	4.09e–06	***

Explanations: SD = standard deviation, z = z-statistic, p-value = probability value, the *, **, *** = levels of significance, $p \le 0.05$, $p \le 0.01$, $p \le 0.001$, φ, Φ, θ, Θ as in Eq. (1).
Source: own study.

pattern of the out of sample rainfall forecast looks like that of measured values, as shown in Figure 4. This shows the model was able to capture the stochastic component in the forecasted monthly rainfall values. Forecast graphs from the XLSTAT software (Fig. 5) have a little bit uniform pattern as the result of almost equal values of rainfall forecast and nearly equal prediction errors (residuals). This shows that the SARIMA models developed using the XLSTAT software were not very good in retaining the stochastic components in the forecasted rainfall. However, as XLSTAT software automatically selects the best forecasting model among several models, it is simple in use by planners with little experience of time series modelling and forecasting procedures. XLSTAT forecast values are close to the long year mean rainfall values and hence, the forecast values can be used for initial planning purposes.

PARAMETERS OF THE SELECTED SARIMA (2, 0, 3) (3, 1, 3)$_{12}$ MODEL

Parameters of the selected model are shown in Table 3. All parameter values of this model are significant and, hence, can be included in the model for forecasting. The four years out of

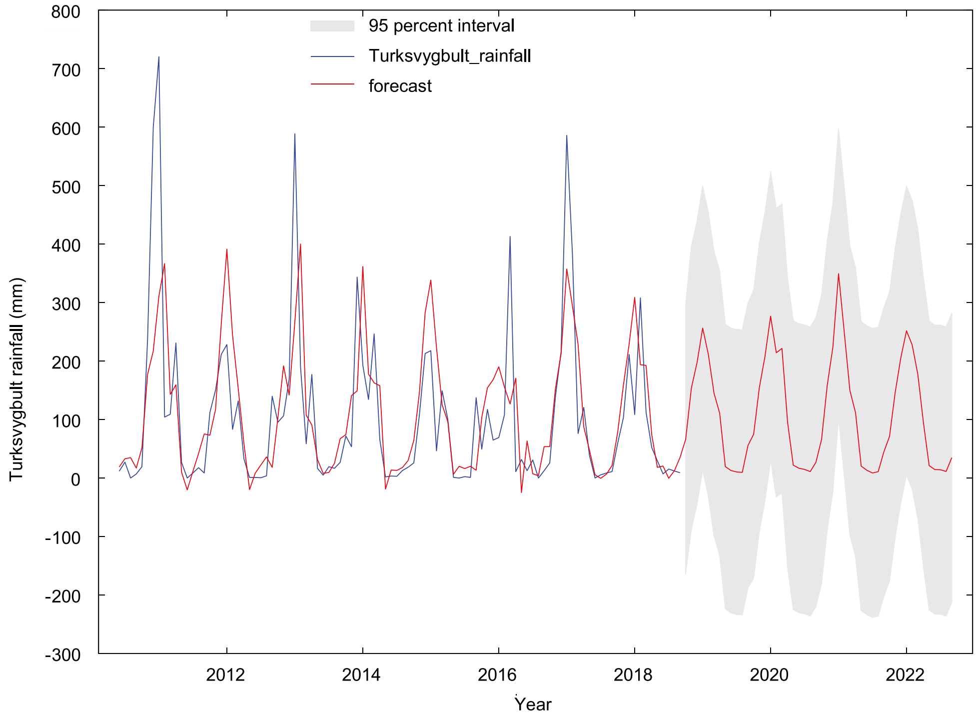

Fig. 4. SARIMA (2, 0, 3) (3, 1, 3)$_{12}$ model forecast of rainfall from 2019 to 2022 using GReTL software; source; own study

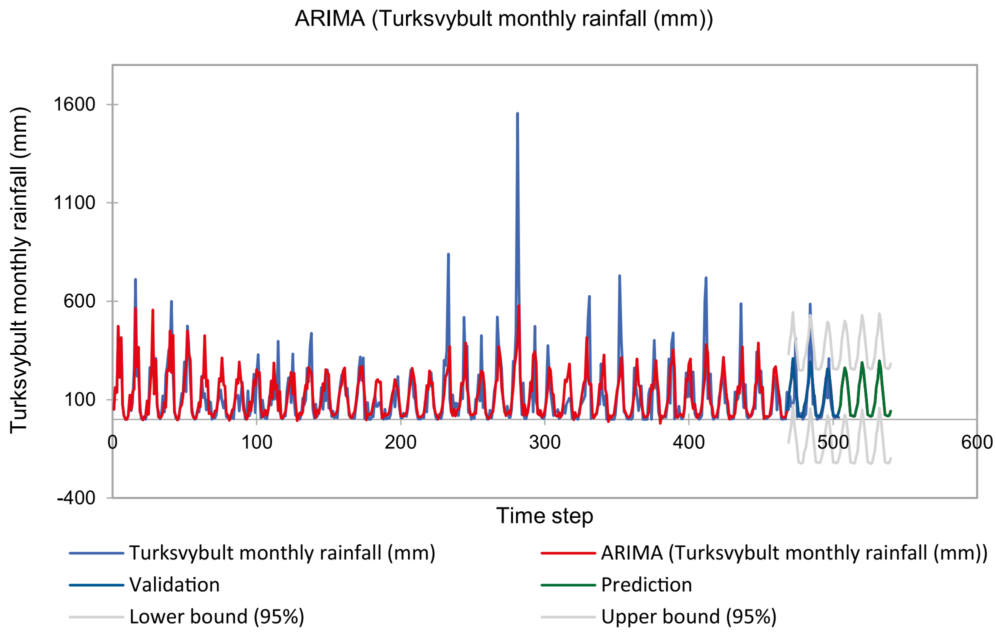

Fig. 5. SARIMA (1, 0, 1) (3, 1, 3)$_{12}$ forecast of rainfall using automatic XLSTAT software; source; own study

sample forecast (2018 to 2022) and in-sample forecasts (2012 to 2018) of the SARIMA (2, 0, 3) (3, 1, 3)$_{12}$ model are shown in Figure 4.

DIAGNOSIS OF FORECAST ACCURACY

In a good SARIMA model, the forecast error or residuals should follow assumptions for the stationarity in the time series model. In other words, residuals should be white noise or independent. To prove this assumption, the autocorrelation of residuals was checked from the plots of residual ACF and PACF (Fig. 6). No significant correlation was found among the residuals as all of them were within the confidence limits of 95% which were close to zero. This shows that the residuals were independent or white noise and their distributions are normal. Moreover, the absence of the autocorrelation of residuals was statistically proved by Ljung–Box Q' test. The Ljung–Box Q' test value of 4.779 at a *p*-value of 0.3107 was greater than 0.05 significance level ($\alpha = 0.05$). Hence, both graphical and statistical diagnosis showed the absence of the residual correlation and proved the highest forecast accuracy of the selected SARIMA (2, 0, 3) (3, 1, 3)$_{12}$ model.

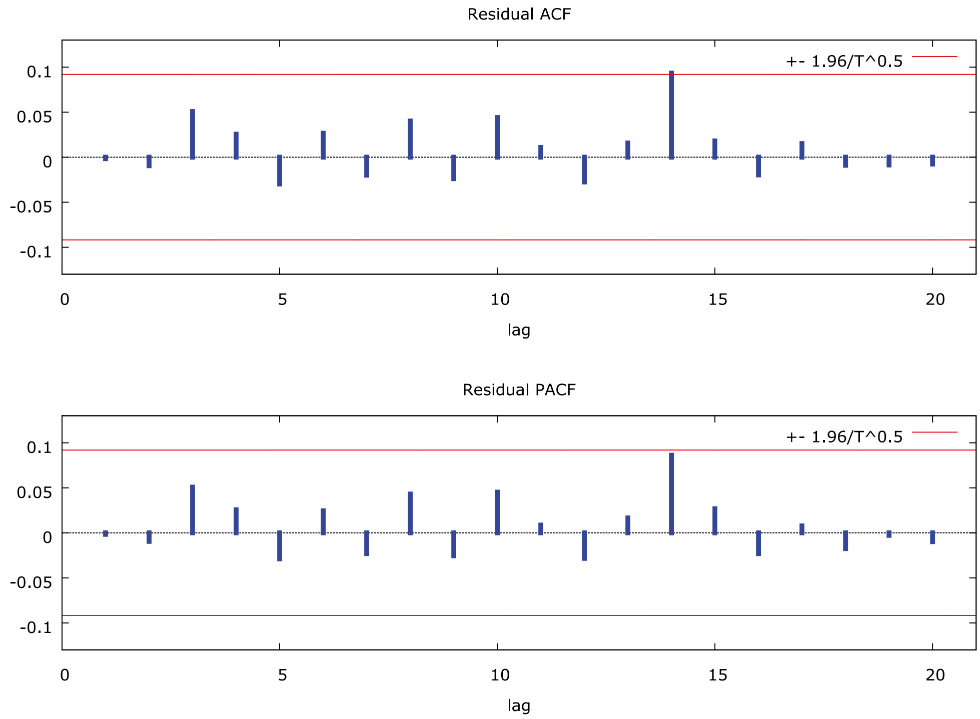

Fig. 6. Residual autocorrelation function (ACF) and partial autocorrelation functions (PACF) plots; source: own study

CONCLUSIONS

The Turksvygbult rainfall was modelled and forecasted using seasonal autoregressive integrated moving average (SARIMA) models. The GReTL and automatic forecasting XLSTAT software solutions were used for modelling and forecasting the rainfall time series. Different statistical and graphical techniques of trend tests revealed the absence of significant long-term monthly rainfall changes in the study area. Among several models, the SARIMA (2, 0, 3) × (3, 1, 3)$_{12}$ model, which was developed using the GReTL software, has shown a minimum Akaike information criteria value of 5642.69. The residuals of this model were found to be white noise (independent) and thereby proved the highest forecast accuracy of the model. Besides, the forecast values of this model retained the stochasticity of the rainfall time series as compared to that of automatic XLSTAT forecast models. A similar pattern of the forecasted graph with that of the recorded rainfall data was good evidence for the availability of stochastic components in the forecasted time series. This helps to capture the rainfall variability in the study area.

The XLSTAT software automatically selected SARIMA (1, 0, 1) (3, 1, 3)$_{12}$ among several models. The advantage of the automatic XLSTAT forecasting is its simplicity. Hence it can be used by those who have little expertise in time series modelling and forecasting. The XLSTAT rainfall forecast values can be used for the initial planning of water resources and agricultural activities.

In conclusion, both SARIMA (2, 0, 1) (3, 1, 3)$_{12}$ and SARIMA (1, 0, 1) (3, 1, 3)$_{12}$ models can be used for rainfall modelling and forecasting at the Turksvygbult area. The rainfall forecast values of this model can be used for planning and management of water harvesting structures, rainfed and irrigated agriculture, and the reservoir water operations for domestic and other uses, thereby improve economic growth in the study area. Moreover, the developed SARIMA modelling and selection procedures can be used in other areas for forecasting rainfall and other time series.

ACKNOWLEDGMENTS

Hereby, we acknowledge the importance of the University of Johannesburg for providing postdoctoral fellowship, Debre Markos University for study leave, Department of Water and Sanitation of South Africa for meteorological data, and the Addinsoft XLSTAT company for providing the XLSTAT version 2019.1 software for time series analysis.

REFERENCES

Akaike H. 1974. A new look at the statistical model identification. IEEE transactions on automatic control. Vol. 19(6) p. 716–723. DOI 10.1109/TAC.1974.1100705.

Anokye R., Acheampong E., Owusu I., Obeng E.I. 2018. Time series analysis of malaria in Kumasi: Using ARIMA models to forecast future incidence. Cogent Social Sciences. Vol. 4(1) p. 1461544. DOI 10.1080/23311886.2018.1461544.

Arumugam P., Saranya R. 2018. Outlier detection and missing value in seasonal ARIMA model using rainfall data. Materials Today. Proceedings. Vol. 5(1) p. 1791-1799. DOI 10.1016/j.matpr.2017.11.277.

Baiocchi G., Distaso W. 2003. GRETL: Econometric software for the GNU generation. Journal of Applied Econometrics. Vol. 18(1) p. 105–110. DOI 10.1002/jae.704.

Box G.E., Jenkins G.M. 1976. Time series analysis: forecasting and control. Michigan. Holden-Day. ISBN 0816211043 pp. 575.

BOX G.E., JENKINS G.M., REINSEL G.C., LJUNG G.M. 2015. Time series analysis: forecasting and control. Hoboken. John Wiley & Sons. ISBN 978-1-118-67502-1 pp. 712.

BROCKWELL P.J., DAVIS R.A., CALDER M.V. 2002. Introduction to time series and forecasting. 2nd ed. New York. Springer. ISBN 0-387-95351-5 pp. 434.

CHEN P., NIU A., LIU D., JIANG W., MA B. 2018. Time series forecasting of temperatures using SARIMA: An example from Nanjing. In: IOP Conference Series: Materials Science and Engineering. Vol. 394(5), 052024. DOI 10.1088/1757-899X/394/5/052024.

CRYER J.D., CHAN K.S. 2008. Time series analysis: With applications in R. 1st ed. New York. Springer. ISBN 978-0-387-75958-6 pp. 491.

DICKEY D.A., FULLER W.A. 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association. Vol. 74(366) p. 427–431. DOI 10.2307/2286348.

HAMILTON J.D. 1994. Time series analysis Princeton. Princeton University Press. ISBN 9780691042893 pp. 816.

HANNAN E.J., QUINN B.G. 1979. The determination of the order of an autoregression. Journal of the Royal Statistical Society. Series B (Methodological). Vol. 41 p. 190-195. DOI 10.1111/j.2517-6161.1979.tb01072.x.

HAUMANN K. 2006. Appendix 2 system operation report Letaba Catchment Reserve determination – Operation report. DWAF. South Africa. Available at: http://www.dwa.gov.za/rdm/documents/Reouce%20Units%20Report_Appendix%202.pdf

HIPEL K.W., MCLEOD A.I., LENNOX W.C. 1977. Advances in Box-Jenkins modeling: 1. Model construction. Water Resources Research. Vol. 13(3) p. 567–572. DOI 10.1029/WR013i003p00567.

KENDALL M. 1975. Multivariate analysis. London. Charles Griffin & Co. Ltd. ISBN 0852642342 pp. 210.

KING J., PIENAAR H. (eds.) 2011. Sustainable use of South Africa's inland waters: A situation assessment of resource directed measures 12 years after the 1998 National Water Act. Water Research Commission Report. No. TT 491/11. ISBN 978-1-4312-0129-7 pp. 245.

KWIATKOWSKI D., PHILLIPS P.C., SCHMIDT P., SHIN Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics. Vol. 54(1–3) p. 159 178. DOI 10.1016/0304-4076(92)90104-Y.

LJUNG G.M., BOX G.E. 1978. On a measure of lack of fit in time series models. Biometrika. Vol. 65(2) p. 297–303. DOI 10.1093/biomet/65.2.297.

MURAT M., MALINOWSKA I., GOS M., KRZYSZCZAK J. 2018. Forecasting daily meteorological time series using ARIMA and regression models. International Agrophysics. Vol. 32(2) p. 253–264. DOI 10.1515/intag-2017-0007.

NYATUAME M., AGODZO S.K. 2018. Stochastic ARIMA model for annual rainfall and maximum temperature forecasting over Tordzie watershed in Ghana. Journal of Water and Land Development. No. 37 p. 127–140. DOI 10.2478/jwld-2018-0032.

PAPALASKARIS T., PANAGIOTIDIS T., PANTRAKIS A. 2016. Stochastic monthly rainfall time series analysis, modeling and forecasting in Kavala City, Greece, North-Eastern Mediterranean Basin. Procedia Engineering. Vol. 162 p. 254–263. DOI 10.1016/j.proeng.2016.11.054.

PHILLIPS P.C., PERRON P. 1988. Testing for a unit root in time series regression. Biometrika. Vol. 75(2) p. 335–346. DOI 10.1093/biomet/75.2.335.

SALAS J.D., DELLEUR J.R., YEVJEVICH V., LANE W.L. 1980. Applied modeling of hydrologic time series. Littleton. Water Resor. Pub. ISBN 0918334373 pp. 484.

SCHWARZ G. 1978. Estimating the dimension of a model. The annals of statistics. Vol. 6(2) p. 461–464. DOI 10.1214/aos/1176344136.

SHAFAEI M., ADAMOWSKI J., FAKHERI-FARD A., DINPASHOH Y., ADAMOWSKI K. 2016. A wavelet-SARIMA-ANN hybrid model for precipitation forecasting. Journal of Water and Land Development. No. 28 p. 27–36. DOI 10.1515/jwld-2016-0003.

SINGH M., SINGH R., SHINDE V. 2011. Application of software packages for monthly stream flow forecasting of Kangsabati River in India. International Journal of Computer Applications. Vol. 20(3) p. 7–14. DOI 10.5120/2416-3231.

TADESSE K.B., DINKA M.O. 2017. Application of SARIMA model to forecasting monthly flows in Waterval River, South Africa. Journal of Water and Land Development. No. 35 p. 229–236. DOI 10.1515/jwld-2017-0088.

TADESSE K.B., DINKA M.O. ALAMIREW T., MOGES S.A. 2017. Evaluation of sseasonal autoregressive integrated moving average models for river flow forecasting. American Journal of Environmental Sciences. Vol. 13(5) p. 378-387. DOI 10.3844/ajessp.2017.378.387.

TAKELE R., GEBRETSIDIK S. 2015. Prediction of long-term pattern and its extreme event frequency of rainfall in Dire Dawa Region, Eastern Ethiopia. Journal of Climatology & Weather Forecasting. Vol. 3 (1) p. 1–15. DOI 10.4172/2332-2594.1000130.