International Journal of Earth, Energy and Environmental Sciences
ISSN: 2517-942X
Vol:12, No:5, 2018

Time Series Modelling and Prediction of River Runoff: Case Study of Karkheh River, Iran

Karim Hamidi Machekposhti, Hossein Sedghi, Abdolrasoul Telvari, Hossein Babazadeh

Abstract**—**Rainfall and runoff phenomenon is a chaotic and complex outcome of nature which requires sophisticated modelling and simulation methods for explanation and use. Time Series modelling allows runoff data analysis and can be used as forecasting tool. In the paper attempt is made to model river runoff data and predict the future behavioural pattern of river based on annual past observations of annual river runoff. The river runoff analysis and predict are done using ARIMA model. For evaluating the efficiency of prediction to hydrological events such as rainfall, runoff and etc., we use the statistical formulae applicable. The good agreement between predicted and observation river runoff coefficient of determination (R^2) display that the ARIMA (4,1,1) is the suitable model for predicting Karkheh River runoff at Iran.

Keywords**—**Time series modelling, ARIMA model, River runoff, Karkheh River, CLS method.

I. Introduction

RAINFALL and runoff are highly nonlinear and complex outcomes of nature which require artificial data analysis for accurate modelling and simulation. Rainfall and runoff modelling is of particular relevance for engineering applications of water as well as in flood disaster management. Flood forecast modelling specifies the trends of river level (rising or falling) and the measure of runoff (low/medium/ high) to assess the likely loss of property and life [18]. Simulation of river runoff series over time is an essential part for real time quantification of water for planning and management of water resource systems. Real time flood predicting is an effective non-structural methodology for flood management. Flood estimation can be calculated in nature to provide sufficient time for hydrologists to implement flood management strategies well. The stochastic time series modelling (ARIMA model) approach can be used to characterize and predict river runoff [13].

In the past, time series studies of rainfall-runoff processes consisted of incorporation of available annual hydrologic data in time dependent or independent stochastic components, and further identification of trends and cycles [10], [21]. The simplest time series model, which deals with only one type of data, has three components to describe the linear stochastic process: Autoregression (AR), Integration (I) and Moving Average (MA) which can be combined as the Autoregressive Integrated Moving Average (ARIMA) model [13], [2]. Various other linear and dynamic regression derivatives of the ARIMA process, including PARIMA, SARIMA, DARIMA, ARMAX, NRL, MLR, and VARMA were developed over the years [7], [9], [13], [16], [20]. These models have long been applied in the modelling of rainfall runoff and forecasting (predicting) stream flows and floods [1], [11], [14].

Graupe et al. used ARIMA model for simulating water flow in Karstic basins [6]. O'Connell (1980) used time series to forecast flood episode consequent of rainfall occurrence [15]. Chang and Tiao [3] attempted an accurate flood forecast systems using time series methods. Chiew et al. [5] and Cheng [4] tried ARIMA models to forecasts rainfall-runoff at different catchments of varying scales. Naill and Momani mixed time series models with other mathematical models to devise new approaches [10]. Shakeel et al. applied Time Series Modelling of AMF (annual maximum flow) Indus River at Sukkur India and suggested that ARIMA (2,1,1) is suitable for studied series [8], [17]. Stojković et al. studied stochastic structure of annual discharges of large European rivers. They suggested that the stochastic flows simulated by the stochastic model can be used for hydrological simulations in river basins [8], [19]. Also, Nigam et al. applied the stochastic ARIMA model for flood (maximum river flow) forecasting on the Kulfo River with mean monthly runoff data. They found that a higher order ARIMA model may produce excellent results for three to six months forecasts [8], [12].

The selection of the modelling and forecasting method used in this work was based upon the hydraulic characteristics of river flow and runoff events. Our assessment of the study of the above cited literatures led us to conclude that ARIMA time series modelling may yield reliable and better forecasts (flow estimation) for the perennial Karkheh River and we therefore applied the ARIMA time series modelling approach to forecast runoff of this river. Also this paper aims to describe an accurate modelling method to forecast the runoff of the Karkheh River, which flows through the western part of Iran. The water in this river is mainly utilized for irrigation and water supply for public consumption. To estimation of the volume of Karkheh River runoff can be effectively employed to plan irrigation plans and water supply projects.

In this study the river runoff of the third largest river in Iran named Karkheh is modelled. The Karkheh River spans Karkheh basin in west of the Iran, located in the central and southern regions of the Zagros mountain range. This basin is

Karim Hamidi machekposhti is the Ph.D. Student in Water Sciences and Engineering Department, Science and Research Branch, Islamic Azad University, Tehran, Iran (phone: 00989197704935- fax: 00982144865179-82; e-mail: karim.hamidi@srbiau.ac.ir).

Hossein Sedghi and Hossein Babazadeh are with Water Sciences and Engineering Department, Science and Research Branch, Islamic Azad University, Tehran, Iran (e-mail: h.sedghi1320@gmail.com, h_babazadeh@srbiau.ac.ir).

Abdolrasoul Telvari is with the Civil Engineering Department, Islamic Azad University, Ahvaz, Iran, (e-mail: telvari@gmail.com).

divided into five subbasins viz. Gamasiab, Qarasou, Kashkan, Saymareh and south Karkheh. The river catchment area is about 51000 km^2 and river long is 900 km. The Karkheh River is mainly utilized for irrigation and public consumption. The river runoff data is collected at Jelogir Majin gauge station in south Karkheh sub basin at upstream of Karkheh dam. Fig. 1 shows the study area location. The data were taken from Iran Water Resources Management Organization (IWRMO).

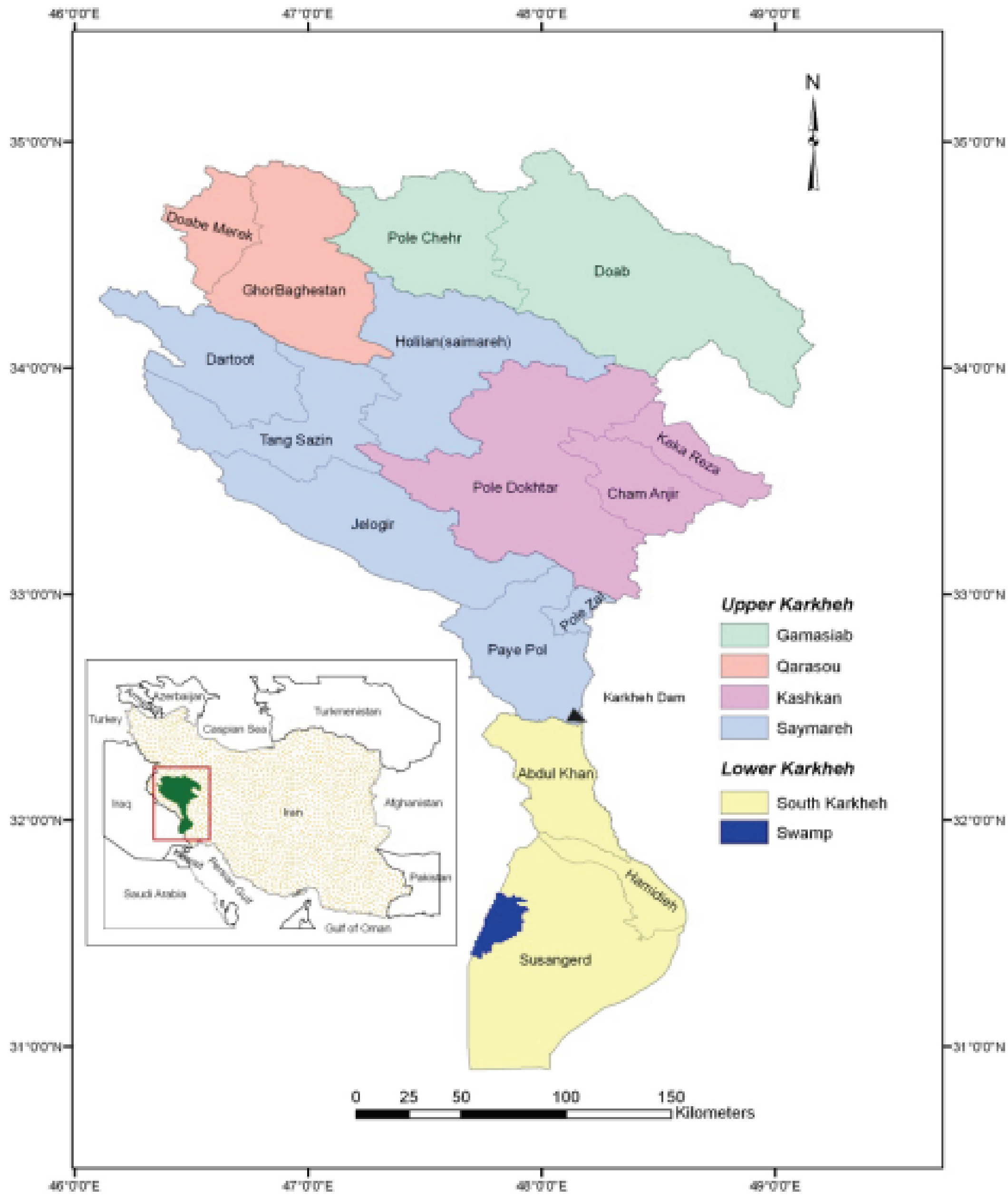

Fig. 1 The study area location

II. Methods

In this study an attempt is made to predict river runoff using annual average data. The univariate time series modelling approach Box et al. (1994) for the prediction of next time ahead predicting is very practical and versatile for observations having more or less linear trends. Box- Jenkins approach makes use of three linear filters: the autoregressive, the integration, and the moving average filter [2]. The general ARIMA equation is given as:

$$(\phi(B))y_t = (\theta(B))a_t \tag{1}$$

An expansion of general form of a nonseasonal univariate model is structurally described by

International Journal of Earth, Energy and Environmental Sciences
ISSN: 2517-942X
Vol:12, No:5, 2018

$$(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)y_t=(1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q)a_t \quad (2)$$

where y_t – stationary series after differencing = (1–B) d Y_t, Y_t – The input variable with reference to time, d – The number of nonseasonal differencing, Φ and θ – The nonseasonal autoregressive and moving average term's order p and q respectively, B –The Backward shift operator, define as Bzt = z_{t-1}, a_t – The residual (white noise term).

The order p and q of AR and MA terms can be determined by examining the ACF and PACF graphs. The three stage standard modelling procedure (identification, parameters estimation and diagnostic check) was used to develop time series models. SAS and SPSS software are used for model development.

The basic methodology of ARIMA modelling is shown in Fig. 2:

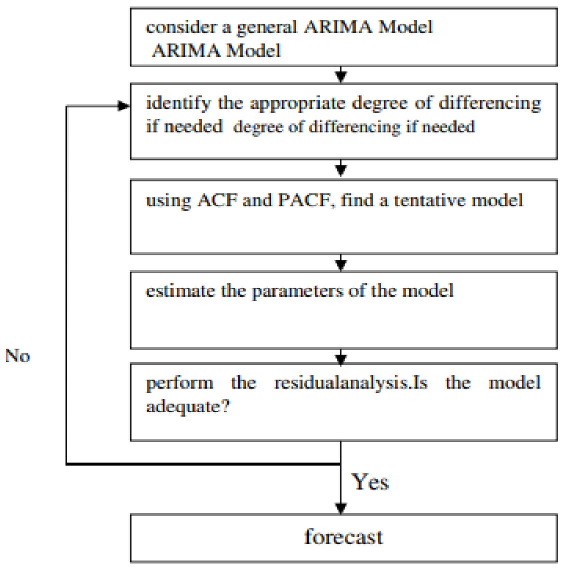

Fig. 2 ARIMA model development

III. Result and Discussion

Statistical analysis of river runoff data: The 48-year time series plot of average annual runoff (year 1958-2005) and the 40-year time series plot of average annual rainfall (year 1966-2005) of Karkheh River are plotted in Figs. 3 and 4. These figures show the time dependent quantitative variations and the patterns of the rainfall and related runoff.

Fig. 5 shows the histogram and normal frequency distribution curve of Karkheh River runoff. Obviously the 48 years runoff frequency distribution is not normal and there exists some extremes runoff events (both low and high) which all fall outside the normal distribution curve.

Table I showed the descriptive statistics of Karkheh River runoff data. There is a vast change in the minimum and maximum Karkheh River runoff values. Small values of skewness and kurtosis are indicative of concentrated and nearly centralized data close to mean data of annual mean runoff in Karkheh River.

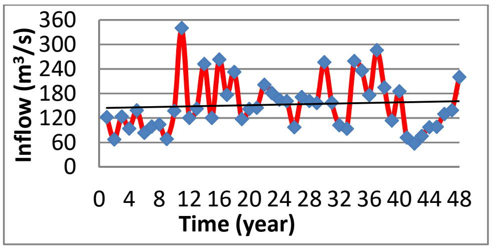

Fig 3 Time series of Karkheh River

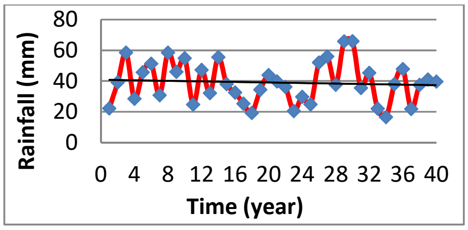

Fig. 4 Time series of Karkheh River

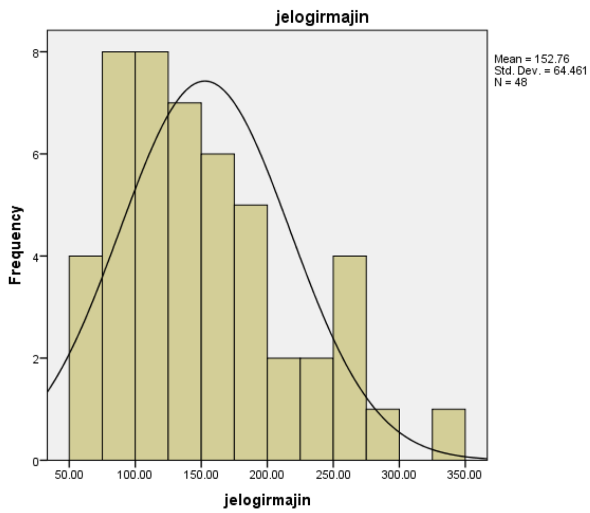

Fig. 5 Histogram and frequency of runoff data of Karkheh River

TABLE I
Descriptive Statistics of Runoff Data

SN	Statistical Parameters	Values
1	Total Variable	48
2	Mean	152.756
3	Std. Error of Mean Deviation	9.304
4	Median	140.23
5	Mode	57.09
6	Std. Deviation	64.46
7	Variance	4155.22
8	Skewness	0.852
9	Kurtosis	0.337
10	Minimum	57.09
11	Maximum	340.25

A. Fitting Box-Jenkins

For select a suitable ARIMA model based on Box- Jenkins approach, we need to identify the order (p,d,q) for a nonseasonal univariate model and also to find out the degree of best fit nonseasonality, which provide a parsimonious representation for both the stochastic component and the total series under consideration. The nonseasonal ARIMA (p,d,q) model, which is called nonseasonal autoregressive integrated moving average model, is fitted to a time series by using a three stages procedure (identification, parameters estimation and diagnostic check). The plotted of observation data against time will show up important features of a hydrologic series such as trend, seasonality, discontinuities and outliers. Fig. 3 shows that there is little increase trend for river runoff. The natural runoff series are not stationary, therefore we use differencing to achieve a stationary series (d=1). This plot shows in Fig. 6. They are stationary and there is no trend.

B. Identification of Representative Models

In identification stage we use the data and any information on how the series was generated, to suggest a subclass of models from the general Box- Jenkins family. In other word, identification stage helps to choice of the order of p, d and q. In practice, the degree (grade) of d assumes 1 using ACF and PACF of the series that is plotted by SPSS or SAS software. The ACF and PACF of the natural runoff for d=0 and d=1 shown in Figs. 7 and 8. The runoff of the Karkheh River shows a nonseasonal pattern, the same can be seen in the ACF and PACF plots, and hence, the flow pattern requires a nonseasonal model. An analysis of significant ACF and PACF plots implies to the first and fourth order nonseasonal ARMA parameterization of runoff series.

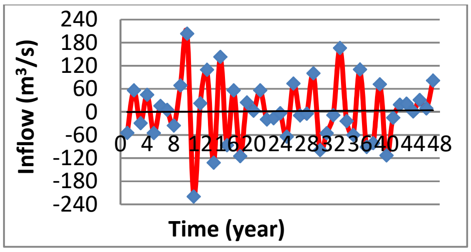

Fig. 6 Time series of Karkheh River (d=1)

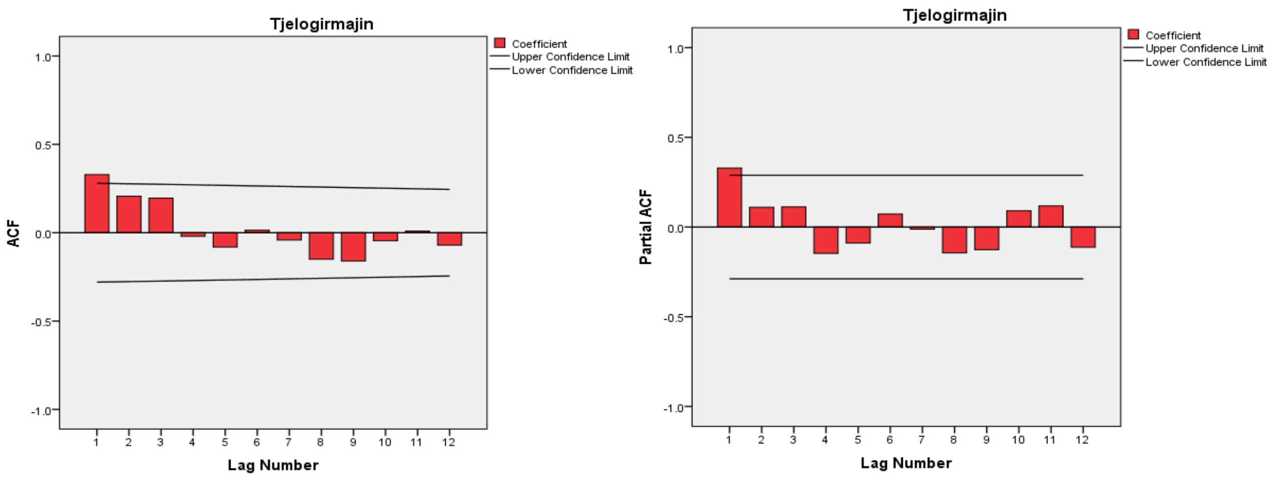

Fig. 7 ACF and PACF for Series (d=0)

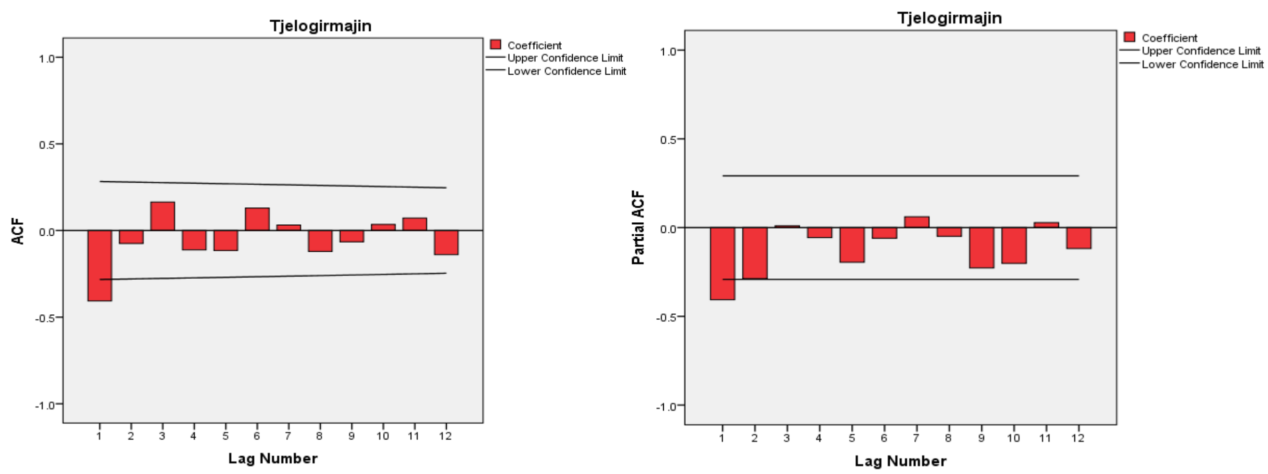

Fig. 8 ACF and PACF for Series (d=1)

C. Estimation of the Model Parameters

In parameter estimation stage there are three methods foe estimating. This method as follows: Maximum Likelihood (ML), Conditional Least Square (CLS) and Unconditional Least Square (ULS) methods. In this study, we used these methods for the estimation of parameter river runoff for three suggested models in previous stage. This models are ARIMA(1,1,0), ARIMA(1,1,1) and ARIMA (4,1,1). The results of values for the selected parameters show in Table II. All three models are suitable for modelling and entrance to next stage because models have two stationary and invertibility conditions. Therefore, we can enter to the diagnostic check stage.

TABLE II
VALUES OF NONSEASONAL ARIMA MODEL PARAMETERS FOR MAD

Estimation Method	Type (Order) and Values of parameters ARIMA(p,1,q)	Std. Error Coefficient	Absolute Value of t	Probability of t	Stationary Condition	Invertibility Condition
ML	p(1) = -0.41932	0.13457	-3.12	0.0018	Satisfy	
	q(0)					
CLS	p(1) = -0.41473	0.13576	-3.05	0.00375	Satisfy	
	q(0)					
ULS	p(1) = -0.42872	0.13480	-3.18	0.0026	Satisfy	
	q(0)					
ML	p(1) = 0.36414	0.15423	2.36	0.0182	Satisfy	
	q(1) = 0.99994	52.247	0.02	0.9847		Not Satisfy
CLS	p(1) = 0.34486	0.15246	2.26	0.0286	Satisfy	
	q(1) = 0.966303	0.0427	22.62	0.0001<		Satisfy
ULS	p(1) = 0.35777	0.14375	2.49	0.0166	Satisfy	
	q(1) = 0.99998	0.29826	3.35	0.0016<		Not Satisfy
ML	p(4) = -0.1636	0.15570	-1.5	0.2933	Satisfy	
	q(1) = 0.66032	0.12124	5.45	0.0001<		Satisfy
CLS	p(4) = -0.17507	0.15914	-1.1	0.2771	Satisfy	
	q(1) = 0.68987	0.11569	5.96	0.0001<		Satisfy
ULS	p(4) = -0.18146	0.15870	-1.14	0.2589	Satisfy	
	q(1) = 0.689354	0.11617	5.93	0.0001<		Satisfy

ML: Maximum Likelihood CLS: Conditional Least Square ULS: Unconditional Least Square

TABLE III
RESULT OF AUTOCORRELATION CHECK OF RESIDUALS

ARIMA Model	Estimation Method	To Lag	Df	Chi-Square	Pr>Chi Square	Adequacy for Modelling
ARIMA(1,1,0)	ML	6	5	8.04	0.1540	Satisfy
		12	11	11.80	0.3792	
		18	17	17.84	0.3990	
		24	23	26.29	0.2875	
	CLS	6	5	7.87	0.1636	Satisfy
		12	11	11.83	0.3764	
		18	17	17.88	0.3963	
		24	23	26.24	0.2898	
	ULS	6	5	8.08	0.1517	Satisfy
		12	11	11.87	0.3733	
		18	17	17.97	0.3906	
		24	23	26.52	0.2771	
ARIMA(1,1,1)	CLS	6	4	3.17	0.5296	Satisfy
		12	10	4.99	0.8920	
		18	16	8.33	0.9386	
		24	22	11.96	0.9583	
ARIMA(4,1,1)	ML	6	4	2.98	0.5618	Satisfy
		12	10	6.31	0.7883	
		18	16	11.02	0.8082	
		24	22	16.88	0.7699	
	CLS	6	4	3.1	0.5416	Satisfy
		12	10	7.19	0.7074	
		18	16	11.74	0.7617	
		24	22	17.32	0.7455	
	ULS	6	4	3.01	0.557	Satisfy
		12	10	6.34	0.7858	
		18	16	10.94	0.8134	
		24	22	16.57	0.7866	

ML: Maximum Likelihood CLS: Conditional Least Square ULS: Unconditional Least Square

D. Diagnostic Check

Diagnostic checking tests are applied to see if the model is adequate or not. There are two statistical tests for this purpose. The tests are Portmanteau test, residual auto-correlation and partial auto-correlation functions tests.

E. Portmanteau Test

It is a test for independency of the residuals and uses the Q statistic defined as:

$$Q = (N - d - DS)\sum_{k=1}^{M} r_k^2 a_t \qquad (3)$$

where r_k (a_t) is the autocorrelation coefficient of the residual (a_t) at lag k, and M is the maximum lag considered (about N/4), ARIMA model is considered adequate if p>chi square is greater than the level of significant 0.05 [8]. Table III shows the results of this test for river runoff. According to this table all three selected models are adequate for runoff data.

F. Residual Autocorrelation Function Test

This test uses for determined independence or dependence of residuals of selected model. The residuals have independence if autocorrelation and partial autocorrelation function (RACF and RPACF) were not significant. RACF and RPACF of the studied series are shown in Fig. 9. Hence, the suggested model can be considered as a suitable model [8]. Table IV displays the goodness of fit statistic. Therefore, the ARIMA (4,1,1) model and, the CLS estimation parameter method is the best model for predicting runoff for Karkheh River.

G. Forecasting

Predicting (forecasting) and observation of river runoff for 10 years (the period 2006-7 to 2015-2016) applying the best model is given in Table IV. Fig. 10 shows the predicted and observed series for the studied data series. Also, Fig. 11 shows the corresponding observed and predicted values. According to the figure, there is a good agreement between observed and predicted values (R^2=0.85). it is confirmed that the selected model is adequate.

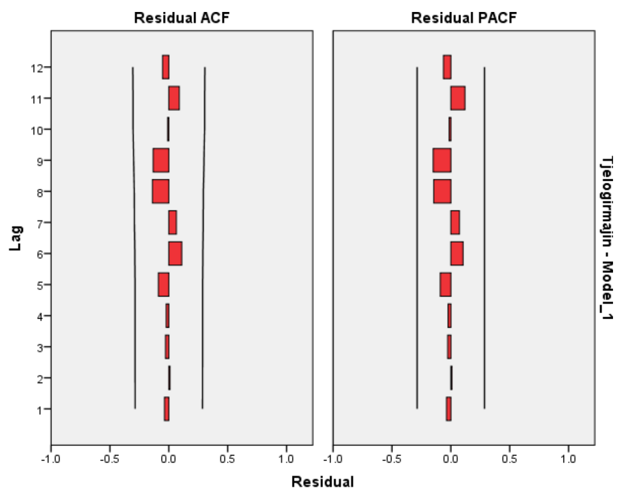

Fig. 9 Correlogram of Residual Series Parameter

IV. Conclusion

The conclusions obtained from this study are as below:

- Time series plots of historical rainfall and runoff data of the Karkheh River in Iran indicate that river runoff rapidly rises when there is precipitation, and quickly decreases after the rain ends, but the river runoff continues to reduce for a long time after the precipitation ends because of contribution from river base-flow. The predict graphs are insensitive to extreme variations in river runoff.

TABLE IV
Forecasts and Observations of Runoff from Period 2006-7 to 2015-16

Period	Forecast	Observation
2006-7	139	154
2007-8	140	150
2008-9	131	127
2009-10	141	150
2010-11	131	125
2011-12	132	127
2012-13	135	130
2013-14	140	148
2014-15	133	140
2015-16	132	130

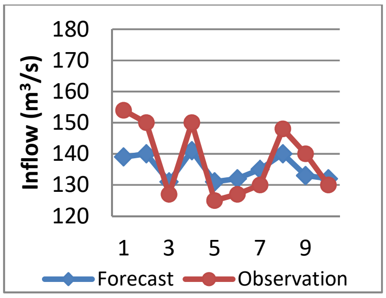

Fig. 10 Comparison of Forecast and Observed data (2006-2015)

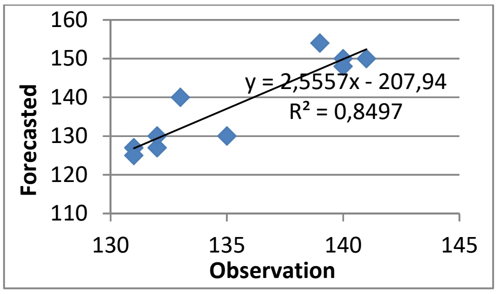

Fig. 11 Correlation between actual values and forecasted values in Karkheh River

- The good appropriate method for modelling the linear and uniform patterns of river runoff is ARIMA process. Due to the highly variable nature of river runoff, a parsimonious model could not always provide best fit predicts for an entire span of data. In most practical cases, insufficient data prevent the use of a parsimonious model. Nevertheless, an ARIMA model can include both the high and low values to provide a reasonable quantification of river runoff during an odd or unusual river flow situation, and these models could be used to predict upcoming abrupt changes in river runoff conditions.

- An ARIMA (4,1,1) model in CLS method is suggested for the estimation of Karkheh River runoff in future and advance warning of floods in Karkheh River at Iran.
- The grade of 4 in model (P=4) for runoff indicates a substantial degree of variability and dependence in the stochastic component. The values of higher order durability on the other hand speak of a fairly uniform degree of dependence in the stochastic components on past events. It is found that the simulated values have a good agreement with the observed runoff for the future 10 years. It is showed that ARIMA modelling is an appropriate approach to model hydrological data which often exhibits autocorrelation with time and cannot be done by simple statistical predicting methods like regression analysis and etc. The Box-Jenkins approach (time series modelling) can show the relationship between variables. If we add effect of rainfall data too in runoff series, i.e. multivariate approach, this will help in to achieve an assumption of a cause effect relationship in time series analysis.
- Factors such as good vegetation of the region and water from snowmelt can be the significant ACF and PACF functions with high orders. In other word, the good vegetation of the region and the forest causes water retention in the soil surface layer and delay in the rise in surface runoff.
- For short term predicting of hydrologic phenomenon like runoff, rainfall, flood and etc. suggest the ARIMA models. The auto-regressive model is a finite memory model, thus it does not fare well in long-term predicting of this hydrologic phenomenon.

References

[1] Bender, M., Simonovic, S. "Time series modeling for long range stream flow forecasting," *Journal of Water Resources Planning and Management,* 1994, 120(6): 857–870.

[2] Box, G.E.P.; Jenkins, G.M. "*Time Series Analysis: Forecasting and Control,*" Holden Day, San Francisco, California, 1976.

[3] Chang, L. C., Tiao, J. "Microcomputer application in stochastic hydrology," *Journal of Hydraulic Division, ASCE,* 1985. 371-375.

[4] Cheng, Y. "Evaluating an autoregressive model for Stream flow forecasting," *Conference Proceeding of Hydraulic Engineering;* 1994, 1105–1109.

[5] Chiew, F. H. S., Stewardson, M. J., McMahon, T. A. "Comparison of six Rainfall-Runoff Modelling Approaches," *Journal of Hydrology;* 1993, 147: 1-36.

[6] Graupe, D., Isailovic, D., Yevjevich, V. "Prediction model for runoff from Karstified catchments," *In Proceedings of the U.S.-Yugoslavian Symposium on Karst Hydrology and Water Resources,* Dubrovnik; 1976, 277–300.

[7] Haltiner, J. P., Salas, J. D. "Short-term forecasting of snowmelt discharge using ARMAX models," *Journal of the American Water Resources Association,* 1988, 24(5):1083- 1089.

[8] Hamidi machekposhti K., Sedghi H., Abdolrasoul Telvari A., Babazadeh, H. "Forecasting by Stochastic Models to Inflow of Karkheh Dam at Iran," *Civil Engineering Journal,* 2017, Vol. 3, No. 5, 340-350.

[9] Hipel, K. W., McLeod, A. I., Lennox, W. C. "Advances in box Jenkins modeling," *Journal Water Res. Res.,* 1977, 13(3): 567-575.

[10] Naill, P. E., Momani, M. "Time Series Analysis Model for Rainfall Data in Jordan: CS," *American J. of Env. Sc.,* 2009, 5(5): 599-604.

[11] Nigam, R.; Nigam, S.; Mittal, S. K. "The river runoff forecast based on the modeling of time series," *Russian Meteorology and Hydrology,* 2014, 39(11), 750-761.

[12] Nigam, R., Bux, S., Nigam, S., Pardasani, K. R., Mittal, S. K., Haque, R. "Time series modeling and forecast of river flow," *Current World Environment, An International Research Journal of Environmental Science;* 2009, 4(1): 79-87.

[13] Nigam, R., Nigam, S., Mittal, S.K. "Modeling tropical river runoff: A time dependent approach," *Sciences in Cold and Arid Regions,* 2014, 6(3): 247-256.

[14] Nigam, R., Nigam, S., Kapoor, S. "Time Series Modeling Of Tropical River Runoff," *International Journal Of Pure And Applied Research In Engineering And Technology (IJPRET),* 2013,1(7): 13-29.

[15] O'Connell, P. E. "Real Time Hydrological Forecasting and Control," *Institute of Hydrology,* UK, 1980, 264-295.

[16] Salas, J. D., Deulleur, J. W., Yevjevich, V., Lane, W. L. "*Applied modeling of hydrologic time series,*" Water Resources publication, Littleton, Colorado; 1980.

[17] Shakeel, A.M.; Idrees, A.M.; Naeem, H.M.; Sarwar, B.M. "Time Series Modelling of Annual Maximum Flow of River Indus at Sukkur," *Pakistan. Journal of Agricultural Sciences,* 1993, 30(1), 36-38.

[18] Singh, V.P., Woolhiser, D.A. "Mathematical modeling of watershed hydrology," *J. of Hydrology Engineering;* 2002, 7(4): 270–292.

[19] Stojković, M.; Prohaska, S.; Plavšić, J. "Stochastic structure of annual discharges of large European rivers," *J. Hydrology and Hydromechanics,* 2015, 63(1), 63–70.

[20] Wang, W. *"Sochasticity, Nonlinearity and Forecasting of Stream flow Processes,"* IOS Press, technical university of Delft, the Netherlands, 2006.

[21] Yevjevich, V. M. *"Stochastic processes in Hydrology,"* Water Resource Pub., FC, Colorado, 1972.