Ismail Y., Atan I.B., Jaafar J., Ashaari Y., Samsudin M.B., Khatib M.R., Baki A.B.M, Mohamed S.N. (2023), Generation of synthetic streamflow data using Arma model, pp. 178-188. In Gastescu, P., Bretcan, P. (edit., 2023), Water resources and wetlands, 6th International Hybrid Conference Water resources and wetlands, 13-17 September 2023, Tulcea (Romania). pp. 287
Available online at http://www.limnology.ro/wrw2023/proceedings.html

6th International Hybrid Conference Water resources and wetlands, 13-17 September 2023, Tulcea (Romania)

GENERATION OF SYNTHETIC STREAMFLOW DATA USING ARMA MODEL

Yulinar ISMAIL [1], Ismail bin ATAN [1], Jurina JAAFAR [1], Yasmin ASHAARI [1], Mohd Bakri SAMSUDIN [2], Mohd Riza KHATIB [2], Aminuddin bin Mohd BAKI [2,3,*], Shahrul Nizam MOHAMED[4]

1. School of Civil Engineering, UiTM,Shah Alam 40450, Selangor, Malaysia
2. Envirab Services, P.O. Box 7866, GPO Shah Alam 40730, Selangor, Malaysia
3. Faculty of Engineering, City University, Menara CityU, Petaling Jaya, Selangor, Malaysia
4. School of Civil Engineering UiTM Johor, Pasir Gudang, Johor, Malaysia

- Corresponding author: aminbaki@msn.com

Abstract. Stochastic models are required in order to generate synthetic series of flows statistically similar to observed ones for use in simulation studies of water resources management. Autoregressive- Moving Average Model is used in this study where a number of parameters are estimated from observed daily flows data for three stations at Pahang River. The parameters are used for generated the synthetic series The ability of the model is measured on the statistical characteristic such as mean, standard deviation, minimum and maximum flow and also the behavior of structure flow that produced by synthetic series. From the analysis, the model is able to preserve the statistical characteristics only for Pahang River at Lubok Paku and Pahang River at Temerloh however it is not satisfactory for Pahang River at Kg Sungai Yap.
Keywords: ARMA model, daily flow, statistical characteristic

1.INTRODUCTION

Synthetic streamflow is very useful in reservoir simulation studies and water resources planning and management. It describes the character of streamflow sequences that may occur in the future. This sequence is widely used to evaluate the ability of the existing or proposed water supply and hydroelectric system. Design, long term planning, and operation planning can benefit from information about the likelihood and likely character of possible drought streamflow.

Sharma et al. (1997) cited that it is very important to generate synthetic streamflow sequences to analyze alternative designs, operation policies, and rules for water resources systems, and that the dependence structure of streamflow sequences is often assumed to be Markovian, that is, dependent on only a fine set of prior values.

Vogel and Stedinger (1988) show that use of stochastic hydrology is likely to result in more precise estimates of over year storage requirement than using just the drought of the record. In order to understand the variability of future streamflow system performance, the alternative is by using the stochastic streamflow model. The model will generate the synthetic flow sequences that are statistically similar to the observed streamflow records or historical streamflow data.

Other researchers including Chakhchoukh (2010), Abo-Hammour *et al.* (2012), Huang *et al.* (2012), Karthikeyan and Kumar (2013), Marelli *et al.* (2013), Bou *et al.* (2013), Cao *et al.* (2013), Laner *et al.* (2013), Rout *et al.* (2014), Zhu *et al.* (2014), Zhu and Li (2015), Zhu *et al.* (2015), Zheng *et al.* (2015), Aghdam *et al.* (2015), Brockwell *et al.* (2016), Triacca (2016), Kumar and Mazumdar (2016), Bertha and Golinval (2017), Xiao *et al.* (2017), Wang *et al.* (2018), Baptista *et al.* (2018), Shen *et al.* (2018), Boubacar Maïnassara, and Saussereau (2018), Singh and Pozo (2019), Hossain *et al.* (2020), Hackstein *et al.* (2020), Moon *et al.* (2021), Berardengo *et al.* (2021), Li *et al.* (2021), Xu *et al.* (2022), You *et al.* (2022) also made similar observations.

The generated synthetic streamflow sequences will augment the performance and description which is provide in historical streamflow record or data; the synthetic sequences and critical periods, then all of the information serve as the basis of reservoir simulations, and possible reservoir systems performance.

On the other hand, the synthetic streamflow and reservoir simulations also perform the realization of drought or wet years that could occur and likely the reservoir systems performance during the events. The synthetic streamflow sequences also can be used to construct a probabilistic description of how the entire system is likely to perform that is not only tied o the particularly events and timing of the drought of record. Lettenmaier *et al.* (1987) said that the synthetic streamflow sequences can also be used to generated refined estimates of the probability that given powers targets can be met without failure due to drought and the capacity that may be available to avoid loss—of –load events in brief emergency situations.

The purpose of this study is to generate the synthetic streamflow using ARMA Models, which will represent the characteristic and statistical parameters that are approximately similar to the historical data.

2.0 FINDING FROM PREVIOUS RESEARCH

2.1 HISTORICAL BACKGROUND

Synthetic streamflow is an important subject in stochastic hydrology and has received a lot of attention in hydrologic literature. Synthetic streamflow was first used by Hazen, (1914) in studies of water supply reliability. Hazen created a 300-year synthetic record by combining the scale-adjusted records of fourteen streams; he then used his synthetic record to compute the probability of supply deficits for several demand levels. Hazen's synthetic record contained significantly less information than a 300-year single-site record, because concurrent flows in his streams were not statistically independent.

Barnes, (1954) used a different approach to create extended streamflows records. Barnes found that the 29 years of observed streamflows at a site in Australia were distributed almost normally, he generated 1000 years of synthetic flow data by selecting flows from a normal distribution with the same mean and variance as the historical series.

Maass *et al*, (1962); Thomas and Fiering, (1962) have developed the models which consider the correlations between consecutive monthly or annual flows. Then, (Fiering, 1967) has extended the models by describe non-normal marginal probability distributions. Other efforts were directed at generating reasonable sequences of concurrent flows at several sites (Beard, 1965b; Matalas, 1967).

Box and Jenkins (1970) established many of the current time series modeling techniques. They have developed a classification scheme for a large family of time series models. In this scheme, the Thomas-Fiering and multivariate Matalas models were denoted as AR(1), or autoregressive models of order 1, because they regress flows in one period on flows in the previous period. Box and Jenkins also discussed autoregressive models of arbitrary order p, or AR(p); moving average models of order q, or MA(q); and combinations of the two, which they called ARMA(p,q).

2.2 MATHEMATICAL FORMULATION

2.2.1 AR Models (Autoregressive Model)

A stationary time series y_t normally distributed with mean μ and variances σ^2, which has an autoregressive (Markovian) correlation (or time dependences structure) with constant parameter. The autoregressive model of order p, denoted y AR(p) representing the variable y_t may generally written as

$$y_t = \mu + \varphi_1 (y_{t-1} - \mu) + \ldots\ldots \ \varphi_p (y_{t-p} - \mu) + \varepsilon_t \qquad (1)$$

Regarding to Fiering (1971); Beard (1967), the autoregressive model are performed as

$$y_t = \mu + \sum_{j=1}^{p} \varphi_j (y_{t-j} - \mu) + \ldots\ldots \ \sigma(1-R)^{1/2} \xi_t \qquad (2)$$

While based on Yevjevich (1972) the equation as follow

$$y_t = \mu + \sigma z_t \qquad (3)$$

$$z_t = \mu + \sum_{j=1}^{p} \varphi_j z_{t-j} + \varepsilon_t \qquad (4)$$

or

$$z_t = \mu + \sum_{j=1}^{p} \varphi_j z_{t-j} + \sigma_\varepsilon \xi_t \qquad (.5)$$

According to Box and Jenkins (1970) the mathematical formulation of AR are as

$$y_t = \mu + \sum_{j=1}^{p} \varphi_j (y_{t-j} - \mu) + \varepsilon_t \quad \text{or} \qquad (6)$$

$$y_t = \mu + \sum_{j=1}^{p} \varphi_j (y_{t-j} - \mu) + \sigma_t \xi_t \qquad (7)$$

2.2.2 ARMA Models (Autoregressive- Moving Average Model)

Consider the originals periodic series $X_{V,T}$, where v denotes the year, $_T$=1,…..w and w is the number of time in the year .Assuming that the distribution of the series is skewed ,an appropriate transformation $X_{V,T}$ to the normal series $Y_{V,\tau}$.Then the periodic ARMA models for $Y_{V,T}$ can be write as

$$Y_{V,\tau} = \mu_\tau + \sigma_\tau z_{V,\tau} \qquad (8)$$

Where μ_τ and σ_τ are the periodic mean and periodic standard deviation need may be represented by an ARMA Models with either constant or time varying (periodic) coefficients.

The ARMA (p,q) model with constant coefficient is

$$z\tau = \sum_{j=1}^{p} \varphi j\, z_{\tau-j} - \sum_{j=1}^{p} \theta i\, \varepsilon_{\tau-j} + \varepsilon_\tau \qquad (9)$$

where τ = (v-1) w+ τ φ and θ are the coefficient of the models and ε_τ is the independent variable.
Tao, P.C and Delluer, J.W. (1976) used the ARMA (p,q) model with time – varying coefficient as

$$z\,\tau = \sum_{j=1}^{p} \varphi_{j,\tau} z_{v,\tau-j} - \sum_{j=1}^{q} \theta_{i,\tau} \varepsilon_{v,\tau-i} + \varepsilon_{v,\tau} \qquad (10)$$

where φ_j, $\theta_{i,\tau}$ are time varying autoregressive and moving average coefficients ,respectively and $\varepsilon_{v,}\tau$ is an independent and identically distributed normal random variable.

3.0 METHODOLOGY

3.1 Generating Synthetic Streamflow Procedures

I. The historical data was transformed using logarithm function where

$$y = \log(x) \qquad (11)$$

The data have to be transformed in order to improve the accuracy of parameter estimates.

II. The seasonality of the transformed data is removed to standardized series with mean equal to zero and standard deviation is one. The standardized flows are obtained from

$$y = \frac{\log(x) - \mu_y}{\sigma_y} \qquad (12)$$

III. The auto-covariance function c_k, the autocorrelation coefficient $r_k = c_k/s^2$, and the partial autocorrelation coefficient φ_k (k) for lags k going from 1 to at least N/4 but less than N is calculated.

IV. From the behavior of autocorrelation and partial autocorrelation functions, it infer the order of model, namely, the values of p and q which are likely to fit the series.

V. The initial estimate of the p autoregressive parameters φ_{1xi} and φ_{2xi-1},……and $\varphi_{pxi-p+1}$ by solving the p Yule-Walker equations

$$x_{i+1} = \varphi_{1xi} + \varphi_{2xi-1} + \cdot\cdot\cdot + \varphi_{pxi-p+1} + \xi_{i+1} \qquad (13)$$

This step is performed by means of MINITAB computer program.

VI. The initial estimates of the q moving average parameter obtained from of the series and also the autovariance function ck' of the zt' series is calculated. It can be calculated as usual. Alternatively, use Box and Jenkins (1976) formula for the c_k' in terms of ck of the z_t series and the φ already available from steps 3 and 4, respectively:
This step also performed by means of MINITAB computer program.

VII. Obtained the maximum likelihood estimate of the parameters. The residual calculate

$$\varepsilon_j = 0 \; ; \; j=1 ,……, \max (p,q)$$

$$j=1,2……N-p$$

the sum of square

$$S = \sum_{t=1}^{N} \varepsilon_t^{\,2} \qquad (14)$$

for several values of φ and θ around the initial estimates and obtain the values of the φ's and θ's for which S is minimum.
The autocorrelation function rk(ε) of the residual series ε_t for the lags k going from 1 to L =N/10 +p+q . The ε_t obtained from the MINITAB computer program. The statistic calculated

VIII. Generation of synthetic series. The series formula
This step is performing by using IMSL (RNARM/DNARM) program.

4.0 DATA ANALYSIS

4.1 Data Collection

In this study, the daily streamflow data of Pahang River which consist of 3 stations site collected from Drainage and Irrigation Department (DID) There are from three gauge stations as numbered station site 3424411, station site 3527410 and station site 4023412. All series of data considered as a historical data.

4.2 Fitting ARMA Models

The order of p and q is determined to identify the appropriate ARMA Model from the behavior of autocorrelation and partial autocorrelation functions presented in table 1. This step is performed by USING MINITAB MACRO Program as in appendices A.

Table 1. The order of AR (p) and MA(q) for each station site

Station site	AR (p)	MA(q)
3424411	3	0
3527410	3	0
4023412	3	0

4.3 Estimation of model parameters.

For this purpose, the method of maximum likehood is used to estimate the parameters. The parameters are presented in table 2.

Table 2. The parameters estimated of each station site

Station Site	AR (1)	AR (2)	AR (3)
3527410	1.3081	-0.5256	0.1749
3527410	1.5264	-0.7888	0.2421
4023412	1.1801	-0.4163	0.1944

4.4 Generating synthetic streamflow

The best model is chosen and it can be used for generation the synthetic data by using IMSL (RNARM/DNARM) program.

4.5 Comparison between the historical data and synthetic data

The final stages, is to compare the descriptive characteristic of historical data and synthetic data generated by the ARMA model such as mean, standard deviation and the minimum and the maximum flow of series.

5.0 RESULT AND DISCUSSION

The accuracy of the model to generate the synthetic data is measured by the ability of the model to reproduce the statistical characteristics similar to that of the historical data. The descriptive characteristic of the observed series after transformation and the generated series are presented in table 3. The mean and the standard deviation value of the observed series for each station site are seen to be well preserved since statistical characteristics of generated series is quite similar to the observed series.

Table 3. The descriptive statistic of transformed data and the generated transformed data

	Site 3424411		Site 3527410		Site 4023412	
	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean
transformedData	1	0	1	0	1	0
Generated transformed data Data	0.937	0.107	0.967	0.1292	0.75	0.1053

The other important characteristics of the historical series that to be considered in water management are the minimum and the maximum flow. The successful of the model to generate the synthetic data also counter by the ability of the model to reproduce the minimum and the maximum flow. It indicates that the minimum and the maximum flow value for the synthetic data are as good as the historical data (less than 20%) for station site 3527410 and station site 3527410.The minimum and the maximum flow for station site 4023412 is not good as expected with between 35% to 40% error.

Another approach to measure the ability of the model in order to generate the synthetic data is by looking at the autocorrelation function (ACF) and partial autocorrelation function (PACF). They indicate the presence of persistence structure in the data. The ACF and PACF also give information about the non-seasonal and seasonal AR and MA operators for a time series.

The identification of the appropriate parametric time series model depends on the shape of the ACF. Based on the Figure 1a, 2a and 3a, the ACF is significantly different from zero, this implies that there is dependence between observations for the historical data. By comparing the Figure 1a and 1b, it shows that the synthetic data have the same structure figure as the historical data. The same goes to by comparing the Figure 2a and 2b and also figure 3a and 3b.

While the PACF for historical data are shown as Figure 4a, Figure 5a, and Figure 6a. The observations indicate significant serial correlations (persistence) associated to daily flows (historical data) at station site 3424411, station site 3527410 and station site 4023412. While, from Figure 4b,5b and 6b, the generated synthetic data are able to reproduce the persistence effect for station site 3424411, station site 3527410 and station site 4023412.The structure flow that producing from the synthetic data from each station site is likely same as the historical data structure.

The result from the comparison between the historical flow and the generated flow are good for the Pahang River at Kg Sungai Paku (station site 3424411) and Pahang River at Temerloh (station site 3527410) but not so satisfactory for Pahang River at Kg Sungai Yap (station site 4023412). However, the method does seem to have some potential and would merit further investigation.

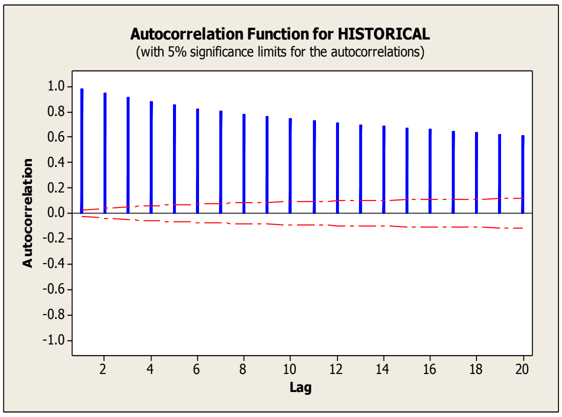
In order to test the ability of the proposed model to reproduces the statistical of the historical data, a synthetic data is generated (simulated) with the sample size equal to the historical data. The figure 7, figure 8 and figure 9 are shown times series plot of the synthetic data (simulation data) for each station site that have been generated by this model. It showed that the synthetic data seems to be adequate for simulating the flows series.

6.0 CONCLUSION

Simulation of the historical flow records are needed for reliable information for many water resources studies. From this study, it is seen that the ARMA model is able to reproduce descriptive characteristics such as mean, standard deviation, minimum and maximum flow of generated series as good (less than 20%) as observation data only for station site 3527410 and station site 3527410. The model does seem to have the potential and would merit further investigation for it allow to be used in water resources and management planning.

ACKNOWLEDGEMENT

Appreciation to the Department of Irrigation and Drainage (JPS) Malaysia for making their data available for this study



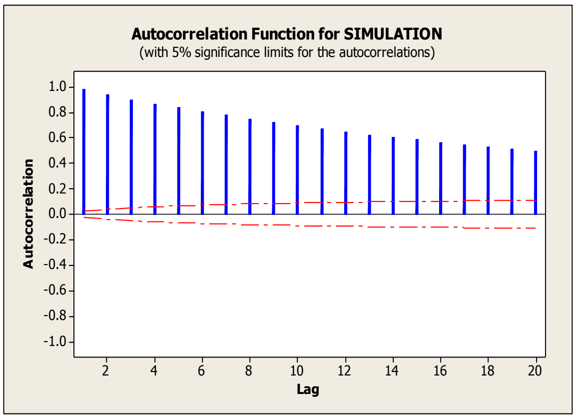
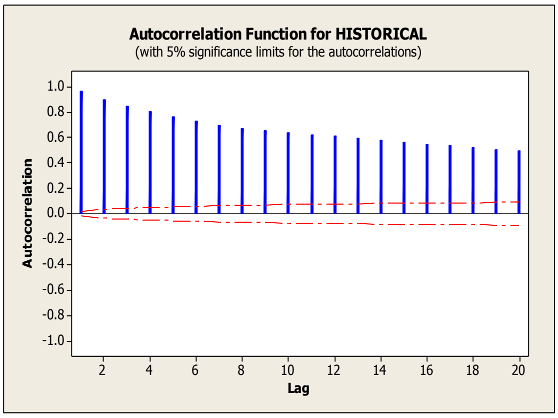


Figure 1a) ACF of Historical Data for for station site 3527410

Figure 1b) ACF of Simulation Data for station site 3527410



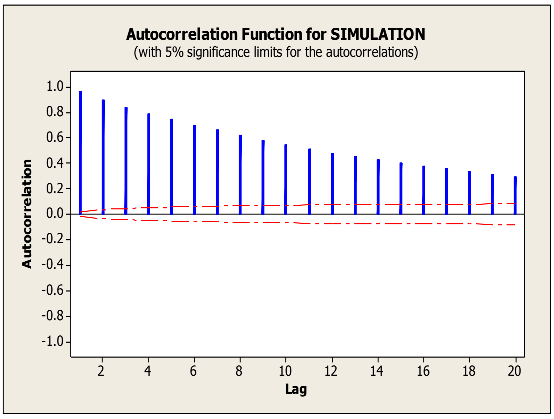
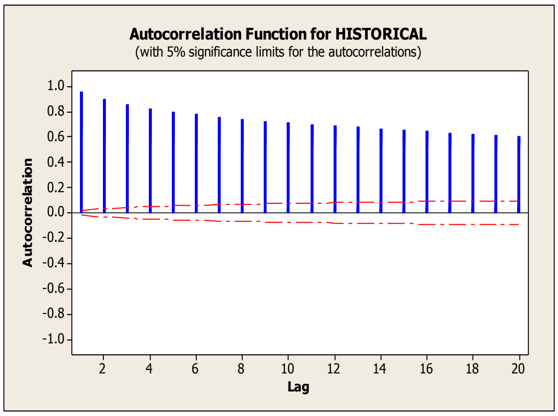


Figure 2a) ACF of Historical Data for station site 3527410

Figure 2b) ACF of Simulation Data station site 3527410



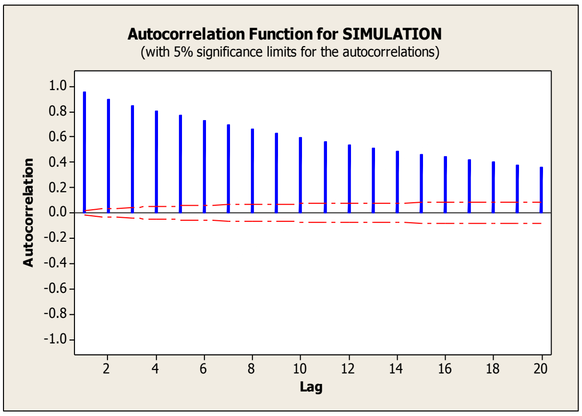
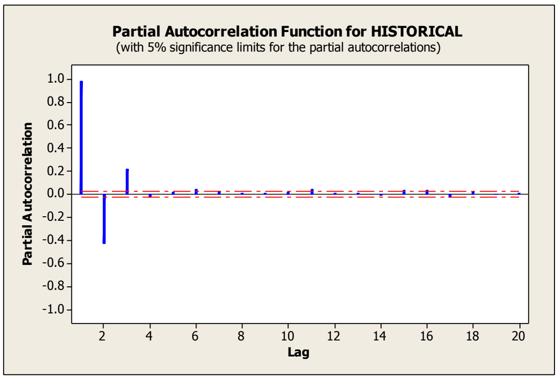


Figure 3a) ACF of Historical Data for station site 4023412

Figure 3b) ACF of Simulation Data for station site 4023412



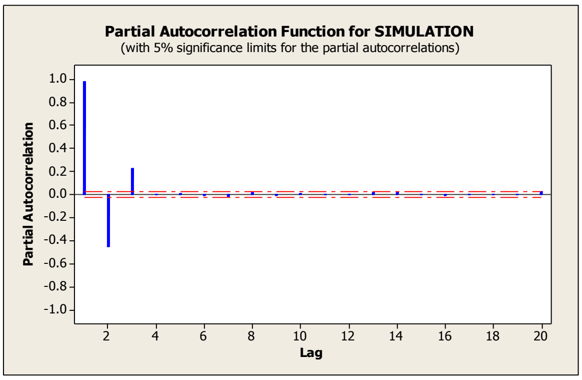
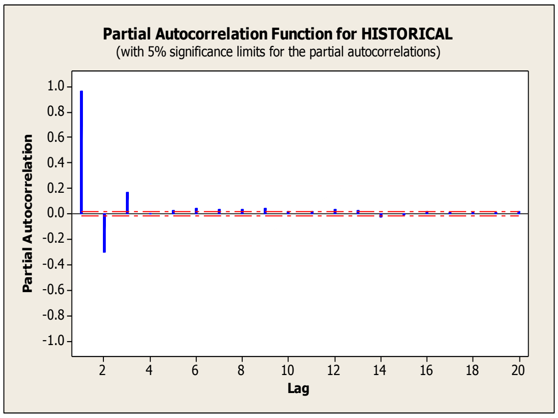


Figure 4a) PACF of Historical Data for station site 3527410

Figure 4b) PACF of Simulation Data station site 3527410



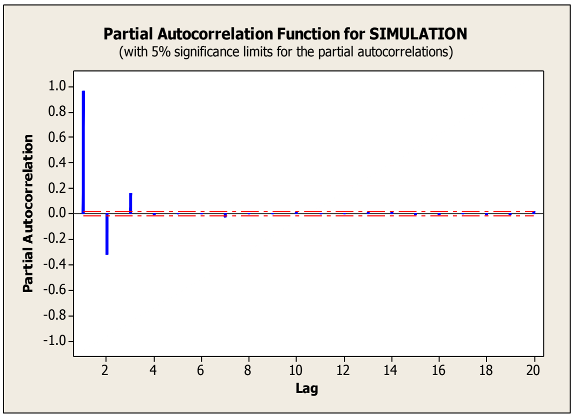
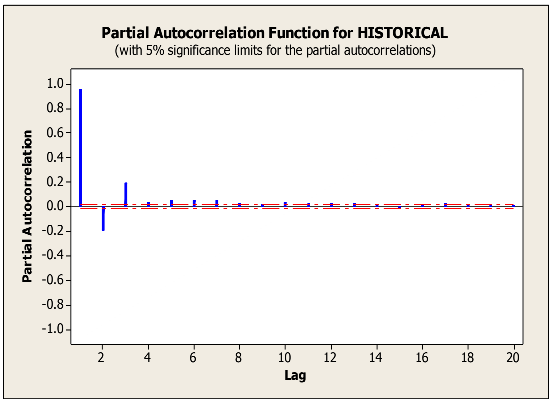


Figure 5a) PACF of Historical Data for station site 3527410

Figure 5b) PACF of Simulation Data station site 3527410



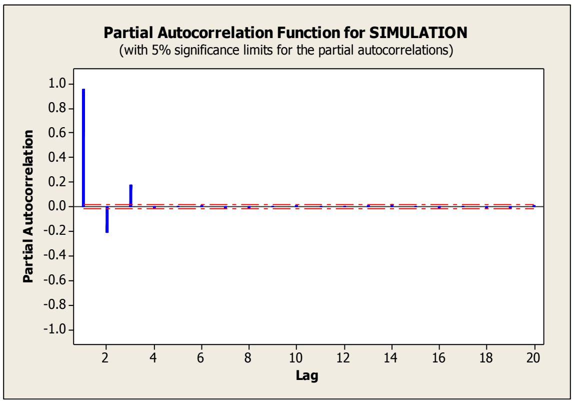


Figure 6a) PACF of Historical Data for station site 4023412

Figure 6b) PACF of Simulation Data for station site4023412

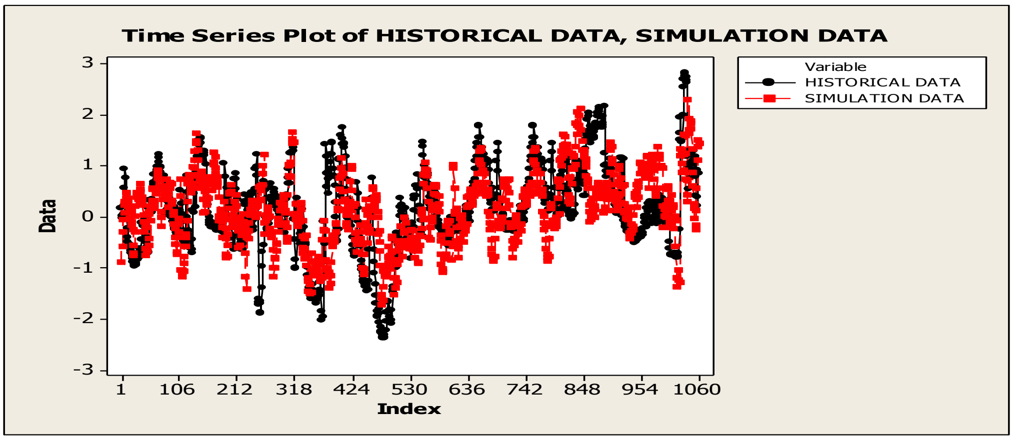


Figure 7. Time Series plot for station site 3424411

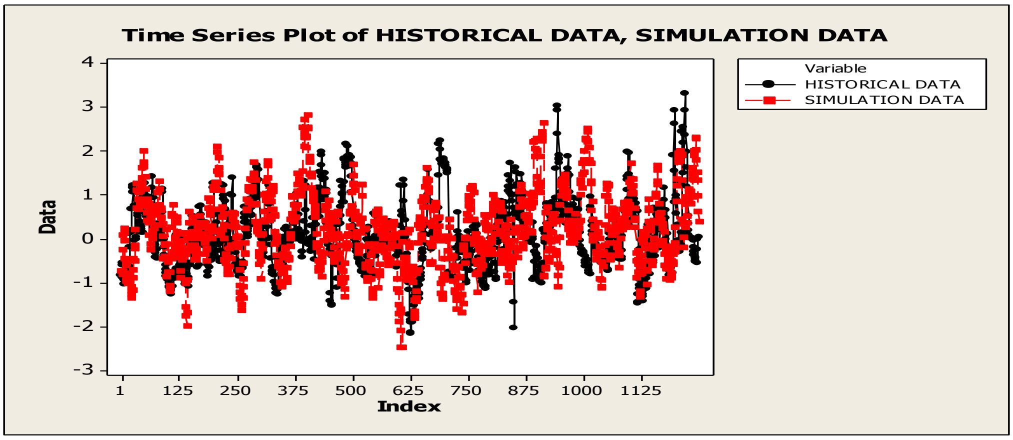


Figure 8. Time Series plot for station site 3527410

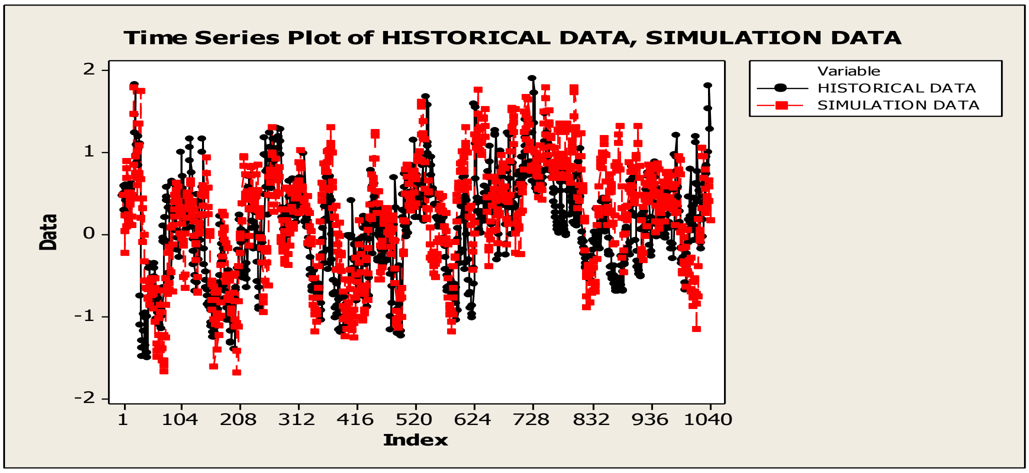


Figure 9. Time Series plot for station site 4023412

REFERENCES

Abo-Hammour, Z. E. S., Alsmadi, O. M., Al-Smadi, A. M., Zaqout, M. I., & Saraireh, M. S. (2012). ARMA model order and parameter estimation using genetic algorithms. *Mathematical and Computer Modelling of Dynamical Systems*, *18*(2), 201-221.

Aghdam, B. H., Vahdati, M., & Sadeghi, M. H. (2015). Vibration-based estimation of tool major flank wear in a turning process using ARMA models. *The International Journal of Advanced Manufacturing Technology*, *76*, 1631-1642.

Bao, C., Hao, H., & Li, Z. X. (2013). Integrated ARMA model method for damage detection of subsea pipeline system. *Engineering Structures*, *48*, 176-192.

Baptista, M., Sankararaman, S., de Medeiros, I. P., Nascimento Jr, C., Prendinger, H., & Henriques, E. M. (2018). Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. *Computers & Industrial Engineering*, *115*, 41-53.

Barnes, F. B. (1954), Storage required for a city water supply, *J. Inst. Eng. Australia 26*(9), 198-203

Beard, L. R. (1965), Use of interrelated records to simulate streamflow, *J. Hydrol. Div., ASCE 91*(HY5), 13-22

Beard, L. R. (1967), Simulation of daily streamflow, *Proceeding of the International Hydrology Symposium,* Ford Collins, Colorado,1, Paper 78

Berardengo, M., Rossi, G. B., & Crenna, F. (2021). Sea Spectral Estimation Using ARMA Models. *Sensors*, *21*(13), 4280.

Bertha, M., & Golinval, J. C. (2017). Identification of non-stationary dynamical systems using multivariate ARMA models. *Mechanical Systems and Signal Processing*, *88*, 166-179.

Boubacar Maïnassara, Y., & Saussereau, B. (2018). Diagnostic checking in multivariate ARMA models with dependent errors using normalized residual autocorrelations. *Journal of the American Statistical Association*, *113*(524), 1813-1827.

Box, G. E. P. and G. W. Jenkins.1970, Time Series Analysis Forecasting and Control, folden-Day, San Francisco,

Brockwell, P. J., Davis, R. A., Brockwell, P. J., & Davis, R. A. (2016). ARMA models. *Introduction to Time Series and Forecasting*, 73-96.

Cao, Y., Liu, Y., Zhang, D., Wang, W., & Chen, Z. (2013, June). Wind power ultra-short-term forecasting method combined with pattern-matching and ARMA-model. In *2013 IEEE Grenoble Conference* (pp. 1-4). IEEE.

Chakhchoukh, Y. (2010). A new robust estimation method for ARMA models. *IEEE Transactions on Signal Processing*, *58*(7), 3512-3522.

Fiering, M.B. (1967), Streamflow Synthesis, *Harvard University Press*, Cambridge, Mass

Hackstein, U., Krickl, S., & Bernhard, S. (2020). Estimation of ARMA-model parameters to describe pathological conditions in cardiovascular system models. *Informatics in Medicine Unlocked*, *18*, 100310.

Hazen, A. (1914), Storage to be provided in impounding reservoirs for municipal water systems, *Trans.Am. Soc. Civ. Eng. 77*, 1539

Hossain, M. B., Moon, J., & Chon, K. H. (2020). Estimation of ARMA model order via artificial neural network for modeling physiological systems. *IEEE Access*, *8*, 186813-186820.

Huang, R., Huang, T., Gadh, R., & Li, N. (2012, November). Solar generation prediction using the ARMA model in a laboratory-level micro-grid. In *2012 IEEE third international conference on smart grid communications (SmartGridComm)* (pp. 528-533). IEEE.

Karthikeyan, L., & Kumar, D. N. (2013). Predictability of nonstationary time series using wavelet and EMD based ARMA models. *Journal of hydrology*, *502*, 103-119.

Kumar, A. S., & Mazumdar, S. (2016, December). Forecasting HPC workload using ARMA models and SSA. In *2016 International Conference on Information Technology (ICIT)* (pp. 294-297). IEEE.

Laner, M., Svoboda, P., & Rupp, M. (2013). Parsimonious fitting of long-range dependent network traffic using ARMA models. *IEEE communications letters*, *17*(12), 2368-2371.

Lettenmaier, D. P. K. M. Latham, R. N. Palmer, J. R. Lund and S. J. Burges (1987), Strategies for coping with drought Part II: Planning techniques for planning and reliability assessment, *EPRI P-5201, Final Report Project*,2194-1

Li, X., Liu, Y., Fan, L., Shi, S., Zhang, T., & Qi, M. (2021). Research on the prediction of dangerous goods accidents during highway transportation based on the ARMA model. *Journal of loss prevention in the process industries*, *72*, 104583.

Maass, A., M. M. Hufschmidt, R. Dorfman, H. A. Thomas, Jr., S. A. Marglin and G. M. Fair (1962). *Design of Water Resource Systems*, Harvard University Press, Cambridge, Mass

Marelli, D., You, K., & Fu, M. (2013). Identification of ARMA models using intermittent and quantized output observations. *Automatica*, *49*(2), 360-369.

Matalas, N. C. (1967), Mathematical assessment of synthetic hydrology, *Water Resources Research*. *3*(4), 937-945

Moon, J., Hossain, M. B., & Chon, K. H. (2021). AR and ARMA model order selection for time-series modeling with ImageNet classification. *Signal Processing*, *183*, 108026.

Rout, M., Majhi, B., Majhi, R., & Panda, G. (2014). Forecasting of currency exchange rates using an adaptive ARMA model with differential evolution based training. *Journal of King Saud University-Computer and Information Sciences*, *26*(1), 7-18.

Sharma, A., Tarboton DG, and Lall U. (1997), Streamflow Simulation: A Nonparametric Approach, *Water Resources Research*, 33, 291-308

Shen, Y., Guo, J., Liu, X., Kong, Q., Guo, L., & Li, W. (2018). Long-term prediction of polar motion using a combined SSA and ARMA model. *Journal of Geodesy*, *92*, 333-343.

Singh, B., & Pozo, D. (2019, September). A guide to solar power forecasting using ARMA models. In *2019 IEEE PES Innovative Smart Grid Technologies Europe (IRIVERT-Europe)* (pp. 1-4). IEEE.

Tao P.C and Delluer, J.W. (1976), Seasonal and non seasonal ARMA Models in hydrology, *ASCE jour. Hydr. Div* , 102, 1541-1500

Thomas, H. A. and Fiering , M. B. , (1962), Mathematical synthesis of streamflow sequences for the analysis of rivers basin by simulation. In Design of water resources system, 459-493.

Triacca, U. (2016). Measuring the distance between sets of ARMA models. *Econometrics*, *4*(3), 32.

Vogel, R.M., and J.R. Stedinger (1988), The value of stochastic streamflow models in over-year reservoir design applications, *Water Resources Research. 24*(9), 1483-90

Wang, J., Zhou, Q., & Zhang, X. (2018, December). Wind power forecasting based on time series ARMA model. In *IOP conference series: Earth and environmental science* (Vol. 199, No. 2, p. 022015). IOP Publishing.

Xiao, H., Huang, D., Pan, Y., Liu, Y., & Song, K. (2017). Fault diagnosis and prognosis of wastewater processes with incomplete data by the auto-associative neural networks and ARMA model. *Chemometrics and Intelligent Laboratory Systems*, *161*, 96-107.

Xu, W., Gu, R., Liu, Y., & Dai, Y. (2015). Forecasting energy consumption using a new GM–ARMA model based on HP filter: The case of Guangdong Province of China. *Economic Modelling*, *45*, 127-135.

Xu, X. H., Ye, L., Pei, Y., Zhao, L., & Wang, J. J. (2022). Research on the Comprehensive Evaluation of the Higher Education System Based on FCE and ARMA Models. *Complexity*, *2022*.

Yevjevich, V. (1972), Probability and Statistics n Hydrology, Water Resources Publications, Fort Collins, Colorado

You, J., Yu, C., Sun, J., & Chen, J. (2022). Generalized maximum entropy based identification of graphical ARMA models. *Automatica*, *141*, 110319.

Zheng, T., Xiao, H., & Chen, R. (2015). Generalized ARMA models with martingale difference errors. *Journal of Econometrics*, *189*(2), 492-506.

Zhu, K., & Li, W. K. (2015). A bootstrapped spectral test for adequacy in weak ARMA models. *Journal of Econometrics*, *187*(1), 113-130.

Zhu, L., Wang, Y., & Fan, Q. (2014). MODWT-ARMA model for time series prediction. *Applied Mathematical Modelling*, *38*(5-6), 1859-1865.