Time Series Analysis And Modeling Of Monthly Rainfall In Saudi Arabia

Abstract:

Managing of water resources is an important future issue. Modeling is fundamental in planning and management of water resource systems. Forecasting of occasions requires identifying proper models to be used in this process. Water is the main source of life on earth. One of the most common and fundamental sources of water on earth supporting the existence of the majority of living organisms is Rainfall. Time series analysis which includes modeling and forecasting constitutes a tool of paramount importance with reference to a wide range of scientific purposes in meteorology (e.g. precipitation, humidity, temperature, solar radiation, floods and draughts). The present work applies the Box-Jenkins approach, employing SARIMA (Seasonal Autoregressive Integrated Moving Average) model is used to perform short term forecasts of monthly time series such as rainfall. Modeling the past observed rainfall time series values which in turn is used to predict the future quantities in accordance to the past. The model is tested by verifying the past rainfall data. In turn, the model generates a reliable future forecast. This model is evaluated by means of both the AIC- and BIC- (SBC-) model.

Keywords: rainfall time Series; Forecasting; Auto Regressive Moving Average Models; Trend; Seasonality; SARIMA Models

1-Introduction and Motivation

Rainfall can be considered as a non-linear natural process raising complexity especially while trying to forecast future values. Because of climate spatial variations, the rainfall is considered as a randomized stochastic process. Rainfall is considered as one of the most important issues which include harvesting, water supply limit plan, flood prediction, and many others. A variety of statistical procedures are often employed to forecast rainfall amounts. ARIMA modeling technique has been applied on a great number of not only financial but also hydrological time series data in order to forecast rainfall data.

2. Methodology Description

2.1. Regional Area of interest

The time series observations for monthly rainfall amounts in, Saudi Arabia were used in this work. The data which is used for this research was for every two weeks / month (more than 500 observations) starting from 1964. Figure 1 and 2 shows a time series plot of the monthly and annual rainfalls at that station respectively.

2.2. Box-Jenkins model building procedure

This model is introduced in 1970 by Box and Jenkins [1]. It is primarily developed for financial time series analysis. The used model deals with stationary time series, and fitting either

autoregressive moving average (ARMA) or autoregressive integrated moving average (ARIMA) or seasonal autoregressive integrated moving average (SARIMA) models with the view to discover the most appropriate match of a time series observations. Also, the most accurate prediction for future forecast.

2.2.1. Identification and selection of the Model

In this phase, the first step, verification the stationary of the variables. Second, locating seasonality if exists, within the time series. This is done by analyzing the plot of autocorrelation and partial autocorrelation functions.

2.2.2. Model parameters estimation

The model is designed with different parameters, seasonal and non seasonal time series. The parameters are selected using iterative algorithm which try all the values and the select the ones with the minimum mean square error between the estimated time series using this model and the original observed values.

2.2.3. Testing and forecasting

This is done by testing whether the estimated model satisfy the stationary univariate process, or not. Specifically, the residuals should not be dependent between each other and must exhibit constancy in terms of mean and variance within the entire length of the time series. ARIMA models have three main types. These types are Autoregressive (AR) and Moving Average (MA) models, Autoregressive Moving Average (ARMA) models, and Seasonal (ARIMA) models.

3. Data Analysis:

The data is being processed via four different phases, namely, data preprocessing, model design, implementation, testing.

3.1 Data Preprocessing:

In order to do a perfect data analysis which includes forecasting with minimum error, some preprocessing operations should be done. First, covering the missed data using linear interpolation or any other techniques is done. In many time series analysis, the log of the data is processed which is easy to be inverted. Five locations in Kingdom of Saudi Arabia are analyzed. In the first location Bahrah, 46 years (1966-2011), the first analysis is done for the total rainfall per each month. Another variable is studied which is the storm day within each year. The calendar is arranged as a day number (the month can be estimated) ranges from (1-365) or (1-366) according to February number of days. Figure 1 illustrates the first time series total rainfall per month within the whole interval (552 values). Figure 2 describes the storm day occurrence for each year within the 46 years. Figure 3 shows the maximum daily rainfall in each year within the interval.

Figure 1: The total rainfall mm/month within the 552 months in Bahrah Location

Figure 2: Storm day number within the years in Bahrah

Figure 3: Maximum daily rainfall mm/day in each year

3.2 Model Design:

Much iteration are done in this phase in order to select the best fit parameters of the model (ARIMA)

ARIMA(0,1,1) Model Seasonally - Integrated with Seasonal MA(12):

Distribution: Name = 'Gaussian'

P1: 13

D1: 1

Q1: 13

Constant: 0

AR1: {}

SAR1: {}

MA1: {NaN} at Lags [1]

SMA1: {NaN} at Lags [12]

Season: 12

Var. : NaN

Second, check the residuals for normality. One assumption of the fitted model is that the innovations follow a Gaussian distribution. Infer the residuals, and check them for normality.

Figure 4: The Quantile, and Quantile plot (QQ-plot) and kernel density

The Quantile, and Quantile plot (QQ-plot) and kernel density estimate show the normality assumption. The next step, the residuals are checked for autocorrelation. The objective is to confirm that the residuals are uncorrelated. Look at the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) plots for the standardized residuals described in figure 5.

Figure 5: Autocorrelation function (ACF) and partial autocorrelation function (PACF) plots for the standardized residuals

The ACF and PACF plots are shown in figure 5. It can be concluded that there is no significant autocorrelation. More formally, conduct a Ljung-Box Q-test at lags 5, 10, and 15, with degrees of freedom 3, 8, and 13, respectively. The degrees of freedom account for the two estimated

moving average coefficients. The Ljung-Box Q-test confirms the sample ACF and PACF results. The null hypothesis that all autocorrelations are jointly equal to zero up to the tested lag which is not rejected (h = 0) for any of the three lags.

3.3 Model Testing:

This phase is concerned with checking the performance of predictive. Use a holdout sample to compute the predictive MSE of the model. In order to achieve this goal, a comparison between the estimated time series and the observed one is shown in figure 6.

ARIMA(0,1,1) Model Seasonally Integrated with Seasonal MA1(12):

Conditional Probability Distribution: Gaussian

Standard t

Parameter Value Error Statistic

Constant 0 Fixed Fixed

MA1{1} -1 0.00879247 -113.734

SMA1{12} -0.830567 0.0213176 -38.9616

Variance 0.146902 0.00496624 29.5802

pmse =

0.13287

Figure 6: Predicted and observed time series-monthly rainfall

The residuals are checked for normality. One assumption of the fitted model is that the innovations follow a normal distribution. Infer the residuals, and check them for normality as shown in figure 7.

Figure 7: Check the normakity of the residuals.

The Quantile, and Quantile plot (QQ-plot) and kernel density estimate confirm the normality assumption. Also, from checking the residuals for autocorrelation, it is clear that the residuals are uncorrelated. Look at the sample autocorrelation function (ACF). Then, partial autocorrelation function (PACF) plots for the standardized residuals. Figure 8 shows a Sample Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF).

Figure 8: Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots for the standard residuals.

These steps are repeated for the maximum rainfall per month in the whole time interval. The

final result is shown in figure 9, observed time series compared with the predicted. Finally the storm day prediction is shown in figure 10.

Figure 9: Maximum Daily Rainfall prediction in the year

Figure 10: Storm day Prediction

4. Conclusions

This work introduces a generic model for time rainfall analysis. The analysis of the research results, emphasize that AIRMA modeling is a promising tool for modeling and simulation of rainfall in arid and semi-arid regions. The proposed models are able to preserve the seasonal statistics of the observed data. The presented work emphasizes the importance of data preprocessing. Also, it is recommended to spend some effort for selecting the best parameters. Testing the models by different methods is shown in this work. Good preprocessing plus selecting the best parameters model result in accurate forecasting.

References:

1. George E. P. Box, Gwilym M. Jenkins, Time series analysis: forecasting and control,Book 1970.
2. B. E. Barry, E. Obuobie, M. Andreini, W. Andah, and M. Pluquet, Comprehensive Assessment of Water Management in Agriculture, 2005, (Comparative Study of River Basin Development and Management.
3. C. Gordon and J. A. Amatekpor, Eds., The Sustainable Integrated Development of the Volta Basin in Ghana, Volta Basin Research Project, Accra, 1999, 159p.
4. C. Leemhuis, G. Jung, R. Kasei, and J. Liebe, The Volta Basin Water Allocation System: Assessing the impact of small-scale reservoir development on the water resources of the Volta basin, West Africa, Advances in Geosciences, 21, 57–62, (2009).
5. M. Long, D. Entekhabi, and S. E. Nicholson, Interannual variability in rainfall, water vapor flux, and vertical motion over West Africa, Journal of Climate, 13, no. 21, 3827–3841, (2000).
6. S. E. Nicholson and I. M. Palao, A re-evaluation of rainfall variability in the Sahel. Part I. Characteristics of rainfall fluctuations, International Journal of Climatology, 13, no. 4, 371–389, (1993).
7. S. E. Nicholson, An overview of African rainfall fluctuations of the last decade, Journal of Climate, 6, no. 7, 1463–1466, (1993).
8. S. E. Nicholson, Historical and modern fluctuations of Lakes Tanganyika and Rukwa and their relationship to rainfall variability, Climatic Change, 41, no. 1, 53–71, (1999).
9. S. E. Nicholson, Climatic and environmental change in Africa during the last two centuries, Climate Research, 17, no. 2, 123–144, (2001).
10. S. E. Nicholson, A. I. Barcilon, M. Challa, and J. Baum, Wave activity on the tropical easterly jet, Journal of the Atmospheric Sciences, 64, no. 7, 2756–2763, (2007).
11. S. E. Nicholson and J. P. Grist, The seasonal evolution of the atmospheric circulation over West Africa and Equatorial Africa, Journal of Climate, 16, no. 7, 1013–1030, (2003).
12. S. Opoku-Duah, D. N. M. Donoghue, and T. P. Burt, Inter comparison of evapotranspiration over the Savannah Volta Basin in West Africa using remote sensing

data, Sensors, 8, no. 4, 2736–2761, (2008).
13. N. Philippon and B. Fontaine, The relationship between the Sahelian and previous 2nd Guinean rainy seasons: a monsoon regulation by soil wetness? Annales Geophysicae, 20, 575–582, (2012).
14. M. Sabry, H. Abd-El-Latif, and N. Badra, Comparison between Regression and ARIMA models in forecasting traffic volume, Australian Journal of Basic and Applied Sciences, 1, no. 2, 126–136, (2007).
15. N. Van de Giesen, J. Liebe, and G. Jung, Adaptation to climate change in the Volta Basin, West Africa, Current Science, 98, no. 8, (2010).
16. Duong Tran Anh , Thanh Duc Dang and Song Pham Van, Improved Rainfall Prediction Using Combined Pre-Processing Methods and Feed-Forward Neural Networks, J 2019, 2, 65–83; doi:10.3390/j2010006 www.mdpi.com/journal/j.
17. M Sidiq, Forecasting Rainfall with Time Series Model, INCITEST, IOP Publishing,IOP Conf. Series: Materials Science and Engineering 407 (2018) 012154 doi:10.1088/1757-899X/407/1/012154.