© 2015 IJSRSET | Volume 1 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN: 2394-4099
Themed Section: Engineering and Technology

# River Flood Prediction using Time Series Model

**Thinn Htet Htet San, Mie Mie Khin**
University of Computer Studies, Mandalay, Myanmar

## ABSTRACT

With the passage of time the impacts of natural hazards continue to increase around the world. The globalization and growth of human societies and their escalating complexity and river flooding will further increase the risks of natural hazards. Flood prediction and control are one of the greatest challenges facing the world today, which have become more frequent and severe due to the effects of global climate change and human alterations of the natural environment. Therefore, it is important to protect people and their property from flooding, helping organizations. By giving warnings of possible floods so that people can make arrangements or move out of the area if is dangerous. Time series model that can predict river flooding using water level data over Ayeyarwady river in Myanmar is proposed. This model focus on the prediction of events and can capture the fact that time flows forward and is applied to model the spatial dependencies not only between the hydrological and hydro-meteorological variables but also between weather stations in Myanmar for river flood prediction. The system predicts water level by daily to daily forecast type for additive models in Time Series using Two-Period, Three-Period and Four-Period moving average. The system compares with the predicted results and actual weather station results. Finally, the best model will be shown for river flood prediction over Ayeyarwady River in Myanmar.

**Keywords:** Flood, Prediction, Time Series, Additive Model

## I. INTRODUCTION

Flood is one of the most damaging natural disasters in this planet that affect many countries in the world year after year. Natural disasters such as flood and tropical cyclones are regarded to be caused by extreme weather conditions as well as changes in global and regional climate. The impact of flooding ranges from the destruction of property, loss of agricultural production and disruption of transports and services to loss of lives. Flood is a natural disaster that results from severe combination of critical meteorological and hydrological conditions which may be grieved by man made causes.
Floods are the most frequent natural hazards globally and the hazard of flooding can be divided into primary, secondary and tertiary effects. The primary effects of floods are those due to direct contact with the flood waters, with the water velocities resulting in floods as the discharge velocity increases. Secondary effects, such as disruption of infrastructure and services and health impacts, are primary effects, while tertiary effects are viewed as the long-term changes that occur, for example changes in the position of river channels.

In this work, the way to predict water level from historical data is to build the prediction models with Time series. These models are constructed by real time of data flow, which is from the recording equipment in testing of complex systems. Weather stations are selected by along with Ayeyarwaddy river which often face local flooding in rainy seasons as our region of interest. Flood level can be modelled by using various time scales such as yearly, monthly, weekly and daily.
The factors that affect flooding in Ayeyarwaddy river will be studied and observed the data for flood prediction in this research. The collections of historical data are used to construct model based on the spatial dependencies between variables which are observable factors contributing to the flood level of Ayeyarwaddy river. The challenge here is to build models for flood prediction of Ayeyarwaddy river in Myanmar. Then the

system evaluates constructed models with river flood prediction.

The organization of this paper is as follows. A briefly survey of many researchers is presented in section 2. The study area and available data is dealt with in section 3. This is followed by the Time series model and its analysis in section 4 and comparison of the proposed models is discussed in section 5. Finally, the system concludes in section 6.

## II. RELATED WORKS

C. Damle [2] has applied Time Series Data Mining to the area of flood forecasting. Three examples of gauging stations, representing high, medium and low flood occurrences were considered. The prediction accuracy is evaluated in terms of $a$, ß, Positive Prediction Accuracy (PPA) and Correct Prediction Accuracy (CPP).

R. Samsudin et al. [8] combined the group method of data handling (GMDH) and the least squares support vector machine (LSSVM) as Novel hybrid forecasting model known as GLSSVM. The GMDH was used to determine the useful input variables which work as the time series forecasting for the LSSVM mode.

M. Kannan et al. [4] predicted the rainfall by using multiple linear regression (MLR) models and computed Pearson coefficient for five years data and then compared with predicted data using regression approach. The predicted values lie below computed values. According to the results, it does not show accuracy but show an approximate value. Predicting flood, cyclone, forest fire detection, and global warming was avoided.

M. A. Kulkarni et al. [5] proposed wind speed prediction using four different statistical techniques; Curve fitting, Auto Regressive Integrated Moving Average Model (ARIMA), Extrapolation using periodic curve fitting and Artificial Neural Networks (ANN). They computed the Root Mean Square Error (RMSE) in prediction of zonal component of wind speed for all the months.

In [12] Song et al. presented key findings and methods for modeling and forecasting from 2000-2008, including time-series, econometric, and combination techniques. This survey identified some new research directions, which included improving the forecasting accuracy through integrating both qualitative and quantitative forecasting approaches, tourism cycles, seasonality analysis, impact assessment, and risk forecasting.

N. Sen presented a long-range summer monsoon rainfall forecast model based on power regression technique with the use of Ei Nino, Eurasian snow cover, north west Europe temperature, Europe pressure gradient, 50 hPa Wind pattern, Arabian sea SST, east Asia pressure and south Indian ocean temperature in previous year. The experimental results showed that the model error was 4% in [9]

S. Nkrintra, et al. [7] described the development of a statistical forecasting method for SMR over Thailand using multiple linear regression and local polynomial-based nonparametric approaches. SST, sea level pressure (SLP), wind speed, EiNino Southern Oscillation Index (ENSO), and IOD were chosen as predictors. The experiments indicated that the correlation between observed and forecast rainfall was 0.6.

T. Sohn, et al. [11] has developed a prediction model for the occurrence of heavy rain in South Korea using multiple linear and logistics regression, decision tree and artificial neural network. They used 45 synoptic factors generated by the numerical model as potential predictors. In [13] W. T. Zaw developed a prediction model for determining Rainfall over Myanmar using multiple linear regressions where 15 predictors has been used. As a result of several experiments, the predicted rainfall amount is close to actual value.

S. Banik et al. (2009) [1] have developed rainfall forecasting model using ANN, ANFIS and GA processes and the results obtained by these models are also compared to the statistical. The ANFIS forecasting model and the GA forecasting model can be used to forecast monthly monsoon rainfall more accurately than the ANN model and the statistical model, forecasting method, namely linear multiple regression model.

D. Nayak et al. (2013) have done survey on rainfall predictions using different neural network architectures over twenty-five years. From the survey it has been found that most of the researchers used back propagation network for rainfall prediction and got significant results. The survey also gives a conclusion that the forecasting techniques that use MLP, BPN, RBFN, SOM and SVM are suitable to predict rainfall than other forecasting techniques such as statistical and numerical methods in [6]. El-Shafie et al. (2011) [10] have developed an adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN) model to forecast the rainfall for Klang River in Malaysia on monthly basis. The result showed performance of ANFIS method is better than ANN method and concluded that ANFIS method is superior to the ANN method in forecasting monthly rainfall.

## III. STUDY AREA AND AVAILABLE DATA

Department of Meteorology and Hydrology (DMH) is the responsible agency for flood forecasting and warning in the country. There are eight major rivers in Myanmar, which are Ayeyarwady, Chindwin, Sittoung, Thanlwin, Bago, Dokehtawady, Shwegyin and Ngawun rivers.

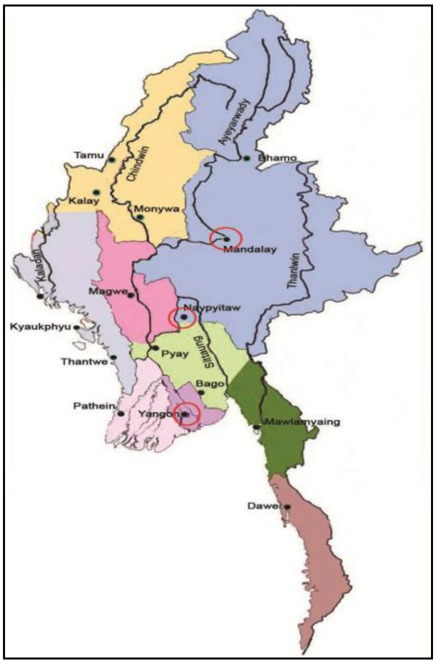


**Figure 1:** Map of Myanmar with Ayeyarwaddy River

The Ayeyarwaddy River starts in the northern part of the country, Kachin State, and flows south to the Andaman Sea on the Ayeyarwaddy Coast as shown in Figure 1. Flood usually occurs in each and every year at one river system or another. The occurrences of floods in Myanmar can be generally expressed as 6% in June, 23% in July, 49% in August, 14% in September and 8% in October. According to the previous 47 years' observation, severe flood years were noted as 1973, 1974, 1976, 1979, 1988, 1991, 1997, 2002, 2004 and 2007. According to historical statistics, heavy monsoon rain in Myanmar especially from mid-July to mid-September have caused flooding along the Ayeyarwaddy, Chindwin, Thanlwin, Sittaung and Yangon rivers and their tributaries [3]. Hence, DMH stations along with Ayeyarwaddy, which is main river of Myanmar and often face local flooding in rainy season, are selected as region of interest. TABLE I shows the historical data are collected from each Met \ Hydro station. These data are observable factors contributing to the river flood level.
Time series model is developed to predict daily flood level prediction for each weather station. The data is collected between 1990 and 2005 of each weather station to train and test the models.

**TABLE I**
DETAIL OF MET/HYDRO STATION

| No | Station Name | Region | Latitude & Longitude |
|---|---|---|---|
| 1 | Myitkyinar | Upper | 25.3'N, 97.4'E |
| 2 | Sagaing | Central | 21.5'N, 95.6'E |
| 3 | Mandalay | Central | 21.9'N, 96.1'E |
| 4 | Magway | Central | 18.83'N, 93.78'E |
| 5 | Aunglan | Lower | 19.38'N, 95.23'E |
| 6 | Pyay | Lower | 18.8'N, 95.21'E |

## IV. TIME SERIES MODEL AND ANALYSIS

A time series is a sequence of data indexed by time, often comprising uniformly spaced observations. It is formed by collecting data over a long range of time at a regular time interval (data points should be at the same interval on the time axis) such that a time series is a collection of data recorded over a period of time—weekly, monthly, quarterly, or yearly. Consecutive points are then linked by means of straight lines to form the series. The model accounts for patterns in the past movements of a variable and uses that information to predict its future movements. For example, an economic variable may be influenced by external factors which we cannot explain like the weather, changes in taste or seasonal cycles in spending. There are many Time series models. Among these this paper will be used Additive model because historical data is different from the range. Time series models are constructed based on weather data (water level). In these models, the river flow, rainfall, tide, snow and *etc* are considered as the system input. Water flow dynamics and water surface level are considered as the system output. Forecasting is the ultimate objective of time series analysis. The system forecasts using trend and seasonality.

To construct the proposed additive model, the historical data at daily (period $t$) water level are collected. Depended on the collected historical daily data, the period trends ($T_t$) are calculated using Moving Average (MA) method. For getting the seasonal factors ($S_t$), the trends are subtracted from the collected data. And then, the environmental effects (irregular $I_t$) or cyclical ($C_t$) component can not be consider in the proposed model. Finally, the forecast ($Y_t$) are obtained based on seasonal factors.

MA is calculated using weather data. The getting MA (trend) is shown by applying two-period, three-period and four-period methods. Two-period is based on 16 historical data records for two days. Three-period is based on 15 historical data records for three days. Four-period is based on 16 historical data records for four days. There are two forecast types: Daily to Daily forecast and Daily to Forecast.

**(1) Daily to Daily Forecast**

Daily to daily forecast is to forecast the historical data for one day i.e. it forecasts the next day using 16 historical data records. When we want to know the forecast for the next day again, the closest 16 records will be applied.

**(2) Daily to Forecast**

Daily to forecast is to forecast for the next three days. The Weather Department usually forecast for two days or three days. But daily to forecast method can give the forecasting results close to the actual results for one day. In remaining days, the difference results with the actual results can be found.

In this paper, we use daily to daily forecast type for prediction.

#### A. Model Evaluation, Comparison and Error Results

In this section, the system evaluates water level of historical weather data from three region such as upper, middle, and lower regions in Myanmar using additive model in time series. Myitkyinar, Sagaing, Mandalay, Magway, Aunglan and Pyay stations in Myanmar are considered to evaluated forecast by one day, two days and three days forecasts.

Firstly, Myitkyinar station at upper region in Myanmar is chosen. This station uses Daily to Daily forecast type and it is forecasted by two periods, three periods and four periods for three days. Then, error rates of three period forecasts are calculated by using Mean Absolute Percentage Error (MAPE). The comparison results of the actual and forecast is as shown in following tables. TABLE II shows the comparison results and error rate for additive model using Two-Period moving average at Myitkyinar station in Myanmar.

TABLE II
COMPARISON RESULTS FOR ACTUAL AND TWO PERIODS FORECASTS AND ERRORS IN MYITKYINAR

| Date | Actual (cm) | Three_Periods (cm) | Two _MAPE |
|---|---|---|---|
| 08/30/1995 | 533 | 536 | 0.56 % |
| 08/31/1995 | 523 | 514 | 1.72 % |
| 09/01/1995 | 503 | 506 | 0.6 % |

The system also forecasts the water level at Myitkyinar station for additive model using Three-Period moving average and compares the actual and forecast results and then calculates the error rates between them as shown in TABLE III.

TABLE III
COMPARISON RESULTS FOR ACTUAL AND THREE PERIODS FORECASTS AND ERRORS MYITKYINAR

| Date | Actual (cm) | Three_Periods (cm) | Three _MAPE |
|---|---|---|---|
| 08/30/1995 | 533 | 538 | 0.94 % |
| 08/31/1995 | 523 | 525 | 0.38 % |
| 09/01/1995 | 503 | 509 | 1.19 % |

TABLE IV presents comparison results and error rates of water level at Myitkyinar station for additive model using Four-Period moving average.

TABLE IV
COMPARISON RESULTS FOR ACTUAL AND FOUR PERIODS FORECASTS AND ERRORS IN MYITKYINAR

| Date | Actual (cm) | Four _Periods (cm) | Four _MAPE |
|---|---|---|---|
| 08/30/1995 | 533 | 529 | 0.75 % |
| 08/31/1995 | 523 | 528 | 0.96 % |
| 09/01/1995 | 503 | 513 | 1.99 % |

## V. COMPARISON OF PROPOSED MODELS

In this section, the additive model in time series is using Two-Period, Three-Period and Four-Period moving averages are compared as shown in Figure 2.

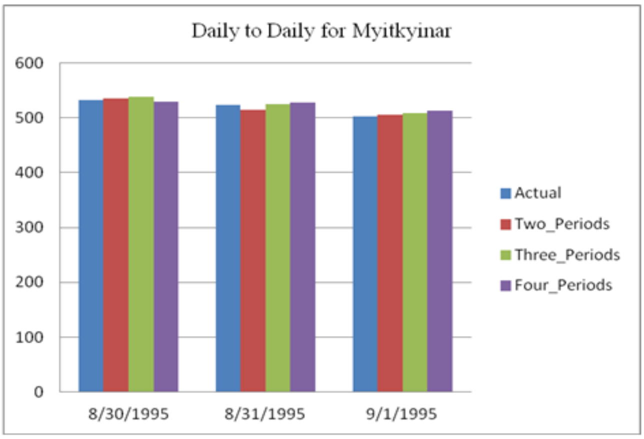


**Figure 2:** Comparison of Daily to Daily for Myitkyinar Station

Figure 2 shows the comparison results with actual for Daily to Daily using two periods, three periods and four periods for three days in 1995. This figures revealed that additive models using Two-Period and Three-Period moving averages for daily to daily forecast method get the better result than the Four-Period moving average.

## VI. CONCLUSION

In this paper, daily flood level prediction for each weather station was implemented by applying additive model in Time series. This model was useful to be acceptable accuracy for flood prediction of Ayeyarwaddy River in Myanmar. This system compared the result of models with actual weather results and also chose the best model for river flood prediction in Myanmar. We also intend to extend and predict such time series data using daily to forecast method.

## VII. REFERENCES


[1] S. Banik, Mohammed Anwer, A.F.M. Khodadad Khan, Rifat Ara Rouf and Farah Habib Chanchary, “Forecasting Bangladeshi monsoon rainfall using neural network and genetic algorithm approaches”, International Technology Management Review, Volume 2 Number 1, 2009.

[2] C. Damle, “Flood forecasting using time series data mining” , University of South Florida (2005).

[3] A. N. Hlaing , “Uncertain Reasoning & Prediction for River Flooding with Bayesian Networks Model”, AICT 2010.

[4] M.Kannan, S.Prabhakaran and P.Ramachandran, “Rainfall Forecasting Using Data Mining Technique”, International Journal of Engineering and Technology, Vol.2 (6), 397-401, 2010.

[5] M. A. Kulkarni, S. Patil, G. V. Rama and P. N. Sen, “Wind speed prediction using statistical regression and neural network”, J. Earth Syst. Sci. 117, No. 4, August 2008, pp. 457-463.

[6] D. Nayak, A. Mahapatra, P. Mishra, “ A Survey on Rainfall Prediction using Artificial Neural Network”, International Journal of Computer Applications “, Volume 72, No.16, 2013.

[7] S. Nkrintra, et al., “Seasonal Forecasting of Thailand Summer Monsoon Rainfall”, in International Journal of Climatology, Vol. 25, Issue 5, American Meteorological Society, 2005, pp. 649-664.

[8] R. Samsudin, P. Saad, and A. Shabri, “ River flow time series using least squares support vector machines ”, University Teknologi Malaysia, (2010).

[9] N. Sen, ” New forecast models for Indian south-west Monsoon season Rainfall”, in Current Science, vol. 84, No. 10, May 2003, pp.1290- 1291.

[10] El-Shafie, O. Jaafer and Seyed Ahmad Akrami, “Adaptive neuro-fuzzy inference system based model for rainfall forecasting in Klang River, Malaysia”, International Journal of the Physical Sciences Vol. 6(12), pp. 2875-2888.

[11] T. Sohn, J. H. Lee, S. H. Lee, C. S. Ryu, “Statistical Prediction of Heavy Rain in South Korea”, in Advances in Atmospheric Sciences, Vol. 22, No. 5, 2005, pp.703-710.

[12] Song, Haiyan, and G. Li, “Tourism demand modelling and forecasting - A review of recent research”, Tourism Management 29.2 (2008): 203-220.

[13] W. T. Zaw, T. T. Naing, “Empirical Statistical Modeling of Rainfall Prediction over Myanmar”, World Academy of Science, Engineering and Technology, Vol: 2 2008-10-27.