21st International Congress on Modelling and Simulation, Gold Coast, Australia, 29 Nov to 4 Dec 2015
www.mssanz.org.au/modsim2015

Statistical correlations between rainfall and climate indices in Western Australia

Iqbal Hossain, H. M. Rasel, Monzur Alam Imteaz and Muhammad Moniruzzaman

Department of Civil and Construction Engineering, Faculty of Science, Engineering and Technology, Swinburne University of Technology, Melbourne, VIC 3122
Email: ihossain@swin.edu.au

Abstract: As the concept of 'sustainable development' is increasing day by day, rainfall has become the most significant and investigated hydro-climate variable in Australia. Rainfall variability is considered as a major economic factor in Australia. It was observed that Australian rainfall is affected by several climate patterns and long-term prediction remains a challenge for many years. However, forecasting rainfall can be beneficial for the design, maintenance and management of water resources infrastructures. Nevertheless, rainfall is the product of complex global atmospheric phenomena. Strong correlations between rainfall and several climate indices have already been observed throughout the world. Any such correlation with climate indices and rainfall afterwards can be used in forecasting long-term rainfall. For the prediction of rainfall in advance, statistical and dynamic systems can be used in practice. However, dynamic systems are too complex and expensive to use in a wide range of situations. This paper focused on the investigation of statistical correlations between rainfall and several climate indices as potential predictor of long-term Western Australian rainfall. Since Australian rainfall is highly variable both in time and space, this analysis was performed on regional scale. Several multiple regression models were investigated using the climate indices as potential predictors of rainfall. The models which satisfied the limits of statistical significance were used to forecast Western Australia rainfall in advance. Historical rainfall data were obtained from Australian Bureau of Meteorology. The rainfall station Roebourne in Western Australia was chosen as a case study. The station was selected based on their long term recorded data having fewer missing values. The major aim was deterministic forecasting of long-term rainfall in terms of climate indices in regional scale. The analysis showed that DMI-ENSO based combined multiple regression models could be used for long-term rainfall forecasting of Western Australia except extreme rainfall.

Keywords: *Climate indices, ENSO, IOD, multiple regression, rainfall forecasting*

1. INTRODUCTION

For the design, planning and development of water resources management strategies, rainfall forecasting is essentially important. It will help to check the water balance between future supply and availability, which will help to ensure in meeting demand. Therefore, rainfall forecasting has become one of the primary goals to water resources managers. Reliable rainfall forecasting could be beneficial in the management of land, watershed and water systems (Anwar et al., 2008) to some extent particularly in Australia, where hydro climatic variability is high (Peel et al., 2001). Nevertheless, the complex global atmospheric phenomena rainfall varies both temporally and spatially. Therefore, researchers have used different modelling techniques to establish relationships between rainfall and climate modes in different parts of the world (Mekanik and Imteaz, 2013).

It is believed that the occurrence of rainfall around the world is influencing by several large scale climate modes. Amongst the climatic variables El Nino Southern Oscillation (ENSO) and Indicant Ocean Dipole (IOD) are well known for their effect on India, North and South America and Australia. A number of studies have been conducted to explore the influence of dominant climate drivers, including ENSO, IOD and Southern Annular Mode (SAM) on Australian rainfall. A part of these researches covered whole Australia Cai et al., (2011); Kirono et al., (2010); Risbey et al., (2009); while others focused on more concentrated on specific region of Australia. For example, Mekanik et al., (2013) focused on South-East Australia using lagged climate indices; Ummenhofer et al., (2008) emphasized on South Western Australia; Nicholls (2010) and Evans et al., (2009) emphasized on South Australia, Verdon et al., (2004) emphasized on East Australia. Most of the studies found that the relationship between climate predictors and rainfall are much complex and single predictors alone are unable to forecast rainfall accurately.

For the prediction of water resources and hydrological variables, multiple regression (MLR) models are commonly used. Many researchers have used MLR models for rainfall forecasting (He et al., 2014) and flood forecasting (Latt et al., 2014). Ihara et al. (2007) used MLR models to investigate the relationship between ENSO and Indian Ocean indices with summer monsoon rainfall. Mekanik et al., (2013) examined the influence of lagged ENSO and IOD on Victorian rainfall using MLR models. However, the predictive capability of currently used models beyond 1 week and shorter than a season is still questionable (Hudson et al., 2011). According to Vitrat (2004), the usual forecast systems generally lost information from the atmospheric initial conditions that are basis for weather forecasts. In the first month of the forecasting period, the ocean state might not change much since the start of the prediction.

This paper presents the statistical correlations between Western Australian rainfall and climate indices. More specifically, the influence of ENSO and IOD on rainfall has been investigated using MLR analysis. Usually IOD is measured by an index called Dipole model index (DMI). Since ENSO and IOD both contribute to the creation of rainfall, this paper investigates the relationship of combined ENSO and IOD on Western Australian rainfall. Rainfall data from Roebourne station has been selected as a case study. Several MLR models were investigated using ENSO and IOD as potential predictors of rainfall. Attempts have been made to forecast rainfall by applying the models that satisfied the statistical significance. These correlations could be used for forecasting long-term seasonal rainfall.

2. DATA COLLECTION AND METHODS

Historical rainfall data was collected from the Australian Bureau of Meteorology website (www.bom.gov.au/climate/data/). The rainfall station Roebourne from Western Australia was selected as the case study. The stations were selected based on the recorded length of data having fewer missing values. Observed monthly rainfall in millimetres was obtained for the selected station from January 1890 to December 2013. Data for the climate indices was obtained from Climate Explorer website (http://climexp.knmi.nl). These data (both rainfall and climate indices) were divided into two sets for the calibration and validation respectively. Data from 1890 to 2008 were used for the calibration and data from 2009 to 2013 was used for the validation of the developed MLR models.

To achieve the objective of this study, MLR modelling technique was used. MLR is linear statistical modelling technique which is used to find out the best relationship between a variable and several other variables through least square method. The general equation for MLR model can be expressed as follows:

$$Y = c_0 + a_1X_1 + a_2X_2 + e \quad (1)$$

Where, 'Y' is the rainfall; X_1 and X_2 are the variables of MLR equation (ENSO and IOD in this case), a_1 and a_2 are the coefficients of the respective variables; c_0 is constant and 'e' is error.

For any developed model, evaluation is necessary to determine whether the initiative is worthwhile in terms of delivering the expected outputs. In this research, the performances of the developed MLR models were evaluated by implementing several error indices and statistical performance tests. The agreement or the disagreement of the observed rainfall data with the developed predicted models were evaluated using widely used statistical methods, e.g. root mean square error (RMSE), Pearson correlation coefficients (R), Willmot index of agreement (d). To check the presence of autocorrelation amongst the samples, Durbin-Watson statistical test was also performed according to Field (2009).

3. RESULTS AND DISCUSSION

In this research individual correlation between Western Australian rainfall and monthly climate indices, DMI, Nino3.4 and Southern Oscillation index (SOI) were investigated. The statistical significant correlations of rainfall with climate indices were further analysed using MLR for rainfall forecasting. The analysis showed that Western Australian rainfall has significant correlations with June, July, August and September DMI, Nino3.4 and SOI. Pearson correlations of the individual climate predictors with spring rainfall of Western Australia are shown in Table 1.

Table 1. Significant Pearson correlation (R) coefficients between rainfall and climate indices

Region	Station	Climate indices	Lagged climate indices			
			Jun.	Jul.	Aug.	Sep.
Western Australia	Roebourne	DMI	-0.18*	-	-0.19*	-0.19*
		Nino3.4	-0.24**	-0.26**	-0.31**	-0.27**
		SOI	0.25**	-	-	0.22*

* Correlation is significant at the 0.05 level; ** Correlation is significant at the 0.01 level

From Table 1, it is clear that rainfall significantly relies on ENSO particularly Nino3.4 compared with DMI. Maximum correlation between rainfall and individual indices was observed -0.27. Rainfall was also significantly influenced by June and September SOI; and maximum correlation was in June which is 0.25.

To assess the combined effect of the climate modes on rainfall, the drivers with significant correlations months were further analysed using MLR modelling technique. MLR was performed to investigate the predictability of rainfall using DMI-Nino3.4 and DMI-SOI combination to find out the potential combined predictors. The test for Durbin Watson (D/W) and Tolerances (T) of the developed combined models were also investigated. Amongst the developed forecasting models of the MLR analysis, the models having lower errors were selected as the best models for rainfall forecasting. The summary of the best MLR models for the station along with the values of regression coefficients are shown in Table 2.

Table 2. Variables and constants extracted from the best MLR models

Region	Station	Models	Const.	Coefficients								R	D/W
				DMI				Nino3.4					
				Jun.	Jul.	Aug.	Sep.	Jun.	Jul.	Aug.	Sep.		
Western Australia	Roebourne	DMI_{Jun}-$Nino3.4_{Jun}$	28.80	-19.92				-13.85				0.27	1.92
		DMI_{Jun}-$Nino3.4_{Jul}$	28.67	-19.69					-14.35			0.28	1.95
		DMI_{Jun}-$Nino3.4_{Aug}$	28.30	-15.54						-16.31		0.32	1.97
		DMI_{Jun}-$Nino3.4_{Sep}$	27.74	-17.21							-13.01	0.29	1.97
		DMI_{Aug}-$Nino3.4_{Aug}$	28.13			-10.41				-16.24		0.32	1.97
		DMI_{Aug}-$Nino3.4_{Sep}$	27.54			-12.1					-12.85	0.29	1.97
		DMI_{Sep}-$Nino3.4_{Aug}$	28.11				-5.39			-16.67		0.31	1.97
		DMI_{Sep}-$Nino3.4_{Sep}$	27.61				-6.80				-13.11	0.28	1.98

From Table 2, it can be seen that D/W statistical tests for all the developed MLR models were around two confirming that the residuals of the predicted models have no autocorrelations and they are independent. Therefore, it assured the statistical goodness-of-fit of the models.

Table 3. Performance of the MLR developed models

Region	Station	Models	R	RMSE	d
Western Australia	Roebourne	DMI_{Jun}-$Nino3.4_{Jun}$	0.26	39.12	0.35
		DMI_{Jun}-$Nino3.4_{Jul}$	0.28	38.97	0.37
		DMI_{Jun}-$Nino3.4_{Aug}$	0.32	38.45	0.42
		DMI_{Jun}-$Nino3.4_{Sep}$	0.29	38.88	0.39
		DMI_{Aug}-$Nino3.4_{Aug}$	0.32	38.50	0.41
		DMI_{Aug}-$Nino3.4_{Sep}$	0.28	38.93	0.38
		DMI_{Sep}-$Nino3.4_{Aug}$	0.31	38.59	0.41
		DMI_{Sep}-$Nino3.4_{Sep}$	0.28	39.04	0.37

Various performances of statistics, such as MLR correlations R, RMSE and index of agreement (d) of the best MLR models for the station are shown in Table 3. It was observed that DMI-Nino3.4 based combined predictor models demonstrated statistically significant results with good forecasting capability of rainfall in Western Australia with R = 0.32.

MLR models in validation stage showed very compatible predictive capability of the selected stations with R close to 0.50 to forecasts the sample test set. The RMSE of the validation data set are much lower compared to the calibration stage shown in Table 4. This indicates good forecasting capability of rainfall with the developed models with a high level of accuracy. All the calculated 'd' values in the validation set are close to 0.50 confirming that the combined climate predictors models are capable of forecasting Western Australia's rainfall.

Table 4. Performance of the MLR models for the test data set

Region	Station	Models	R	RMSE	d
Western Australia	Roebourne	DMI_{Jun}-$Nino3.4_{Jun}$	0.49	22.66	0.40
		DMI_{Jun}-$Nino3.4_{Jul}$	0.46	22.84	0.51
		DMI_{Jun}-$Nino3.4_{Aug}$	0.57	21.75	0.64
		DMI_{Jun}-$Nino3.4_{Sep}$	0.51	22.59	0.57
		DMI_{Aug}-$Nino3.4_{Aug}$	0.50	21.27	0.65
		DMI_{Aug}-$Nino3.4_{Sep}$	0.42	21.98	0.58
		DMI_{Sep}-$Nino3.4_{Aug}$	0.51	21.55	0.64
		DMI_{Sep}-$Nino3.4_{Sep}$	0.42	22.40	0.57

The best predicted models were selected considering the lower errors, higher R and d values. The best predicted model developed for the Roebourne station is shown by the following equation:

$$Rainfall = 28.30 - 15.54DMI_{Jun} - 16.31Nino3.4_{Aug} \quad (2)$$

Figures 1 to 2 present the outputs from the best developed regression models. From the Figures, it is clear that the developed MLR models were capable of producing the observed rainfall except the extreme rainfall. As we know that rainfall is the final result of complex global atmospheric global phenomena, prediction of extreme rainfall remains challenge. Instead of only two climate indices, influences of other factors were intense during the extreme rainfall years.

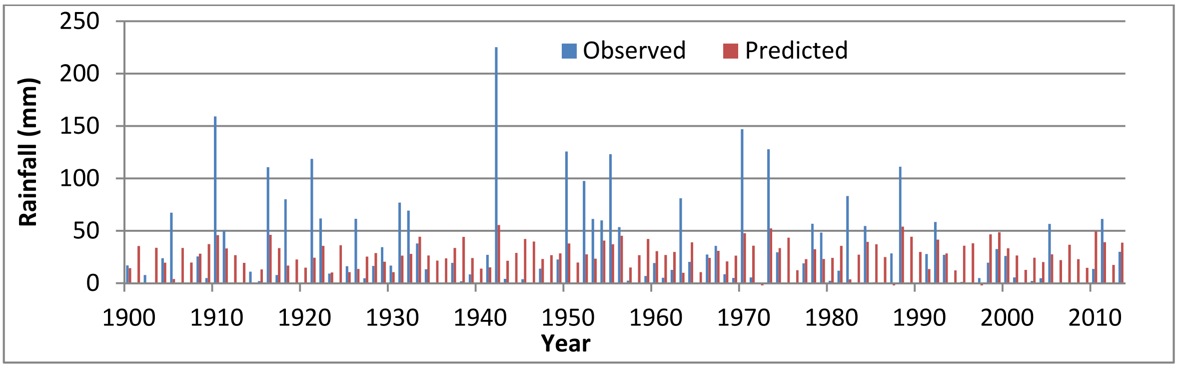

Figure 1. DMI_{Jun}-$Nino3.4_{Aug}$ Combined Modelling outputs

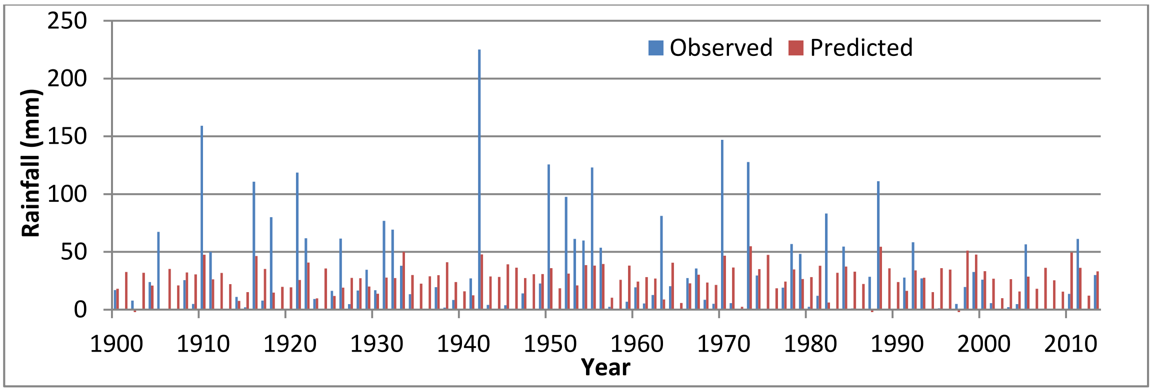

Figure 2. DMI_{Sep}-Nino3.4_{Aug} Combined Modelling outputs

The prediction results, various statistical evaluation parameters as well as statistical significances demonstrated the capability of developed DMI-ENSO based combined MLR models in forecasting Western Australian rainfall with good accuracy. However, some parts of the duration the models are over estimating and some parts they are under estimating from the actual observation. These variations in the prediction are due to the other climatic drivers.

4. CONCLUSIONS

Attempt has been made for the prediction of Western Australian monthly rainfall by considering single and combined climate indices DMI and ENSO as potential predictors. In this study, climate indices DMI and Nino3.4 were selected. The Pearson correlation coefficients of rainfall with each individual climate indices were used to select combination of indices for further analysis with MLR technique. It was discovered that rainfall exhibits significant correlations with four months climate indices (June, July, August and September). The outputs of the analysis showed that discrete impacts of Nino3.4 (ENSO climate drivers) and DMI both have strong influence at Roebourne in Western Australia.

Furthermore, the developed MLR models were validated to investigate the predictive capability of rainfall with separate data set. The statistical errors (R and RMSE) of the validation period for the developed MLR models were lower compared to the calibration period of the data set. Moreover, all the ‘d’ values in the validation stage are very much close to 0.5. This indicates that DMI-ENSO based models could be improved forecasting model for the prediction of West Australian rainfall. Further investigation of MLR technique should be performed in this region to suggest a generalize model for forecasting monthly rainfall in this region.

REFERENCES

Anwar, M.R., Rodriguez, D., Liu, D.L., Power, S. & O’leary, G.J., Quality and potential utility of ENSO-based forecasts of spring rainfall and wheat yield in south-eastern Australia. Australian Journal of Agricultural Research, 59: 112–126, 2008.

Cai, W., van Rensch, P., Cowan, T. & Hendon, H. H., Teleconnection pathways of ENSO and the IOD and the mechanisms for impacts on Australian rainfall. Journal of Climate, 24(15), pp. 3910-3923, 2011.

Evans, A. D., Bennett, J. M. & Ewenz, C. M., South Australian rainfall variability and climate extremes. Climate Dynamics, 33, pp. 477-493, 2009.

Field, A. P. (2009). Discovering statistics using SPSS: SAGE publications Ltd.

He, X., Guan, H., Zhang, X. and Simmons, C. T. (2014). A wavelet-based multiple linear regression model for forecasting monthly rainfall. International Journal of Climatology, 34, 1898-1912.

Hudson D, Alves O, Hendon HH, Wang G. 2011. The impact of atmospheric initialisation on seasonal prediction of tropical Pacific SST. Climate. Dynamics. 36: pp. 1155–1171.

Ihara, C., Kushnir, Y., Cane, M. A. and De La Pena, V. H. 2007. Indian summer monsoon rainfall and its link with ENSO and Indian Ocean climate indices. International Journal of Climatology, 27, 179-187.

Kirono, D. G. C., Chiew, F. H. S., & Kent, D. M., Identification of best predictors for forecasting seasonal rainfall and runoff in Australia. Hydrological Processes, 24(10), pp. 1237-1247, 2010.

Latt, Z. Z. and Wittenberg, H. (2014). Improving Flood Forecasting in a Developing Country: A Comparative Study of Stepwise Multiple Linear Regression and Artificial Neural Network. *Water Resources Management,* 28, 2109-2128.

Mekanik, F., Imteaz, M.A., Gato-Trinidad, S. and Elmahdi, A. (2013). Multiple linear regression and artificial neural network for long-term rainfall forecasting using large scale climate modes. Journal of Hydrology, 503: pp. 11-21.

Mekanik, F. and Imteaz, M.A. (2013) Analysing lagged ENSO and IOD as potential predictors for long-term rainfall forecasting using multiple regression modelling, 20th International Congress on Modelling and Simulation MODSIM 2013, Adelaide, December.

Peel, M.C., McMahon, T.A., Finlayson, B.L. & Watson, T.A., Identification and explanation of continental differences in the variability of annual runoff. Journal of Hydrology 250, pp. 224–240, 2001.

Risbey, J. S., Pook, M. J., McIntosh, P. C., Wheeler, M. C., & Hendon, H. H., On the remote drivers of rainfall variability in Australia. Monthly Weather Review, 137(10), pp. 3233-3253, 2009.

Ummenhoffer, C.C., Gupta, Pook, M.J. and England, M.H. (2008). Anomalous rainfall over southwest Western Australia forced by Indian Ocean Sea surface temperatures. Journal of Climate, 21(19): pp. 5113-5134.

Verdon, D.C., Wyatt, A.M., Kiem, A.S. and Franks, S.W. (2004). Multidecadal variability of rainfall and streamflow: Eastern Australia. Water Resources Research, 40(10): W10201.

Vitart F. 2004. Monthly forecasting at ECMWF. Monthly Weather Review 132: pp. 2761–2779.