Pertanika J. Sci. & Technol. 26 (2): 691 - 706 (2018)

PERTANIKA JOURNALS

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Statistical Modelling of Daily Rainfall Variability Patterns in Australia

Bright Emmanuel Owusu[1,2]* and Nittaya McNeil[1]

[1]*Department of Mathematics and computer Science, Faculty of Science and Technology, Prince of Songkla University, Mueang Pattani, 94000, Thailand*

[2]*Department of Information and Communication Technology/Mathematics, Faculty of Science and Technology, Presbyterian University College, Ghana*

ABSTRACT

The present study uses statistical methods, specifically multiple regression (MR) and gamma generalised linear models (Gamma GLM), for the modelling and analysis of period rainfall variability in Australia during the period 1950-2013. The data from 92 observational stations were collected from the Australian Bureau of Meteorology. Factor analysis was first used to group the 92 stations into eight rainfall regions. The consecutive five-day mean rainfall in a year (period) in each region is then modelled using MR and the Gamma GLM. The results show that the models fit the data quite well in all the regions, but the MR model did better than the Gamma GLM in some of the regions. The MR models revealed three rainfall groupings, and each group had diverse rainfall patterns and trends. Significant decreasing annual rainfall trend was revealed in the southwest and the north southeast regions. In contrast, significant increasing annual rainfall trends were found in the north and the northwest regions.

Keywords: Australia, factor analysis, gamma generalised linear models, multiple regression, rainfall

INTRODUCTION

Increasing severe weather conditions such as flood and drought in many areas of the globe have been attributed to climate change (O'Gorman & Schneider, 2009). Intensified extreme rainfall events are not the only concerns of climate change. Other features of rainfall such as annual mean, occurrence and number of events or the time of the year in which they occur and the sequence and duration of these events can also be affected (Kumar, 2013). Consistent water sources and supply are essential, particularly to the increasing population of human society. They can be rigorously affected by long-term changes in rainfall. The fluctuations in the spatial distribution of rainfall will increase

Article history:
Received: 31 January 2017
Accepted: 27 September 2017

E-mail addresses:
bright.owusu@presbyuniversity.edu.gh (Bright Emmanuel Owusu)
nchirtki@gmail.com (Nittaya McNeil)
*Corresponding Author

ISSN: 0128-7680 © 2018 Universiti Putra Malaysia Press.

differences between dry and wet areas (Held & Soden, 2006; Allan et al., 2010). Therefore, modelling rainfall variability is essential for the determination of the various types of life in the wet tropical regions and mainly for production in dry areas in some countries.

Numerous authors have studied daily rainfall variability and patterns using various statistical techniques. One of the simplest models considered in the analysis and modelling of rainfall is MR. Mekanik, Imteaz, Gato-Trinidad and Elmahdi (2013) applied MR and Artificial Neural Networks (ANN) to estimate long-term seasonal spring rainfall in Victoria. The ANN models revealed greater correlation relative to the MR models, showing that ANN models explained the pattern and trend of rainfall well compared with the MR models.

In contrast, Jeong, St-Hilaire, Ouarda and Gachon (2012) evaluated MR, robust regression, ridge regression and ANN models to establish suitable transfer functions in statistical downscaling models for daily maximum and minimum temperature and daily precipitation occurrence and amounts. They discovered that the monthly MR, annual ANN and annual MR models performed better than the robust and ridge regression models based on the modified Akaike information criterion. In regression modelling, observation must be independent and normally distributed; this assumption may not be true for most time series data such as rainfall. Rainfall data are highly skewed even after aggregation, and in modelling such data, transformation is normally applied. Meng, Zhang, Dai and Yan (2007) applied a logarithmic transformation on monthly rainfall to attain normality. However, application of this transformation to the summer monsoon rainfall amounts in Asia by Mooley (1973) revealed poor results.

Alternatively, transformation of the data to normality can be overlooked and it can directly be modelled using non-normal distributions. In the modelling of periods of non-zero rainfall amount, some authors have revealed that the Gamma GLM fits the data quite well (Coe & Stern, 1982; Stern & Coe, 1984; Kenabatho et al., 2012). Rainfall data between adjacent observational stations are spatially correlated, and numerous methods can be used to manage this concern. Factor analysis has currently gained attention in climate research to describe the variability of correlated observations with possibly fewer unobserved variables called factors. This method has been used in studies such as analysis of the variability of temperature (McNeil & Chooprateep, 2014; McNeil & Chirtkiatsakul, 2016; Wanishsakpong & McNeil, 2016 ;), solar radiation (Cheung et al., 2015) and rainfall (Wickramagamage, 2010). It was applied to put the period rainfall of the meteorological stations into groups. This study uses MR and Gamma GLM models in describing the variability and patterns of five-daily rainfall observations in Australia during the period 1950 to 2013.

MATERIALS AND METHOD

Daily accumulated rainfall from 92 stations throughout Australia for the period between 1950 and 2013 were acquired from the Australian Bureau of Meteorology. The data were analysed using statistical methods. The period rainfall for each station was computed to minimise the skewness and thus, reduce the number of observations to 73 for each year after removing data for 29 February. Data for leap years were excluded to obtain equal observations for all years. Moreover, the period rainfall was deemed appropriate to use because some of the stations had

data on aggregation for as many as five days. The computed mean was based on the number present in cases where data were missing. After computing the five-day average rainfall, only one of the stations had more than 6% (7.3%) missing data. The remaining stations had mostly less than 1% missing data. However, after finding the five-day average region-wise, only one region had 0.2% missing data, and it was imputed by linear regression in fitting the MR model and GLM in fitting the Gamma GLM models. Factor analysis was initially used to put all the stations into groups, and this was done by conducting maximum likelihood factor analysis on the correlation matrix of the period rainfall for each station. The model with p factors can be written as:

$$y_q = \mu_q + \sum_{m=1}^{p} \lambda_q^{(m)} \phi^{(m)} \quad [1]$$

Where, μ_q, is the mean over the five days for each station, λ_q is the factor loading and ϕ is the common factor. The factors were recognised via their loading values. Promax, an oblique rotation method, was applied to control the loading (Venables & Ripley, 2002). A station is assigned to a factor if its loading is at least 0.33 (Hair et al., 1998).

The MR and Gamma GLM models were then fit to the non-zero period rainfall observations in the groups revealed by the factor analysis. These methods used the mean period rainfall observations from each region as the dependent variable and the year and period factors as the independent variables in fitting the models. Brief accounts of these methods are given below.

Let μ_{kst} be the mean period rainfall observation for each region k (k = 1,...8) at period s and year t, then the MR model in this paper is

$$h(\mu_{kst}) = \mu_{0k} + \sum_{i=1}^{64} \eta_{ik} x_{ti} \sum_{j=1}^{73} \beta_{jk} w_{sj} + h(\mu_{t(s-1)k}) + \varepsilon_{kst} \quad [2]$$

where, $h(.)$ represents the dependent variable fourth root transformation, which is defined as $h(\mu_{tsk}) = \sqrt[4]{\mu_{tsk}}$, ε_{tsk} , which are random error terms that are normally distributed, having a mean 0 and variance σ^2. x_{ti} is the predictor for relating years t and year i, w_{sj} is the predictor for relating periods s and period j. Thus, x_{ti} becomes 1 if $t = i$ and 0 elsewhere and w_{sj} becomes 1 if $s = j$ and 0 elsewhere. The year 1950 and period 1 are the starting year and period, respectively. For region k, μ_{0k} denotes the mean period rainfall, which is the starting year and period, η_{ik} is the year i effect and β_{jk} the period j effect. Fit values $h(\mu_{kst})$ from the model [2] have to be transformed backwards to give fit values of μ_{kst}. Thus, the backward-transformed value can be written as μ_{kst}. Let $\hat{\mu}$ be the fit value for μ and $h(\mu)$ the above defined transformation. Then $\hat{\mu}$ can be estimated from the fit value $h(\hat{\mu})$ as $\hat{\mu} = (h(\hat{\mu}))^4$.

The probability density function of the Gamma distribution is normally written as:

$$f(x \mid \alpha, \gamma) = \frac{1}{\Gamma(\alpha)} \gamma^{\lambda} x^{\lambda-1} e^{-\gamma x}, x, \alpha, \gamma > 0 \quad [3]$$

where, α and γ are the shape and the scale parameters, respectively and x is the non-zero period rainfall amount and α is assumed to be constant throughout the dataset for each

region. If μ_s is the mean rainfall amount for period s related to predictors ξ_s then the mean period rainfall amount is estimated by Equation 4, where τ_0 is constant and τ's and ω are slopes.

$$\ln(\mu_s) = \tau_0 + \tau_1\xi_1 + \tau_2\xi_2 + ... + \tau_k\xi_k + \omega\mu_{s-1} \quad [4]$$

R programming language is used for the data analysis and graphical displays (R Development Core Team, 2015).

RESULTS AND DISCUSSION

The factor model categorised the 92 stations into eight groups (factors) via the factor loading. Table 1 gives the loading of the factors organised in descending order within each factor and the uniqueness values excluding the mixed factors.

Table 1
Factor loading of the 92 stations with the identified dominating factors in bold

Stations	Factor Loading								Uniqueness
	F1	F2	F3	F4	F5	F6	F7	F8	
15085	**0.828**								0.381
15086	**0.810**								0.401
15015	**0.788**								0.464
37043	**0.744**			0.115	0.163	-0.132			0.435
14815	**0.731**	-0.145				0.101		0.109	0.405
15005	**0.728**				0.127	-0.106			0.510
2012	**0.703**				-0.119	0.205			0.400
14902	**0.690**	-0.145			-0.103		0.159	0.105	0.309
29008	**0.620**			0.119	0.116	-0.117	0.196		0.489
38003	**0.602**	0.111		0.192	0.156	-0.146			0.516
15590	**0.584**	0.316		0.106			-0.214	-0.107	0.512
3027	**0.562**	-0.114			-0.118	0.286			0.489
14015	**0.512**	-0.118			-0.219	0.107	0.205	0.104	0.349
38020	**0.347**	0.260				-0.114			0.746
14042	**0.499**	-0.112		-0.122	-0.178	0.107	0.325	0.122	0.280
18012	-0.204	**0.804**	0.148						0.353
16005		0.804							0.461
18002	-0.234	**0.775**				0.106			0.449
16031		**0.746**	-0.126			0.127			0.536
16000		**0.739**	-0.124						0.533
18091	-0.205	**0.712**	0.221						0.386

Table 1 (*continue*)

Stations	Factor Loading								Uniqueness
	F1	F2	F3	F4	F5	F6	F7	F8	
11003	-0.205	**0.655**				0.212			0.616
24003		**0.650**	0.176	-0.107				0.254	0.364
17031	0.198	**0.646**	-0.101						0.520
20017		0.593						0.187	0.566
16047	0.317	0.518					-0.161		0.604
46043		0.434		0.163				0.321	0.495
48115		0.346		0.249				0.259	0.578
98003			0.886						0.188
90015			0.874						0.210
91034	0.107		0.860				-0.102		0.293
91033	0.111		0.772						0.387
90059	-0.103	0.228	**0.755**						0.232
99005			0.752					0.207	0.382
94027		-0.184	**0.714**				-0.124	-0.162	0.557
97006		-0.174	**0.688**		-0.121		-0.126	-0.279	0.549
89002		0.197	**0.661**					0.290	0.263
92018			0.444					0.253	0.728
35029				0.806					0.414
41100				0.803				0.152	0.395
43020				0.792					0.377
54023				0.742			-0.173	0.297	0.373
48031				0.721				0.249	0.357
35065			0.115	**0.682**			0.298	-0.128	0.411
39020				0.580			0.344		0.534
40141	-0.145		-0.124	**0.559**			0.207		0.594
59000		-0.118	-0.254	**0.552**		0.109		0.199	0.520
44181	0.144	0.225		0.380		-0.112		0.168	0.569
10620		-0.169			0.918				0.282
8230		-0.126			**0.849**	0.137			0.333
8004					**0.809**	0.102			0.336
9518					**0.775**	-0.156			0.299
12011					**0.736**	0.265			0.420
9594		0.165			0.554				0.563
7059						0.724			0.511
7103		0.113				0.718			0.514

Table 1 (*continue*)

Stations	Factor Loading								Uniqueness
	F1	F2	F3	F4	F5	F6	F7	F8	
12108		0.171				0.674			0.538
13002		0.183			-0.153	**0.637**			0.588
4006	0.175				-0.156	**0.628**			0.536
7049					0.267	**0.615**			0.526
5001	0.134				-0.116	**0.601**			0.572
5008						0.594			0.630
6029					0.129	**0.554**	0.107		0.651
4019	0.290				-0.101	**0.487**			0.615
33001		0.128		0.152			0.741		0.394
33013				0.284			**0.711**	-0.118	0.401
32001	0.163	0.125	-0.108				0.698		0.353
31062	0.181	0.129	-0.150	-0.201			0.645		0.471
27042	0.328			-0.132	-0.188		0.582		0.308
27005	0.261			-0.139			0.477		0.401
73025		0.146	0.218	0.124				**0.592**	0.352
69018			-0.193	0.268				**0.542**	0.565
84016			0.114	0.116				**0.530**	0.661
80044		0.333	0.329					**0.455**	0.348

The eight factors were made up of the following regions of Australia: Factor 1 (F1), the north (N), Factor 2, (F2) the central south (CS), Factor 3, (F3) the south southeast (SSE), Factor 4, (F4) the central east (CE), Factor 5, (F5) the southwest (SW), Factor 6, (F6) the northwest (NW), Factor 7, the northeast (NE) and Factor 8, (F8) the north southeast (NSE). Figure 1 shows the stations as grouped by factor analysis. The classification reveals the spatial rainfall distribution of the 92 stations used in the study. A similar classification was uncovered by factor analysis (Cheung et al., 2015) in the analysis of spatial and temporal patterns of solar absorption by clouds and the modelling of daily maximum temperature in Australia by Wanishsakpong and McNeil (2016).

Further concern for the modelling of time series observation is the dependency on response observation, which contravenes independent error assumption. Figure 2 is the autocorrelation function (ACF) that is used to evaluate the autocorrelation of mean period rainfall. Some of the sample lag values (the black vertical lines) are beyond the 95% confidence interval line (the horizontal dotted lines), which indicates significant dependency among the response variables. The dependency among the period rainfall in each region was minimised by adding the AR(1) term (Chatfield, 1996) to both the MR and Gamma GLM models (Figure 3). Most of the sample lag values are within the 95% confidence interval, indicating a drastic reduction of serial autocorrelation in the models. Also, multicollinearity was assessed using the variance inflation factor (VIF), which is defined as $1 / (1-R^2)$, where

R^2 is the coefficient of determination. The response variables have been noted to have multicollinearity if VIF >10 (Lin, 2008). The values of the VIF were below 5 in all the regions, which revealed no indication of multicollinearity.

The pattern of period rainfall amount for all the eight regions as modelled by Equations 2 and 4 are shown in Figure 4. Both models showed an obvious seasonal periodic pattern, and this is clearly seen especially in the N and NE regions, where over 52% of the variations were explained by the models (coefficient of determination of 0.54 and 0.52, respectively). Moreover, the fluctuations and their timing in the data were well displayed by the models as well as the 64 years' mean. However, the models underestimated the magnitude of the fluctuations. Relatively, there is not much difference between the models from both methods in the remaining regions as shown in Figure 4. The eight factor regions display diverse rainfall variability and trends for all the fit models.

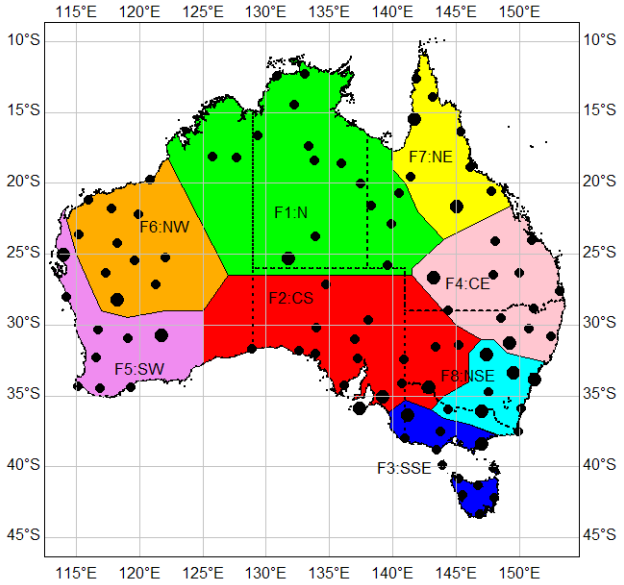

Figure 1. The locations of the stations used in the study and the groups (colours) revealed by factor analysis. The large points represent mixed factors. The state boundaries are also included (dotted lines)

The N and NE, which is in the tropical region of Australia, have similar period rainfall variation as in the NW and the CE. The pattern in CS and NSE was also similar to that in the SW and SSE. In general, the NE region relative to the other regions had the highest amount of period rainfall in the seasonal span, which can be due to a substantial number of temporal meteorological conditions such as fronts and low pressure systems, which possibly can disturb rainfall in this area. The area is additionally more susceptible to severe weather systems such as tropical cyclones and a great deal of convective clouds entrenched, resulting in greater proportion of cloud cover relative to other regions (Cheung et al., 2015).

The models were assessed by plotting residuals alongside the theoretical quantiles (Figure 5). The figure shows that the residual plots had low variations with respect to the expected line of best fit, indicating that both models fit the data quite well in the factor regions except for

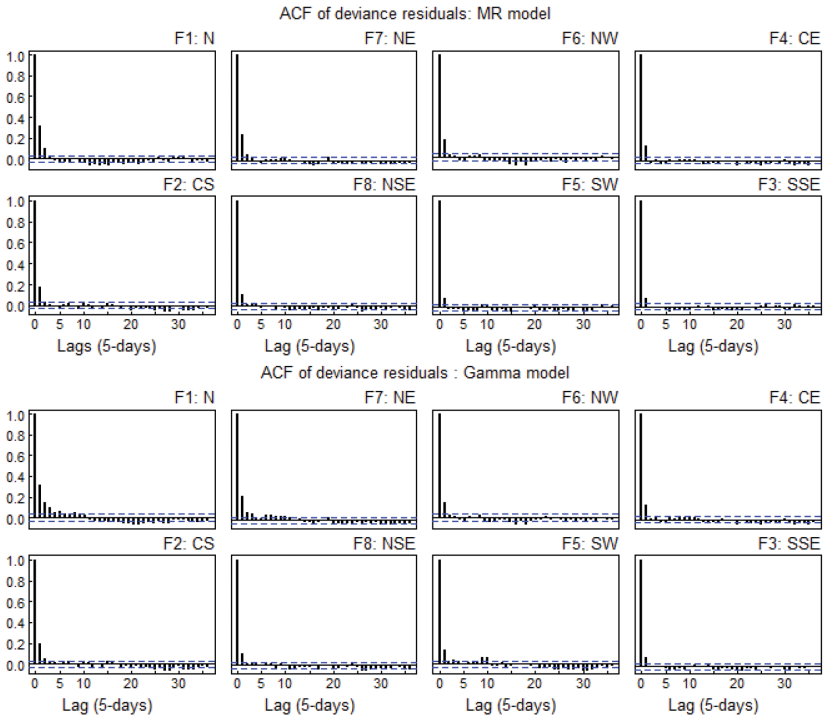

Figure 2. The autocorrelation function for the models. The top and bottom panels relate to before minimising serial correlation

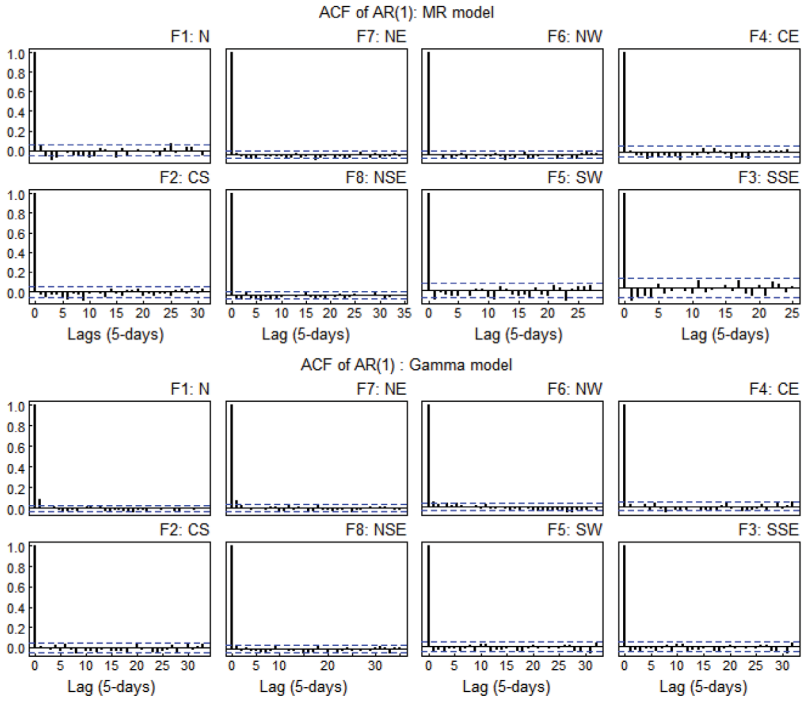

Figure 3. The autocorrelation function for the models. The top and bottom panels relate to after minimising serial correlation

little departures at the upper and lower tails of some of the models. The MR model seemed to perform better than the Gamma GLM model in the N, NE and SW regions. Analysis of the deviance residuals and the fit values of the models did not show any patterns and outliers.

Further, the MR was used to model the 64 years' mean period to show the monthly and annual rainfall patterns for all the regions. This model was applied because it is one of the simplest methods to describe monthly rainfall patterns (Hughes & Saunders, 2002; Oettli & Camberlin, 2005). It also fit the period mean rainfall better than the Gamma GLM models in most of the regions.

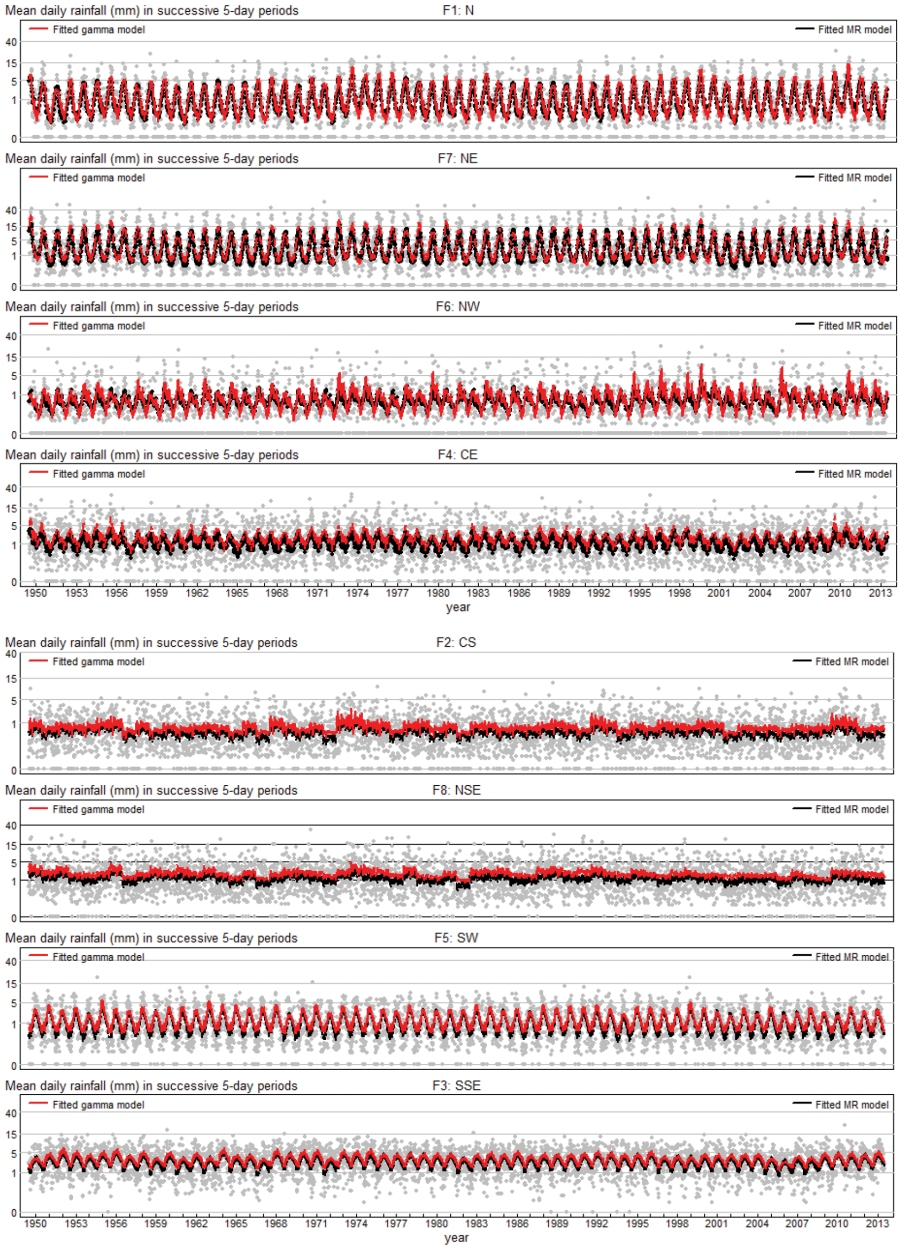

Figure 4. The plots of period rainfall and their estimated mean patterns in eight regions during 1950-2013. The grey dots represent the observed period's rainfall, while the black and red curves represent their estimated mean from models 2 and 4, respectively

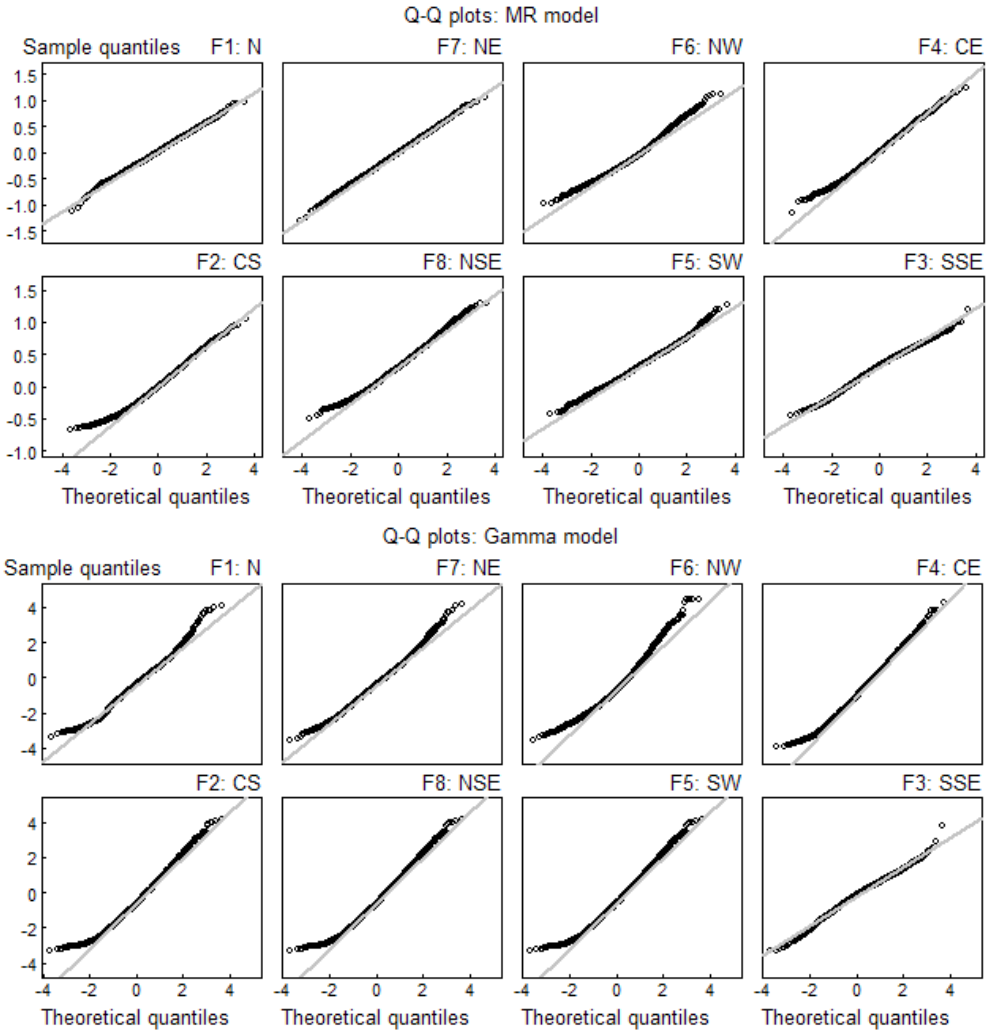

Figure 5. The residual quantile-quantile (Q-Q) plots for the models. The top graph shows that of the MR (response variable is fourth root transformed) and the bottom shows the Gamma. Both models were filtered with AR(1) term

Figure 6 shows the results of the estimated rainfall variations for the 64 years for all the regions. Analysis of the graph indicated three main categories regarding the period's rainfall estimates. The models in the N, NE, CE and NW regions can be grouped as one category (first group) of items possessing similar patterns, while the models in the CS and NSE can be grouped together (second group).

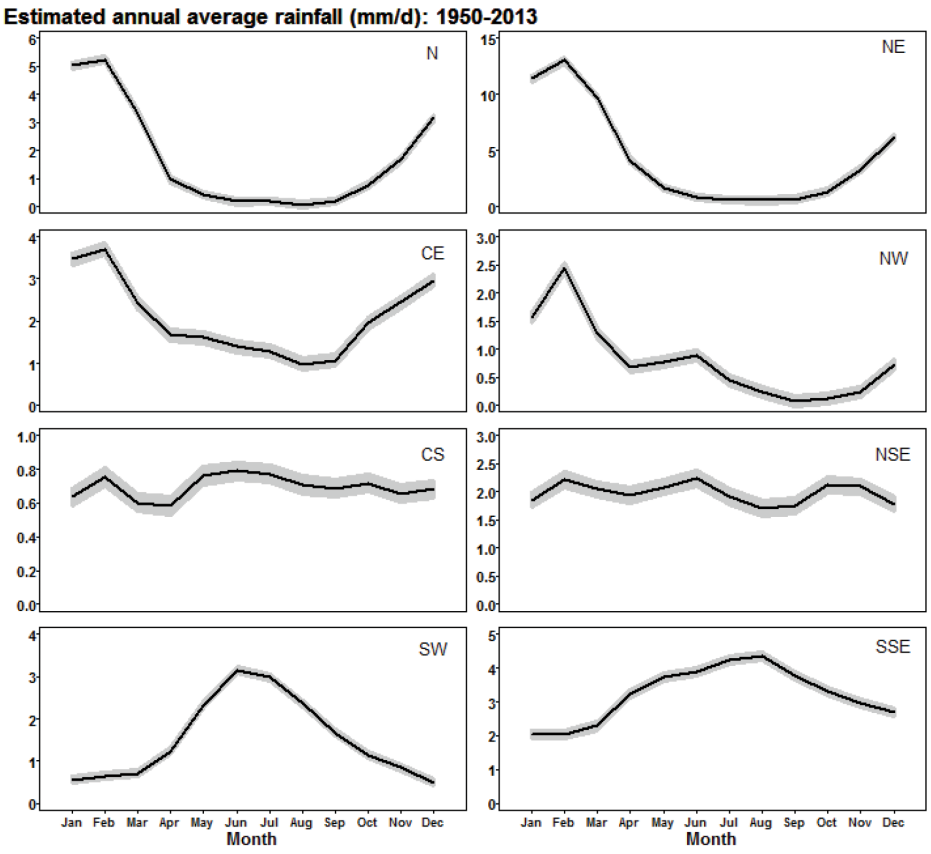

Figure 6. The estimated monthly rainfall (line) and the 95% confidence interval (shaded) in each month of a year in the factor regions

The SW and SSE also form a group (final group) with similar patterns. In the first group, the period rainfall rises sharply from January to March, reaches its maximum between this period, then decreases steadily to its minimum between August and September. The period rainfall observed in the second group is quite stable throughout the year (it varies from 0.6 to 2.2 mm d^{-1}). The largest estimated period rainfall in these regions occurred between May and July, but the minimum was observed during March and April in the CS and occurred between August and October in the NSE. On the other hand, the final group received their minimum period rainfall between January and February. The period rainfall increased steadily from February until it attained the largest between June and August. Generally, the estimated period rainfall is different in each region: the NE, which comprises equatorial and tropical regions, has the highest in the Austral summer during the active part of the monsoon, while the mid-latitude region of SSE has the most during the Austral winter.

Figure 7 shows the results of the estimated inter-annual rainfall pattern. In all eight regions, these patterns fluctuated between 0.3 and 7.1 mm d^{-1} and attained the lowest value between 1950 and 1975 in the NW. However, the highest value occurred during 1950 and 1960 in the

NE. In general, the highest annual rainfall values in most of the regions also occurred between 1950 and 1970 except in the N, CS and NW regions, where it occurred between 1970 and 2012. The estimated annual rainfall amount was low, especially in CS and the NW relative to the remaining regions.

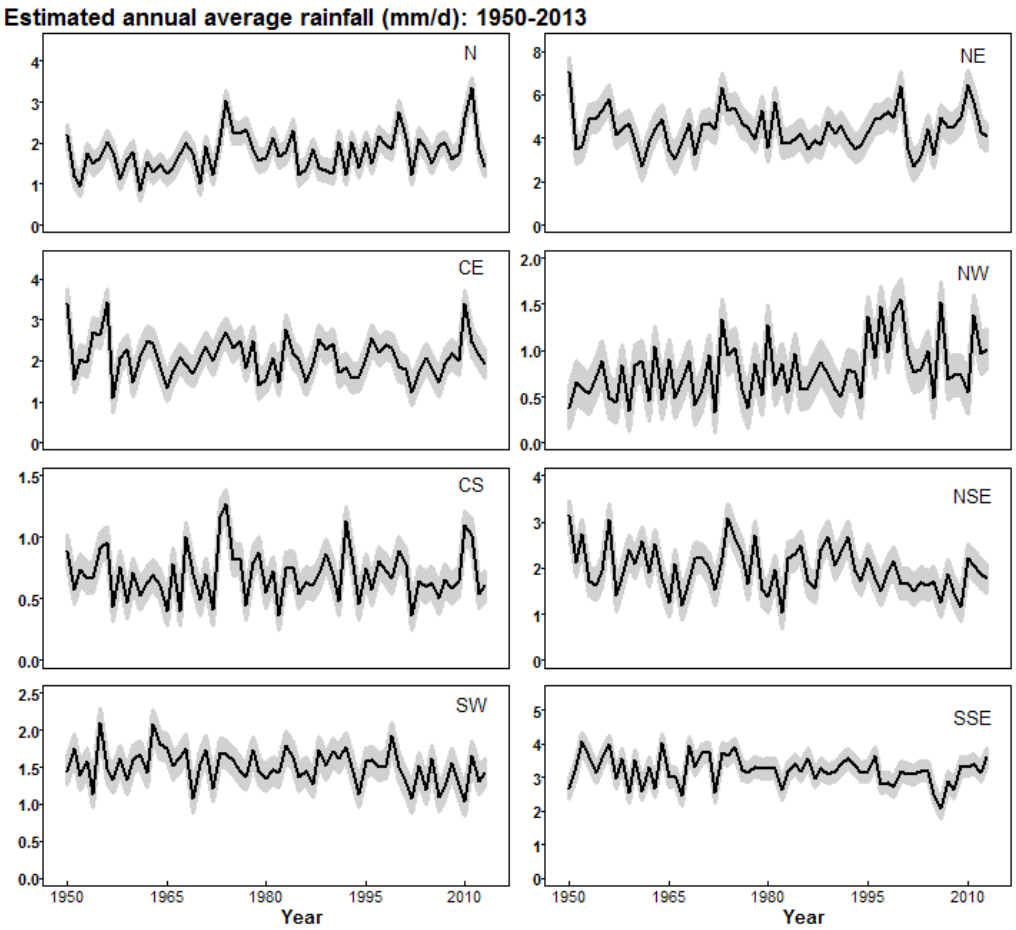

Figure 7. The estimated annual rainfall (line) and the 95% confidence interval (shaded) for factor regions from 1950-2013

Examination of the 64 years' annual rainfall time series by linear regression indicated negative and positive slopes for the regions (Table 2). Analysis of Table 2 revealed that the CE, CS, NSE, SW and SSE regions were characterised by negative slopes, which indicated decreasing trends. Significant decreasing trends were evident in the NSE and SW regions. The IOCI (2002) also reported long-term decline in winter (May-October) rainfall of 15-20% since the 1970s in the southwestern corner of Western Australia. Average spring rainfall during 1997-2006 in the southeast of Australia was revealed to be below average and resulted in the

problem of recurring drought in this part of Australia (Murphy & Timbal, 2008). Hope et al. (2010) revealed similar results in the analysis of temporal relationship concerning rainfall variability in the SW and southeast of Australia.

Timbal and Fawcett (2013) also observed unique shortfalls in rainfall at southeastern Australia. The reported rainfall deficit was evident in pre-winter and early winter rainfall. These declining trends may be partly due to changes in large-scale atmospheric circulation (Nicholls, 2006; Timbal & Fawcett, 2013) including a pole-ward movement of the westerly winds and increasing atmospheric surface pressure. Changes in anthropogenic greenhouse gases and ozone levels (Delworth & Zeng, 2014) and expansion of the Southern Hemisphere Hadley cell (Post et al., 2014) can affect decreasing rainfall trends in these regions.

Table 2
Use of linear regression for analysis of annual rainfall trends

Region	N	NE	CE	NW	CS	NSE	SW	SSE
Slope	0.0076	0.0065	-0.0029	0.0001	-0.0001	-0.0075	-0.0035	-0.0050
p-value	0.0302	0.8757	0.4725	0.0026	0.7898	0.0235	0.0444	0.0964

In contrast, significant increasing annual rainfall trends were evident in the N and NW regions of Australia, while a non-significant increasing trend at 95% confidence level was evident in the NE region (Table 2). Unlike the N and the NW regions, the NE region is mostly liable to intense weather conditions, for instance, tropical cyclones that bring a convergence of convective clouds. This explains the greater part of cloud cover in the N and NW regions (Cheung et al., 2015). It could also explain part of the increasing rainfall trend in this region since cloud formation is related to rainfall.

CONCLUSION

This study used statistical models in describing the period rainfall in Australia during 1950-2013. It is based on 64 years' daily accumulated observation obtained from 92 stations distributed over Australia. Factor analyses identified eight factors, which corresponded to eight geographical rainfall regions. The eight regions include the north, northeast, central east, northwest, central south, north southeast, southwest and south southeast. The rainfall from each region exhibits diverse overall means with various seasonal variations and were explored using Gamma generalised linear and multiple regression models.

These methods were applied to model the period rainfall obtained from each of the regions. The period rainfall observations in successive periods were serially correlated, and this was minimised using the AR(1) technique in fitting the models. Both models fit the data quite well in all the regions, particularly in the tropical region (lower latitudes) where the atmospheric internal variability is small relative to the forced change. However, the multiple regression models did better than the Gamma generalised linear model in the north, northeast and the southwest regions.

The estimated period and annual rainfall means by the multiple regression models revealed three different groups, and each of the groups showed diverse monthly or annual patterns. In general, the eight identified regions of the factor analysis showed three main periods of rainfall category, with each having a distinctive dominating pattern. The understanding and knowledge of these patterns are essential for planning the rainfall resources in numerous regions of Australia. The northeast received the highest annual rainfall followed by the south southeast. In contrast, central south and the northwest received low rainfall with a similar pattern during this period. The decreasing annual rainfall observed in the southern regions may be due to a reduction in Austral autumn and winter rainfall over the southern parts of Australia, particularly South Western Australia.

ACKNOWLEDGEMENT

The work was supported by the Higher Education Research Promotion and Thailand's Education Hub for the Southern Region of ASEAN Countries under grant no. THE-034. We are most grateful to Emeritus Prof. Don McNeil.

REFERENCES

Allan, R. P., Soden, B. J., John, V. O., Ingram, W., & Good, P. (2010). Current changes in tropical precipitation. *Environmental Research Letters*, *5*(2), 1–7.

Chatfield, C. (1996). *The analysis of time series*. Melbourne, Australia: Chapman & Hall.

Cheung, K. K., Chooprateep, S., & Ma, J. (2015). Spatial and temporal patterns of solar absorption by clouds in Australia as revealed by exploratory factor analysis. *Solar Energy, 111*, 53–67.

Coe, R., & Stern, R. D. (1982). Fitting models to daily rainfall data. *Journal of Applied Meteorology*, *21*(7), 1024–1031.

Delworth, T. L., & Zeng, F. (2014). Regional rainfall decline in Australia attributed to anthropogenic greenhouse gases and ozone levels. *Nature Geoscience*, *7*(8), 583–587.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate Data Analysis* (5th ed.). New York, NY: Prentice Hall International.

Held, I. M., & Soden, B. J. (2006). Robust responses of the hydrological cycle to global warming. *Journal of Climate*, *19*(21), 5686–5699.

IOCI. (2002). Climate variability and change in south west Western Australia. *Indian Ocean Climate Initiative Panel.* Perth.

Hope, P., Timbal, B., & Fawcett, R. (2010). Associations between rainfall variability in the southwest and southeast of Australia and their evolution through time. *International Journal of Climatology*, *30*(9), 1360–1371.

Jeong, D. I., St-Hilaire, A., Ouarda, T. B. M. J., & Gachon, P. (2012). Comparison of transfer functions in statistical downscaling models for daily temperature and precipitation over Canada. *Stochastic Environmental Research and Risk Assessment*, *26*(5), 633–653.

Kenabatho, P. K., McIntyre, N. R., Chandler, R. E., & Wheater, H. S. (2012). Stochastic simulation of rainfall in the semi-arid Limpopo basin, Botswana. *International Journal of Climatology*, *32*(7), 1113–1127.

Kumar, P. (2013). Hydrology: Seasonal rain changes. *Nature Climate Change*, *3*(9), 783–784.

Lin, F. J. (2008). Solving multicollinearity in the process of fitting regression model using the nested estimate procedure. *Quality and Quantity*, *42*(3), 417–426.

Lloyd-Hughes, B., & Saunders, M. A. (2002). Seasonal prediction of European spring precipitation from El Niño-Southern oscillation and local sea-surface temperatures. *International Journal of Climatology, 22*(1), 1–14.

McNeil, N., & Chirtkiatsakul, B. (2016). Statistical models for the pattern of sea surface temperature in the North Atlantic during 1973-2008. *International Journal of Climatology, 36*(11), 3856-3863.

McNeil, N., & Chooprateep, S. (2014). Modeling sea surface temperatures of the North Atlantic Ocean. *Theoretical and Applied Climatology*, *116*(1–2), 11–17.

Mekanik, F., Imteaz, M. A., Gato-Trinidad, S., & Elmahdi, A. (2013). Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. *Journal of Hydrology, 503*, 11–21.

Meng, W. G., Zhang, Y. X., Dai, G. F., & Yan, J. H. (2007). The formation and development of a heavy rainfall mesoscale convective system along southern China coastal area. *Journal of Tropical Meteorology*, *23*(6), 521-529.

Mooley, D. A. (1973). Gamma distribution probability model for Asian summer monsoon monthly rainfall. *Monthly Weather Review*, *101*(2), 160–176.

Murphy, B. F., & Timbal, B. (2008). A review of recent climate variability and climate change in southeastern Australia. *International Journal of Climatology, 28*(7), 859–879.

Nicholls, N. (2006). Detecting and attributing Australian climate change: A review. *Australian Meteorological Magazine*, *55*(3), 199–211.

Oettli, P., & Camberlin, P. (2005). Influence of topography on monthly rainfall distribution over East Africa. *Climate Research, 28*(3), 199–212.

O'Gorman, P. A., & Schneider, T. (2009). The physical basis for increases in precipitation extremes in simulations of 21st-century climate change. *Proceedings of the National Academy of Sciences*, *106*(35), 14773–14777.

Post, D. A., Timbal, B., Chiew, F. H., Hendon, H. H., Nguyen, H., & Moran, R. (2014). Decrease in southeastern Australian water availability linked to ongoing Hadley cell expansion. *Earth's Future, 2*(4), 231–238.

R Development Core Team. (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, URL http://www.R-project.org

Stern, R. D., & Coe, R. (1984). A model fitting analysis of daily rainfall data. *Journal of the Royal Statistical Society. Series A (General), 147*(1), 1–34.

Timbal, B., Arblaster, J., Braganza, K., Fernandez, E., Hendon, H., Murphy, B., & Wheeler, M. (2010). *Understanding the anthropogenic nature of the observed rainfall decline across South Eastern Australia*. Centre for Australian Weather and Climate Research.

Timbal, B., & Fawcett, R. (2013). A historical perspective on southeastern Australian rainfall since 1865 using the instrumental record. *Journal of Climate*, *26*(4), 1112–1129.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (4th Ed.). Queensland: Springer.

Wanishsakpong, W., & McNeil, N. (2016). Modelling of daily maximum temperatures over Australia from 1970 to 2012. *Meteorological Applications*, *23*(1), 115–122.

Wickramagamage, P. (2010). Seasonality and spatial pattern of rainfall of Sri Lanka: Exploratory factor analysis. *International Journal of Climatology*, *30*(8), 1235–1245.