Songklanakarin J. Sci. Technol.
41 (4), 870-878, Jul. – Aug. 2019

http://www.sjst.psu.ac.th

Original Article

Statistical modelling of 5-day average rainfall in Australia during 1950-2013

Bright Emmanuel Owusu[1, 2] Nittaya McNeil[1, 3*] and Mayuening Eso[1, 3]

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

[2] *Department of Mathematics, Faculty of Science and Technology, Presbyterian University College Ghana, Abetifi-Kwahu, Ghana*

[3] *Centre of Excellence in Mathematics, Commission on Higher Education, Ministry of Education, Ratchathewi, Bangkok, 10400 Thailand*

Received: 3 September 2017; Revised: 24 February 2018; Accepted: 18 April 2018

Abstract

Many regions across Australia have high rainfall variability, which has various effects on water and food availability. This study examines rainfall patterns over consecutive 5-day periods during 1950-2013 using data from 92 observational stations in Australia. The first model used factor analysis to classify the stations into distinct geographical regions. Gamma generalised linear model (GLM) was then fitted to describe the rainfall amount in each category with season and year factors as the predictors.

Factor analysis revealed eight factors, which represent eight geographical regions of Australia. Analysis of the similarities in the seasonal evolutions between regions revealed three seasonal rainfall groupings. The GLM models fitted the data quite well in all the areas and showed that 5-daily rainfall is significantly affected by the period of the year and its annual average in most of the regions. The models could be used to simulate rainfall data for the areas with inadequate rainfall records.

Keywords: rainfall variability, classification, generalised linear models, factor analysis

1. Introduction

Many areas of the globe, particularly in the tropics, are extremely susceptible to changes in rainfall patterns. A study by Trenberth, Dai, Rasmussen, and Parsons (2003) revealed that, for the past few decades, the global hydrological cycle has been undergoing significant changes, which include rainfall amount, frequency and duration. These changes in rainfall have resulted in severe floods or droughts, which often have diverse effects on human life and health, food and provision of potable water, ecosystems, and infrastructure (Herring *et al.*, 2016).

*Corresponding author
Email address: nittaya.ch@psu.ac.th

Modelling and analysis of rainfall with appropriate statistical methods have diverse applications not just for predictive purposes but also in hydrological analysis, agricultural production and meteorological planning. They also facilitate sound understanding of spatial-temporal rainfall distribution patterns. Thus, classifying and model fitting analysis of various rainfall events, such as the probability of occurrence, the mean amount in a month or year, and distribution patterns, will be beneficial to different stakeholders in their policy planning.

Many authors have analysed rainfall amount with various mathematical models. Many have used methods including artificial neural networks (Deo & Shahin, 2015) and the *K*-Nearest-Neighbours (Toth, Brath, & Montanari, 2000). Other studies also used multiple linear regression (Shukla, Tripathi, Pandey, & Das, 2011) and support vector machines

for regression (Lin & Jhong, 2015) methods. Bagirov, Mahmood, and Barton (2017) evaluated a model with clusters linear regression (CLR) of the rainfall data in Australia and observed the superiority of CLR over the other models.

Rainfall observations are known to be skewed; the gamma generalised model has been found by many researchers (Coe & Stern, 1982; Kenabatho, McIntyre, Chandler, & Wheater, 2012; Stern & Coe, 1984;) to fit rainfall observations reasonably well. Cai, Cowan and Thatcher (2012) have revealed rainfall reductions over semi-arid parts of the Southern Hemisphere, such as southern coastal Chile, southern Africa, and southeastern Australia. There have been various studies on different aspects of Australian rainfall. For instance, Drosdowsky (1993) has studied the seasonal rainfall and has revealed that it is highly variable throughout the country.

Trend analyses of rainfall (Chambers, 2003; Lavender & Abbs, 2013; Taschetto & England, 2009) have shown significant observed changes, with increases in the northwest and decreases in the east during the period of 1970-2009. These rainfall changes could be partially associated with tropical cyclones and other low pressure systems as revealed by Lavender and Abbs (2013). Similar studies (Beecham, Rashid, & Chowdhury, 2014; Srikanthan & Pegram, 2009;) have used a multi-site two parts model to describe the daily variability at some sites. They described the occurrences and amounts with different models and observed acceptable model fits in all the sites.

Hasan and Dunn (2010) have used a Poisson-gamma model to describe monthly rainfall observations for agricultural planning. They used monthly observations from 220 stations and observed acceptable model fits on modelling the mean rainfall amount and the probability of occurrence. Knowledge of the trends in the spatial-temporal rainfall variability and their physical explanations are essential in climate change assessment (Kenabatho *et al.*, 2012).

Many scholars (Bärring, 1987, 1988; Um, Yun, Jeong, & Heo, 2011; Wickramagamage, 2010;) have studied rainfall regions classification. Most of these studies applied Principal Component Analysis (PCA) and factor analysis to describe the spatial-temporal variability based on classification into distinct regions. These methods delineated rainfall regions reasonably well in all their study areas. However, the latter gave easily interpretable results relative to the PCA. Chambers (2003) used PCA and cluster analysis to group rainfall in South Australia and observed that both methods gave similar clusters. Factor analysis models have been widely used in several meteorological applications to classify multiple outcomes, of data collected from stations spread over a region of considerable size such as a continent or nation (Bukantis, 2002; Cheung, Chooprateep, & Ma, 2015; Unkel, Trendafilov, Hannachi, & Jolliffe, 2010). The use of the method by these authors may be motivated by its ability to account for spatial correlations in the observations. The present paper fits appropriate statistical models to classify and describe the variability patterns of the 5-day rainfall amount in Australia.

2. Materials and Methods

This study used 64 years (1950-2013) daily aggregated rainfall data for 92 observational weather stations spread over Australia. The data were obtained from Australian Bureau of Meteorology (http://www.bom.gov.au/climate/data). These 92 stations were selected to give a sample covering the whole area as evenly as possible, and they have continuous rainfall records that extend over a period of 64 years. Data recorded on leap years were omitted to maintain regular year with 365 days observations, and thus each station had 23,360 observations from the 64 years of records. For each site, the 5-day average rainfall defined as the average over successive 5-days, was computed and used in this study.

This choice in the statistical analysis has the following advantages; the proportion of missing data is substantially reduced, and daily patterns are represented graphically well. Correlation between data in consecutive periods is also considerably reduced, and the number of parameters in the model is reduced by a factor of 5, facilitating data management, computations and providing smoother model fits. Where data are missing, the average is estimated based on the number present. If all the observations in a 5-day period are missing, the GLM model, described in the next section, is used for the imputation of these missing values.

Statistical methods were used to analyse the data. The first method used factor analysis to classify the 92 weather stations into factors, which correspond to different geographical regions. The factor model with m factors $(f_1, f_2, f_3, ..., f_m)$ can be written as;

$$y_t = \mu + \sum_{j=1}^{m} \lambda_j f_j, \quad (1)$$

where y_t is the rainfall for time t, μ is the average rainfall amount of each station, λ_j is the factor loadings on the j^{th} factor while f_j is the j^{th} common factor. The model also provides uniqueness values to the variables, and those variables with high uniqueness values cannot be assigned to any factor (Rencher, 2002).

The factor analysis technique was then applied to the correlation matrix of the 5-day period rainfall observations and the parameters were estimated by the maximum-likelihood method. The factors were rotated by the "Promax" one of the "oblique" rotation techniques to achieve simple structure patterns of loadings, and each factor is identified by the estimated loadings (Venables & Ripley, 2002). A stations was assigned to a factor if the loading was at least 0.33 and this was also the largest among the factor loadings (Hair, Anderson, Tatham, & Black, 1998).

The second model involves fitting gamma GLM to the non-zero rainfall amount for these regions revealed by the factor analysis to generate its amount on wet days. The shape parameter of the gamma distribution is assumed constant through data set for each region. For instance, if μ_i is the mean rainfall amount for any period i connected with predictor vector x_i, then the average 5-day rainfall amount is estimated by equation (2) where α is the coefficient vector.

$$\ln(\mu_i) = x_i\alpha. + \mu_{i-1}, \quad (2)$$

where μ_{i-1} denotes the AR(1) term. A significant concern in fitting models to time series data such as rainfall is the dependence among the response variable (serial correlation) which violates the independence assumption. The serial correlation was minimised by the introduction of AR(1) term in the modelling. The predictors used in this study were the 5-day rainfall periods (seasonal), the annual daily rainfall (year factors) and the AR(1) term. Therefore, 73 seasonal, 64 year factors and the AR(1) terms were employed in the models. The treatment contrast is usually used to fit non-Gaussian GLMs to obtain model coefficients that are easy to interpret. The models were thus fitted using the treatment contrast which makes the first coefficient for each factor a reference; such that each coefficient represents a comparison of that coefficient with the first factor (Venables & Ripley, 2002).

The fitted models were assessed by the residuals quantile-quantile (Q-Q) plots. The quantile-quantile plot is one of the measures used to check the model fit at each observation for GLMs. The sum contrasts (Tongkumchum & McNeil, 2009) were used to obtain 95% confidence intervals (CI) to compare the fitted model means with the overall rainfall means. This contrast gives criteria to classify levels of the factor into three groups, according to whether each relating CI exceeds, matches, or is below the overall mean. All data analysis and graphical displays were carried out in R version 3.2.3 (R Development Core Team, 2013).

3. Results and Discussion

Factor analysis grouped Australia into eight distinct geographical rainfall regions and these regions are made up of groups of stations that load with just one factor. The analysis was conducted on several factors, but the factor loadings showed that eight factors were adequate and accounted for 52% of the observed correlations in the data with correlation values of at least 0.34. Also, 15 stations out of the 92 correlated with factor 1 (F1), 13 with factor 2 (F2), 10 with factors 3, 4 and 6 (F3, F4, and F6) and 7 with factor 5 (F5). Six stations correlated with factor 7 (F7) and 4 with factor 8 (F8). A similar classification was revealed by Wanishsakpong and McNeil (2016) in the modelling of daily maximum temperature in Australia. However, 17 stations correlated to more than one factor in this study (mix-factors). The seventeen stations with mixed factor loadings were not used in the GLM modelling. These mixed factor stations, most of them over southeast Australia, may be due to contributions from different climate regimes to the rainfall in this region, including tropical and subtropical processes. A similar reason may apply to the three mixed factor stations observed over Western Australia as well. Figure 1 depicts the map of Australia, locations of the observational stations used in the study and the eight regions as determined by the factor model. These regions include F1: North (N), F2: Central South (CS), F3: South-Southeast (SSE) and F4: Central East (CE). The remaining regions are F5: Southwest (SW), F6: Northwest (NW), F7: Northeast (NE) and F8: North-Southeast (NSE).

Australian Bureau of Meteorology has categorised the climate in Australia into six regimes, namely tropical, grassland, desert, equatorial, temperate and subtropical group. These administrations were revealed using the modified Köppen scheme (Köppen, 1918), which identified climate

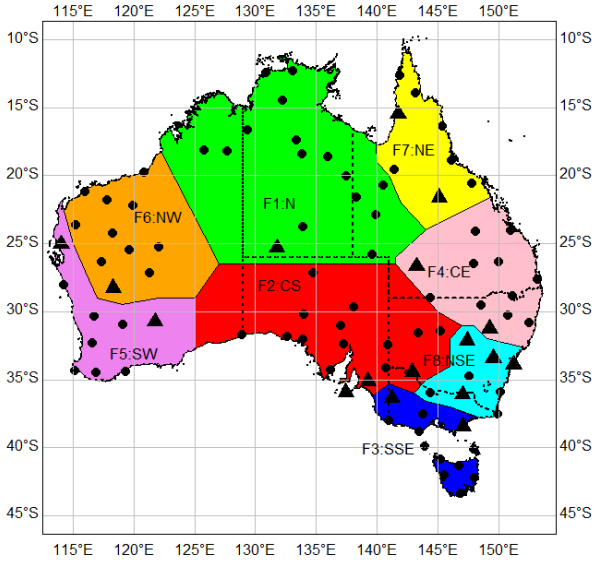

Figure 1. Map of rainfall factor regions of Australia during 1950-2013. The small sized points show stations that correlated with only one factor and the triangle points show stations that correlated with more than one factor (mixed factors).

boundaries with a mixture of natural landscape topographies and parts of human experience. The tropical climate system is mainly in Northern Australia while the grassland climate system is to the south with the desert classification in central Australia. The equatorial area is the pointed part of Cape York, and Bathurst and Melville Islands in the north of Darwin, and temperate regime mainly exist in southeastern Australia. The subtropics are more central coast Australia, and the northeastern part is more tropical (northern Queensland).

The eight geographical regions revealed by the factor analysis do not follow the climate regime directly. Comparatively, the factor region N revealed by the factor analysis mostly comprises desert, grassland and tropical areas, while the NE is made up of mostly the equatorial and the tropical areas. The CS stands out as part of the desert area while the SW, NSE and the SSE are mostly part of the temperate regime. The CE is observed to be a mixture of the grassland and subtropical categories. Besides, the NW comprised the grassland and the desert land categories.

3.1. Annual rainfall

The annual rainfall amount on 5-day period and its probability of occurrence defined as the likelihood of non-zero annual 5-day mean rainfall on any particular year in all the eight regions are shown in Figure 2. There are no distinct annual patterns revealed by the graph. Among all the regions, the NE received the highest average of 4.43 mm d^{-1} followed by SSE with a mean of 3.22 mm d^{-1} while the CE received an average of 2.08 mm d^{-1}.

NSE, N and SW followed in that order with averages 1.98, 1.76 and 1.51 mm d^{-1} respectively. The NW received 0.79 mm d^{-1} while the smallest amount of 0.67 mm d^{-1} was received in the CS. Also, the NE received considerably high annual rainfall relative to the other regions followed by

the SSE. In contrast, the NW and CS received relatively low annual rainfall. Similar probability patterns in some of the regions were also evident (Figure 2). These probability patterns show the likelihood of obtaining the corresponding 5-day rainfall in that particular year.

The analysis of the probability curve revealed that SSE was anticipated to experience rainfall throughout the year, whereas SW, NSE, CS and CE showed varying probabilities between 0.8-1.0. However, N and NE had a probability between 0.59 and 1.0. The least annual rain probability of 0.59 was observed with high variations throughout the 64 years in NW.

3.2. Seasonal rainfall

The observed seasonal rainfall amount and the probability of occurrence which is the likelihood of non-zero 5-day rainfall on any particular 5-day period of the year in Australia from 1950-2013 are shown in Figure 3. Each bar corresponds to 5-day average in each particular month. It revealed four distinct seasonal rainfall distribution patterns in Australia, but out of the four seasonal rainfall groups, some regions have similar patterns with different seasonal rainfall amounts. Among the regions, the CS and NSE have similar seasonal patterns while the N and NE regions have a similar distribution. Rainfall in these regions is mostly influenced by tropical cyclones that convey masses of convective clouds. Besides, NW and CE have similar distributions, but apparent two rainfall peaks were observed in NW. The apparent peaks in the seasonal rainfall in the NW may be due to summer monsoon and the cold winter fronts pushing northwards. Seasonal rainfall distribution was observed to be similar in the SW and SSE. These patterns had very low variations in the CS and NSE, but the remaining regions experienced high seasonal variations.

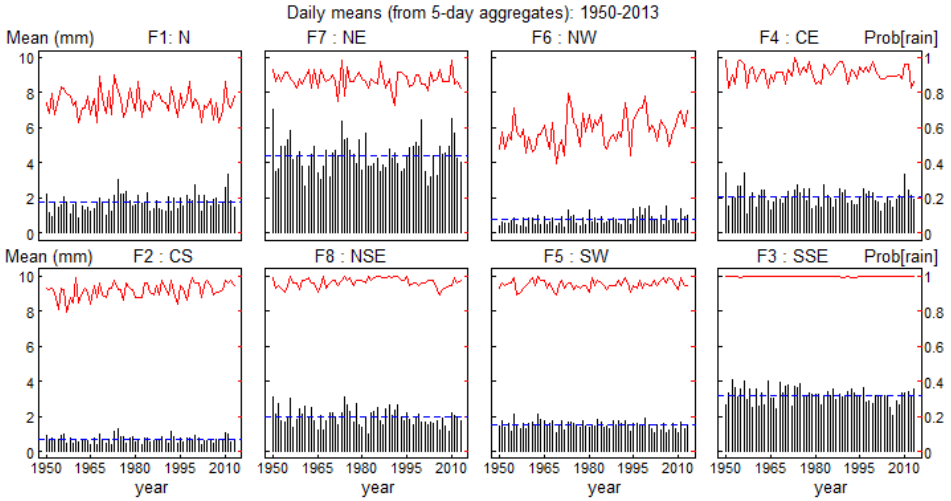

Figure 2. Annual rainfall amount and the probability of occurrences during 1950-2013. The black bars denote the rainfall amount; the horizontal blue dashed lines the overall mean, and the red curve denotes the probability of its occurrence.

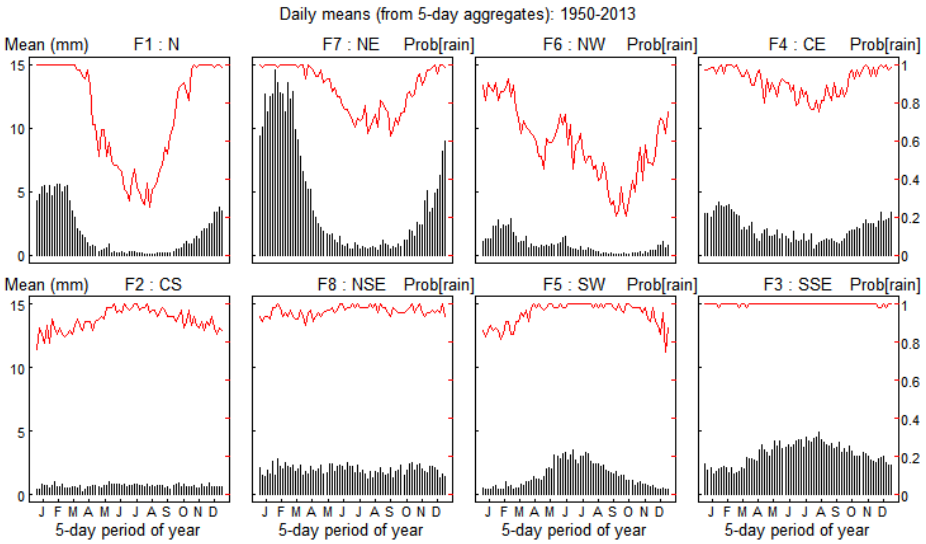

Figure 3. Seasonal rainfall amount and probability of occurrences in Australia during 1950-2013. The black bars denote the rainfall amount while the red curve indicates the probability of its occurrence.

Analysis of the similarities in the seasonal evolutions between regions by correlation analysis revealed three main seasonal rainfall groupings. This result is consistent with the earlier study by Peel, Finlayson, and Mcmahon (2007) which revealed three main climate patterns throughout Australia. The first group consisting of the N, NE, NW and the CE regions which were seen to be strongly positively correlated, have apparent summer monsoon rainfall with dry winters. The second group comprised the CS and the NSE which were also seen to be strongly correlated. These regions observe sporadic rainfall with the apparent seasonal pattern. The final group comprised the SW and SSE regions which have similar seasonal patterns with a correlation of 0.82. On the other hand, the SW and SSE have evident peak rainfall during the winter season. The observed peak in winter may be as a result of climate variability mode which develops in June and peaks around October. The southwest of Australia is predominantly influenced by weather fronts, often related to the Southern Annular Mode (SAM), whereas the southeast is influenced by the SAM and IOD (and associated weather fronts).

Seasonal rainfall probability of 1.0 was observed in the N and NE from December-March. The observed seasonal probability of 1.0 shows that it is certain that it will rain during this period. However, this probability dropped to 0.3 in the N in August and 0.6 in NE during September. In NW, rainfall probability varied from 0.2-0.9.

3.3. The analysis of the fitted gamma GLM models

The cube root transform of the response variable was used in fitting the models for all regions. The predictors were seen to be significant in the patterns in the N, SSE, and CE but the models in SW, NW and NE were significantly affected by the seasonal effect only. In CS and NSE the models were seen to be significantly affected by the year factors.

Figure 4 shows the results of the fitted gamma GLM model of the rainfall of each region. The top panel shows the models in N, NE, NW, and CE while the bottom panel shows the models in CS, NSE, SW, and SSE. The grey points are the observed rainfall, and the black lines denote the fitted gamma

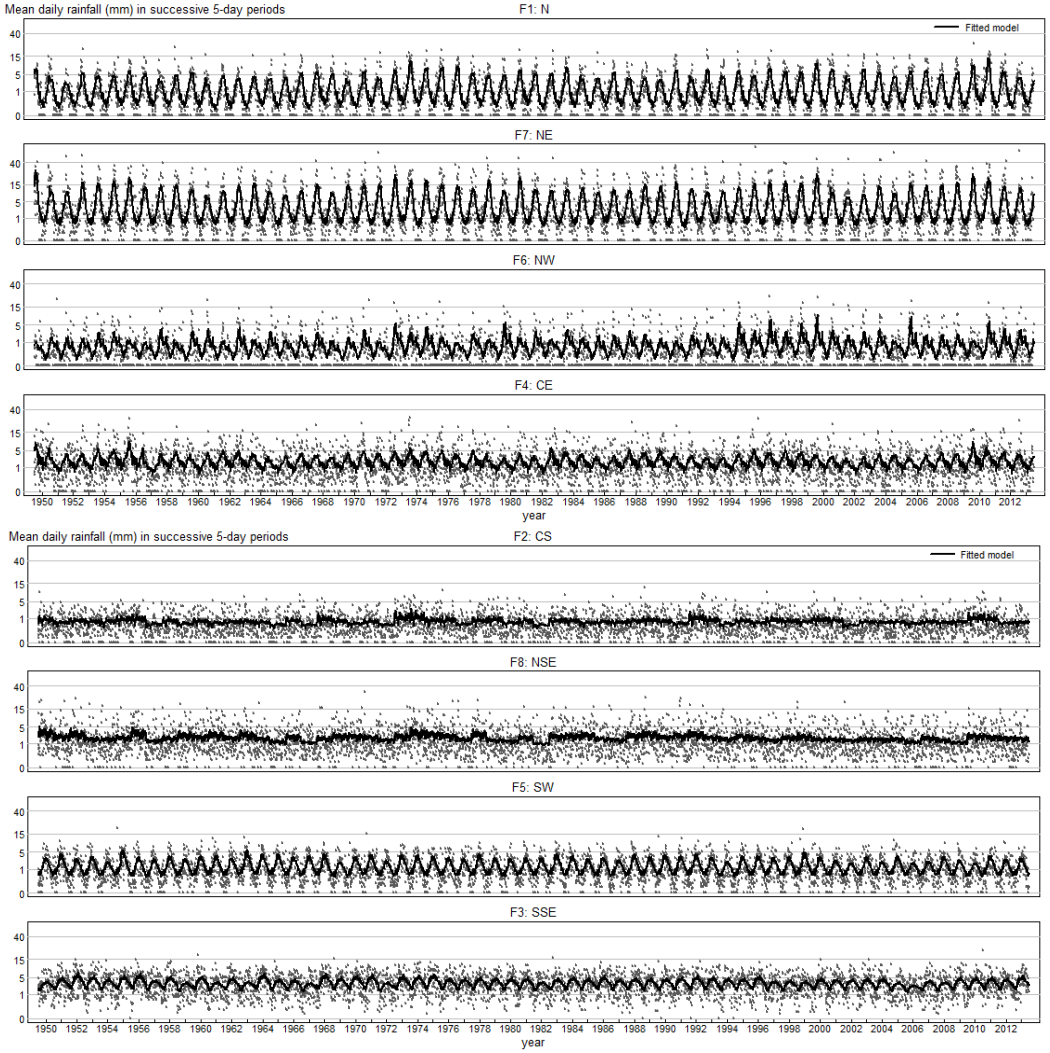

Figure 4. The results of the fitted gamma model for the 5-day rainfall means in the eight factor regions in Australia during 1950-2013. The grey points denote the 5-day rainfall and the black line denotes the fitted 5-day rainfall mean by the GLM.

model. One noticeable feature is the well-marked seasonal periodic pattern. The eight factor regions revealed different rainfall variability throughout the fitted models. The fitted models vary throughout the 64 years in all the regions with minimum variations observed in CS and NSE from 1950-2013 (the CS and NSE experiences uniform rainfall throughout the season). However, high temporal variations were observed in the N and NE. Besides, Figure 4 shows similar patterns of the fitted models in some of the regions. For instance, the mean rainfall patterns in N and NE were similar while that of NW was also similar to CE. Also, not much difference was observed between the patterns in CS and NSE, and the rainfall patterns in SSE and SW were similar.

The fitted gamma models were assessed using the residuals quantile-quantile (Q-Q) plots from the models (Figure 5). The residuals plots had low variations relative to the expected line of best fit that indicated the models fitted the data quite well in most of the factor regions except for little departures at the upper and lower tails of some of the models. These departures at the lower and upper tails of some of the models may be as a result of rainfall extremes or other neglected influential predictors, which indicates that the gamma GLM models may only work well for non-extreme rainfall values. Similar results were seen by Hasan and Dunn (2010) using simple Poisson-gamma modelling of monthly rainfall in Australia and in a similar study by Kenabatho *et al.* (2012) in semi-arid Limpopo Basin in Botswana. However, Kenabatho *et al.* attributed the departures of the models from a normal distribution to the data quality. They established that GLM model is sensitive to data quality of rainfall observations and good quality data can improve the performance of the model. Also, the evaluation and analysis of the deviance residuals versus the fitted values of the models did not display any distinctive patterns and outliers, signifying no indication of poor fit of the models. Interestingly, this study revealed that the GLM model performed better in modelling the observations over a short-term period (considering the first thirty-two or last thirty-two years data separately). The model captured the periodic seasonal patterns quite well as regards the timings of periods. Interestingly, the seasonal variations were observed to be higher in all the factor regions than their annual rainfall variations. The fitted models revealed significant mean rainfall amount in the NE followed by SSE and lowest in CS. Moreover, in N, NE, NW and CE, most of the annual rainfall was observed in the summer months, and these results were seen by Lavender and Abbs (2013) over north of Australia and they attributed it partially to the tropical cyclones and other low pressure systems. Considering all the eight factor regions, N, NE, NW and CE receives a substantial amount of rainfall from December-March (wettest months) while uniform monthly rainfall distributions were observed in CS and NSE. Similar results were seen by Hasan and Dunn (2010) on the analysis of monthly rainfall from 1912-1971 with Poisson and gamma models. On the other hand, SW and SSE have their wettest months between August and September.

3.4. The 95% confidence intervals for the adjusted means

The 95% confidence intervals for the adjusted means delineated the temporal variations of the fitted rainfall means (seasonal and annual) from the overall mean between 1950-2013 in Australia (Figure 6). All the models were within the 95% CI signifying that the models represented quite well all the regions, but with high variations in most parts. Seasonal rainfall increases sharply from January to February where it attains the maximum, decreases to a minimum in August and increases gradually until December in the N and NE.

The NW receives above average rainfall mostly between January and June and experiences below average until December while in the CE below average rainfall mostly occurs between April and October. However, no clear seasonal patterns were observed in the CS and the NSE regions. In the SW and SSE regions, seasonal rainfall increases sharply from February to June and decreases from July to December. These seasonal patterns were significant in all the factor regions ($p < 0.001$) except CS and NSE regions (first and second panels of Figure 6b). Also, annual rainfall has a very low variation concerning the overall rainfall mean in the N, NE, NW and the SW regions ($p < 0.05$ in N, but $p > 0.05$ in the NE, NW and SW regions). Annual rainfall observed between the remaining regions varies considerably from the overall means during 1950-2013.

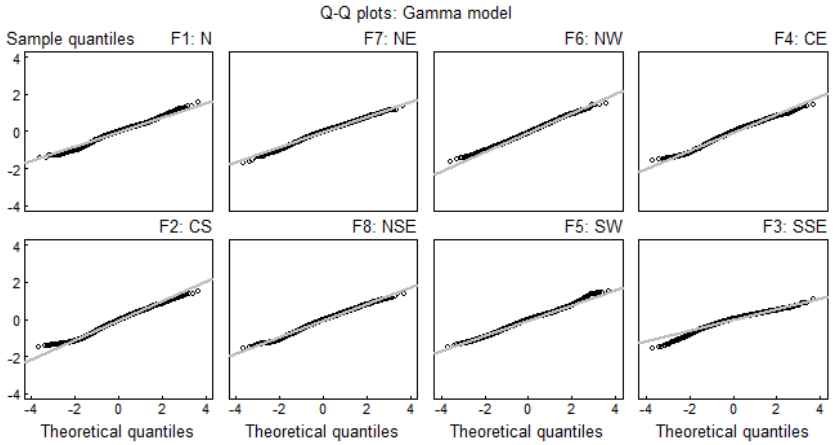

Figure 5. The residuals quantile-quantile (Q-Q) plots based on the gamma GLM for the fitted 5-day rainfall means for the eight regions during 1950-2013. The black lines denote the residuals while the grey line indicates the line of best fit for the correct gamma model.

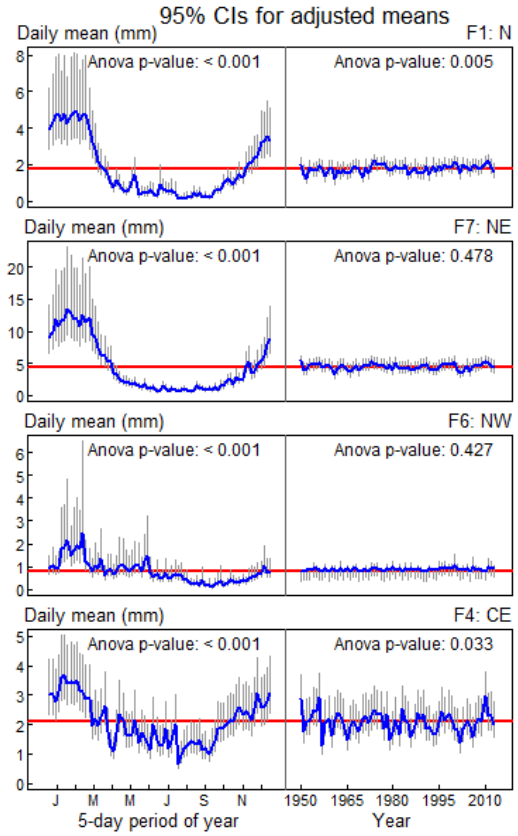

Figure 6a. The 95% confidence interval plots for the adjusted means of the 5-day rainfall from in N, NE, NW, and CE during 1950-2013. The left panels show the 5-day periods effects while the right panels show the annual effects. The horizontal red line indicates the overall mean and the blue curve is the fitted rainfall mean. The grey bars are the 95% CIs.

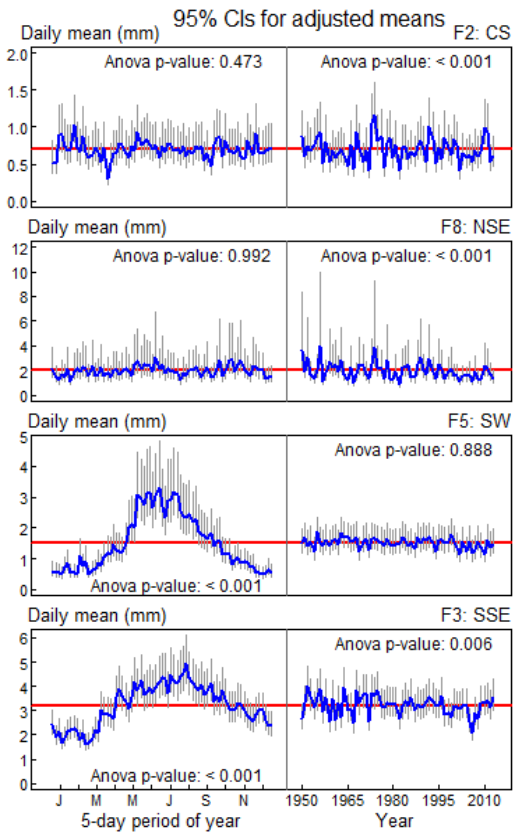

Figure 6b. The 95% confidence interval plots for the adjusted means of the 5-day rainfall in CS, NSE, SW and SSE during 1950-2013. The left panels show the 5-day periods effects while the right panels show the annual effects. The horizontal red line indicates the overall mean and the blue curve is the fitted rainfall mean. The grey bars are the 95% CIs.

5. Conclusions

This study showed how a class of statistical models could be used to analyse daily rainfall to provide essential information on its variability patterns. In the first model, the estimated 5-day average rainfall based on the data in each station over 64 years (1950-2013) collected from 92 observational stations in Australia were grouped into distinct climatic regions by using factor analysis. The result showed that eight factors were adequate and accounted for about 52% of the observed correlations in the data, thus grouping Australia into eight geographical rainfall regions. The model revealed evident spatial rainfall variability during the 64 years. The analysis of similarities in the seasonal evolutions between the eight regions by correlation analysis revealed three seasonal rainfall distributions. These seasonal patterns were observed to be similar in the N, NE, NW and CE regions. The CS and NSE regions were seen to have similar seasonal distributions while that of SW and SSE were also observed to be related.

In general, the observed eight climatic regions have three main seasonal groups, and each has its unique dominating factor. The knowledge of these seasonal groups may be essential in the planning of rainfall resources in Australia.

The gamma GLM was then used to model the non-zero rainfall amount with the seasonal and the annual factors as predictors. The gamma GLM fitted the data quite well in most of the factor regions except little departures at the upper and lower tails of some of the models. The model could be used to simulate rainfall data for the areas with inadequate rainfall records.

Further studies could explore the mixed factors, extreme rainfall values, and examine the relationship between the seasonal rainfall in the factor regions with climate variables such as humidity and temperature, climate indices/tele connection, and model the probability pattern of each region.

Acknowledgements

This work was supported by the Higher Education Research Promotion and Thailand's Education Hub for Southern Region of ASEAN Countries Project Office of the Higher Education Commission under grant no. THE/034. Thanks to Centre of Excellence in Mathematics, Commission on Higher Education, Thailand. We are also grateful to BOM for the provision of the daily accumulated data. Our tremendous thanks go to Emeritus Prof. Dr. Don McNeil for his comments and support during this study.

References

Bagirov, A. M., Mahmood, A., & Barton, A. (2017). Prediction of monthly rainfall in Victoria, Australia: Clusterwise linear regression approach. *Atmospheric Research*, *188*, 20-29.

Bärring, L. (1987). Spatial patterns of daily rainfall in central Kenya: Application of Principal Component analysis, common factor analysis and spatial correlation. *International Journal of Climatology*, *7*, 267-289. doi:10.1002/jjoc.3370070306.

Bärring, L. (1988). Regionalization of daily rainfall in Kenya by means of common factor analysis. *International Journal of Climatology*, *8*, 371-389. doi:10.1002/ joc.3370080405.

Beecham, S., Rashid, M., & Chowdhury, R. K. (2014). Statistical downscaling of multi-site daily rainfall in a South Australian catchment using a Generalised Linear Model. *International Journal of Climatology*, *34*, 3654-3670. doi:10.1002/joc.3933.

Bukantis, A. (2002). Application of factor analysis for quantification of climate-forming processes in the eastern part of the Baltic Sea region. *Climate Research*, *20*, 135-140. doi:10.3354/cr020135.

Cai, W., Cowan, T., & Thatcher, M. (2012). Rainfall reductions over Southern Hemisphere semi-arid regions: the role of subtropical dry zone expansion. *Scientific Reports*, *2*, 1-5. doi:10.1038/srep00702.

Chambers, L. E. (2003). *South Australian rainfall variability and trends* (BMRC Research Report No.92). Melbourne, Australia: Commonwealth of Australia, Bureau of Meteorology.

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. doi:10.1016/j.solener. 2014.10.014.

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

Deo, R., Shahin, M., (2015). Application of the artificial neural network model for prediction of monthly standardized precipitation and evapotranspiration index using hydrometeorological parameters and climate indices in Eastern Australia. *Atmospheric Research*, *161*, 65-81.

Drosdowsky, W. (1993). An analysis of Australian seasonal rainfall anomalies: 1950-1987. II: temporal variability and teleconnection patterns. *International Journal of Climatology*, *13*, 111-149. doi:10.1002/ joc.3370130102.

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1998). *Multivariate data analysis*. London, England: Prentice Hall International.

Hasan, M. M., & Dunn, P. K. (2010). A simple Poisson-gamma model for modelling rainfall occurrence and amount simultaneously. *Agricultural and Forest Meteorology*, *150*, 1319-1330. doi:10.1016/j.agrfor met.2010.06.002.

Herring, S. C., Hoell, A., Hoerling, M. P., Kossin, J. P., Schreck, C. J III, & Stott, P. A. (2016). Explaining extreme events of 2015 from a climate perspective. *Bulletin of the American Meteorological Society*, *97*(12), S1-S145.

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*, 1113-1127. doi:10.1002/joc.2323.

Köppen, W. P. (1918). Klassification der Klimate nach Temperatur, Niederschlag und Jahreslauf. *Petermmans*, *64*, 243-248.

Lavender, S. L., & Abbs, D. J. (2013). Trends in Australian rainfall: Contribution of tropical cyclones and closed lows. *Climate Dynamics*, *40*, 317-326. doi: 10.1007/s00382-012-1566-y.

Lin, G.-F., & Jhong, B.-C. (2015). A real-time forecasting model for the spatial distribution of typhoon rainfall. *Journal of Hydrology*, *521*, 302-313.

Peel, M. C., Finlayson, B. L., & Mcmahon T. A. (2007). Updated world map of the K¨oppen-Geiger climate classification. *Hydrology and Earth System Sciences Discussions*, *European Geosciences Union*, *11*(5), 1633-1644.

R Development Core Team (2015). R: A language and environment for statistical computing (Version 3.2.3). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from *http://www.R-project.org/*.

Rencher, A. C. (2002). *Methods of multivariate analysis* (2nd ed.). New York, NY: John Wiley and Sons.

Shukla, R., Tripathi, K., Pandey, A., & Das, I. (2011). Prediction of Indian summer monsoon rainfall using Nino indices: a neural network approach. *Atmospheric Research*, *102*(1-2), 99-109.

Srikanthan, R., & Pegram, G. G. S. (2009). A nested multi-site daily rainfall stochastic generation model. *Journal of Hydrology*, *371*, 142-153. doi:10.1016/ j.jhydrol. 2009.03.025.

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

Taschetto, A. S., & England, M. H. (2009). An analysis of late twentieth century trends in Australian rainfall. *International Journal of Climatology*, *29*, 791-807. doi:10.1002/joc.1736.

Tongkumchum, P., & McNeil, D. (2009). Confidence intervals using contrasts for regression model. *Songklanakarin Journal of Science and Technology*, *31*, 151-156.

Toth, E., Brath, A., & Montanari, A. (2000). Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology*, *239*(1), 132-147.

Trenberth, K. E., Dai, A., Rasmussen, R. M., & Parsons, D. B. (2003). The changing character of precipitation. *Bulletin of the American Meteorological Society*, *84*(9), 1205-1217. doi:10.1175/BAMS-84-9-1205

Um, M. -J., Yun, H., Jeong, C.-S., & Heo, J. -H. (2011). Factor analysis and multiple regressions between topography and precipitation on Jeju Island, Korea. *Journal of Hydrology*, *410*, 189-203. doi:10.1016/j.jhydrol.2011.09.016.

Unkel, S., Trendafilov, N. T., Hannachi, A., & Jolliffe, I. T. (2010). Independent exploratory factor analysis with application to atmospheric science data. *Journal of Applied Statistics*, *37*, 1847-1862. doi:10.1080/02664760903166280.

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

Wanishsakpong, W., & McNeil, N. (2016). Modelling of daily maximum temperatures over Australia from 1970 to 2012. *Meteorological Applications*, *23*(1), 115-122. doi:10.1002/met.1536

Wickramagamage, P. (2010). Seasonality and spatial pattern of rainfall of Sri Lanka: Exploratory factor analysis. *International Journal of Climatology*, *30*, 1235-1245. doi:10.1002/joc.1977.