Drought Predictions: Applications in Australia

Geraldine H. Wong

Thesis submitted for the degree of

Doctor of Philosophy

in

Statistics

at

The University of Adelaide

Discipline of Statistics, School of Mathematical Sciences,

Faculty of Engineering, Computer and Mathematical Sciences

January 9, 2010

Chapter 7

Predicting low rainfall for pixels throughout Australia with climatic indicators

The previous chapter demonstrated the performance of stochastic modelling in predicting short and longer term monthly rainfall at specific rainfall gauges. This chapter examines the Rainman software, which is developed by Stone [119] and used by many farmers, businesses and governmental agencies for example the Australian Bureau of Meteorology, to carry out rainfall forecasting. This software analyzes records of historical monthly and daily rainfall dating back to 1832, for monthly and daily patterns and forecast seasonal rainfall based on the SOI.

This chapter investigates the performance of this software across Australia, by applying significant tests on the June to October rainfall probability distributions associated with both the extreme phases defined by Stone. June to October is traditionally the growing period in Australia and is especially crucial for agriculture. Another forecasting strategy which only depend on the SOI of May is also compared with the classification that Rainman uses, through statistical tests. A third method where the association between the rainfall distribution and the ENSO classification of the current water year, April to March, is also examined here. This does not, in itself, provide a prediction, but it could be incorporated with physically based climate models, which are capable of predicting ENSO states.

The latter section of this chapter examines the correlations between the climatic index,

sea-surface temperatures, with rainfall in selected regions of Australia. Appropriate multiple regression models are fitted using the highest correlated sea-surface temperature pixels. The predicted rainfall obtained from the regression model are compared to the observed rainfall and recommendations are proposed.

7.1 Study area and data

Monthly rainfall grid data from 1900 to 2003 across Australia are obtained from BoM. BoM has recognized that rainfall variability in the country has a large impact on the economy. This has led to the revision of how rainfall data is recorded in Australia. Previously, Australia was divided into rainfall districts, where the size of each district varied with station density and population. Recall in Chapter 4, district rainfall data was employed in copula modelling of drought.

In order to investigate spatial rainfall patterns, gridded data sets were generated from historical rainfall records from rainfall stations, using the Barnes two-dimensional analysis [64]. Each grid is a regular 0.25° (approximately 25km) by 0.25° latitude-longitude with a latitudinal range from 10 to 44.5°S, and a longitudinal range from 112 to 156°E [58]. There are 178 intervals in the longitudinal and 139 intervals in the latitudinal directions, which is stored in an array of 24,742 pixels, of which 11,225 are over land.

Annual mean precipitation across Australia is shown in Figure 7.1 for 1900 to 2003. The maximum mean annual rainfall recorded is 3389mm in north-east Queensland, while the minimum recorded is in central Australia with 128mm. Most parts of eastern Australia receive more rainfall than the western and southern parts. This difference in rainfall can be attributed to physical landforms such as the Great Dividing Range, which extends along the eastern coastline from Queensland through NSW to Victoria, and ocean currents and atmospheric circulation such as the SO and SST.

The SOI data used here is obtained from the Long Paddock website [94]. The SOI given in this website is calculated using a climatic base period of 1887 to 1989, in contrast to the National Climate Centre (NCC), which calculates the SOI using the climatic base period of 1933 to 1992. The monthly SOI data provided by the BoM, which was used in Chapter 6, is calculated using this base period. The time series plot of the SOI data from 1900 to 2003 from the Long Paddock website is shown in Figure 7.2. There is no apparent trend in the SOI time series. The autocorrelogram of the SOI from 1900

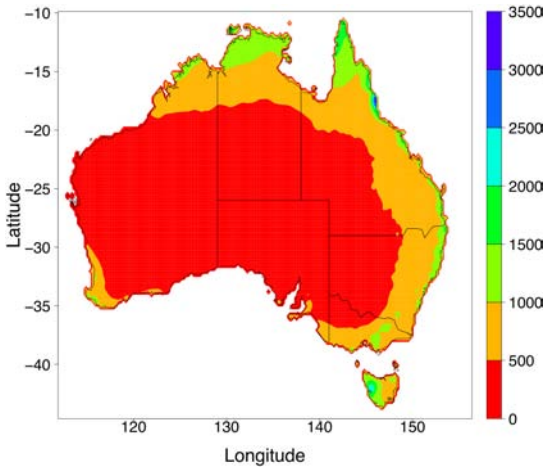


Figure 7.1: Annual mean precipitation across Australia, 1900-2003

to 2003 in Figure 7.3 shows that there is slow decay in the autocorrelation of the SOI, indicating there is substantial correlation between lagged values of the SOI.

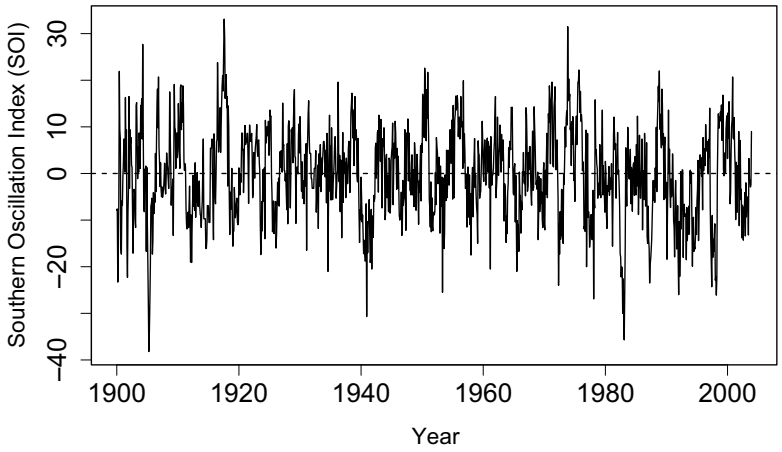


Figure 7.2: Time series plot of SOI from Long Paddock website, 1900-2003

The Earth System Research Laboratory, run by the U.S. National Oceanic and Atmospheric Administration (NOAA), maintains an online reconstructed SST data set, using the International Comprehensive Ocean-Atmosphere Data Set (ICOADS) SST data. The worldwide monthly SST anomalies from 1854 to 2003 are recorded in an array, with 180 intervals in the longitudinal direction and 89 intervals in the latitudinal direction. Each grid represents 2.0° latitude by 2.0° longitude across the entire globe which spans from 88.0°N to 88.0°S in latitude direction and 0.0°E to 358.0°E in the longitudinal direction. The mean monthly SST for every grid from 1900 to 2003 is calculated and Figures 7.4 shows the SST during January and July respectively. The

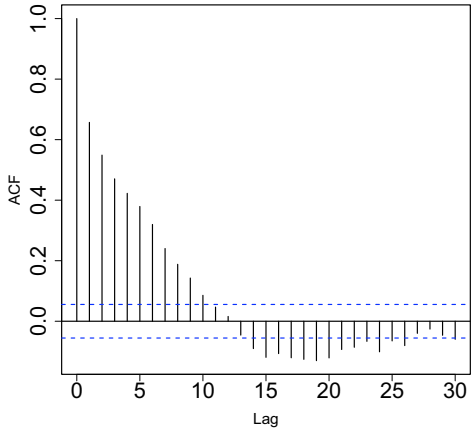


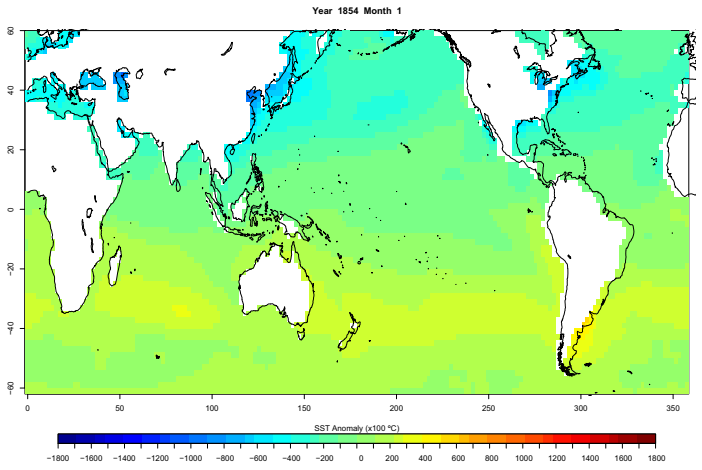
Figure 7.3: Autocorrelogram plot of SOI, 1900-2003

SST changes are more pronounced in the extreme hemispheres and coincides with the summer and winter seasons.

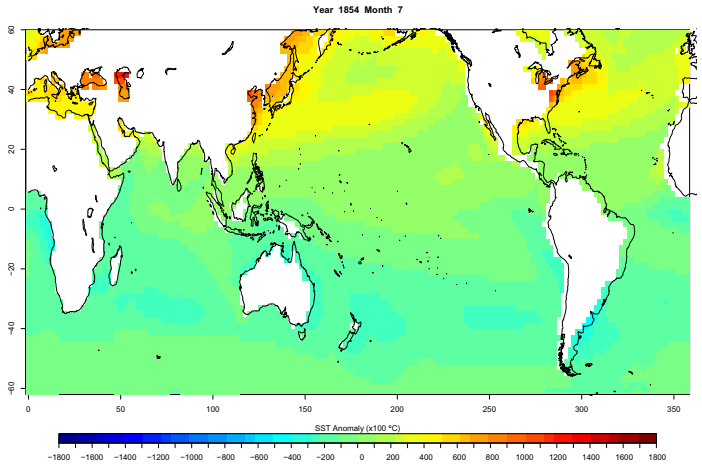
7.2 SOI phases

In Chapter 3, research by Stone *et. al.* [119] discussed the potential of predicting the amount of rainfall through the lag-relationship between rainfall and SOI patterns. This method is recapitulated here as the following sections compares the effectiveness of this method with other proposed techniques.

Stone [118] analyzed the SOI data from 1882 to 1989 in his dissertation and found an association between changes in the SOI April and May with June to October rainfall distributions. Based on these SOI values and their one-month lag relationships, five phases were established and this classification is shown in Figure 7.5. Rainfall from the subsequent five months were tabulated and their probability distributions corresponding to each of the SOI phases were produced. Results show a causal relationship between the SOI phase and rainfall probability distribution. Cumulative probability distributions of subsequent rainfall associated with the SOI phases from a number of locations show that the probability of obtaining rainfall above the median rainfall is higher following a 'rapid rise' SOI phase than a 'rapid fall' phase. Figure 7.6 illustrates this using the example of Goondiwindi in Eastern Australia. Goondiwindi has latitude 28.3°S and longitude of 150.2°E and is located near the NSW border, 350km south west of Brisbane. The short-dashed line denotes the rainfall distribution corresponding to the SOI phase 'rapid fall' in April and May, while the long-dashed line shows the



(a)



(b)

Figure 7.4: Mean monthly sea-surface temperatures for (a) January and (b) July,1990 - 1999

distribution following a 'rapid rise' phase. The solid line represents the 'all-years' distribution. According to this diagram, the probability of obtaining more than 200mm of rainfall is 25% following a 'rapid fall' phase while the probability of obtaining the same amount of rainfall is 70% following a 'rapid rise' phase.

In the next section, non-parametric statistical tests, as described in Appendix E, are introduced to determine if the rainfall probability distributions obtained from the extreme SOI phases are significantly different from each other at a given significance level. These tests are performed for every grid in Australia and the number of grids that are significant are counted. The effectiveness of the method is measured based on the number of grids across Australia that show a significant difference between the

'rapid rise' and 'rapid fall' phase.

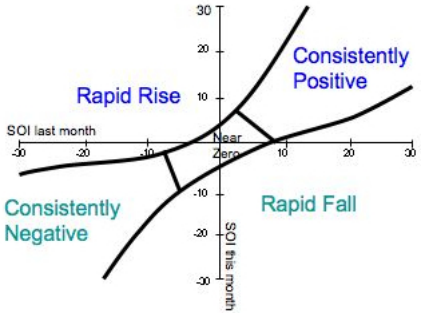


Figure 7.5: Classification of SOI values into five SOI phases, based on current and previous month.

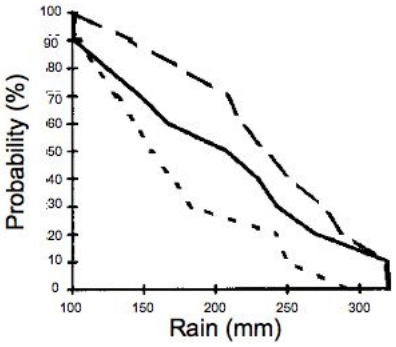


Figure 7.6: Cumulative probability distribution of rainfall for June-October at Goondiwindi, for rapid rise (long dash), rapid fall(short-dash) and overall (solid)

7.3 Predicting drought using SOI

This section introduces three methods of categorizing SOI. The rainfall probability distributions corresponding to the respective SOI categories are obtained and statistical significance testing on these distributions are performed to evaluate the prediction skill of each of the three classification methods.

7.3.1 Categorization using SOI phases

The SOI values from April and May of each year starting from 1900 to 2003 are being categorized into the following five phases as shown in Figure 7.5: consistently negative, consistently positive, rapid fall, rapid rise and near zero. The April and May SOI values were used by Stone *et. al.* [119] for rainfall prediction, as the SOI of these months were shown to be highly associated with rainfall.

Table 7.1 shows the distribution of the years corresponding to the SOI phases. The phases are then associated to the subsequent grid rainfall from June to October of the same year. Significance testing at the 5% level is carried out between the rainfall associated to the extreme phases of rapid rise and rapid fall. These statistical significance test are described in detail in Appendix E. These tests are performed on this set of data so as to be consistent with the study done by Stone *et. al.* [119]. Figure 7.7 displays the plot of Australia, with significance between the extreme phases shaded in black and non-significance shaded in grey.

Table 7.1: Number of observations corresponding to each phase

SOI phase	Number of years in each phase
Consistently negative	16
Consistently positive	22
Rapid fall	14
Rapid rise	26
Near zero	26

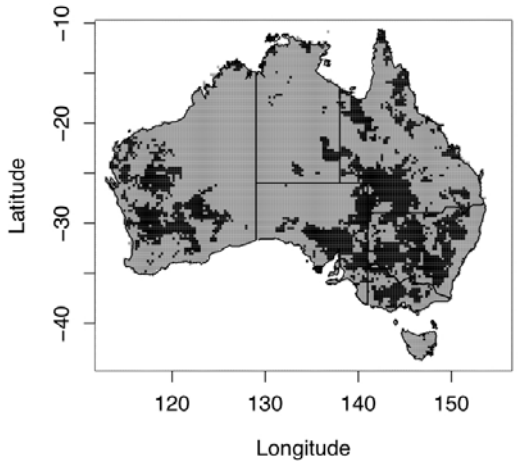


Figure 7.7: Statistically significant (5% level) differences in June-October rainfall associated with 'rapid rise' SOI phase and 'rapid fall' phase shown in black.

Evidently, there is an observable significant difference on eastern Australia which gives rise to a clustering effect in that region. There is also a noticeable clustering in the western to southwest of Australia. A striking feature in Figures 7.7 and 7.8 is the relatively slight association between the April and May SOI phases and June to October rainfall, to the east of the Great Divide. It may be that the SOI has more effect on summer coastal rainfall, which is increasingly higher than the winter rainfall as you

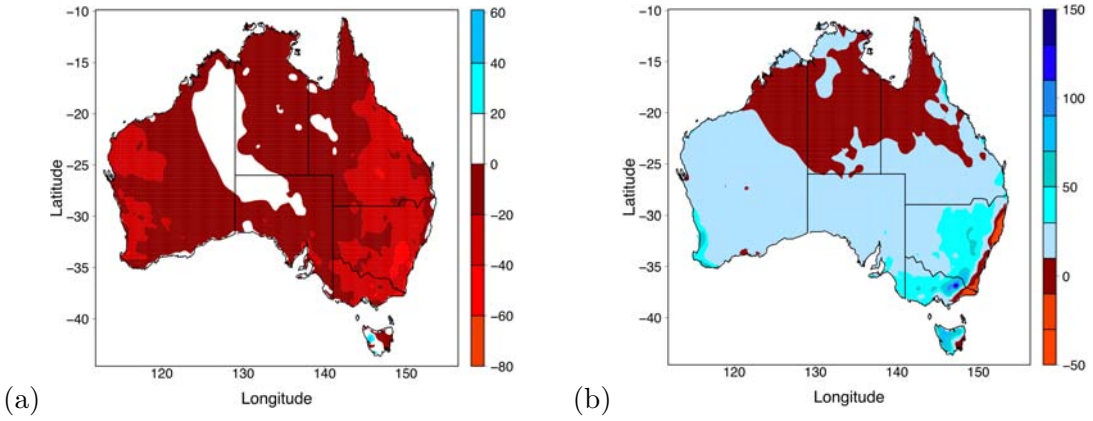


Figure 7.8: Difference in rainfall means (mm) between: (a) Rapid fall SOI phase and overall distribution; (b) Rapid rise SOI phase and overall distribution

move north. Given the evidence of a difference in rainfall between rapid rise and rapid fall phases at many grid points, the increase or decrease relative to overall mean is shown in absolute terms in Figures 7.8(a) and (b) and as a proportion in Figures 7.9(a) and (b).

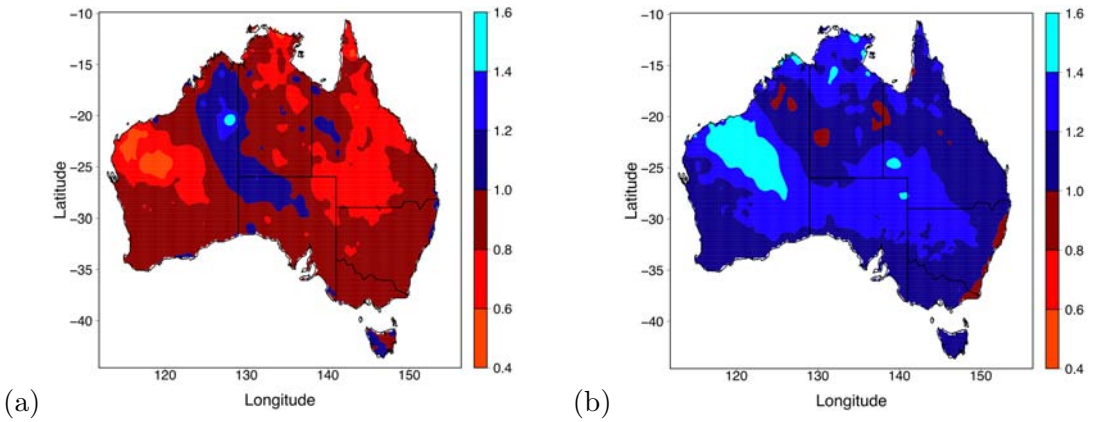


Figure 7.9: Ratio of mean rainfall between: (a) Rapid fall SOI phase and overall distribution; (b) Rapid rise SOI phase and overall distribution

Figure 7.8(a) shows that in most parts of Australia, there is a decrease in rainfall associated with the rapid fall phase. There is only a slight increase in rainfall in central and parts of northwest Australia. For the rapid rise phase, there is a general increase in rainfall over Australia, with the exception of a few grid squares over the east coast.

Some general observations can be made from Figures 7.9(a) and (b). It appears that

the rapid fall phase has somewhat more influence on June to October rainfall than the rapid rise phase. Also, the rapid fall phase is generally associated with a decrease in June to October rainfall whereas the rapid rise is associated with an increase.

The rapid fall and rapid rise phases of the April and May SOI categorisation were chosen for the above comparisons because they had the greatest effect on June to October rainfall. This approach is justified because there is evidence, from the Kruskal-Wallis test, that the general categorisation into five phases does distinguish different rainfall regimes. Grid squares for which the Kruskal-Wallis test is statistically significant at the 5% level are shown in black in Figure 7.10. There is extensive amount of clustering over eastern Australia and scattered clusters over western and northern Australia.

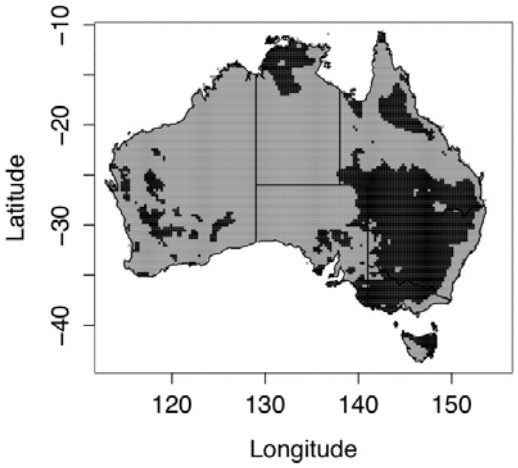


Figure 7.10: Kruskal-Wallis test of significance at 5% level of June-October rainfall with statistically significant pixels in black.

There is a technical issue concerning multiple comparisons. As there are 11,225 pixels, there are 11,225 hypotheses tests involved in this plot. If these tests were independent and performed at the 5% level, the number of statistically significant pixels X, would be distributed as binomial with 11,225 trials and probability of success 0.05. We would expect 5% to be statistically significant and X would have a mean of 561.25 and a standard deviation of 23.09. But, given the spatial correlation of rainfalls, tests for neighbouring pixels are not independent, and this has the effect of increasing the standard deviation of X, although the mean remains the same. An informal explanation, for the increase in the standard deviation of X, is that if statistical tests were confined to a subset of grid squares, that were sufficiently separated to be approximately independent, there would be considerably fewer of the tests. However, the number of statistically significant pixels is 3005 (27%) and this is substantially

greater than the expected number of 561.25. An overall significance level could be obtained by simulation.

7.3.2 Categorization using SOI May

In this section, this method of classification is applied to the SOI May values of each year, since the rainfall period between June to October is being analyzed. The classification differs from the study by Jin *et. al.* [56] due to the distribution of the SOI May values. Studies done in Japan also showed that there is statistically significant correlation between SOI and precipitation in the subsequent months, using categorization method of the SOI [60].

For consistent comparisons to be made with the SOI phases, the SOI May values is being categorized into five groups: Strong El-Niño (SOI < -8.6); Weak El-Niño ($-8.6 \leq$ SOI < -1.9); Normal Condition ($-1.9 \leq$ SOI ≤ 1.9); Weak La-Niña ($1.9 <$ SOI ≤ 8.6); Strong La-Niña (SOI > 8.6). The subsequent June to October rainfall totals are then categorized accordingly. Table 7.2 displays the number of observations in each category. As a parallel comparison to the SOI phases in 7.3.1, there are 19 observations corresponding to the Strong El-Niño category and 16 observations in the Strong La-Niña category, as compared to 14 in the rapid fall phase and 26 observations in the rapid rise phase.

Table 7.2: Number of observations corresponding to each ENSO state

SOI May category	Number of years in each category
Strong El-Niño	19
Weak El-Niño	19
Normal Condition	19
Weak La-Niña	31
Strong La-Niña	16

Figure 7.11 displays the results from K-S test between rainfall associated to SOI May categorized as Strong El-Niño and rainfall associated to SOI May categorized as Strong La-Niña year. There is a considerable amount of clustering in eastern Australia, consistent with results in Section 7.3.3. However, the significant pixels on the southwest of Australia are more scattered.

Comparisons between the rainfall probability distributions associated to SOI May value

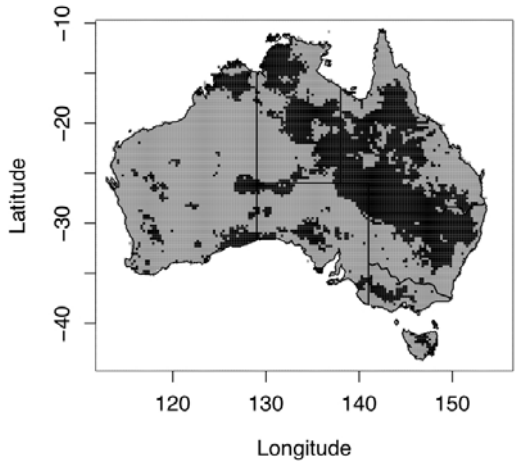


Figure 7.11: Statistically significant (5% level) differences in June-October rainfall associated with SOI May categorized as Strong El-Niño and Strong La-Niña shown in black.

categorized as Strong El-Niño and 'all-years' distribution are made. The same is carried out with the Strong La-Niña category. Figures 7.12(a) and (b) show the difference in absolute means between rainfall associated with Strong El-Niño and Strong La-Niña and overall average rainfall respectively. Figure 7.12(a) shows that in most parts of Australia, the rainfall associated with Strong El-Niño is around 0-60mm lower than the overall average rainfall. In scattered parts of southern Tasmania, the rainfall associated to the Strong El-Niño category is higher than the overall mean rainfall.

When mean rainfall associated with Strong La-Niña category is compared to overall average rainfall, it is generally higher over the country, with southern parts of Tasmania again receiving a higher amount of rainfall than other parts of Australia. These observations are in agreement with Figures 7.13(a) and (b), where the proportion of mean rainfall during the Strong El-Niño years is mainly below 1, and rainfall during the Strong La-Niña years is mostly above 1 for the whole country.

7.3.3 Categorization based on ENSO classification

The method of ENSO classification used by Ropelewski and Halpert [103], as discussed in Chapter 2, was applied to the monthly SOI values and the years 1900 to 2003 are classified into El-Niño, La-Niña or Neutral. Rainfall totals from June to October of the same year are calculated and categorized according to which ENSO state that year is classified into. Table 7.3 displays the distribution of years according to the ENSO state. Similar significance testing at the 5% level is performed between rainfall distributions

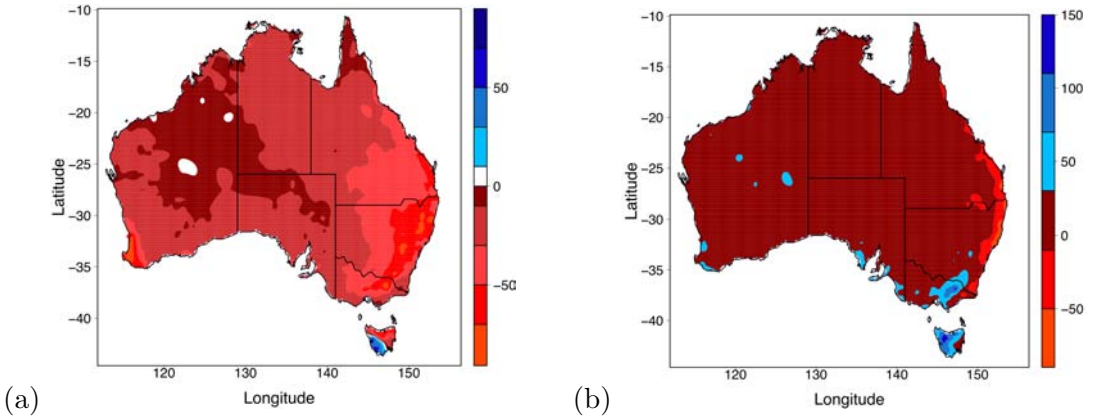


Figure 7.12: Difference in means (mm) between: (a) Strong El-Niño and overall distribution; (b) Strong La-Niña and overall distribution, associated with SOI May categorization.

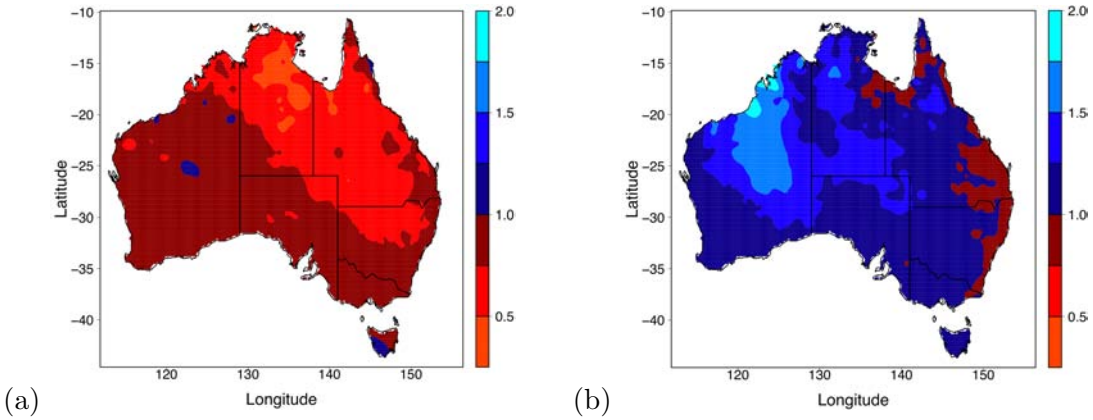


Figure 7.13: Ratio of mean rainfall between: (a) Strong El-Niño and overall distribution; (b) Strong La-Niña and overall distribution, associated with SOI May categorization.

associated with El-Niño and La-Niña years. Figure 7.14 shows the plot of Australia with significant p-values shaded in black.

This approach of classification sees a substantial increase in the number of pixels that are statistically significant. This is consistent with the claim that El-Niño events are often associated to lower than usual rainfall while La-Niña brings about a higher amount of rainfall to most parts of Australia.

Although this form of classification gives a higher number of pixels as compared to SOI phase categorization, this categorization is not immediately useful for forecasting, since monthly SOI values for the entire year from April to March have to be known before

Table 7.3: Number of observations corresponding to each ENSO state

ENSO state	Number of years in each category
El-Niño	32
La-Niña	29
Neutral	43

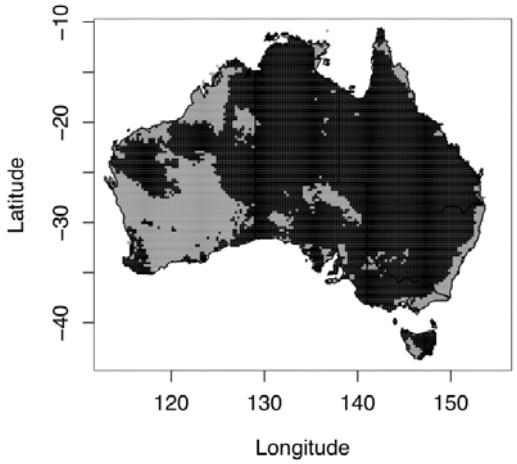


Figure 7.14: Test of significance at 5% level between June-October rainfall associated with El-Niño, and La-Niña with statistically significant pixels in black.

the association of June to October rainfall can be made. However, recent research claims to have developed physically based climate models that can predict ENSO state up to 2 years [21]. This model-based prediction of El-Niño depends largely on initial conditions such as SST, sea level and winds. Chen *et. al.* [21] demonstrate that the model is able to predict most of the warm and cold events in the past 148 years, although the model's prediction accuracy is higher when predicting stronger El-Niño and La-Niña events.

Nevertheless, this model would allow the ENSO state to be predicted for the following year and rainfall probability distributions can be obtained based on the ENSO state. Given that there is a significant difference in probability rainfall distributions between El-Niño and La-Niña events in most parts of Australia, this form of categorization has the potential to provide farmers a forecast of one year ahead based on the probability of obtaining a certain amount of rainfall given the ENSO state.

7.4 Summary of results

In the first half of this chapter, the two prediction methods, SOI phases and SOI May categories are both limited to years in which the phase is rapid rise or rapid fall or the SOI May category is Strong El-Niño or Strong La-Niña. There is little evidence of any substantial association between rainfall and intermediate phases or categories. So predictions would only be made with credibility during 40, for phases, and 35, for categories, years out of 104 years. In this respect, the SOI phases are slightly more effective. However, in terms of percentage of statistically significant pixels, the SOI phases are significant over a smaller area of Australia (26.8%) compared with (33.6%) for categories (Table 7.4). Figure 7.15 shows the density plot of the rainfall distributions associated with rapid rise and rapid fall from a pixel in Queensland that was found to have significantly different distributions for both phases.

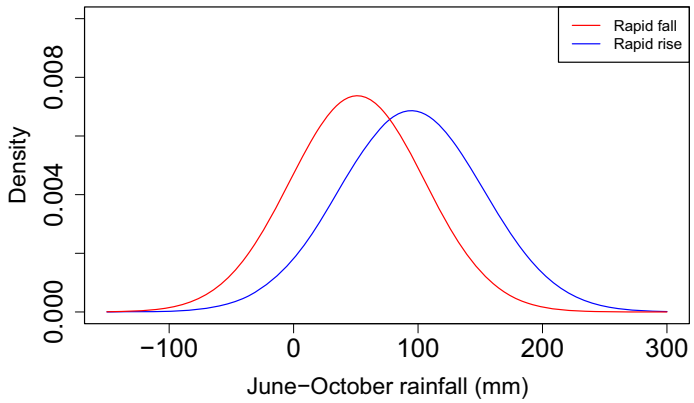


Figure 7.15: Density plot of rainfall distributions associated with rapid rise (blue) and rapid fall (red) phases

If the two measures are combined, 26.8% of 40 is 10.8 which is slightly less than 33.6% of 35 which is 11.7. There is no evidence that the SOI phases are better than the SOI May categorisation. The latter categorisation has advantages of simplicity and of being on an ordinal scale (Table 7.2). The El-Niño and La-Niña year classification is not directly comparable because it cannot be used for predictions unless combined with physically based climate models. This strategy remains to be investigated.

Eastern Australia and western to southwest Australia shows a higher number of significant pixels than the rest of Australia. Hence, rainfall forecasts can be made based on the associations to the phases or the ENSO states. This could eventually lead to more reliable forecasts of droughts, which could be useful for farmers by providing them an

estimate of the increase or decrease in the amount of rainfall they should expect given a certain SOI phase or SOI May category. This finding is consistent with the fact that drought usually develops during an El-Niño, especially in the eastern two-thirds of the continent due to the Walker circulation shifting eastwards [135].

Table 7.4: Summary table for significance testing between SOI phases

ENSO state	Number of years	Percentage of statistically significant pixels
SOI April and May categorized as rapid rise vs. SOI April and May categorized as rapid fall	26 vs. 14	26.8%
El-Niño years vs. La-Niña years	32 vs. 29	75.6%
SOI May categorized as Strong El-Niño vs. SOI May categorized as Strong La-Niña	19 vs. 16	33.6%

7.5 Predicting drought using SST anomalies

7.5.1 Model and Application

In Chapter 2 and 3, SST anomalies is considered an influential climatic indicator of the ENSO phenomenon, which affects the occurrence of droughts. The relationship between all of the 16020 SST pixels and the monthly rainfall of a selected pixel in Australia is investigated in this chapter. For illustration, monthly rainfall data from Warragamba pixel is used. Warragamba is located in the state of NSW, with longitude 156°E and latitude 33.8°S. This region is also home to the Warragamba Dam, which is the primary water supply for Sydney, with a catchment of 9050 square kilometers. Figure 7.16 shows the location of the Warragamba Dam, in relation to Sydney.

To provide a spatial perspective of the correlations between the SST and the rainfall of this region, Figure 7.17 shows the location of Warragamba and the location of the five highest correlated pixel relating to each month's rainfall. These figures illustrates the changes in spatial correlations with SST across the year. To avoid selecting neighbouring pixels of the strongest correlated ones, a mask boundary has been constructed to identify the next highest correlated pixel which is not located in close proximity to previously chosen pixels.

NOTE:
This figure is included on page 118 of the print copy of
the thesis held in the University of Adelaide Library.

Figure 7.16: Location of Warragamba Dam [123]

In general, there is a noticeable change in the area where the strongest correlated pixels are found for every month. In January, highly positive correlated SST pixels are located close to the eastern and western coast of Australia, and most of the southern hemisphere. This gradually develops to a more negatively correlated area in July. Given the monthly changes in correlation, the highest correlated SST pixel relating to Warragamba rainfall is different for each month.

To account for temporal correlations for each month's rainfall, the correlations are calculated for different lag SST months. For instance, the correlations between March rainfall and the SST prior to March are calculated. For each lag, the five strongest correlated SST pixel relating to Warragamba's rainfall is selected, with a similar mask boundary attached. Based on the selected SST pixels, the following regression models with and without interactions are developed.

$$\begin{aligned} Y_t = \beta_0 + \beta_1 SST1_{t-k} + \beta_2 SST2_{t-k} + \beta_3 SST3_{t-k} + \beta_4 SST4_{t-k} \\ + \beta_5 SST5_{t-k} + e_t \end{aligned} \tag{7.1}$$

$$Y_t = SST1_{t-k} \times SST2_{t-k} \times SST3_{t-k} \times SST4_{t-k} \times SST5_{t-k} + e_t \tag{7.2}$$

where Y_t is the monthly Warragamba rainfall in month t, and $SST1_{t-k}$ is the highest

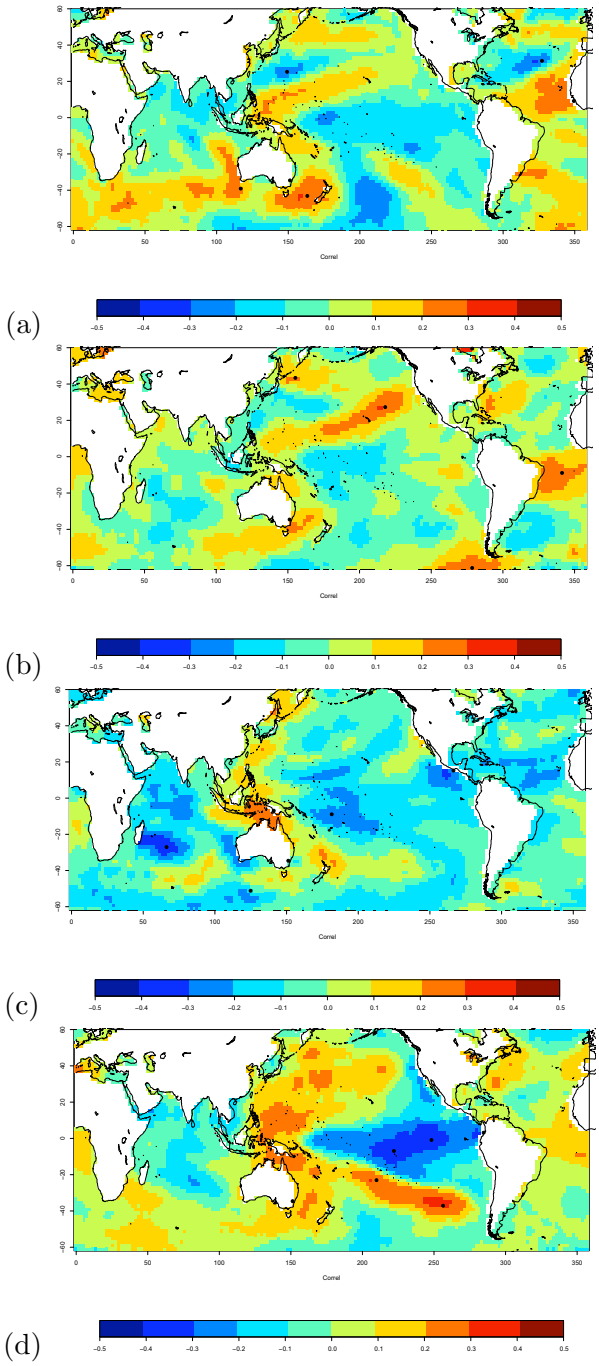


Figure 7.17: Correlation plots for (a) January rainfall and SST; (b) April rainfall and SST; (c) July rainfall and SST and (d) November rainfall and SST,1990 - 2003

correlated SST pixel in month $t-k$ where k represents the lag. $SST2$ represents the next strongest correlated SST pixel and so on. The residuals of the model are represented as e.

7.5.2 Coefficient of determination

The coefficient of determination, R^2 is commonly used statistic output given by most statistical software, as a measure of strength of a relationship fitted by a regression model. It provides an indication of how well future outcomes are likely to be predicted by this model

Usually, the adjusted R^2 function is also provided and is a modification of the R^2 which penalizes R^2 as more predictors are included into the model:

$$\begin{aligned} R^2_{adj} &= 1 - \frac{\sum_{i=1}^{n} r_i^2/(n-p-1)}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2/(n-1)} \\ &= 1 - \frac{s_e^2}{s_Y^2} \end{aligned}$$

where n is the sample size and p is the number of predictors in the regression model, while s_e^2 and s_Y^2 indicates the estimates of the variances of the errors and of the observations, respectively. With large data sets, relatively few predictions and convincing fits, the distribution may not be important. However, in the context of rainfall prediction using SST, the distinction is crucial. Relatively impressive R^2 values as displayed in Sharma [111] for example, become much less so when the loss of degrees of freedom is considered. For example, an R^2 value of 57% recorded for forecasting rainfall from one of the multiple regression models looks promising until the adjusted R^2 of only 38% is taken into consideration.

7.5.3 Results and predictions

Regression models in Equations 7.1 and 7.2 are fitted with the strongest correlated SST pixels for a particular time lag. The adjusted R^2 values for the models without interaction terms and with interaction terms are shown in Tables 7.5 and 7.6 respectively. In general for most time lags, regression models with interactions terms generally perform better in explaining variability as compared to the models without interaction terms. For the purpose of analyzing the predictive ability of SST, the highest adjusted

R^2 value from Table 7.6 is chosen. This corresponds to the regression model with March rainfall regressed on the previous November five highest correlated SST pixels, giving an adjusted R^2 of 0.419.

Table 7.5: Adjusted R-squared values from regression models without interaction terms.

Lag of SST/ Month of rainfall	**Jan**	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**	**Dec**
1	0.206	0.216	0.167	0.215	0.208	0.126	0.160	0.247	0.195	0.279	0.163	0.092
2	0.198	0.292	0.176	0.177	0.165	0.158	0.133	0.285	0.158	0.187	0.140	0.246
3	0.194	0.203	0.127	0.172	0.161	0.107	0.150	0.253	0.188	0.183	0.175	0.171
4	0.190	0.132	0.208	0.239	0.258	0.161	0.111	0.236	0.300	0.268	0.177	0.210
5	0.223	0.166	0.152	0.303	0.186	0.148	0.176	0.196	0.217	0.348	0.184	0.140
6	0.145	0.198	0.327	0.141	0.139	0.197	0.221	0.141	0.108	0.255	0.155	0.216
7	0.120	0.180	0.179	0.217	0.098	0.182	0.193	0.144	0.220	0.222	0.241	0.133
8	0.173	0.163	0.175	0.209	0.152	0.212	0.170	0.156	0.242	0.250	0.149	0.140
9	0.196	0.164	0.230	0.184	0.201	0.179	0.179	0.119	0.198	0.146	0.341	0.184
10	0.175	0.184	0.159	0.115	0.295	0.135	0.212	0.168	0.292	0.347	0.350	0.269
11	0.226	0.106	0.177	0.096	0.123	0.107	0.127	0.220	0.204	0.261	0.213	0.129
12	0.201	0.161	0.134	0.203	0.186	0.184	0.138	0.159	0.255	0.247	0.207	0.113

Table 7.6: Adjusted R-squared values from regression models with interaction terms.

Lag of SST/ Month of rainfall	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0.186	0.163	0.161	0.221	0.204	0.053	0.312	0.267	0.169	0.395	0.122	-0.001
2	0.227	0.194	0.230	0.104	0.297	0.202	0.042	0.307	0.146	0.259	0.202	0.252
3	0.401	0.131	0.198	0.196	0.101	0.208	0.210	0.283	0.076	0.293	0.167	0.208
4	0.235	0.094	0.419	0.105	0.200	0.057	0.130	0.168	0.282	0.305	0.316	0.115
5	0.193	0.155	0.301	0.233	0.176	0.179	0.201	0.210	0.176	0.357	0.192	0.125
6	0.235	0.071	0.386	0.179	0.147	0.105	0.326	0.030	0.019	0.191	0.136	0.226
7	0.151	0.118	0.217	0.178	0.160	0.123	0.081	0.186	0.249	0.271	0.224	0.060
8	0.232	0.126	0.162	0.249	0.271	0.183	0.212	0.182	0.204	0.249	0.093	0.091
9	0.252	0.279	0.237	0.377	0.189	0.138	0.284	0.058	0.211	0.149	0.351	0.071
10	0.355	0.211	0.275	0.095	0.349	0.134	0.219	0.270	0.269	0.238	0.361	0.212
11	0.248	0.087	0.255	0.057	-0.05	0.113	0.110	0.074	0.299	0.254	0.246	0.244
12	0.187	0.041	0.300	0.175	0.174	0.133	0.196	0.099	0.313	0.234	0.205	0.03

The first 90 years of March rainfall data from 1901 to 1990 is regressed against the five November SST pixels from 1900 to 1989. Based on this model, the rainfall for March 1991 is predicted using November 1990 SST anomalies. This process is repeated as each new March rainfall observation is added and the regression model is re-fitted with new parameters for the previous year's November SST anomalies.

Predictions from the regression model is then compared with the observed rainfall and the residuals are calculated. To evaluate the performance of the rainfall predictions, the median of the observed rainfall distribution is calculated and this is compared to the predicted rainfall. The probability that the predicted rainfall will exceed the median is calculated using the following equation:

$$\mathbb{P}(\text{predicted rainfall exceeds median rainfall}) \;=\; \Phi\left(\frac{\hat{y}_{t+1} - m_y}{s_{\hat{y}}}\right) \qquad (7.3)$$

where $\hat{y}_{t+1}$ is the predicted March rainfall at time $t+1$, m_y represents the median of the observed rainfall up to time t and $s_{\hat{y}}$ is the standard deviation of the predicted rainfall distribution.

Both predicted and median rainfall values are considered against the observed March rainfall and recommendations are made based on the accuracy of each value relative to the observed rainfall. Table 7.7 displays the rainfall prediction results for 1991 to 2003, where $p_{\text{exceed } m_y}$ denotes the probability of exceeding the median rainfall as expressed in Equation (7.3) and 'M' refers to the recommendation of using the median over the prediction and 'P' refers to the preference of using predicted rainfall over the median. The time series plot of predicted rainfall (dashed line) and observed rainfall (solid line) is provided in Figure 7.18(a), along with the median rainfall (dotted line), while Figure 7.18(b) shows their correlation plot.

Overall, all of the regression models fitted in the above table have an adjusted R-squared of between 0.42 and 0.47. In all of the thirteen predictions, the model only produced two favorable predictions that were closer in value to the observed rainfall than the median rainfall. More often than not, the model produces extremely high rainfall predictions. There is a weak correlation between the observed and predicted rainfall of -0.25 and also by observing Figure 7.18(b). Hence, despite the high correlations and adjusted R-squared which exist between the selected SST pixels and Warragamba rainfall, results from this application indicate that the regression model is often inadequate in producing four-month ahead predictions.

Table 7.7: Summary table of predictions for Warragamba March rainfall, 1991-2003.

Rainfall year(t)	**Observed Rainfall**(y_t)	**Predicted Rainfall**($\hat{y}_t$)	**Residual**	**Median** (mm)	$p_{\text{exceed } m_y}$	**Recommendation** (M/P)
1991	12.8	75.8	-63.0	43.4	0.83	M
1992	58.2	84.3	-26.1	43.2	0.88	M
1993	88.9	27.1	61.8	43.4	0.32	M
1994	39.2	1.08	38.1	43.6	0.11	M
1995	13.7	42.7	-29.0	43.4	0.49	P
1996	36.5	75.5	-39.0	43.2	0.82	M
1997	29.7	108	-78.4	42.75	0.97	M
1998	6.5	91.3	-84.8	42.3	0.92	M
1999	77.4	42.8	34.6	42.1	0.51	P
2000	106.7	23.3	83.4	42.3	0.29	M
2001	55.4	196	-141	42.75	1.00	M
2002	35.0	75.3	-40.3	43.2	0.82	M
2003	55.6	85.7	-30.1	42.75	0.89	M

(a) (b)

Figure 7.18: (a) Time series plot of predicted and observed Warragamba March rainfall; (b) Correlation plot between observed and predicted rainfall, 1991-2003.

7.6 Summary of Chapter

This chapter explores the prediction capability of two climatic indicators known to affect Australia's rainfall significantly; the SOI and SST. Previous studies have investigated the relationship of rainfall with SOI and SST. This chapter adds to earlier work by re-examining the phase method introduced by Stone *et. al.* [119] and also the probabilistic forecasts procedure using the SST anomalies by Sharma *et. al.* [111].

The investigation of the SOI phase method was extended to all of Australia's rainfall in the first half of this chapter. The SOI phase method was then compared with the SOI May categories, but there was no evidence that the former is better for predicting

June to October rainfall than the latter.

The SOI phase method effectively reduces to a claim that a rapid fall in SOI over April and May is associated with below average June to October rainfall, and that a rapid rise in April and May SOI is associated with an above average June to October rainfall. Forty out of 100 years fall into one or the other of these phases. The difference between mean rainfall during the rapid fall and rapid rise is statistically significant at the 5% level for 27% of the pixels in Australia. In a similar way, the difference between mean rainfall for SOI May categorized as Strong El-Niño and Strong La-Niña is statistically significant at the 5% level for 34% of pixels but only 35 out of 104 years. Once the number of years in which credible prediction can be made is allowed, there is little difference in the performance of the two forecasting strategies. It is perhaps more useful to consider using a combination of both, which does at least allow plausible predictions of low rainfall to be made approximately 25% of the time. Table 7.8 shows the distribution of years given both categorization methods.

Table 7.8: Number of years in each categorization

Categories	Strong El-Niño	Other	Strong La-Niña
SOI Rapid rise	0	18	8
Other	11	45	8
SOI Rapid fall	8	6	0

Large-scale correlations between all SST pixels and a selected rainfall pixel were calculated in the later part of this chapter. A routine which selects the five highest correlated SST pixels allowed for separate multiple regression models to be fitted for one to twelve month time lags. The adjusted R^2 values for each of these multiple regression models with and without interaction terms, were used to measure the strength of the relationship, since this statistic takes into account the number of terms used in the model. Results indicate that there is a general improvement in adjusted R^2 values when the interaction terms are included in the regression models. To demonstrate the application of these models, the model with the highest adjusted R^2 value was chosen. The resulting predictions performed badly against the observed rainfall, despite high correlations between the climatic indicator and the rainfall.

Both applications in this chapter have taken advantage of the associations of the climatic indicators and rainfall to carry out predictions in Australia. These relationships are useful since an estimate of the increase or decrease in the amount of rainfall could

assist farmers in making better informed decisions about which crops to plant and environmental agencies to manage water supplies and plan restrictions on use.

Chapter 8

Improving drought predictions using copulas

Modelling of drought characteristics using copulas and the influence of ENSO on the dependence structure were explored in Chapter 4 and 5. The relationship between rainfall around Australia and two prominent climatic indicators were examined in detail using several methods of categorization in Chapter 7.

In this chapter, rainfall sum from June to October in two contiguous rainfall districts are predicted from April and May SOI using the trivariate Gumbel-Hougaard copula conditioned on the SOI April state being classified as positive or negative. Statistical tests show a significant difference between the fitted copulas in both SOI states. This copula approach can be extended to modelling for more than two sites and the rainfall prediction can be used to provide an estimate of drought risk.

8.1 Study area and data

Winter cropping season in eastern Australia is from June to October and hence predicting rainfall over this period would be of great financial value. June to October rainfall from 1913 to 2002 from two neighbouring districts lying west of the Great Dividing Range, District 62 (Central Tablelands north) and 64 (Central Western Slopes) in NSW have been selected to illustrate this. Both rainfall districts are an important region in producing wool, beef, oats and wheat.

NSW has a temperate climate with large climatic variations depending on the prox-

imity to the coast and mountains. Figure 8.1 shows the location of the neighbouring districts 62 and 64. District 62 is located within the central parts of NSW, with latitude 31.73° to 33°S and longitude 148.97° to 150.4°E. The region has a varied elevation of between 170m to 1101m, with maximum monthly rainfall being 323.1mm and minimum being 0mm. Lying north of District 62 is District 64 positioned at a latitude of 31.2° to 32.28°S and longitude of between 148.7° and 148°E. The elevation range in this district is between 213.4m to 860m, with maximum monthly rainfall of 309.1mm. Both Districts 62 and 64 have roughly similar mean annual rainfall of 642.6mm and 625.8mm respectively. Figures 8.2 display the time series plot of June to October rainfall sum for Districts 62 and 64. There is no evidence of a linear trend in both districts. The monthly SOI data over this time period is obtained from the Queensland Government Long Paddock website [94], which is also used in Chapter 7 where the time series plot is also provided .

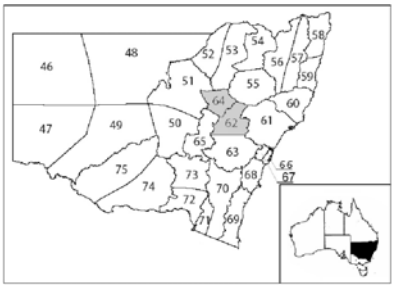


Figure 8.1: Location of District 62 and 64, NSW

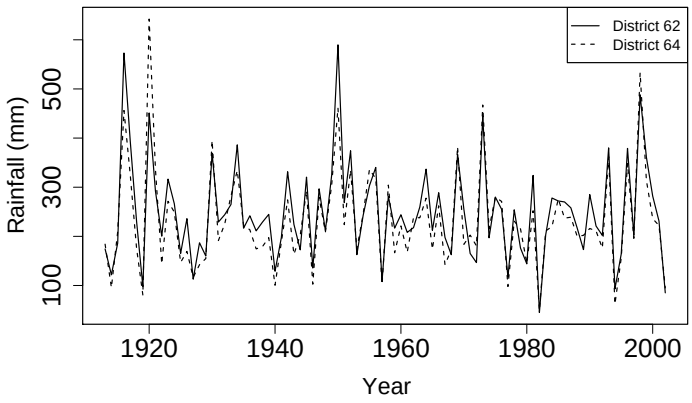


Figure 8.2: Time series plot of June to October rainfall sum for Districts 62 and 64, 1913-2002.

8.2 Methodology and Application

8.2.1 Identifying and categorizing of copula variables

In Chapters 2 and 3, negative values of the SOI were found to be highly correlated with El-Niño events in Australia, during which droughts occur more frequently. Changes between the SOI April and May states were associated with subsequent June to October rainfall total and this was discussed in Chapter 7. This relationship was shown to be effective in predicting drought occurrence and its corresponding risk. Furthermore, drought risk for a particular rainfall region can be better determined, since rainfall from neighbouring districts are generally spatially correlated.

Based on these established relationships with June to October rainfall sum, three variables are chosen to model their dependence structure by the trivariate Gumbel-Hougaard copula, introduced in Chapter 4. The three variables are: District 62 June to October rainfall total; District 64 June to October rainfall total; and SOI May. For each year from 1913 to 2002, these data triplets are categorized into two groups based on the SOI April value; positive (+ve) or negative (−ve). Table 8.1 provides the basic statistics of the variables in the two groups and Table 8.2 shows the correlations between the copula variables in both groups. There is a general decrease in mean June to October rainfall total in both districts associated with SOI April being negative. It is also more likely for SOI May to be negative when SOI April is negative. Similarly, when SOI April is in a positive state, there is a higher tendency for SOI May to be positive. Correlations between rainfall in both districts are roughly similar for both SOI April states, however correlations between the respective district rainfall totals and SOI May in SOI April negative state is overall higher as compared to that in the positive state.

Table 8.1: Means and Standard deviations of copula variables given SOI April state

Variable	District 62 Jun-Oct rainfall sum (mm)		District 64 Jun-Oct rainfall sum (mm)		SOI May	
	SOI −ve	SOI +ve	SOI −ve	SOI +ve	SOI −ve	SOI +ve
Number of years	50	40	50	40	50	40
Mean	237.4	262.6	226.2	239.1	-2.7	3.5
Standard deviation	110.8	83.7	120.6	70.5	10.1	7.1

Table 8.2: Correlations between copula variables given SOI April state

Correlations	Pearson	
	SOI April −ve	SOI April +ve
District 62 Jun-Oct rainfall, District 64 Jun-Oct rainfall	0.93	0.91
District 62 Jun-Oct rainfall, SOI May	0.37	0.11
District 64 Jun-Oct rainfall, SOI May	0.31	0.10

8.2.2 Fitting marginal distributions to copula variables

Similar approach used in Chapters 4.2 and 5.4, are applied here to fit appropriate marginals to each copula variable for both SOI April states. To determine the most appropriate distribution requires the use of the AIC criterion again.

Tables 8.3 shows the best fitted distribution for each copula variable with estimated parameters for the respective distributions. The estimated means for the corresponding distributions for Districts 62 and 64 are both lower when SOI April is negative. The largest difference in estimated means between both SOI April states is observed for District 62 June to October rainfall total, where the standard deviation of the rainfall total is larger in the SOI April negative state (9.99) compared to the SOI April positive state (7.01).

Table 8.4 computes the correlation between the copula variables following the transformation to uniform variables. Comparing these correlations to those in Table 8.2, it is evident that the correlations between the rainfall totals in both districts remain largely unaffected in both SOI April states, which is not surprising since rainfall in neighbouring districts are usually highly correlated.

However, correlations between rainfall totals from both districts and SOI May have reduced in the SOI April positive state, while an increment is observed in the SOI April negative state. Following this transformation, correlations between rainfall totals from both districts and SOI May are found to be highly similar for both SOI April states. For example, before transformation when SOI April is in a negative state, correlations between June to October rainfall total from District 62 and SOI May is 0.37, while correlations between that of District 64 and SOI May is 0.31. The corresponding correlations for the same set of data after transformation was found to be 0.46 and 0.45 respectively. These highly similar correlations are essential when the dependence structure of these variables are modelled by the asymmetric Gumbel-Hougaard copula.

This is ideal for this form of Archimedean copula since, a restriction of this copula is that the correlations between the inner variables and the outer variable of the copula are identical, as mentioned in Chapter 4.

Table 8.3: Parameters of marginal distributions of copula variables

Variable	Distribution	Estimated Parameters
SOI April positive		
District 62 Jun-Oct rainfall	Normal	$\mu = 3.53, \sigma = 7.01$
District 64 Jun-Oct rainfall	Lognormal	$\mu = 5.53, \sigma = 0.29$
SOI May	Lognormal	$\mu = 5.44, \sigma = 0.29$
SOI April negative		
District 62 Jun-Oct rainfall	Normal	$\mu = -2.72, \sigma = 9.99$
District 64 Jun-Oct rainfall	Lognormal	$\mu = 5.36, \sigma = 0.49$
SOI May	Lognormal	$\mu = 5.29, \sigma = 0.53$

Table 8.4: Correlations between transformed uniform copula variables given SOI April state

Correlations	Pearson	
	SOI April −ve	SOI April +ve
District 62 Jun-Oct rainfall, District 64 Jun-Oct rainfall	0.94	0.89
District 62 Jun-Oct rainfall, SOI May	0.46	0.02
District 64 Jun-Oct rainfall, SOI May	0.45	0.04

8.2.3 Modelling dependence structure using trivariate Gumbel-Hougaard copula

The asymmetric trivariate Gumbel-Hougaard copula from Equation (4.12) in Chapter 4 is used here. Recall that one of the property of this copula is that the dependence parameter of nested variables in the copula are the highest correlated pairs. In this case, the rainfall totals between District 62 and 64 are the highest among all the variable pairs. Due to this prerequisite, the uniform variables of District 62 June to October rainfall total are taken to be u_1, District 64 June to October rainfall total to be u_2 and SOI May to be u_3 in Equation (4.12).

For both trivariate Gumbel-Hougaard copulas in their respective SOI April states, the estimation of parameters θ_1 and θ_2 are estimated using MLE, introduced in Chapters 4 and 5. These estimated parameters for both SOI April states are shown in Table 8.5,

along with their standard errors given in brackets. Both estimated copula parameters are higher in the SOI April negative state. This is due to the higher correlations observed in Tables 8.3 and 8.4 during the SOI April negative state.

Table 8.5: Parameters of trivariate Gumbel-Hougaard copula given SOI state

Parameter	SOI April −ve	SOI April +ve
$\hat{\theta}_1$	1.26	1.04
	(0.13)	(0.09)
$\hat{\theta}_2$	4.52	3.21
	(0.52)	(0.41)

In order to assess if a significant difference exists between the dependence structure of variables from SOI April positive and negative states, the null hypothesis of equality between the copula parameters of both SOI states is assumed. Statistical significance tests at the 5% level are performed on the estimates of the dependence parameter θ. The test statistics for testing the hypotheses of the equality of parameters in the two SOI states are 1.38 for θ_1 and 1.96 for θ_2, which corresponds to p-values of 0.18 and 0.05 respectively.

Results indicate a very strong evidence that θ_2 is higher during the SOI April negative state than positive state. There is also weak evidence that θ_1 is higher in SOI April negative state than positive state. These results justify allowing for different dependence parameters for different climatic conditions.

8.2.4 Goodness-of-fit tests

Once again, the fitted Gumbel-Hougaard copulas have to be assessed for a satisfactory fit to the observed data. Similar approach taken in Chapters 4 and 5 is used here. Figure 8.3 shows the plot of the probabilities derived from the fitted copula against the probabilities calculated from the empirical copula, where the empirical copula is given in Equation 4.17 in Chapter 4. Recall that the distance between the plotted points and the diagonal line should be minimized for an optimal fit. For both SOI April states, the points are relatively close to the diagonal line and hence satisfy this test.

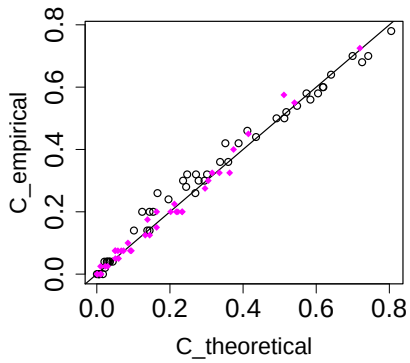


Figure 8.3: Plot of values from theoretical Gumbel-Hougaard copula against empirical copula for SOI April positive (diamonds) and SOI April negative (circles)

8.2.5 Simulating from the Gumbel-Hougaard Copula

Random variables are simulated from the respective trivariate Gumbel-Hougaard copula fitted with their estimated parameters obtained above, using the algorithm described in Chapter 4.2.3. Simulations of 100,000 are generated for the three variables from their corresponding copula given SOI April state. Table 8.6 shows the correlations between pairs of simulated variables. The correlations derived from the simulations are close to those derived from the observed data (Table 8.2) when the SOI April is negative. The difference in the correlations are negligible when similar comparisons are made for the SOI April positive state. These simulations justify that the fitted trivariate Gumbel-Hougaard copulas for both SOI April states are adequate in modelling the dependence structure of the observed data.

Copula density plots of District 62 June to October rainfall against SOI May for both SOI April states are constructed from the simulations, and are displayed in Figures 8.4(a) and (b) respectively. Figures 8.5(a) and (b) shows the copula density plots of District 64 June to October rainfall total and SOI May, given their SOI April states.

Table 8.6: Correlations between simulated variables given SOI April state

Correlations	SOI April −ve	SOI April +ve
u_1, u_2	0.93	0.87
u_1, u_3	0.31	0.06
u_2, u_3	0.31	0.06

Both Figures 8.4 and 8.5 show that different SOI April conditions have an effect on

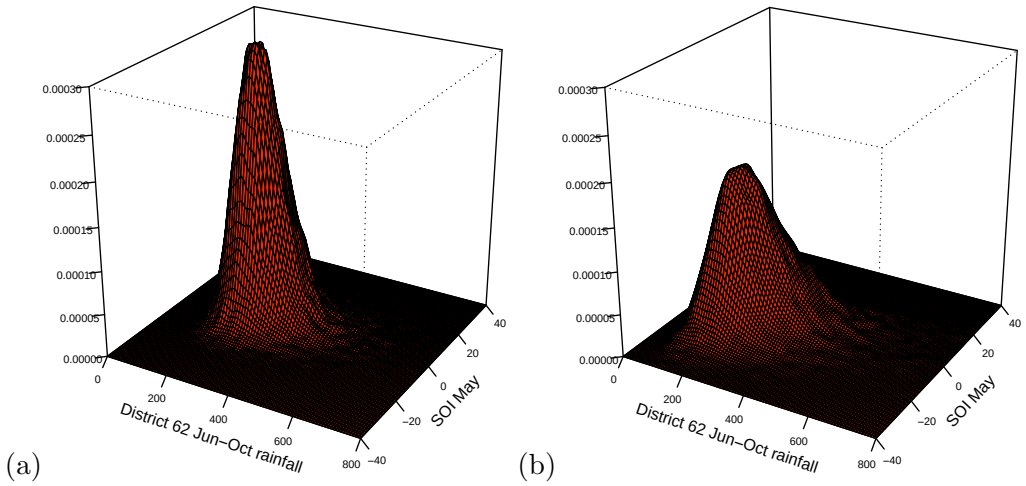


Figure 8.4: Copula densities of District 62 sum of June to October rainfall and SOI May during (a) SOI April positive; and (b) SOI April negative states.

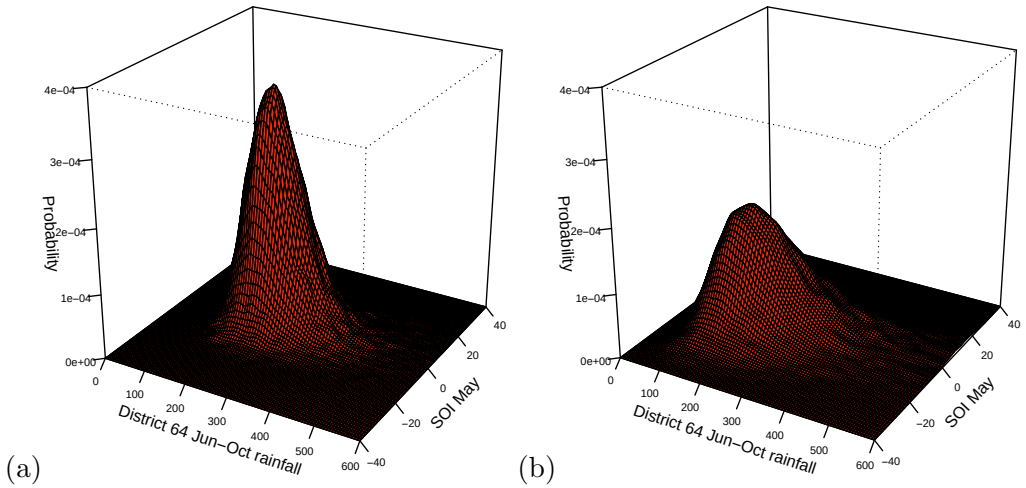


Figure 8.5: Copula densities of District 64 sum of June to October rainfall and SOI May during (a) SOI April positive; and (b) SOI April negative states.

the copula densities between both district rainfall totals and the SOI May. Given this difference in density plots and copula dependence parameters for both SOI April states, the probability of exceeding the drought threshold and consequently a drought occurring, can be derived based on these distribution plots and the use of the SPI. The SPI of -1 or less indicates a drought, hence rainfall amounts equivalent to this SPI value in both districts can be computed and used to assess the probability of drought risk in these districts.

8.2.6 Forecasting using copula

The fitted copulas for whichever SOI April state pertains, can be used for forecasting the June to October rainfall distributions of Districts 62 (u_1) and 64 (u_2) given the

SOI May (u_3). First, the copula $C(u_1, u_2, u_3)$ has to be partially differentiated with respect to u_1, u_2 and u_3 to give its *pdf* $c(u_1, u_2, u_3)$. Then given the SOI May as u_3^*, we can simulate from $c(u_1, u_2 \mid u_3^*)$ by using the Metropolis-Hastings algorithm [23] on $c(u_1, u_2, u_3^*)$. This will generate the bivariate rainfall distribution of u_1 and u_2 given $u_3 = u_3^*$ and whether SOI April is positive or negative.

8.3 Summary of chapter

This chapter extends the application of copulas from modelling drought characteristics in Chapters 4 and 5 to providing a framework for the assessment of drought risk through the established relationship between preceding SOI states and rainfall in neighbouring rainfall districts (Chapter 8). This model is applied to two neighbouring districts, Districts 62 and 64, in NSW. Two groups of triplets consisting of June to October rainfall total in Districts 62 and 64 and SOI May, are obtained by segregating them according to the positive or negative values of SOI April.

The dependence structure of the triplets for each SOI April category is then modelled through the Gumbel-Hougaard copula, where the copula parameters are estimated through MLE. Goodness-of-fit tests performed indicate no evidence of a lack of fit using the Gumbel-Hougaard copula to model these set of historical data. Significance tests and the copula density plots from the simulations for both SOI April positive and negative demonstrate the difference between dependence structures of the triplets given the SOI value from the preceding month. Based on these results, drought risk through the use of rainfall predictions may be improved with the inclusion of the SOI from preceding months in the copula model.

Chapter 9

Conclusions and future research directions

9.1 Conclusions

The aim of this thesis was to predict drought occurrence spatially and temporally in Australia, by first considering multivariate models of drought risk and then examining the practicability of short term predictions of droughts. Outcomes from this research will provide drought risk information which will be useful for drought risk assessment and short-term drought predictions. This information could assist the agricultural community make better informed decisions on appropriate crops to plant, as well as environmental and governmental agencies to manage water supplies and plan restrictions on use.

Much of the literature concentrates on droughts defined solely in terms of their duration, which is a univariate approach, and relies on aggregation of rainfall over a defined area. Examples of such areas are: a square pixel, a river basin or a rainfall district, or a larger region such as a state. A multivariate approach allows for further characteristics of droughts, such as their peak and average intensities, to be modelled and also provides scope for more detailed spatial modelling. The most common approach to multivariate modelling is the multivariate Gaussian distribution, which has the advantage of being mathematically tractable. There are standard results for conditional distributions of one sub-set of variables on another subset.

The apparent restriction of Gaussian marginal distributions can easily be circum-

vented by transforming any marginal *cdf* to a uniform *cdf* and then taking the inverse Gaussian distribution, which is equivalent to fitting a Gaussian copula. A popular alternative is to take a Box-Cox transform of the variables. However, the multivariate Gaussian distribution has a unique correlation structure which is not realistic for many applications especially for extreme values. In particular, the multivariate Gaussian distribution does not have tail dependence. Copulas offer a highly versatile approach to modelling all multivariate distributions, and there has been renewed research interest in this area over the past 10 years.

In Chapter 4, a single rainfall district in NSW was considered and copulas were fitted to three drought characteristics: duration; severity and peak intensity. Comparison was made between the Gaussian and Gumbel-Hougaard copulas and it was demonstrated that the tail dependence of the latter provided a more realistic model.

In Chapter 5, the ideas of Chapter 4 were extended to the comparison of two rainfall districts, east and west of the Great Dividing Range, conditioned on El-Niño and La-Niña climatic states. Furthermore, similar to the Gaussian copula, the t-copula, allows for separate correlation parameters and also allows for tail dependence, was introduced. Following this, a comparison of the copula forms was made in terms of return periods. This chapter provided a comprehensive demonstration of the copula model and showed that the calculation of return periods on a univariate scale can result in misleading conclusions and analysis of drought risk when correlations are present between hydrological variables. The availability of such information is especially crucial when making recommendations in drought mitigation and water resource planning.

Attention of this study turned from modelling drought behaviour to short-term predictions of expected SPI values in Chapter 6, methods for forecasts with lead times of one to twelve months. The forecasting performance of various stochastic models was evaluated using rainfall data from three station gauges in NSW. From the RMSE results, it was observed that using rainfall in a climatic regression model which included climatic indices, performed the best in identifying a one-month ahead drought when compared to the SPI(3). For predictions to be of practical use to the agricultural community, they will have to be made at a longer time lag. Coupling statistical modelling with numerical weather prediction models is a method recommended to address this issue. Autoregressive models with climatic indicators and weighted regression models were employed to predict six-month ahead monthly forecasts. The inclusion of weights in

the climatic regression model reduced the RMSE value based only on drought periods, substantially. This chapter also validated the significant influence of climatic indices on droughts, by improving drought forecasts with the inclusion of climatic indicators.

Chapter 7 moved on to predicting the short-term probability rainfall distributions through establishing relationships between rainfall, SOI and SST variables. In particular, this study has examined the performance of a widely-used rainfall prediction software, Rainman [119], and compared this to other forms of categorization techniques. Firstly, this study proposed alternate and simpler strategy for forecasting June to October rainfall using only SOI from the preceding month. This strategy was demonstrated for all of Australia. Statistical significance testing between the extreme categories of each method were carried out for all pixels in Australia and their forecasting performance were evaluated based on the percentage of pixels that were significantly different at the 5% level.

This thesis has presented results that show that the classification technique which only uses the SOI from the preceding month, is equally capable of providing probabilistic rainfall forecasts as Rainman. A third classification procedure categorized rainfall of the same year according to the ENSO state that year was in. This study showed promising results of an association between June to October rainfall and ENSO state, hence the potential of using this association for rainfall predictions, given an accurate prediction of an ENSO state, can be made well in advance. Such a prediction would need to come from a numerical weather prediction model. In addition, this analysis has demonstrated that such classification methods and software, should be used with extreme care when making predictions, especially in regions where no significant difference in rainfall between categories were found, as this may cause misleading conclusions.

This research continued investigating the potential of probabilistic rainfall forecasting based on relationships with another influential climatic indicator, SST anomalies. This analysis follows from the study by Sharma [111], which showed that using highly correlated pixels of SST anomalies as predictors of quarterly rainfall, are an improvement to three commonly used ENSO indices. Utilizing rainfall data from the same region as Sharma [111] and selecting the corresponding five highest correlated SST pixels, this section has demonstrated that although these pixels gave relatively impressive R^2 values from the fitted model which was considered by Sharma [111], these R^2 values were less impressive once degrees of freedom is taken into account which was reflected

by the adjusted R^2 values. Furthermore, the rainfall predictions made from these regression models performed poorly and it was therefore recommended by this study to use the median rainfall instead.

Finally, in Chapter 8, the thesis combined two concepts from earlier chapters and extended it to improve drought predictions. This chapter took advantage of the established June to October rainfall relationship with SOI April and May discussed in Chapter 7, and the copula concept employed to model drought characteristics. Rainfall during this period from two neighbouring NSW rainfall districts were fitted to the trivariate Gumbel-Hougaard copula along with the SOI in May. Separate trivariate copulas were fitted for different SOI April conditions. Results from this analysis showed a significant difference between the fitted copula parameters, and was further supported by the copula densities from both ENSO states. In the case where a significant difference exists, these copula densities were shown to be valuable in drought risk assessment. Findings from this thesis have opened a number of possible research areas that could contribute to improvements in drought predictions and they are described in the following section.

9.2 Future research directions

This research has presented results which are instrumental in understanding the nature of droughts in Australia and may lead to further development in the modelling and prediction of droughts. Some of these results have led to the identification of future possible research avenues which are discussed below.

9.2.1 Further modelling of drought characteristics using copulas

Though the application of copulas in the area of hydrology has been recent, this concept represents a significant addition to the modelling of drought characteristics and drought predictions. The field of copula research is rather wide and there are many other copula forms for which their applications to drought remains to be explored. Within the Archimedean copula family, a range of copulas are capable of modeling a variety of dependence structures, for example the Cook-Johnson and Frank copula. Furthermore, limited research on the application of elliptical copulas to droughts has been carried out, mainly due to not having closed-formed solutions. Results from the

application of t-copulas in Chapter 5 have been promising, and it is possible that this copula may be extended to fit a larger number of variables. One of the drawbacks of the t-copula is its strong symmetry, where the level of tail dependence in the upper tail is the same as the lower tail. Several extensions of the t-copula were introduced in [30] to introduce more asymmetry, for example the skewed t-copula allows for heavier marginal tails and the grouped t-copula allows for different levels of tail dependence for each subvector of the vector $\mathbf{X}$. These modified versions of the t-copula may provide more flexibility in modelling specific extreme drought events.

Another aspect of the copula approach which has recently gained increasing attention is the use of distortions. Some authors label this as a transformation of the existing copulas [35] and few studies have focused on this technique [34, 79, 5]. Recall the basic construction of an Archimedean bivariate copula is based on the following:

$$C(u_1, u_2) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2)) \tag{9.1}$$

Then for $\gamma : [0,1] \to [0,1]$, Durrleman *et. al.* [35] defines the transformation of the copula $\mathbf{C}_\gamma$ on $[0,1]^2$ as

$$\mathbf{C}_\gamma(x, y) = \gamma^{-1}(\mathbf{C}(\gamma(x), \gamma(y))) \tag{9.2}$$

The distorted copula is also a copula and possesses properties such as concavity on $[0,1]$, being twice differentiable and continuous from $[0,1]$ to $[0,1]$. Morillas [79] provides a list of expressions that the transforming generator functions γ can take. This introduction of distortions to copulas has paved the way for the construction of a variety of new copulas, which can be more versatile for drought modelling and predicting.

Part of the IFM approach discussed in Chapter 4 was to fit a continuous marginal distribution to each hydrological variable to transform them onto a uniform scale. Drought characteristics in this thesis have been calculated on both monthly and daily time scales. However, the use of months to derive drought duration results in step or discrete-looking data. Issues have been raised over whether the use of continuous marginal distributions on duration should be permitted. Shiau [113] justifies fitting an exponential distribution to drought duration defined in months [149]. Utilizing daily rainfall in this thesis have been proposed to overcome this restriction. The discrete copula family is another alternative and applications of this copula form have been recent [43]. The application of this copula family to hydrology is worth investigating

and may improve the existing approach.

9.2.2 Improving predictions of climatic indices and the inclusions of other indices

This thesis has concentrated on the relationships between the ENSO phenomenon and its associated indices, the SOI and SST, with drought to make predictions. In Chapter 7, two categorization techniques which relied on SOI from preceding months were used. Further work will investigate the relationship between rainfall and a linear combination of preceding SOI, using a canonical correlation technique. Furthermore, the third categorization which displayed a significant difference between rainfall during El-Niño and La-Niña states throughout most of eastern and central Australia. Thorough investigation into the performance of rainfall forecasts could be conducted, based on combining predictions of annual ENSO with relatively high correlation between June to October rainfall and annual ENSO, using physically based climate models of Chen *et. al.* [21].

The drought variability in Australia is largely due to its location, which is influenced by the ocean-atmospheric interactions in the Indian, Southern and Pacific Ocean. Other global climatic indices worth investigating are the IOD and MEI, which could be incorporated in the development of drought forecasting models. Background description of the IOD in Chapter 2, associated the positive phase of the IOD with dry conditions in parts of Australia. There is still substantial work required to develop a forecasting framework to predict IOD phases, by improving the predictions of the corresponding index, DMI. Luo *et. al.* [65] has shown promising results using a coupled model to show that two positive IOD phases can be predicted up to 3 or 4 seasons ahead. Also, the influence of the IOD on the SOI, SST, and consequently ENSO could be examined. A relationship between the IOD and SOI were suggested by Behera and Yamagata [16]. Hence, a mixture of these indices incorporated in a model could be used to predict drought. It has also been widely reported that temperature increase in recent years, has affected drought in Australia. The association of temperature with these climatic indices could also be investigated and this relationship may provide further insight into the prediction of droughts.

9.2.3 Producing predictions for wider spatial extent and applications of drought predictions

The modelling and prediction techniques in this research have so far only been applied to mainly NSW and the pixels in Australia in Chapter 7. One possible extension would be to increase the spatial extent of drought predictions, by way of modelling all rainfall pixels as a single multivariate distribution and forecasting multi-site drought. Another challenge when predicting droughts in Australia, is to predict both spatially and temporally. Spatio-temporal modelling of rainfall has only been recently considered in the literature [108, 2]. Sanso and Guenni [108] used the multivariate dynamic linear model to model time-varying long-range spatial dependence and fitted this model using Markov Chain Monte-Carlo method, which allowed for a predictive distribution of rainfall to be obtained. These methods could be applied to the spatio-temporal modelling of drought using the SPI values instead of rainfall data. Another proposed extension that may take into account both aspects would be the copula approach. Figure 9.1 displays 9 selected grids in time t_1 and t_2 for a small site. The shaded pixel is the pixel for which rainfall predictions are made in time t_2. Given that rainfall in neighbouring pixels are highly correlated with the shaded pixel, a copula that incorporates rainfall from neighbouring pixels in time t_1 and t_2, rainfall of shaded pixel in time t_1 and t_2, could be modelled. The joint distribution function and conditional distribution function may provide probabilistic rainfall information for the shaded pixel. It is possible that this application could be expanded for a larger area, by combining rainfall from 2 or more pixels.

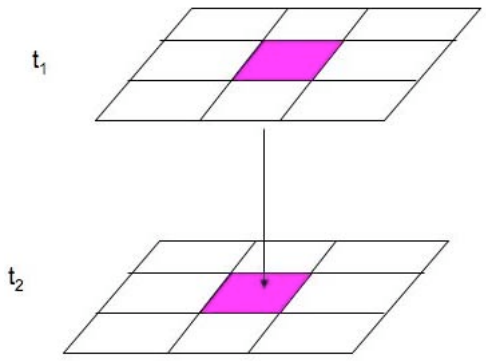


Figure 9.1: Grid rainfall diagram for a small region

As mentioned, the agricultural community is especially affected by the loss in revenue when a drought occurs. A worthwhile area of investigation would be to calculate the

definitive cost attached to each drought episode. The relationship between drought indices such as SPI and global climatic indices with the economic cost could be examined through the application of the multivariate distribution using copulas. By evaluating this relationship, a system of economic valuation, for instance, the Conditional Value at Risk (CVaR) and the Valuation at Risk (VaR) which is commonly used in finance, could be laid out. This will give farmers and governmental agencies an overview of the estimated cost to set aside, to cope with the effects of drought.

The economic cost of drought analysis would also provide a system of valuation for prevention measures to be implemented based on cost benefit analysis. Cost benefit analysis is an essential technique in project appraisal. This involves attaching monetary weights to the total expected costs and comparing these costs to the total expected benefits of one or more actions, in order to choose the most profitable option. Further research could examine the feasibility of various drought mitigation efforts, for example, dams which have been used worldwide or the construction of desalination plants.

9.3 Conclusion

This thesis has presented a comprehensive analysis of drought modelling and predictions in Australia and has established a framework for more important investigations in drought predictions. Ultimately, it is hoped that this research will be beneficial to both public and private sectors directly affected by future droughts.

Appendix A

Derivation of Conditional Archimedean copula

This appendix gives an example of deriving the conditional copula, using a trivariate asymmetric Gumbel-Hougaard copula. From Equation (4.12), the trivariate asymmetric Gumbel-Hougaard copula is given as:

$$C_1(C_2(u_1,u_2),u_3) \;=\; \exp[-([(-\ln u_1)^{\theta_2}+(-\ln u_2)^{\theta_2}]^{\frac{\theta_1}{\theta_2}}+(-\ln u_3)^{\theta_1})^{\frac{1}{\theta_1}}]$$

By the definition of Equation (4.9),

$$C_2(u_2 \mid u_1) \;=\; \frac{\partial C_2(u_1,u_2)}{\partial u_1}\Big/\frac{\partial C_1(u_1)}{\partial u_1}$$

$$C_3(u_3 \mid u_1,u_2) \;=\; \frac{\partial^2 C_3(u_1,u_2,u_3)}{\partial u_1 \partial u_2}\Big/\frac{\partial^2 C_2(u_1,u_2)}{\partial u_1 \partial u_2}$$

where $C_3(u_1,u_2,u_3) = C_1(C_2(u_1,u_2),u_3)$. The derivation of $C_2(u_2 \mid u_1)$ is

$$\begin{aligned}\frac{\partial C_2(u_1,u_2)}{\partial u_1}\Big/\frac{\partial C_1(u_1)}{\partial u_1} &= \exp(-((-\ln(u_1))^{\theta_2}+(-\ln(u_2))^{\theta_2})^{1/\theta_2}) \\ &\times (((-\ln(u_1))^{\theta_2}+(-\ln(u_2))^{\theta_2})^{(1/\theta_2)-1}\times((1/\theta_2) \\ &\times((-\ln(u_1))^{(\theta_2)-1}\times((\theta_2)\times(1/u_1)))))\end{aligned}$$

and $\frac{\partial C_1(u_1)}{\partial u_1}=1$.

The derivation of $C_3(u_3 \mid u_1,u_2)$ first requires the derivation of $\frac{\partial^2 C_3(u_1,u_2,u_3)}{\partial u_1 \partial u_2}$ and

$\frac{\partial^2 C_2(u_1,u_2)}{\partial u_1 \partial u_2}$ separately. The result from deriving $\frac{\partial^2 C_3(u_1,u_2,u_3)}{\partial u_1 \partial u_2}$ is

$$
\begin{aligned}
\frac{\partial^2 C_3(u_1,u_2,u_3)}{\partial u_1 \partial u_2} &= \exp(-(((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{\frac{\theta_1}{\theta_2}} \\
&+ (-\ln(u_3))^{\theta_1})^{1/\theta_1}) \times ((((-\ln(u_1))^{\theta_2} \\
&+ (-\ln(u_2))^{\theta_2})^{\theta_1/\theta_2} + (-\ln(u_3))^{\theta_1})^{(1/\theta_1)-1} \\
&\times ((1/\theta_1) \times (((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{(\theta_1/\theta_2)-1} \\
&\times ((\theta_1/\theta_2) \times ((-\ln(u_2))^{\theta_2-1} \times (\theta_2 \times (1/u_2)))))) \\
&\times ((((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{\theta_1/\theta_2} \\
&+ (-\ln(u_3))^{\theta_1})^{(1/\theta_1)-1} \times ((1/\theta_1) \times (((-\ln(u_1))^{\theta_2} \\
&+ (-\ln(u_1))^{\theta_2})^{(\theta_1/\theta_2)-1} \times ((\theta_1/\theta_2) \\
&\times ((-\ln(u_1))^{(\theta_2)-1} \times (\theta_2 \times (1/u_1))))))) - \exp(-(((-\ln(u_1))^{\theta_2} \\
&+ (-\ln(u_2))^{\theta_2})^{\theta_1/\theta_2} + (-\ln(u_3))^{\theta_1})^{1/\theta_1}) \\
&\times ((((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{\theta_1/\theta_2} \\
&+ (-\ln(u_3))^{\theta_1})^{(1/\theta_1)-1} \times ((1/\theta_1) \times (((-\ln(u_1))^{\theta_2} \\
&+ (-\ln(u_2))^{theta2})^{((\theta_1/\theta_2)-1)-1} \times (((\theta_1/\theta_2) - 1) \\
&\times ((-\ln(u_2))^{(\theta_2)-1} \times ((\theta_2) \times (1/u_2)))) \times ((\theta_1/\theta_2) \\
&\times ((-\ln(u_1))^{(\theta_2)-1} \times ((\theta_2) \times (1/u_1)))))) + (((-\ln(u_1))^{\theta_2} \\
&+ (-\ln(u_2))^{\theta_2})^{\theta_1/\theta_2} + (-\ln(u_3))^{\theta_1})^{((1/\theta_1)-1)-1} \\
&\times (((1/\theta_1) - 1) \times (((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{(\theta_1/\theta_2)-1} \\
&\times ((\theta_1/\theta_2) \times ((-\ln(u_2))^{(\theta_2)-1} \times ((\theta_2) \times (1/u_2)))))) \\
&\times ((1/\theta_1) \times (((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{(\theta_1/\theta_2)-1} \\
&\times ((\theta_1/\theta_2) \times ((-\ln(u_1))^{(\theta_2)-1} \times ((\theta_2) \times (1/u_1)))))))
\end{aligned}
$$

and the result obtained from $\frac{\partial^2 C_2(u_1,u_2)}{\partial u_1 \partial u_2}$ is

$$\begin{aligned}\frac{\partial^2 C_2(u_1,u_2)}{\partial u_1 \partial u_2} &= \exp(-((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{1/\theta_2}) \\ &\times (((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{(1/\theta_2)-1} \times ((1/\theta_2) \\ &\times ((-\ln(u_2))^{(\theta_2)-1} \times ((\theta_2) \times (1/v))))) \times (((-\ln(u_1))^{\theta_2} \\ &+ (-\ln(u_2))^{\theta_2})^{(1/\theta_2)-1} \times ((1/\theta_2) \times ((-\ln(u_1))^{(\theta_2)-1} \\ &\times ((\theta_2) \times (1/u_1))))) - \exp(-((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{1/\theta_2}) \\ &\times (((-\ln(u_1))^{\theta_2} + (-\ln(u_2))^{\theta_2})^{((1/\theta_2)-1)-1} \times (((1/\theta_2) - 1) \\ &\times ((-\ln(u_2))^{(\theta_2)-1} \times ((\theta_2) \times (1/u_2)))) \times ((1/\theta_2) \\ &\times ((-\ln(u_1))^{(\theta_2)-1} \times ((\theta_2) \times (1/u_1)))))\end{aligned}$$

With these results, the derivation of $C_3(u_3 \mid u_1, u_2)$ is straightforward.

Appendix B

Regression results of annual drought impact from Chapter 5

Results from regressing annual drought impact on time and ENSO indicators are provided for both District 58 and 63. Here, ENSO_ind2 refers to the La-Niña state while ENSO_ind3 refers to the Neutral state.

District 58:

```
Call:
lm(formula = yearly_sum ~ year + ENSO_ind)

Residuals:
    Min      1Q  Median      3Q     Max
-1.5186 -0.7940 -0.2953  0.4310  3.9245

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.821197   0.321566   5.664 1.93e-07 ***
year        -0.010806   0.004903  -2.204   0.0502 .
ENSO_ind2   -0.650093   0.331565  -1.961   0.0532 .
ENSO_ind3   -0.569780   0.302126  -1.886   0.0627 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 1.205 on 86 degrees of freedom
Multiple R-squared: 0.1009,Adjusted R-squared: 0.06957
F-statistic: 3.218 on 3 and 86 DF,  p-value: 0.02671
```

District 63:

```
Call:
lm(formula = yearly_sum ~ year + ENSO_ind)

Residuals:
    Min      1Q  Median      3Q     Max
-2.2315 -0.7292 -0.4120  0.3500  6.3797

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.296480   0.397826   5.773 1.21e-07 ***
year        -0.005014   0.006066  -0.827 0.410787
ENSO_ind2   -1.311595   0.410197  -3.197 0.001941 **
ENSO_ind3   -1.415353   0.373776  -3.787 0.000282 ***
---
Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1

Residual standard error: 1.49 on 86 degrees of freedom
Multiple R-squared: 0.1648,Adjusted R-squared: 0.1357
F-statistic: 5.656 on 3 and 86 DF,  p-value: 0.001386
```

Appendix C

Copula parameter estimation

This appendix shows the MLE program in **R**, for the purpose of determining the copula parameters θ_1 and θ_2 of the asymmetric trivariate Gumbel-Hougaard copula.

Following the result given in Appendix A, the MLE function is

```
MLE_parameter = function(p) {

L = (exp(-(((-log(u))^(p[2]) + (-log(v))^(p[2]))^(p[1]/p[2]) +
    (-log(w))^(p[1]))^(1/p[1])) * ((((-log(u))^(p[2]) +
    (-log(v))^(p[2]))^(p[1]/p[2]) + (-log(w))^(p[1]))^((1/p[1]) -
    1) * ((1/p[1]) * ((-log(w))^((p[1]) - 1) * ((p[1]) *
    (1/w))))) * ((((-log(u))^(p[2]) + (-log(v))^(p[2]))^(p[1]/p[2]) +
    (-log(w))^(p[1]))^((1/p[1]) - 1) * ((1/p[1]) *
    (((-log(u))^(p[2]) + (-log(v))^(p[2]))^((p[1]/p[2]) - 1)
    * ((p[1]/p[2]) * ((-log(v))^((p[2]) - 1) * ((p[2]) * (1/v)))))))
    - exp(-(((-log(u))^(p[2]) + (-log(v))^(p[2]))^(p[1]/p[2])
    + (-log(w))^(p[1]))^(1/p[1])) *  ((((-log(u))^(p[2])
    + (-log(v))^(p[2]))^(p[1]/p[2]) + (-log(w))^(p[1]))
    ^(((1/p[1]) - 1) - 1) * (((1/p[1]) - 1) * ((-log(w))
    ^((p[1]) - 1) * ((p[1]) * (1/w)))) * ((1/p[1]) * (((-log(u))
    ^(p[2]) + (-log(v))^(p[2]))^((p[1]/p[2]) -1) * ((p[1]/p[2])
    * ((-log(v))^((p[2]) - 1) * ((p[2]) * (1/v)))))))) *
    ((((-log(u))^(p[2]) + (-log(v))^(p[2]))^(p[1]/p[2])
    + (-log(w))^(p[1]))^((1/p[1]) -1) * ((1/p[1]) * (((-log(u))
```

```
^(p[2]) + (-log(v))^(p[2]))^((p[1]/p[2]) - 1) * ((p[1]/p[2])
* ((-log(u))^((p[2]) - 1) * ((p[2]) * (1/u)))))) -
exp(-(((-log(u))^(p[2]) + (-log(v))^(p[2]))^(p[1]/p[2]) +
(-log(w))^(p[1]))^(1/p[1])) * ((((-log(u))^(p[2]) +
(-log(v))^(p[2]))^(p[1]/p[2]) + (-log(w))^(p[1]))^((1/p[1]) -
1) * ((1/p[1]) * (((-log(u))^(p[2]) + (-log(v))^(p[2]))
^((p[1]/p[2]) -  1) * ((p[1]/p[2]) * ((-log(v))^((p[2]) - 1)
* ((p[2]) * (1/v)))))) * ((((-log(u))^(p[2])+ (-log(v))
^(p[2]))^(p[1]/p[2]) + (-log(w))^(p[1]))^(((1/p[1]) - 1) - 1)
* (((1/p[1]) - 1) * ((-log(w))^((p[1]) - 1) * ((p[1])* (1/w))))
* ((1/p[1]) * (((-log(u))^(p[2]) + (-log(v))^(p[2]))^((p[1]/p[2])
- 1) * ((p[1]/p[2]) * ((-log(u))^((p[2]) - 1) * ((p[2]) *
(1/u)))))) - (exp(-(((-log(u))^(p[2]) + (-log(v))^(p[2]))^(p[1]/p[2])
+ (-log(w))^(p[1]))^(1/p[1])) * ((((-log(u))^(p[2]) +
(-log(v))^(p[2]))^(p[1]/p[2]) + (-log(w))^(p[1]))^((1/p[1]) - 1)
* ((1/p[1]) * ((-log(w))^((p[1]) - 1) * ((p[1]) * (1/w)))))
* ((((-log(u))^(p[2]) + (-log(v))^(p[2]))^(p[1]/p[2]) +
(-log(w))^(p[1]))^((1/p[1]) - 1) * ((1/p[1]) * (((-log(u))^(p[2])
+ (-log(v))^(p[2]))^(((p[1]/p[2]) - 1) - 1) * (((p[1]/p[2]) - 1)
* ((-log(v))^((p[2]) - 1) * ((p[2]) * (1/v)))) * ((p[1]/p[2]) *
((-log(u))^((p[2]) - 1) * ((p[2]) * (1/u)))))) + (((-log(u))^(p[2])
+ (-log(v))^(p[2]))^(p[1]/p[2]) + (-log(w))^(p[1]))^(((1/p[1])
- 1) - 1) * (((1/p[1]) - 1) * (((-log(u))^(p[2]) + (-log(v))^(p[2]))
^((p[1]/p[2]) - 1) * ((p[1]/p[2]) * ((-log(v))^((p[2]) - 1)
* ((p[2]) * (1/v)))))) * ((1/p[1]) * (((-log(u))^(p[2]) + (-log(v))
^(p[2]))^((p[1]/p[2]) - 1) * ((p[1]/p[2]) * ((-log(u))^((p[2]) - 1)
* ((p[2]) * (1/u)))))) - exp(-(((-log(u))^(p[2]) + (-log(v))^(p[2]))
^(p[1]/p[2]) + (-log(w))^(p[1]))^(1/p[1])) * ((((-log(u))^(p[2])
+ (-log(v))^(p[2]))^(p[1]/p[2]) + (-log(w))^(p[1]))^((((1/p[1])
- 1) - 1) - 1) * ((((1/p[1]) - 1) - 1) * ((-log(w))^((p[1])
- 1) * ((p[1]) * (1/w)))) * (((1/p[1]) - 1) * (((-log(u))^(p[2])
+ (-log(v))^(p[2]))^((p[1]/p[2]) - 1) * ((p[1]/p[2])
* ((-log(v))^((p[2]) - 1) * ((p[2]) * (1/v)))))) * ((1/p[1])
* (((-log(u))^(p[2]) + (-log(v))^(p[2]))^((p[1]/p[2])
```

```
    - 1) * ((p[1]/p[2]) * ((-log(u))^((p[2]) - 1) * ((p[2]) *
    (1/u)))))) + (((-log(u))^(p[2]) + (-log(v))^(p[2]))^(p[1]/p[2])
    + (-log(w))^(p[1]))^(((1/p[1]) - 1) - 1) * (((1/p[1]) - 1)
    * ((-log(w))^((p[1]) - 1) * ((p[1]) * (1/w)))) * ((1/p[1]) *
    (((-log(u))^(p[2]) + (-log(v))^(p[2]))^(((p[1]/p[2]) - 1)
    - 1) * (((p[1]/p[2]) - 1) * ((-log(v))^((p[2]) -1) * ((p[2])
    * (1/v)))) * ((p[1]/p[2]) * ((-log(u))^((p[2]) - 1) * ((p[2])
    * (1/u))))))))

sum(-log(L))
}

MLE_parameter.nlm = nlm(MLE_parameter, p=c(2,3.5),hessian=TRUE)
```

Here, u_1, u_2 and u_3 have been represented as u, v and w respectively. θ_1 and θ_2 are represented by the vector $\mathbf{p}$.

Appendix D

Regression results from Chapter 6

This appendix provides the regression results from fitting a linear trend to both SOI and MEI respectively.

```
> time <- c(1:length(SOI))
> summary(lm(SOI ~ time))

Call:
lm(formula = SOI ~ time)

Residuals:
     Min       1Q   Median       3Q      Max
-32.2975  -6.5270   0.2228   6.8294  33.0623

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.9328988  0.6129669   3.153 0.001659 **
time        -0.0034862  0.0009824  -3.549 0.000404 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.07 on 1078 degrees of freedom
Multiple R-squared: 0.01155,Adjusted R-squared: 0.01063
F-statistic: 12.59 on 1 and 1078 DF,  p-value: 0.0004037
```

```
> time <- c(1:length(MEI))
> summary(lm(MEI ~ time))

Call:
lm(formula = MEI ~ time)

Residuals:
     Min       1Q   Median       3Q      Max
-1.95202 -0.71914 -0.02792  0.64517  2.98650

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.458321   0.074549  -6.148 1.38e-09 ***
time         0.001556   0.000199   7.818 2.20e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9478 on 646 degrees of freedom
Multiple R-squared: 0.08643,Adjusted R-squared: 0.08501
F-statistic: 61.11 on 1 and 646 DF,  p-value: 2.196e-14
```

Appendix E

Statistical significance tests

This appendix gives a brief background of the statistical tests used to determine if two or more probability distributions are from the same distributions.

Kolmogorov-Smirnov (K-S) test

The Kolmogorov-Smirnov (K-S) test is used to decide whether there is evidence that two empirical distribution functions are from different probability distributions, at a given significance level [44]. Let two random samples of sizes m and n from continuous populations F_X and F_Y have order statistics $X_{(1)}, \ldots, X_{(m)}$ with empirical distribution denoted by $S_m(X)$ and $Y_{(1)}, \ldots, Y_{(n)}$, with empirical distribution $S_n(X)$ respectively. The null hypothesis is that the population distributions are identical. The K-S test criterion is the maximum absolute difference between the two empirical distribution, $D_{m,n}$:

$$D_{m,n} \quad = \quad \max_x \mid S_m(x) - S_n(x) \mid$$

To calculate the probability of $D_{m,n}$ greater than or equal to the observed, that is $P(D_{m,n} \geq d)$ where d is the observed of $\max_x \mid S_m(x) - S_n(x) \mid$, the combined sample of $m+n$ observations are first arranged in increasing order of magnitude. On a coordinate system, this arrangement can be represented as a path. Then, d is the largest of the differences $\mid u/m - v/n \mid = \mid nu - mv \mid /mn$ where u and v are coordinates of all points on the path.

The total number of arrangements of mX and nY random variables is $\binom{m+n}{m}$, under

H_0 each of the corresponding paths is equally likely. The probability of an observed $D_{m,n}$ not less than d is the number of paths which have points at a distance from the diagonal not less than nd, defined as $A(m,n)$, divided by $\binom{m+n}{m}$. Hence

$$P(D_{m,n} \geq d) \;=\; 1 - P(D_{m,n} < d) = 1 - A(m,n)/\binom{m+n}{m}$$

Small values of this probability are evidence that the population distributions differ.

Kruskal-Wallis (K-W) test

The K-W test is a distribution-free alternative to a one-way analysis of variance, where the data are replaced by their ranks. Gibbons [44] provide a detailed explanation. Under H_0, there is a single sample of size N from the same population. The N observations are combined into a single ordered ascending sequence, keeping track of which observation belongs to which sample. Ranks are assigned to this sequence. If adjacent ranks are evenly distributed among the k samples, which would be true for a random sample from a single population, the total sum of ranks, $\sum_{i=1}^{N} i = N(N+1)/2$, would be divided proportionally according to sample size among k samples. The expected sum of ranks for the ith sample, containing n_i observations would be $\frac{n_i}{N}\frac{N(N+1)}{2} = \frac{n_i(N+1)}{2}$.

Gibbons [44] denote the actual sum of ranks assigned to the elements in the ith sample as R_i. A useful test criterion is a weighted sum of squares of deviations, with the reciprocals of the respective sample sizes used as weights. The Kruskal-Wallis (K-W) test statistic is defined as:

$$H \;=\; \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{1}{n_i} \left[R_i - \frac{n_i(N+1)}{2} \right]^2$$

If the ranked sample data is recorded in a table of k columns, where the entries in the ith column are the n_i ranks occupied by the elements in the ith sample, then R_i is the ith-column sum. The hypothesis rejected for $H \geq \chi^2_{\alpha,k-1}$. Both tests are usually applied when the underlying population locations or means are unknown.

Bibliography

[1] C. Abbe. Drought. *Monthly Weather Review*, 22:323–324, 1894.

[2] P. Ailliot, C. Thompson, and P. Thomson. Space-time modelling of precipitation by using a hidden Markov model and censored Gaussian distributions. *Journal of the Royal Statistical Society: Series C*, 58(3):405–426, 2009.

[3] R. Allan, J. Lindesay, and D. Parker. *El-Niño Southern Oscillation and Climatic Variability*. CSIRO Publishing, Collingwood, Victoria, Australia, 1996.

[4] W. M. Alley. The Palmer Drought Severity Index: Limitations and Assumptions. *Journal of Applied Meteorology*, 23(7):1100–1109, 1984.

[5] E. Alvoni, P. L. Papini, and F. Spizzichino. On a class of transformations of copulas and quasi-copulas. *Fuzzy Sets and Systems*, 160(3):334–343, 2009.

[6] American Meteorological Society. Meteorological drought-Policy statement. *Bulletin of the American Meteorological Society*, 78:847–849, 1997.

[7] K. Ashok, Z. Guan, and T. Yamagata. Influence of the Indian Ocean Dipole on the Australian winter rainfall. *Geophysical Research Letters*, 30(15):1821–1824, 2003.

[8] Australian Bureau of Agricultural and Resource Economics. Drought update: Australian crop and livestock report. Technical report, ABARE, Canberra, Australia, 2006.

[9] Australian Bureau of Meteorology. Bureau of Meteorology Rainfall Districts [online]. Available from: `http://www.bom.gov.au/climate/cdo/about/rain-districts.shtml` [cited 19th June 2009].

[10] Australian Bureau of Meteorology. Climate Glossary: Southern Oscillation Index [online]. Available from: `http://www.bom.gov.au/climate/glossary/soi.shtml` [cited 18th November 2009].

[11] Australian Bureau of Meteorology. Climate variability and change: Time series graphs [online]. Available from: `http://www.bom.gov.au/climate/change/` [cited 19th June 2009].

[12] Australian Bureau of Meteorology. The "Federation Drought", 1895-1902 [online]. Available from: `http://www.bom.gov.au/lam/climate/levelthree/c20thc/drought1.htm` [cited 2nd August 2009].

[13] Australian Bureau of Meteorology. Southern Oscillation Index (SOI) Archives [online]. Available from: `http://www.bom.gov.au/climate/current/soihtm1.shtml` [cited 19th June 2009].

[14] Australian Bureau of Meteorology. The World War II droughts 1937-45 [online]. Available from: `http://www.bom.gov.au/lam/climate/levelthree/c20thc/drought3.htm` [cited 19th June 2009].

[15] Australian Bureau of Statistics. Australia economic indication: Impact of 2003-2004 farm season on Australia production. Technical report, ABS, Canberra, Australia, 2004.

[16] S. K. Behera and T. Yamagata. Influence of the Indian Ocean Dipole on the Southern Oscillation. *Journal of the Meteorological Society of Japan*, 81(1):169–177, 2003.

[17] Richard Black. Climate water threat to millions. Available at http://news.bbc.co.uk/2/hi/science/nature/6068348.stm, October 2006.

[18] W. Breymann, A. Dias, and P. Embrechets. Dependence structures for multivariate high-frequency data in finance. *Quantitative Finance*, 3:1–14, 2003.

[19] S. Cambanis, S. Huang, and G. Simons. On the Theory of Elliptically Contoured Distributions. *Journal of Multivariate Analysis*, 11, 368-385 1981.

[20] C. Chatfield. *The Analysis of Time Series: An Introduction.* Number Sixth Edition in 1. Chapman and Hall, New York, 2003.

[21] D. Chen, M. A. Cane, A. Kaplan, S. E. Zebiak, and D. Huang. Predictability of El-Niño over the past 148 years. *Nature*, 428, 2004.

[22] Umberto Cherubini, Elisa Luciano, and Walter Vecchiato. *Copula methods in finance*. John Wiley, Chichester, UK, 2004.

[23] S. Chib and E. Greenberg. Understanding the Metropolis-Hastings Algorithm. *The American Statistician*, 49(4):327–335, 1995.

[24] F. H. S. Chiew, T. C. Piechota, J. A. Dracup, and T. A. McMahon. El-Niño/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting. *Journal of Hydrology*, 204(1):138–149, 1998.

[25] T. B. Cline. Selecting seasonal streamflow models. *Water Resources Research*, 17(4):975–984, 1981.

[26] I. Cordery. Long range forecasting of low rainfall. *International Journal of Climatology*, 19:463–470, 1999.

[27] C. De Michele and G. Salvadori. A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas. *Journal of Geophysics Research*, 108(D2):4067, 2003.

[28] C. De Michele, G. Salvadori, M. Canossi, A. Petaccia, and R. Rosso. Bivariate statistical approach to check adequacy of dam spillway. *Journal of Hydrologic Engineering*, 10(1):50–57, 2005.

[29] C. De Michele, G. Salvadori, G. Passoni, and R. Vezzoli. A multivariate model of sea storms using copulas. *Coastal Engineering*, 54(10):734–751, 2007.

[30] S. Demarta and A. J. McNeil. The t-copula and related copulas. *International Statistical Review*, 73(1):111–129, 2005.

[31] H. F. Diaz and V. Markgraf. *El-Niño and the Southern Oscillation: Multiscale Variability and Global and Regional Impacts*. Cambridge University Press, 2000.

[32] J. A. Dracup, K. S. Lee, and E. G. Paulson Jr. On the definition of droughts. *Water Resources Research*, 16:297–302, 1980.

[33] W. Drosdowsky. Potential predictability of winter rainfall over southern and eastern Australia using Indian Ocean sea-surface temperature anomalies. *Australian Meteorological Magazine*, 42:1–6, 1993.

[34] F. Durante and C. Sempi. Copula and semicopula transforms. *International journal of mathematics and mathematical sciences*, 2005(4):645–656, 2005.

[35] V. Durrleman, A. Nikeghbali, and T. Roncalli. A simple transformation of copulas. Technical Report 1-15, Groupe Rech. Oper. Credit Lyonnais, 2000.

[36] P. Embrechets, F. Lindskog, and A.J. McNeil. *Handbook of heavy tailed distributions in finance*, chapter Modelling dependence with copulas and applications to applications to risk management, pages 1–48. Elsevier, 2003.

[37] A-C. Favre, S. El Adlouni, L. Perreault, N. Thiemonge, and B. Bobee. Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, 40(W01101), 2004.

[38] D. A. K. Fernando and A. W. Jayawardena. Generation and forecasting of monsoon rainfall data. In *Proceedings of the 20th WEDC Conference*, pages 310–313, Colombo, Sri Lanka, 1994.

[39] C. K. Folland. Relative influences of the Interdecadal Pacific Oscillation and ENSO on the South Pacific Convergence Zone. *Geophysical Research Letters*, 29(13), 2002.

[40] D. M. Frick, D. Bode, and J. D. Salas. Effects of drought on urban water supplies. I: Drought analysis. *Journal of Hydraulic Engineering*, 116(6):733–753, 1990.

[41] D.G. Friedman. *The prediction of long-continuing drought in south and southwest Texas.* Occasional papers in meteorology No. 1. Hartford, Conn. : Travelers Insurance Co., 1957.

[42] C. Genest and A-C. Favre. Everything you always wanted to know about copula modelling but were afraid to ask. *Journal of Hydrologic Engineering*, 12(4):347–368, 2006.

[43] C. Genest and J. Neslehova. A primer on copulas for count data. *Astin Bulletin*, 37(2):475–515, 2007.

[44] J. D. Gibbons. *Nonparametric Statistical Inference.* Marcel Dekker Inc., New York, 1985.

[45] M.H. Glantz. *Currents of Change: Impacts of El-Niño and La-Niña on Climate and Society.* Cambridge, 2001.

[46] S. Grimaldi and F. Serinaldi. Asymmetric copula in multivariate flood frequency. *Advances in Water Resources*, pages 1155–1167, 2006.

[47] N. B. Guttman. Comparing the Palmer Drought Index and Standardized Precipitation Index. *Journal of American Water Resources Association*, 34(1):113–121, 1998.

[48] N. B. Guttman. Accepting the Standardized Precipitation Index: A Calculation Algorithm. *Journal of American Water Resources Association*, 35(2):311–322, 1999.

[49] J. Hassan. *A history of water in modern England and Wales*. Manchester University Press, 2001.

[50] M. J. Hayes, M. D. Svoboda, D. A. Wilhite, and O. V. Vanyarkho. Monitoring the 1996 Drought using the Standardized Precipitation Index. *Bulletin of the American Meteorological Society*, 80:429–438, 1999.

[51] R. R. Heim. A Review of Twentieth-Century Drought Indices Used in the United States. *Bulletin of the American Meteorological Society*, 83(8):1149–1166, 2002.

[52] A. J. Henry. Climatology of the United States. *Bulletin Q. U.S. Weather Bureau Bulletin*, 361:51–58, 1906.

[53] Intergovernmental Panel on Climate Change. IPCC Climate Change 2001 Synthesis report [online]. Available from: `http://www.acer-acre.org/ClimateChangeCD/sec5/511c.htm` [cited 2nd August 2009].

[54] Intergovernmental Panel on Climate Change. IPCC Climate Change 2001 Synthesis report [online]. Available from: `http://www.acer-acre.org/ClimateChangeCD/sec5/511c.htm` [cited 2nd August 2009].

[55] K. Jajuga. *Innovations in Classification, Data Science, and Information Systems*, chapter 6, pages 446–453. Springer Berlin Heidelberg, 2005.

[56] Y. H. Jin, A. Kawamura, K. Jinno, and R. Berndtsson. Detection of ENSO-influence on monthly precipitation in South Korea. *Hydrological Processes*, 19(20):4081–4092, 2005.

[57] H. Joe. *Multivariate models and dependence concepts*. Chapman and Hall, New York, 1997.

[58] D. Jones and G. Weymouth. Technical Report 70: An Australian Monthly rainfall Dataset. Technical report, Bureau of Meteorology, Melbourne, Australia, 1997.

[59] P. R. Julian and H. C. Fritts. On the possibility of quantitatively extending climatic records by means of dendroclimatological analysis. In *Proceedings First Statistical Meteorology Conference*, pages 76–82, Hartford, CT, 1968. American Meteorological Society.

[60] A. Kawamura, S. Eguchi, and K. Jinno. Cross-correlation between southern oscillation index and precipitation/temperature in Fukuoka, Japan. In *Proceedings of Fresh Perspectives Symposium*, Christchurch, New Zealand, 2000.

[61] J. Keyantash and J. Dracup. The Quantification of Drought: An Evaluation of Drought Indices. *American Meteorological Society*, 83:1167–1180, 2002.

[62] A. S. Kiem and S. W. Franks. Multi-decadal variability of drought risk, eastern Australia. *Hydrological Processes*, 18(11):2039–2050, 2004.

[63] A.S. Kiem and S. W. Franks. On the identification of ENSO-induced rainfall and runoff variability: a comparison of methods and indices. *Hydrological Sciences Journal*, 46(5):715–727, 2001.

[64] S. E. Koch, M. DesJardins, and P. J. Kocin. An Interactive Barnes Objective Map Analysis Scheme for Use with Satellite and Conventional Data. *Journal of Climate and Applied Meteorology*, 22(9):1487–1503, 1983.

[65] J. J. Luo, Y. Behera, Y. Masumoto, H. Sakuma, and T. Yamagata. Successful prediction of the consecutive IOD in 2006 and 2007. *Geophysical Research Letters*, 35:L14S02, 2008.

[66] S. Marcovitch. The measure of droughtiness. *Monthly Weather Review*, 58:113, 1930.

[67] R. Mashal, M. Naldi, and A. Zeevi. On the dependence of equity and assest returns. *RISK*, 16:83–87, 2003.

[68] J. L. McBride and N. Nicholls. Seasonal relationships between Australian rainfall and the Southern Oscillation. *Monthly Weather Review*, 111:1998–2004, 1983.

[69] T. B. McKee, N. J. Doesken, and J. Kliest. Drought monitoring with multiple time scales. In *Proceedings of the 9th Conference on Applied Climatology*, pages 233–236, Dallas, Texas, 1995. American Meteorological Society.

[70] T.B. McKee, N. J. Doesken, and J. Kliest. The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*, pages 179–184, Boston, MA, USA, January 1993. American Meteorological Society.

[71] M. J. McPhaden, A. J. Busalacchi, R. Cheney, J. R. Donguy, K. S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G. T. Mitchum, P. P Niiler, J. Picaut, R. W. Reynolds, N. Smith, and K. Takeuchi. The Tropical Ocean-Global Atmosphere observing system: A decade of progress. *Journal of Geophysical Research*, 103(C7):14169–14240, 1998.

[72] A. V. Metcalfe. *Statistics in Civil Engineering*. A Hodder Arnold Publication, London, 1997.

[73] A. K. Mishra and V. R. Desai. Drought forecasting using stochastic models. *Stochastic Environmental Research and Risk Assessment*, 19(5):326–339, 2005.

[74] A. K. Mishra and V. R. Desai. Spatial and temporal drought analysis in the Kansabati river basin, India. *International Journal of River Basin Management*, 3(1):31–41, 2005.

[75] A. K. Mishra and V. R. Desai. Drought forecasting using feed-forward recursive neural network. *Ecological Modelling*, 198(1-2):127–138, 2006.

[76] A. K. Mishra, V. R. Desai, and V. P. Singh. Drought forecasting using a hybrid stochastic and neural network model. *Journal of Hydrologic Engineering*, 12(6):626–638, 2007.

[77] R. Modarres. Streamflow drought time series forecasting. *Stochastic Environmental Research and Risk Assessment*, 21(3):223–233, 2007.

[78] S. Morid, V. Smakhtin, and M. Moghaddash. Comparison of seven meteorological indices for drought monitoring in Iran. *International Journal of Climatology*, 26:971–985, 2006.

[79] P. M. Morillas. A method to obtain new copulas from a given one. *Metrika*, 61(2):169–186, 2005.

[80] T. T. Munger. Graphic method of representing and comparing drought intensities. *Monthly Weather Review*, 44:642–643, 1916.

[81] R. B. Nelsen. *An Introduction to Copulas*. Springer-Verlag, New York, 2nd edition, 2006.

[82] N. Nicholls. The stability of empirical long range forecasting techniques: a case study. *Journal of Applied Meteorology*, 23:143–147, 1984.

[83] N. Nicholls. Sea surface temperatures and Australian winter rainfall. *Journal of Climate*, 2:965–973, 1989.

[84] H.K. Ntale and T. Y. Gan. Drought indices and their application to East Africa. *International Journal of Climatology*, 23:1335–1357, 2003.

[85] J. C. Ochoa-Rivera. Prospecting droughts with stochastic artificial neural networks. *Journal of Hydrology*, 352:174–180, 2008.

[86] Y. Opoku-Ankomah and I. Cordery. Temporal variation of relations between New South Wales rainfall and the southern oscillation. *International Journal of Climatology*, 13:51–64, 1993.

[87] W. C. Palmer. Meteorological drought. Technical Report 45, Weather Bureau, US Department of Commerce, Washington, D. C., 1965.

[88] F. Pearce. Water anger everywhere but no one stops to think [online]. 1995. Available from: `http://www.independent.co.uk/news/uk/water-anger-everywhere-but-no-one-stops-to-think-1598235.html` [cited 29th May 2009].

[89] S.G. Philander. *El-Niño, La-Niña and the Southern Oscillation*, volume 46. Academic Press, 1990.

[90] A. Poulin, D. Huard, A-C. Favre, and S. Pugin. Importance of tail dependence in bivariate frequency analysis. *Journal of Hydrologic Engineering*, 12(4):394–403, 2007.

[91] S. Power, T. Casey, C. Folland, A. Colman, and V. Mehta. Inter-decadal modulation of the impact of ENSO on Australia. *Climate Dynamics*, 15:319–324, 1999.

[92] C. H. B. Priestley. Rainfall-SST associations on the New South Wales coast. *Australian Meteorological Magazine*, 47:15–25, 1964.

[93] C. H. B. Priestley and A. J. Troup. Droughts and wet periods and their association with SST. *Australian Journal of Science*, 29:56–57, 1966.

[94] Queensland Government. The Long Paddock: Climate Management Information for Rural Australia. Available at http://www.longpaddock.qld.gov.au/SeasonalClimateOutlook/SouthernOscillationIndex/SOIDataFiles/index.html.

[95] R Development Core Team. R: A language and environment for statistical computing, 2007. Available from: `http://www.R-project.org`.

[96] W. D. Rase. Volume preserving interpolation of a smooth surface from polygon related data. *Journal of Geographic System*, 3:199–213, 2001.

[97] E. M. Rasmusson and T.H. Carpenter. The relationship between eastern equatorial Pacific sea surface temperatures and rainfall over India and Sri Lanka. *Monthly Weather Review*, 111:517–528, 1983.

[98] J. A. Raynal-Villasenor and J. D. Salas. Multivariate extreme value distributions in hydrological analyses. In *Water for the Future: Hydrology in Perspective*, number 164 in 1, pages 111–119, Washington, D. C., 1987. IAHS Publication.

[99] B. Renard and M. Lang. Use of a gaussian copula for multivariate extreme value analysis: Some case studies in hydrology. *Advances in Water Resources*, 30(4):897–912, 2006.

[100] C. F. Ropelewski and M. S. Halpert. North American precipitation and temperature patterns associated with the El-Niño Southern Oscillation. *Monthly Weather Review*, 114:2352–2362, 1986.

[101] C. F. Ropelewski and M. S. Halpert. Global and regional scale precipitation patterns associated with the El-Niño Southern Oscillation. *Monthly Weather Review*, 115:1606–1626, 1987.

[102] C. F. Ropelewski and M. S. Halpert. Precipitation patterns associated with the high index phase of the Southern Oscillation. *Journal of Climate*, 2(3):268–284, 1989.

[103] C. F. Ropelewski and M. S. Halpert. Quantifying Southern Oscillation-precipitation relationships. *Journal of Climate*, 9(5):1043–1059, 1996.

[104] N. H. Saji, B. N. Goswami, P. N. Vinayachandran, and T. Yamagata. A dipole mode in the tropical Indian Ocean. *Nature*, 401:360–362, 1999.

[105] N. H. Saji and T. Yamagata. Interference of teleconnection patterns generated from the tropical Indian and Pacific oceans. *Climate Research*, 25:151–169, 2003.

[106] G. Salvadori and C. De Michele. Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resources Research*, 40(W112511):1–17, 2004.

[107] G. Salvadori and C. De Michele. On the use of copulas in hydrology: Theory and practice. *Journal of Hydrologic Engineering*, 12(4):369–380, 2007.

[108] B. Sanso and L. Guenni. A Nonstationary Multisite Model for Rainfall. *Journal of the American Statistical Association*, 95(452):1089–1100, 2000.

[109] F. Serinaldi and S. Grimaldi. Fully nested 3-copula: Procedure and Application on hydrological data. *Journal of Hydrologic Engineering*, 12(4):420–430, 2007.

[110] A. Sharma. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 1 - A strategy for system predictor identification. *Journal of Hydrology*, 239:232–239, 2000.

[111] A. Sharma, K. C. Luk, I. Cordery, and U. Lall. Seasonal to interannual rainfall probabilistic forecasts for improved water supply management: Part 2 - Predictor identification of quarterly rainfall using ocean-atmosphere information. *Journal of Hydrology*, 239:240–248, 2000.

[112] D. Shepard. A two dimensional interpolation function fro irregularly data. In *Proceedings ACM National Conference*, pages 517–524, 1968.

[113] J.T. Shiau. Fitting drought duration and severity with two-dimensional copulas. *Water Resources Management*, 20:795–815, 2006.

[114] J. Shukla and D. A. Paolino. The Southern Oscillation and long range forecasting of the summer monsoon rainfall over India. *Monthly Weather Review*, 111:1830–1837, 1983.

[115] A. Sklar. Fonctions de répartitions à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8:229–231, 1959.

[116] D. I. Smith, M. F. Hutchinson, and R. J. McArthur. Australian climate and agricultural drought: Payments and policy. *Drought Network News*, 5(3):11–12, 1993.

[117] PXK Song. Multivariate dispersion models generated from gaussian copula. *Scandinavian Journal of Statistics*, 27(2):305–320, 2000.

[118] R. C. Stone and A. Auliciems. SOI Phase relationships with rainfall in eastern Australia. *International Journal of Climatology*, 12:625–636, 1992.

[119] R.C. Stone, G.L. Hammer, and T. Marcussen. Prediction of global rainfall probabilities using phases of the Southern Oscillation Index. *Nature*, 384:252–255, 1996.

[120] N. A. Streten. Southern hemisphere sea surface temperature variability and apparent associations with Australian rainfall. *Journal of Geophysical Research*, 86:485–497, 1981.

[121] N. A. Streten. Extreme distributions of Australian annual rainfall in relation to sea surface temperature. *Journal of Climatology*, 3:143–153, 1983.

[122] R. Suppiah. Trends in the Southern Oscillation Phenomenon and Australian rainfall and changes in their relationship. *International Journal of Climatology*, 24:269–290, 2004.

[123] Sydney Catchment Authority. Location of warragamba dam [online]. Available from: `http://www.sca.nsw.gov.au/__data/assets/image/0018/1359/DamLocations.jpg` [cited 1st December 2009].

[124] C. W. Thornthwaite. The Climates of North America: According to a New Classification. *Geographical review*, 21:633–655, 1931.

[125] C. W. Thornthwaite. An approach toward a rational classification of climate. *Geographical Review*, 38(1):55–94, 1948.

[126] C. W. Thornthwaite and J. R. Mather. The water budget and its use in irrigation. In *Water-Yearbook of Agriculture 1955*, pages 346–358. U. S. Dept. of Agriculture, 1955.

[127] K. E. Trenberth. The Definition of El-Niño. *Bulletin of the American Meteorological Society*, 78(12):2771–2777, 1997.

[128] UN Atlas of Oceans. Indian Ocean Dipole [online]. Available from: `http://www.oceansatlas.org` [cited 8th June 2009].

[129] United Nations. World Population Prospects: The 2008 Revision [online]. Available from: `http://www.un.org/esa/population/` [cited 2nd August 2009].

[130] United Nations. Trends in the total Migrant Stock: The 2005 Revision [online]. 2006. Available from: `http://www.un.org/esa/population/publications/migration/migration2005.htm` [cited 2nd August 2009].

[131] A. Wahlquist. Longest, hottest drought on record, says Bureau of Meteorology [online]. 2008. Available from: `http://theaustralian.news.com.au/story/0,25197,24478392-11949,00.html` [cited 11th October 2008].

[132] G. T. Walker. Correlation in seasonal variations of weather, VII. *Memorandum of Indian Meteorological Department*, 24:75–131, 1923.

[133] G. T. Walker. World weather IX. *Memorandum of Indian Meteorological Department*, 24:275–332, 1924.

[134] L. A. Walls and A. Bendel. Time series model in reliability. *Reliability Engineering*, 18(239-265), 1987.

[135] G. Wang and H. H. Hendon. Sensitivity of Australian rainfall to inter-El Niño variations. *Journal of Climate*, 20(16):4211–4226, 2007.

[136] D. A. Wilhite and M.H. Glantz. Understanding the drought phenomenon: the role of definitions. *Water International*, 10(3):111–120, 1985.

[137] K. Wolter and M. S. Timlin. Monitoring ENSO in COADS with a seasonally adjusted principal component index. In *Proceedings of the 17th Climate Diagnostics Workshop*, pages 52–57, Norman, Oklahoma, 1993.

[138] S. A. Wooldridge, S. W. Franks, and J. D. Kalma. Hydrological implications of the Southern Oscillation: variability of the rainfall-runoff relationship. *Hydrological Sciences Journal*, 46(1):73–88, 2001.

[139] H. Wu, M. J. Hayes, A. Weiss, and Q. Hu. An evaluation of the standardized precipitation index, the china-Z index and the statistical z-score. *International Journal of Climatology*, 21:745–758, 2001.

[140] J. Yan. Enjoy the Joy of Copulas: with a package copula. *Journal of Statistical Software*, 21(4):1–21, 2007.

[141] V. M. Yevjevich. An objective approach to definitions and investigations of continental hydrologic droughts. Hydrology Paper 23, Colorado State University, Fort Collins, CO, 1967.

[142] S. Yue. The gumbel logistic model for representing a multivariate storm event. *Advances in Water Resources*, 24(2):179–185, 2000.

[143] S. Yue. The gumbel mixed model applied to storm frequency analysis. *Water Resources Management*, 14(5):377–389, 2000.

[144] S. Yue. Joint probability distribution of annual maximum storm peaks and amounts as represented by daily rainfalls. *Hydroscience Journal*, 45(2):315–326, 2000.

[145] S. Yue. A bivariate extreme value distribution applied to flood frequency analysis. *Nordic Hydrology*, 32(1):49–64, 2001.

[146] S. Yue. A bivariate gamma distribution for use in multivariate flood frequency analysis. *Hydrological Processes*, 15(6):1033–1045, 2001.

[147] S. Yue. The bivariate lognormal distribution to model a multivariate flood episode. *Hydrological Processes*, 14:2575–2588, 2001.

[148] K. Yurekli, K. Kurunc, and F. Ozturk. Application of linear stochastic models to monthly flow data of Kelkit Stream. *Ecological modelling*, 183(67-75), 2005.

[149] E. Zelenhastic and A. Salvai. A method of streamflow drought analysis. *Water Resources Research*, 23(1):156–168, 1987.

[150] L. Zhang and V. P. Singh. Bivariate flood frequency analysis using the copula method. *Journal of Hydrologic Engineering*, 11(2):150–164, 2006.

[151] L. Zhang and V. P. Singh. Bivariate rainfall frequency distributions using archimedean copulas. *Journal of Hydrology*, 332(1-2):93–109, 2006.

[152] L. Zhang and V. P. Singh. Gumbel-hougaard copula for trivariate rainfall frequency analysis. *Journal of Hydrologic Engineering*, 12(4):409–419, 2007.

[153] L. Zhang and V. P. Singh. Trivariate flood frequency analysis using the gumbel-hougaard copula. *Journal of Hydrologic Engineering*, 12(4):431–439, 2007.

[154] X. G. Zhang and T. M. Casey. Long-term variations in the southern oscillation and relationships with australian rainfall. *Australian Meteorological Magazine*, 40:211–225, 1992.