USING LOCAL AND HISTORICAL DATA TO ENHANCE UNDERSTANDING OF SPATIAL AND TEMPORAL RAINFALL PATTERNS

DEREK YATES

A thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

2014

Faculty of Agriculture and Environment

The University of Sydney

New South Wales

Australia

Certificate of Originality

This thesis is submitted to the University of Sydney in fulfilment of the requirements of the degree of Doctor of Philosophy.

The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text.

I hereby declare that I have not submitted this material, either in full or part, for a degree at this or any other institution.

Wednesday, 29 January 2014

Abstract

Farmers face uncertainty in their businesses from many factors, but rainfall is a key determinant of both the nature of the production system and variation in financial returns. Currently, various weather forecasting services are available from the Australian Bureau of Meteorology (BoM) based on about 7000 stations covering all of Australia. Seasonal Climate Forecasts are seen as another tool that can help to improve farm productivity.

It is well known that many farmers keep their own rainfall records, and likely that the farmers have a high degree of confidence in their own records. Australian Bureau of Statistics figures indicate that there were possibly 7000 grain related 'agricultural businesses' in NSW alone in 2009/10 indicating that there is the potential to increase data density by up to an order of magnitude.

This project is part of a broader study to improve rainfall predictions for grain farmers using data collected locally to the users (crowd sourcing). The data is collected directly on farm, and from other sources which may be available. The focus is on the historical data, its collection and analysis, in terms of discerning patterns in time and space which may help provide a local framework, within which coarser scale forecasts can be interpreted and understood. Data will be stored on secure database systems at the University of Sydney.

Results indicate that farm data does provide more local detail, temporally and spatially. Deficit and surplus analysis demonstrates the predictive capacity of the local temporal data, despite limited data precluding the definition of ideal criteria and parameters for predictive 'similar year' selection.

The spatial data demonstrates quantifiable site specific differences from institutional data. Testing across more climate types may allow these differences to be defined within and across regions.

Tests for an indicator time period show that farm rainfall in the early part of the growing season (April and May) may indeed be indicative of seasonal condtions, while more data is needed to confirm this. The use of southern oscillation life cycle information to select appropriate years considerably improved the relationships revealed, with a doubling of relationship strength across all climatic types, although the strength of the relationships differed across the climatic types, and the strongest relationships were split between the months of April and May. More extensive analysis, with more data across more BoM districts (and therefore climate classes) will be required to confirm this conclusion, but it appears that farm rainfall records and SOI information can provide an indicator time period to help farmers interpret, refine and utilise seasonal forecasts.

Acknowledgements

I would like to acknowledge the financial support for this project provided by the Australian Research Council and GrainGrowers.

I would also like to acknowledge :

- the invaluable support of my family during this project, my wife Rachel and children Darcy, Oliver and Charlotte.
- the support, encouragement and enthusiasm of my supervisors, Associate Professor Willem Vervoort, Professor Alex McBratney and Associate Professor Budiman Minasny.

Contents

List of figures

List of tables

1 Introduction

From the earliest days of European settlement of Australia (1788), wheat has been an essential part of Australian culture. By 1790 wheat was being cultivated at Rose Hill, about 25km west of the main settlement of Sydney, as well as being grown in and around Sydney on public land Tench (1793). Climatic variability was also apparent from early days. Tench notes that the climate was 'changeable beyond any other I have ever heard of'. Rainfall proved lower in amount and uniformity than might have been hoped for, with Tench describing it as 'not a sufficiency, but torrents of water sometimes fall'.

Over the next fifty or so years, wheat production spread throughout the various Australian settlements, with Tasmania supplying New South Wales' needs in the early nineteenth century. There was one early year with a total net export balance for Australia, in 1833, but it was not until 1845 that wheat exports became a regular event. After federation (1901) wheat exports became common and were increasing Dunsdorfs (1957). The first official mention of wheat exports appears to have been in 1844 (Gipps 1846), with a request for minimal import duties into the United Kingdom by George Gipps, then governor of NSW. Wheat cultivation continued to expand, and during 2009/10 approximately 13.9 million hectares was planted to wheat in Australia, on about 25,000 properties, with 21.8 million tonnes being harvested (ABS 2012). From this production, about 14.7 million tonnes (67%), worth about $AUD 3,675 million (estimates from ABARE (2012) figures) was exported.

Farmers face uncertainty in their businesses from many factors, but rainfall 'is a key determinant of both the nature of the production system and variation in financial returns.' (Hammer, Hansen et al. 2001). All farmers need to plan their growing seasons carefully, with non irrigated grain growers making decisions about what to plant, when to plant, whether to fertilise and if so when, amongst many others. Numerous authors have described the intuitively obvious importance of rainfall to crop growth, for example 'The water environment of crop ecosystems has long been recognized as a major determinant of yield' Nix and Fitzpatrick (1969). Rainfall forecasts are, therefore, clearly a vital part of a farmer's planning apparatus, especially given that Australia is the driest inhabited continent, that large seasonal variability exists in both rainfall and temperature(ABS 2012), and that there is evidence that rainfall distribution patterns are changing (SOE 2011).

Currently, various weather forecasting services are available from the Australian Bureau of Meteorology (BoM). For example, quality controlled weather forecasts in graphical or map form out to 7 days are provided by the Bureau from the website using the 'Forecast Explorer'[1]. However, the Forecast Explorer uses a six kilometre square interpolation grid (3600ha) based on about 7000 stations (Jones, Wang et al. 2009) covering all of Australia. Even if all the stations were in NSW, this would only represent one recording station per 114km² (114000ha). Seasonal Climate Forecasts (SCF's) are seen as another tool that can help to improve farm productivity (Hayman, Crean et al. 2007), but SCF's are provided by the BoM on a district basis (the BoM divides Australia into forecast districts, for example NSW is divided into seventeen districts), with a fairly coarse resolution.

It is well known that many farmers keep their own rainfall records, and likely that the farmers have a high degree of confidence in their own records. It is more likely that 'no rain' days will be missed than rainfall events, considering the value of the rainfall to the farmers. Perhaps not surprisingly, some famers (graziers) preferred to use their own records rather than rely on seasonal forecasting, and intended to test forecasts against them (Hayman, Crean et al. 2007).

Australian Bureau of Statistics (ABS 2012) figures indicate that there were possibly 7000 grain related 'agricultural businesses' in NSW alone in 2009/10 (4971 grain / sheep or beef, and 3190 'Other grain growing'). According to the ABS figures 13530000 hectares were planted to wheat in NSW in 2009/10. Using the estimate of 7000 businesses from the ABS, and assuming 15% participation from businesses, the data coverage (sample area / station) would be approximately 12885ha, as compared to the 114000 ha noted above for BoM data. Clearly, the more sites providing data the better the data density (stations/ha) becomes, but this simple exercise shows that there is the potential to increase data density by an order of magnitude, which would improve the spatial resolution of rainfall forecasts, or the application of forecasts to the local (participant's) environment. This project will be collecting daily rainfall data, and it is intended that analyses will be performed, where appropriate, using daily data. Once again, this will allow greater detail to be drawn from forecasts, assisting management decisions on a suitable time scale, such as weekly or daily.

From this background, this project is part of a broader study (hereafter called the program) to improve rainfall predictions for grain farmers using data collected locally to the users (crowd

---

[1] 'Forecast Explorer subsequently replaced by 'MetEye' see - http://www.bom.gov.au/australia/meteye/. MetEye alos uses gridded data with a three or six kilometre grid, depending on the size of the state

sourcing or citizen science). Because the project is involved with the improved use of on farm or local rainfall data, methods of rainfall analysis associated with irrigated grain growing will not be examined. The project data will mainly be collected directly on farm, and other sources which may be available will be investigated. This project is focussed on the historical data, its collection and analysis in terms of discerning patterns in time and space which may help provide a local framework, within which coarser scale forecasts can be interpreted and understood. It is hoped that the farm historical rainfall records will display a predictive capacity by relating current to previous scenarios, which will add to farmers' instruments for seasonal management and decision making.

Farmer data will be accepted 'as is', with only basic error checking and unit conversions applied as necessary. As noted above, it is believed that farmer data will record rainfall events with reasonable accuracy, and one goal is to demonstrate the value of this data, not discourage famers on the basis that their collection techniques may be imperfect. Clearly, error checking will be required before analyses are performed using farmer data, so that missing data or data that is wildly different from averages can be corrected or excluded, but it is hoped that once communication with the farmer has been established, these errors can be minimised or eliminated.

Farmers who provide data to the project (project participants) will be authorized to access a web based interactive predictive system including local data and data specific to their property, under the broader program. This project will provide analysis of historical data trends, defining the relationship of the participant's property and data to regional conditions, developing and demonstrating whatever predictive or interpretive capacity the farm rainfall data contains.

Crop yield data will not be collected during this project, despite the obvious desirability of this data. It is hoped that this project will demonstrate the value of the farm rainfall data, paving the way for future work to investigate the relationship between any parameters developed and farm crop yields.

Certain climate change effects, such as warming of the climate system, are now considered unequivocal, but analysing for changes in the frequency and intensity of precipitation is more difficult than for climatic averages (IPCC 2007)[2].

[2] Physical data for the 5th IPCC report, dated 2013, available in draft form, was not available for citing or quotation at the time of wrtitng

The IPCC report states that:

- Increases in the amount of precipitation are *very likely*[3] in high-latitudes, while decreases are *likely* in most subtropical land regions … continuing observed patterns in recent trends.
- An increase in the frequency of heavy precipitation events over most areas is *Very likely*
- An increase in the area affected by droughts is *Likely*

This project is investigating trends and relationships in historic farm rainfall data, mostly on seasonal or annual time scales, and so will not attempt to include climate change effects on the data.

An important partner in this project, including contributing to funding, is Grain Growers Limited (GrainGrowers). GrainGrowers is Australia's member-based, grain producer organisation, and its membership will provide the basis for contact and initial data gathering operations. All data, while remaining the property of the participants and GrainGrowers, will be stored on secure database systems at the University of Sydney.

1.1 Aims

1. To develop a data gathering, secure storage, primary analysis and reporting system for rainfall data already collected and being collected by farmers, to simplify and encourage user participation in the data gathering process, and to develop new methods of displaying user rainfall data which relate to user needs and interests.

2. By exploring patterns in time and space in historic rainfall data, and the relationship between farm data and other available data, test the usefulness of the farm data, and test whether the farm data can have value in a predictive or interpretive role, and therefore aid in farm management decisions.

3. To test the idea of an indicator period or season, which can be used to predict annual or seasonal rainfall, and whether such indicator periods persist across regions.

[3] The terms *Likely* and *Very Likely* are part of the IPCC uncertainly assessment, expressing probabilities of occurrence of >66% and >90% respectively.

4. To test for effects on such an indicator of seasonal or annual rainfalls, of external influences and climate drivers.

1.2 Hypotheses

1. Farm rainfall historical data is a valuable, existing resource, that can be collected and used to enhance local understanding of spatial and temporal patterns in rainfall, and rainfall variability

2. Incorporation of community collected farm based rainfall data will increase the spatial and temporal detail in climate data, despite inherent uncertainty

3. Incorporation of farm based rainfall data will increase the ability to predict spatial and temporal patterns in climate data, despite inherent uncertainty

2 Literature review

The aims of this project focus on the collection and analysis of historical farm daily rainfall data, to allow users to improve their understanding and utilization of existing rainfall forecast information.

This literature review is an attempt to follow the developing use of rainfall data in Australia, especially in relation to wheat crops, the availability and use of data at different intervals, and the issue of single parameter indices for rainfall or yield prediction. The review will therefore focus on the direct use of rainfall data as an index for, or in relation to other climatic parameters or forecasts and grain production. Many authors have covered issues such as rainfall trends in Australia which will not be re-examined here, for example Cornish (1949); (1950), Nicholls and Lavery (1992), Gallant, Hennessy et al. (2007), Chowdhury and Beecham (2010), and standard texts such as Maidment (1993), Dingman (2002) detail rainfall data collection and methods of analysis.

2.1 Grain growing and the use of climate information

According to the FAO (Food and Agriculture Organization of the United Nations), wheat is the most important grain food crop consumed by humans (FAO 2002). It has been grown for some 8000 years, and world production is around 650 million tonnes per annum (MTpa). Production was 655.8 MT in 2008 (FAO 2009). Australian production averages around 20 MTpa, with 21.8 MT being produced in 2009/10 (ABARE 2012).

Standard texts indicate that the world’s wheat is predominantly grown in semi-arid and sub humid climates (Nix 1975). Nix goes on to state that for Australian conditions, the water regime plays a dominant role in defining boundary conditions for successful wheat cultivation. He also notes that temperature sets further limits on both the duration and timing of the crop cycle.

The stages of development of a wheat plant are commonly grouped into growth stages defined as germination to emergence (E); growth stage 1 (GS1) from emergence to double ridge; growth stage 2 (GS2) from double ridge to anthesis (flowering); and growth stage 3 (GS3), which includes the grainfilling period, from anthesis[4] to maturity. The time span of these development stages depends on genotype, temperature, day-length and sowing date (FAO 2002). The environmental factor which

[4] Anthesis: blossom, full bloom OED (2013). Oxford English Dicitnary. Oxford, Oxford University Press.

controls the sowing date at any location is temperature (Slafer and Rawson 1994), to satisfy any vernalization[5] responses. However, a moisture content of 35 to 45 percent by weight in the grain is also required for germination (FAO 2002).

Possibly the earliest connection between rainfall and agriculture in Australia, including wheat growing, was developed by G. W. Goyder, drawing on his knowledge of the country gained through extensive field assessment of the South Australian region, (Goyder 1857). In response to particularly severe droughts in 1864 and 1865, the South Australian parliament instructed Goyder, the then surveyor general, to delineate the line between areas in drought and areas where rainfall had continued (Meinig 1961). The resulting boundary, drawn by Goyder in 1865 became known as Goyder's line. There seems to be some confusion about the actual interpretation of the line, as being related to rainfall, vegetation type and condition, or some combination of both (Meinig 1961), (Sheldrick 2010), (Taylor 1918). According to Sheldrick, Goyder always intended the line to represent the limit of reliable rainfall, inside which small farms could consistently expect successful crops. It appears that Goyder understood how various vegetation types responded to drought conditions (Sheldrick 2010) and used this knowledge, extensive field inspections, and no doubt the small amount of rainfall data available (Bureau of Meteorology records indicate that there were thirty-two stations operating in South Australia at the time) to define the boundary. Whatever Goyder's exact method of defining the line, it became widely accepted as the limit of reliable agriculture in South Australia, and by the early twentieth century was also accepted as marking the ten inch (254 mm) isohyet.

Other early work relating climatic information to crops, also starting in the late 1800's, involved workers examining the concept of periodicity in various natural phenomena. In a review article, Hooker (1921) refers to work by Professor W. Stanley Jevons, in 1875, investigating the relation between wheat crops and sun spots, and refers to Jevons as among the best known of the earlier writers.

Hooker noted that the increasing availability of crop statistics had allowed some work relating crop yields to preceding meteorological conditions. He quotes the work of '[Sir] Rawson W. Rawson'

---

[5] Veranlization: The technique of exposing seeds, young plants, etc., to low temperatures in order to hasten subsequent flowering; also, the natural process induced by cold weather which this technique imitates. (OED 2013)

(Rawson 1874), the then Governor of Barbados, as providing some of the first equations for estimating a crop in advance. Rawson was using rainfall data to predict the sugar crop on Barbados.

Hooker claims to have (possibly) pioneered the use of correlation in examining the relationship between weather and crops, referring to work in 1908, relating wheat to rainfall and temperature in England.

Hooker refers to a study by Wallén (1917), examining correlation coefficients between rainfall and temperature and wheat, barley, oats and rye grown in Sweden during the years 1881 to 1910. Hooker describes Wallén's work as the most up to date and complete of the methods of research into relationships between the weather and crops.

The first maps of rainfall variability and reliability in Australia were produced by G W Taylor (Taylor 1918), in a study into the influence of rainfall on development and production in Australia. In summarising Taylor's work, Visher (1919) noted that the degree of variability (he described the rainfall series as erratic) questioned the value (biological significance) of rainfall averages. Taylor maintained that rainfall differences are more significant than temperature differences for vegetation in Australia.

Other early work focused on the fact that maps of vegetation species showed a zonal distribution. Specific species are dominant at the centre of such zones, with their dominance diminishing with distance from the centre towards other zones (Transeau 1905). It was understood that while specific types of plants did not necessarily originate at specific points and spread outwards from these points, certain types of vegetation flourished under particular climatic conditions, at specific locations. In the words of Transeau (1905) 'the complex of climatic factors most favourable to the development of this type of vegetation is here localized'. He labelled locations which typified a set of vegetation habitats and climatic conditions centres of distribution. The zones defined by the centres were understood to be dynamic in time and space, and in terms of species distribution. In an attempt to understand the spatial distribution of such centres of distribution, Trumble (1939) investigated the relationship between the locations of the centres and climatic variables. He indicated that while the nature and distribution of agricultural plants was influenced by day length, temperature, rainfall and evaporation, the period of growth was a function of moisture availability.

Clearly, wheat production relies on a sequence of optimal climatic conditions, and the environmental factors which primarily influence the timing and success of the crop are temperature and moisture availability. The use of climate information in grain growing in Australia has developed from an understanding of limitations imposed by climatic conditions (Goyder 1857), (Taylor 1915) to sophisticated seasonal forecasts and predictions supplied by the Bureau of Meteorology[6]. The literature involving this development is further examined below, focussing on the use of seasonal data and the search for single index and/or seasonal growth indicators, specifically relating to wheat.

2.2 Use of growing season or periodic data

A season can be defined as any of the periods into which the year is divided by the earth's changing position in regard to the sun (OED 2013). Traditionally, the seasons refer to the four periods; Spring, Summer, Winter and Autumn, which are related to the movement of the sun between solstices and equinox. There are, however, other seasons relevant to different cultures and environments, for example it is quite common to refer to wet and dry seasons in tropical areas, and the Australian Bureau of Meteorology (BoM) gives examples of indigenous Australian calendars on its web site. It is common in literature concerning agricultural systems to refer to a growing season, which usually relates to the period from sowing to harvest for a particular crop, which may cross other seasonal boundaries.

Wheat has been grown in Australia since the days of early settlement, for example yield records are available for New South Wales (NSW) from 1792 (Dunsdorfs 1956). However, there was no overview of the distribution of wheat, or indeed any agricultural products, until 1915 (Taylor 1915). In his study to correlate climate with the distribution of cattle, sheep and wheat, Griffith Taylor notes that the chief controller of these commodities is 'undoubtedly' rainfall. Griffith Taylor is mainly interested in absolute limits for the commodities, in terms of isohyets defining existing and likely boundaries to production, and divides the sites into winter and summer rainfall categories. He notes the importance of rainfall during the growing season, which he defines as from April to October, initially in reference to Western Australian conditions. Griffith Taylor also notes the significance of topographic factors, making this work one of the earliest to document the significance of seasonal and spatial climatic variation to wheat production.

Prescott (1934), reviewed the relationship between climatic conditions and the distribution of plants, animals and soils, in the light of finding the most efficient single value function to define such

[6] http://reg.bom.gov.au/climate/data/

distributions. Prescott acknowledged that mean annual temperature and rainfall values were the most common single value indicators, but noted that these values in themselves were not widely applicable. Prescott examined the use of annual values, and what he termed rainfall efficiency, that is the influence of rising temperatures in reducing rainfall effectiveness. Prescott notes that some workers (Perrin 1931, Andrews and Maize 1933 as quoted by Prescott 1934) used monthly values of a precipitation – temperature ratio when discussing vegetation types and aridity. Prescott concluded that the most promising approach was one based on the Meyer ratio (Meyer 1926, as quoted by Prescott 1934) of precipitation to saturation deficit of atmospheric humidity (P/SD) using annual values, which he claimed related well to general vegetation zone boundaries.

Davidson (1934), studying the distribution of insects, concluded that a monthly precipitation – evaporation index was a useful single factor index which may be valuable in defining aridity in Australia. Davidson related monthly values of saturation deficit to free water evaporation, and introduced the notion that months and areas could be defined for which rainfall exceeded evaporation. He developed maps dividing Australia into bioclimatic zones using evaporation factors, and defined the Precipitation to Evaporation (P/E) ratio of 0.5 as a critical value.

Davidson (1936) stated that mean annual temperatures have little value in an ecological sense, and concluded that the mean monthly precipitation – evaporation ratio (P/E), is a useful index for mapping moisture zones in Australia. He goes on to further define P/E = 0.5 as the value below which adequate moisture will not be available for general plant growth at the soil surface. He also describes the months with P/E greater than 0.5 as the growing period , and months with P/E less than 0.5 as the dormant period.

Prescott (1936), referred to earlier work (Prescott 1934, Davidson 1936), as having developed the idea of the usefulness of monthly values. Prescott continued his work examining the Meyer ratio, and concluded that monthly values of the ratio (precipitation to saturation deficit of atmospheric humidity, P/SD) provided a good estimate of aridity. A value of the ratio of 5 was sufficient to maintain the moisture content of the surface soil at about wilting point. Prescott also noted that a value of this ratio of 5 is roughly equivalent to a P/E ratio of 0.3.

Trumble (1937) noted that the annual rainfall was the most frequently quoted climatic index for agricultural purposes, and that the incidence, reliability and effectiveness of the rainfall were recognized as important but poorly defined. He also noted that the realization of the significance of rain during the growing period had led to the arbitrary choice of rainfall during April to October or

April to November as a measure of the seasonal precipitation. Trumble reported that little investigation had occurred into annual rainfall variability. Cornish (1936) had shown that while the annual rainfall for a specific site may be statistically unchanged, there were quantifiable variations in the pattern of rainfall within the years under investigation. Trumble (1937) claimed that Cornish's work and widespread recognition of within season variation indicated that the mean annual precipitation was not a satisfactory guide to local climatic variation, and continued his work based on monthly rainfall and evaporation data.

Wark (1941) credited Trumble (1937, 1939) with developing the concept of a period of 'influential rainfall', being that period when monthly rainfall exceeded one-third the monthly evaporation. He went on to examine rainfall data for fifty years in South Australia, investigating the frequency with which favourable conditions (of rainfall) occurred in different areas of the state. He concluded that the technique, based on mean monthly rainfall and evaporation, was useful in determining the suitability of areas for specific agricultural activities.

Prescott (1949), in examining soil leaching, noted that improved measurement techniques had shown that simple P/E or P/SD ratios were not adequate for defining climatic or soil zones. However, probably due to the time scale of his (soil formation) study, he persisted with annual average values of precipitation and evaporation. He concluded that the most efficient single value climatic index was of the form $P/E^{\,m}$, with m being an empirical constant varying from 0.67 to 0.80.

Cornish (1949) examined trends in wheat yields in South Australia between 1896 and 1941. He states that monthly rainfall data constituted the most convenient and available form for immediate use (at the time). Cornish noted that Trumble's (1937) effective rainfall periods of five and seven months defined the areas of successful wheat cultivation, and used the period from April to November inclusive as covering most options for seeding and harvest. He then decided that the four periods: April and May; June, July and August; September and October; and November, could be defined as rainfall variates allowing for seasonal variation. Continuing his work regarding yield, Cornish (1950) noted that while many workers were seeking critical periods or months which would most strongly influence wheat yields, he could not justify any such conclusion, He claimed that rainfall has a strong influence on yield throughout a season, and that its influence could not be confined to specific, short intervals of time.

Trumble (1952), in a study of grassland agronomy, once again noted that climatic data were best used as sequential calendar months grouped into units. Trumble noted that while evaporation values varied little from year to year, rainfall (at Adelaide) varied widely.

By this period (1950's) it is apparent that monthly time intervals are the most widely used in examining rainfall data, even if technological advances meant that higher frequency data was available. Modelling relationships between yield and climatic conditions was still based on monthly data, or periods ranging over some weeks, even if data was aggregated to create these periods. For example Millington (1961) relates yield to rainfall in the month following seeding.

Slatyer (1962) conducted a study into the climate of the Alice Springs area, with largely empirical data. Slatyer had access to daily rainfall records from the BoM. He recognized the significance of rainfall variation over time, and notes the importance of sporadic rainfall to plant growth. While Slatyer developed comparative criteria based on rainfall over daily and weekly periods, he also reported initial and continuing carryover, or effective rainfall as monthly values.

The use of daily data seems to have been more common in what could be termed process models. For example Baier and Robertson (1966) used daily rainfall data in a model to estimate soil moisture conditions.

Fitzpatrick, Slatyer et al. (1967), investigating periodic plant growth, aggregated daily rainfall data in five day 'pentads', which they deemed as appropriate to their study, allowing more realistic modelling of evaporative effects in an arid environment.

Fitzpatrick and Nix (1969) again used Bureau of Meteorology (BoM) daily rainfall records in work developing a soil water model. However, they describe other available data as 'meagre', and use weekly time steps for the model.

Nix and Fitzpatrick (1969) used the model from Fitzpatrick and Nix (1969) to develop an index of water stress for wheat and sorghum, again using daily rainfall data aggregated into weekly model time steps.

Regional yield or plant breeding studies still required the use of averaged seasonal or monthly data Nix and Fitzpatrick (1969), Boyd, Goodchild et al. (1976), French and Schultz (1984), Wynen (1984), Stephens, Walker et al. (1994), Stephens and Lyons (1998). Boyd, Goodchild et al. (1976) claimed that plant breeders could only expect a generalized understanding of the growing-season conditions. They therefore used average monthly rainfall data, grouped into periods such as: annual; May to

October; when rainfall was likely; and the growing season. Fitzpatrick and Nix (1970) modelling plant responses to light, temperature and water stress, claimed that monthly values only provided low resolution information of seasonal trends, and so transformed this data into long term mean weekly values.

By the late 1960's and early 1970's, continuous rainfall data was being used; Gangopadhyaya and Sarker (1965), Fitzpatrick and Nix (1969), Taylor and Gilmour (1971), Taylor, Storrier et al. (1974), Seif and Pederson (1978), but in most cases the rainfall data was aggregated: into weeks (Nix and Fitzpatrick 1969); specified periods such as 'Fallow rain' (Seif 1978); or seasonal rain (Taylor 1974). As Gangopadhyaya and Sarker (1965) noted, 'Rainfall is essentially a discontinuous function in point of time', and so it was being treated by aggregation into various time units which will, it was hoped, relate to the development of particular crops.

By the 1980's, the availability of sufficient data, and possibly technological advances, had encouraged workers to experiment with continuous data analyses; Doraiswamy and Thompson (1982), Boer, Fletcher et al. (1993). Doraiswamy and Thompson developed a crop (wheat) moisture stress index for use in crop phenology[7] modelling using daily data. McMahon (1983), investigating the usefulness of wheat crop modelling, declared that daily climate data should be used by such models, and that the time step for crop growth modelling should not be less than one day. McMahon does not stipulate model output periods, but suggest that they should relate to field scale issues. Boer, Fletcher et al. (1993), examining rainfall patterns in time and space in wheat growing areas in Australia, used monthly aggregated daily data.

Studies into estimating evaporation, water use efficiency and crop coefficients for water balance and crop water use models; French and Schultz (1984), Choudhury, Ahmed et al. (1994), Allen, Clemmens et al. (2005), Allen, Pereira et al. (2005), Allen, Pruitt et al. (2005), Burt, Mutziger et al. (2005). Temesgen, Eching et al. (2005) used daily and sub daily time steps and rainfall data. However, many of these works; French and Schultz (1984), Allen, Clemmens et al. (2005) also report monthly aggregated results to facilitate evaluations over differing crops and areas. While investigations into the nature of rainfall used daily (Sadras, Roget et al. 2002) and sub daily down to minute time steps

[7] Phenology : The field of science concerned with cyclic and seasonal natural phenomena, esp. in relation to climate and plant and animal life. OED 2013

(Peters and Christensen 2002), authors reverted to monthly and seasonal time periods when the studies were crop related (Sadras and Rodriguez 2007).

The trend to use aggregated data to allow relationships to growing seasons or other periods of interest seems to have become the norm by the 1990's and continues up to the time of writing; Yuan, Luo et al. (2004), Pan, Wood et al. (2008), Robinson, Silburn et al. (2010), Sadras and Rodriguez (2007). Hooker (1921) encapsulated this method quite succinctly, by noting that while daily records were important to discover the 'statistics of the most favourable conditions', the 'totality of the crop' related to some aggregate of the 'various atmospheric elements'. It may be timely to investigate the relationship between crops and single index climatic data on a continuous data basis. While the end of season yield will be the end focus, it may be that variability in time and space cannot be properly examined on a seasonal basis. Thornthwaite (1948), in a study into climatic classification, stated that 'In the present study, vegetation is regarded as a physical mechanism by means of which water is transported from the soil to the atmosphere; it is the machinery of evaporation as the cloud is the machinery of precipitation.' While the behaviour of vegetation is probably not as dynamic as that of clouds, critical events will certainly be occurring on a diurnal rather than seasonal scale.

2.3 Seasonal growth indicators

2.3.1 Crop (wheat) climatic constraints

Standard texts such as 'Australian Field Crops' (Lazenby and Matheson 1975) indicate that water is the key factor in defining boundaries for wheat cultivation throughout the world's wheat growing areas, with temperature limits further refining such boundaries.

As a result, the areas most suited to dryland (non irrigated) wheat growing in Australia, are in southern and eastern Australia, with a winter (May to October) seasonal rainfall component (Nix 1975). Other areas are considered too hot and wet or too dry. Many efforts have been made to define the limits of wheat production, which, according to Nix (1975) merely outlined areas of existing wheat cultivation at the in time.

Historically, with regard to indices relating to the water availability during the crop cycle, much attention has been paid to the availability of water from preceding periods through fallowing techniques (Nix, 1975). Clearly, the soil moisture status, as influenced by preseason rainfall as well as rainfall during the cropping season is an important factor in successful crop production.

The seasonal water regime in Australia, by way of absolute limits in water supply, defines the boundaries of wheat production. Within these boundaries, successful wheat production relies on matching growth and development patterns to the following water and temperature constraints (Nix 1975):

- the timing of sowing rains
- the duration of the midwinter depression in temperature and radiation values
- the timing of the earliest safe ear emergence as determined by frost conditions, and
- the rapid increase in temperature and evaporation rate during the late spring and early summer months

The seasonal temperature regime is intuitively less variable than the rainfall regime. Seasonal temperatures are controlled by incoming solar radiation, which is much more constant than the atmospheric cloud system controlling rainfall. Nix (1975) describes the water regime as being the 'largest random component of the crop system.'

2.3.2 Crop climatic indices

Ideally, one single indicator would encapsulate the effects of all climatic and soil variability. At the stage of writing, no such indicator has been developed, but much work has been done attempting to define the relationship between climate and plant production.

Early work focused on the perceived zonal distribution of vegetation species. Transeau (1905) defined centres of distribution, as previously discussed.

Examinations of monthly, seasonal and annual temperature and rainfall maps indicated that neither factor alone showed spatial distributions which matched the centres. The knowledge that transpiration was connected with plant adaptation lead Transeau (1905) to suggest that the ratio of rainfall (or precipitation, P) to evaporation (E) from a free water surface would provide a method of mapping which would involve several climatic factors. This ratio became known as the Transeau ratio. Transeau claimed that the strong relationship between this ratio and plant distribution was explained by the fact that 'such ratios involve four climatic factors which are of the greatest importance to plant life, *viz* temperature, relative humidity, wind velocity and rainfall.' The influence of other factors, such as soil or topography, was seen as being manifest in the arrangement of species within and between areas primarily controlled by climatic factors.

Other early work involving agricultural production (Taylor 1915), or yield prediction (Hooker 1921) focussed on rainfall as the most relevant single parameter.

The Transeau and Meyer ratios were to become widely accepted as useful single value climatic indices relating rainfall to agriculture and plant ecology (Prescott 1946).

Prescott (1934) noted that Transeau (1905) first studied climatic factors involving evaporation, and described how subsequent workers (for example Szymkiewizc 1923, 1925 as quoted by Prescott) had further examined the calculation of evaporation and its reliance on saturation deficit, that is the difference between vapour pressure at an evaporating surface and the vapour pressure of the surrounding air. He concluded that the most promising approach was one based on the Meyer ratio. Prescott (1936) seems to have introduced the notion of rainfall efficiency, in his studies into climatic influences of Australian bioclimatic zones, to describe rainfall that was available for plant production, after evaporative losses had been accommodated.

Trumble (1937) examining climatic control of agriculture in South Australia, concluded that a P/E ratio of 0.3, with evaporation being measured from an exposed standard tank, could be used to define the time interval over which rainfall influences the growth of plants. He also noted that this roughly coincided with a Meyer ratio of P/SD = 5. He referred to this period of effective rainfall (during which the soil moisture level will be at or above wilting point) as the period of influential rain. This period provided a measure of the effective rainfall season, which does not necessarily coincide with the full growing season due to the possibility of subsoil water use by plants. Trumble also argued that evaporation measured from a standard tank compared favourably with saturation deficit, representing a full day's conditions as opposed to an instantaneous reading, being the best single parameter related to transpiration or soil evaporation, and being expressed in the same units as rainfall. Trumble asserted that the amount of rainfall during the influential rain period, being, by definition, effective rainfall (for plants), may prove a good instrument for climate comparison for agricultural purposes.

Prescott (1938) examined the subject of indices in agricultural climatology, and noted that indices involving monthly averages of rainfall, temperature and saturation deficit would be of some value. He concluded that it was too early to expect that a single universal index would be discovered, and suggested that each agroclimatological problem should be studied individually to determine the most suitable index to meet its needs.

Wark (1941) used Trumble's Influential Rainfall Period (when monthly $P > 0.3E$) to divide South Australia into climatic zones. He concluded that the method was of value in determining the suitability of areas to specific crops, and noted that there was a need for studies into variability (of rainfall and evaporation) within each (defined) zone.

While the Transeau ratio was widely accepted, it became clear over time that the ratio would not be constant for all climatic regions. For example Prescott (1946) remarked that the ratios appropriate to tropical conditions were generally lower than ratios for temperate conditions. Prescott (1946) proposed a new index (I), involving the ratio of precipitation (P) to some power of evaporation (E):

$$\frac{P}{E^m} = I$$ **Equation 2-1**

Prescott further defined the ratio, as applying where edaphic (soil) conditions were constant, and found that the index held for boundaries separating major soil zones in Australia, with m = 0.7.

Both of these ratios, or indices, clearly reflect the acceptance that the growth of plants is primarily controlled by rainfall, but also that other climatic variables such as temperature and wind will influence the availability of incident rainfall to plants. These other factors are represented most simply by an evaporation term.

C. E. Hounam authored a series of reports (Hounam 1947; Hounam 1947a; Hounam 1950), examining the suitability of areas in the wheat growing regions of Australia for further (post world war two) development. Information about suitability for wheat production at the time was apparently based primarily on rainfall. For example Hounam (1947), reported that the Bureau of Agricultural Economics indicated that a total of eight inches (203 mm) of rain during the months of April to August inclusive was considered desirable for satisfactory (wheat) production.

The optimal monthly distribution of the rain (in Western Australia) was :

April	1½ to 2 inches	(38.1 to 50.8 mm)
May	1½ inches	(38.1 mm)
June	1½ inches	(38.1 mm)
July	2 inches	(50.8 mm)
August	1 to 1½ inches	(25.4 to 38.1 mm)

Hounam examined rainfall figures over a growing period (from April to October in Western Australia) and specifically from April to August. He noted that using an eight inch (203 mm) isohyet to define limits to productive areas excluded some already successful centres. He further noted that an eight inch (203 mm) map does not take into account the specific monthly requirements of wheat. And that one or two individual monthly totals may be sufficient to yield an eight inch (203 mm) seasonal total, irrespective of deficiencies over the remainder of the period or season.

Hounam developed an indicator which he called the 'Seasonal Growth Factor' (SGF). The SGF was the product of the respective probabilities of receiving the optimal amounts of rain, such as those mentioned above, for the growing season (in this example from May to August). This factor represents the 'theoretical chance of receiving the desired amount or more during every month of the period.' Hounam tested the relationship between the SGF and yield in his various reports. An example of the relationship for the South West Wheat Belt Hounam (1947a) is presented below in Figure 2-1. The data for this example is for 34 stations throughout southern NSW for the seasons 1929-30 to 1939-40, and is presented in full in Appendix 1. Hounam describes the correlation as fairly good, but does not present a curve fitting equation or any statistics on the goodness of fit.

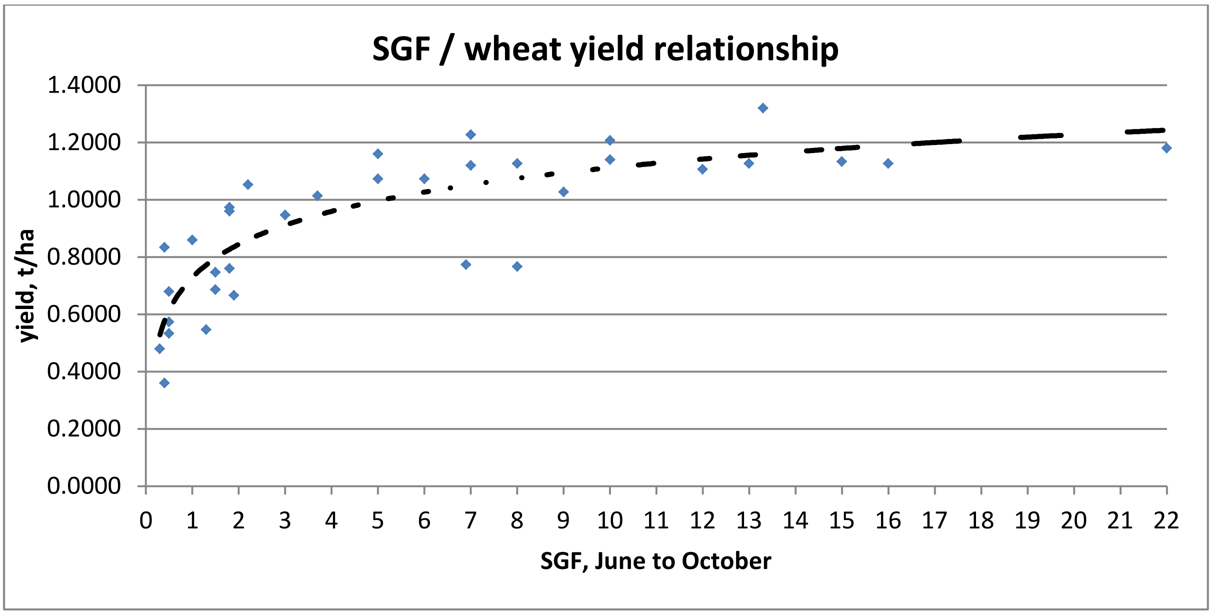


Figure 2-1 An example of the Seasonal Growth factor / yield relationship, after Hounam 1947, yield units converted

Hounam recognized that while plant growth depended on rainfall, when considering moisture available to plants, some thought must be given to evaporation Hounam (1947a). Following Prescott's work, he noted that early work on P/E ratios had been developed into the equation

$$\frac{P}{E^{0.7}} = 0.54$$ Where: P is now the *effective* rainfall in inches, and **Equation 2-2**

E is the evaporation in inches

Hounam (1947a) defined effective rainfall as 'the rainfall necessary to start germination and maintain growth above the wilting point'. This equation applied to natural growth of pastures, and so could be applied to cereals when methods to conserve water were being used. The equation defines the minimum amount of rainfall required, under the specific conditions of evaporation, for germination and development to occur. Hounam used annual average evaporation figures to calculate periods of the year when the rainfall exceeds the effective amount for specific areas. He defined these periods as 'influential'. The period of influential rain therefore being the time over which the surface soil is above the wilting point for plants. As with his earlier SGF, Hounam calculated the probabilities of receiving influential rain for each month. For periods of interest, such as the critical months for wheat in Western Australia of April to August, the percentage frequencies for influential rain are multiplied together. The result is 'theoretically the chance of receiving monthly rainfalls greater than the effective amounts throughout the period.' Hounam termed this value the 'Influential Rainfall Factor' for a specific place and period, and tested the relationship between the factor and yield. An example of the relationship is shown in Figure 2-2.

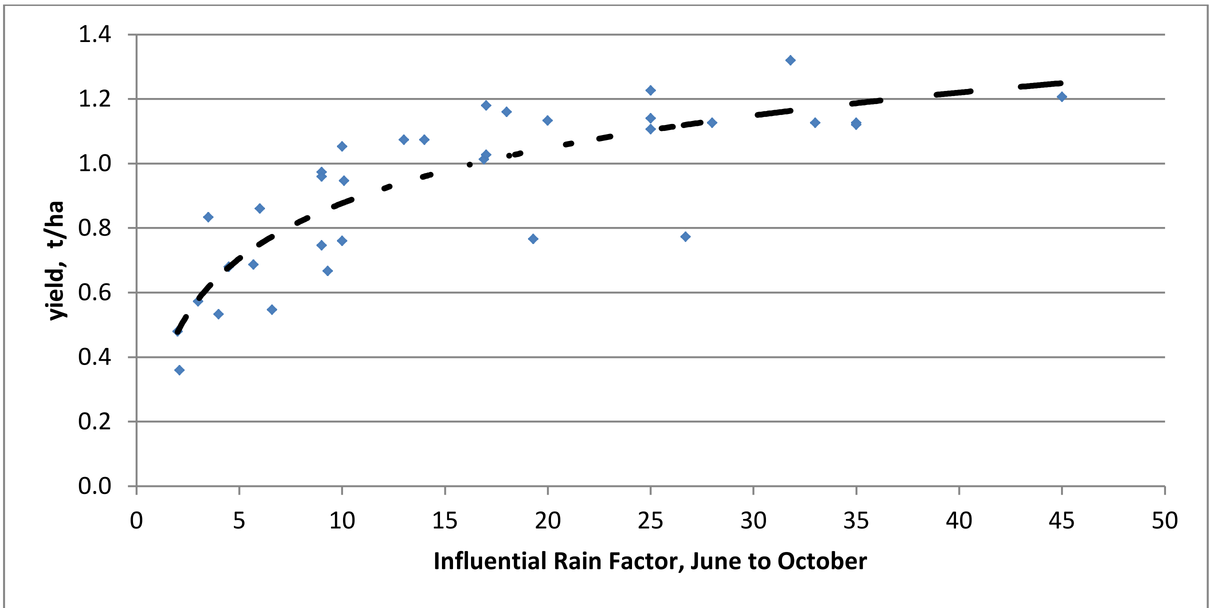


Figure 2-2 An example of the Influential Rain Factor / yield relationship, after Hounam 1947, yield units converted

The data for this example is for 34 stations throughout southern NSW for the seasons 1929-30 to 1939-40, and is presented in full in Appendix 1.

Hounam asserted in his various reports that the Influential Rainfall Factor could be used to successfully define areas suitable or appropriate to wheat production.

Cornish (1949), (1950) studied yield trends and the influence of rainfall on yield in South Australia. While he noted that Trumble's work (1937) helped delineate wheat cultivation, he decided that the most relevant parameter to relate yield and seasonal conditions would be the rainfall in certain defined periods of the year. Cornish performed multiple regressions involving yield, three seasonal rainfall variates and time. Cornish concluded that the relation between yield and rainfall was dominant over yield versus time, and that 'it can be claimed that a rainfall record provides a sufficiently accurate index of the seasonal conditions in this environment.' While Cornish's regressions demonstrated the relationships between yield and rainfall, the respective regressions were only relevant to specific localities and environments.

Slatyer (1962) investigated the climate of the Alice Springs area, using critical amounts of rainfall for a specific time period and area, the initial effective rainfall and effective carryover rainfall of White (1955). Slatyer extended White's method to include evaporation, and used cumulative precipitation excess (> $0.4E$ over weekly and > $0.2E$ over four weekly initial periods) to estimate soil moisture storage suitable for pasture growth.

Attempts to model soil water regimes were becoming more common in the 1960's (Fitzpatrick, Slatyer et al. 1967), (Fitzpatrick and Nix 1969), with workers mainly focussing on direct rainfall and evaporation or evapotranspiration measurements, and improving evaporation estimates (Fitzpatrick 1963).

Investigating drought in the United States, Plamer (Palmer 1965) developed a series of indicators (Z index, Crop Moisture Index, Drought Severity Index), all of which werer water balance indices considering various parameters as well as precipitation. Palmer's indcies realated to the probabiliteis of reciving 'near-normal' rainfall over spefified periods, unlike som previous work where crop water requirements wre the focus.

Gangopadhyaya and Sarker (1965) devised location specific response curves for predicting wheat yields using seasonal rainfall distributions at locations in India, based on the work of Fisher (1925). While he claimed that 75% of the total variation in wheat yield could be accounted for by the regression equations he developed, he noted that the discontinuous nature of rainfall could not be adequately represented by continuous mathematical functions. The location specific nature of his response curves, as with previous regression models, precludes their wider application.

Nix and Fitzpatrick (1969), studying crop water stress using weekly averaged data, concluded that the best index for plant stress related to an excess of available water over potential evaporation, during critical periods of plant development. They evaluated simpler weather data and yield relationships, and concluded that simpler weather data was not significantly related to grain yield in the study area (Qld, Australia). They did note, however, that rainfall between sowing and heading approached significance.

By the mid 1960's, the focus of research into climate issues seems to have moved to investigating regional scale subjects, especially in relation to global climatic patterns such as the El Niño phenomenon (see section 2.4). While teleconnections[8] had been known about for some time Walker (1924), it was during the 1960's that a dependence on two way coupling between atmosphere and ocean dynamics was proposed (Bjerknes 1969), with work progressing until the 1980's when successful ENSO[9] modeling was achieved (Cane, Zebiak et al. 1986). Work into rainfall patterns had continued eg (Hutchinson and Bischof 1983) but the link with crops was not continued.

The emphasis in crop related work was had also become model based by the 1960's. For example Baier and Robertson (1968) studying wheat yields in Canada concluded that wheat yields were modelled better by the use of a soil moisture budget than climatological data. Baier and Robertson also claimed that 'Rainfall per se was found to be unsuitable as (a) basis for yield estimation.' In later work (Baier 1973), Baier used a combination of climatic data such as temperature and derived data such as potential evapotranspiration to develop a crop-weather analysis model to relate to wheat yields.

In developing a cropping systems model, McCown and Williams (1989) noted that such models in use in the 1980's could be divided into two classes :

---

[8] causal links between patterns of weather in two locations, or between two atmospheric occurrences, which are very far apart, such the El Niño Southern Oscillation, ODG (2004). Oxford Dictionary of Georgraphy, Oxford University Press.

[9] ENSO is an acronym for the 'El Niño Southern Oscillation', which is a climatic phenomenon widely accepted as influential in the Australian climate (Jones 2009). It will be further discussed in section 2.4

1. models aiming to simulate yield across a range of environmental conditions and genotypes, and
2. models where the emphasis is on soil loss and degradation

McCown and Williams went on to develop AUSIM, a modelling system with sub models for crops, soil water, nitrogen and phosphorus, all of which had weather inputs as variables.

Many models relating to crop production and / or soil erosion were developed during the 1980's and 1990's. For example PERFECT (Littleboy, Silburn et al. 1992) related soil erosion to long term crop production, noting that 'Variability in climate dominates agricultural production in the subtropical region of Australia.' CREAMS (Knisel 1980; 1982) predicted pollution from agricultural areas. McCown, Hammer et al. (1996) claimed that a systems approach was needed to reflect modern agricultural production systems, and that new priorities were needed. Among the priorities was the gathering of good field data. (McCown, Hammer et al.) presented a new software system, the 'Agricultural Production Systems Simulator' (APSIM) designed to meet their perceived needs.

Revisiting drought definition, McKee (McKee 1993) devised a Standardised Precipitation Index (SPI), a probability index considering only precipitation. The SPI uses transformed monthly data to provide the probability of recording a given amount of precipitation, standardized so that an index of zero indicates the median precipitation amount. The index is negative for drought, and positive for wet conditions. The SPI is used by the United States government as a way of measuring drought.[10]

The development of complex modelling techniques involving multiple climatic and other environmental parameters, and relating to crop production systems, meant that the literature no longer involved crop growth indices, to this author's knowledge. The next section of this literature review therefore takes up the topic of climate forecasting and possible climatic indicators.

2.4 External climatic influences and indictors

Investigating rainfall variability in Indian monsoons, Walker (1924) reported on the relationship between regional atmospheric pressure fluctuations and seasonal rainfall. He described a system of east-west circulation in the atmosphere above the Pacific and Indian oceans which became known

[10] See, for example http://www.ncdc.noaa.gov/oa/climate/research/prelim/drought/spi.html

as the Walker circulation. The typical Walker circulation circumstance is for high atmospheric pressure systems located over the eastern Pacific Ocean to drive surface air flows towards the western Pacific. Thus, generally, high atmospheric pressures in the Pacific Ocean tend to relate to low pressures in the Indian Ocean, and lower temperatures and higher rainfalls in Australia.

Walker developed the notion of oscillations (weakening or even reversals of typical Walker circulation) in regional atmospheric pressure patterns and associated seasonal rainfalls, and suggested that these depended on sea currents ‘leaving S. America’, in the case of the southern hemisphere oscillation (SO). Walker later produced an index for the strength of the oscillations (Walker and Bliss 1937) related to seasonal mean values of rainfall, surface temperature, and sea-level atmospheric pressure.

According to Troup (1965) the southern oscillation phenomenon had received little attention since Walker’s development of the concept, while he believed that it was promising in terms of seasonal forecasting. Troup noted that Walkers’ index, which had also been called ‘the southern oscillation’, was quite complex, but claimed that pressure was the factor dominating the index. Troup determined to call a new index, based only on pressure differences, the ‘Southern Oscillation Index’ (SOI).

Versions of this index, the mean sea level pressure (MSLP) difference between Papeete (Tahiti) and Darwin (Australia) have developed into the most common way of describing the intensity of the southern oscillation. Of the differing methods of calculation, the Australian Bureau of Meteorology (BoM) uses what is referred to as ‘the Troup SOI, which is the standardised anomaly of the Mean Sea Level Pressure difference between Tahiti and Darwin.’[11]

$$SOI = 10 * \frac{\lfloor P_{diff} - P_{diffav} \rfloor}{SD(P_{diff})}$$

Where P_{diff} : (average Tahiti MSLP for the month) - (average Darwin MSLP for the month),

P_{diffav} : long term average of P_{diff} for the month in question, and

$SD(P_{diff})$: long term standard deviation of P_{diff} for the month in question.

[11] http://reg.bom.gov.au/climate/glossary/soi.shtml

The BoM includes a multiplication factor of 10 to allow the SOI to be quoted as a whole number, with a range from about –35 to about +35. The BoM uses 1933 to 1992 as the climatology period, and notes that daily or weekly SOI values can fluctuate markedly due to short-lived, day-to-day weather patterns, with the index therefore being best represented by monthly (or longer) averages (BoM 2012).

The relationship between Australian rainfall and the southern oscillation has been extensively examined and confirmed, by, for example, Pittock (1975) with one hundred and seven rainfall stations, and McBride and Nicholls (1983) for the same stations.

The term El Niño originally referred to the warm current off western South America marking the end of the fishing season, which commonly occurred around Christmas, El Niño being Spanish for the Christ child. Since at least the 1980's, the term has been applied to events when 'anomalously warm surface waters cover not only the coastal zone of South America but also most of the tropical Pacific Ocean.' (Philander 1983). During these periods the pressure differential between Tahiti and Darwin decreases (or reverses), resulting in negative phases of the SOI, decreases in the strength of the Pacific Trade Winds and a reduction in winter and spring rainfalls across much of eastern Australia (BoM 2012). The reverse of this condition, with positive SOI, strong trade winds and warmer seas to the north of Australia, is referred to as a La Niña event, La Niña being Spanish for 'girl child', and hence opposite to El Niño. La Niña events are associated with increased rainfall over much of northern and eastern Australia (BoM 2012).

The BoM refers to the El Niño Southern Oscillation (ENSO) as the oscillation between El Niño and La Niña conditions. Accordingly, there are three possible phases for the ENSO (BoM 2012) :

1. The neutral phase, (neither El Niño nor La Niña), when trade winds blow east to west across the surface of the tropical Pacific Ocean, bringing warm moist air and warmer surface waters towards the western Pacific and keeping the central Pacific Ocean relatively cool

2. El Niño, (weakened Walker circulation), trade winds weaken or may even reverse, allowing the area of warmer than normal water to move into the central and eastern tropical Pacific Ocean

3. La Niña event, the Walker Circulation intensifies with greater convection over the western Pacific and stronger trade winds.

The earlier manifestations of the phases remain, El Niño indicating the enhanced probability of drier conditions for eastern Australia, and La Niña indicating the enhanced probability of wetter conditions for eastern Australia.

A large body of working relating the ENSO phenomenon to rainfall in Australia exists, for example Pittock (1975); McBride and Nicholls (1983); Rasmusson and Wallace (1983); Nicholls (1985); Ropelewski and Halpert (1987); Whetton (1988); Stone and Auliciems (1992); Zhang and Casey (1992); Opokuankomah and Cordery (1993); Suppiah and Hennessy (1996); Kane (1997); Diaz, Hoerling et al. (2001); Chiew and Leahy (2003); Suppiah (2004); Cai and Cowan (2008); Beecham and Chowdhury (2010); Chowdhury and Beecham (2010); Kamruzzaman, Beecham et al. (2013). It is not the intention here to review this aspect of the literature. Instead, selected works that are more specific to this project will be examined

While noting the good correlations between SOI and Australian rainfall obtained by Troup and Pittock, Stone and Auliciems (1992); Stone, Hammer et al. (1996) examined phases of the SOI life cycle in relationship to rainfall, suggesting changes in SOI could be important. Stone and Auliciems (1992) identified five SOI phases, depending on SOl values in the current and immediately preceding month, identified as follows (also showing the relationship to rainfall in eastern Australia):

1. consistently negative — rainfalls below long term median
2. consistently positive — rainfalls above long term median
3. rapid fall — rainfalls below long term median
4. rapid rise — rainfalls above long term median
5. consistently near zero — neutral

Hammer, Holzworth et al. (1996) noted improvements in climate forecasting, and the development of various seasonal climate forecasting systems. They examined existing forecast methodologies, and considered 'the potential value of improved forecast quality', concluding that 'The forecasting system giving greatest value was the Southern Oscillation Index (SOI) phase system of Stone and Auliciems (1992).

Examining the potential benefits of climate forecasting to agriculture, Jones, Hansen et al. (2000) noted that many workers, in many different areas, had already shown good correlations between ENSO activity and agricultural production. Jones was using the three phase (El Niño, La Niña and neutral) ENSO model in common use. He concluded that the potential value of improved management practices was 'about US$ 15 ha^{-1}' for the locations studied, a figure he considered low, reflecting weather variability and ENSO based forecast uncertainty. He also noted that ‘there is considerable potential for adjusting crop management if climate forecasts can be improved.’

Sadras, Roget et al. (2002) studied crop management strategies in the Mallee region of southeast Australia. They claimed that the value of ENSO based forecasts in southern Australia was still unconfirmed, noting that local predictions were needed for management purposes. They concluded that April rain could anticipate seasonal rain, and be of significant use to farmers in the region. They did not, however, test this conclusion in different localities.

Hansen, Potgieter et al. (2004) examined the use of general circulation models (GCM’s) as input to a crop growth model to forecast regional wheat yields in northeast Australia, and concluded that the GCM forecasts were better yield predictors than either three phase (BoM) ENSO systems or five phase (Stone et al.) systems. They noted that the system also predicted yields better than rainfall, despite the (predicted) rain being oused to predict yield. They suggested that model (yield) responses to factors beyond seasonal rain, such as antecedent conditions could account for this.

Wang, McIntosh et al. (2009) used an agricultural systems model, with 114 years of historical climate data to investigate the value of: (1) historical climate knowledge, (2) a perfect climate forecast, and (3) various forecasts of targeted variables. While they concluded that well tested agricultural systems models were better management tools than SOI phase systems, they did note that the SOI (five) phase system did have ‘skills to separate rainfall’.

Chowdhury and Beecham (2010) examined the relationship between the (three phase) SOI and rainfall trends using high quality BoM data for ten sites in Australian cities. The discovered that the SOI was influential in three out of seven sites exhibiting increasing or decreasing trends. While the results were not conclusive, they did note that while previous studies had shown a strong relationship between the SOI and Australia’s eastern and northern regions, they believed their results showed that the relation also exists in the southern region.

Kamruzzaman, Beecham et al. (2013) investigated eleven different climatic indices, functions of sea surface temperature (SST) and sea level pressure (SLP) differentials, and their relation to rainfall and runoff in the Murray Darling Basin (MDB) in south-eastern Australia. They used indices from the Pacific, Atlantic, Indian and Southern Oceans, including the already mentioned three phase SOI index. They found that different models proved best for rainfall and runoff for different catchments, and noted in their conclusions that while it would be worthwhile using computerized optimization to improve the analysis, 'for general purposes, a simpler model with linear and quadratic terms of time, cosine, sine and SOI might suffice.'

2.5 Conclusions

There early use of climatic data in Australia was to categorize existing and likely areas for production, with production not surprisingly being the emphasis in the developing agricultural sector. It was soon noted that annual average figures were too coarse to adequately define cropping systems, and as better data became available, various data intervals were used as deemed appropriate to specific studies or crops. The availability of continuous data has allowed accurate crop and soil process models to be developed, although there still seems to be disagreement about appropriate reporting intervals. The most suitable reporting intervals may depend more on user requirements than the perceived modelling or analytical accuracies.

The search for a crop or climatic index was possibly initiated in an attempt to understand or simplify the complex relationships between crops and the environment which were being discovered. While simple yield and rainfall regressions proved promising, it was apparent that within season variation and site specificity limited their applicability. Including evaporative terms, while quite beneficial, divided researchers about the most appropriate method. Clearly no single index had proved universally applicable, while some acceptance of early growing season rains as an indicator seems common in specific areas (Sadras, Roget et al. 2002). The search for such an index seems to have faded with growing confidence in the capacity of crop production and soil process models, with possibly the most significant outcomes being an acceptance of the variability of rainfall, the value of a soil moisture relationship and the importance of preceding or continuous conditions.

The development of complex climatic models, and the availability of quality data for agricultural systems models, has allowed some quantification of the relationship between crops and climate, especially on a seasonal basis. However, such models are quite complicated, and at this stage probably not appropriate as farm level tools. There is little doubt about the relationship between the climatic southern oscillation phenomenon and seasonal rainfall patterns in Australia, while its absolute definition remains elusive. It is likely that as climate models progress, the understanding of such phenomena will allow much greater detail to be revealed and predicted. It seems that of the potential indices revealed, the five phase SOI system of Stone et al warrants investigation. Research involving the three phase SOI system seems less conclusive, while the SOI life cycle and lag relationship of the five phase system could promise a good fit to the carry over systems already noted as significant. Many of the other indices relate to sea surface conditions and energy transfer involving the southern Pacific Ocean, and so a single index, which derives from and encapsulates these systems, is accessible to local users, and applicable to their own data has some appeal.

3 Data

3.1 Aim & intent

As previously noted, the aims of this project involve the analysis of historical rainfall data collected by farmers and others (Crowdsourced data), along with rainfall data from the Bureau of Meteorology (BoM). From that background, the following two chapters of this report detail work performed to satisfy the first project aim :

- To develop a data gathering, secure storage, primary analysis and reporting system for rainfall data already collected and being collected by farmers, to simplify and encourage user participation in the data gathering process, and to develop new methods of displaying user rainfall data which relate to user needs and interests.

This chapter is focussed on data gathering and management.

Historical daily rainfall data from the BoM is easily obtained, with data being available from their online data services web site[12]. It was decided to use the data on DVD which was available at the time the project started, which has since become data on USB[13]. The data set obtained consists of 17384 stations with daily rainfall records Australia wide, all of which have associated descriptive metadata with information such as dates of data availability, quality of data, identifiers, location and the like. While high quality stations were available in most of the areas available for analysis, the completeness of the site records was a useful filter for selections of sites to include in analyses. For the data set used in this study the total numbers of stations per state or territory are as follows: Antarctica 4; Islands 40; New South Wales 4955; Northern Territory 688; Queensland 3955; South Australia 1590; Tasmania 1045; Victoria 2286 and Western Australia 2821. The 152 high quality (HQ) stations used by the BoM in climate change monitoring, while also listed separately on the BoM website, are included on the DVD.

Another source of historical daily rainfall data in New South Wales (NSW) is the NSW Office of Water (formerly Department of Water and Energy). A Historic data CD/DVD 'Pinneena' is available,

[12] http://www.bom.gov.au/climate/data-services/
[13] While the term USB in fact refers to the Universal Serial Bus (a system of data transfer), it is commonly used for portable data storage devices using this communications architecture

containing 'the majority of information stored in the government's continuously monitored water archive'[14]. For this project, version 9.2 was available, including 2640 years of continuous rainfall data from 188 stations in NSW. As with the BoM data, descriptive metadata is included on the DVD, also including completeness of site record information.

Historic daily rainfall data was also sought from institutions such as University research groups and other government agencies. While some information was made available from individuals within some agencies, no extra data collections were made available, and so individual data sets would be processed as they became available, in a manner similar to farm data sets.

It was recognised that data from an assortment of sources would have a variety of formats, and as electronic or paper records. It was intended that a system of data input would be developed to allow users to input electronic data via the internet. Clearly paper data would require transcribing, but it was expected that sufficient electronic data would be available to allow analytical systems to be developed and proceed before reliance on paper data was necessary.

While no precise target was specified, initial discussions suggested one thousand sites (farms) as a potentially achievable goal, representing about five percent of the farmer membership (19000 in 2012) of GrainGrowers, the project industry partner, (GGA 2012). This number was also seen as significant in a statistical sense, being well above those suggested as appropriate for a population of about 20000 (Krejcie and Morgan 1970).

It was also anticipated that groups of farms in reasonably close proximity would be available, allowing investigation of spatial relationships between farms and nonfarm data sites.

The majority of the sites were expected to be in Victoria, New South Wales and Queensland, those states being home to most of the GrainGrowers membership.

3.2 Collection

To establish contact with GrainGrowers members, a promotional article including a questionnaire was inserted into GrainGrowers' newsletter in March 2010 (see Appendix 2, Data gathering). GrainGrowers collected responses and forwarded information to the University. Responses to this

[14] http://waterinfo.nsw.gov.au/pinneena/cm.shtml

questionnaire seemed promising initially, with thirteen early responses and the first set of electronic data via email by August 2010. Another twenty eight respondents made contact by December 2010, but only three more sets of electronic data became available in this time. By December 2010 seven sets of paper data had also been delivered. The GrainGrowers member information was stored in a database, which included a linked table with information about supplied data.

In addition to direct email of electronic data sets, an internet data collection service was established, becoming operational in August 2010. Between August and December 2010 two sets of data were sent via the internet site, with three more delivered in early 2011.

During 2010 it was decided that a more direct contact system was necessary, and was therefore attempted. In most cases where personal contact had been made, data was made available, and so it was hoped that by contacting some of the farmers individually, more data would be obtained. As no further contacts were forthcoming from GrainGrowers, in early 2011 I attended two Grains Research & Development Corporation (GRDC) research updates, in central NSW (Trangie) and southern Queensland (Goondiwindi), to distribute printed information about the project and discuss the project with interested parties. The GRDC research updates are meetings involving growers and grower advisors, with GrainGrowers having a marketing presence at the various meetings. Possibly because there was no opportunity for direct contact or presenting information to the farmers, no extra contacts were secured from these meetings.

In early 2011 (March) a temporary data entry operator was appointed to speed up conversion of electronic data to appropriate formats and begin work on paper data transcription. Unfortunately, health issues forced the retirement of the operator before any additional data sets were processed.

In May 2011 another temporary data operator was appointed, with the first set of processed data delivered in June.

By late 2011 nineteen sets of daily rainfall data were available. These data sets were spread across Victoria, New South Wales and Queensland, with no real clusters of farm data locations. The best grouped data set was data from previous research near Cox's Creek in the Liverpool plains area (New South Wales).

By early 2012 it was clear that a much more concerted effort was required if the anticipated farmer data sets were to be gathered. Direct contact still seemed the best option. To achieve this, additional staff were appointed to the project, and areas with groups of properties with existing contact details from GrainGrowers were identified. While some progress was made contacting farmers and obtaining copies of paper data, no further electronic data sets were available by the end of 2012.

While development of analytical techniques was progressing, efficient data gathering, storage and processing systems, being as yet superfluous, were not proceeding.

In 2013, it was decided that, for the purposes of this project and report, a deadline must be set on the inclusion of new data sets for analysis. By March 2013, up to sixty new farmer data sets were available from existing and new contacts, with some new data sets being processed by the end of March. This report details procedures and analysis including the data available at that time (March 2013), some thirty one farmer data sets, the BoM data and Pinneena data.

3.3 Operating data sets

As previously noted, BoM and other institutional data (Pinnneena) is readily available. The only area with a cluster of farm data available was that in the vicinity of Gunnedah, in northern NSW. Figure 3-1 shows the locations of the thirty one farm data sets, and Table 3-1 lists farm site locations.

For the purposes of analysis, any data sites within an area of ±0.5 degrees longitude and latitude of the user's property (or farm selected for analysis) was designated as 'near' the user's property.

While the relationship between latitude, longitude and ground distance in kilometres varies due to the curvature of the earth's surface, the variation is not great over the range of locations of interest. For example one degree of latitude equals 111.0 kilometres, and one degree of longitude equals 77.9 kilometres at Ballarat (Vic), while at Dalby (Qld) the values are 110.9 and 91.0 respectively. Thus the area 'near' the user's property is about one hundred by one hundred kilometres for the purposes of analysis. In a study into semi-arid rangelands in the USA, Augustine found that while sites up to 8 kilometres apart could be expected to be similar, beyond this 'the potential difference in precipitation between a gauge and an unmeasured location increases linearly up to a separation distance of 160 km' (Augustine 2010). Sites within distances of about 100 kilometres, may therefore be expected to be related in some possibly linear, discernible manner, and should be within the threshold of inter-regional variability. Jones (Jones, Wang et al. 2009) reported similar results for Australian conditions.

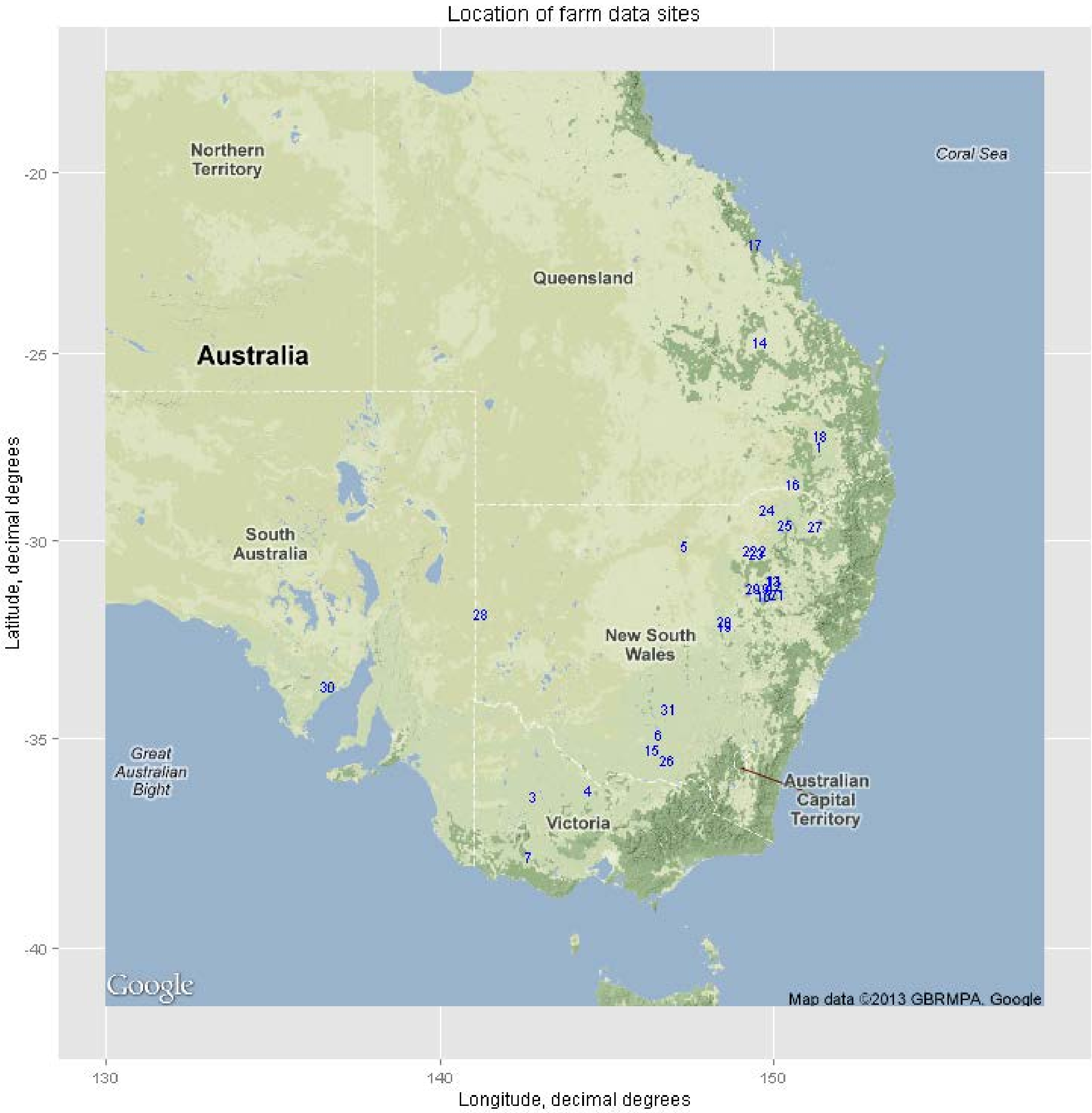


Figure 3-1 Location of farm data sites, as listed in table

Table 3-1 Farm site locations

	ID	Property	Nearby town	X (longitude) (decimal degerees)	Y (latitude)
[1,]	A00013	Springfields	Dalby	151.395	-27.4906
[2,]	A00029	Woolangabba	Woolangabba	149.712	-30.2413
[3,]	A00036	Laen	Horsham	142.833	-36.4167
[4,]	A00039	Pine Grove	Lockington	144.442	-36.25
[5,]	A00040	Lancewood	Brewarrina	147.341	-30.1314
[6,]	A00042	Belalie	Narrandera	146.5983	-34.8934
[7,]	A00046	Coolana	Chatsworth	142.664	-37.8694
[8,]	A00053	Elkay	Mullaley	149.901	-31.243
[9,]	A00054	Mount Nombi	Mullaley	149.787	-31.1982
[10,]	A00055	Mentone	Purlewaugh	149.711	-31.4044
[11,]	A00056	Womponia	Mullaley	150.035	-31.0372
[12,]	A00057	Dimberoy	Mullaley	150.049	-31.144
[13,]	A00058	Beulah	Mullaley	149.973	-31.0128
[14,]	A00059	Unumgar	Mullaley	149.594	-24.6736
[15,]	A00060	Sunrise Park	Urana	146.38	-35.28
[16,]	A00061	Wondalli	Goondiwindi	150.587	-28.5
[17,]	A00063	Erinvale	Millmerran	149.432	-21.9636
[18,]	A00068	Armaroo	Dalby	151.3783	-27.2292
[19,]	A00081	Cooreena Park	Dubbo	148.557	-32.1783
[20,]	A00085	Ashlee	Dubbo	148.531	-32.055
[21,]	A00086	Nowley	Spring Ridge	150.1095	-31.3521
[22,]	A00087	Wilgafields	Wee Waa	149.3	-30.23
[23,]	A00088	Willow Glen	Narrabri	149.472	-30.304
[24,]	A00090	Urella	Moree	149.811	-29.1921
[25,]	A00094	Avondale	Gravesend	150.3279	-29.5793
[26,]	A00095	Munyalba	Wagga Wagga	146.8222	-35.53
[27,]	A00096	Poolbrook	Nullamanna	151.2666	-29.6027
[28,]	A00097	Silverton	Broken Hill	141.23	-31.889
[29,]	A00098	Warrigal	Connabarabran	149.4083	-31.2338
[30,]	A00099	Wevenor Rocks	Cleve	136.6557	-33.7084
[31,]	A00100	Greendale	Bechkam	146.8769	-34.2669

3.4 Management

It is assumed in this document that readers are familiar with PC (desktop personal computer) style computers using the MS (Microsoft) Windows operating system. MS EXCEL (spreadsheet software) is also used by many farmers, and so was an easy option for most users who had electronic data.

Project data is stored using an SQL database. SQL is sometimes referred to as *Structured Query Language*, and is based on database software created by IBM (International Business Machines Corporation) in the 1970’s called SEQUEL, an acronym for *Structured English Query Language*. SQL is a special-purpose programming language designed for managing data in relational database management systems (RDBMS, systems that use separate tables rather than one large one). The SQL databases are managed using ‘MySQL’ software. MySQL (My S-Q-L officially, but also called My Sequel) is the world's most used open source relational database management system, that runs as a server providing multi-user access to a number of databases. It is named after co-founder Michael Widenius' daughter, ‘My’ (SQL 2012).

Data analysis and reporting is performed using ‘R’ software. R is an integrated suite of open source software facilities for data manipulation, calculation and graphical display (Venables 1990). One aim of the overall program, of which this project is a part, is to provide an interactive rainfall analysis and prediction system. For a system to be made available to a user group, in this case farmers, the components of the system must be readily available to the users, or at least available with minimal assistance. It is also conceivable that users, advisors or staff associated with the project may wish to create new analytical tools within the framework of the systems provided, and so components of the system should be accessible for inspection and possible modification. To this end free open source software is being used, open source meaning the source code for the software is available for editing according to specific needs (SQL 2012).

Data used in the project is organised to allow portability and a clear structure. While an internet based interactive system is the eventual aim, many users may be better served, at least initially, by partial installations on local computers. The organisation is therefore based on a standard directory structure and simple naming procedures. The analytical scripts (R) have been written to make use of the standard directory structure. In this initial manifestation of the management structure, the topmost project data directory is called ‘rain project data tables’. The directory ‘rain project data tables’ is used as the base directory for project files on the author’s computer, and is transportable between computers (including sub directories), allowing analytical or organisational processes to operate correctly. Wherever the various R scripts are run, the directory ‘rain projects data tables’ must be found by automatic searching or user selection. Initial setup involves R scripts including user (computer) system and site specification, after which the location of the directory is saved for future reference.

Throughout this document, directory separators are shown as backslashes. Some software, for example R, uses slashes instead of the normal Windows backslashes as separators, or double backslashes to indicate a backslash.

3.4.1 Farmer data

The starting point for the farmer data system is a simple list of membership details for members supplying data to the project (users or participants), provided by GrainGrowers in MS EXCEL format. This file was saved as a text file and imported into a table in MySQL. My SQL provides a unique identifier to each data item (user), which in this case, was prefixed with an 'A' to facilitate processing by R. For example, the first farm data site entered is therefore assigned the project identifier A00001. This table, entitled 'farmer_data_info' contains the users' name and address, property location, type of data and data dates. The same membership information from GrainGrowers also provides another MySQL table with more detailed user contact and address information. Farmer data, in whatever format, was processed to provide standard files for use in R analytical scripts. shows a schematic chart of the steps used to generate user data and descriptive metadata files.

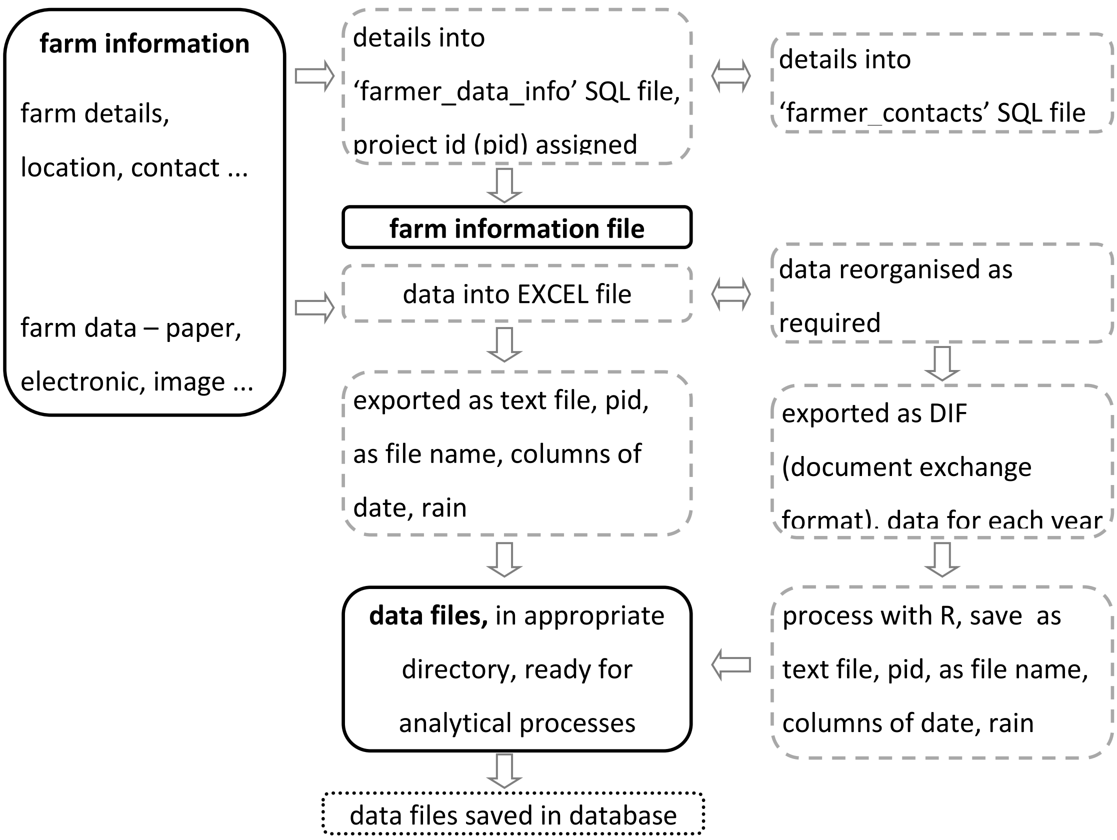


Figure 3-2 Schematic of farm data information processing

See Appendix 3, Farmer data, for more detailed information regarding file processing and directory structure.

3.4.2 Bureau of Meteorology (BoM) data

The Bureau of Meteorology data is very well organised and therefore easier to prepare for use in the analyses. The data is provided on DVD disc[15], or possibly other formats as technological changes occur. Whatever the means of delivery, the data consists of files for each station specified, and descriptive metadata information to allow the user to interpret the files. For this project all available sites with daily rainfall data were used, as described earlier.

The BoM data, being readily available, was not saved in the SQL database, and so was processed with R scripts to read metadata and station data files, and produce text files with station information and data files of date, rainfall and quality information. When the real time, interactive system for the broader program is operational, BoM data files for relevant stations can be downloaded directly from the BoM website. See Appendix 4, BoM data, for more detailed information regarding file processing and directory structure.

3.4.3 Pinneena data

The Pinneena data is equally well organised, if not as easy to access as the BoM data. The data is provided in database format on DVD disc[16], and includes software to extract information about stations as required. For this project all available sites with daily rainfall were used.

The Pinneena data, being readily available, was not saved in the SQL database, and so was processed with R scripts to read metadata and station data files, and produce text files with station information and data files of date, rainfall and quality information. See Appendix 5, Pinneena data, for more detailed information regarding file processing and directory structure.

3.5 Conclusions

The data gathering process has proved to be an extremely difficult one.

[15] http://www.bom.gov.au/climate/data-services/
[16] http://waterinfo.nsw.gov.au/pinneena/cm.shtml

Farmers are usually quite busy and pragmatic people, and need some incentive to devote their time to activities not perceived as directly related to their businesses.

While the project outcomes will hopefully allow users (suppliers of data to the project) to appreciate how the analyses and presentation of their own data can be used to benefit their business activities, there was no method of demonstrating this at the start of the data collection process.

Some of the more technically advanced farmers can simply send data electronically with little effort. Many of the farmers, however, have records on paper or are not every day users of technology such as email or the internet. The effort required for these potential users to contribute to the project is significant, and the value to them was not immediately apparent. This is confirmed by the readiness with which many farmers supplied information on a face to face basis, when the potential benefits of the project could be presented to them.

Apparently (as reported by data gathering team) some farmers' data is committed to agronomists or commercial software systems, and was therefore unavailable for this project. The operations, analyses and/or outcomes of such systems were also unavailable to this project. It would be anticipated that data collected by any future project would remain the property of the farmer, and could be used as input to agricultural modelling systems as required.

The aims of a data gathering and management system have been achieved. It is hoped that as the analytical and reporting systems develop, more famers will become participants. It is recognized, however, that a slow and potentially difficult process of continued data gathering may be necessary to realize the potential of this project and demonstrate the value of the farm data.

4 Initial reporting (the feedback report)

One outcome of the broader program, of which this project is a part, is to develop an interactive rainfall predictive system incorporating user's data, to be available as an internet product.

Part of the first aim of this project is the development of new methods of displaying or interpreting the historical data collected by farmers which will relate to their farm operations and other needs or interests.

Periodic reporting about progress on the program (program updates) would be made to participants and interested parties, and examples of the type of information to be included in the final system would be supplied in these updates, as analytical methods developed. It was also intended that a more specific report, which became known as the feedback report, would be made available to participants, that is landholders who had contributed data to the program.

4.1 Aim

To address the second part of the first project aim mentioned above, there were two objectives to be addressed by the feedback report :

1. To provide feedback to the participants, in the form of:
    a. temporal information, a historical perspective of their data, in relation to data from their local area, in a new and relevant form
    b. spatial information, a representation of their data in relation to their local area, and data from their local area

The data provided as feedback should be an improvement over that readily available to users from existing forecasting systems. Completion of this aim is the subject of the rest of this chapter. It was hoped that as users were provided with such feedback, more users would become aware of the project, and participate.

4.2 Design

The feedback report was intended for all participants, with varying levels of technical and scientific knowledge. It was therefore important that the report should not contain material that could not be easily understood by participants, such as complex statistical analyses, and that all language and terminology should relate to participant experience, not be presented as a disengaged reference document. All data processing and output was performed using the R software suite.

The broader program predictive information is intended to be interactive and operate in real time. It was therefore important to have the feedback report capable of real time operation. It is anticipated that participants with real time data, such as weather station output, could have such equipment (eventually) providing information to reporting and predictive systems, and so the capacity to work with real time data should be part of the feedback report system. While automatic input and analysis are beyond the intention of versions of the feedback report discussed here, up to current date analysis and display must be an option, within the limits of the available data.

While some farmer data sets were up to date (for example to 2013 at the time of writing), very few of the available nonfarm data sets were similarly current. It would have been possible to continually update the data from the Bureau of Meteorology (BoM), but this would mean continually updating analytical scripts and data management systems. As the broader program progresses, facilities will need to be developed to handle this issue, on at least an annual basis for the type of analyses covered by this project. When using real time farm data, concurrent BoM data would also be required, and so downloads on a monthly update system could be appropriate. This data is currently available online from the BoM[17]. For this project, it was decided that a cut off date at the end of 2008 would be used for data in analyses, a date chosen to match the BoM data sets. This is not considered detrimental, as this project is dealing with historical data, covering spans of at least twenty years (for example).

4.3 Report output

4.3.1 Site definition

The starting point for the report is the definition of the area of interest, the area near the participant's property (the 'user'). As noted previously, a rectangular area of ±0.5 degrees longitude and latitude surrounding the user's property is identified in analyses relevant to the property. All data sites within this area are initially selected, and then a subset of sites with data in the same years as the farm is extracted for inclusion in analyses. As an example, the 'Dimberoy' data set locations are shown in Figure 4-1.

[17] http://www.bom.gov.au/climate/data/index.shtml?bookmark=200

Dimberoy (project id A00057) is one of the farms in the cluster of data sites on the Liverpool plains (NSW), and will be used to demonstrate various analytical procedures. It was decided that a map such as the one shown in Figure 4-1 would be the simplest way to display the local area, reassuring the user that the system was using the correct location, highlighting local topographic features and showing the available data sites nearby. The system generated map is also the easiest way make users are aware of other local users, and hopefully see the density of local users develop over time.

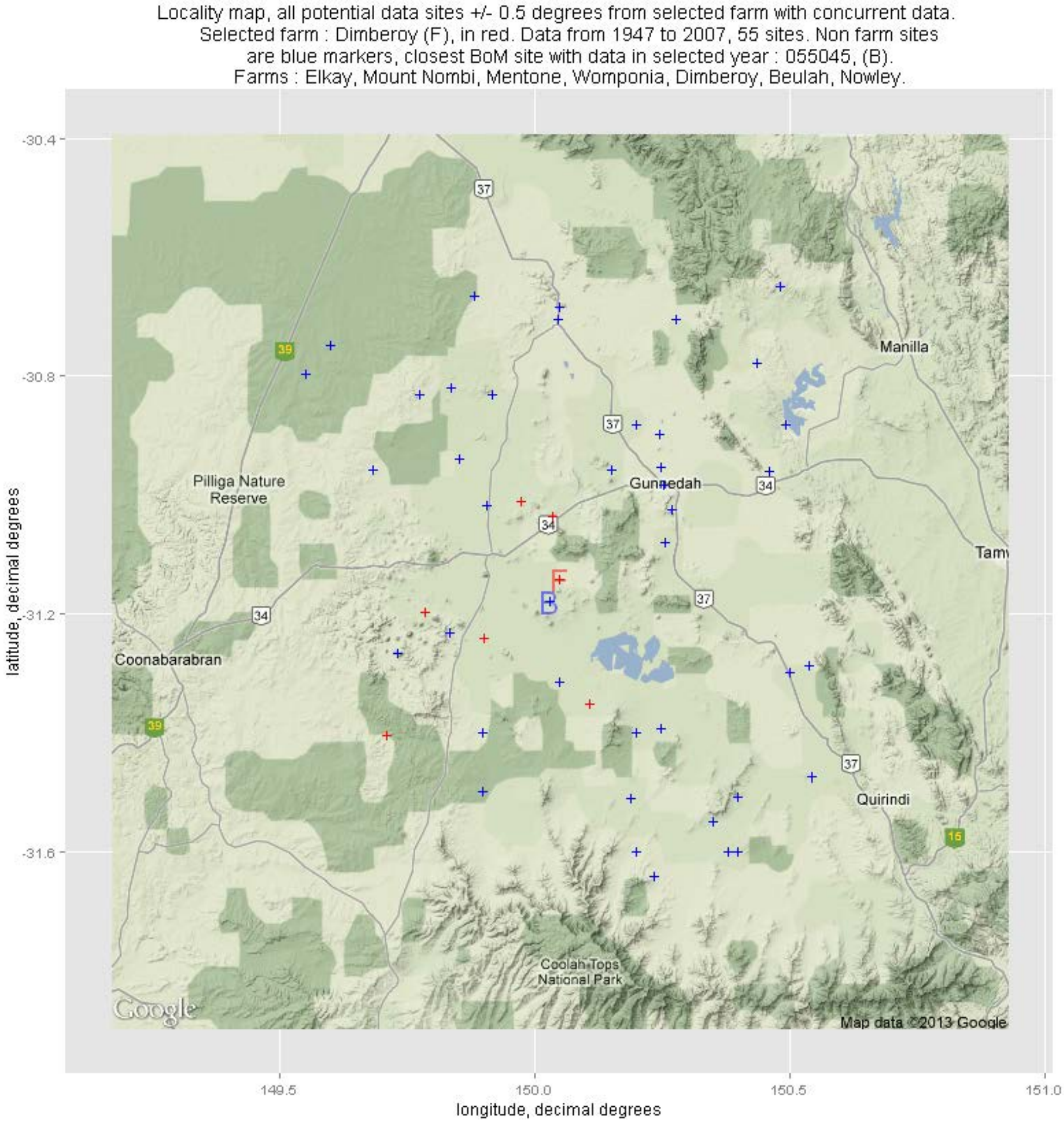


Figure 4-1 Example of Feedback report output, locality map and potential data sites for selected farm

The BoM station chosen for comparison in this plot is simply the closest one to the selected farm. Depending on the analysis being performed, the nearest high quality (HQ) BoM station, or the BoM station with the longest record within about ±0.25 degrees longitude and latitude was used as a representative BoM station.

In all cases, the BoM stations used for comparison are not intended to be equivalent to, or represent, BoM forecasts in any form. They are intended as a quality controlled local comparison, so that the user can note any points of interest or concern. While farms contributing data to analyses are listed in the caption on (example) maps such as Figure 4-1, their locations were not specified on this version of the map.

4.3.2 Temporal data

The most basic display of rainfall time series data is a plot of rain for specific intervals over the time period in question. This information is of some interest, displaying extreme time periods or events, and possibly allowing comparison with other sites of interest. It is also very easy to develop and display. Figure 4-2 shows an example of such a time series plot for the Dimberoy farm and nearby BoM site.

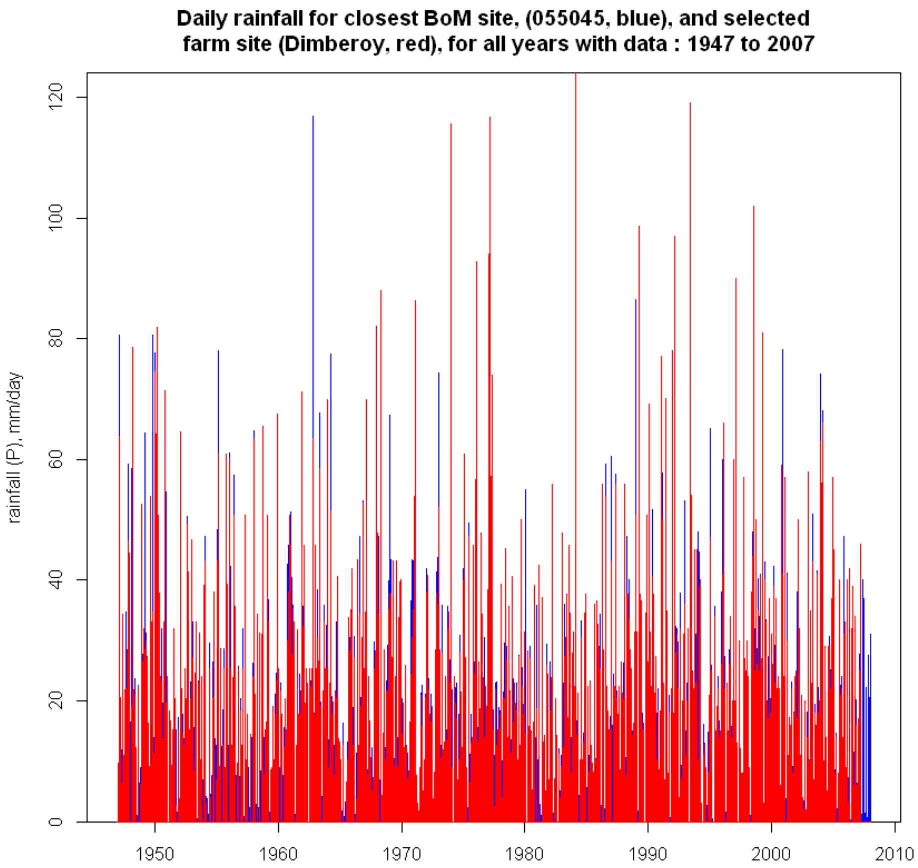


Figure 4-2 All available daily rainfall data for selected farm and BoM station

As indicated, the plot is of daily rainfall for the selected farm (in this case 'Dimberoy') and the closest Bureau of Meteorology (BoM) station (BoM station 055045 at Curlewis, NSW), for the period over which the selected farm has data available. Annual average plots, as shown in Figure 4-3 present similar information to the daily data plot, but the plot is clearer and differences are easier to distinguish. These plots are useful for comparison, presentation of extreme events and a historical overview of the user's data. Figure 4-3 is presented as a display option, scaled to cover the data range, to enhance comparison between the farm and BoM data. The apparent zero point for the BoM data is actually a null data point, with neither stations displaying zero annual rainfall for the period shown.

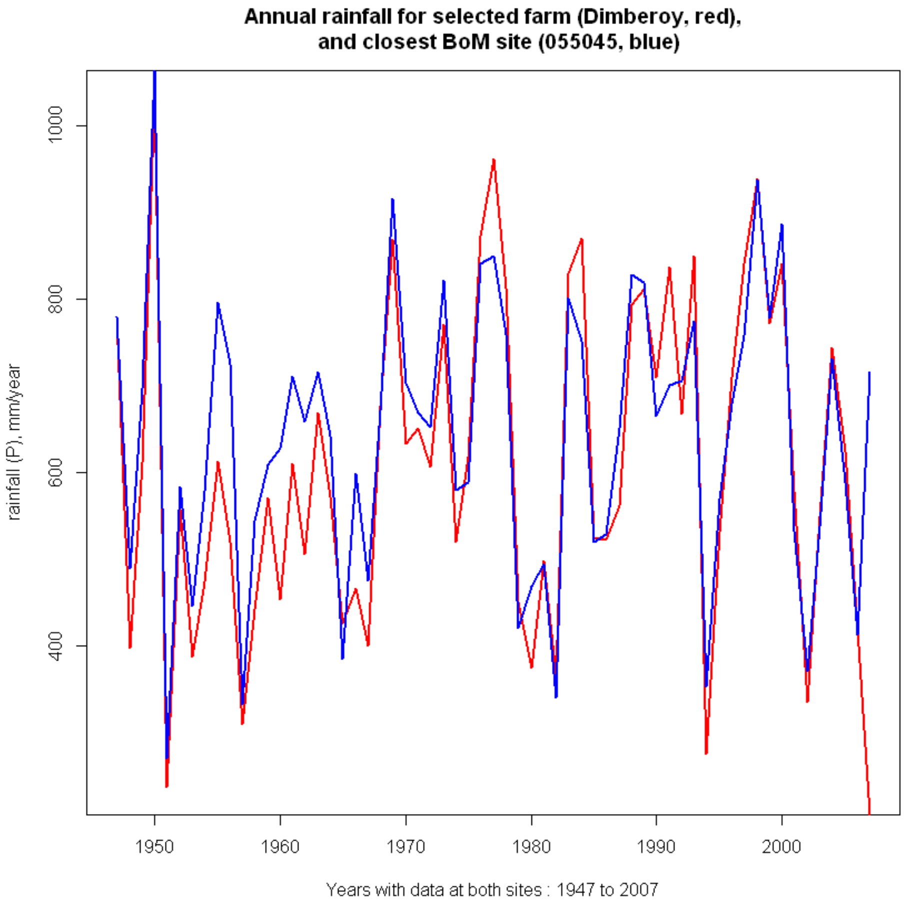


Figure 4-3 Annual average rainfalls for Dimberoy and BoM station 55045

While such plots do not provide information beyond that possibly already available to participants, and do not add value to the information provided by the user, they are an essential component of the feedback system.

Cumulative data provides a more progressive display, showing how a period (year, season) is developing, possibly in comparison to other good or bad periods, and possibly in relation to other sites. This information could be a useful tool for management purposes, allowing judgements to be made on the relative quality of a season, and hence the level of risk for specific crops or operations. However, problems will occur with cumulative data analysis if there is missing or suspect data. While such data can be excluded from averaging techniques, some sort of gap filling or estimation is required with cumulative data. Therefore, cumulative data may prove problematic with incomplete farmer data sets. In the first instance, farmer data is accepted as is, and so incomplete sets will not be discarded. Analyses where specific time series can be used will still be possible for such sites, but some other form of presentation is needed.

With seasonal or periodic comparisons in mind, some link to average conditions could be a useful indicator of the potential for crop success or failure at that time. According to Radinović and Ćurić (2009) 'The percent of normal precipitation is one of the simplest measures of rainfall for a location'. Once again, this information, or information very similar, is readily available to participants, for example from the BoM website 'Rainfall Ranges'[18] as shown in Figure 4-4.

Figure 4-4 shows cumulative data, the link to average, or median conditions, and a predictive element, and so could be useful in planning the season's activities. The information presented in Figure 4-4 is not particularly easy to interpret, but, for example 2013 is predicted to have above average monthly rainfall for the forecast period (at Gunnedah NSW), and even if the rainfall for the coming months (scenarios beyond the forecast date) is low, the overall average will still be 'Above median'. This data is only available covering six month periods (five month span) as shown in Figure 4-4, and only available for specific start dates (January, April, July and September). The information is also only available at specific locations (BoM sites). While this is useful if users are within a small distance of the relevant sites, the information is subject to the same data density and spatial variability uncertainties as previously discussed.

It is worth noting that, while Gunnedah is about 25km from 'Dimberoy', the comparison station 055045 is only about 3.5km away. In this case, station 055045 is a high quality station, and so may

[18] http://www.bom.gov.au/watl/rainfall/ranges.shtml

be part of the data set used to generate information such as shown in Figure 4-4. This, of course, begs the question as to the value of farm data if a high quality station is so close to the farm.

While the area used for the comparison, on the Liverpool Plains in NSW, is fortunate to have high quality productive land, and a high level of focus from the Bureau of Meteorology and other scientific organisations, such proximity to HQ data is not necessarily the case elsewhere, as explained previously.

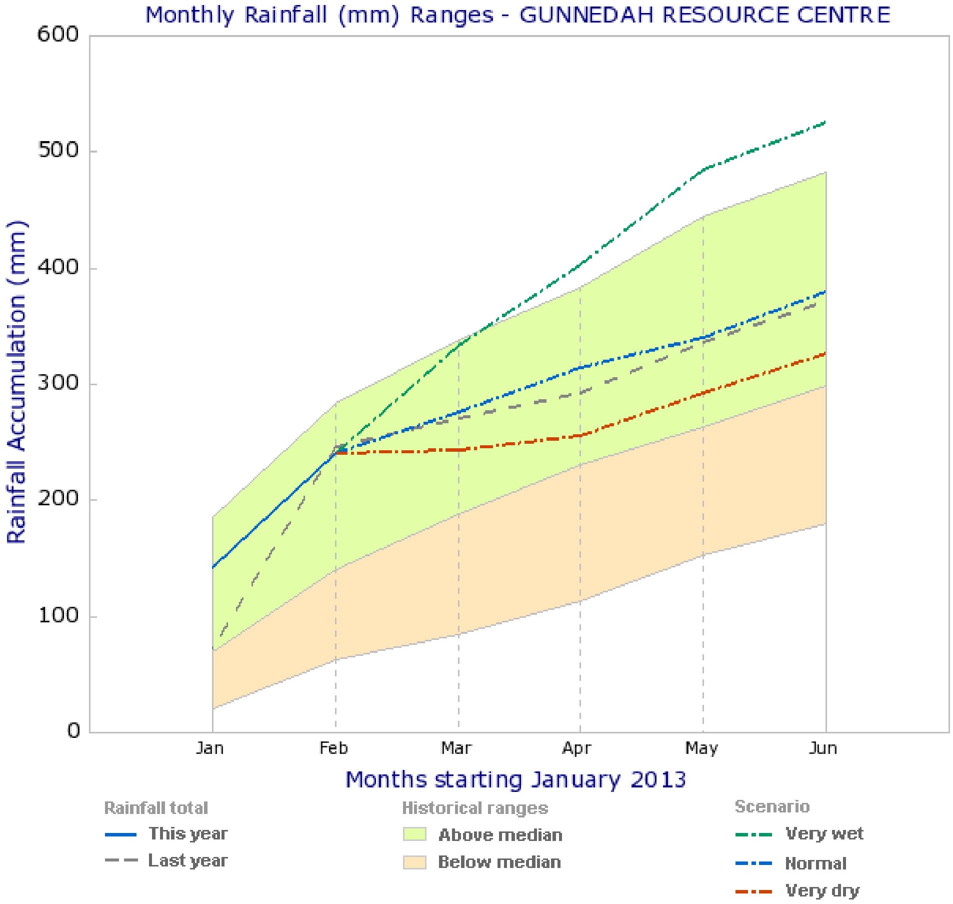


Figure 4-4 BoM 'Rainfall Ranges' information

It is also worth noting that even with the simplest examination of the data, such as that shown in Figure 4-3, there are significant differences between the data from the two sites. While the simple plot does not provide useful comparative data, it demonstrates the potential value of the farm data.

Another method of data access and presentation is provided by 'Australian CliMate', an app (application) for touch devices such as tablet computers and smart phones[19]. This information is very accessible and provides climate and weather information for various locations around the country, including daily summations, forecasts and analyses. This newer style of weather information presentation provides more rapid access to weather information than what may be termed traditional BoM forecasts and websites, and is probably indicative of future user interfaces. It does require reasonable internet access, not yet guaranteed in many rural areas, and, as with other information based on the BoM stations, it is also site specific.

To benefit users beyond what they may be able to access from BoM or similar sources, and derive advantage from the use of daily data, it was believed a continuous analysis should be used. One such analysis, which relates to average as well current conditions, is the cumulative daily deficit and surplus of precipitation, with respect to the average daily precipitation (DSP). This measure displays the cumulative effect of above or below average precipitation for a specific period, and can therefore be related to potential soil moisture storage or refill requirements. While this term is cumulative, it is controlled by values from the time period chosen, and so missing or problematic time periods can be avoided. Clearly, as real time data becomes available this problem becomes irrelevant, and up to date data analyses can automatically include the latest data available.

Radinović and Ćurić (2009) were interested in drought assessment, and used the surplus or deficit technique to define drought periods with interpolated monthly data. They provided a simple equation for daily deficit and surplus of precipitation $\left(D_d\right)_i$ (equation 3, Radinović and Ćurić (2009):

$$\left(D_d\right)_i = T_D + \sum_{i=1}^{n} (P_d - \overline{P_d})_i \qquad \textbf{Equation 4-1}$$

where T_{Di} is the deficit or surplus carried from the previous time period

P_d is the observed precipitation

$\overline{P_d}$ is the 'potentially expected' or average precipitation

and subscripts i and d refer to time steps and daily (as opposed to monthy or other time interval) values

[19] CliMate is a suite of climate analysis tools delivered on the Web, iPhone, iPad and iPod touch devices, developed by the Managing Climate Variability' program of the Commonwealth of Australia (government), 2013

Figure 4-5 shows an example of a DSP plot for the Dimberoy data set for 1980. The year 1980 was chosen arbitrarily for analysis, and is shown on the DSP plot and Figure 4-3 as a fairly dry year.

It is believed that this technique, applied to daily farm rainfall data, provides a continuous assessment of the moisture status of a season or other period of interest, and will supplement the user's management system. For example, if the DSP indicates neutral conditions, a farmer may be able to decide, based on the DSP and other seasonal information, whether specific crop or fertilization plans are appropriate, while if the DSP is in deficit, decisions about the viability of a crop, or the effectiveness of potential seasonal rains in refilling the system may be possible.

The plot includes traces for the wettest and driest years (by annual totals) for the farm data set, and a plot of the same DSP analysis using the (same) BoM reference station, with the endpoints of the DSP analysis noted. As might be expected, with the reference BoM station being close to the property the time series plots and Figure 4-5 show strong similarities between the sites.

Figure 4-6 shows a similar DSP plot, using 'year to date' data, that is starting the analysis 365 days prior to the date of analysis.

This type of plot will become more relevant when users can feed real time data into their systems. This can provide a different impression of conditions, with the year to date data plot showing a more negative DSP than the annual plot for the same date (about day 77 on the annual plot). This information, potentially relating to carryover (antecedent) conditions, may provoke a quite different management response.

As well as annual or year to date analyses, users may prefer to use an analysis for a particular period, based on their own crop management systems or strategies. For example if fallowing is being used, a 'Summer to Summer' analysis may be preferred, reflecting carry over moisture. While interactive user manipulation is clearly not available in the early printed versions of the report, it is envisaged that the online versions will have this capacity.

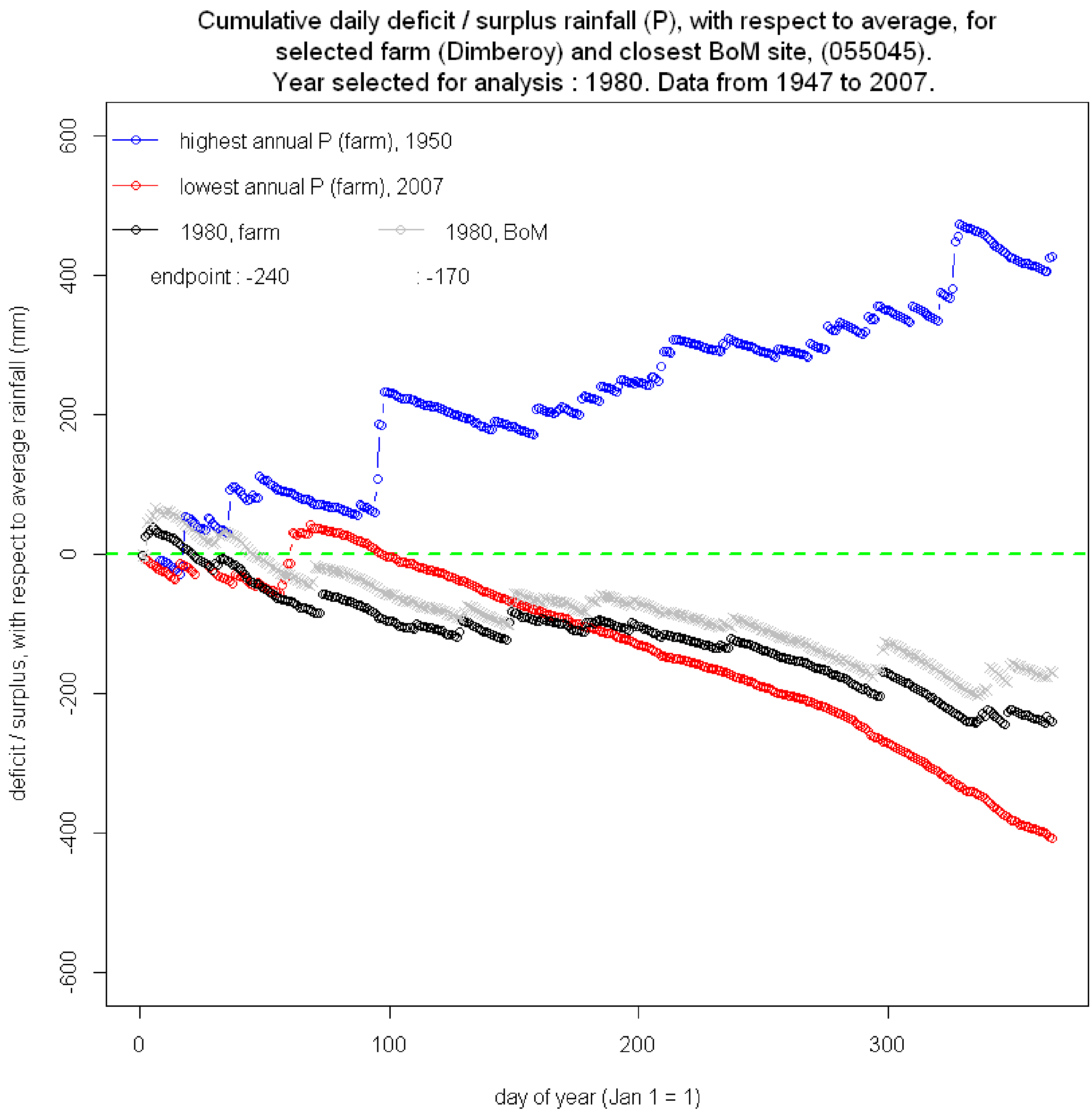


Figure 4-5 Example of Feedback report output, DSP data for selected farm and BoM station

The DSP analysis is not readily available to users (to the author's knowledge), and provides a new, simple illustration of seasonal or periodic rainfall which could assist users' understanding of their property's moisture status. The DSP is consequently a key part of the feedback report, encompassing the temporal data, with a suitable explanation included.

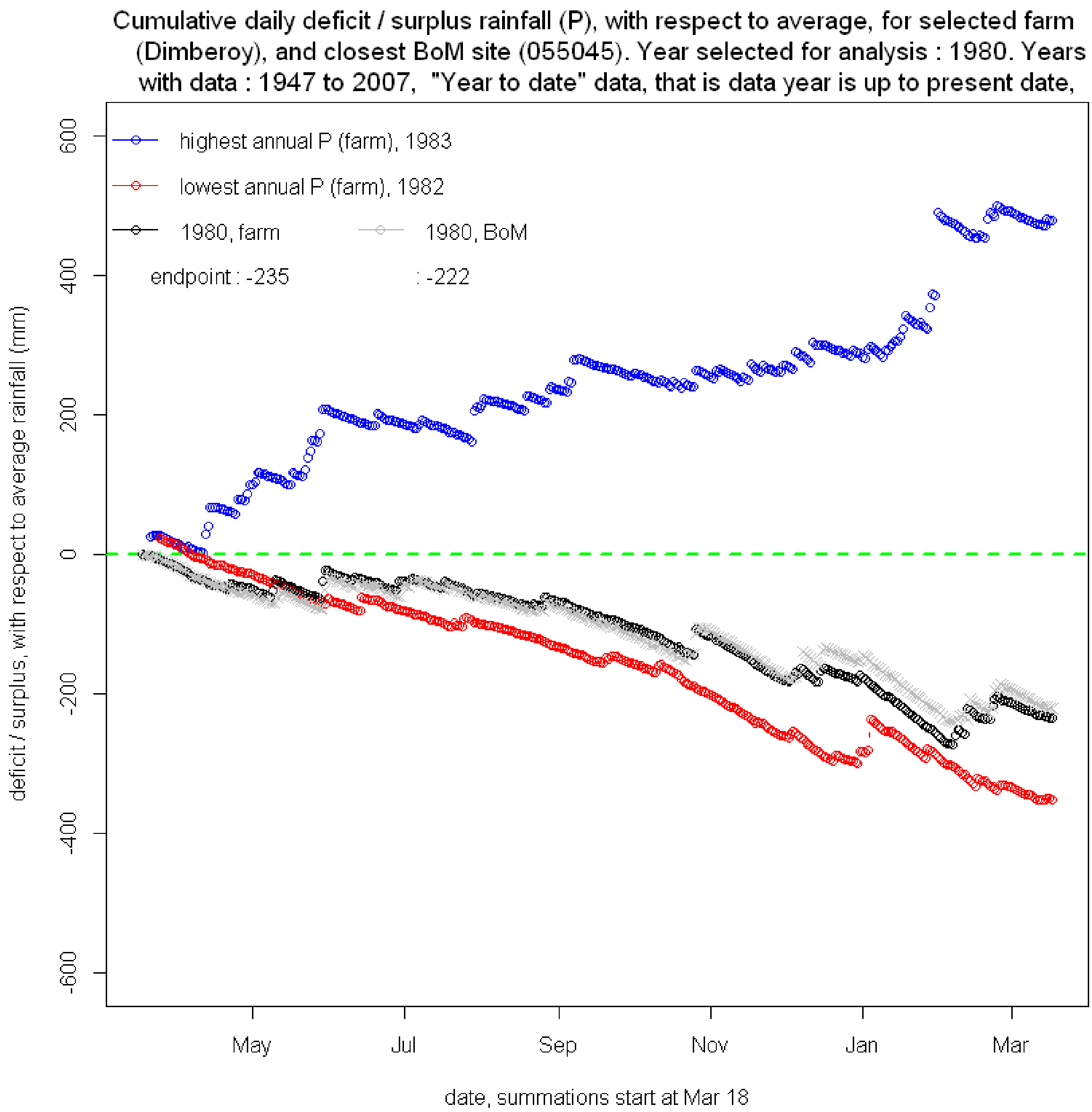


Figure 4-6 Example of Feedback report output, DSP data for selected farm and BoM station, 'year to date' data

4.3.3 Spatial data

Presenting the site data from the area of interest around a user's property is intended to give the user some understanding of how rainfall varies in the locality and how regional forecasts may be interpreted at the farm level.

As noted earlier, it was hoped to avoid statistical or presentation techniques which may be confusing to users. While the method used to derive a map of spatially interpolated values may at first appear confusing, it was believed that the map itself would be of value to users, and the intention of the presentation quite clear. Spatial estimation (or geostatistical prediction) quantifies the intuitive notion that places ‘close to one another tend to have similar (environmental variable) values, whereas ones that are further apart differ more on average’ (Webster and Oliver 2007). The derivation of spatial estimates at different scales depends on the number of data points available for analysis, for example high resolution estimates would require high resolution networks, such as 13 gauges per 35km² (Dirks, Hay et al. 1998), which equates to over 3000 gauges for the Dimberoy data set area of about 10000km². As shown in Figure 4-1, there are only 55 potential data sites in the Dimberoy data set.

A number of methods are available for interpolation of rainfall, possibly the simplest being that of Thiessen (1911), in which weights are assigned to rain gauges (stations) based on the areas deemed to be represented by the gauge. Other methods follow similar lines, assigning weights to a rain gauge based on the distance between the gauges, such as the Shepard’s inverse distance method (Shepard 1968), or inverse distance squared methods (Dingman 2002). The methods are used for estimating areal rainfall, but can be used to specify rainfall at points in the landscape. Goovaerts (2000) concluded that geostatistical techniques (kriging), based on spatial correlations between neighbouring sites, provided better estimates of rainfall than distance based methods, which he noted confirmed other workers’ findings. Another advantage of kriging is the ability to complement sparsely sampled primary data with more densely sampled secondary data, which may be useful in the broader program, and so it was decided to use kriging to develop the interpolated rainfall maps for the report. A simple plot of rainfall for the selected year, of farm sites and one HQ BoM site, allows users to exclude any suspicious farm data. One farm site was excluded due to extremely low values being displayed in this test. Table 4-1 lists the sites from the Dimberoy set which had data for 1980, used in the spatial analyses. The UTM[20] coordinates shown in Table 4-1 are used to simplify the spatial analyses by use of linear distance coordinates.

Geostatistics treats spatial data as a sample from a random process, and a summary statistic for spatial correlation is the sample variogram, which estimates the variance of the data at increasing intervals of distance (Webster and Oliver 2007). Figure 4-7 shows such a variogram for the Dimberoy

[20] The Universal Transverse Mercator (UTM) grid, a world wide plane coordinate system was developed in the 1940's by the Corps of Engineers, U.S. Army, (Dracup 2006).

data set, clearly displaying in increasing trend of variance with respect to distance. As reported earlier (section 3.3) Augustine (2010), (Jones, Wang et al. 2009) reported approximately linear increases in rainfall variance up to 160km.

Given knowledge of the way a property varies in space and time through the variogram, kriging provides an estimate of the value of the property at unsampled points. 'In its original formulation a kriged estimate at a place was simply a linear sum or weighted average of the data in its neighbourhood.' Kriging is also known as an exact interpolator, which is an interpolator that maintains the exact values of the data points (Webster and Oliver 2007). Whilst the procedure may seem complicated, such an explanation could be understood by less technically minded users.

Table 4-1 Sites and annual precipitation (P) in Dimberoy data set for 1980

project id	P mm	Y – latitude decimal degrees	X - longitude decimal degrees	Y – km, UTM zone 55	X – km, UTM zone 55
A00053	77.500	-31.2430	149.9010	6539.837	776.2972
A00055	66.040	-31.4044	149.7110	6522.400	757.7541
A00057	77.216	-31.1440	150.0490	6550.436	790.7020
B053082	94.600	-30.7973	149.5526	6590.075	744.2393
B055002	69.600	-31.2342	149.8345	6540.978	769.9862
B055006	62.800	-31.6405	150.2356	6494.862	806.8800
B055007	99.400	-30.7056	150.0458	6599.063	791.7262
B055014	70.800	-31.1168	150.2682	6552.856	811.7006
B055017	75.100	-31.5711	149.7762	6503.759	763.4872
B055018	87.000	-31.1711	149.6456	6548.421	752.1548
B055023	91.000	-30.9841	150.2540	6567.614	810.7772
B055024	75.200	-31.0261	150.2687	6562.915	812.0448
B055029	69.400	-31.2667	150.1000	6536.693	795.1849
B055034	91.600	-30.7454	150.0557	6594.623	792.5545
B055036	65.000	-31.4135	150.4234	6519.498	825.4889
B055037	70.800	-31.5077	150.3986	6509.123	822.8058
B055038	72.800	-31.0976	149.9114	6555.935	777.7126
B055039	77.000	-31.3946	150.2488	6522.099	808.9419
B055044	117.800	-30.7044	150.2767	6598.572	813.8569
B055045	71.000	-31.1791	150.0312	6546.590	788.8975
B055046	60.200	-31.5271	150.4296	6506.879	825.6844
B055055	83.600	-30.9604	150.4601	6569.648	830.5538
B055064	49.800	-31.6001	150.3811	6498.925	820.8259
B055239	65.200	-31.5879	150.4804	6499.982	830.2965
B055263	102.800	-30.9582	149.6844	6571.939	756.4253
B055264	94.300	-31.5103	150.1900	6509.431	802.9745
B055268	108.600	-30.8325	149.7745	6585.666	765.3820
B055273	98.400	-30.8202	149.8366	6586.881	771.3588
B055274	94.000	-30.7784	150.4359	6589.907	828.8624
B055275	60.400	-31.2874	150.5394	6533.137	836.9744
B055276	86.600	-30.8828	150.4928	6578.158	833.9495

A typical set of sample points (for this project) is shown in Table 4-1, with thirty one points for the Dimberoy data set. This number of data points is, unfortunately, well below that considered appropriate for a precise estimate of the variogram. Variograms based on less than fifty data points are often display little structure, and greater than one hundred points is desirable (Webster and Oliver 2007). The kriging procedure is still seen as the best option for generating the spatial data, especially in the light of potentially greater sample numbers or co kriging with other, higher density parameters. However, while the spatial interpolations provide a best estimate grid for unknown data points, it is probably inappropriate to draw conclusions from analytical variogram parameters, which is why they were not included here or on Figure 4-7. The increasing variance with sample separation distance, in accord with other quoted work, is therefore the only information presented in relation to the variogram.

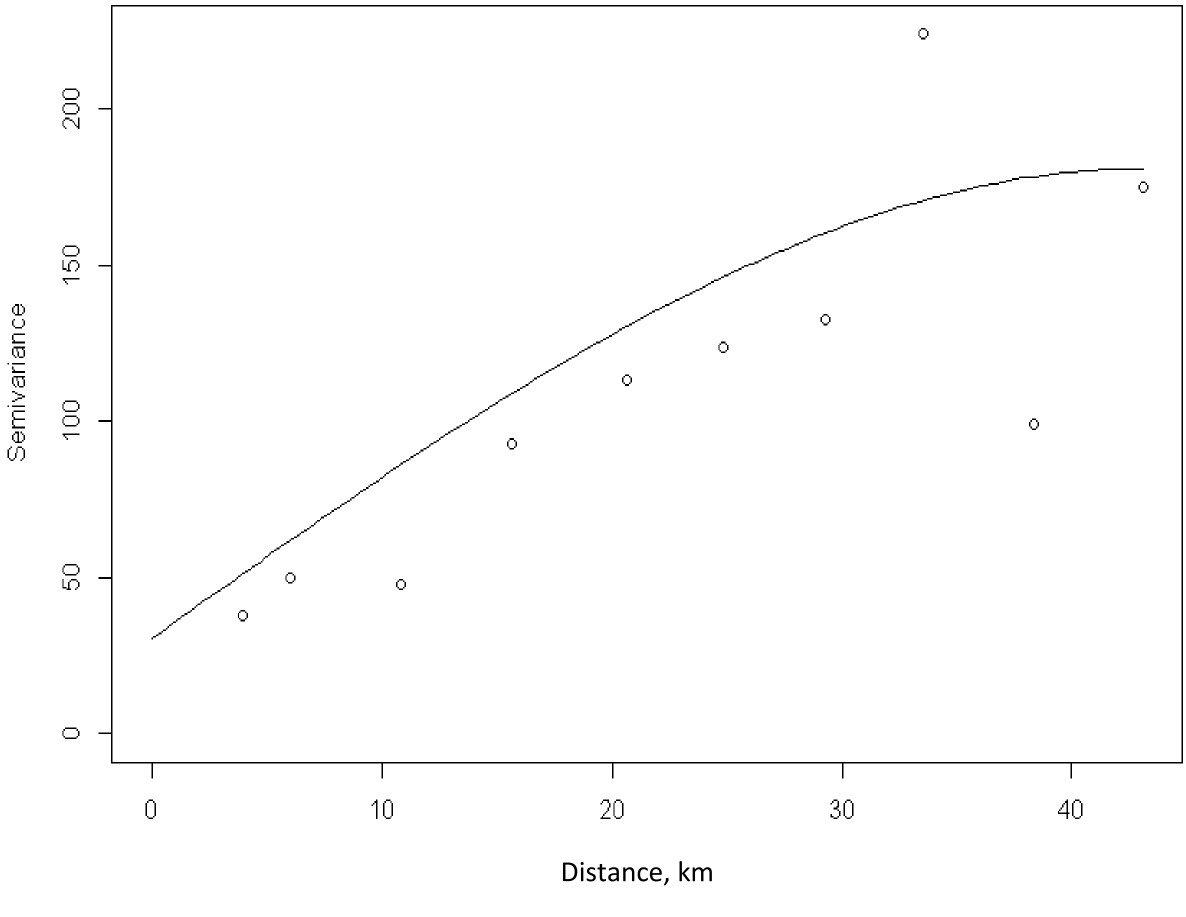


Figure 4-7 Variogram for April and May rainfall, Dimberoy data set 1980

The choice of April and May rainfall for this example analysis and spatial plot was arbitrary.

Figure 4-8 shows an example of the spatial data estimates provided by kriging for the Dimberoy data set.

The interpolation maps show the user the patterns of rainfall in the local area, displaying the variability and highlighting patterns which may help interpret regional forecasts. For example if there is some topographic influence which is manifest in consistently higher or lower rainfall at the user's location. Once again, when the reporting system is interactive, users will be able to specify the month or season and year or simply the year of interest, therefore allowing users to specify such things as the planting season for analysis.

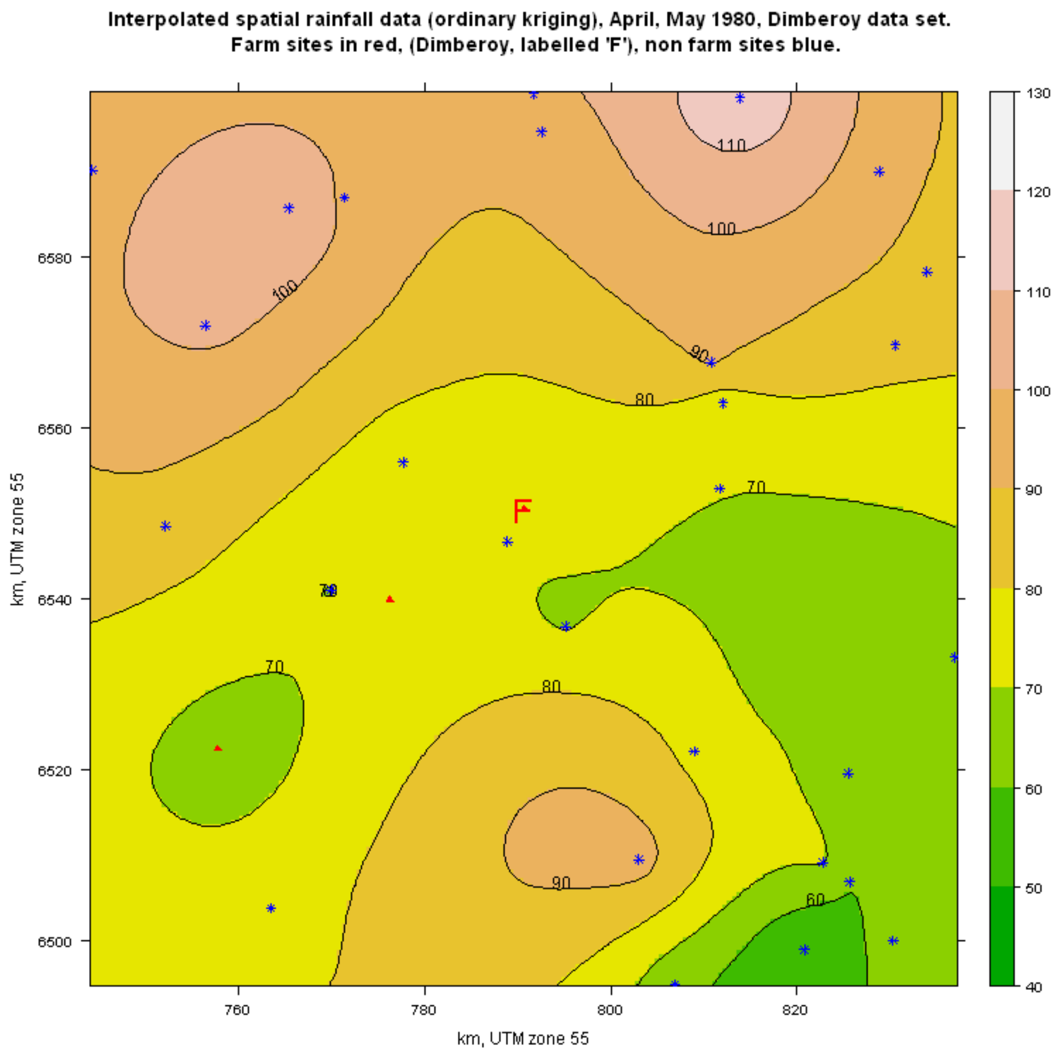


Figure 4-8 Example of Feedback report output, spatial interpolation of data for selected farm and nearby area

As far as the author is aware, spatial information at the level of detail provided by interpolation maps, as shown in Figure 4-8, is not available to users of BoM or other forecasting and interpreting systems, while proprietary systems may be available through advisor or consultants' services at some cost.

4.4 Conclusions

I believe that the aim of this chapter, that the data provided as feedback should be an improvement over that readily available to users from existing forecasting systems, has been achieved.

The DSP analysis demonstrates a new, simple illustration of seasonal or periodic rainfall. In future reports, the DSP should probably be presented as an indicator of moisture status in the system, which users can readily relate to potential seasonal outcomes and therefore management decisions.

The spatial information provided in the report has greater detail than is currently available, and provides information about local variability which would be of use in interpreting existing forecasts. By inspecting various scenarios, users may be able to quantify differences in local rainfall patterns characterised by seasonal, event intensity or directional circumstances, which could be useful in management decisions.

The information in the feedback report can provide users with a unique insight into the patterns and variability of rainfall at their property, and this can be used as knowledge to interpret regional forecasts and predicted seasonal or periodic average values. This will be improved with an online system and the use of (their own) real time data by users.

An example of one of the feedback reports generated for a users' farm is presented in Appendix 6, The Feedback Report.

5 Adding value to forecasts with farm data, temporal data

The second aim of this project :

- By exploring patterns in time and space in historic rainfall data, and the relationship between farm data and other available data, test the usefulness of the farm data, and test whether the farm data can have value in a predictive or interpretive role, and therefore aid in farm management decisions.

is the subject of the following two chapters, this chapter dealing with temporal data and the next with spatial data.

For historic farm data to have some additional value to users, it is believed that some application to current local conditions and some predictive role should be viable, which would enhance understanding or interpretation of seasonal or regional forecasts already available to farmers. The farm data could have two possible roles in terms of predictive capacity :

1. as a scaling factor for forecasts and predictions based on BoM data,
   that is displaying some constant difference from or relationship to BoM data;
2. as a unique predictor, based on patterns in the farm historical data that allow comparison between similar seasons or periods, that is 'what happened last time it was like this'

It is interesting to note that if the role as a scaling factor proves to be the most appropriate, the value of the farm data would possibly be diminished.

5.1 Scaling factor tests

As already discussed, farmers have access to regional, seasonal and, more recently, weekly rainfall data and predictions, usually based on high quality Bureau of Meteorology (BoM) data, and specific to BoM stations which will be of varying distances to users' properties. It was shown in the previous chapter that the farm rainfall data displayed differences from nearby Bureau of Meteorology (BoM) station data, and so to test the first, 'scaling factor' relationship it will be necessary to test the differences or similarities between the farm and BoM data.

As previously, the Dimberoy data set will be used to demonstrate the analyses performed, and other sites will be included in the analyses to test spatial similarities or differences. The first step is to examine the data and select a working range of sites and dates. An inspection of farm data start and finish dates (included in the farm metadata information) suggested dates from 1980 to 2006 as including the highest number of farm data sites while maintaining at least twenty years of record. The potential farm data is displayed in Figure 5-1.

Figure 5-1 Dimberoy data set, farm rainfalls and closest BoM station to Dimberoy

From Figure 5-1 site A00054 is quite different from the other sites. The low values may indicate incorrectly labelled or incorrectly converted data, such as from points (hundredths of an inch) to mm. Inspection of the data showed that converting units did not provide data consistent with the other sites. A00054 will therefore be excluded from the analysis. The other data sets are all similar, with A00053 having missing data during the 1990's. The resulting data sites, with non farm sites at least 95% complete, are displayed in Figure 5-2. The distances from the BoM site closest to farm A00057, ('Dimberoy') to the farm sites are listed below.

Distance from BoM station 055045 to	farm	Dimberoy A00057	:	13.077 km
	farm	Elkay A00053	:	27.862 km
	farm	Mentone A00055	:	56.659 km

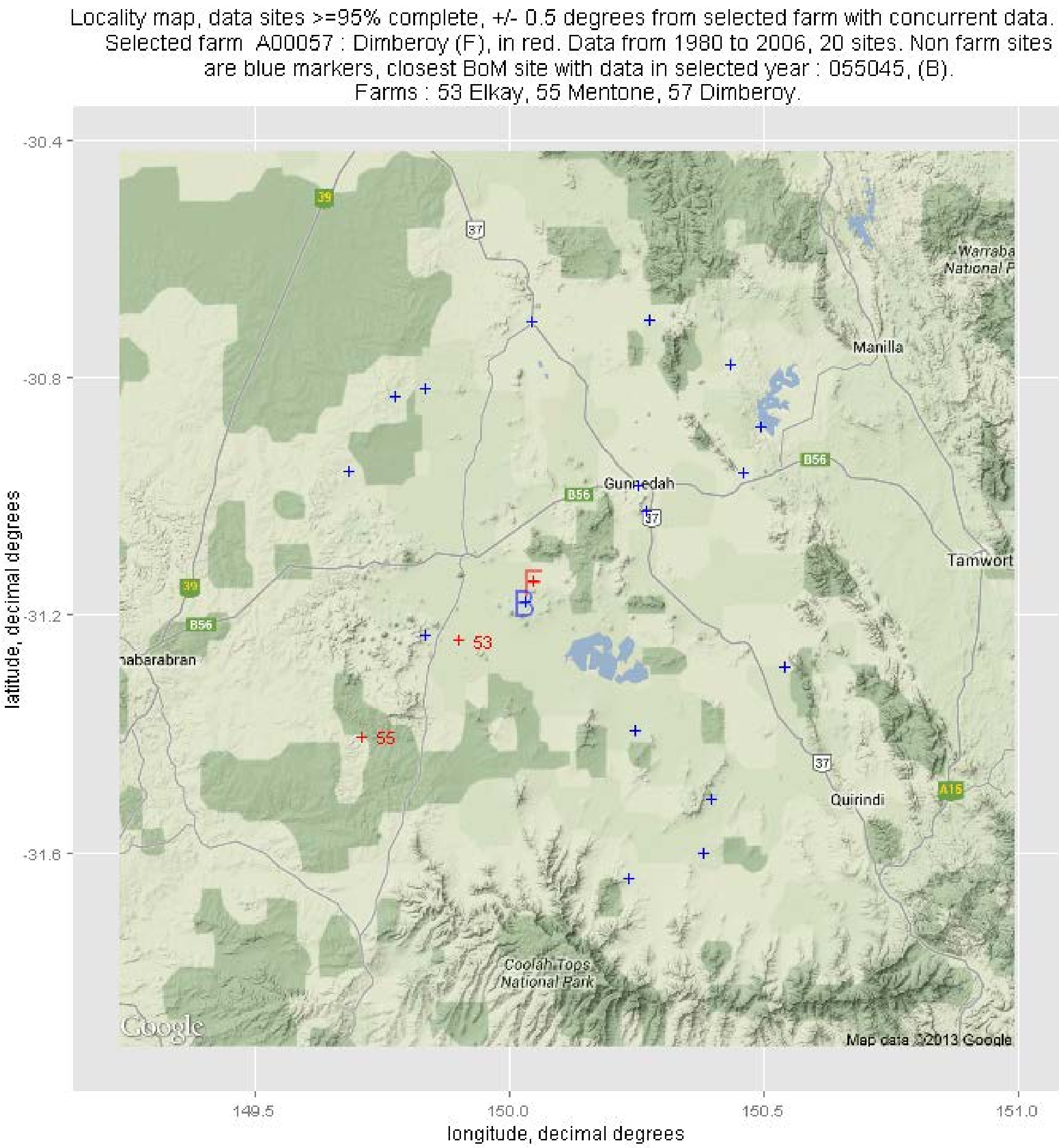


Figure 5-2 Locality map for sites used in scaling factor tests, Dimberoy data set

To avoid problems in comparisons between sites with possibly varying absolute rainfall ranges, it was decided to compare differences between farm sites and nearby BoM sites in the scaling factor tests. While the farm sites and nearby BoM sites in each farm data set are all within limits considered to be similar or be linearly related (Augustine 2010), there are significant differences between sites across the wheat growing areas.

To test for differences between farm and BoM sites, daily averages over some time period will be used, with missing data excluded from the averages. The daily averages from 1980 to 2006 are shown in Figure 5-3. The averages do not show one site consistently different from the others, while (qualitatively) Dimberoy appears to display the highest peak rainfalls in high rainfall events in summer, and Elkay appears to have slightly lower rainfall overall.

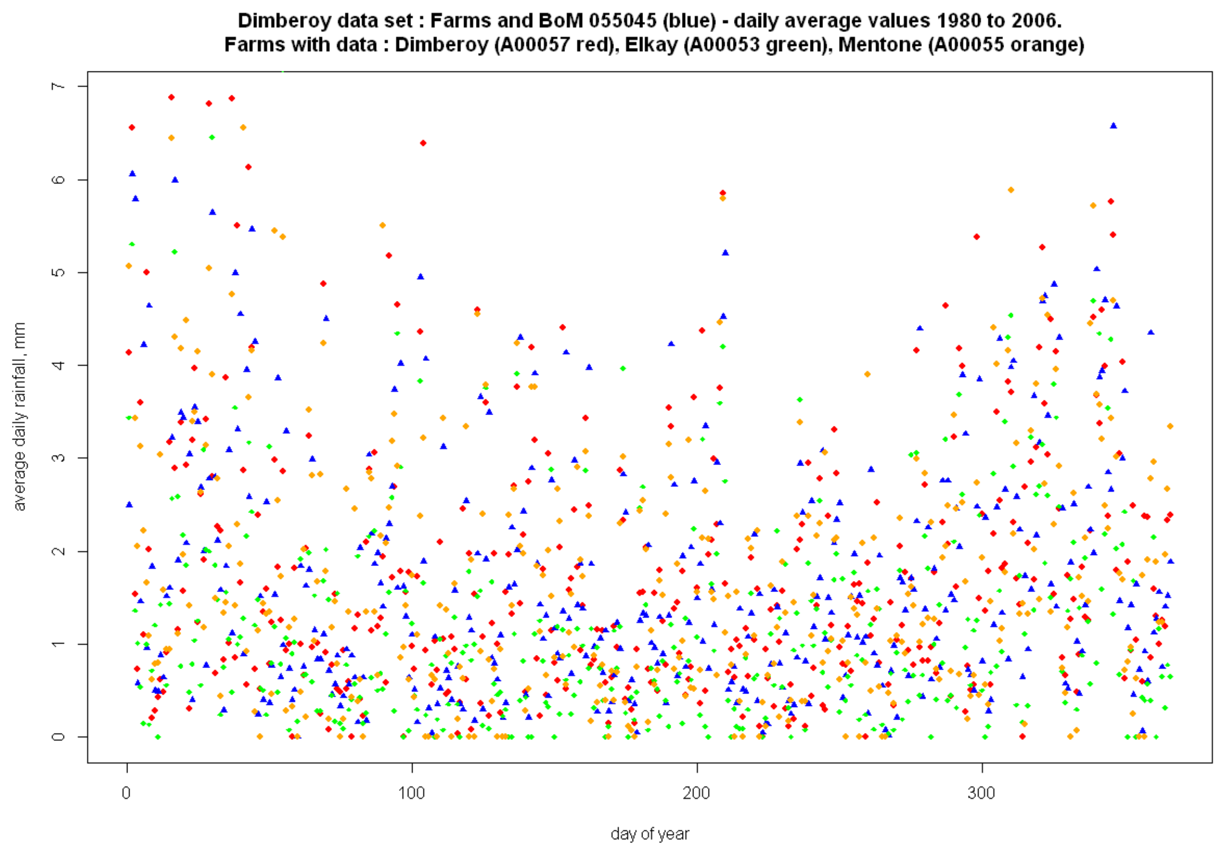


Figure 5-3 Daily rainfall averages from the Dimberoy data set

A box and whisker plot[21] of the same data demonstrates that the bureau station (blue points in Figure 5-4) is similar to the averages of the farm sites (red boxes in Figure 5-4).

[21] Plots in which the boxes enclose the interquartile ranges and the whiskers extend to some proportion of the data extremities (Webster 2001)

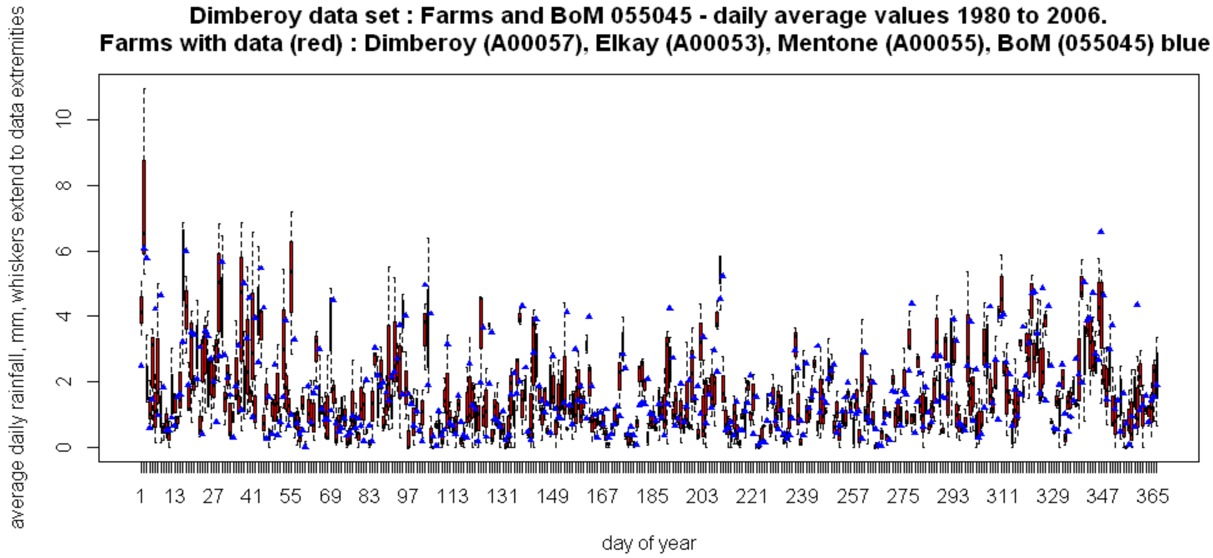


Figure 5-4 Box and Whisker plot of Dimberoy daily rainfall data

The first difference calculation, daily differences between Dimberoy and the closest BoM station, highlights a problem with this simple method. As can be seen in Figure 5-5, there are difference extremes, some of which appear to be on consecutive days or quite closely spaced. Figure 5-6 shows the same data, with the 'whiskers' extended to cover the data extremities. This plot shows the scale of differences to be similar to the range of the daily rainfall data shown in Figure 5-1.

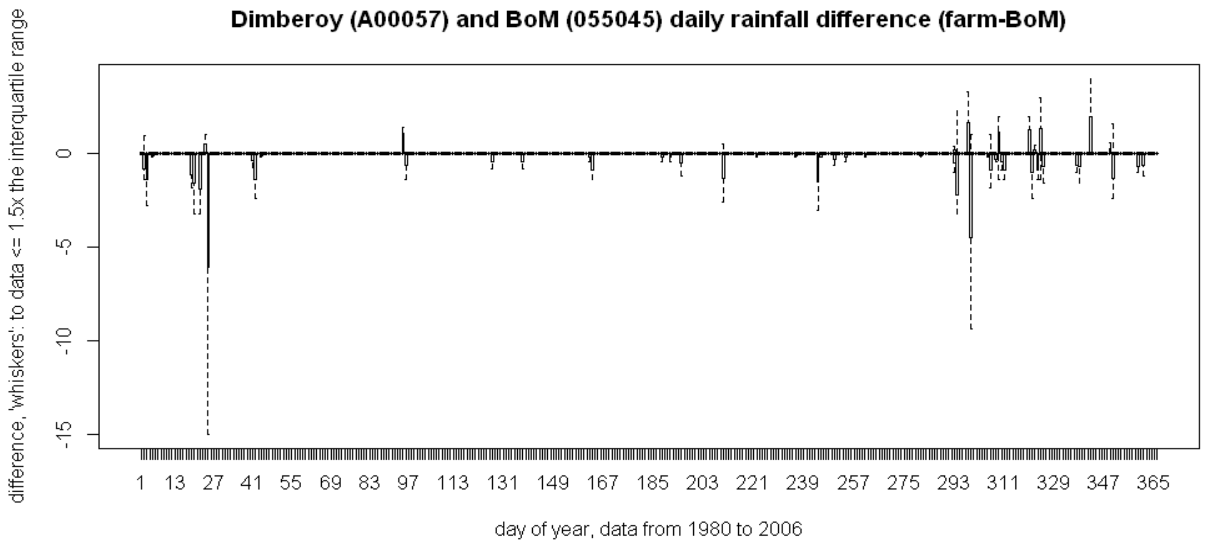


Figure 5-5 difference box and whisker plot – Dimberoy data set

A subset of the data for Dimberoy and BoM station 055045 for January 1980 is shown in Table 5-1. The table shows that the BoM station is regularly reporting rainfall the day after it is recorded on the

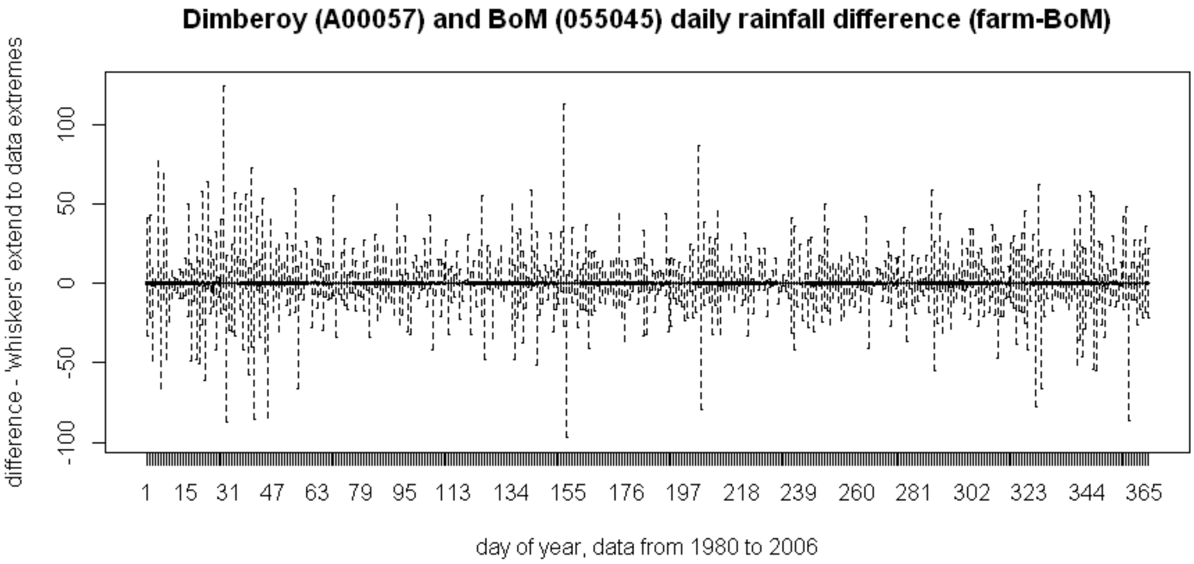


Figure 5-6 Daily difference box and whisker plot - Dimberoy data set, whiskers to data extremities

farm, resulting in a positive difference on the first day followed by a negative difference of similar magnitude on the following day. One response to this may be the rejection of the farm data as unreliable, but I believe some cases may be a genuine temporal variation, or the variation may simply be the result of differences in the procedures at the two sites.

Table 5-1 Dimberoy and closest BoM rainfall events, Jan 1980

	month	day	A00057	BoM	difference
1	1	1	0.000	0.0	0.0
2	1	2	31.496	0.0	31.496
3	1	3	6.096	55.0	-48.904
4	1	4	9.652	8.0	1.652
5	1	5	5.080	8.0	-2.92
6	1	6	0.000	12.0	-12
10	1	10	0.000	2.2	-2.2
11	1	11	0.000	4.4	-4.4
12	1	12	4.064	1.6	2.464
14	1	14	0.000	3.0	-3
15	1	15	0.000	0.0	0.0
30	1	30	14.732	6.2	8.532
31	1	31	0.000	15.0	-15

The BoM site will probably have been read at the same time (commonly 0900) each day, while the farm readings may have been recorded at a consistent time, or may not. I have no reason to doubt the accuracy of the farm data, but it could have been recorded (for example) after rain on the day it occurred, or as afternoon or evening readings when other tasks are finished, which could explain the discrepancy. Some farm records were collected by relatives or previous owners, and so no checking or corrections are possible.

With the exception of standardisation, there is also no reason to accept that regular recording has more value than event based recording, for these analyses, restricted to on farm use. The fact that not all events have the same lag, such as the rain on the eleventh and twelfth of January, when the farm rain is reported on the day after the BoM site, encourages confidence in the farm data. Clearly, this issue cannot be resolved, except in unusual circumstances such as farmers (or others responsible for recordings) taking (and keeping) detailed notes.

The only real solution to such problems would be continuous, automatic rainfall recording at all sites. At the time of writing, there are 716 automatic weather stations in the BoM system[22]. While on farm weather stations are becoming more popular, no figures are available regarding numbers.

Using weekly rainfall data should (for the most part) avoid this problem, at the sake of potential loss of detail in individual events. Figure 5-7 shows a boxplot of average weekly rainfall for the Dimberoy data set.

Once again the BoM station is similar to the farm data, both in seasonal trends and events such as week 30 1980, shown in Figure 5-7.

The differences between Dimberoy and the closest BoM site using weekly data from 1980 to 2006 are shown in Figure 5-8. The figure shows a range of differences, with maximum interquartile and absolute ranges at the start and end of the year, in summer.

[22] BoM website - http://www.bom.gov.au/climate/data-services/#tabs=4

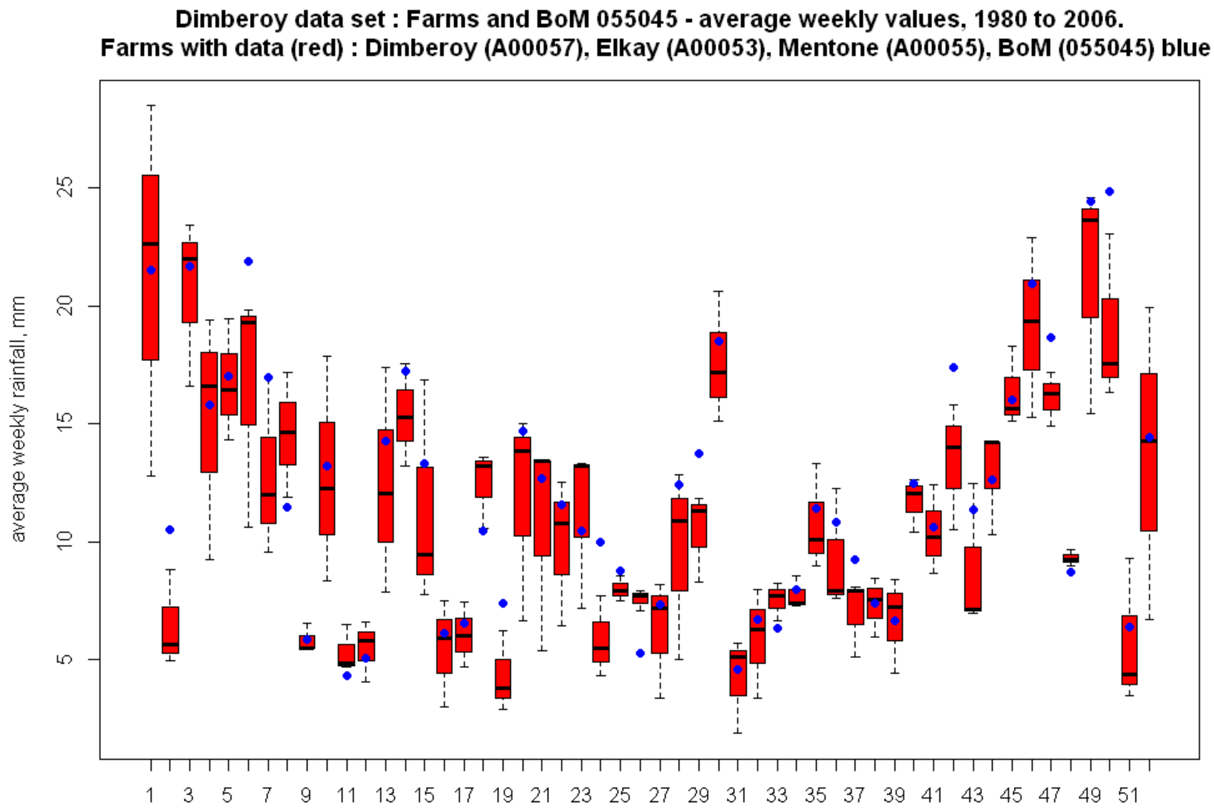


Figure 5-7 Boxand whisker plot of average weekly values, Dimberoy data set 1980 to 2006

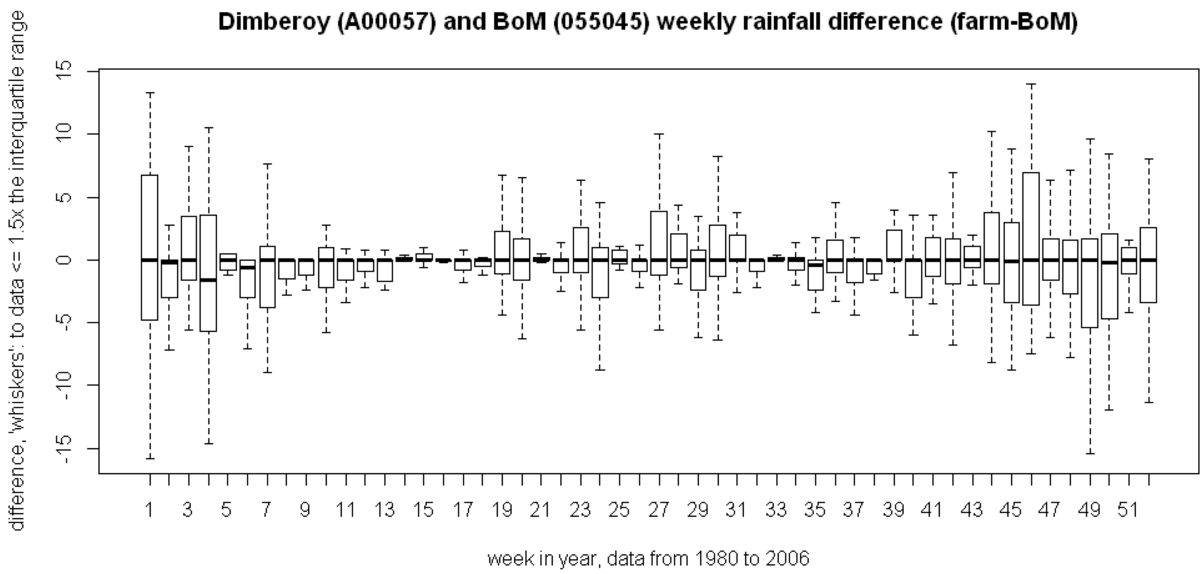


Figure 5-8 Box and whisker plot of difference between weekly rainfalls, Dimberoy data set

While it is tempting to suggest this shows a seasonal pattern, the scale of the rainfall events at different times of the year must be considered, with the same summer maxima displayed in weekly rainfall totals, as shown in Figure 5-7.

Figure 5-9 shows standardised weekly difference values for the Dimberoy data set. Standardisation has been achieved by dividing the weekly difference values (farm – BoM) by the weekly BoM values. Figure 5-9 shows three plots, for the different farms in the data set, but with respect to the same BoM station, on the same scale. The distances of the respective farms to the BoM station are :

BoM station 055045 to	farm	A00057	13.077 km
	farm	A00053	27.862 km
	farm	A00055	56.659 km

Figure 5-9 indicates that there is no outstanding seasonal pattern to the difference figures for the Dimberoy data set. The average values (A00057=0.436, A00053=0.115, A00055=0.691) for the whole difference data set at each site do not indicate significant differences between sites, with A00053 possibly being reduced due to missing data (see Figure 5-1). The slightly higher average difference for A00055 may be due to the greater distance from the BoM station, which would be expected to result in a greater difference, but with limited data this cannot be confirmed. There are clearly many questions that could be asked about the relationships between the farm and BoM data, for example involving variability with respect to scale in wetter and drier times, or distance effects across the data set, but time and resources prohibit further examination in this project. It is hoped that such examinations will be available in future studies, and that by that time more farm data will be available to use in analyses.

Figure 5-10 shows a summary of simple (individual) linear model analyses of three farms in different areas of NSW and the nearby BoM stations, for weekly rainfalls from 1980 to 2006. The models use the BoM data as a predictor for the farm data.

Figure 5-10 shows, from top to bottom, linear model slopes, intercepts and R squared values for the three sites, with Dimberoy in blue, Willow Glen in red and Munyalba in green. Dimberoy and Willow Glen are in the Liverpool Plains area of northern NSW (near Narrabri), and Munyalba is in southern NSW (near Wagga Wagga). The selection process outlined previously in section 3.3 was used, providing nearby BoM stations to farm sites for analysis. An example of the site locations for the Dimberoy data set is shown in Figure 5-1.

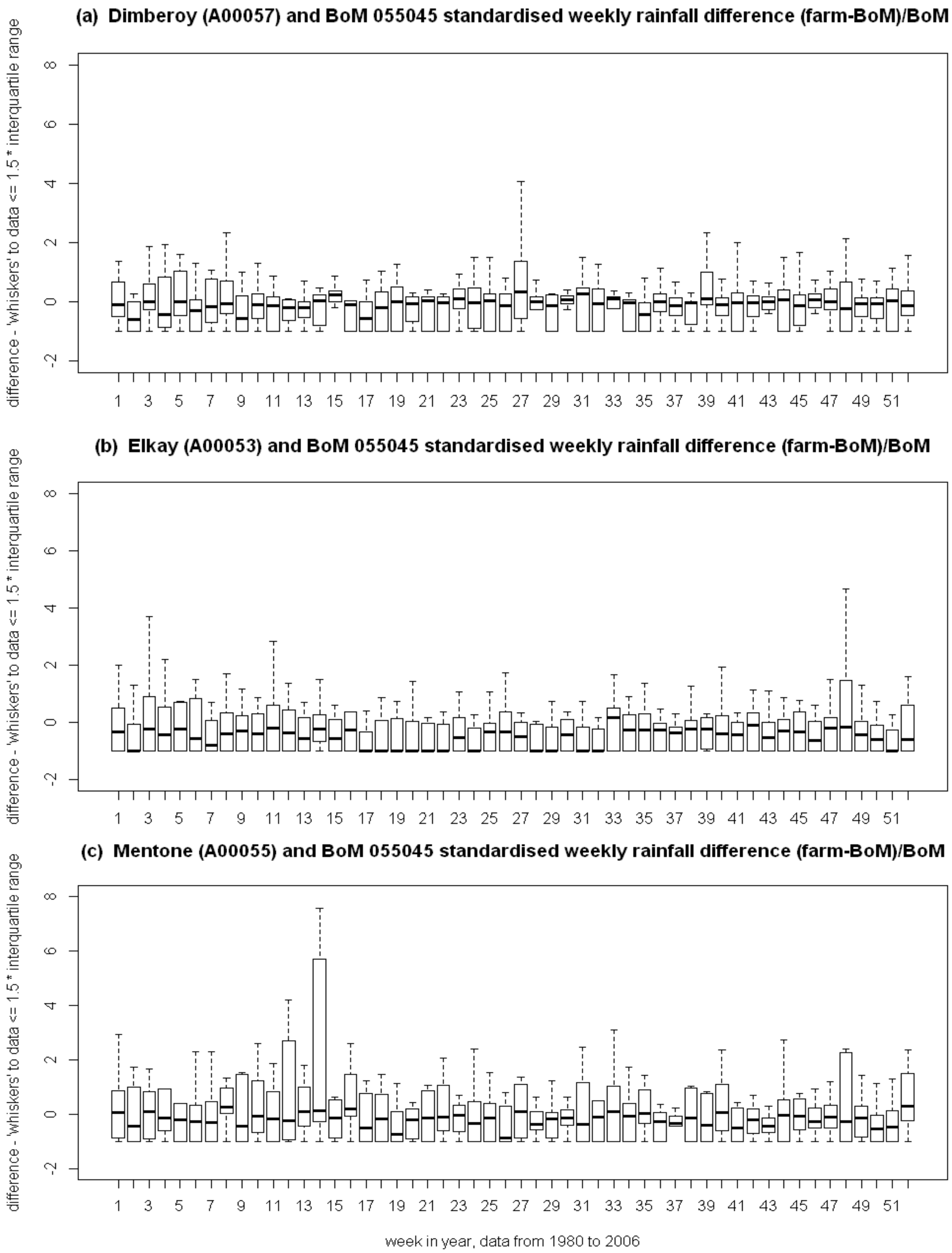


Figure 5-9 Standardised weekly rainfall differences, farms in Dimberoy data set

Figure 5-10 Linear model results for vaious farms and nearby BoM sites

To check the significance of the R squared vales, t values can be calculated based on sample sizes of 20, 11 and 17 sites for Dimberoy (see Figure 5-1), Willow Glen and Munyalba respectively. The t values show that critical R squared values for significance are 0.16, 0.28 and 0.19 for Dimberoy, Willow Glen and Munyalba respectively, indicating that all the R squared values, while of varying strengths, are significant (Webster and Oliver 2007).

Once again, while it is tempting to infer that the model intercepts imply slightly different relationships in wetter and drier times, this does not hold up with greater inspection. The higher R squared values seem to relate to more uniform slope and intercept values, and, given that the rainfall at Munyalba in southern NSW is more uniform in winter and summer than the Liverpool plains,[23] which would experience a summer maximum, the similarities between Dimberoy and Munyalba argue against this inference.

While the differences between individual farm sites and BoM stations is discernible for particular events (Figure 5-3 Daily rainfall averages from the Dimberoy data set) and for periodic summations as shown by the previous chapter, there does not appear to be any consistent pattern which would

[23] http://www.bom.gov.au/jsp/ncc/climate_averages/rainfall/

allow reliable prediction of farm data from differences to BoM sites. This highlights the value of the farm data in fine tuning seasonal or other forecasts for local conditions.

5.2 Unique predictor tests

Section 4.3.2 introduced the Deficit / Surplus (DSP) method of displaying temporal data to provide a continuous display of 'difference from normal' information. While the DSP is a novel and more useful way to display such information, the hypotheses here is that the data collected on their own properties can be used, with the DSP, to provide predictive information for rainfall on the property. The DSP method, by virtue of being a cumulative method, relates to soil moisture storage, that is how empty (deficit) or full (surplus) the soil moisture store is. The DSP can be thought of as a potential parameter in a water balance model for the system. While the relationship between DSP and soil moisture store will not be further examined or quantified in this study, it is believed that users will understand the DSP concept, and so a new version of the DSP with some predictive capacity would be useful. The DSP plots used in the feedback report (copy below) show a selected year farm DSP and boundaries defined by high and low rainfall years.

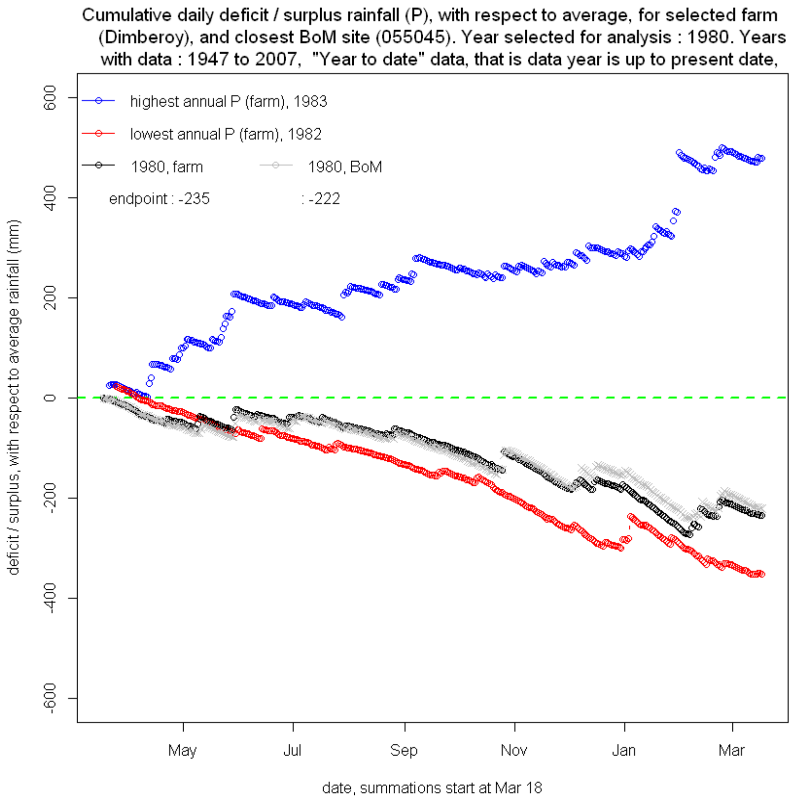


Figure 5-11 Copy of Figure 4-6, 'Example of Feedback report output, DSP data for selected farm and BoM station, 'year to date' data'

Figure 5-11 shows 'year to date' data, that is an analysis that takes data for 365 days prior to the date of analysis as the data 'year'. The rationale is to allow for real time data analysis, when data is available, and also to provide users a full year of analysis irrespective of the date of analysis. While the year for analysis would normally be the current year when using real time data, users would be able to select any year and specify the period of analysis, such as seasonal, in interactive or real time versions.

5.2.1 Predictor selection

To test a predictive analysis, an 'analogue year' (or years) will be selected which will be used to match, or 'predict' the known DSP for a selected calendar year, with the prediction progressing from the date of analysis forward to the end of the calendar year selected, using the Dimberoy data and a comparison site for the analysis. Figure 5-12 shows an example of the resulting predictive analysis, for the arbitrarily selected year of 1980, as used previously. Figure 5-12 shows predictive 'boundary' lines, in black and grey in the figure, which are simply the DSP values for the previously drawn wettest (blue) and driest (red) years redrawn, starting from the date of analysis ('year to date').

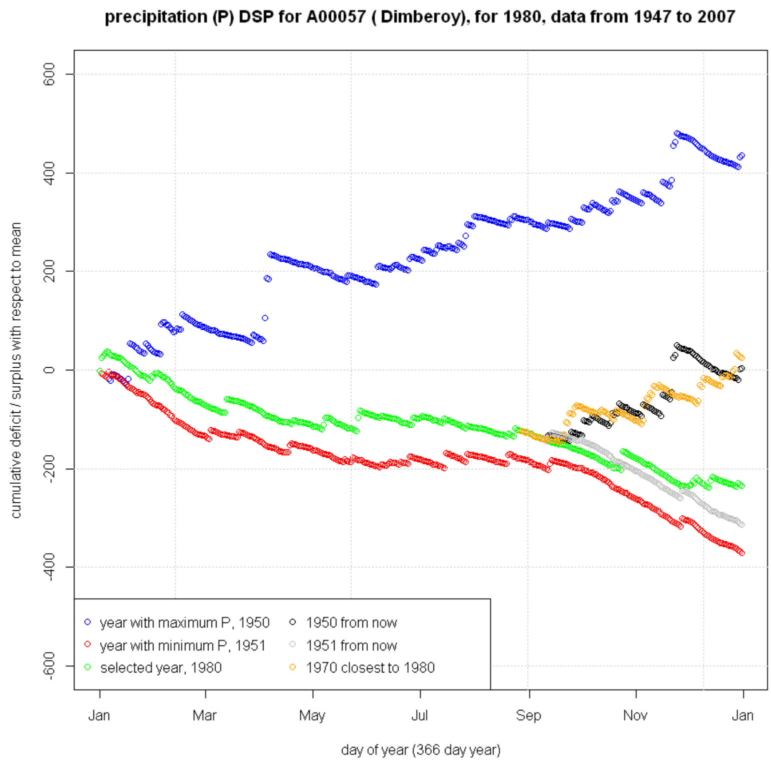


Figure 5-12 DSP predictive plot, Dimberoy data set 'Closest ytd P'

The ‘analogue or predictor year’ (orange line) in Figure 5-12 is the year 1970, the year with the cumulative total rainfall value (P) closest to that of the year selected for analysis at the date of the analysis (‘year to date’). The DSP for the selected year going forward from the date of analysis (green line) is the line the predictive data (orange line) should match. The green line (future DSP) would clearly not be present in a real time analysis, but shows the closeness of the analogue year match.

It was initially thought that years with similar rainfall values to the selected year ‘year to date’ may have similar patterns for the remainder of the year, and so provide the required analogue. A number of other criteria were selected to test the validity of this option, and possible improvements, as listed in Table 5-2 below. The ‘mean, 5% and mean, 10%’ criteria were used to select years for which average DSP values would be used as analogue.

Table 5-2 Analogue year selection criteria

label	detail
closest P	the DSP for the year with the ‘year to date’ total rainfall value closest to the ‘year to date’ total rainfall for the selected year
mean, 5% P	the mean DSP for years with total ‘year to date’ rainfall within five percent of the total rainfall for the ‘year to date’ value for the selected year
mean, 10% P	the mean DSP for years with total ‘year to date’ rainfall within ten percent of the total rainfall for the ‘year to date’ value for the selected year
closest ytd DSP	the DSP for the year with the ‘year to date’ DSP value closest to the ‘year to date’ DSP for the selected year
10% P, ytd SOIp	the DSP for the year with the total ‘year to date’ rainfall value within ten percent of the value of the total ‘year to date’ rainfall for the selected year, and with the average monthly ‘year to date’ Southern Oscillation Index phase(SOIp) value (Stone and Auliciems 1992) closest to the average monthly ‘year to date’ SOI phase value for the selected year

The criteria in Table 5-2 are based on similarities between the rainfall for selected year and other years, with the Southern Oscillation Index phase (SOIp) value providing input from an external climatic driver. While these criteria are quite simple, it was hoped that evidence of a predictive capacity would be demonstrated within the limited time and data constraints of this project. A sequence of years at five year intervals from 1962 to 2007 was chosen for analysis. This period was selected to allow similar analyses at a different location, 'Cooreena Park', near Dubbo (NSW), with a different climate type[24], with data from 1962 to 2010, for comparison with the Dimberoy analyses.

5.2.2 Results and discussion

Figure 5-13 and Figure 5-14 show some examples of predictive DSP plot results for Dimberoy and Cooreena Park, using the 'closest P' prediction year. Figure 5-13 shows plots with what appear by eye to be quite good matches to the selected year values, while Figure 5-14 shows some plots with fairly poor matches to the selected year data.

The better matches of Figure 5-13 are very encouraging, demonstrating that the DSP analysis can demonstrate a predictive capacity. Unfortunately the selection of an analogue for the DSP values is not simple, and the 'closest P value' is clearly not appropriate in all cases.

For example, in Figure 5-14, poor matches are apparent whenever the selected year or analogue year is near the extremes of the range (b, c and d), and even if the match is quite good at some points (a, b) the variations between analogue and selected year values can be large. It would be expected that extreme value years would have few equivalent years, and so this result would also be expected. Figure 5-14 (b) also shows that variability between years can allow the match to be quite good for one or two months and then diverge (c), or run with an apparently constant offset from the selected year data (a).

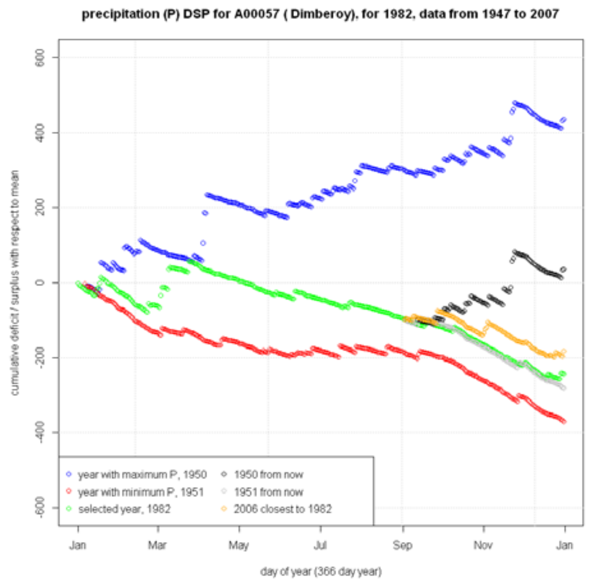
Table 5-3 lists R squared values from a simple linear model with the selected year DSP as the response, and the year with the closest P value at 'year to date' as the prediction term. With about four months of daily DSP data used in the linear models, the R squared values, while once again indicating variable strength relationships have a threshold value for significance of 0.029 according to t value tests based on sample size (Webster and Oliver 2007). Significant R squared values in tables 5-43 to 5-7 are highlighted.

The following tables (Table 5-4 to Table 5-7) list R squared values from a simple linear model with the selected year DSP as the response, using the various other prediction term options, as described in Table 5-2.

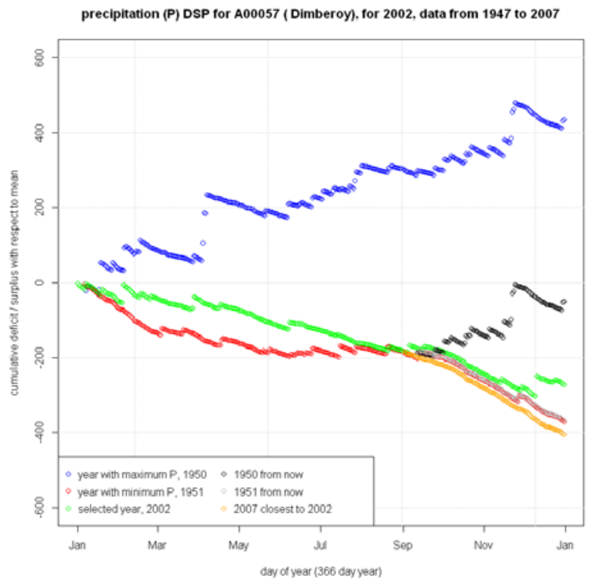
[24] http://www.bom.gov.au/jsp/ncc/climate_averages/climate-classifications/index.jsp?maptype=seasgrpb

Figure 5-15 shows all the R squared values for the DSP predictions, and the annual rainfalls for the Dimberoy and Cooreena Park data sets. The farm initials, D and CP, are used as suffixes to the prediction term labels in the tables and Figure 5-15 to indicate the data set being used.

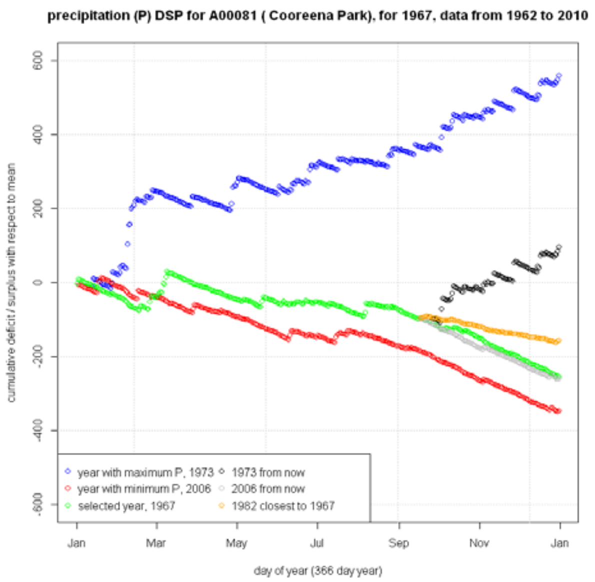
a



b



c



d

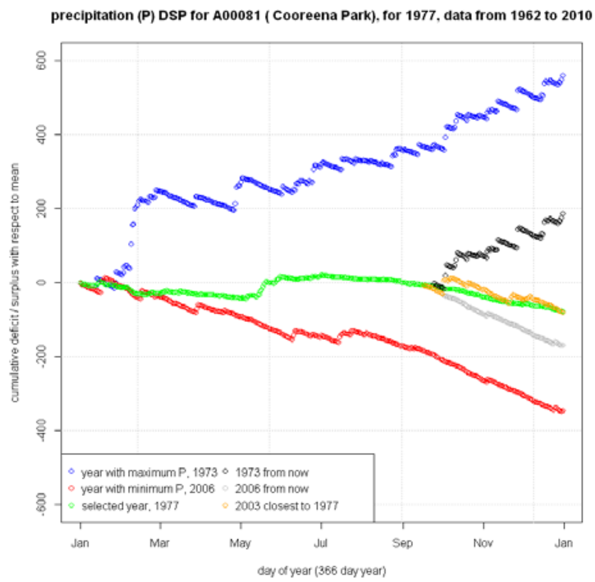


Figure 5-13 Examples of predictive DSP analysis, using closest P value to select prediction year, for years 1982 (a) and 2002 (b) at Dimberoy 1967 (c) and 1977 (d) at Cooreena Park

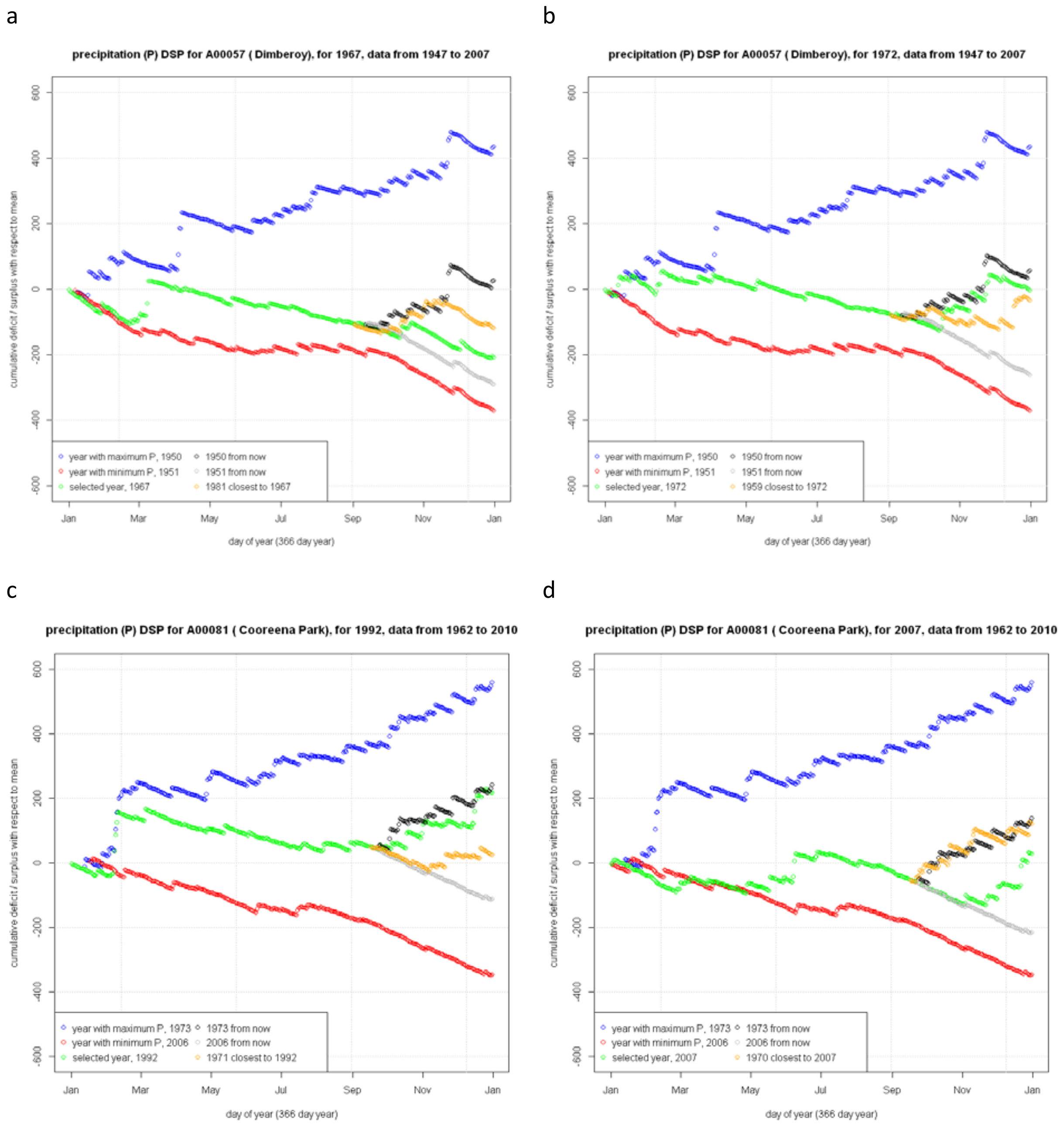


Figure 5-14 Examples of predictive DSP analysis, using closest P value to select prediction year, for years 1967 (a) and 1972 (b) at Dimberoy, and 1992 (c) and 2007 (d) at Cooreena Park

Table 5-3 R squared values from linear model, closest P, for Dimberoy (D) and Cooreena Park (CP) data sets

year	1962	1967	1972	1977	1982	1987	1992	1997	2002	2007
prediction year (D)	1966	1981	1959	1950	2006	1952	1974	1999	2007	2002
R squared	0.136[25]	0.020	0.022	0.733	0.900	0.687	0.616	0.165	0.813	0.813
prediction year (CP)	1982	1982	1966	2003	1962	1964	1971	2004	2008	1970
R squared	0.606	0.982	0.154	0.784	0.606	0.023	0.003	0.022	0.478	0.006

Table 5-4 R squared values from linear model, Dimberoy data set, mean, 5% P

year	1962	1967	1972	1977	1982	1987	1992	1997	2002	2007
No. years	6	7	1		5	4	6	4	2	1
R squared	0.038	0.051	0.523		0.133	0.626	0.110	0.222	0.499	0.514

Table 5-5 R squared values from linear model, Dimberoy data set, mean, 10% P

year	1962	1967	1972	1977	1982	1987	1992	1997	2002	2007
No, years	10	9	11		9	10	14	7	4	2
R squared	0.306	0.485	0.544		<0.001	0.315	0.871	0.750	0.684	0.417

Table 5-6 R squared values from linear model, Dimberoy data set, closest ytd DSP

year	1962	1967	1972	1977	1982	1987	1992	1997	2002	2007
year	1954	1981	1959	1998	2006	1986	1974	1969	1957	1960
R squared	0.013	0.688	0.512	0.365	0.718	0.300	0.209	0.063	0.667	0.017

Table 5-7 R squared values from linear model, 10% rain, ytd SOIp, for Dimberoy (D) and Cooreena Park (CP) data sets

year	1962	1967	1972	1977	1982	1987	1992	1997	2002	2007
prediction year (D)	1966	1961	2006		1981		1997	1992		1979
R squared	0.136	0.286	0.792		0.194		0.229	0.229		0.066
prediction year (CP)	1966	2009	2004		2005				1976	1979
R squared	0.741	0.113	0.011		0.883				0.433	0.043

[25] Statistically significant R squared values highlighted

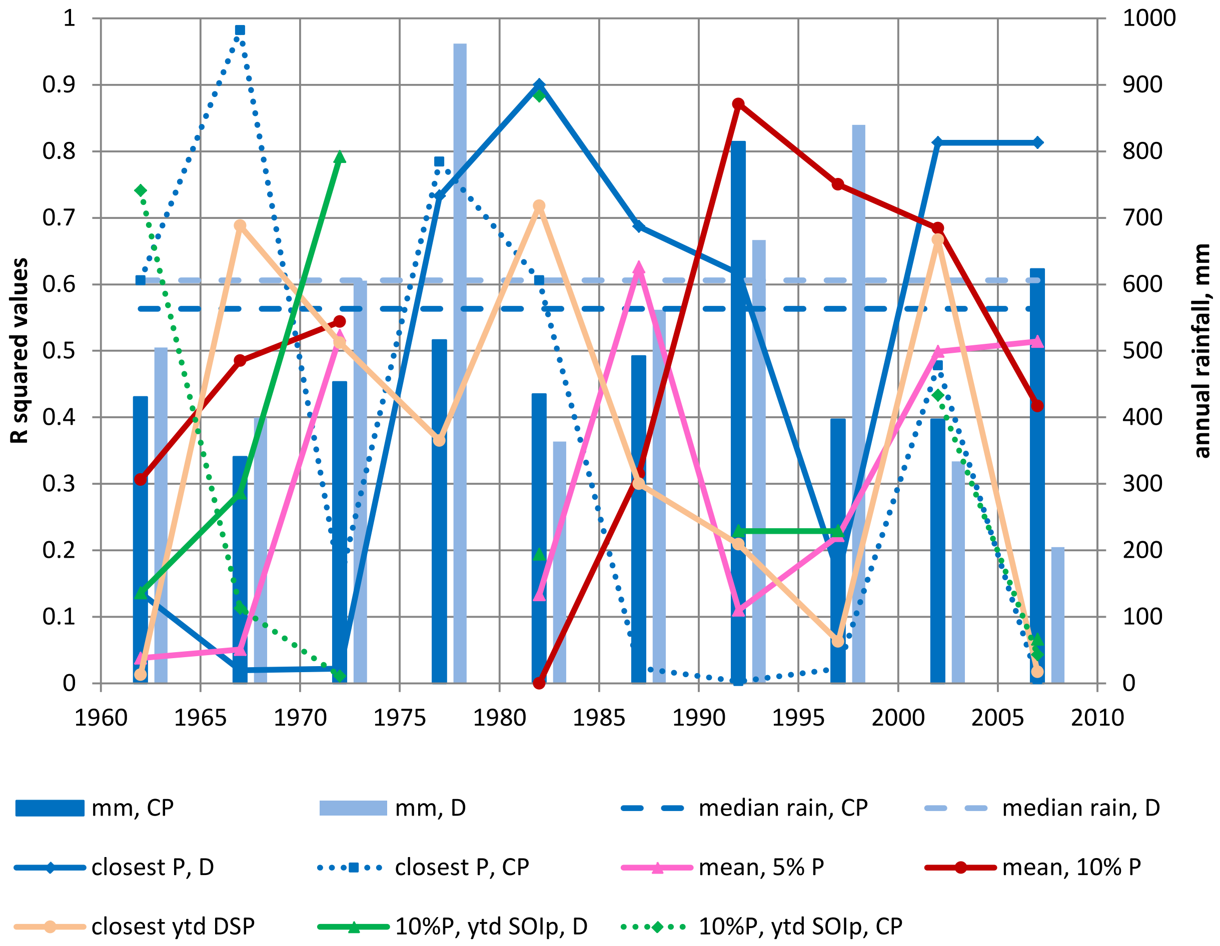


Figure 5-15 R squared values from various linear models for DSP predictors and annual rainfalls, Dimberoy and Cooreena Park data sets

Figure 5-15 and the tables above show quite variable matches and correlations between the response values and the predicted response values. Some very good matches demonstrate the predictive capacity of the DSP analysis using local data, while the poorer matches indicate that the ideal prediction year or option has not yet been discovered. It is possible that a nearest neighbour or fuzzy clustering time series analysis will identify the ideal predictor for the DSP predictive process. This analysis is not within the scope of this study, but is potentially part of the broader program of which this study is a part (Montazerolghaem, Vervoort et al. 2012), (Plain, Minasny et al. 2008).

It is interesting to note that from Figure 5-15 that the simplest, initial method of selecting a predictor year, based on rainfall similarity alone, selects predictor years with matches as good as other, more complex methods of predictor selection.

5.3 Conclusions

It appears from the scaling factor tests that there is not a simple pattern which could be used to define a local relationship between individual farm sites and nearby BoM stations, once again demonstrating the value of the farm data in characterising the local conditions.

The lack of a simple relationship between the sites, across different climatic zones once again reflects the spatial and temporal variability of daily rainfall. It would be interesting to re-examine these relationships when more data is available, possibly allowing detection of scale or event intensity effects.

The unique predictor tests show very promising results, while the selection of the ideal predictor presents an exciting challenge. This further demonstrates the value of the local rainfall data, and allows a conviction that the historic rainfall data alone will provide a predictive capacity to increase the relevance and usefulness of the DSP analysis. The predictive DSP analysis can provide valuable, unique information to users, adding to their tools for farm operation and risk management strategies.

An interesting area for examination would be the applicability of predictive DSP analysis to extreme events, which may be increasingly important under the influences of climate change.

6 Adding value to forecasts with farm data, spatial data

An increase in the spatial density of rainfall data will, intuitively, allow regional forecast information to be interpolated to a finer scale, giving more detailed information for specific locations. As noted in the previous chapter, the second aim of the project :

- By exploring patterns in time and space in historic rainfall data, and the relationship between farm data and other available data, test the usefulness of the farm data, and test whether the farm data can have value in a predictive or interpretive role, and therefore aid in farm management decisions.

involves potential predictive or interpretive roles for the historic rainfall data. The purpose of this chapter is to report on the potential of the historic rainfall data for revealing spatial patterns which could help farmers to interpret current forecasts.

6.1 Interpretive value tests

To demonstrate an interpretive potential, it is necessary to test whether the analytical systems developed demonstrate variation at the scale of interest, providing a system which can be confidently used by farmers to interpret current forecasts, as well as historical data.

The kriging method for spatial interpolation of rainfall data was introduced in section 4.3.3 (see, for example, Figure 4-8 Example of Feedback report output, spatial interpolation of data for selected farm and nearby area). This method was used to generate the spatial interpolations of rainfall data for the feedback report previously discussed, which will be further developed here. It is worth repeating that kriging is an exact interpolation method, meaning that the method maintains the known site data at the various rainfall sites, as shown in the interpolation maps.

A simple technique to quantify the spatial variation is to compare the interpolated values close to the selected farm, when including farm data in the interpolation as opposed to an interpolation excluding farm data.

The same Dimberoy data set as previously described will be used for the analyses discussed here. As with the previous chapters, 1980 will be used in the procedures described here, with the period of April and May as a seasonal subset, being of interest as representing the planting season for much of the grain growing area in south east Australia. Figure 6-1 shows the data sites available for the analysis, nineteen non farm sites and three farm sites, as compared to the 55 potential data sites with seven farm sites sown in Figure 4-1. The sites excluded did not have data in the years or period selected for analysis. The Cooreena Park data set, also presented previously, will be used as a comparison site for the Dimberoy data set, being a site with a different climate type[24]. Details and maps for the Cooreena Park analysis are included in Appendix 7, Spatial data for chapter 6.

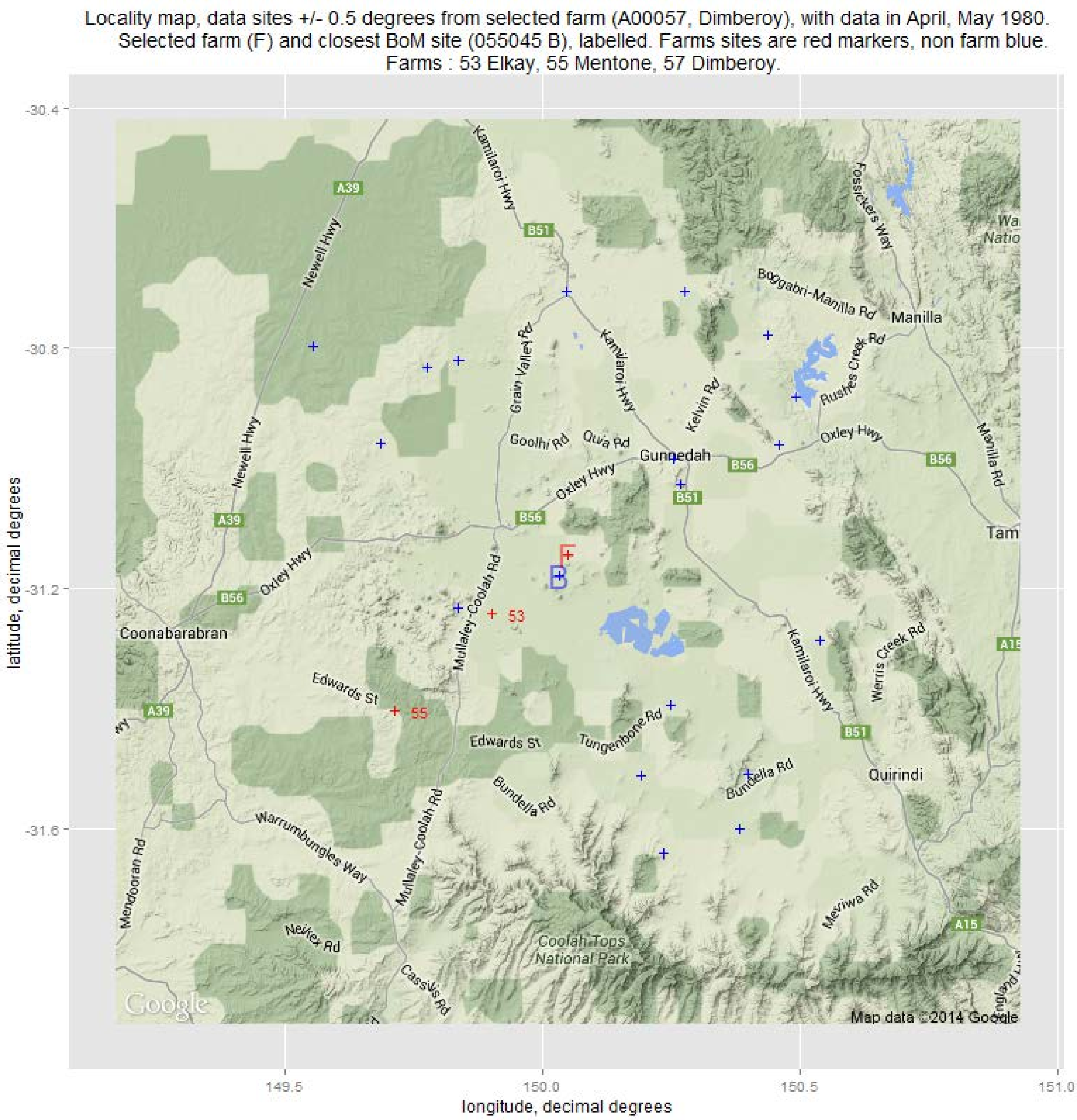


Figure 6-1 Map of data sites for spatial interpolation analysis

Figure 6-2 shows a map of spatially interpolated rainfall for the area near the selected farm, in April and May 1980, using all the available data for the analysis. Included on Figure 6-2 is a box marking an area surrounding the selected farm of twenty by twenty kilometres.

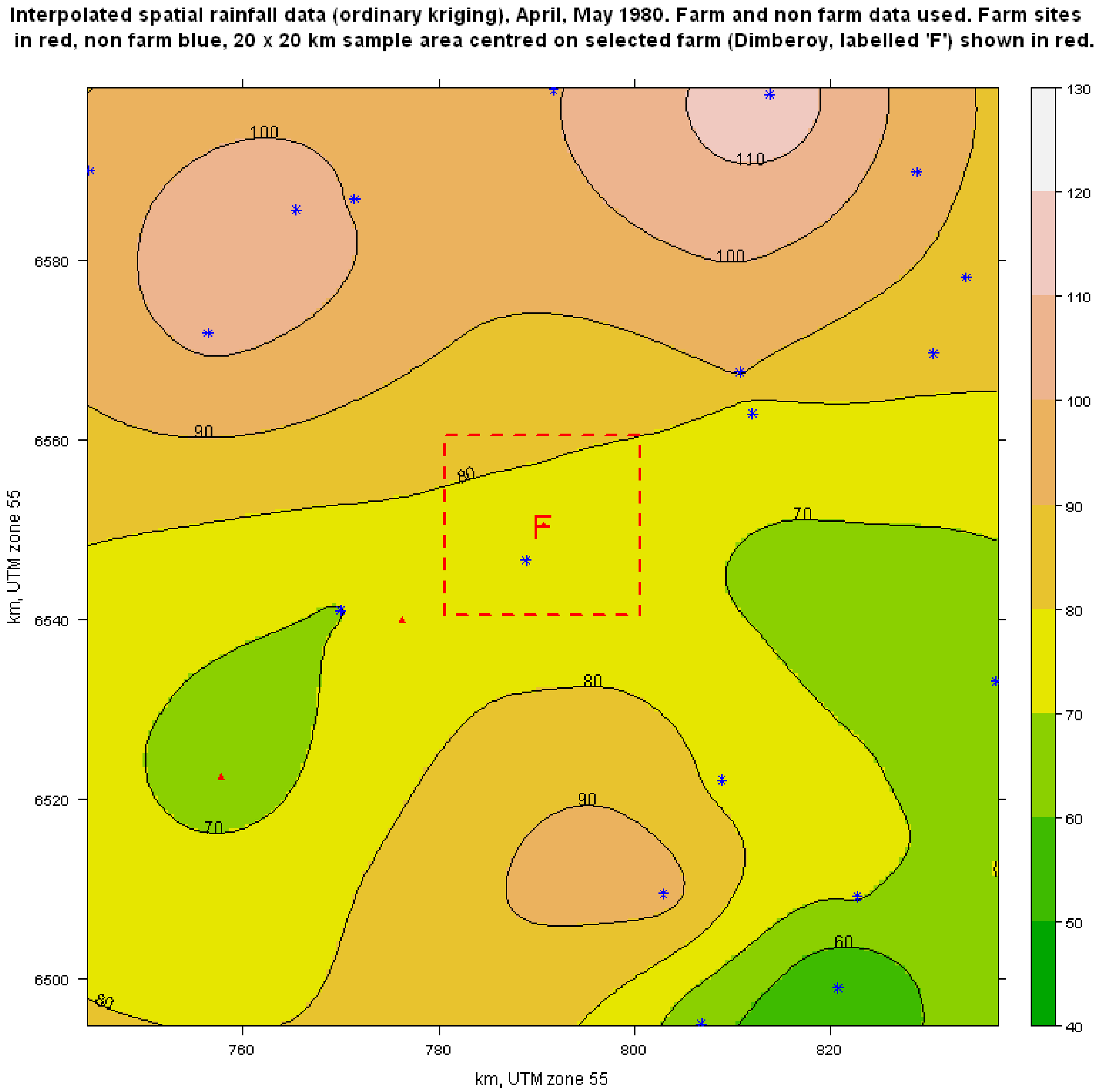


Figure 6-2 Spatial interpolation of data for selected farm and nearby area, all data

Figure 6-3 shows another map of interpolated rainfall for the area near the selected farm. Unlike the previous map, this interpolation was performed excluding farm data.

As with Figure 6-2, Figure 6-3 also shows an area surrounding the selected farm. Data from the bounding areas has extracted from the interpolation data sets for comparison.

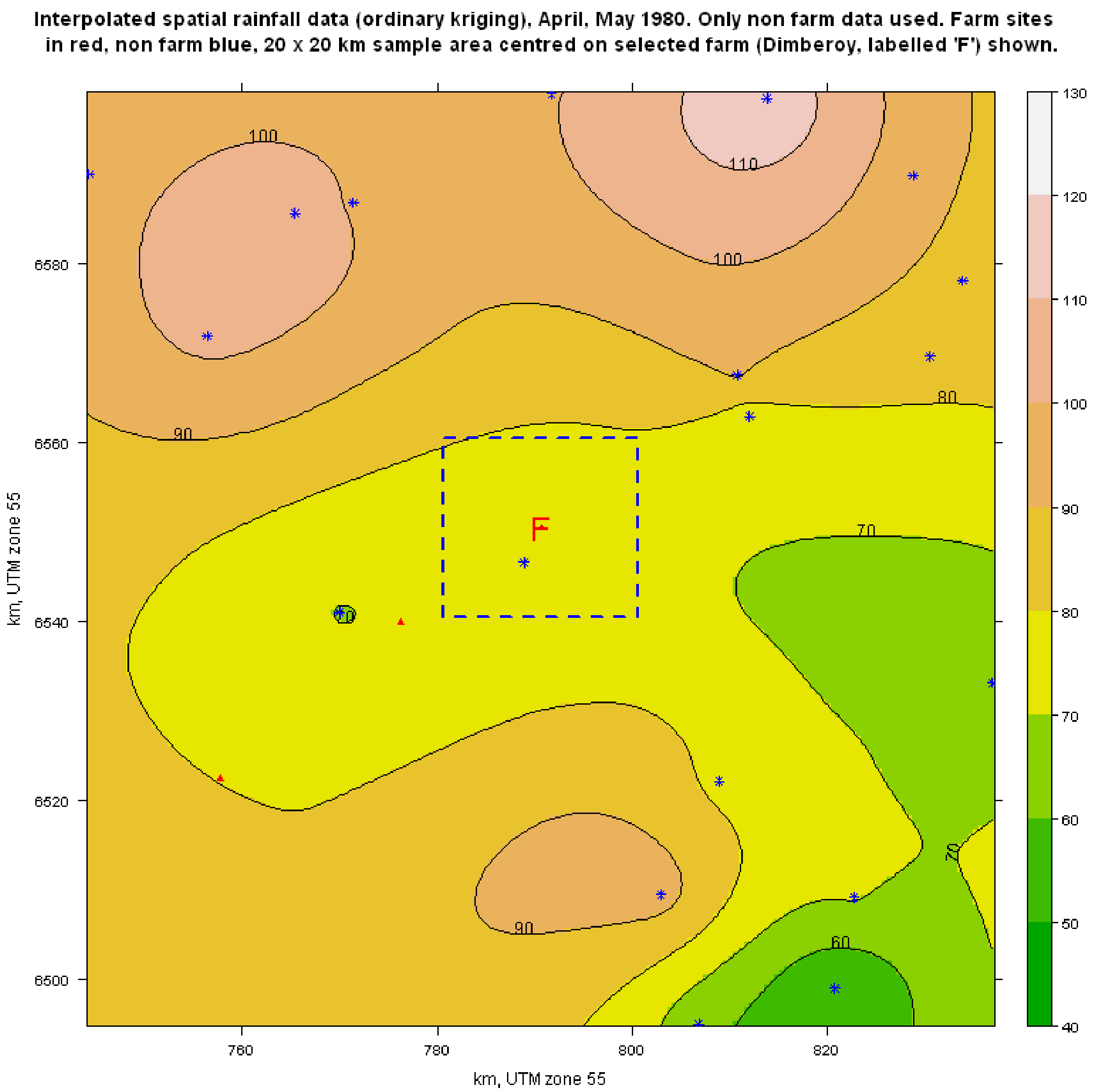


Figure 6-3 Spatial interpolation of data for selected farm and nearby area, non farm data only

6.2 Results and discussion

The data extracted by the method described above can be displayed as a histogram, the number of interpolated rainfall sites in categories partitioning the size of the events, as shown in Figure 6-4. The interpolated data set (kriging) generates interpolated data sites at a much finer grid than the original data points. The selected farm rainfall for the period is also shown on Figure 6-4, by the red marker on the rainfall (x) axis.

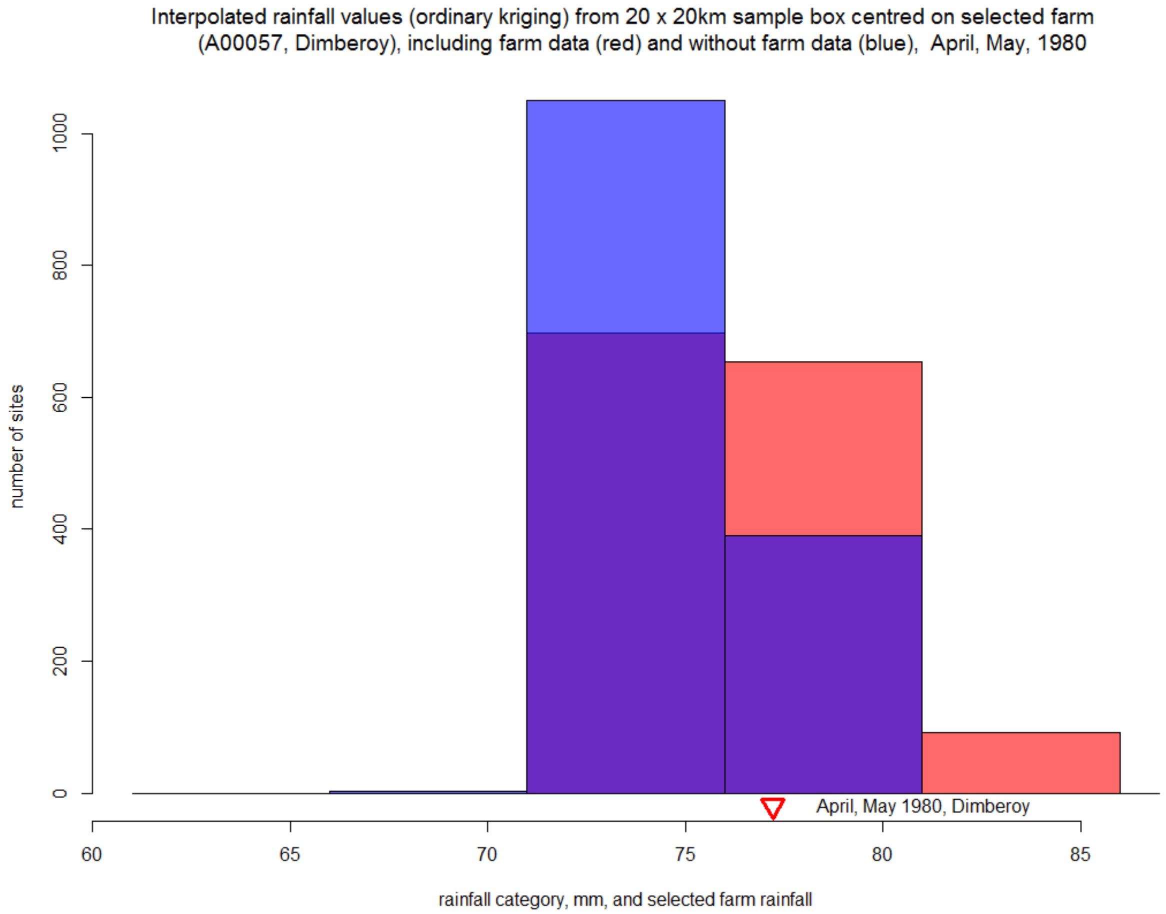


Figure 6-4 Histogram of interpolated rainfall values near the selected farm, with (red) and without farm data (blue)

For this example, the Dimberoy data set, the counts (number of interpolation data points or sites) for the respective categories are shown in Table 6-1.

Table 6-1 Counts of interpolated data points (sites) in histogram analysis

rainfall category, mm	interpolation with all data	interpolation excluding farm data
61 – 66	0	0
66 – 71	0	3
71 – 76	698	1050
76 – 81	654	391
81 – 86	92	0
86 - 91	0	0

From this data, it can be seen that excluding the farm data from the interpolation causes a 51% increase in the lower intensity rainfall sites (up to 76mm) at the expense of higher intensity sites.

While the total rainfall near the farm is not shown as being significantly different, the pattern of rainfall values across the sites is demonstrably different. The interpolation plots do not show peaks or troughs around the farms in the data set, and during the analysis the operator has the opportunity to remove data sites that appear to differ from the other data, hence there is no reason to reject the finding as simply displaying inaccuracies in the farm data.

To further test this result, similar analyses for three wet and three dry years at Dimberoy are summarised in the figures below. The figures are reduced in size to allow for comparisons. See Appendix 7 for larger plots and associated data.

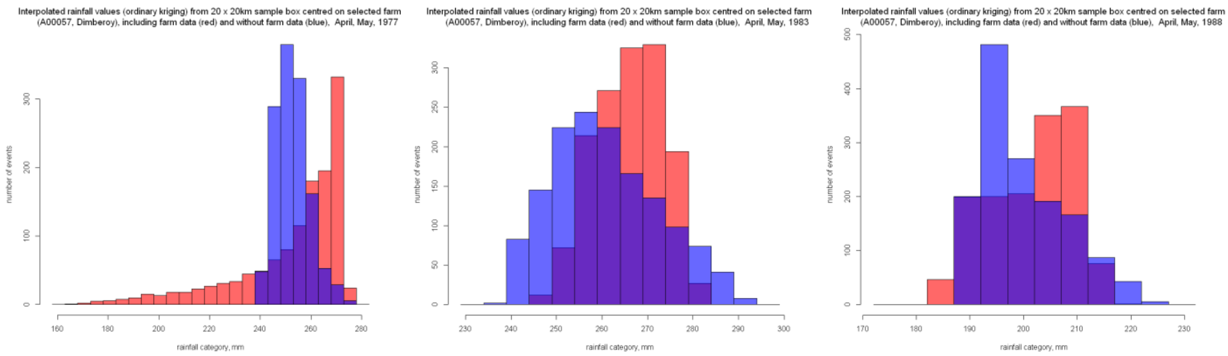


Figure 6-5 Histograms of interpolated data values from near Dimberoy, for (left to right) wet years 1977, 1983, 1988

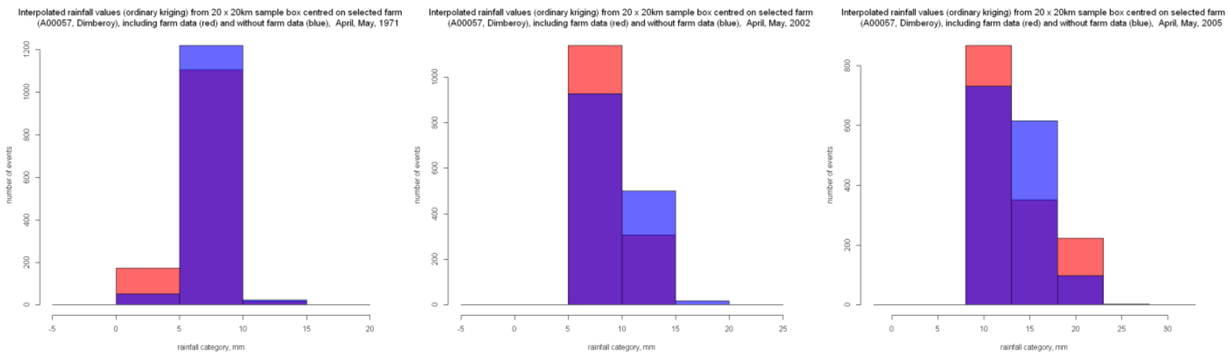


Figure 6-6 Histograms of interpolated data values from near Dimberoy, for (left to right) dry years 1971, 2002, 2005

The number of years (three) selected for these analyses, and the years selected, were chosen to include as many data sites as possible from within the wettest and driest years. Table 6-1 shows the rainfall totals for the season (April and May) for the sites and years selected. As can be seen, not all sites had data in the relevant periods, and so years were selected from those of the wettest or driest five which had at least two data sites.

Table 6-2 Seasonal (April and May) rainfall totals for Dimberoy, nearby farms and closest BoM site with data in respective years, 1977, 1983, 1988 wet, 1971, 2002, 2005 dry

Site / year	Dimberoy A00057	Elkay A00053	Mount Nombi A00054	Mentone A00055	Womponia A00056	Beulah A00086	closest BoM to Dimberoy 055045
1977	272.5	134.0					248.5
1983	274.8	285.2		285.2			268.6
1988	208.3	175.5		189			197.2
1971	4.8		8.89				5.1
2002	11.0	5.5		20.3	5.5		11.6
2005	63	28.5		58.4	44.5	37	10.8

Kolmogorov-Smirnov (two sided k-s) test performed on the extracted data sets (including and without farm data), indicated that the data sets could be treated as different distributions for all the years tested (Johnson 1999), (Marsaglia, Tsang et al. 2003), with p values from the Dimberoy tests all less than 1e-05.

Figure 6-7 shows a similar summary for wet years for sites near Cooreena Park (near Dubbo, NSW).

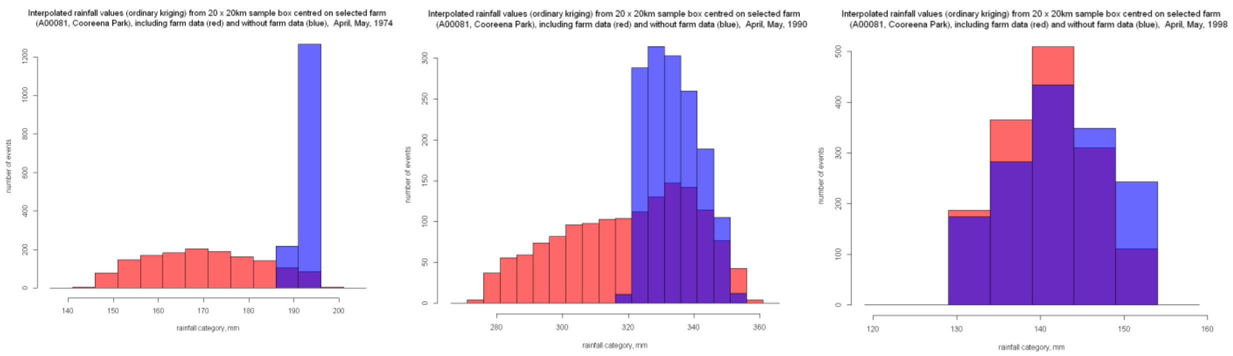


Figure 6-7 Histograms of interpolated data values from near Cooreena Park, for (left to right) wet years 1974, 1990 and 1998

This site was used for comparison because it represents a different environment from Dimberoy (BoM 'uniform rainfall' as compared with 'summer rainfall' at Dimberoy)[26], and also because the Riverina site used in previous analyses (Munyalba) did not have another nearby farm site with data in sufficient years for comparison.

Table 6-3 Seasonal (April and May) rainfall totals for Cooreena Park, nearby farms and closest BoM site with data in respective years, 1974, 1990 and 1998, wet

site / year	Cooreena Park A00081	Ashlee A00085	closest BoM to Cooreena Park 065012
1974	165	138.4	192.8
1990	329	255.3	345.7
1998	135	154.4	153.4

The plots in Figure 6-7 are similar in form to those in the original comparison shown in Figure 6-5 for Dimberoy. The figures show that excluding the farm data from the interpolation causes an increase in the lower intensity rainfall sites at the expense of higher intensity sites for the wet years. Figure 6-6, showing the analysis for the dry years at Dimberoy, does not show any real differences when the farm data is excluded from the analysis. This could be due to the small number of events and the scale of the events. It is possible that some differences may emerge with more data, which would allow a finer scale interpolation and analysis.

The fact that Figure 6-7 shows an opposite trend (compared with Dimberoy wet years) in two out three wet years strengthens confidence in the farm data. That is, while clear differences in the analyses are indicated, the differences are not similar at different locations, which could imply that the farm data consistently under or overestimated rainfall values, further implying a consistent error across farm data.

6.3 Conclusions

I believe that it is quite clear that the farm data provides information which can help users interpret regional or seasonal forecasts in greater detail to supplement farm management operations. Many farms have extensive rainfall records, for example over one hundred years at Dimberoy, and so such

[26] http://www.bom.gov.au/jsp/ncc/climate_averages/climate-classifications/index.jsp?maptype=seasgrpb

analyses should provide reliable results. Using the example above, knowledge of the possible patterns of event rainfalls could allow operators to plan activities such as heavy traffic operations and timing.

The information presented here shows how the use of farm data can help understand how a season or period may develop by comparison with historical data, and how local conditions can influence the spatial distribution of rainfall. Both these factors could be useful in helping farmers schedule their operations or aid in risk management strategies. I believe it is important to demonstrate this to the farming and wider community, to make evident the value of existing farm data, and encourage its use. The differences displayed in the spatial data are very encouraging, and I believe that if more farm data were available, it would be possible to better define these differences, and how they vary in space across the regions.

The addition of extra data sites will clearly improve the opportunity for detailed analysis and the precision of such analyses. As previously noted, variogram precision is increased with data density, with data sets iudeally in above one hundred sites (Webster and Oliver 2007). The increased precision will allow more robust analyses and understanding of errors, uncertainties and variability in the results. Increases data density also allows examination of rainfall variability at temporal different temporal, scales, both at individual stations and spatially. The evaluation of cyclic behaviour at different temporal scales could be related to regional climate drivers and therefore be significant in climate change research (Beecham and Chowdhury 2010), (Chowdhury and Beecham 2010).

7 Indicator time period

7.1 Aim

This aim of this section is to test the third aim of the project:

- To test the idea of an indicator period or season, which can be used to predict annual or seasonal rainfall, and whether such indicator periods persist across regions.

If a single month or period of time is representative of the annual or seasonal rainfall at one or more locations, the specified time periods could also correlate with grain yields for the year or season. Various authors have reported or specified particular time periods as being more influential on wheat yields than other periods or average values. This study is focused on the link between rainfall in specific periods and annual or seasonal rainfall, once again emphasizing the use of existing farm data.

Trumble (1937) noted that the realization of the significance of rain during the growing period had led to the arbitrary choice of April to October or April to November rainfall as a measure of the seasonal precipitation. Hounam (1947), reported that eight inches (203 mm) of rain during the months of April to August inclusive was considered desirable for satisfactory (wheat) production, with ideal amounts in each calendar month being specified. Cornish (1949) decided that the four periods: April and May; June, July and August; September and October; and November, could be defined as rainfall variates allowing for seasonal variation. Later however, Cornish (1950) claimed he could not justify the conclusion (by 'many workers') of critical periods or months which would most strongly influence wheat yields. Sadras, Roget et al. (2002) concluded that April rain anticipated seasonal rainfall in western Victoria and northeast South Australia, and could allow growers to 'derive thresholds to discriminate between likely wet and likely dry seasons on the basis of their own rainfall records'. Sadras claimed that the significance of April rainfall could be justified in a climatic sense as being 'consistent with the early autumn shift from cyclonic to frontal rainfall'. Later Sadras and Rodriguez (2007) reverted to using seasonal and annual data in an investigation into spatial rainfall patterns.

None of the literature (read by the author) has fully examined the claims of any relationship between monthly or periodic and seasonal or annual data, while there is general acceptance on monthly or multiple month intervals being of relevance in relation to yield data (Trumble, 1937, 1939, Wark 1941, Hounam 1947, 1950, Fitzpatrick 1970, Sadras 2002, 2007).

7.2 Data

In order to test for a relationship between the rainfall for some time period and seasonal or annual rainfall, it will be necessary to select which time periods will be relevant, and then test these periods for seasonal or annual relationships at more than one location. The Bureau of Meteorology (BoM) has ninety nine rainfall districts covering the Australian continent, which were developed 'partly in an effort to group sites with relatively similar rainfall climates'[27], which will provide a convenient way to select data sites. The locations of the BoM rainfall districts are shown in Figure 7-1.

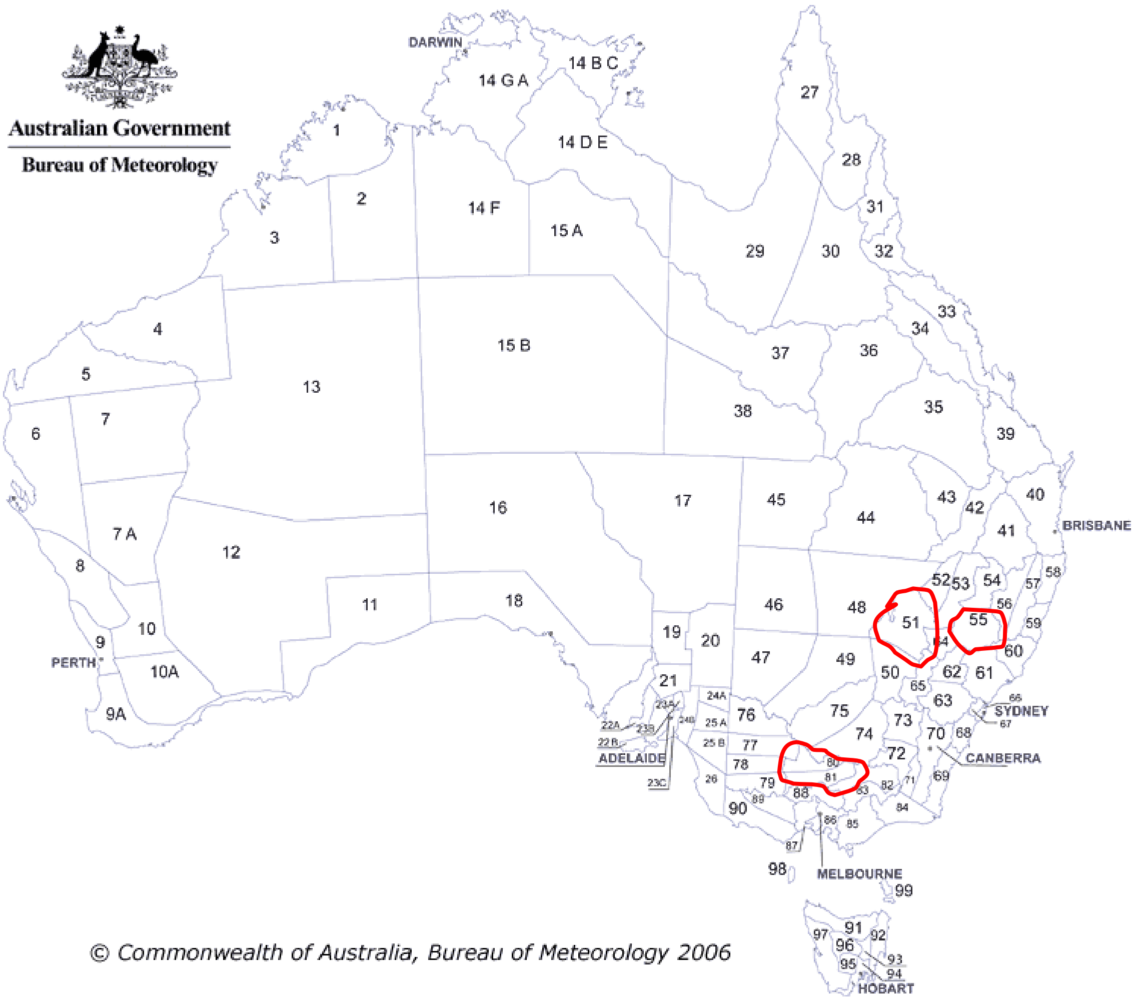


Figure 7-1 BoM rainfall districts,[16] districts 51, 55 and 80 outlined

[27] http://www.bom.gov.au/climate/cdo/about/rain-districts.shtml

While there will certainly be differences between sites in a district due to topography, the sites will be similar in terms of annual or seasonal rainfall patterns and totals. A list of the district names is provided in Appendix 8, Bureau of Meteorology rainfall districts. Figure 7-1 highlights the locations of the three rainfall districts (51, 55 and 80) which were selected for comparison purposes in the analyses presented in this chapter, due to the availability of farm data, and to represent different major seasonal rainfall zones, as listed in Table 7-1.

Table 7-1 Details of BoM rainfall districts used in the analyses

district name and number		climate class	seasonal rainfall	town
Central Western Plains (NSW)	51	uniform	uniform rainfall	Dubbo
Northwest Slopes (S) (NSW)	55	summer	wet summer and low winter rainfall	Tamworth
Lower North (VIC)	80	winter	wet winter and low summer rainfall	Echuca

An example of data sites in the BoM district containing the Dimberoy data set, BoM district 55, is shown in Figure 7-2. Locality maps for BoM district 51 and 80 are shown in Appendix 8.

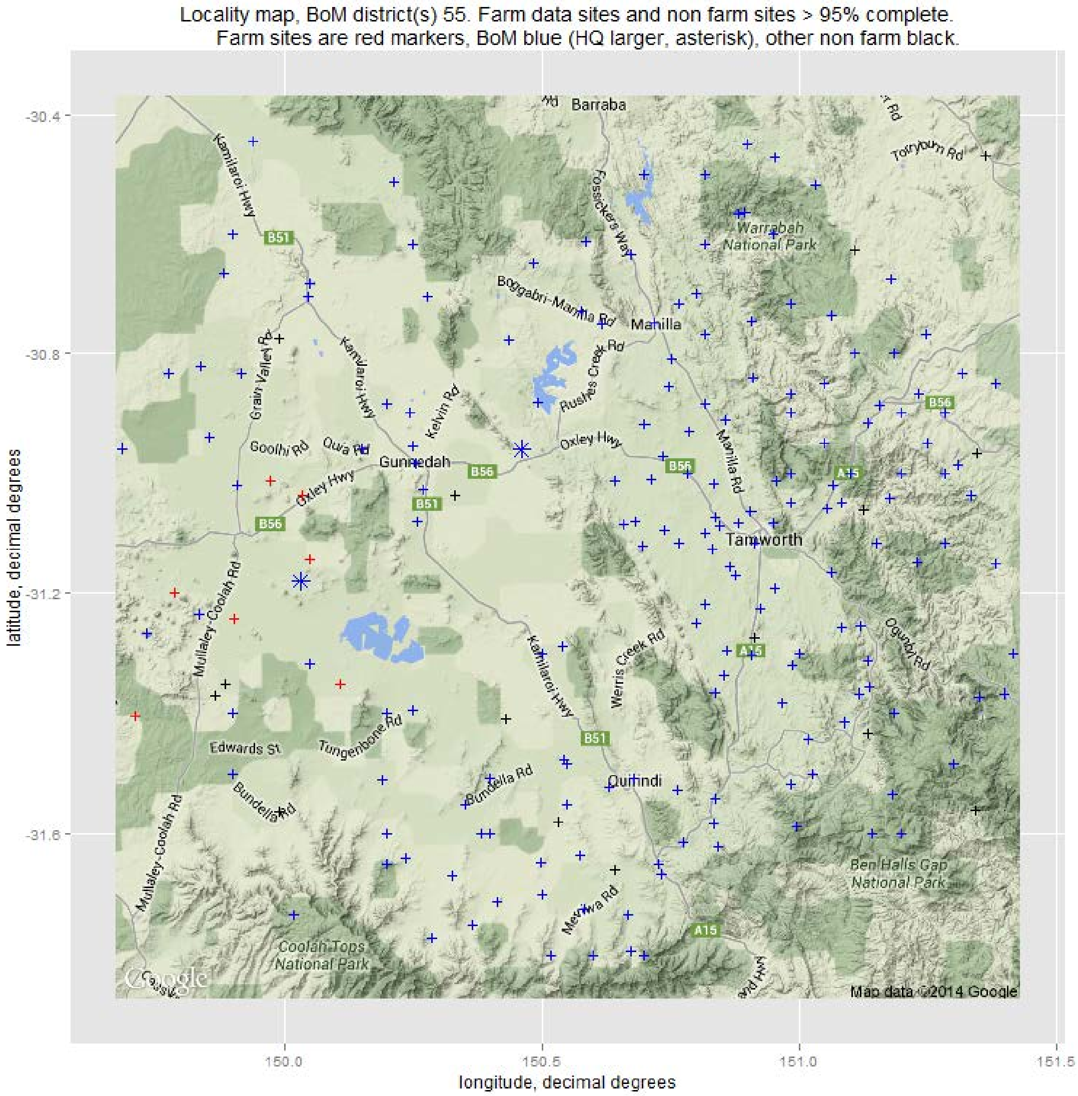


Figure 7-2 Locality map, potential data sites in BoM district 55

Figure 7-2 shows two hundred and two potential data sites in BoM district 55 (Northwest Slopes, South). A subset of the potential sites with a reasonable time span of data (80 years) reduces the number of sites significantly.

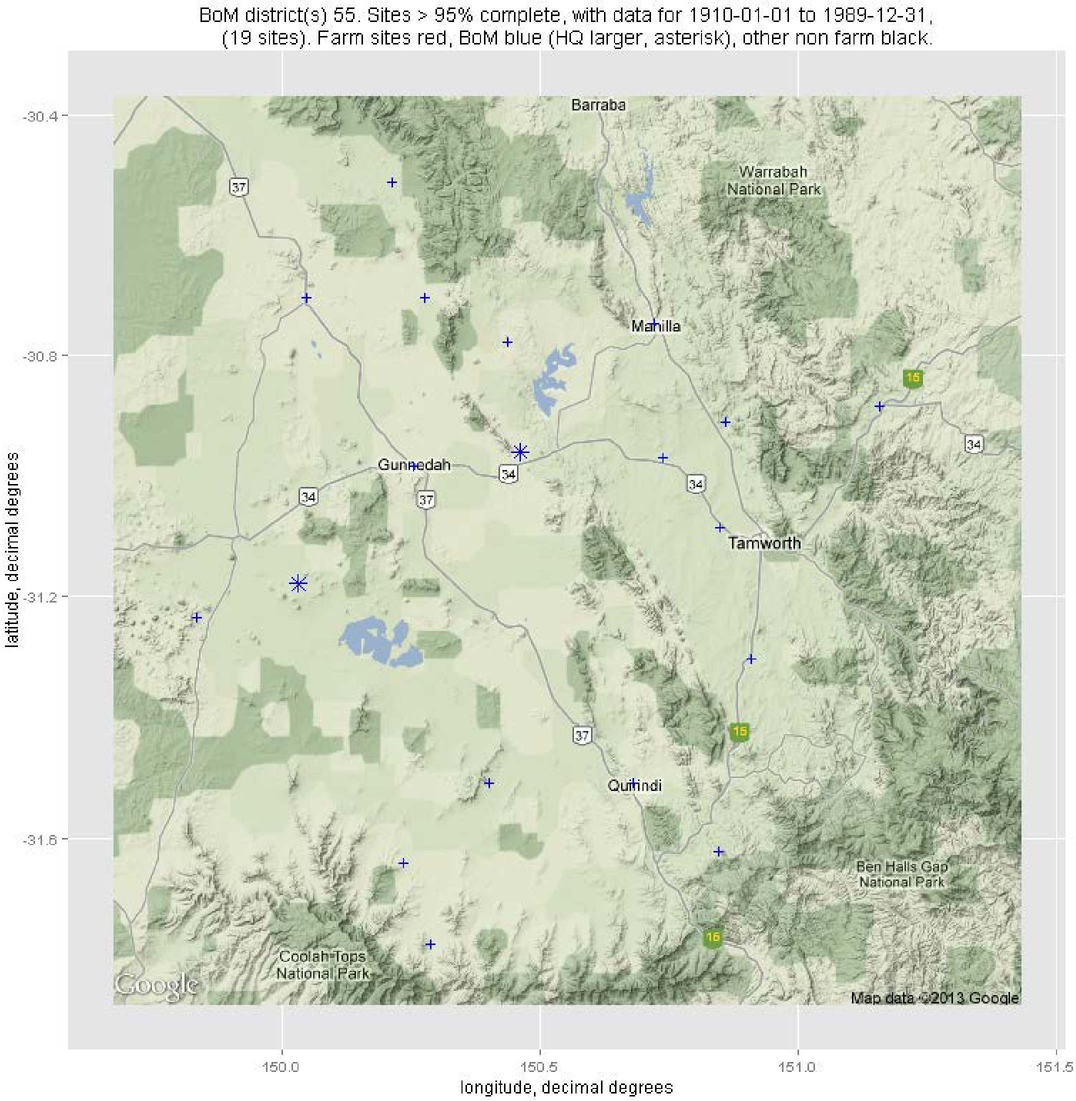


Figure 7-3 Locality map, BoM district 55 data sites for 1910 to 1989

Figure 7-3 shows the subset of sites (nineteen sites) from BoM district 55, with greater than 95% complete records for the years from 1910 to 1989. An eighty year span was chosen for analysis after inspecting the data set, as having a reasonable number of sites with the longest time span possible. The BoM seasonal climate classes (or major seasonal rainfall zones) for Australia are based on a hundred year period from 1900 to 1999[28], but the number of sites available with data completely

[28] http://www.bom.gov.au/jsp/ncc/climate_averages/climate-classifications/index.jsp?maptype=seasgrpb

covering this time range reduced the numbers of stations from 17, 19 and 12 (BoM districts 51, 55 and 80 respectively) to 13, 15 and 10.

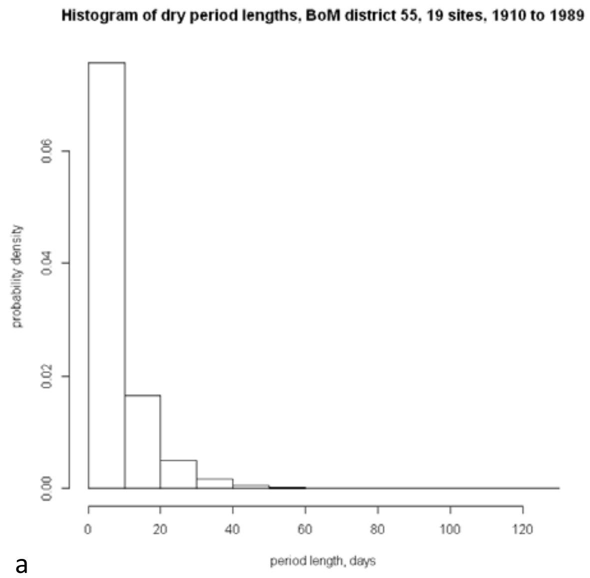
While time periods of months have been quoted as being significant in relation to yield (Cornish 1949), (Sadras, Roget et al. 2002), there is no justification provided (to the author's knowledge) as to why a calendar month would be more significant than any other thirty day period, except perhaps in relation to lunar cycles.

Therefore, instead of testing for a relationship between rainfall in calendar months and seasonal or annual rainfalls, it is preferable to investigate the relevance of a thirty day period, and the possible relationships between the rainfall in any thirty day period during a year and seasonal or annual rainfalls.

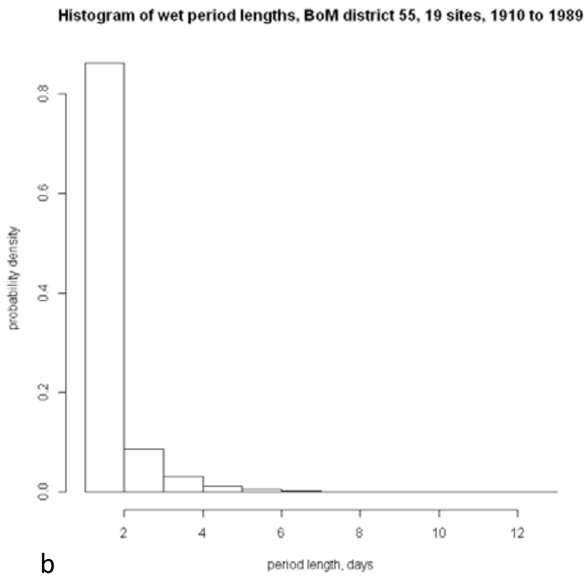
7.3 Length of period

If a time period is relevant in terms of cycles or patterns of rainfall events, the length of the period should cover at least one full cycle of conditions. In this case the possible conditions are rainfall events and dry times between rainfall events.

Figure 7-4 shows probability histograms for wet and dry period lengths for BoM district 55. Table 7-1 and Table 7-2 present the values for the histograms shown, and corresponding values for rainfall districts 51 and 80. The densities in the histogram figures sum to 0.1 (Figure 7-4a, due to the category widths being 10 not 1) and 1 (Figure 7-4b), but are presented as percentage values for simplicity in Table 7-1 and Table 7-2. Figure 7-4 shows that for BoM district 55 the majority of dry periods are ten or less days, and that the majority of wet periods are one or two days. The data tables indicate that the uniform rainfall district (district 51) has more dry periods exceeding thirty days than the others, and also has longer extremes. The winter maximum rainfall district (80) despite having the lowest annual average rainfall (for the seventeen sites and the period 1910 to 1989), appears to have more prolonged rainfall events. Quantitatively, the percentages of wet periods with duration three days or less are 95.8, 94.8 and 94.6 for the three BoM rainfall districts 51, 55 and 80 respectively. The percentages of dry periods with duration thirty days or less are 94.27, 97.16 and 97.08 for the three BoM rainfall districts respectively. Given the minor variations mentioned, it seems clear that a period of thirty days will encompass most wet and dry cycles, and therefore is an appropriate period for investigation.



a



b

Figure 7-4 Histograms of dry (a) and wet (b) period lengths, BoM district 55, 1910 to 1989

Table 7-2 Histogram dry period densities (as %) and counts

BoM district (sites)	51 (17)		55 (19)		80 (12)	
Ave ann P, 1910-89	464.7 mm		628.9 mm		402.3	
range	density	count	density	count	density	count
0-10	67.47	22674	75.58	42239	78.76	30087
10-20	19.07	7848	16.55	9249	14.08	5377
20-30	7.73	3179	5.03	2810	4.24	1618
30-40	3.14	1294	1.82	1018	1.71	652
40-50	1.48	611	0.66	370	0.69	262
50-60	0.54	224	0.216	115	0.27	103
60-70	0.26	108	0.11	61	0.13	51
70-80	0.13	52	0.02	13	0.07	26
80-90	0.09	36	0.01	9	0.03	10
90-100	0.05	22	< 0.01	1	0.02	8
100-110	0.01	6	< 0.01	1	0.01	4
110-120	< 0.01	4	0	0	< 0.01	1
120-130	< 0.01	2	< 0.01	1	<0.01	1
130-140	0	0				
140-150	< 0.01	1				

Table 7-3 Histogram wet period densities (as %) and counts

BoM district (sites)	51 (17)		55 (19)		80 (12)	
Ave ann P, 1910-89	464.7 mm		628.9 mm		402.3	
range	density	count	density	count	density	count
1-2	88.23	36114	86.04	47833	86.31	32930
2-3	7.60	3112	8.72	4850	8.26	3152
3-4	2.45	1003	3.17	1763	3.33	1269
4-5	1.09	445	1.22	678	1.27	485
5-6	0.47	192	0.59	329	0.49	185
6-7	0.07	29	0.15	81	0.24	91
7-8	<0.01	26	0.08	43	0.06	23
8-9	<0.01	8	0.02	11	0.02	9
9-10	<0.01	3	0.01	5	0.02	6
10-11			<0.01	1	0.01	2
11-12			<0.01	2	0	0
12-13			<0.01	1	0	0
13-14					<0.01	1
14-15					<0.01	1

7.4 Rolling thirty day periods

While a thirty day period is an appropriate period covering the lengths of the majority of wet and dry periods or cycles, there is no intuitive support for the thirty day periods being best represented by calendar months. Current calendar months are approximately equally spaced intervals subdividing the calendar year, originally based on the lunar month, with the lunar month being twenty nine and a half days long. The system in use 'in the West' using thirty and thirty-one day months (excluding February) dates from eight BC (BCE), and is atributed to Augustus, first Emperor of Rome, (OED 2013).

There has been limited work investigating the influence of lunar cycles on rainfall (Bradley, Brier et al. 1962), (Adderley and Bowen 1962), (Roy 2006). Roy concluded that rainfall maxima ocurred 'a few days after the full moon', in India, with the lunar influence on meteoric dust or upper atmosphere conditions being seen as possible causes. This work is very inconclusive, and has apparently not been continued, and no more recent studies have been found (by this author).

To test for a relationship between any thirty day period and the seasonal or annual rainfall, a simple linear model (see section 7.6) using the period rainfall as a predictor and the seasonal or annual rainfall as the reponse was tested. All possible thirty day periods starting in a given year were to be tested, including those after starting day number 335, which extended into the following year. Note that the word ‘rolling’ does not imply carryover effects or smoothing, but the incremental progression of the starting point to cover all possible starting points.

If calendar months or lunar cycles are of significance, and all possible thirty day periods are tested, these cycles will be represented by one or more of the thirty day periods during the year.

7.4.1 Data and method

1. As for the previous analysis, the data consists of daily rainfalls for each of the sites :
   - BoM district 51, seventeen sites, 1910 to 1989
   - BoM district 55, nineteen sites, 1910 to 1989
   - BoM district 80, twelve sites, 1910 to 1989

2. The total periodic rainfall for all possible thirty day periods during a year (three hundred and sixty five periods) was extracted

3. The periodic rainfall for all years (80), at one site, was used as predictor variable (x) in a linear model with annual rainfall or seasonal rainfall for each year at the same site as response variable (y)

4. The model R squared values were extracted for comparison

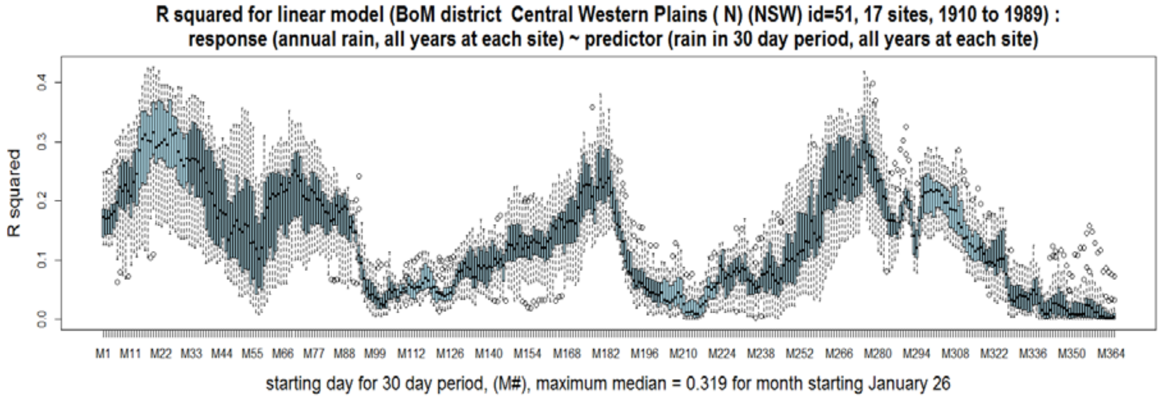
The above process was followed for all years and sites.

The process was repeated for subsets of wet (above median rainfall) and dry (below median rainfall) years.

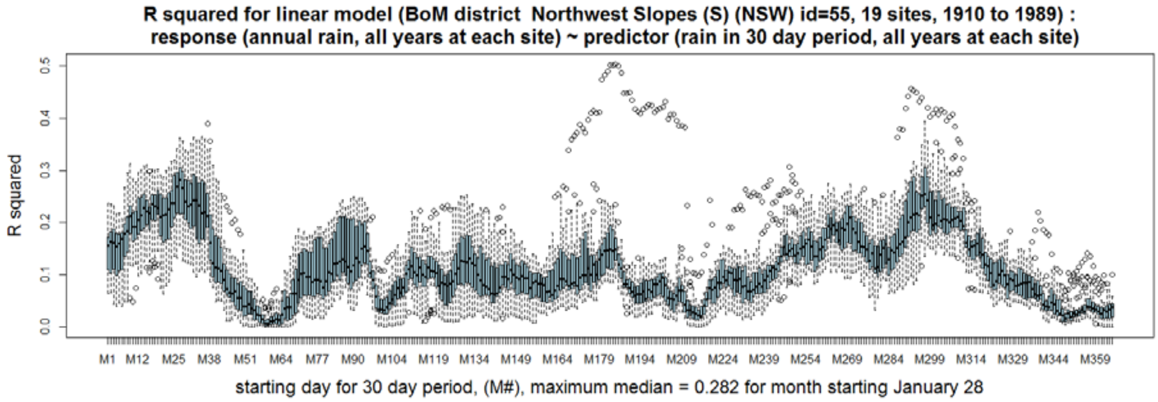
The ‘season’ for use in the seasonal comparison was from April to October. This period was chosen due to its previous definition as being a relevant growth season for wheat crops (Trumble 1937), (Hounam 1947; Hounam 1950).

7.5 Results and discussion

a



b



c

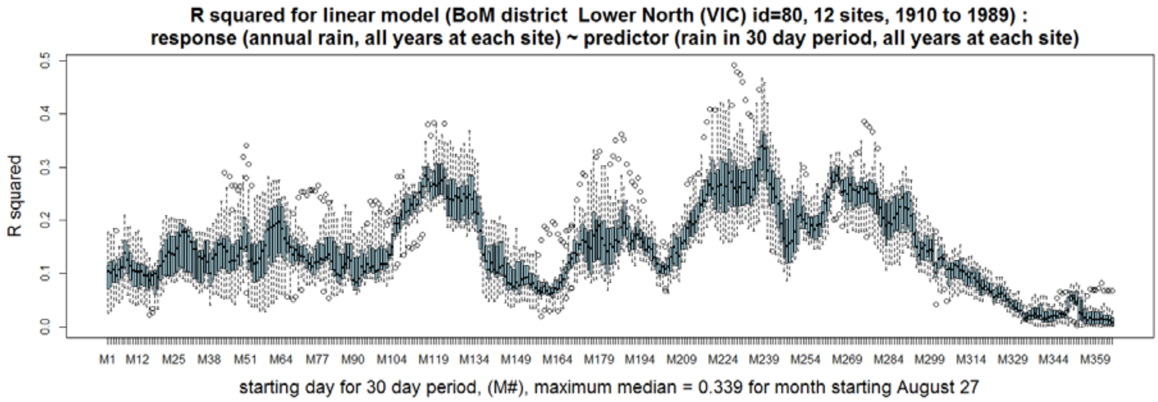


Figure 7-5 Linear model R squared boxplots for BoM districts 51(a), 55(b) and 80(c), all years, annual rain, data from 1910 to 1989

Figure 7-5 shows the results of the analysis described above for each of the BoM districts (51, 55 and 80) for all the years (1910 to 1989), using annual rainfall totals. The horizontal (x) axis is labelled as the starting day for the thirty day periods, These start days will coincide with calendar months as listed below.

day number :	1	32	60(61)	91	121	152	182	231	244	274	305	335
start of month :	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec

Months after February will start one day later in leap years, as shown by the bracketed value for March, but this is not significant in terms of the approximate 'positions' of calendar months along the axis. Figure 7-6 shows similar results for the seasonal analysis. Also noted in the captions of Figure 7-5 and Figure 7-6 are the maximum median R squared values and the period in which the maximum ocurred for each analysis.

a

b

c

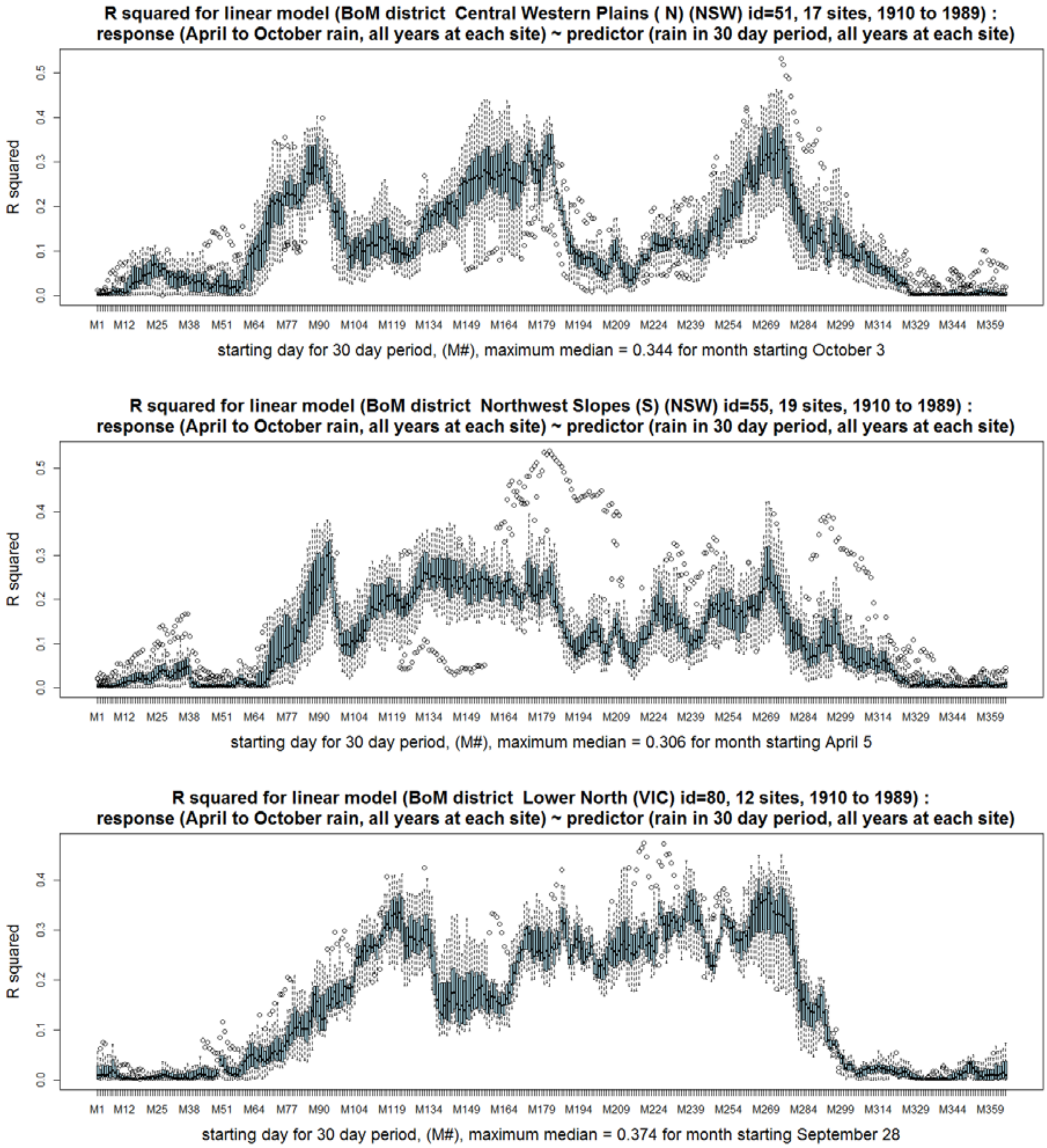


Figure 7-6 Linear model R squared boxplots for BoM districts 51(a), 55(b) and 80(c), all years, seasonal rain, data from 1910 to 1989

Figure 7-7 shows a summary plot for the figures above, comparing median R squared data from the plots above, for annual and seasonal rainfall.

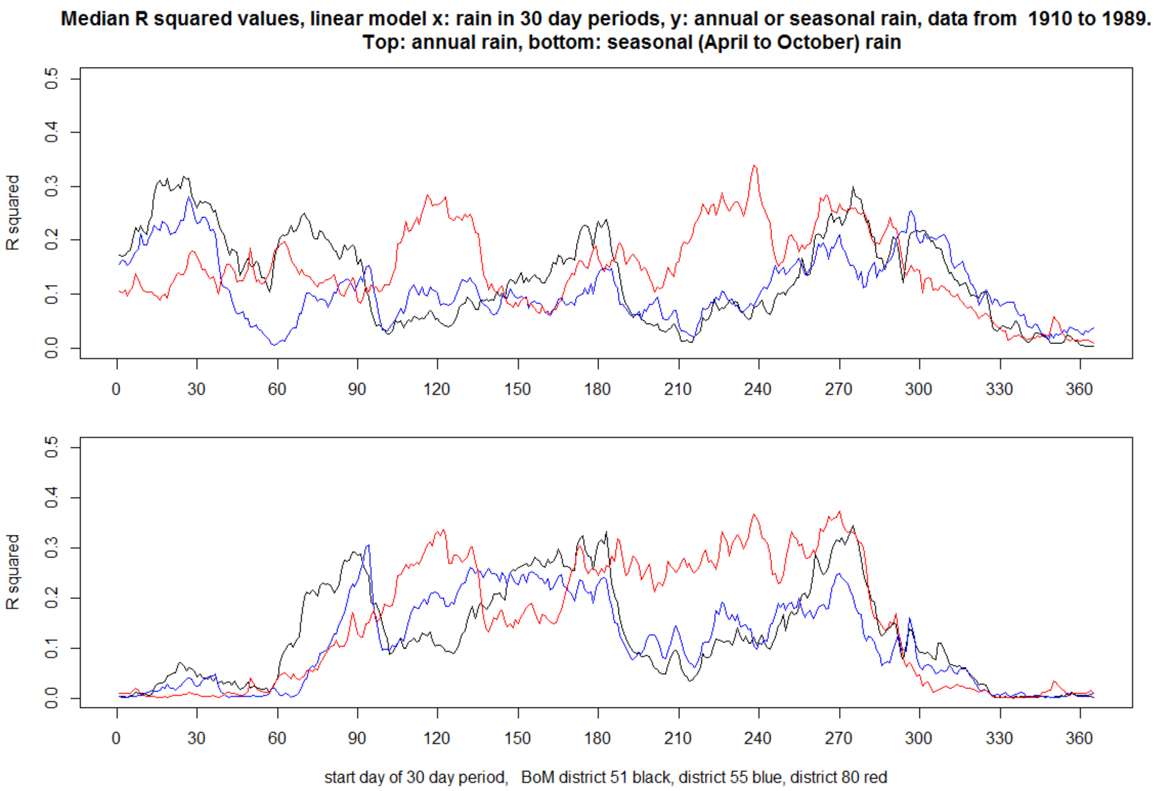


Figure 7-7 R squared summary plot, 30 day periods and annual and seasonal rain BoM district 51 (black line), BoM district 55 (blue line) and BoM district 80 (red line), data from 1910 to 1989

To check the significance of the R squared vales, t values can be calculated based on sample sizes of 17, 19 and 12 sites for BoM districts 50, 51 and 80 respectively. The t values show that critical R squared values for significance are 0.19, 0.17 and 0.26 for BoM districts 50, 51 and 80 respectively. The summary plots above show that the peak R squared values are significant (Webster and Oliver 2007).

The figures above do not show any particular 30 day period with a single, outstanding high correlation for all seasons at all sites. However, it appears that there are quite a good correlations between the periods starting at about 90, 180 and 270 days (April, July and October), and both the annual and seasonal rainfall for districts 50 and 51, the uniform and summer maximum rainfall districts, with the seasonal relationship being slightly stronger. The relationship between the 90 day period and seasonal rainfall (R squared 0.292 and 0.306 for districts 51 and 55 respectively) is of particular interest, seeming to support the significance of April rainfall noted by others (Sadras, Roget et al. 2002). The value of such a relationship lies partly in the fact that April is near the start of the growing season, and so may be a useful indicator of the season ahead, unlike the 180 and 270 day values which may be similarly related. While district 80 (winter maximum rainfall) does not display the same relationship, it does show a good correlation between the period start at about 120 days (May) and both annual and seasonal rainfall. As with the other districts, the seasonal relationship is slightly stronger, with an R squared of 0.335. While other periods are in fact slightly better correlated (R squared 0.368 and 0.374 for periods atarting at 238 and 269 days respectively), the May relationship, also being near the start of the growing season, is the one of most interest. It is also noteworthy that the relaitonships are not as good for periods adjacent to the 90 and 120 day periods, indicating that these periods are quite distinct periods with better correlations, not just part of a constant correlation pattern. Figure 7-8 and Figure 7-9 show summary plots for the analyses comparing median R squared data, for annual and seasonal rainfall, for wet (greater than medain rainfall) and dry (less than median rainfall) years for the three BoM districts.

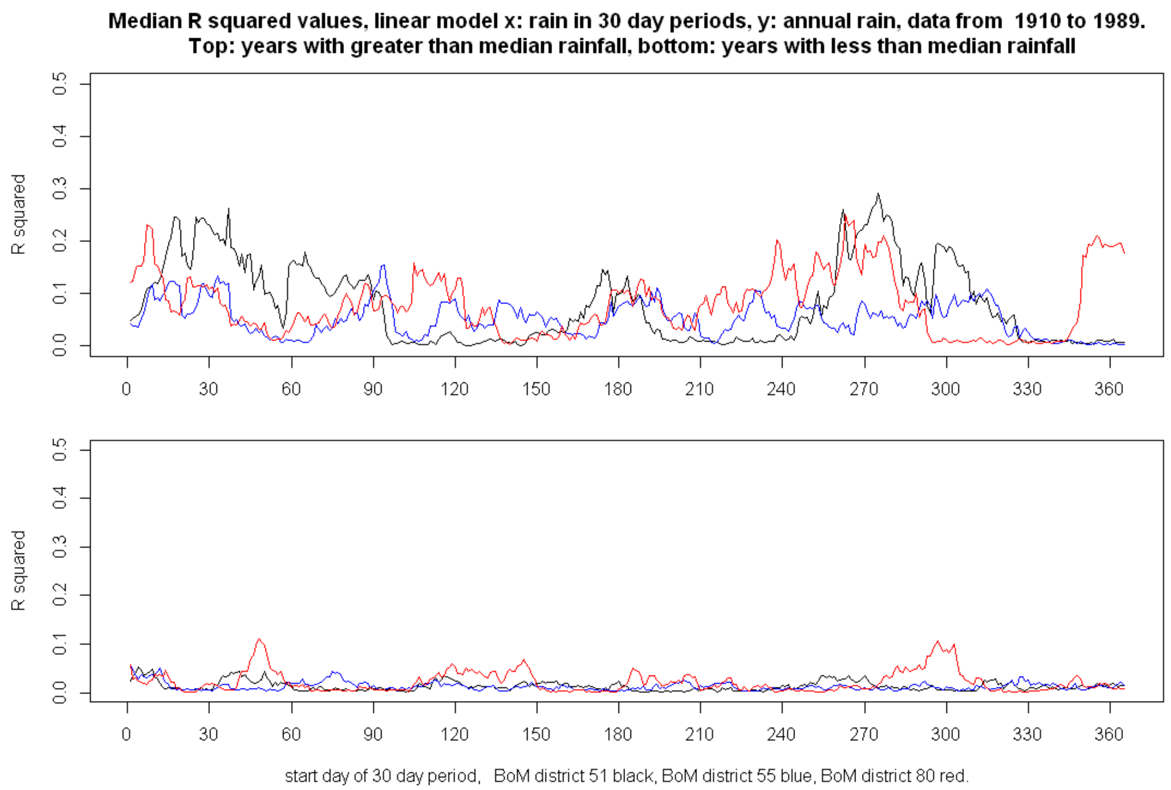


Figure 7-8 R squared summary plot, 30 day periods and annual rain, wet (top) and dry (bottom) years BoM district 51 (black line), BoM district 55 (blue line) and BoM district 80 (red line), data from 1910 to 1989

Figure 7-8, the analysis for annual rainfalls for wet and dry years, show consistently lower correlations than the analyses using all the annual data. It is probably not surprising that the dry years show the poorest correlations. Qualtiatively, it could be expected that many years with consistently low rainfall values, and many zero rainfall periods could make the analysis difficult. The analyses for the wet years show some reasonable correlations at about 0 to 30 days and 270 days (period start) for BoM districts 51 and 80. District 55 shows fairly poor correlations throughout, but does show a peak at 90 days. While the correlations around 270 days (October) are of interest, they do not represent a useful indicator, being late in the season and not uniform across the districts.

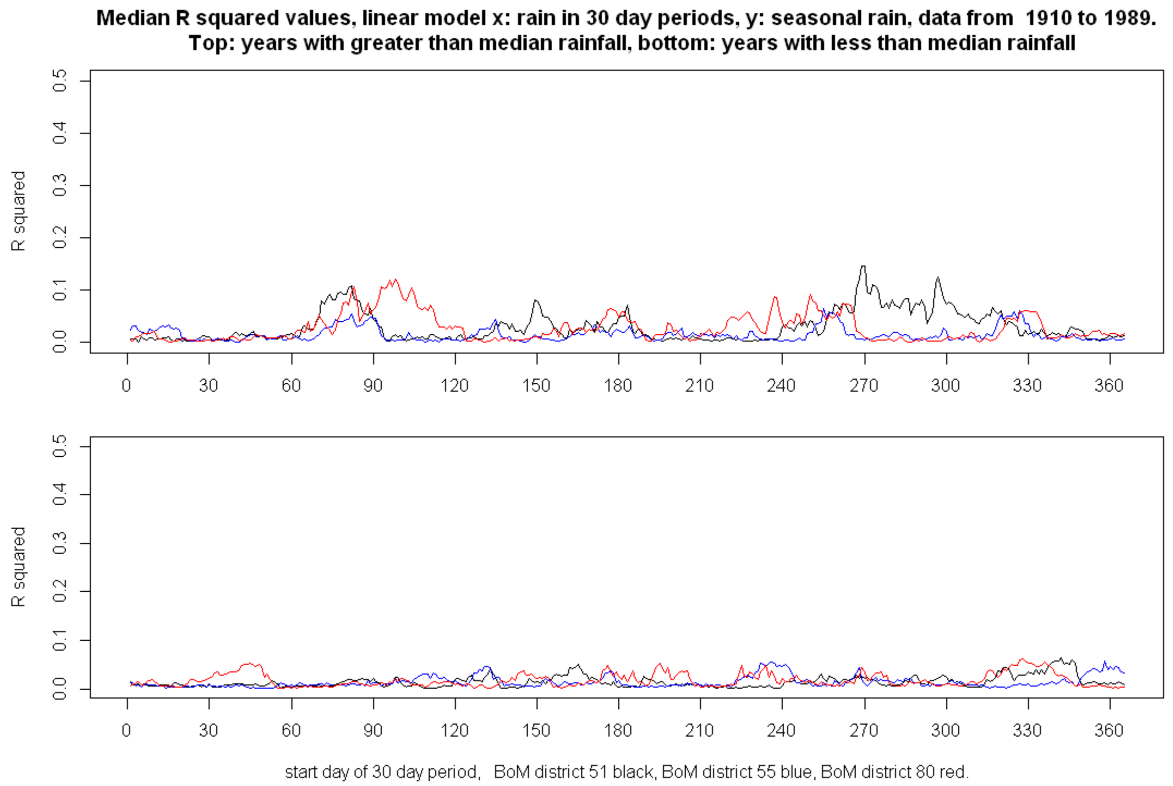


Figure 7-9 R squared summary plot, 30 day periods and seasonal rain, wet (top) and dry (bottom) years BoM district 51 (black line), BoM district 55 (blue line) and BoM district 80 (red line), data from 1910 to 1989

Figure 7-9, the analysis for seasonal rainfalls for wet and dry years, show consistently lower correlations than the analyses using all the annual data. It is worth noting, however, that the pattern of higher correlations in the early part of the year for wet years is similar to the pattern for seasonal rainfall using all the years. That is, the better correlations for BoM district 51 and 55 occur about 30 days before district 80, with the better correlations for the seasonal rain wet data analysis occuring about 15 days before the analysis using all the annual rainfall data.

7.6 Linearity of rolling month relationships

While the linear model used in the analyses above is a convenient, if simple, method of testing the relationship between the 30 day period rainfalls and the annual or seasonal rainfall, it is necessary to check that the assumption if linearity is valid or at least reasonable.

To check the linearity assumption, the data for rainfalls for specific 30 day periods can be compared against annual rainfalls for a particular site. BoM district 55 has two high quality sites, stations 055055 (Carroll, The Ranch) and 055045 (Curlewis, Pine Cliff), which can be used for this purpose. The sites are shown on the locality map (Figure 7-3). Figure 6-1 identifies site 055045 as the site to the south west of Gunnedah, and site 055055 is the site north east of Gunnedah on Figure 7-3.

A simple plot of annual rainfall at the BoM sites and selected 30 day periods was used for comparison. A similar analysis to those describe above was performed to select data for specific 30 day periods, that is :

1. The daily data was grouped into all possible thirty day periods during a year (threehundred and sixty five goups)

2. The total periodic rainfall for all possible thirty day periods during a year (three hundred and sixty five periods) was extracted

3. The periodic rainfall for all years, at one site, was used as predictor variable (x) in a linear model with annual rainfall for each year at the BoM site as response variable (y)

4. The model R squared values were extracted for comparison, with ten periods being used in the plots shown below, consisting of :
   a. two 30 day periods with high R squared values from the analysis for site 055055 (M27, M297)
   b. two 30 day periods with the low R squared values from the analysis for site 055055 (M59, M215)
   c. four randomly selected 30 day periods

Figure 7-11 shows the resulting plot of annual rainfall versus periodic rainfall for selected sites for BoM site 055045.

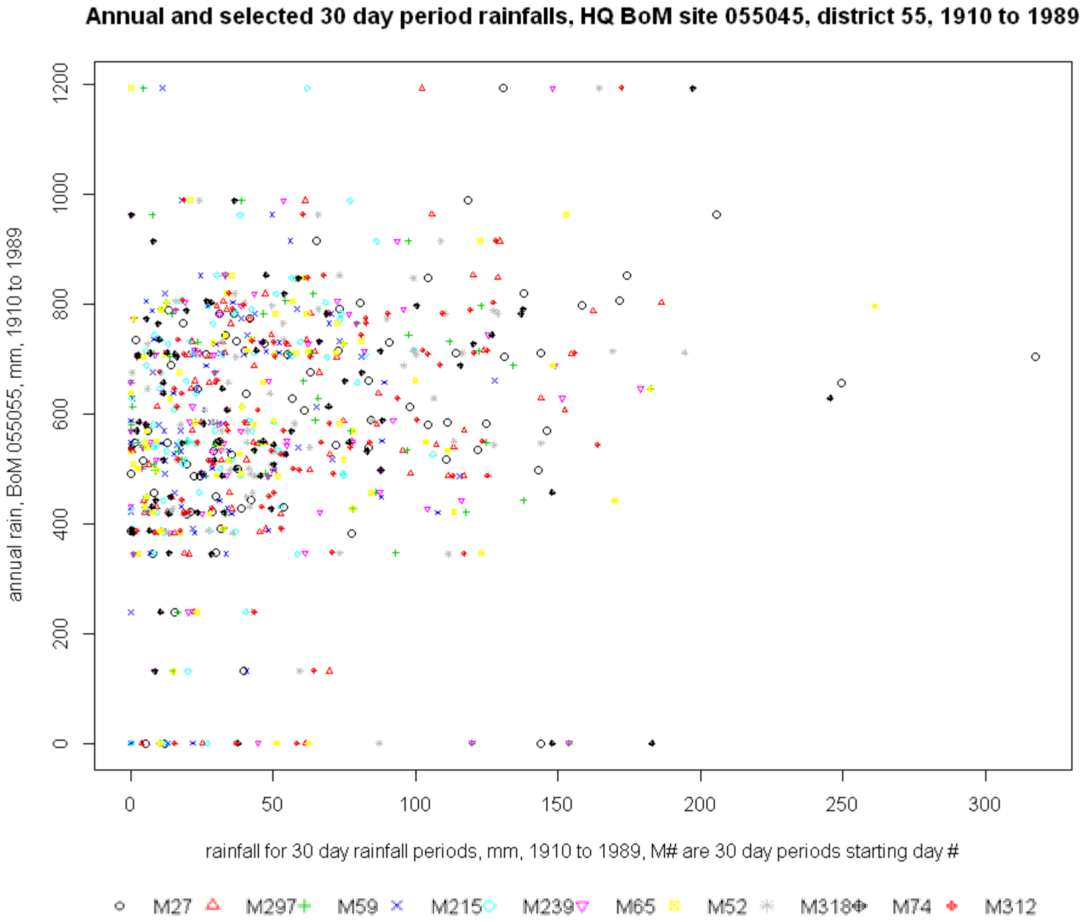


Figure 7-10 Annual rainfalls and selected 30 day period rainfalls for BoM site 055045

Figure 7-10 does not show any clustering or divisions in the plot pattern. The lack of such patterns, while not necessarily proving that a linear model is the best choice, shows that there is no reason to reject the linear model as a descriptive tool. The selected 30 day periods used are listed in Table 7-3, with some statistical values from the linear model fit of annual rainfall (response) and periodic rainfall (predictor). The periods in the table are referred to by the descriptors previously used, for example M27 to represent the thirty day period stating on the 27th day of a 366 day year. The statistics in the table show that, according to the p values, the linear model fit is significant (<0.05, Johnson (1999) for half of the selected thirty day periods (highlighted). It is worth noting that the periods with higher R squared values (M27, M297, M182 and M238) would also be significant at a higher level than the other periods, with these higher R squared value periods shown to be periods of relevance in the previous tests. While the lowest mean periodic rainfall (M102) also shows a low (not significant) R squared, and the highest mean rainfalls show the highest R squared values, there is a reasonable spread of rainfall means across the R squared range.

Table 7-4 Selected statistics for linear model fit, Annual rainfalls and selected 30 day period rainfalls (predictor) for BoM site 055045

	M27	M297	M59	M215	M174	M102	M98	M182	M108	M238
multiple R sq	0.139	0.128	0.004	0.022	0.063	0.002	0.014	0.091	0.019	0.115
p value	0.001	0.001	0.598	0.187	0.026	0.702	0.302	0.007	0.217	0.002
annual average P						590.69				
Selected month (M) values :										
Median	41.1	46.1	31.4	28.3	34.6	23.2	24.5	31.5	22.8	22.7
Mean	66.5	58.0	44.9	34.8	41.1	31.7	32.6	39.5	34.2	34.0
Max	104.2	79.8	64.9	49.7	50.9	43.5	48.5	56.8	49.5	50.0
number of NA's					2					1

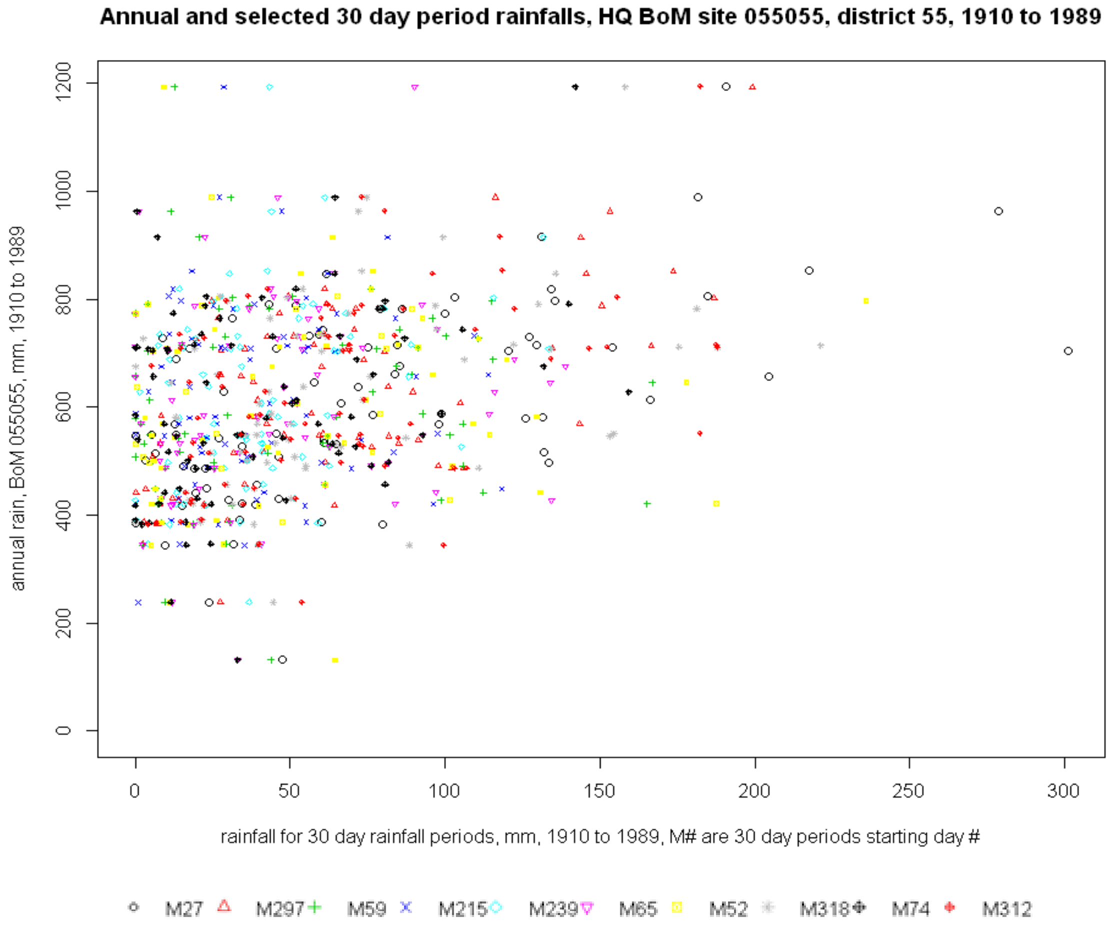


Figure 7-11 Annual rainfalls and selected 30 day period rainfalls for BoM site 055055

Figure 7-11 shows the resulting plot of annual rainfall versus periodic rainfall for selected sites for BoM site 055055, and Table 7-4 shows corresponding model statistics, with significant values once again highlighted.

Table 7-5 Selected statistics for linear model fit, Annual rainfalls and selected 30 day period rainfalls (predictor) for BoM site 055055

	M27	M297	M59	M215	M174	M102	M98	M182	M108	M238
multiple R sq	0.323	0.367	0.005	0.011	0.121	0.021	0.061	0.132	0.065	0.086
p value	0.000	0.000	0.533	0.360	0.002	0.211	0.032	0.001	0.027	0.010
annual average P					590.69					
Selected month (M) values :										
Median	75.2	62.8	45.5	39.4	44.8	31.2	32.6	44.8	36.1	40.6
Mean	102.9	81.7	76.2	55.2	52.2	43.1	47.4	53.7	52.3	50.1
Max	3	4	4	4	4	4	4	4	4	4
number of NA's	0.323	0.367	0.005	0.011	0.121	0.021	0.061	0.132	0.065	0.086

The statistics in the table show that, according to the p values, the linear model fit is significant (<0.05, Johnson (1999) for seven out of ten of the selected thirty day periods (highlighted). Once again, the periods with higher R squared values (M27, M297, M182 and M238) would also be significant at a higher level than the other periods. While both sites are high quality BoM stations, site 055055 has a higher number of significant R squared values than site 055045. This may be due to the more central location of 055055, making 055055 more typical of the district as a whole.

7.7 Conclusions

While it is not suggested that the linear model is the best possible option for describing the data, the results presented show that there is no reason to reject the model as inappropriate.

These thirty day indicator time period results are very encouraging, especially in light of the small data sets available for analysis. It appears that rainfall in the early part of the growing season (April and May) may indeed be indicative of seasonal condtions. I believe that a more extensive analysis, with more data across more BoM districts (and therefore climate classes) will be required to confirm or disprove this conclusion. While these results are not definitive, it apperas that farm rainfall records alone may be able to provide an indicator time period which will provide another tool for farmers to help interpret and utilise seasonal forecasts. It appears at this stage that the relationships for wet (above median) rainfall years and seasonal rainfalls are quite strong, while relationships for dry years are not. This may be due to data scaling or handling of zero data value points, issues which provide avunues for future study.

There is good potential for future study based on these results. This type of analysis could be performed and refined for all the BoM climate zones, and also by expanding the potential indicator periods to include periods which may be of singificance such as previous summer or wet seasons. With the availablility of more data, similar analyses could include farm data, with associated benfits of increased data density. The potentail benefit to famers could lie in the relationship of regional or seasonal specific indicator periods based on local information, which relate to approaching seasons with or without other forecasts. This information could be important in regions where existing data is scarce or spatial variability is high.

8 External climatic influences or indicators

8.1 Aim

The focus of this chapter is the fourth and final aim of the project :

- To test for effects on such an indicator of seasonal or annual rainfalls, of external influences and climate drivers.

The motivation of the project is to find new ways for farmers to make use of the rainfall data they are collecting. Therefore, while this chapter involves external climatic drivers, any analyses and parameters need to be easily accessible to users, relevant to their needs and within their capabilities to apply and understand. Ideally this translates to a single indicator which users know or can relate to, and is applicable to their own rainfall data to help forecast events or add value to existing forecasts.

As mentioned in the literature review, Jones, Hansen et al. (2000) noted that many workers had shown 'high correlation between ENSO activity and agricultural production in many parts of the world'. More recent work Kamruzzaman, Beecham et al. (2013) reconfirms the value of the SOI in relation to rainfall and runoff in the Murray Darling Basin in south-eastern Australia. This suggests that a parameter based on the ENSO or SOI may provide a useful tool for indicating seasonal or annual rainfalls, simplifying the multiple parameters required by general circulation models or agricultural systems models, or the many other indices relating to sea surface temperatures and or pressures. While the SOI may not be a parameter that users would necessarily feel comfortable defining or calculating, it is one they are probably familiar with, and is readily available, for example on the BoM internet sites. Stone and Auliciems (1992) pointed out that simple correlations between wet or dry ENSO phases may be missing the effect of changes in the life cycle of the SOI, and that changes in the SOI may relate to rainfall trends before (or after) the SOI can be categorized as part of the three phase SOI system. Previous work in this project has indicated that April or May rainfall may be indicative of seasonal conditions (section 7.5), but more work is needed to clarify this and test different locations and climatic types. The five phase SOI system (SOIp) of Stone and Auliciems (1992) will therefore be tested as an indicator which can improve or replace the relationship between April and May rainfalls and annual or seasonal rainfalls. The SOIp is a single parameter which should appeal to users, it is readily available on the Queensland Government internet site[29], and can be calculated from SOI data if required. As already noted in the literature review, Stone identified five SOI phases, depending on SOl values in the current and immediately preceding month, categorized as follows (also showing the relationship to rainfall in eastern Australia):

1.	consistently negative	rainfalls below long term median
2.	consistently positive	rainfalls above long term median
3.	rapid fall (or just 'falling')[a]	rainfalls below long term median
4.	rapid rise (or just 'rising')[a]	rainfalls above long term median
5.	consistently near zero	neutral

[a] (Stone and Auliciems 1992)

8.2 Methods

8.2.1 SOI phase testing

Before testing the SOI phases in relation to seasonal or annual farm data, it is necessary to test the assertion that the phases relate to rainfalls as described by Stone and Auliciems (1992), listed above. A simple way to test the relationship is to plot the monthly rainfalls categorized by SOI phases (SOIp), for a number of BoM rainfall districts, thereby covering a variety of climate classes and seasonal rainfall patterns. A location map of the districts selected for this process was previously presented in section 7.2, and will not be duplicated here. The same BoM districts used in the tests for potential indicator periods (chapter 7) will be used for this analysis, and so these district details are repeated below. An extra BoM district will be included, district 41. While this district is in the same major classification as district 55, it may have a higher annual rainfall[30], and, being summer dominant may have a higher level of seasonality which could influence the relationship.

Table 8-1 Details of BoM rainfall districts used in the analyses

district name and number		climate class	seasonal rainfall	town
East Darling Downs	41	summer	wet summer and low winter rainfall	Dalby
Central Western Plains (NSW)	51	uniform	uniform rainfall	Dubbo
Northwest Slopes (S) (NSW)	55	summer	wet summer and low winter rainfall	Tamworth
Lower North (VIC)	80	winter	wet winter and low summer rainfall	Echuca

[29] http://www.longpaddock.qld.gov.au/seasonalclimateoutlook/rainfallprobability/phases.php
[30] http://www.bom.gov.au/jsp/ncc/climate_averages/climate-classifications/index.jsp?maptype=seasb

The data period chosen for analysis is the one hundred year period from 1907 to 2007, with only sites having greater than ninety-five percent complete records being included in the analysis. These specifications were chosen to provide the maximum number of sites across the various locations covering a reasonable time period. While the districts chosen were also ones which included farm data, the specifications excluded all the farm sites.

8.2.2 SOI phases in relation to farm data

To test the possible value of the SOIp to users' data and situation, the SOIp will be treated as a possible filtering mechanism in the use of monthly data as a seasonal or annual rainfall indicator. Work so far has shown that April or May rainfall may be indicative of seasonal conditions, and useful due to the early season timing. This result will be investigating by testing the relationship between April and May farm rainfalls and seasonal or annual rainfalls for 'wet' years, and then retesting with a subset of the monthly data defined by the SOIp. Remembering that this analysis is aimed at practical, real time use by farmers, and should therefore be included in a 'next generation' feedback report, the analysis should be kept as simple as possible, and so the 'test' will simply be the reported R squared value from a linear model fit of monthly or subset data to seasonal and annual farm rainfalls. The analysis will be performed on wet years (above median seasonal or annual total rainfalls), as it is believed the ability to predict wet seasons will be of most value to farmers. Clearly future work should include all options. The procedure can be outlined as follows :

1. Select farm data, all years available, subset for wet (above median) years
   (earlier work showed the relationships with dry periods were not strong)
2. Run linear model, monthly rain as predictor for seasonal or annual rainfall, keep R squared
3. Subset, years with SOIp being 2 or 4 (the phases associated with above median rainfalls)
4. Rerun linear model, subset monthly rain as predictor, keep R squared
   The 'season' for use in the seasonal comparison was from April to October. This period was chosen due to its previous definitions as being a relevant growth season for wheat crops (Trumble 1937), (Hounam 1947; Hounam 1950).
5. Repeat for all districts
6. Compare R squared values

The farms from the respective BoM districts available for analysis are :

district name and number		climate class	farm	data interval
East Darling Downs	41	summer	Springfields	1975 - 2009
Central Western Plains (NSW)	51	uniform	Ashlee	1954 - 2009
Northwest Slopes (S) (NSW)	55	summer	Dimberoy	1947 - 2006
Lower North (VIC)	80	winter	Pine Grove	1965 - 1999

8.3 Results and discussion

8.3.1 SOI phase testing

A plot of the relationship between the SOI phases and monthly rainfall for individual stations in BoM district 55 is shown in Figure 8-1. Of the BoM stations listed station 055055 is a high quality station. Figure 8-1 shows the data extremes.

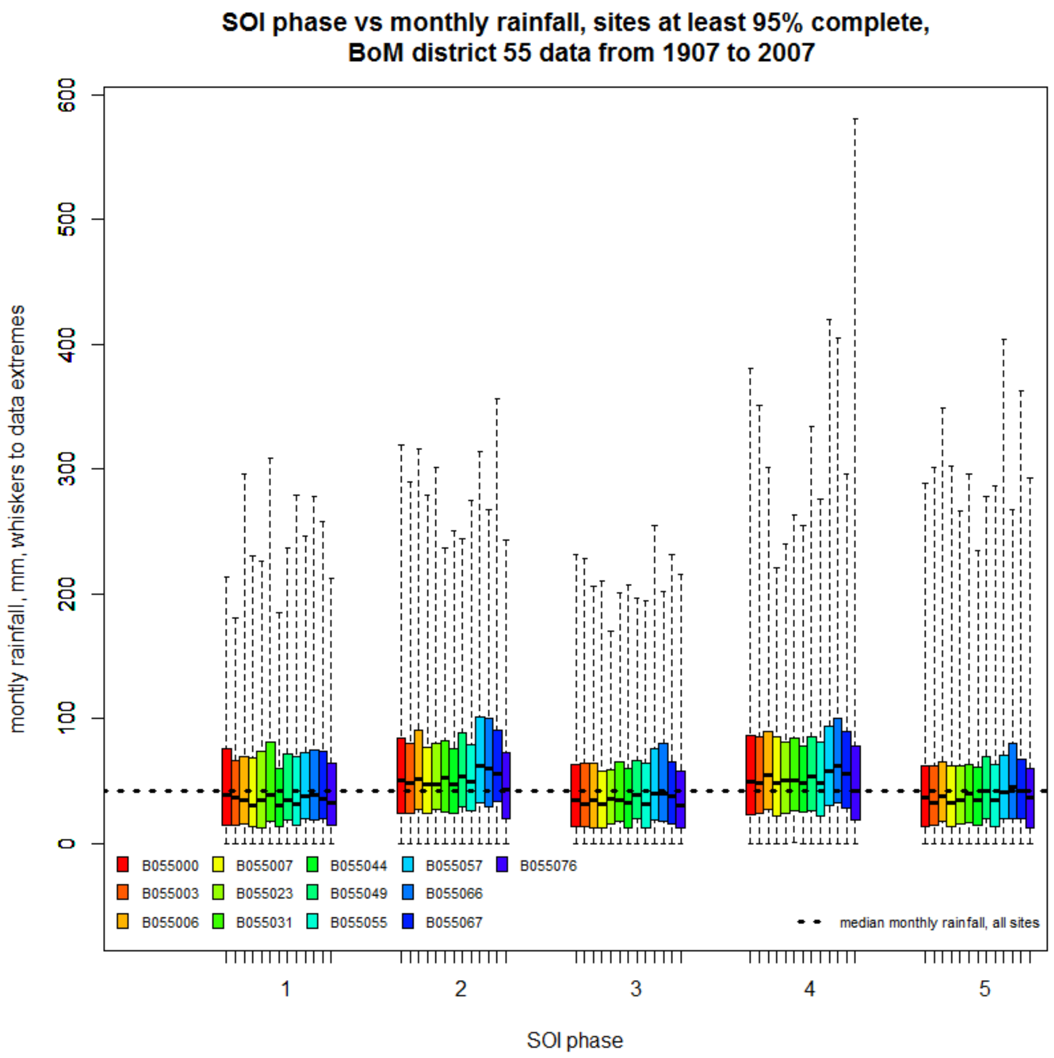


Figure 8-1 SOI phase and monthly rainfalls for all sites, BoM district 55

Similar plots for all the districts will not be shown here, but a summary plot for all districts is shown in Figure 8-2. Both figures show the SOIp data confirming the relationship with monthly rainfalls suggested by Stone et al, across four of the possible six climate classes (one being arid) accepted by the BoM. In terms of the analysis of farm data, phases two and four are confirmed as being associated with rainfalls above the long term median. It is possibly worth noting that the station with the poorest relationship with the SOIp (especially in phases two and four), station 055076, also has the highest extreme data point, which may indicate a data quality issue.

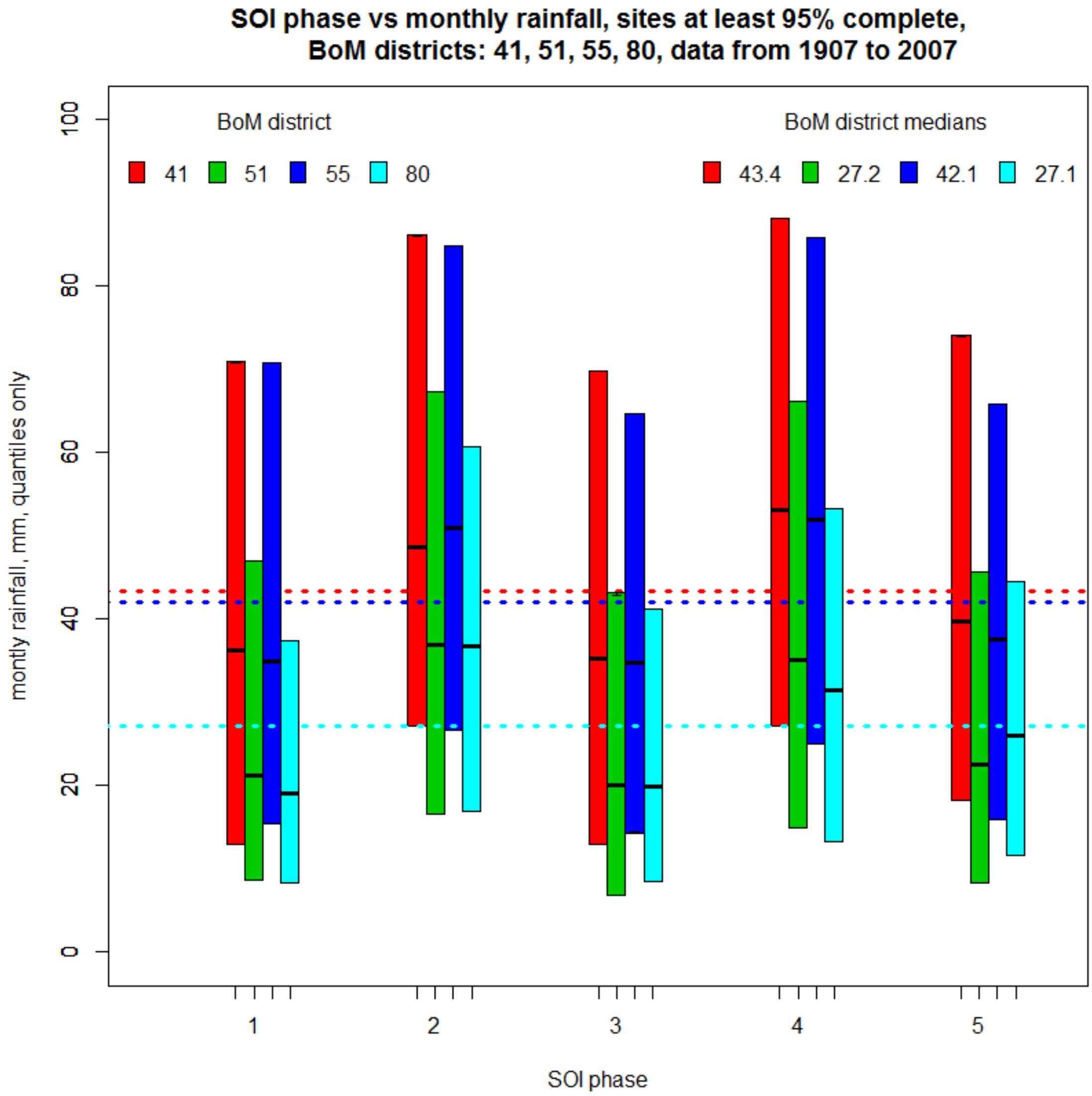
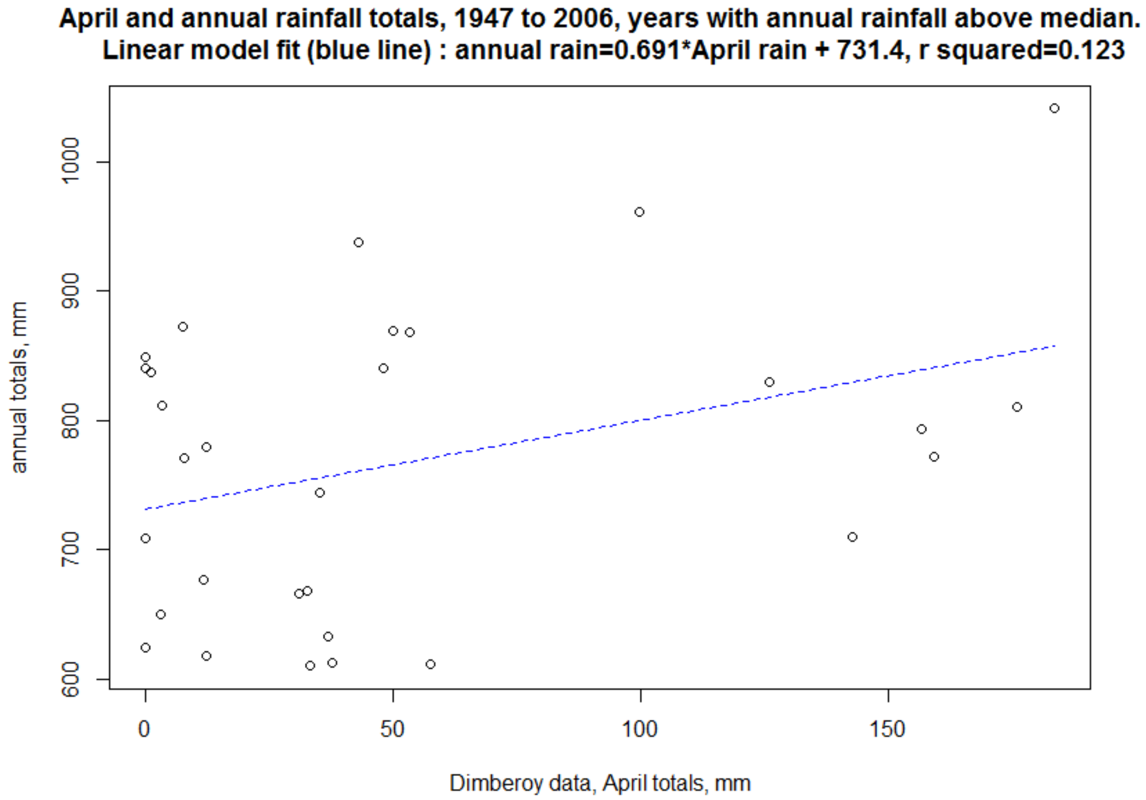


Figure 8-2 SOI phase and monthly rainfalls, BoM districts 41, 51, 55 and 80. Dashed lines indicate median values, as listed on plot (district 51 median invisible behind district 80 median)

8.3.2 SOI phases in relation to farm data

Figure 8-3 shows an example of the SOI phases in relation to farm data. Figure 8-3 (a) shows the relationship between April rainfall and annual rainfall for wet years, for the Dimberoy data. Figure 8-3 (b) shows the relationship between April rainfall and annual rainfall for wet years, for years when the SOI phase was two or four, for the Dimberoy data. The coefficients for a linear model fit to the data sets are also shown (slope, intercept and R squared). Figure 8-4 shows the results of a similar analysis using seasonal data.

a



b

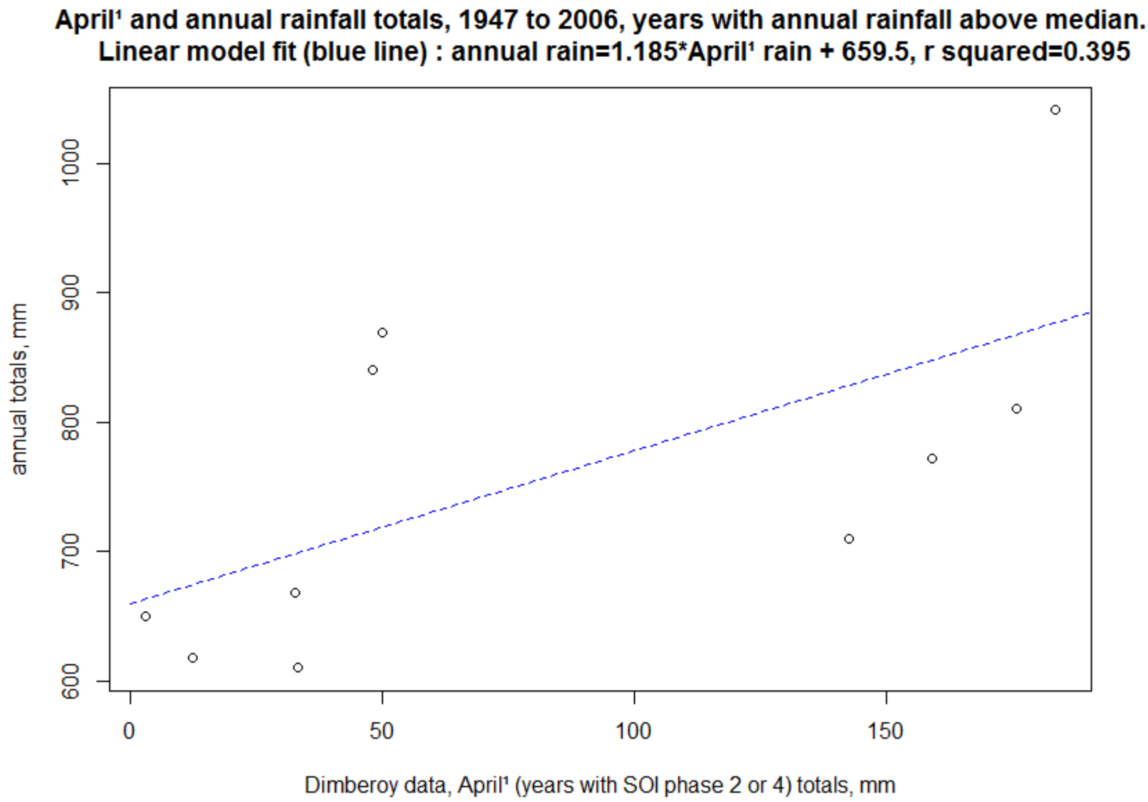
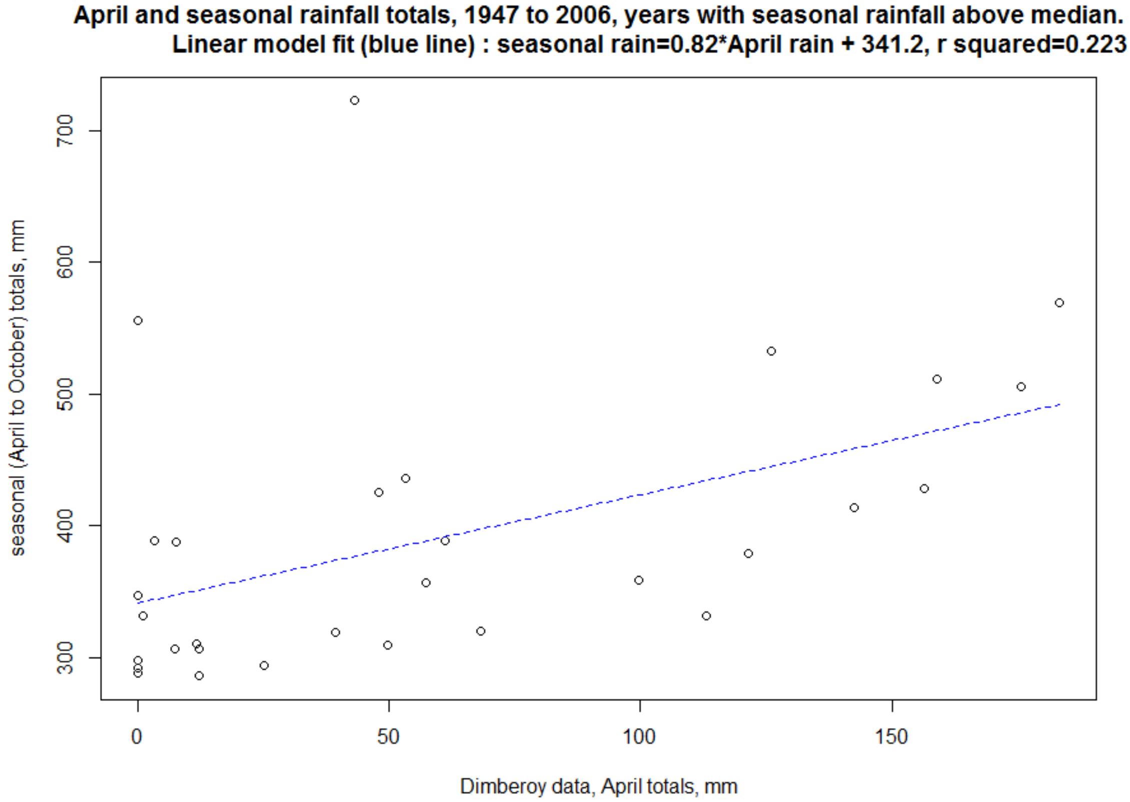


Figure 8-3 April (a) and SOI phase 2, 4 April (b) monthly rainfalls and annual rain, Dimberoy

a



b

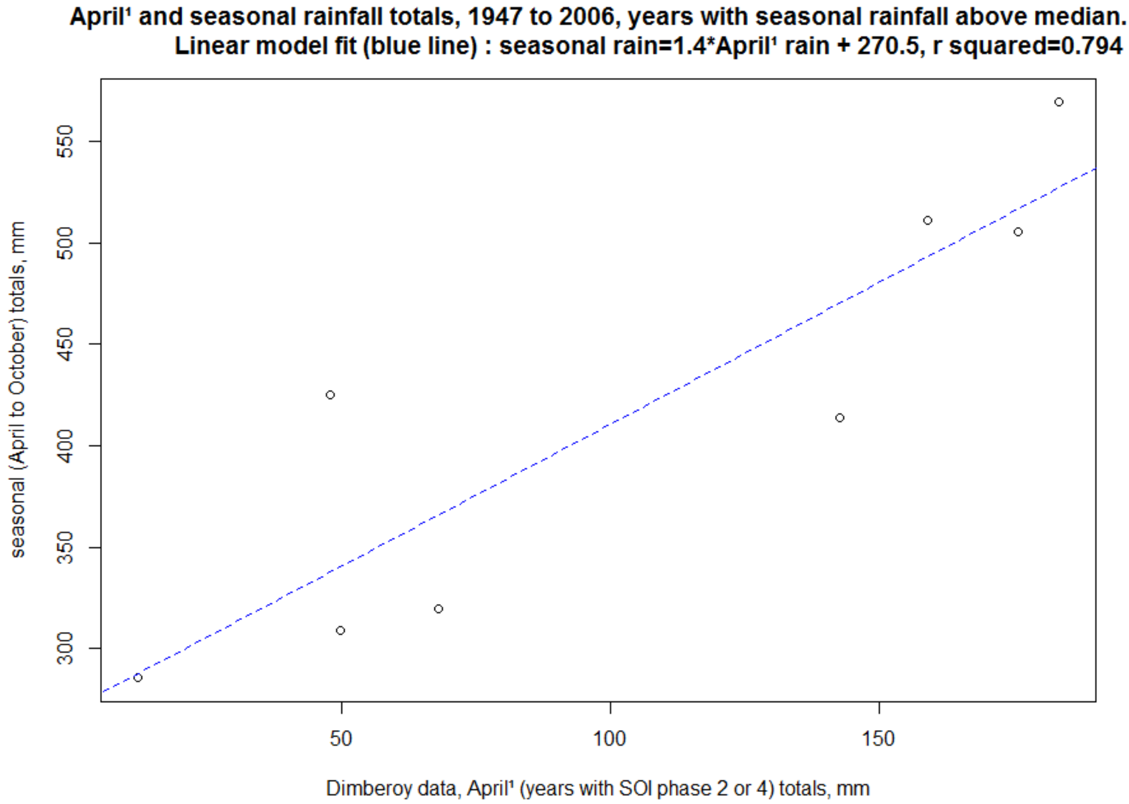


Figure 8-4 April (a) and SOI phase 2, 4 April (b) monthly rainfalls and seasonal rain, Dimberoy

Table 8-1 summarises the analyses, with the maximum change (between all data and SOIp equals two or four) results highlighted, and so more results will not be presented in graphical format.

As with previous results (chapter 7) dealing with indicator time periods within the rainfall data alone, R squared values were not better than about 0.3 to 0.4. Using the years with related SOIp values some much improved R squared values emerge. Districts 51 and 55 were indicated as being similar, with April being possibly the most significant for annual and seasonal relationships, based on rainfall alone. District 80 had the highest R squared values in May, also for annual and seasonal relationships.

The use of SOI phase to select years has provided a much improved relationships:

- district 41 results are inconclusive, showing May to have the best relationship with annual rainfall, unaffected by SOIp data, and also the best relationship with seasonal rainfall, enhanced by the SOIp;

Table 8-2 Summary of monthly and SOI phase and seasonal or annual farm rainfall linear model analyses

month		March		April		May		June	
		all	SOIp 2 or 4	all	SOIp 2 or 4	all	SOIp 2 or 4	all	SOIp 2 or 4
site	BoM district	R squared values							
annual rainfall									
Springfields	41	0.007	0.005	0.261	0.001	0.541	0.413	0.092	0.0
Ashlee	51	0.065	0.48	0.006	0.009	0.051	0.019	0.023	0.004
Dimberoy	55	0.022	0.072	0.123	0.395	0.156	0.11	0.055	0.108
Pine Grove	80	0.123	0.149	0.245	0.723	0.299	0.429	0.004	0.017
seasonal rainfall (April to October)									
Springfields	41	0	0.074	0.275	0.003	0.117	0.537	0.418	0.092
Ashlee	51	0.049	0.362	0.069	0.147	0.109	0.076	0.039	0.049
Dimberoy	55	0.001	0.015	0.223	0.794	0.054	0.118	0.245	0.243
Pine Grove	80	0.102	0.24	0.264	0.445	0.412	0.312	0.036	0.007

a good relationship with June and seasonal rainfall is negated by inclusion of the SOIp
the negative effect of the SOIp for June cannot be explained at this stage, but it is worth noting that district 41 and district 80 have smaller data sets than the other districts;
other results have low R squared values

- district 51 shows improved relationships between March rainfall and seasonal or annual rain, with improvement due to the SOIp greater for annual rains;
  the relationship between April and seasonal rain is improved by the SOIp;
  other results have low R squared value
- district 55 shows improvement with the use of SOIp data for March and April and annual or seasonal rains;
  the greatest improvement is for April and seasonal rain;
  other results have low R squared values
- district 80 also shows general improvement in the relationships between March And April and annual or seasonal rains with the use of SOIp data;
  the best improvement is for April and annual rainfalls;
  other results have low R squared values

Table 8-2 shows the p values for the linear model analyses maximum change values presented in table Table 8-1. The ‘improved’ relationship values using SOIp, are all significant (Johnson 1999) for at least the 0.1 level with most at the 0.05 level. Of the relationships without SOIp, two would be considered significant and two not. The two poorer relationships are for the sites with about thirty years of data, as compared with the other sites which have over fifty years of data.

Another point of interest can be seen by inspecting the rainfall, SOI and SOIp data for some of the example years. The four wettest years shown during the analysis time period at Dimberoy are 1950, 1989, 1990 and 1999.

Table 8-3 shows the relevant data for these years. The BoM considers an SOI above +8 to indicate La Niña conditions (likely wet), and SOI less than -8 El Niño conditions (likely dry)[11]. While 1950, 1989 and 1999 conform to this rule, 1990 does not.

Table 8-3 Significfance (p) values for and SOI phase and seasonal or annual farm rainfall linear model analyses (Table 8-1)

month		March		April		May		June	
		all	SOIp 2 or 4	all	SOIp 2 or 4	all	SOIp 2 or 4	all	SOIp 2 or 4
site	BoM district	p values							
		annual rainfall							
Ashlee	51	0.191	0.009						
Pine Grove	80			0.043	0.068				
		seasonal rainfall (April to October)							
Springfields	41					0.195	0.025		
Dimberoy	55			0.008	0.003				

Note that there is no suggestion here that the BoM 'got it wrong', in fact one year out of four 'incorrect' is probably within the acceptable accuracy. The BoM lists 1989 to 2001 as a moderate, and 1950 as a strong La Niña event[31], showing that the SOI is not the only factor used by the BoM to define conditions. However, the data here demonstrates that the SOIp term appears to define the conditions accurately, contrary to what may be inferred from the BoM classification, confirming the conclusions of Stone and Auliciems (1992). While the BoM describes conditions from 1989 to 2001 as waxing and waning intensity La Niña, it is apparent from these results that farmers may be able to use the SOIp data in combination with their own rainfalls to improve on the available rainfall forecast information, in this case going from variable to 'wet'.

Table 8-4 Rainfall, SOI and SOIp data for wet years, Dimberoy

year	April rain, mm	April SOIp	Annual rain, mm	Annual mean SOI	5 month to April mean

[31] http://www.bom.gov.au/climate/enso/lnlist/

					SOI
1950	183.388	2	569.214	14.2	11.26
1989	175.768	4	505.460	6.3	10.9
1990	142.494	4	413.512	-2.7	-6.9
1999	159.000	4	511.000	7.3	11.6
2000	48.000	2	425.000		

8.4 SOIp data only

As a cross check that the rainfall data is imparting value beyond a relationship based on the SOIp data alone, similar linear models to those described above were run using the same sites with better relationships from the tables above and years which had SOIp values of two or four (indicating wet conditions). The SOIp only results are shown below, and indicate that while the SOIp value is a good indicator on its own, the inclusion of farm data does have a positive effect. Interestingly, the sites with the longer data sets have better relationships and the effect of the farm data is not as strong. Possible influences on these relationships such as size of data set, seasonality of rainfall and possibly timing and duration of events would make an interesting study when more data becomes available.

farm	SOIp and farm data (tables above)		SOIp only data	
	R squared	p value	R squared	p value
Ashlee (March, annual rain)	0.48	0.009	0.475	0.000
Pine Grove (April, annual rain)	0.723	0.068	0.549	0.035
Springfields (May, seasonal rain)	0.537	0.025	0.267	0.071
Dimberoy (April, seasonal rain)	0.794	0.003	0.738	0.000

8.5 Conclusions

Considering the small amount of data available for analysis, these results are very encouraging. While earlier results showed an indicator month is a possibility, the relationships based on rainfall in the indicator periods were not particularly strong. The inclusion of the SOIp terms has shown much stronger relationships, although consistency across all BoM climate classes is not demonstrated.

Stone, Hammer et al. (1996) indicated that the SOIp term could be successful in forecasting median rainfalls for the subsequent three month period. The work presented above demonstrates that, if associated with specific indicator periods, the SOIp term may be useful in predicting entire seasonal rainfall medians, extending its value by two to three months. This capacity is demonstrated to cross three climatic zones, but clearly more work will be needed to confirm this and define more precise relationships.

Future work should also consider similar terms to the SOIp, such the Sea Surface Temperature (SST), (Drosdowsky and Chambers 2001), which provide a similar, categorised index relating to the 'large scale climate state'. At the time of writing, SST class or phase data was not as accessible as the SOIp data.

While the results show some influences which cannot be explained here, they show the value the farm rainfall data can add to existing forecast information, and indicate some areas where further study would be of interest.

Improving the differentiation of dry years has not transpired with the methods and data available in this study. Once again, this invites further study into possible timing and lag relationships between periodic and seasonal rainfall, or the refinement of methods to handle low value data, amongst many other issues.

The value of the farm data, in potentially redefining a season as wet, instead of fluctuating or uncertain, will, I believe, be of great significance to users.

9 Conclusions

This project has been about demonstrating the value of historic daily rainfall data already collected and being continuously collected by farmers in Australia. There was a technical component dealing with data gathering and management, discussed in Chapter 3. The research focus has been on discovering new ways for farmers to use their data which will assist them in understanding their local climate and interpret available forecast material, but more importantly derive some predictive capacity from the data to help in planning their seasonal activities. It is worth restating that the work in this thesis is not aimed at replacing existing rainfall forecasts provided by (for example) the Bureau of Meteorology, or to compete with complex climate or agricultural system models. This document details an examination of the use of rainfall data, as collected by farmers, to help explain spatial and temporal rainfall patterns.

In the literature there has been analysis of issues such as rain gauge density, and the idea that more data will increase the spatial resolution of analyses or models, but there seems to have been little effort in extracting value or information from individual data sites.

An idea developed from the literature review was the desirability of a single parameter or very simple rainfall indicator or index. Ideally this indicator should relate strongly to seasonal or annual rainfall. The idea of a rainfall or crop index has been popular over time, but was dismissed as being impractical as reliance on computerised crop, soil and climate systems developed. Another appealing aspect of such an index is its simplicity, and the transparency of this type of analysis to the farming community. Unlike a complex modelling approach, an uncomplicated index of this type could demonstrate a simple relationship between farmer rainfall data and seasonal conditions for crop growth.

9.1 Data

The collection of farm rainfall data proved to be the most problematic aspect of the project. I believe that the aim of the project in this regard has been met, and a framework is in place for data collection and management. Initial reporting (feedback) systems have been created and are continuing.

The development of a new method of interpretation and display for this data, continuous rainfall deficit / surplus analysis, reveals the relationship between local rainfall and the developing season (or other time period) in terms of likely boundary conditions and averages specific to the users property.

The presentation of local spatial rainfall mapping gives the users an opportunity to explore (spatial) rainfall patterns that may exist on and around their property, once again based on their own rainfall and local conditions. The analysis can be performed for any season or time period, allowing examination of extreme events, wet or dry seasons and more average conditions. While these analyses are of general interest, they may also be useful in planning for specific occasions such as extreme events, where any knowledge of unusual patterns may be helpful.

9.2 Value adding

The next aim of the project was to test the usefulness of the rainfall data in a predictive or interpretive capacity, hence adding value to the rainfall records already held by many farmers.

The lack of a simple relationship between the sites, across different climatic zones once again reflects the spatial and temporal variability of daily rainfall. It would be interesting to re-examine these relationships when more data is available, possibly allowing detection of scale or event intensity effects, especially if they persist across regions.

The extension of the deficit / surplus analysis into a new role as a predictive tool has demonstrated the value of the rainfall data, and how it can be used to forecast the likely deficit / surplus situation for a given season or time period. The predictive deficit / surplus concept could be refashioned to relate to 'how wet or dry is it?', or 'how full or empty is the system' models, both of which should be useful and easily understood by users. The analysis has the capacity to operate in real time, so that the possible development of a season going forward can be shown, which could be an extremely useful tool. It is important to note that while the analysis has universal application, the time scale of the analysis, for example prediction one or more months forward or the whole season, and its specificity to the users conditions, means that it can provide information not available in existing regional or seasonal forecasts.

The partial quantification of the spatial data analyses can allow users to examine specific events, once again such as extreme events, and explore patterns relevant to their property. This analysis is also universally applicable, but provides information to allow users to interpret existing seasonal or regional forecast in more detail, specific to their property.

I believe that these analyses achieve the aim mentioned above, and also confirm the hypotheses listed in chapter 1. I believe it is quite clear that the farm data can provide very useful information to farmers, unavailable from any other source, despite apparent (or perceived) quality issues. A statistical approach, with multiple large data sets, is possibly a good option for dealing with the uncertainties in the rainfall systems.

9.3 Rainfall indicator or index

The analysis of the relationship between all possible (annual) thirty day periods and seasonal or annual rainfall produced some very interesting results. There were various periods which appeared to be reasonably well correlated to seasonal and annual rainfall, in April or May, June or July, and October. While not particularly well defined or applicable across all climate types, these periods did exhibit relationships which could therefore be used in an indicative role. These relationships persisted for seasonal and annual rainfalls, and for wet years. Perhaps not surprisingly, the low rainfall values in dry times did not show similar relationships. While the relationships were not particularly strong, the existence of a relationship at the start of the growing season in most areas, in April or May, was very promising. A genuine indicator of seasonal conditions based on early season rain has been suggested in the past, but not confirmed across different climatic areas.

To further test this idea, an external climatic influence, the five phase southern oscillation index system (SOIp) was used in a focussed analysis of the relationship between rainfall in early season calendar months and seasonal or annual rainfall. The relationship between monthly rainfalls related to years defined as wet (above median annual rainfall) was compared with the same relationship with a subset of years based on the SOIp for the relevant month. The inclusion of the SOIp considerably improved the relationships revealed, with a doubling of relationship strength across all climatic types, although the strength of the relationships differed across the climatic types, and the strongest relationships were split between the months of April and May. While the SOIp subset would probably include mostly high rainfall years, inspection of the data did not show that the subset consisted only of high or extreme rainfall years when high monthly values and therefore high relationships may be expected. Some of the strong relationships occurred when other indicators would have been neutral or variable, emphasising the potential value of this combined indicator. The inclusion of the farm data improved relationships that were developed using SOIp data alone.

Most farmers would have some knowledge of the relevance of the SOI to Australian climate systems. The SOIp values are readily available or calculable, and may therefore provide a parameter, which, when combined with (their own) early season rainfall values, provides and indicator of seasonal rainfall. Perhaps more importantly, the indicator allows the more accurate differentiation of years or seasons likely to be wet, as opposed to neutral or seasons with indeterminate conditions. I believe this indicator is simple enough to be practical and would be readily understood by farmers.

9.4 Future work

Crowdsourcing and citizen science, both terms which are popular at the moment, refer to the collection and or analysis of data from a large number of people, typically by the internet (OED 2013), and commonly from people not part of the scientific community. Farmers were perhaps among the first purveyors of crowdsourced data, being by necessity observers of the natural world and in some cases recorders. There is a wealth of rainfall, and probably other, data held by farmers in Australia. Its collection, documentation and protection should be a high priority for the scientific, environmental and educational communities. This project has demonstrated the potential value to farmers of this data, and its intrinsic value exists as baseline data for climatic and other studies, verification data sets for climatic and hydrologic modelling and as a time series for climatic model testing. It is conceivable that one day soon computerised climate and weather modelling will be functional at a farm or paddock scale, over weekly or similar time steps. Until that time, the ability to harness the information in existing data sets is very valuable. Clearly the continuation of the collection of this data is of prime concern, despite the difficulties displayed by this project. It is hoped that the results of this project may provide some incentive to continue the process.

An immediate opportunity arises from the collection of further data, which is to understand and analyse the current uses and value or perceived value (to the farmer) of the data being collected (by the farmer). While some proprietary systems apparently use local data, it is unclear what role an individual farmer’s data has in such systems, or in any other role. There was no evidence of data sharing within farmer groups, which is another area of potential significance.

Quality control is an issue which should be addressed in the farm data. A set of standard codes, perhaps similar to those already used by the Bureau of Meteorology would be appropriate. While there is strong evidence to support the value of the farm data, and good evidence confirming the accuracy of most of the data, the data sets should have good metadata to help future users.

The promise of the predictive deficit / surplus analysis suggests that the rainfall data itself may be adequate to provide the ideal predictive reference.

The definition of a predictor year (or other period) presents an exciting project in itself, with time series clustering and pattern matching a likely starting point, and with long term records providing the necessary data.

An investigation into the effects of differing time scales for the analysis would be beneficial. For example the effect of running any analyses over two year time periods instead of one, or using some earlier wet or dry point as the starting point. The accuracy of predictive time ranges would also be of interest. While it might be expected that a minimal predictive time step would allow greatest accuracy, based on the similarity between the predictor year and analytical year conditions, there could be an ideal time frame which relates best to user requirements such as weekly or monthly management decisions, while maximising accuracy. The examination of different, shorter, time frame predictions, with automatic updating in interactive systems could negate problems with inaccuracy in longer term predictions, and be appropriate for within season management decsisions.

The strong indication that early season farm rainfall data, conditioned by the SOIp, can indeed forecast seasonal conditions is an exhilarating prospect that needs further testing and development. As with the deficit / surplus analysis, the results obtained with limited data were better than anticipated, and provide a good incentive to continue the task. The use of the farm rainfall and SOIp data to refine the information currently available, such as from indeterminate to wet or dry, should prove very useful.

The use of the predictive capacity of the farm rainfall data should be explored in relation to water balance, crop yields other agricultural management systems. Precision agriculture data use, decision making and risk management systems may all be structures where the data and analytical methods could be incorporated. In examining an expanded role of rainfall analysis into such areas as crop yields, new analyses would need to include other important parameters. For example, given that temperature can be an important, if secondary, limiting factor for wheat production, these effects should also be examined and included as appropriate.

In conclusion, despite the difficulty experienced in obtaining data, I believe this project has conclusively fulfilled the stated aims, and demonstrated the value of farm rainfall data, both archival and being continuously collected. The results demonstrate that with the unique analyses developed, farmers can use their own data (also both archival and being continuously collected) to interpret seasonal and regional forecast material available to them, and as a method of predicting likely seasonal rainfall outcomes unique to their property, with greater specificity than methods currently available to them.

10 Appendix 1

10.1 C E Hounam data – Seasonal Growth Factor

Table IX hounam 1947

location	average yield bushels/acre	yield t /ha	SGfactor june - october
Molong	16.9	1.1267	8
Parkes	14.2	0.9467	3
Peak Hill	10.3	0.6867	1.5
Forbes	10	0.6667	1.9
Burrangong	16.6	1.1067	12
Young	19.8	1.3200	13.3
Murrumburrah	16.9	1.1267	16
Holbrook	17.7	1.1800	22
Hume	17	1.1333	15
Illabo	18.4	1.2267	7
Junee	11.6	0.7733	6.9
Jindalee	17.1	1.1400	10
Kyeamba	18.1	1.2067	10
Mitchell	16.8	1.1200	7
Temora	15.2	1.0133	3.7
Weddin	16.1	1.0733	5
Grenfell	11.5	0.7667	8
Condoblin	5.4	0.3600	0.4
Berrigan	14.6	0.9733	1.8
Carrathool	7.2	0.4800	0.3
Conargo	12.5	0.8333	0.4
Coolamon	15.8	1.0533	2.2
Coreen	16.1	1.0733	6
Corowa	15.4	1.0267	9
Culcairn	16.9	1.1267	13
Jerilderie	12.9	0.8600	1
Lockhart	17.4	1.1600	5
Murray	11.2	0.7467	1.5
Deniliquin	8.2	0.5467	1.3
Moama	11.4	0.7600	1.8
Murrumbidgee	10.2	0.6800	0.5
Urana	14.4	0.9600	1.8
Wakool	8	0.5333	0.5
Windouran	8.6	0.5733	0.5

10.2 C E Hounam data – Influential Rain Factor

Table XI hounam 1947

location	average yield bushels/acre	yield t /ha	ir factor june - october
Molong	16.9	1.1267	28
Parkes	14.2	0.9467	10.1
Peak Hill	10.3	0.6867	5.7
Forbes	10	0.6667	9.3
Burrangong	16.6	1.1067	25
Young	19.8	1.3200	31.8
Murrumburrah	16.9	1.1267	35
Holbrook	17.7	1.1800	17
Hume	17	1.1333	20
Illabo	18.4	1.2267	25
Junee	11.6	0.7733	26.7
Jindalee	17.1	1.1400	25
Kyeamba	18.1	1.2067	45
Mitchell	16.8	1.1200	35
Temora	15.2	1.0133	16.9
Weddin	16.1	1.0733	14
Grenfell	11.5	0.7667	19.3
Condoblin	5.4	0.3600	2.1
Berrigan	14.6	0.9733	9
Carrathool	7.2	0.4800	2
Conargo	12.5	0.8333	3.5
Coolamon	15.8	1.0533	10
Coreen	16.1	1.0733	13
Corowa	15.4	1.0267	17
Culcairn	16.9	1.1267	33
Jerilderie	12.9	0.8600	6
Lockhart	17.4	1.1600	18
Murray	11.2	0.7467	9
Deniliquin	8.2	0.5467	6.6
Moama	11.4	0.7600	10
Murrumbidgee	10.2	0.6800	4.5
Urana	14.4	0.9600	9
Wakool	8	0.5333	4
Windouran	8.6	0.5733	3

11 Appendix 2, Data gathering

11.1 GrainGrowers promotion December 2009

Industry News

Grant to investigate improved seasonal rain prediction

FOLLOWING THE UNIVERSITY of Sydney's successful funding application to the Australian Research Council (ARC), GGA is now an industry partner to a project that will look at improved seasonal rainfall prediction for grain growers using farm level data and novel modelling.

Growers will know that successful grain production depends heavily on reliable seasonal forecasts. However, the highly variable climate means that current forecasts lack detail in space and time.

Using a combination of 'fuzzy' classification and artificial neural networks, this project will develop a locally detailed continuously updating data-driven seasonal forecast system. The system will use high density climate data that GGA members will have the opportunity to provide, as well as climate drivers such as sea surface temperature from the Bureau of Meteorology. After validation against observed data, the forecasts will be delivered via a web-based portal to users.

The project is expected to commence in the first few months of 2010 and we will be seeking cooperation from GGA members to assist in providing farm rainfall records to assist this project.

Congratulations to Dr Rutger W Vervoort, Dr Budiman Minasny, and Prof Alexander B McBratney from the Faculty of Agriculture, Food and Natural Resources at the University of Sydney for developing this project and the support of ARC. We look forward to working with them to assist growers with improved seasonal forecast development.

NSW Grain Freight Review

IN NOVEMBER 2008 the Australian Government initiated a review into NSW grain freight. The review assessed the 'efficiency of grain supply chains in NSW' and focused on the future of the branch line network and the wider implications for road management if rail is no longer available.

The key messages from the report focus on five key areas, being:

1. Retention of the majority of branch lines
2. Government investment should be contingent on industry contributions
3. The need for long term planning and investment certainty
4. Road access and investment must be considered as part of a complete supply chain
5. Ongoing improvement across the whole supply chain.

The report noted that the grain flows in NSW had three major markets — domestic stock feed, domestic human consumption, and exportable surplus with four main supply chains — direct exports (44%); net interstate flows (18%); NSW animal feed (23%) and NSW human consumption (15%).

The following table (see right) provides a brief summary of the review's recommendations.

Members interested in reading the complete Review and its findings and recommendations in more detail can download it by visiting www.nationbuildingprogram.gov.au

11.2 GrainGrowers promotion and questionnaire March 2010

Industry News

A novel way to improve seasonal rainfall prediction

EACH YEAR IN Australia 25,000 grain farmers face the same big and difficult decision: how much to invest in seed, fertiliser and other inputs for the coming season?

Increasing uncertainty regarding climate exacerbates this decision, however with improved forecasts of the coming season such decisions could be made more favourably and with less risk.

GGA, in partnership with the University of Sydney and supported by funding from the Australian Research Council, has commenced a project that will combine the weather data collected on-farm by most grain farmers with existing data from the Bureau of Meteorology (and elsewhere), to develop a new type of very local seasonal climate predictor. These more sophisticated individualised forecasts will then be delivered to many thousands of grain farmers.

What we need now is to enlist the support of growers to provide GGA with access to your individual rainfall and other weather information.

The project will run over a number of years but as a first step we would like to develop an initial list of growers willing to participate. If you're interested in being involved, please complete the form below and return it to GGA.

This form is also available on the GGA website by visiting **www.graingrowers.com.au**

☐ **Yes, I am willing to participate in this project and share my weather records to assist to improve long term seasonal forecasts.**

Name: .... GGA Membership No: (if known) ....

Address: ....

Phone: (business hours) .... (mobile): ....

Fax: .... Email address: ....

Do you keep the following weather records:

1. a) Rainfall: ☐ Yes ☐ No
   b) Date records commenced: / /
2. a) Temperature: ☐ Yes ☐ No
   b) Since when? / /
3. a) Wind speed and direction? ☐ Yes ☐ No
   b) Since when? / /
4. a) Relative humidity? ☐ Yes ☐ No
   b) Since when? / /
5. Do you have an automatic weather station? ☐ Yes ☐ No
6. How do you record your weather details?:

   ☐ Paper ☐ Spreadsheet (electronic)
   ☐ Other: (please detail any other records you have)

   ....
   ....
   ....

Fax back to GGA on 02 9886 229 or email membership@graingrowers.com.au. For further information call GGA on 1800 620 519

12 Appendix 3, Farmer data

Typical data location : ‘ ...\rain project data tables\farmer\ ... ‘

12.1 Farmer data input 1. farm information

Farmer contact, location and data delivery information is stored as an SQL data base ‘...\rain project data tables\sql\rain1.sql’. Data from this data base is exported as ' ... \rain project data tables \farmer_dat_info.csv'.

R script files (eg: ..... \farm data set.r) read this farm information file for use as required.

12.1.1 farmer_data_info table

User id and basic information:

- references SQL farmer_contacts table for identifer
- updated file exports to file ‘C:\ ... rain project data tables\sql\farmer_data_info.csv’
- file above (exported csv file) is also used to import new data from GrainGrower files, then used to repopulate sql file

Table :

Column heading	Data
ID	project ID (pid) – item number with preceding ‘A’
mem_num	GrainGrowers membership number
fname	property owner first name
lname	property owner last name
town	property address or location reference if address unavailable
property	name of property
	NOTE property name MUST be correctly capitalized, eg ‘Derwent Park’, not derwent park
latitude_dd	latitude, decimal degrees
longitude_dd	longitude, decimal degrees
altitude	altitude, m
digital_daily_d	existence of digital daily data – ‘ddd’ (Y\N and possibly comment such as ‘paper’)

ata_ddd	
ddd_loaded	date ddd saved as date and rain only text file, project code eg 'F00046.txt' in C:\Documents and Settings\dyat2433\My Documents\rain project\data tables\farmer daily txt'
ddd_start	first year of ddd
dd_years	number of years of ddd
digital_weekly _data_dwd	existence of digital weekly data – 'dwd' (Y\N and possibly comment such as 'points')
dwd_loaded	
dwd_start	first year of dwd
dd_years	number of years of dwd
digital_monthl y_data_ddd	existence of digital monthly data – 'dmd' (Y\N and possibly comment such as 'paper')
dmd_loaded	date dmd saved as date and rain only text file, project code eg 'F00046.txt' in C:\Documents and Settings\dyat2433\My Documents\rain project\data tables\farmer daily txt'
dmd_start	first year of dmd
dmd_years	number of years of dmd

12.2 Farmer data input 2. farm data

Basic data consists of rainfall records from farms. Various file types and data formats are possible, from digital daily data to scanned paper images and retyped tables in word processing and Portable Document Format (PDF).

12.2.1 Digital data

Original data files are commonly created in MS EXCEL, for each property, and stored as ' ... \rain project data tables\farmer\PROPERTY.xlsx', where 'PROPERTY' is the name or locality id for the farm if no name is available.

From these files text files with daily rainfall are created manually or by R scripts for interpreting standard EXCEL files. Text data files, as for other data files (eg BoM) are simple date and rainfall files,

For example :

\rain project data tables\farmer\Example Farm.xlsx

Provides data for :

\rain project data tables\farmer daily txt\Example Farm.txt

If the data is provided as (or easily convertible to) simple date and rainfall data), a text file with the project identifier (A00XXX, being supplied by the SQL identifier in 'farmer_data_info'), is generated.

In both cases the resulting text files are saved in the '\rain project data tables\farmer daily txt' directory and contain :

Example Farm date	mm
1876/12/8	0
1876/12/9	10.4

12.2.1.1 Monthly data

Sites with monthly and daily data are treated the same as sites with daily data only, with the monthly data being included in the generated text file. Dates for monthly data are listed as the last day of the relevant month.

Sites with only monthly data and treated similarly to sites with daily data, with the generated text data files, consisting of columns of date (year and month number only) and monthly total rainfall being saved to '\rain project data tables\farmer monthly txt\ ...

12.2.2 Paper data

Paper data is stored in various forms as (for example) ' ... \rain project data tables\paper data\farmer\PROPERTY.*', and, when transcribed into EXCEL, saved as '' ... \rain project data tables\ farmer\PROPERTY.xlsx'. Transcribed files include data suitable for saving as text files for use in R, or in standard EXCEL file interpretable by R.

For example :

Data consisting of photographs of rainfall record sheets, in this case for 'Springfields' farm, are saved as: \rain project data tables\paper data\farmer\Springfields\1975.jpg, 1976.jpg, ... 1997.jpg, 1998.jpg ...

and provide data for :

\rain project data tables\farmer\Springfields.xlsx

which in turn provides data for :

\data tables\farmer daily txt\A00013.txt, with properties as described above.

12.3 SQL files

Data base : (for example) 'localDB\rain1.sql'

Database consists of two descriptive metadata tables; 'farmer contacts' and 'farmer data info', and all daily rainfall and monthly text data in separate tables.

12.3.1 farmer_data_info table

Identical to 'farmer_data_info.csv' table discussed above.

12.3.2 farmer_contacts table

This table contains user (farmer) names, addresses and contact information, plus a chronology of contact details or contact attempts.

based on GrainGrowers excel file 'C:\ ... \Members responses as at DD-MM-YYY.xlsx'
updated file exported to file 'C:\ ... rain project data tables\sql\farmer_contacts.csv'
file above (exported csv file) is also used to import new data from GrainGrower files, then used to repopulate sql table

Table :

Column heading	data
ID	project ID – item number with preceding F
mem_num	GrainGrowers membership number
fname	property owner first name
lname	property owner last name
property	name of property NOTE property name MUST be correctly capitalized, eg 'Derwent Park', not derwent park
town	property address or location reference if address unavailable
phone	phone number
mobile	mobile phone number
email	email address
data_in	type of data received

	eg digital daily, digital weekly or paper
contact1	type and date of first contact
	eg email 18 Oct 2010, GGA if GGA (now GG) made (unrecorded) contact
contact2	subsequent contacts
contact3	
contact4	
contact5	

13 Appendix 4, BoM data

13.1 BoM data input 1. station information

13.1.1 Rainfall files

File for NSW sites from BoM disc is (for example):

\rain project data tables\BoM\disc\sites\dr_nsw_sites.txt

R script (eg : ..... \BoM disc stn info.r) reads BoM standard station information files and writes reformatted list to file, using BoM station ID with preceding 'B' as in column 1 below, as project id.

For example 'BoM_P_stninfo.txt', (not all columns) :

st	st_num	dist	st_name	open	close	lat
B046000	046000	46	TIBOOBURRA (BINERAH DOWNS)	1939-01-01	1952-12-31	-29.0333
B046001	046001	46	WHITE CLIFFS (BOOTRA)	1890-01-01	1944-12-31	-30.0167

Table below lists data provided on BoM disc.

Column heading	data
st	project record identifier
st_num	Bureau of Meteorology Station Number.
Dist	Bureau of Meteorology Rainfall district code
st_name	Bureau of Meteorology Station Name.
open	month\year site opened. (YYYY\MM\DD)
close	month\year site closed. (YYYY\MM\DD)
lat	latitude to 4 decimal places, in decimal degrees
long	longitude to 4 decimal places, in decimal degrees
meth	method by which latitude\longitude was derived
state	state
alt	height of station above mean sea level in metres
alt_bar	height of barometer above mean sea level in metres
WMO	WMO (World Meteorological Organisation) Index, a number assigned to a site that makes international weather reports every day

first_year	first year of data supplied in data file
last_year	last year of data supplied in data file
pc_comp	percentage complete between first and last records
pc_Y	percentage of values with quality flag 'Y'
pc_N	percentage of values with quality flag 'N'
pc_W	percentage of values with quality flag 'W'
pc_S	percentage of values with quality flag 'S'
pc_I	percentage of values with quality flag 'I'
eor	# symbol, end of record indicator

13.1.2 BoM quality flag descriptions

* QUALITY FLAG DESCRIPTIONS

____

Y: quality controlled and acceptable

N: not quality controlled

W: quality controlled and considered wrong

S: quality controlled and considered suspect

I: quality controlled and inconsistent with other known information

X: no quality information available

13.1.3 Evaporation files

File for evaporation station information files (one file) : ...\BoM disc evap\IDCJDC05_stations.txt (for example, name depends on BoM).

R script (eg : ..... \BoM disc stn info.r) reads BoM standard station information files and writes reformatted list to file ‘BoM_E_stninfo.txt’, using BoM station ID with preceding ‘BE’ .

Formats and columns are as for rainfall files.

13.2 BoM data input 2, station data

13.2.1 Rainfall

Data from Bureau of Meteorology (BoM) sites is from BoM data disc, saved as (for example) ‘... \data tables\BoM\disc\NSW\dr_055023.txt’ for BoM station number 055023, ‘Gunnedah Pool’.

R script file (eg: ..... \ BoM disc rain.r) reads BoM station rainfall files in disc format, and saves a table with date, rainfall (mm) and quality data in single station text files, with project id as filename.

For example :

\rain project data tables\BoM\disc\NSW\dr_055023.txt

Provides data for :

\rain project data tables\BoM disc daily txt\B055023.txt

containing :

GUNNEDAH POOL date	rain_mm	qual
1876/12/8	0	Y
1876/12/9	10.4	Y

13.2.2 Evaporation

Similar to rainfall file processing, data from Bureau of Meteorology (BoM) sites is from BoM data disc, saved as (for example) '...\data tables\BoM disc evap\data\NSW\dr_055024' for BoM station number 055024, 'Gunnedah resource centre'.

\rain project data tables\BoM disc evap\data\NSW\dr_055024.txt

Provides data for :

\rain project data tables\BoM disc daily txt\BE055024.txt

14 Appendix 5, Pinneena data

(waterinfo.nsw, formerly NSW Dept Water Environment and Energy)

Typical data location :

C:\ ..... \rain project data tables\Pinneena NSW\ ...

14.1 Pineena data input 1. station information

File 'SITE.CSV' is output from Pinneena site data manager, exported as CSV

R script (eg : ..... \ Pinneena stn info.r) reads 'SITE.CSV', writes reformatted list to file, using Pinneena station ID with preceding 'P' as in column 1 below, as project id.:

For example Pinneena_NSW_stninfo.txt, (not all columns) :

proj_id	Pin_id	Stn_name	start	end	num_days	lat_dd	long_dd	elev_m
P012001	012001	HOMESTEAD CREEK AT FOWLERS GAP	1976-10-21	1995-12-16	6994	-31.0864	141.6961	0
P012002	012002	FRIESLICH CREEK AT FRIESLICH DAM	1977-03-19	2003-05-31	9569	-31.0676	141.6677	0
P070168	070168	STEEPLE FLAT-(CBM)	1957-09-17	1987-12-23	11054	-36.56	149.3766	0

Table below lists data provided by Pinneena data manager.

Column heading	data
proj_id	project record identifier
Pin_id	Pinneena station number.
Stn_name	station name.
start	month\year site opened. (YYYY\MM\DD)
end	month\year site closed. (YYYY\MM\DD)

num_days	number of days of record
lat_dd	latitude to 4 decimal places, in decimal degrees
long_dd	longitude to 4 decimal places, in decimal degrees
ll_datum	datum for latitude and longitude
elev_m	height of station above mean sea level in metres.
elev_acc	elevation accuracy (Pinneena supplied)
pc_bad	% bad or missing data (Pinneena supplied)

14.2 Pinneena data input 2. station data

R script file :..... \read HYCSV.r

Read data from multiple individual station files output by Pinneena (9.2) HYCSV program, write to comma separated values files with location in header, date in first column, followed by sequence of columns for individual station rainfalls, file names are station ID.

For example :

\rain project data tables\Pinneena NSW\HYCSV single\P419032.txt

becomes :

\rain project data tables\Pinneena NSW daily txt\P419032.txt

containing :

COXS CREEK AT BOGGABRI date	rain_mm
1996/03/02	NULL
1996/03/03	0

15 Appendix 6, The Feedback Report

This example is purely representational, the data has not been checked or corrected in any way. The accompanying text was edited between different versions of the report, and so is also purely indicative of the reports.

FACULTY OF AGRICULTURE AND ENVIRONMENT

THE UNIVERSITY OF SYDNEY

BETTER RAINFALL FORECASTS FOR GRAIN GROWERS

Dear *Name*,

Thank you being involved with this project and for supplying us with your rainfall data.

This update contains a number of algorithms through which we have run your data. I would like to emphasise that the algorithms are examples of what we can produce and if would like any specific data to be run through them, we would be more than happy to oblige.

Throughout this update, reference is made to “your local BOM site” – This isn’t necessarily the closest BOM site to your property. Data from “your local BOM site” will either come from the closest high quality BOM site or a local BOM site with the longest data records.

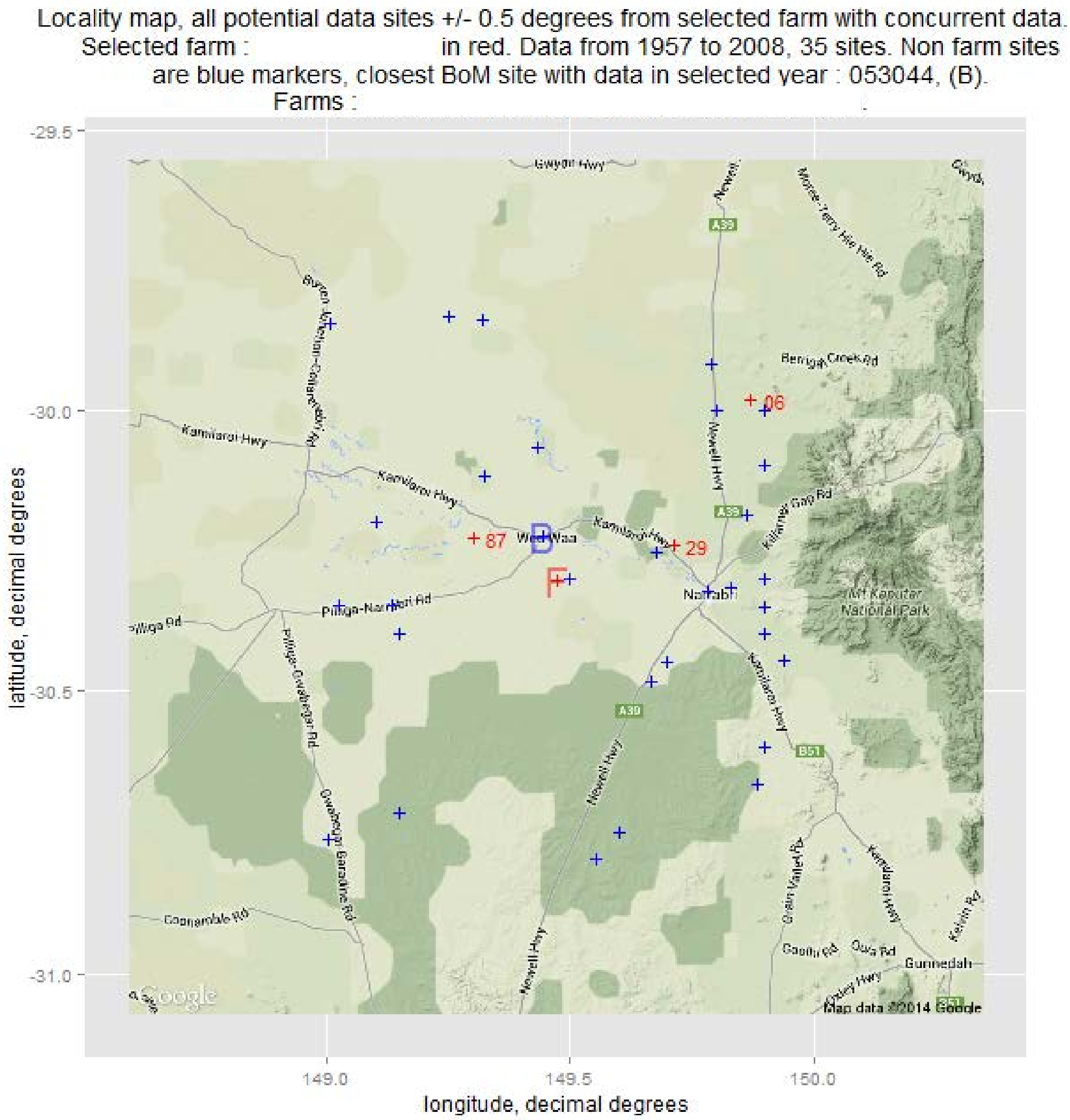
We are currently working on a number of innovative initiatives such as an on-line portal for instant feedback as well as improving the current algorithms and generating new developments which will assist in transforming your rainfall data into more accessible, informative and valuable information.

In the meantime, we shall continue to supply you with updates as the project progresses. We would like to remind you though, that the information you receive about your property will be greatly enhanced with additional data-points, therefore if you could encourage others to get in contact with us, it will not only benefit them, but you as well. Accordingly, attached to this update is a promotional flyer which you can pass around.

If you have any questions about your data, this update or the project, please don’t hesitate to contact me.

Regards,

E. agricgga@sydney.edu.au

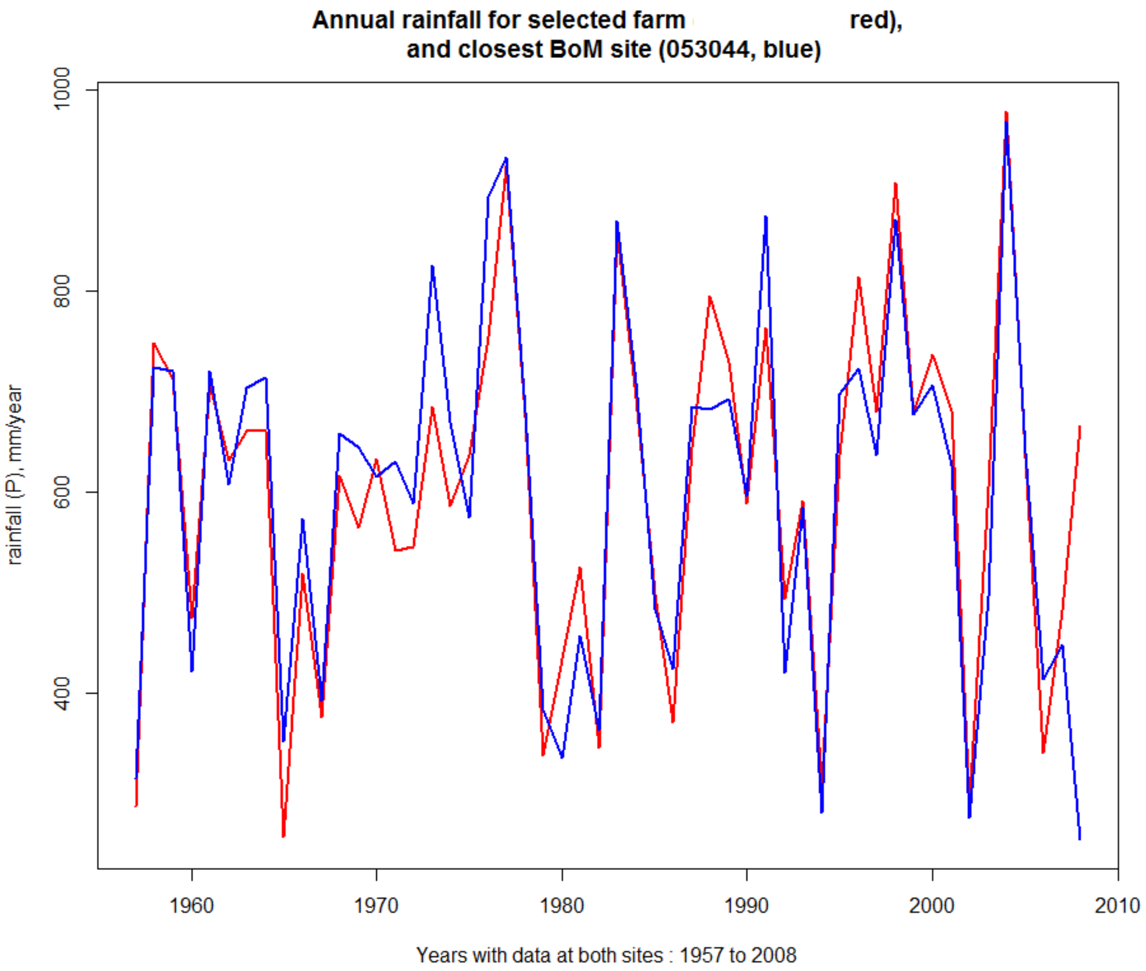


The locality map shows *Property Name* in the center (marked red).

The black dots represent other data points around *Property Name*.

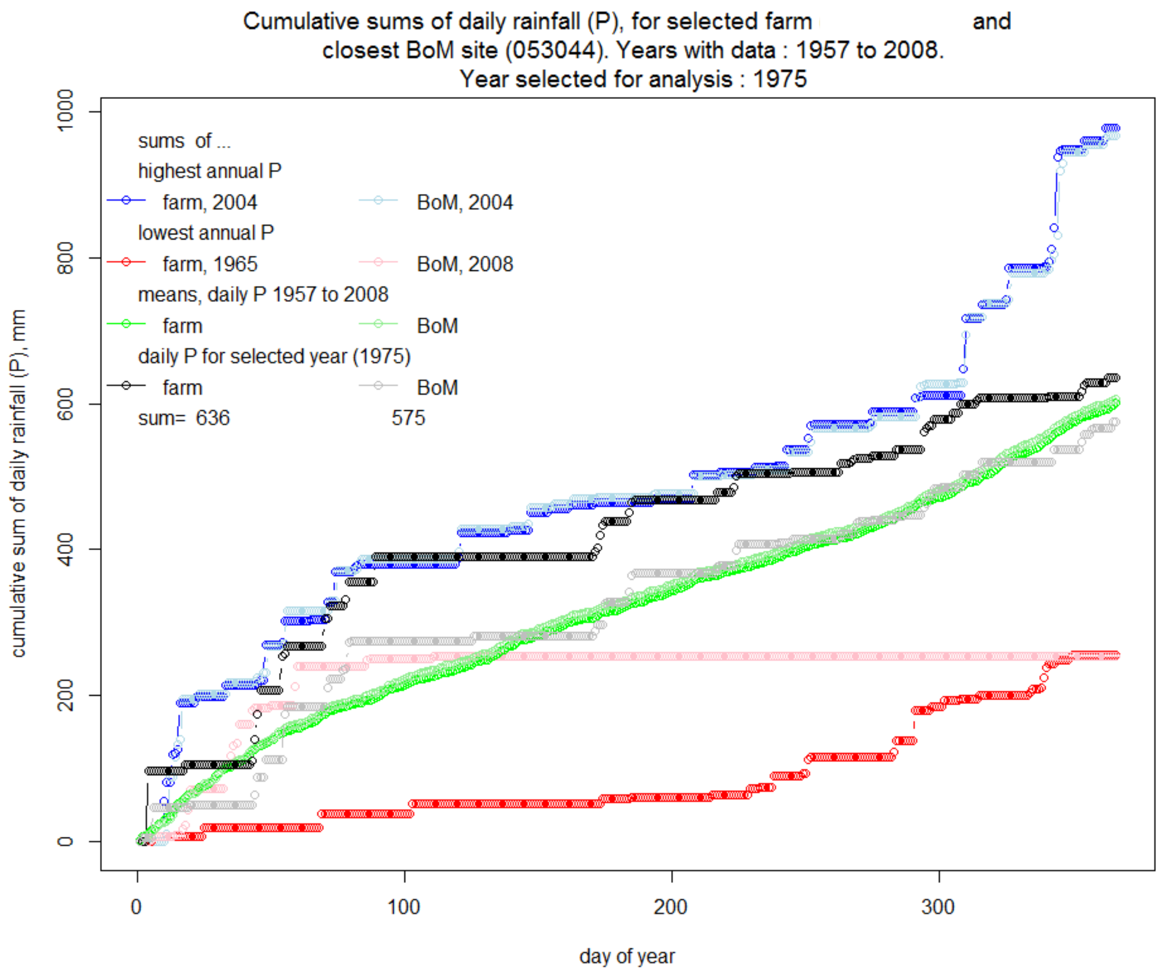
These data points are taken from Bureau of Meteorology (BOM) sites, other organisations and private records.

The greater the density around your property, the more accurate and detailed your feedback will be.



The annual rainfall map displays a comparison of *Property Name* with your local BOM site.

From this, you may be able to decipher patterns regarding the rainfall at your property compared with that of the BOM site and possibly better interpret BOM seasonal outlooks. The relationship between your farm records and the BoM site in extreme events could be of particular interest.

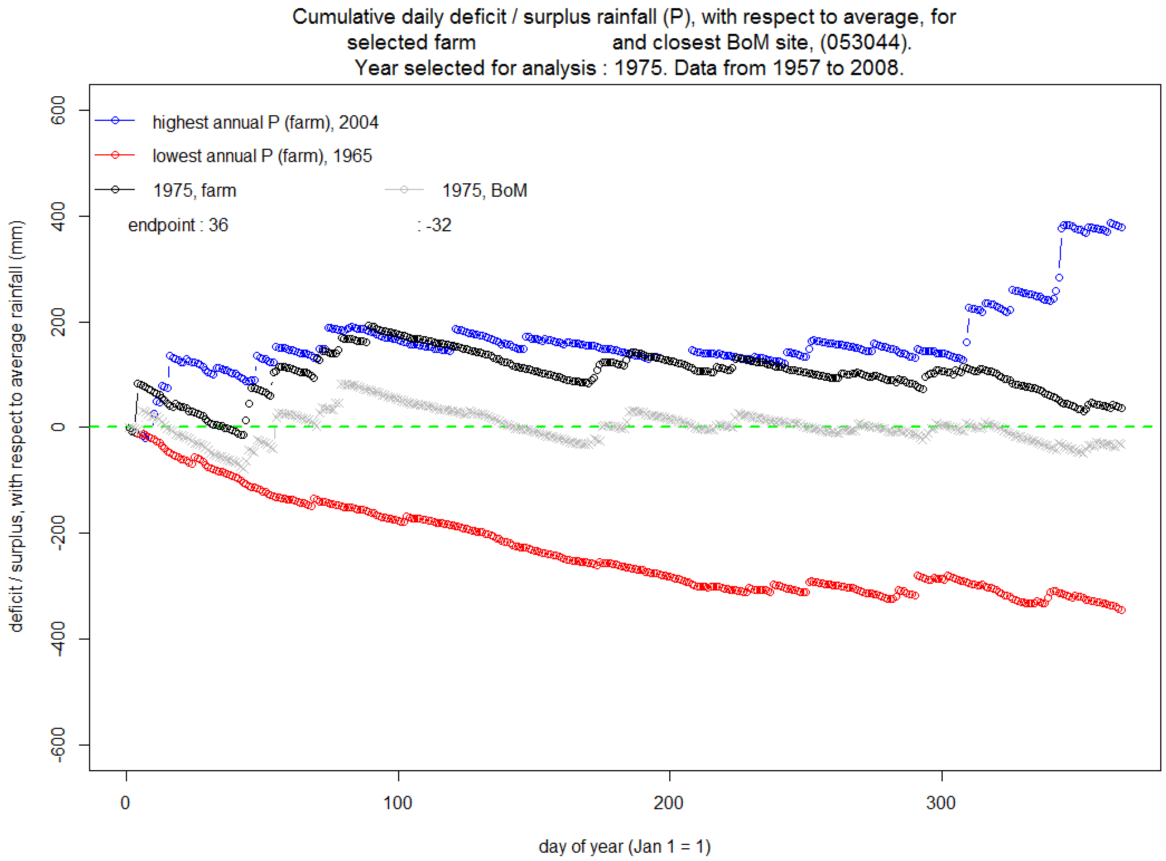


The cumulative sums of daily rainfall shows the cumulative rainfall for all reported years.

This graph considers *Property Name*.

The blue and red lines represent the wettest and driest years respectively. The green line represents your average rainfall and the black and grey lines highlight a selected year, in this case 1975. You can select any year for comparison (within the range of your farm data).

Considering these graphs, you may interpret specific periods throughout the year which are historically dry or wet, and how a selected year compares with your historical record.



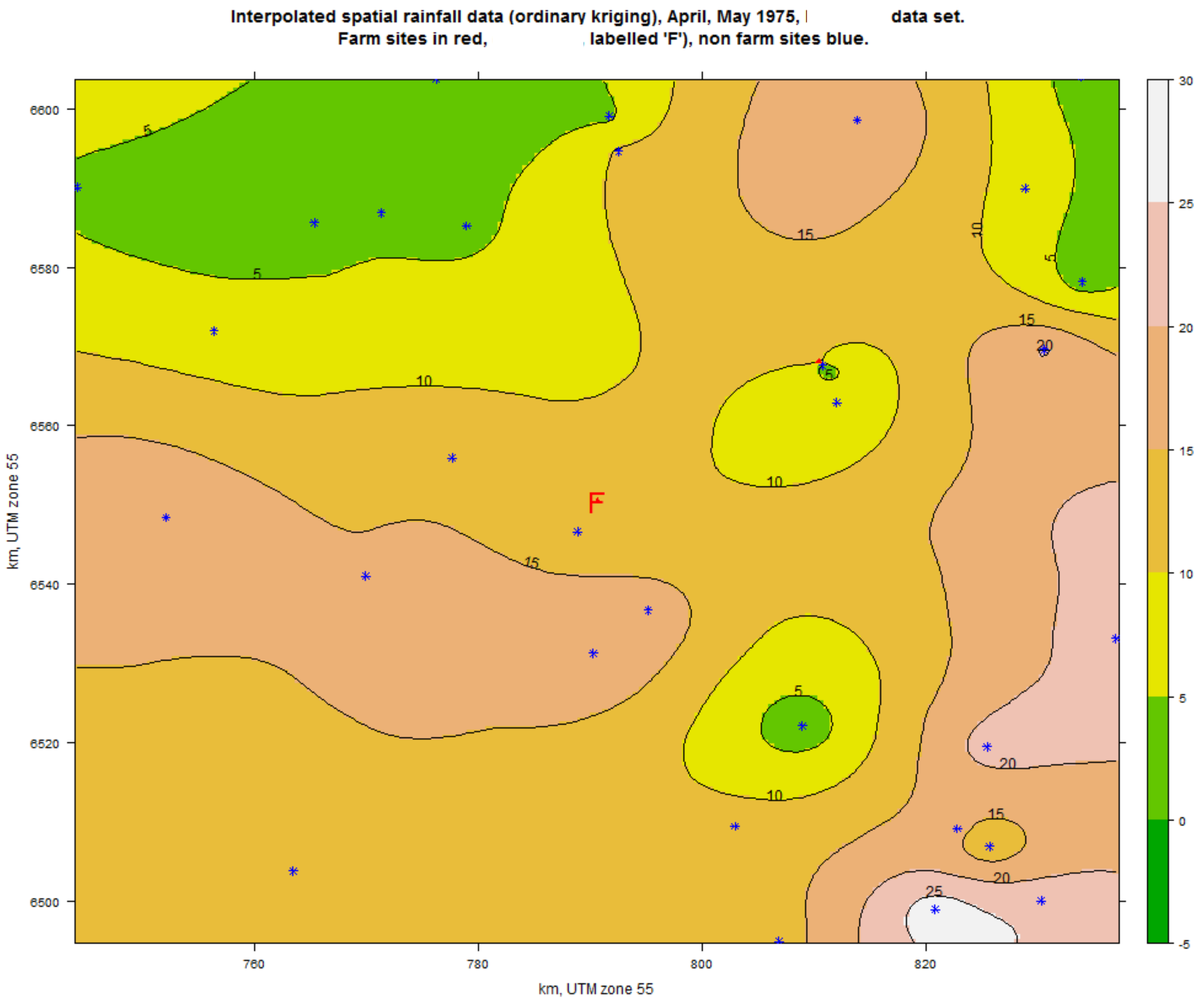
The cumulative daily deficit/surplus (DSP) map shows the cumulative rainfall, with respect to the average for a given year. This graph considers *Property Name*.

The green line represents the average rainfall, calculated on a daily basis. The black and grey lines represent the selected year for the farm and BoM site respectively, in this case, 1975. The blue and red lines represent the highest and lowest rainfall years for the farm respectively.

If at a selected time, your property has received above average rainfall, the black line will be above the green line. The black line will likewise be below the green line at any time that the cumulative rainfall is less than the cumulative average.

Because the DSP is a cumulative plot, the value can be interpreted as an indicator of how full (wet) your system is, or how much any predicted rain will fill the system if it is in deficit.

We are currently working on going live with this algorithm. That is, working to set up an on-line portal where you can enter rainfall data as it happens and view how the cumulative rainfall is progressing with respect to the average, min and max. With this, you will be able to see, in real-time, how the yearly rainfall on your farm is progressing.



This map represents the rainfall variation around your property. Again, *Property Name* is located in the center of the map.

This map is for the period April and May, 1975.

As we only have data for specific points (ie the exact locations of the rain gauges) this algorithm calculates the most likely rainfall experienced between each point.

One application for this map is that you can bring up rainfall patterns of past years or seasons to compare to current predictions. That is, if the current year is predicted to have high rainfall, you can look at the patterns for previous wet years to estimate what areas may receive more rain than others, and vice-versa for dry years.

16 Appendix 7, Spatial data for chapter 6

16.1 Dimberoy data set

Site data, latitude, longitude and UTM coordinates

project id	P mm	Y – latitude decimal degrees	X - longitude decimal degrees	Y – km, UTM[32] zone 55	X – km, UTM zone 55
A00053	77.500	-31.2430	149.9010	6539.837	776.2972
A00055	66.040	-31.4044	149.7110	6522.400	757.7541
A00057	77.216	-31.1440	150.0490	6550.436	790.7020
B053082	94.600	-30.7973	149.5526	6590.075	744.2393
B055002	69.600	-31.2342	149.8345	6540.978	769.9862
B055006	62.800	-31.6405	150.2356	6494.862	806.8800
B055007	99.400	-30.7056	150.0458	6599.063	791.7262
B055014	70.800	-31.1168	150.2682	6552.856	811.7006
B055017	75.100	-31.5711	149.7762	6503.759	763.4872
B055018	87.000	-31.1711	149.6456	6548.421	752.1548
B055023	91.000	-30.9841	150.2540	6567.614	810.7772
B055024	75.200	-31.0261	150.2687	6562.915	812.0448
B055029	69.400	-31.2667	150.1000	6536.693	795.1849
B055034	91.600	-30.7454	150.0557	6594.623	792.5545
B055036	65.000	-31.4135	150.4234	6519.498	825.4889
B055037	70.800	-31.5077	150.3986	6509.123	822.8058
B055038	72.800	-31.0976	149.9114	6555.935	777.7126
B055039	77.000	-31.3946	150.2488	6522.099	808.9419
B055044	117.800	-30.7044	150.2767	6598.572	813.8569
B055045	71.000	-31.1791	150.0312	6546.590	788.8975
B055046	60.200	-31.5271	150.4296	6506.879	825.6844
B055055	83.600	-30.9604	150.4601	6569.648	830.5538
B055064	49.800	-31.6001	150.3811	6498.925	820.8259
B055239	65.200	-31.5879	150.4804	6499.982	830.2965
B055263	102.800	-30.9582	149.6844	6571.939	756.4253
B055264	94.300	-31.5103	150.1900	6509.431	802.9745
B055268	108.600	-30.8325	149.7745	6585.666	765.3820
B055273	98.400	-30.8202	149.8366	6586.881	771.3588
B055274	94.000	-30.7784	150.4359	6589.907	828.8624
B055275	60.400	-31.2874	150.5394	6533.137	836.9744
B055276	86.600	-30.8828	150.4928	6578.158	833.9495

[32] The Universal Transverse Mercator (UTM) grid, a world wide plane coordinate system was developed in the 1940's by the Corps of Engineers, U.S. Army, (Dracup 2006).

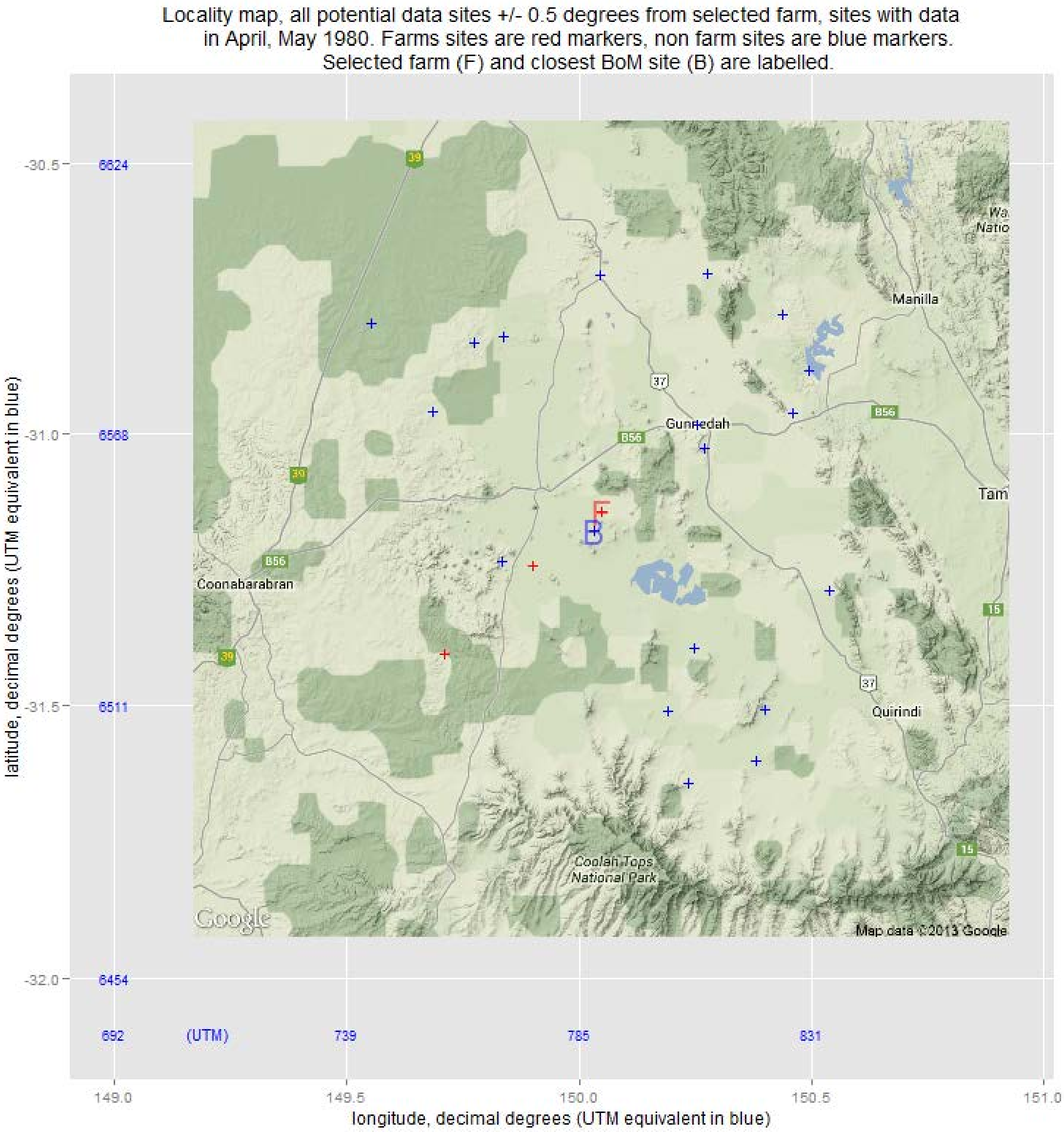
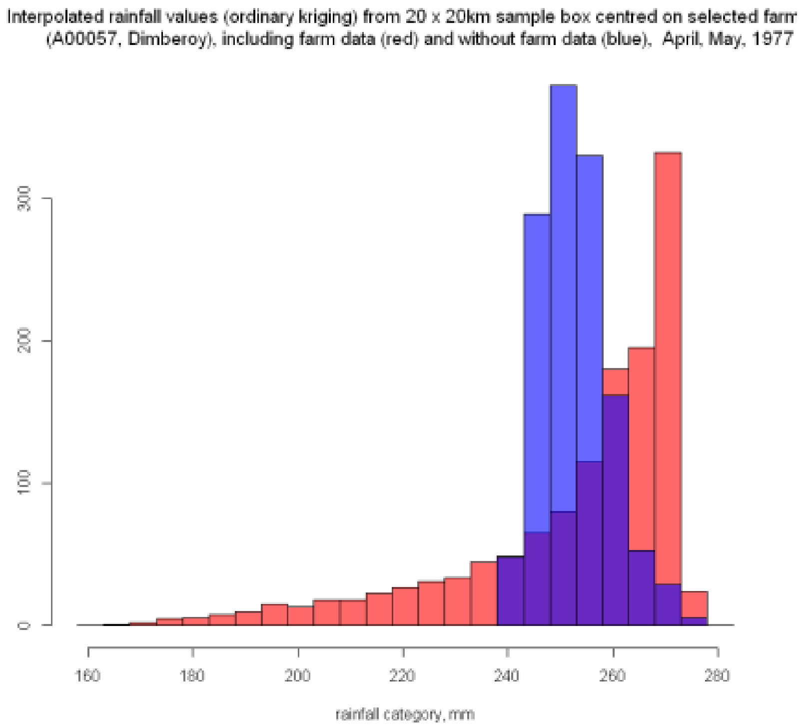


Figure 16-1 Location map for Dimberoy data set showing latitude, longitude and UTM coordinates

Figures were presented in Chapter 6 in a reduced form to allow for comparisons between plots. They are re-presented here at normal scale, along with class intervals and counts for the histograms and rainfall figures for the years and sites with data.



16-2 Histogram of interpolated rainfall values for April and May 1977 (wet), at Dimberoy

Table 16-1 Seasonal (April and May) rainfall totals for Dimberoy, nearby farms and closest BoM site with data in respective years, 1977, 1983, 1988 wet, 1971, 2002, 2005 dry

Site / year	Dimberoy A00057	Elkay A00053	Mount Nombi A00054	Mentone A00055	Womponia A00056	Beulah A00086	closest BoM to Dimberoy 055045
1977	272.5	134.0					248.5
1983	274.8	285.2		285.2			268.6
1988	208.3	175.5		189			197.2
1971	4.8		8.89				5.1
2002	11.0	5.5		20.3	5.5		11.6
2005	63	28.5		58.4	44.5	37	10.8

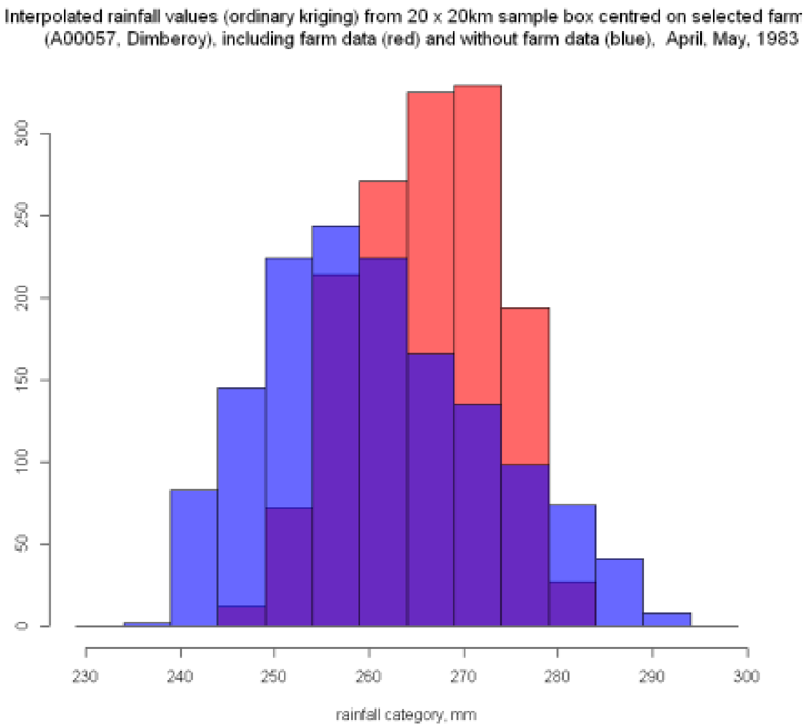


Figure 16-3 Histogram of interpolated rainfall values for April and May 1983 (wet), at Dimberoy

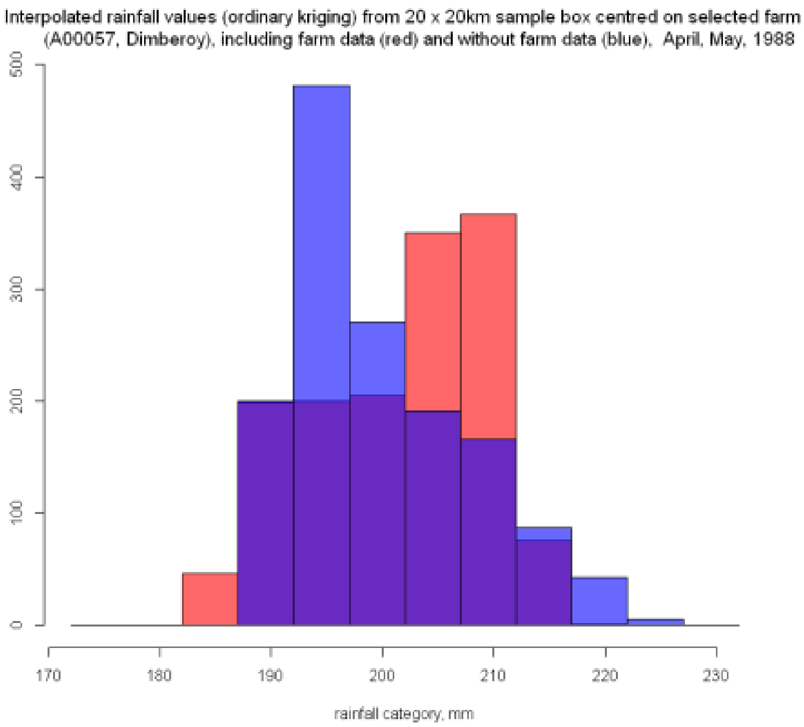


Figure 16-4 Histogram of interpolated rainfall values for April and May 1988 (wet), at Dimberoy

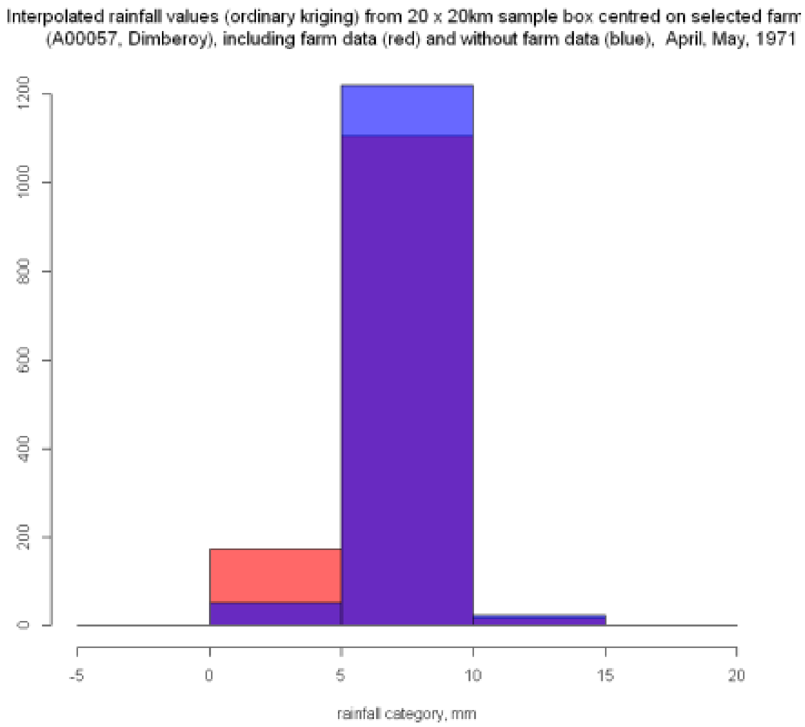


Figure 16-5 Histogram of interpolated rainfall values for April and May 1971 (dry), at Dimberoy

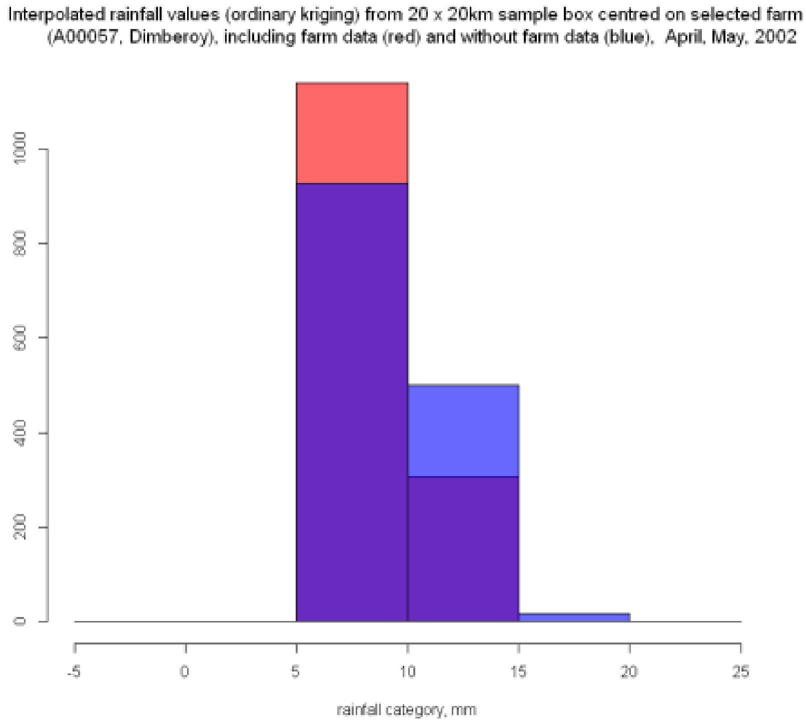


Figure 16-6 Histogram of interpolated rainfall values for April and May 2002(dry), at Dimberoy

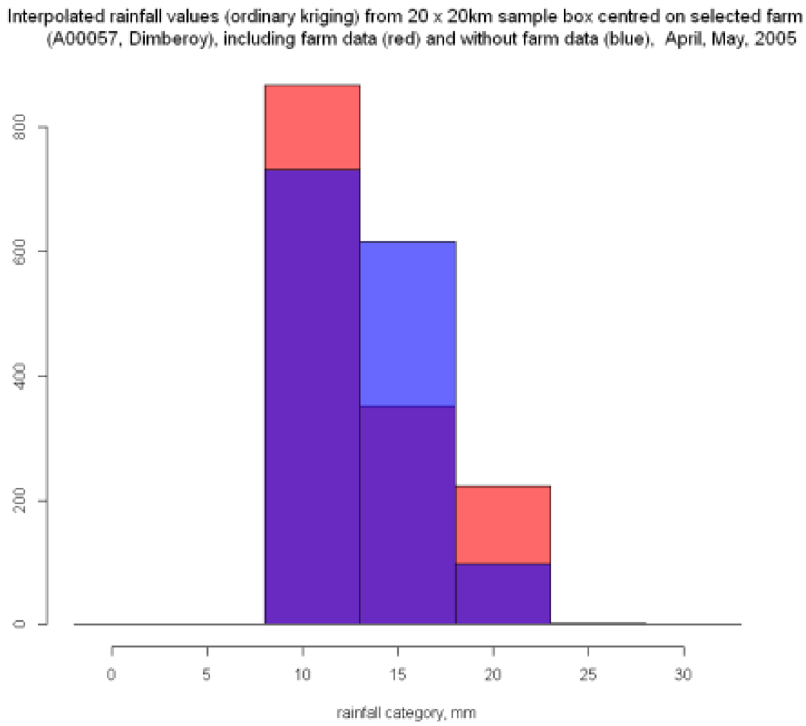


Figure 16-7 Histogram of interpolated rainfall values for April and May 2005 (dry), at Dimberoy

Table 16-2 Histogram classes and counts, interpolated rainfall near Dimberoy for years used in histograms above

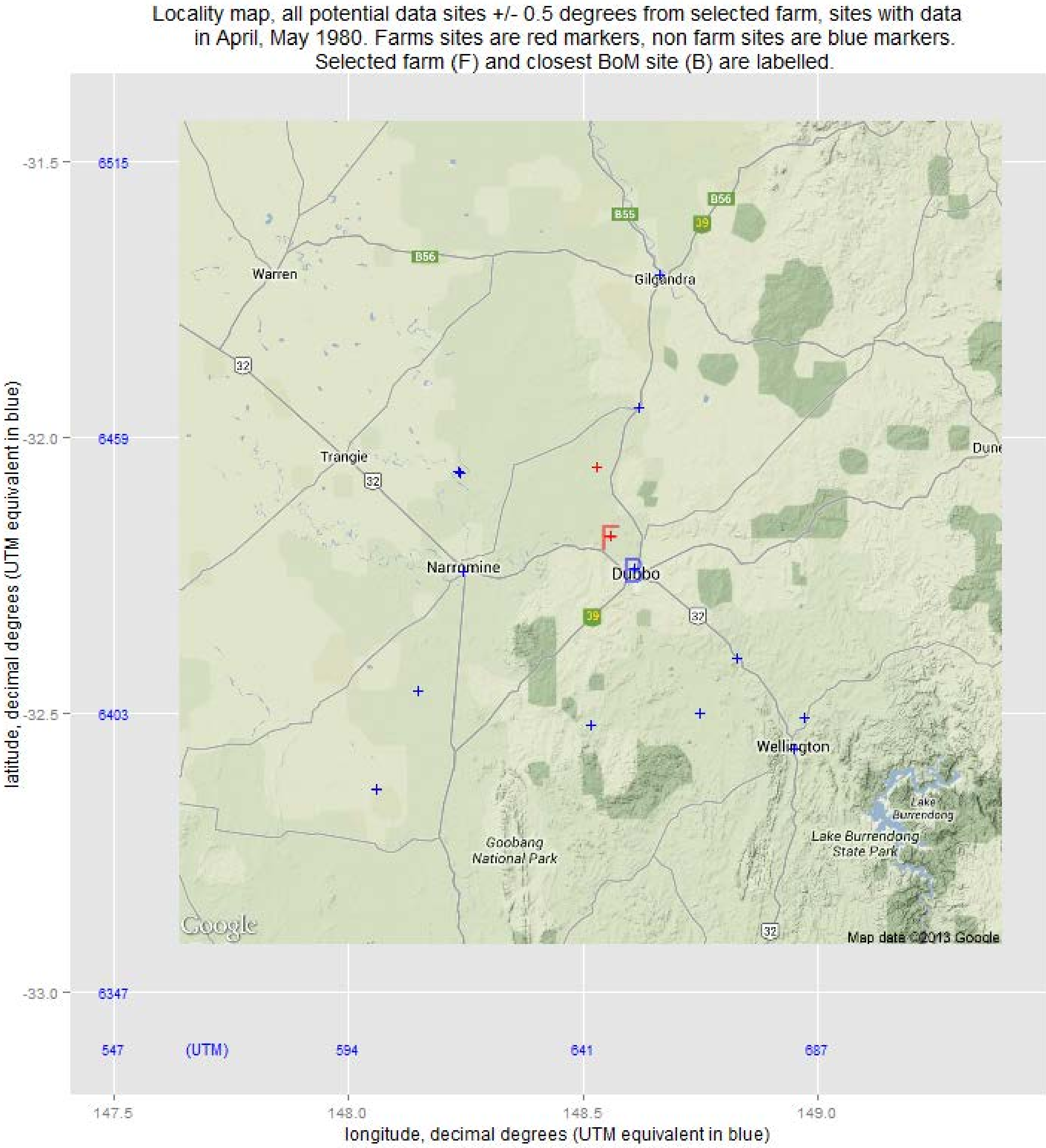
wet years

1977			1983			1988		
range, mm	count all data	count nonfarm	range, mm	count all data	count nonfarm	range, mm	count all data	count nonfarm
158-163	0	0	229-234	0	0	172-177	0	0
163-168	1	0	234-239	0	2	177-182	0	0
168-173	2	0	239-244	0	83	182-187	46	0
173-178	5	0	244-249	12	145	187-192	199	200
178-183	6	0	249-254	72	224	192-197	200	481
183-188	8	0	254-259	214	244	197-202	205	270
188-193	10	0	269-264	271	224	202-207	350	191
193-198	15	0	264-269	325	166	207-213	367	166
198-203	13	0	269-274	329	135	213-217	76	87
203-208	18	0	274-279	194	98	217-222	1	43
208-213	18	0	279-284	27	74	222-227	0	6
213-218	23	0	284-289	0	41	227-232	0	0
218-223	27	0	289-294	0	8			
223-228	31	0	294-299	0	0			
228-233	34	0						
233-238	45	0						
238-243	49	48						
243-248	65	289						
248-253	80	379						
253-258	115	330						
258-263	180	162						
263-268	195	53						
268-273	332	29						
273-278	24	6						
278-283	0	0						

dry years

1971			2002			2005		
range, mm	count all data	count nonfarm	range, mm	count all data	count nonfarm	range, mm	count all data	count nonfarm
0-5	172	52	0-5	0	0	3-8	0	0
5-10	1107	1220	5-10	1138	927	8-13	867	731
10-15	17	24	10-15	306	500	13-18	351	615
15-20	0	0	15-20	0	17	18-23	223	98
			20-25	0	0	23-28	3	0
						28-33	0	0

16.2 Coreena Park Data set



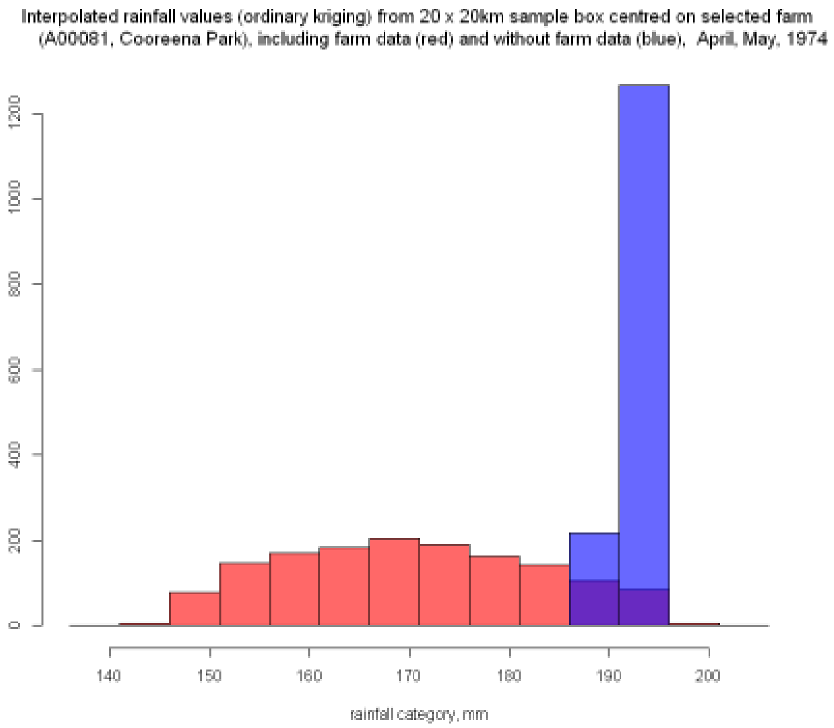


Figure 16-8 Histogram of interpolated rainfall values for April and May 1974 (wet), at Cooreena Park

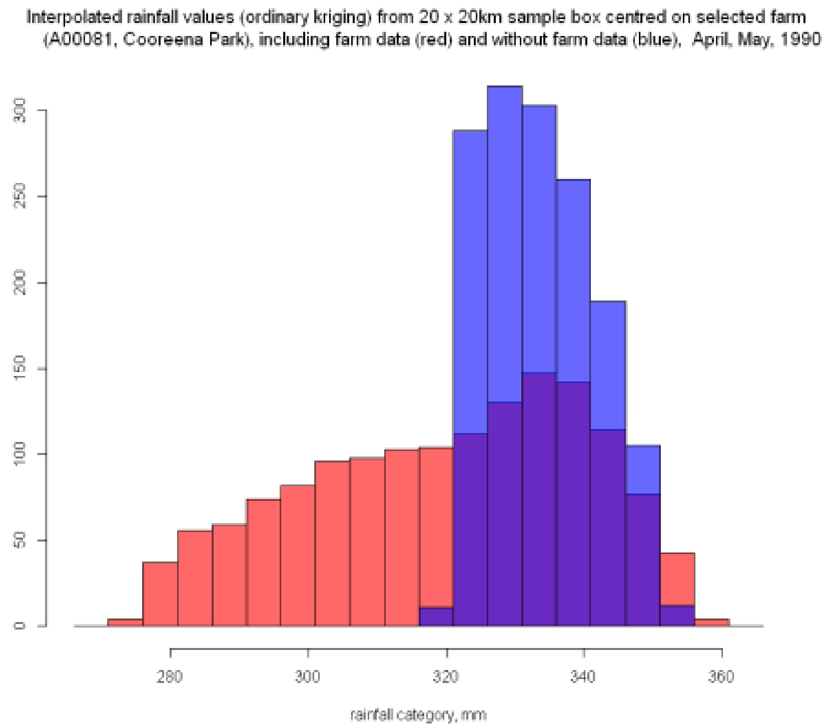


Figure 16-9 Histogram of interpolated rainfall values for April and May 1990 (wet), at Cooreena Park

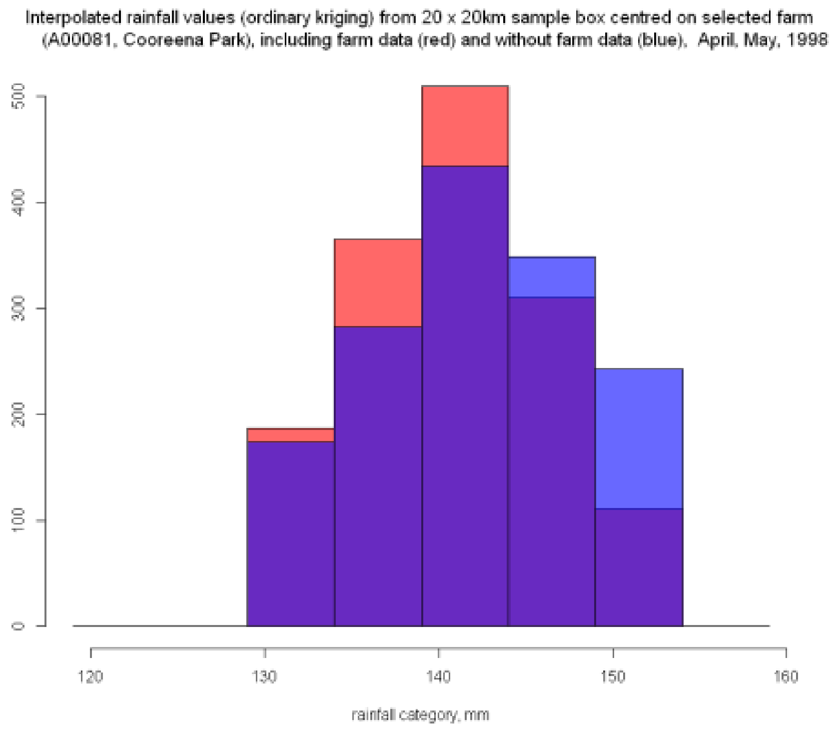


Figure 16-10 Histogram of interpolated rainfall values for April and May 1998 (wet), at Cooreena Park

Table 16-3 Seasonal (April and May) rainfall totals for Cooreena Park, nearby farms and closest BoM site with data in respective years, 1974, 1990 and 1998, wet

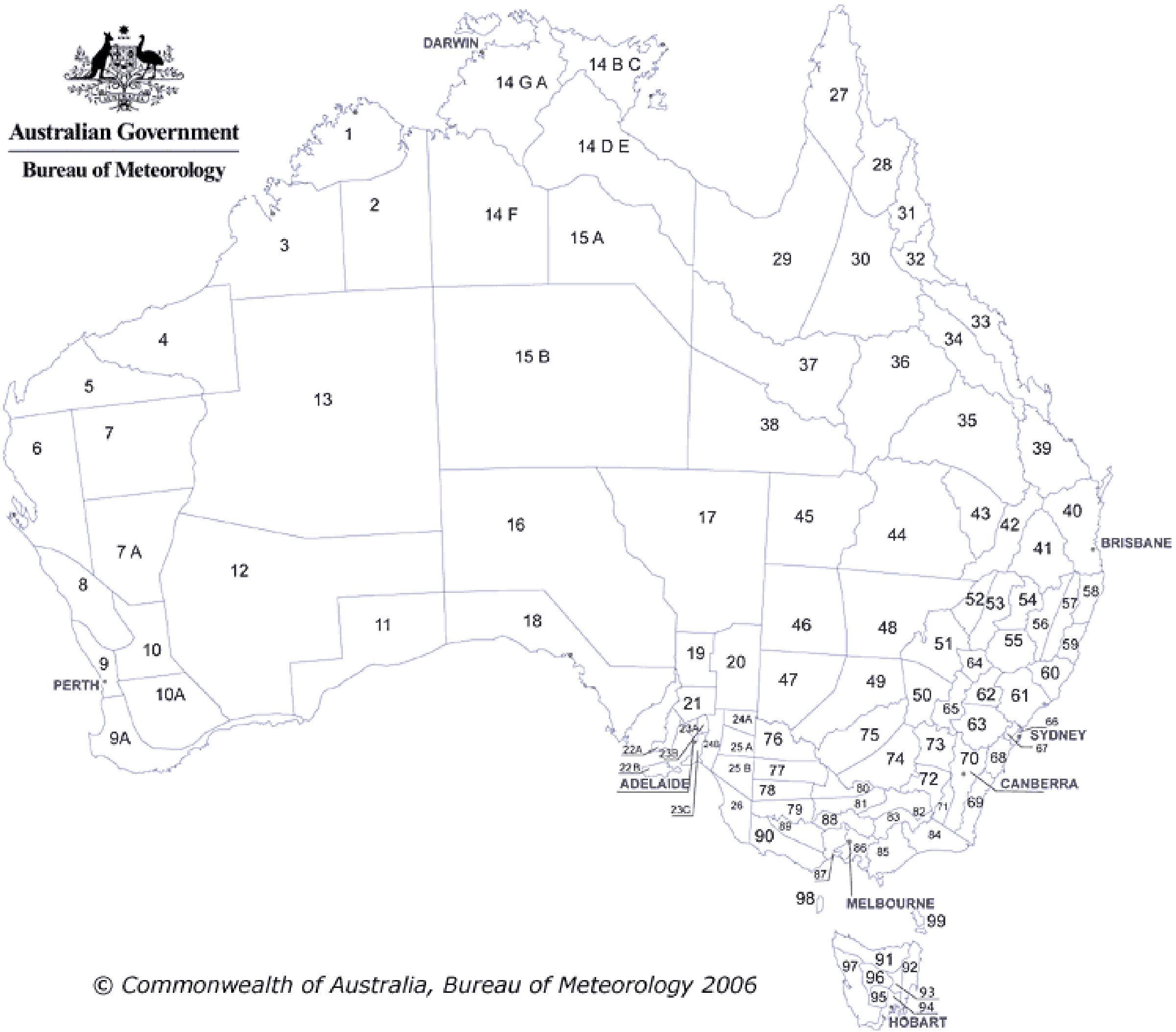
site / year	Cooreena Park A00081	Ashlee A00085	closest BoM to Cooreena Park 065012
1974	165	138.4	192.8
1990	329	255.3	345.7
1998	135	154.4	153.4

Table 16-4 Histogram classes and counts, interpolated rainfall near Coreenal Park for years used in histograms above

wet years

1974			1990			1998		
range, mm	count all data	count nonfarm	range, mm	count all data	count nonfarm	range, mm	count all data	count nonfarm
136-141	0	0	266-271	0	0	119-124	0	0
141-146	5	0	271-277	4	0	124-129	0	0
146-151	79	0	277-281	37	0	129-134	187	174
151-156	147	0	281-286	56	0	134-139	365	282
156-161	171	0	286-291	69	0	139-144	509	434
161-166	184	0	291-296	74	0	144-149	310	349
166-171	204	0	296-301	82	0	149-154	111	243
171-176	190	0	301-306	96	0	154-159	0	0
176-181	162	0	306-311	98	0			
181-186	143	0	311-316	103	0			
186-191	106	217	316-321	104	11			
191-196	86	1265	321-326	112	288			
196-201	5	0	326-331	130	314			
201-206	0	0	331-336	147	303			
			336-341	142	260			
			341-346	114	189			
			346-351	77	105			
			351-356	43	12			
			356-361	4	0			
			361-366	0	0			

17 Appendix 8, Bureau of Meteorology rainfall districts



Source : http://www.bom.gov.au/climate/cdo/about/rain-districts.shtml

Western Australia
1. North Kimberley
2. East Kimberley
3. West Kimberley
4. De Grey
5. Fortescue
6. West Gascoyne
7. East Gascoyne
7A. Murchison
8. North Coast
9. Central Coast
9A. South Coast
10. North Central
10A. South Central
11. Eucla
12. Southeast
13. Northeast

Northern Territory
14BC. Arnhem
14DE. Roper-McArthur
14F. Victoria
14GA. Darwin-Daly
15A. Barkly
15B. Alice Springs

South Australia
16. Northwest
17. Far North
18. Western Agricultural
19. Upper North
20. Northeast
21. Lower North

South Australia (cont.)
22. West Central (22A + 22B)
22A. Yorke Peninsula
22B. Kangaroo Island
23.East Central (23A + 23B + 23C)
23A. Adelaide Plains
23B. County Light
23C. Mt. Lofty Ranges
24. Murray River (24A + 24B)
24A. Upper Murray Valley
24B. Lower Murray Valley
25A. Murray Mallee
25B. Upper Southeast
26. Lower Southeast

Queensland
27. North Peninsula
28. South Peninsula
29. Lower Carpentaria
30. Upper Carpentaria
31. Barron North Coast
32. Herbert North Coast
33. East Central Coast
34. West Central Coast
35. Central Highlands
36. Central Lowlands
37. Upper Western
38. Lower Western
39. Port Curtis
40. Moreton
41. East Darling Downs
42. West Darling Downs
43. Maranoa
44. Warrego
45. Far Southwest

New South Wales
46. Western (Far Northwest)
47. Western (Lower Darling)
48. Western (Upper Darling)
49. Western (Southwest Plains)
50. Central Western Plains (S)
51. Central Western Plains (N)
52. Northwest Plains (W)
53. Northwest Plains (E)
54. Northwest Slopes (N)
55. Northwest Slopes (S)
56. Northern Tablelands (W)
57. Northern Tablelands (E)
58. Upper North Coast
59. Lower North Coast
60. Manning
61. Hunter
62. Central Tablelands (N)
63. Central Tablelands (S)
64. Central Western Slopes (N)
65. Central Western Slopes (S)
66. Metropolitan (E)
67. Metropolitan (W)
68. Illawarra
69. South Coast
70. Southern Tablelands
(Goulburn-Monaro)
71. Southern Tablelands
(Snowy Mountains)
72. Southwest Slopes (S)
73. Southwest Slopes (N)
74. Riverina (E)
75. Riverina (W)

Victoria
76. North Mallee
77. South Mallee
78. North Wimmera
79. South Wimmera
80. Lower North
81. Upper North
82. Lower Northeast
83. Upper Northeast
84. East Gippsland
85. West Gippsland
86. East Central
87. West Central
88. North Central
89. Western Plains
90. West Coast

Tasmania
91. Northern
92. East Coast
93. Midlands
94. Southeast
95. Derwent Valley
96. Central Plateau
97. West Coast
(Mountain Region)
98. King Island
99. Flinders Island

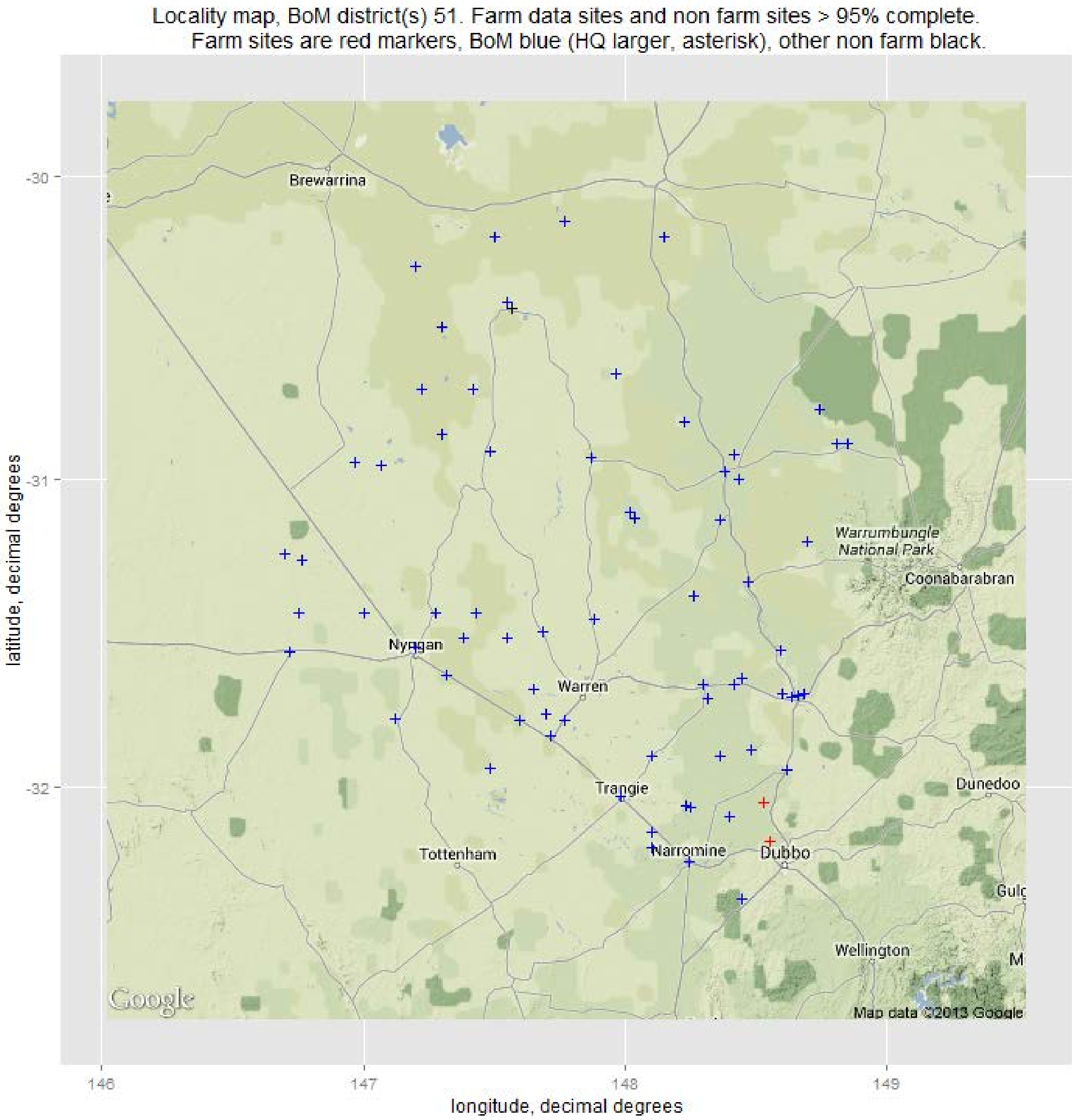


Figure 17-1 Locality map for BoM district 51

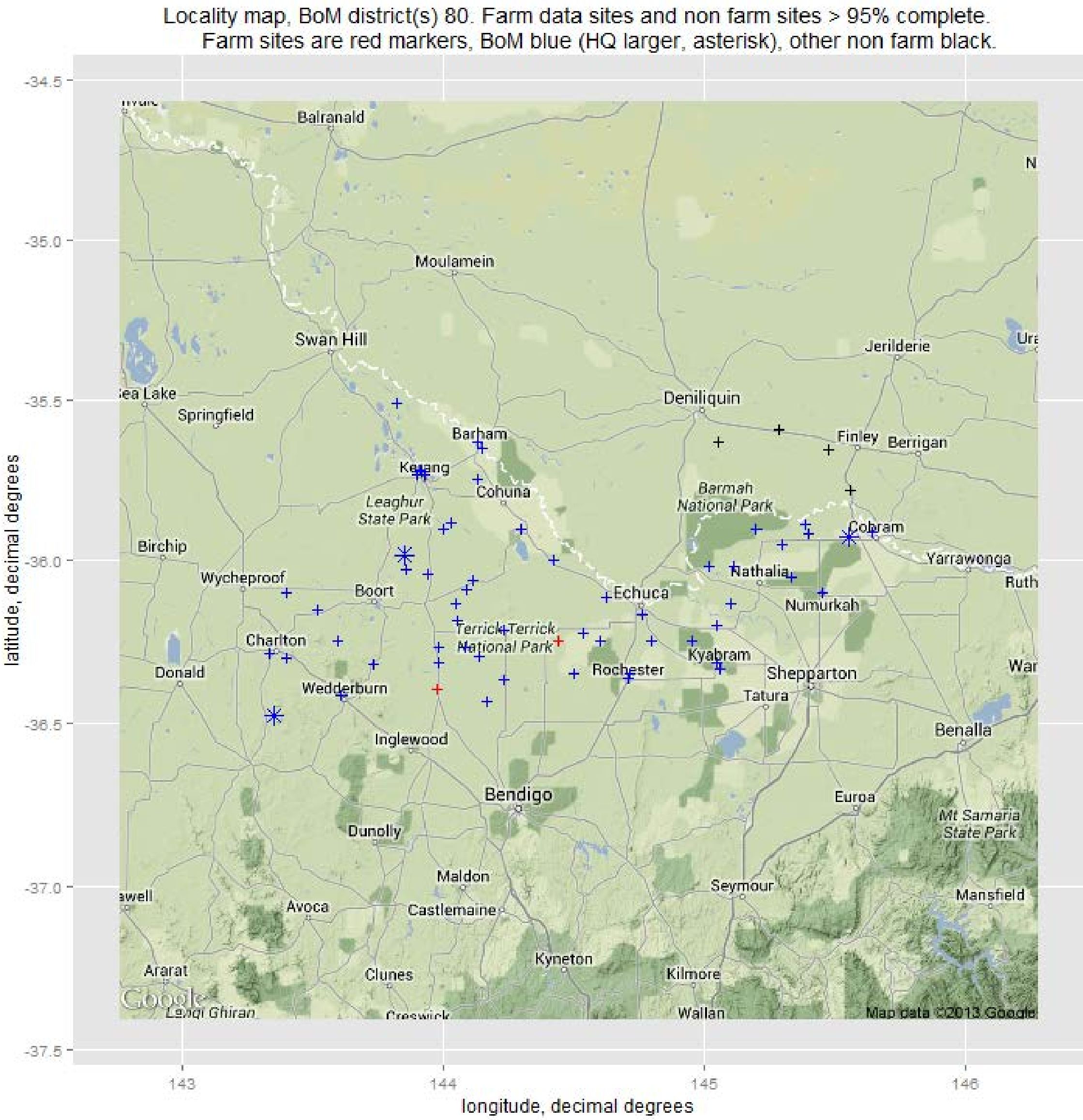


Figure 17-2 Locality map for BoM district 80

18 References

ABARE (2012). Australian Crop Report.

ABS (2012). 2012 Year Book Australian

Adderley, E. E. and E. G. Bowen (1962). "Lunar Component in Precipitation Data." Science **137**(3532): 749-&.

Allen, R. G., A. J. Clemmens, et al. (2005). "Prediction Accuracy for Projectwide Evapotranspiration Using Crop Coefficients and Reference Evapotranspiration." Journal of Irrigation and Drainage Engineering **131**(1): 24-36.

Allen, R. G., L. S. Pereira, et al. (2005). "FAO-56 Dual Crop Coefficient Method for Estimating Evaporation from Soil and Application Extensions." Journal of Irrigation and Drainage Engineering **131**(1): 2-13.

Allen, R. G., W. O. Pruitt, et al. (2005). "Estimating Evaporation from Bare Soil and the Crop Coefficient for the Initial Period Using Common Soils Information." Journal of Irrigation and Drainage Engineering **131**(1): 14-23.

Augustine, D. J. (2010). "Spatial versus temporal variation in precipitation in a semiarid ecosystem." Landscape Ecology **25**(6): 913-925.

Baier, W. (1973). "Crop-Weather Analysis Model: Review and Model Development." Journal of Applied Meteorology **12**(6): 937-947.

Baier, W. and G. Robertson (1966). "A NEW VERSATILE SOIL MOISTURE BUDGET." Canadian Journal of Plant Science **46**(3): 299-&.

Baier, W. and G. W. Robertson (1968). "The performance of soil moisture estimates as compared with the direct use of climatological data for estimating crop yields." Agricultural Meteorology **5**(1): 17-31.

Beecham, S. and R. K. Chowdhury (2010). "Temporal characteristics and variability of point rainfall: a statistical and wavelet analysis." International Journal of Climatology **30**(3): 458-473.

Bjerknes, J. (1969). "Atmospheric teleconnections from the equatorial Pacific." Monthly Weather Review **97**(3): 10.

Boer, R., D. Fletcher, et al. (1993). "Rainfall Patterns in a Major Wheat-Growing Region of Australia." Aust. J. Agric. Res **44**: 16.

BoM (2012). Record-breaking La Niña events, An analysis of the La Niña life cycle and the impacts and significance

of the 2010–11 and 2011–12 La Niña events in Australia. Melbourne, Bureau of Meteorology.

Boyd, W., N. Goodchild, et al. (1976). "An analysis of climatic environments for plant-breeding purposes." Australian Journal of Agricultural Research **27**(1): 19-33.

Bradley, D. A., G. W. Brier, et al. (1962). "Lunar Synodical Period and Widespread Precipitation." Science **137**(3532): 748-&.

Burt, C. M., A. J. Mutziger, et al. (2005). "Evaporation Research: Review and Interpretation." Journal of Irrigation and Drainage Engineering **131**(1): 37-58.

Cai, W. and T. Cowan (2008). "Dynamics of late autumn rainfall reduction over southeastern Australia." Geophysical Research Letters **35**(9): L09708.

Cane, M. A., S. E. Zebiak, et al. (1986). "EXPERIMENTAL FORECASTS OF EL-NINO." Nature **321**(6073): 827-832.

Chiew, F. H. S. and M. J. Leahy (2003). "Inter-decadal pacific oscillation modulation of the impact of El Nino/Southern Oscillation on Australian rainfall and streamflow." Modsim 2003: International Congress on Modelling and Simulation, Vols 1-4: 100-105.

Choudhury, B. J., N. U. Ahmed, et al. (1994). "Relations between evaporation coefficients and vegetation indices studied by model simulations." Remote Sensing of Environment **50**(1): 1-17.

Chowdhury, R. K. and S. Beecham (2010). "Australian rainfall trends and their relation to the southern oscillation index." Hydrological Processes **24**(4): 504-514.
Cornish, E. A. (1936). "On the secular variation of the rainfall at Adelaide, South Australia." Quarterly Journal of the Royal Meteorological Society **62**: 481-490.
Cornish, E. A. (1949). "Yield Trends in the Wheat Belt of South Australia during 1896-1941." Australian Journal of Scientific Research Series B-Biological Sciences **2**(2): 83-137.
Cornish, E. A. (1950). "The Influence of Rainfall on the Yield of Wheat in South Australia." Australian Journal of Scientific Research Series B-Biological Sciences **3**(2): 178-218.
Davidson, J. (1934). "Climate in Relation to Insect Ecology in Australia 1 Mean monthly Precipitation and Atmospheric Saturation Deficit in Australia." Transactions Roy. Soc. S. Australia **58**: 197-210.
Davidson, J. (1936). "Climate in Relation to Insect Ecology in Australia 3 Bioclimatic Zones in Australia." Transactions Roy. Soc. S. Australia **60**: 88-92.
Diaz, H. F., M. P. Hoerling, et al. (2001). "ENSO variability, teleconnections and climate change." International Journal of Climatology **21**(15): 1845-1862.
Dingman, S. L. (2002). Physical Hydrology. New Jersey USA, Prentice Hall.
Dirks, K. N., J. E. Hay, et al. (1998). "High-resolution studies of rainfall on Norfolk Island Part II: Interpolation of rainfall data." Journal of Hydrology **208**(3-4): 187-193.
Doraiswamy, P. C. and D. R. Thompson (1982). "A crop moisture stress index for large areas and its application in the prediction of spring wheat phenology." Agricultural Meteorology **27**(1-2): 1-15.
Drosdowsky, W. and L. E. Chambers (2001). "Near-global sea surface temperature anomalies as predictors of Australian seasonal rainfall." Journal of Climate **14**(7): 1677-1687.
Dunsdorfs, E. (1956). "Historical Statistics of the Autralian wheatgrowing industry.": 94.
Dunsdorfs, E. (1957). The Australian Wheat Growing Industry, 1788-1948. Melbourne, Melbourne University Press.
FAO (2002). Bread Wheat. FAO Plant Production and Protection Series. Rome, FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS.
FAO (2009). Food Outlook. The Global Information and Early Warning System on Food and Agriculture. F. a. A. O. o. t. U. Nations. Rome, Food and Agriculture Organization of the United Nations**:** 92.
Fisher, R. A. (1925). "The Influence of Rainfall on the Yield of Wheat at Rothamsted." Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character **213**(402-410): 89-142.
Fitzpatrick, E. A. (1963). "Estimates of Pan Evaporation from Mean Maximum Temperature and Vapor Pressure." Journal of Applied Meteorology **2**(6): 780-792.
Fitzpatrick, E. A. and H. A. Nix (1969). "A Model for Simulating Soil Water Regime in Alternating Fallow-Crop Systems." Agricultural Meteorology **6**(5): 303-&.
Fitzpatrick, E. A. and H. A. Nix (1970). The Climatic Factor in Australian Grassland Ecology. Australian Grasslands. R. Miltonn Moore. Canberra, Australian National University Press**:** 30-026.
Fitzpatrick, E. A., R. O. Slatyer, et al. (1967). "Incidence and duration of periods of plant growth in Central Australia as estimated from climatic data." Agricultural Meteorology **4**(6): 389-404.
French, R. and J. Schultz (1984). "Water use efficiency of wheat in a Mediterranean-type environment. I. The relation between yield, water use and climate." Australian Journal of Agricultural Research **35**(6): 743-764.
Gallant, A. J. E., K. J. Hennessy, et al. (2007). "Trends in rainfall indices for six Australian regions: 1910-2005." Australian Meteorological Magazine **56**(4): 223-239.
Gangopadhyaya, M. and R. P. Sarker (1965). "Influence of rainfall distribution on the yield of wheat crop." Agricultural Meteorology **2**(5): 331-350.
GGA (2012). Grain Growewrs Annual report 2012.

Gipps, G. (1846). Historical records of Australia. Governors' despatches to and from England.
Goovaerts, P. (2000). "Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall." Journal of Hydrology **228**(1-2): 113-129.
Goyder, G. W. (1857). "Report on the Country between Mount Serle and Lake Torrens, South Australia." Proceedings of the Royal Geographical Society of London **2**(1): 16-30.
Hammer, G. L., J. W. Hansen, et al. (2001). "Advances in application of climate prediction in agriculture." Agricultural Systems **70**(2-3): 515-553.
Hammer, G. L., D. P. Holzworth, et al. (1996). "The value of skill in seasonal climate forecasting to wheat crop management in a region with high climatic variability." Australian Journal of Agricultural Research **47**(5): 717-737.
Hansen, J. W., A. Potgieter, et al. (2004). "Using a general circulation model to forecast regional wheat yields in northeast Australia." Agricultural and Forest Meteorology **127**(1-2): 77-92.
Hayman, P., J. Crean, et al. (2007). "How do probabilistic seasonal climate forecasts compare with other innovations that Australian farmers are encouraged to adopt?" Australian Journal of Agricultural Research **58**(10): 975-984.
Hooker, R. H. (1921). "Forecasting the crops from the weather." Quarterly Journal of the Royal Meteorological Society **47**: 75-99.
Hounam, C. E. (1947). Studies in Applied Climatology Western Australian Pamphlet No. 1

Climate of the West Australian Wheat Belt with special reference to Rainfall over Marginal Areas. C. o. A. Commonwealth Meteorological Bureau. Melbourne. **1**.
Hounam, C. E. (1947a). Studies in Applied Climatology New South Wales Pamphlet No. 2

Climate of the South West Wheat Belt of N.S.W. with special reference to Rainfall over Marginal Areas. C. o. A. Commonwealth Meteorological Bureau. Melbourne, Government Printer. **2**.
Hounam, C. E. (1950). Studies in Applied Climatology

Climate of the North-West Wheat Belt of N.S.W. with special reference to Rainfall over Marginal Areas. C. o. A. Commonwealth Meteorological Bureau. **4**.
Hutchinson, M. F. and R. J. Bischof (1983). "A new method for estim ating the spatial distribution of mean seasonal and annual rainfall applied to the Hunter Valley, New South Wales." Australian Meteorological Magazine **31**(3): 5.
IPCC (2007). Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. R. K. Pachauri and A. Reisinger. Geneva, Switzerland: 104.
Johnson, D. H. (1999). "The insignificance of statistical significance testing." Journal of Wildlife Management **63**(3): 763-772.
Jones, D. A., W. Wang, et al. (2009). "High-quality spatial climate data-sets for Australia." Australian Meteorological and Oceanographic Journal **58**(4): 233-248.
Jones, J. W., J. W. Hansen, et al. (2000). "Potential benefits of climate forecasting to agriculture." Agriculture Ecosystems & Environment **82**(1-3): 169-184.
Kamruzzaman, M., S. Beecham, et al. (2013). "Climatic influences on rainfall and runoff variability in the southeast region of the Murray-Darling Basin." International Journal of Climatology **33**(2): 291-311.
Kane, R. P. (1997). "On the relationship of ENSO with rainfall over different parts of Australia." Australian Meteorological Magazine **46**(1): 39-49.
Knisel, W. G. (1980). CREAMS - A field-scale model for chemicals, runoff, and, erosion from agricultural management systems, US Department of Agriculture, Science and Education: 643.
Knisel, W. G. (1982). "Systems for Evaluating Non-Point Source Pollution - an Overview." Mathematics and Computers in Simulation **24**(2-3): 173-184.

Krejcie, R. V. and D. W. Morgan (1970). "Determining sample sizes for research activities." Educational and Psychological Measurement **30**: 4.
Lazenby, A. and E. M. Matheson (1975). Australian Field Crops: Wheat and other temperate cereals, Angus and Robertson.
Littleboy, M., D. Silburn, et al. (1992). "Impact of soil erosion on production in cropping systems .I. Development and validation of a simulation model." Soil Research **30**(5): 757-774.
Maidment, D. R. (1993). Handbook of Hydrology, McGraw Hill.
Marsaglia, G., W. W. Tsang, et al. (2003). "Evaluating Kolmogorov's Distribution." Journal of Statistical Software **08, issue i18** (i18).
McBride, J. L. and N. Nicholls (1983). "Seasonal Relationships between Australian Rainfall and the Southern Oscillation." Monthly Weather Review **111**(10): 1998-2004.
McCown, R. L., G. L. Hammer, et al. (1996). "APSIM: A novel software system for model development, model testing and simulation in agricultural systems research." Agricultural Systems **50**(3): 255-271.
McCown, R. L. and J. Williams (1989). AUSIM: A cropping systems model for operational research. SSA IMACS 1989 Biennial Conference on Modelling and Simulation. Australian National University, Canberra, 25-27th September.
McKee, T. B. (1993). "The relationship of drought frequency and duration to time scales." Proceedings, Eighth Conference on Applied Climatology: 179 184.
McMahon, T. A. (1983). A General Wheat Crop Model for Australia. Agricultural Engineering Report No. 67/83. Melbourne, Department of Agricultural Engineering, University of Melbourne. **67:** 52.
Meinig, D. W. (1961). "Goyder's Line of Rainfall: The Role of a Geographic Concept in South Australian Land Policy and Agricultural Settlement." Agricultural History **35**(4): 207-214.
Millington, R. (1961). "Relations between yield of wheat, soil factors, and rainfall." Australian Journal of Agricultural Research **12**(3): 397-408.
Montazerolghaem, M., W. Vervoort, et al. (2012). Spatiotemporal fuzzy based climate classification for Southeastern Australia. 34th Hydrology and Water Resources Symposium. Sydney.
Nicholls, N. (1985). "IMPACT OF THE SOUTHERN OSCILLATION ON AUSTRALIAN CROPS." Journal of Climatology **5**(5): 553-560.
Nicholls, N. and B. Lavery (1992). "Australian rainfall trends during the twentieth century." International Journal of Climatology **12**(2): 153-163.
Nix, H. A. (1975). The Australian Climate and its Effects on Grain Yield and Quality, Angus and Robertson.
Nix, H. A. and E. A. Fitzpatrick (1969). "An Index of Crop Water Stress Related to Wheat and Grain Sorghum Yields." Agricultural Meteorology **6**(5): 321-&.
ODG (2004). Oxford Dictionary of Georgraphy, Oxford University Press.
OED (2013). Oxford English Dicitnary. Oxford, Oxford University Press.
Opokuankomah, Y. and I. Cordery (1993). "Temporal Variation of Relations between New-South-Wales Rainfall and the Southern Oscillation." International Journal of Climatology **13**(1): 51-64.
Palmer, W. C. (1965). "Meteorological Drought." United States Weather Bureau, Research Paper No. 45. .
Pan, M., E. F. Wood, et al. (2008). "Estimation of regional terrestrial water cycle using multi-sensor remote sensing observations and data assimilation." Remote Sensing of Environment **112**(4): 1282-1294.
Peters, O. and K. Christensen (2002). "Rain: Relaxations in the sky." Physical Review E **66**(3): 036120.
Philander, S. G. H. (1983). "El-Nino Southern Oscillation Phenomena." Nature **302**(5906): 295-301.
Pittock, A. B. (1975). "Climatic Change and Patterns of Variation in Australian Rainfall." Search **6**(11-1): 498-504.

Plain, M. B., B. Minasny, et al. (2008). "Spatially explicit seasonal forecasting using fuzzy spatiotemporal clustering of long-term daily rainfall and temperature data." Hydrology and Earth System Sciences **5**: 16.
Prescott, J. A. (1934). "Single value climatic factors." Transactions roy. Soc. S. Aust. **58**: 48-61.
Prescott, J. A. (1936). "The climatic control of the Australian deserts." Trans, Roy. Soc. S. Aust. **60**: 2.
Prescott, J. A. (1938). "Indices in agricultural climatology." Journal of the Australian Institute of Agricultural Science **4**: 33-40.
Prescott, J. A. (1946). "A Climatic Index." Nature **157**(3991): 555-555.
Prescott, J. A. (1949). "A Climatic Index for the Leaching Factor in Soil Formation." Journal of Soil Science **1**(1): 9-19.
Radinović, D. and M. Ćurić (2009). "Deficit and surplus of precipitation as a continuous function of time." Theoretical and Applied Climatology **98**(1): 197-200.
Rasmusson, E. M. and J. M. Wallace (1983). "Meteorological Aspects of the El-Nino Southern Oscillation." Science **222**(4629): 1195-1202.
Rawson, C. B. (1874). Report upon the rainfall of Barbados and upon its influence upon the sugar crops, 1847-71,. Barbados.
Robinson, J. B., D. M. Silburn, et al. (2010). "Modelling shows that the high rates of deep drainage in parts of the Goondoola Basin in semi-arid Queensland can be reduced with changes to the farming systems." Australian Journal of Soil Research **48**(1): 58-68.
Ropelewski, C. F. and M. S. Halpert (1987). "Global and Regional Scale Precipitation Patterns Associated with the El-Nino Southern Oscillation." Monthly Weather Review **115**(8): 1606-1626.
Roy, S. S. (2006). "Impact of lunar cycle on the precipitation in India." Geophysical Research Letters **33**(1).
Sadras, V. c., D. Roget, et al. (2002). "Dynamic cropping strategies for risk management in dry-land farming systems." Agricultural Systems **76**(3): 929-948.
Sadras, V. O. and D. Rodriguez (2007). "The limit to wheat water-use efficiency in eastern Australia. II. Influence of rainfall patterns." Australian Journal of Agricultural Research **58**(7): 657-669.
Seif, E. and D. Pederson (1978). "Effect of rainfall on the grain yield of spring wheat, with an application to the analysis of adaptation." Australian Journal of Agricultural Research **29**(6): 1107-1115.
Sheldrick, J. (2010). "1855-1856: George Goyder's Long Ride to Mapping Reliable Rainfall." Globe, The(65): 28-40.
Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. Proceedings of the 1968 23rd ACM national conference, ACM**:** 517-524.
Slafer, G. A. and H. M. Rawson (1994). "Sensitivity of Wheat Phasic Development to Major Environmental-Factors - a Reexamination of Some Assumptions Made by Physiologists and Modelers." Australian Journal of Plant Physiology **21**(4): 393-426.
Slatyer, R. O. (1962). PART III. Climate of the Alice Springs Area. General Report on Lands of the Alice Springs Area, Northern Territory, 1956-57. R. A. Perry. Melbourne, CSIRO**:** 109-128.
SOE (2011). Australia State of the Environment.
SQL (2012). Information technology -- Database languages -- SQL -- Part 1: Framework (SQL/Framework). ISO/IEC 9075-1:2008: Information technology – Database languages – SQL – Part 1: Framework (SQL/Framework). http://www.iso.org/iso/catalogue_detail.htm?csnumber=45498.
Stephens, D. J. and T. J. Lyons (1998). "Rainfall-yield relationships across the Australian wheatbelt." Australian Journal of Agricultural Research **49**(2): 211-224.
Stephens, D. J., G. K. Walker, et al. (1994). "Forecasting Australian wheat yields with a weighted rainfall index." Agricultural and Forest Meteorology **71**(3-4): 247-263.

Stone, R. and A. Auliciems (1992). "SOI phase relationships with rainfall in eastern Australia." International Journal of Climatology **12**(6): 625-636.

Stone, R. C., G. L. Hammer, et al. (1996). "Prediction of global rainfall probabilities using phases of the Southern Oscillation Index." Nature **384**(6606): 252-255.

Suppiah, R. (2004). "Trends in the southern oscillation phenomenon and Australian rainfall and changes in their relationship." International Journal of Climatology **24**(3): 269-290.

Suppiah, R. and K. J. Hennessy (1996). "Trends in the intensity and frequency of heavy rainfall in tropical Australia and links with the Southern Oscillation." Australian Meteorological Magazine **45**(1): 1-17.

Taylor, A. and A. Gilmour (1971). "Wheat protein prediction from climatic factors in southern NSW." Australian Journal of Experimental Agriculture **11**(52): 546-549.

Taylor, A., R. Storrier, et al. (1974). "Nitrogen needs of wheat. 1. Grain yield in relation to soil nitrogen and other factors." Australian Journal of Experimental Agriculture **14**(67): 241-248.

Taylor, T. G. (1915). The climatic control of Australian production (an attempt to gauge the potential wealth of the Commonwealth). Melbourne**:** 36.

Taylor, T. G. (1918). The Australian environment (especially as controlled by rainfall). A. C. o. S. a. Industry. Melbourne, Advisory Council of Science and Industry**:** 98.

Temesgen, B., S. Eching, et al. (2005). "Comparison of Some Reference Evapotranspiration Equations for California." Journal of Irrigation and Drainage Engineering **131**(1): 73-84.

Tench, W. (1793). A complete account of the settlement at Port Jackson.

Thiessen, A. H. (1911). "Precipitation averages for large areas." Monthly Weather Review **39**(7): 1082-1089.

Thornthwaite, C. W. (1948). "An Approach toward a Rational Classification of Climate." Geographical Review **38**(1): 55-94.

Transeau, E. N. (1905). "Forest centers of eastern America." American Naturalist **39**: 875-889.

Troup, A. J. (1965). "Southern Oscillation." Quarterly Journal of the Royal Meteorological Society **91**(390): 490-&.

Trumble, H. C. (1937). "The climatic control of agriculture in South Australia." Transactions Roy. Soc. S. Australia **61**: 41-62.

Trumble, H. C. (1939). "Climatic factors in relation to the agricultural regions of southern Australia." Transactions Roy. Soc. S. Australia **63**: 36-43.

Trumble, H. C. (1952). Grassland Agronomy in Australia. Advances in Agronomy. **4:** 1 - 65.

Venables, W. N. (1990). An Introduction to R. http://www.r-project.org/index.html.

Visher, S. S. (1919). "THE AUSTRALIAN ENVIRONMENT ESPECIALLY AS CONTROLLED BY RAINFALL, by G W Goyger, Review - Summary." Monthly Weather Review.

Walker, A. (1924). "Correlations in seasonal variaitons of weather, IX. A further study of world weather." Memoirs of the India Meteorological Department **24**(9): 59.

Walker, G. T. and E. W. Bliss (1937). "World Weather VI." Memoirs of the Royal Meteorological Society **4**(39): 21.

Wallén, A. (1917). "not available." STOCKHOLIN K SVENSKA **57**

Wang, E. L., P. McIntosh, et al. (2009). "Quantifying the value of historical climate knowledge and climate forecasts using agricultural systems modelling." Climatic Change **96**(1-2): 45-61.

Wark, D. C. (1941). "The variability of the length of the rainfall season and the amount of influential rainfall in South Australia." Trans, Roy. Soc. S. Aust. **65**(2): 6.

Webster, R. and M. A. Oliver (2007). Geostatistics for Environmental Scientists. Chichester, West Sussex PO19 8SQ, England, John Wiley & SOns.

Whetton, P. H. (1988). "A synoptic climatological analysis of rainfall variability in south-eastern Australia." Journal of Climatology **8**(2): 155-177.

White, R. L. C. (1955). Drought and effective rainfall frequency in pastoral New South Wales, west of the wheat belt. Meteorological study. Melbourne, Commonwealth Bureau of Meteorology (Australia). **5**.
Wynen, E. (1984). "Variables Affecting Wheat Yields in New-South-Wales 1945-46 to 1968-69." Journal of the Australian Institute of Agricultural Science **50**(3): 167-172.
Yuan, G. F., Y. Luo, et al. (2004). "Evaluation of a crop water stress index for detecting water stress in winter wheat in the North China Plain." Agricultural Water Management **64**(1): 29-40.
Zhang, X.-G. and T. M. Casey (1992). "Long-term variations in the Southern Oscillation and relationships with Australian rainfall." Australian Meteorological Magazine **40**(4): 14.