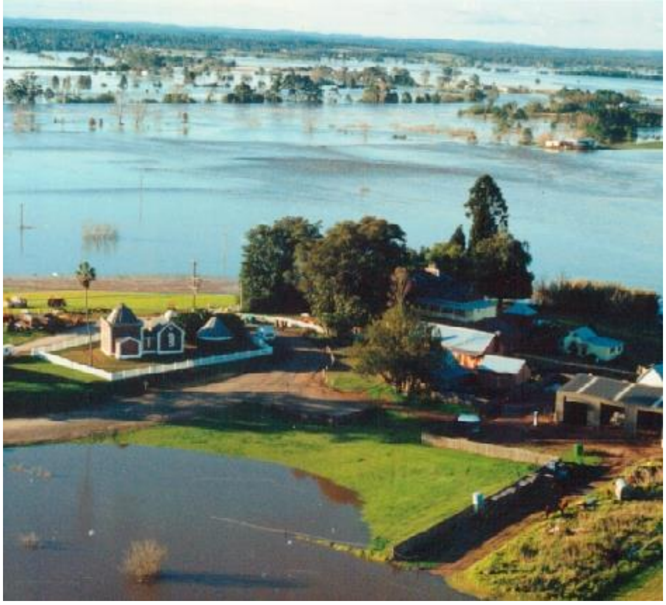


Australian Rainfall & Runoff

Revision Projects

REVISION PROJECT 2

SPATIAL PATTERNS OF RAINFALL

Short Duration Areal Reduction Factors

STAGE 2 REPORT

P2/S2/019

FEBRUARY 2014

Australian Government

ENGINEERS AUSTRALIA
Water Engineering

AUSTRALIAN RAINFALL AND RUNOFF
REVISION PROJECT 2: SPATIAL PATTERNS OF RAINFALL

SHORT DURATION AREAL REDUCTION FACTORS

STAGE 2 REPORT

FEBRUARY, 2014

Project Project 2: Spatial Patterns of Rainfall	**AR&R Report Number** P2/S2/019
Date 6 February 2014	**ISBN** 978-085825-9614
Contractor WMAwater Pty Ltd	**Contractor Reference Number** 28060-10
Authors Peter Stensmyr, Mark Babister, Monique Retallick	**Verified by** James Ball

COPYRIGHT NOTICE

This document, Project 2: Spatial Patterns of Rainfall 2014, is licensed under the Creative Commons Attribution 4.0 Licence, unless otherwise indicated.

Please give attribution to: © Commonwealth of Australia (Geoscience Australia) 2014
We also request that you observe and retain any notices that may accompany this material as part of the attribution.

Notice Identifying Other Material and/or Rights in this Publication:
The authors of this document have taken steps to both identify third-party material and secure permission for its reproduction and reuse. However, please note that where these third-party materials are not licensed under a Creative Commons licence, or similar terms of use, you should obtain permission from the rights holder to reuse their material beyond the ways you are permitted to use them under the 'fair dealing' provisions in the *Copyright Act 1968*.

Further Information
For further information about the copyright in this document, please contact:
Intellectual Property and Copyright Manager
Corporate Branch
Geoscience Australia
GPO Box 378
CANBERRA ACT 2601
Phone: +61 2 6249 9367 or email: copyright@ga.gov.au

DISCLAIMER

The Creative Commons Attribution 4.0 Licence contains a Disclaimer of Warranties and Limitation of Liability.

ACKNOWLEDGEMENTS

This project was made possible by funding from the Federal Government through the Department of Climate Change and Energy Efficiency. This report and the associated project are the result of a significant amount of in kind hours provided by Engineers Australia Members.

ENGINEERS
AUSTRALIA
Water Engineering

Contractor Details

WMAwater
Level 2, 160 Clarence Street
Sydney, NSW, 2000

Tel: (02) 9299 2855
Fax: (02) 9262 6208
Email: wma@wmawater.com.au
Web: www.wmawater.com.au

FOREWORD

AR&R Revision Process

Since its first publication in 1958, Australian Rainfall and Runoff (ARR) has remained one of the most influential and widely used guidelines published by Engineers Australia (EA). The current edition, published in 1987, retained the same level of national and international acclaim as its predecessors.

With nationwide applicability, balancing the varied climates of Australia, the information and the approaches presented in Australian Rainfall and Runoff are essential for policy decisions and projects involving:

- infrastructure such as roads, rail, airports, bridges, dams, stormwater and sewer systems;
- town planning;
- mining;
- developing flood management plans for urban and rural communities;
- flood warnings and flood emergency management;
- operation of regulated river systems; and
- prediction of extreme flood levels.

However, many of the practices recommended in the 1987 edition of AR&R now are becoming outdated, and no longer represent the accepted views of professionals, both in terms of technique and approach to water management. This fact, coupled with greater understanding of climate and climatic influences makes the securing of current and complete rainfall and streamflow data and expansion of focus from flood events to the full spectrum of flows and rainfall events, crucial to maintaining an adequate knowledge of the processes that govern Australian rainfall and streamflow in the broadest sense, allowing better management, policy and planning decisions to be made.

One of the major responsibilities of the National Committee on Water Engineering of Engineers Australia is the periodic revision of ARR. A recent and significant development has been that the revision of ARR has been identified as a priority in the Council of Australian Governments endorsed National Adaptation Framework for Climate Change.

The update will be completed in three stages. Twenty one revision projects have been identified and will be undertaken with the aim of filling knowledge gaps. Of these 21 projects, ten projects commenced in Stage 1 and an additional 9 projects commenced in Stage 2. The remaining two projects will commence in Stage 3. The outcomes of the projects will assist the ARR Editorial Team with the compiling and writing of chapters in the revised ARR.

Steering and Technical Committees have been established to assist the ARR Editorial Team in guiding the projects to achieve desired outcomes. Funding for Stages 1 and 2 of the ARR revision projects has been provided by the Federal Department of Climate Change and Energy Efficiency. Funding for Stages 2 and 3 of Project 1 (Development of Intensity-Frequency-Duration information across Australia) has been provided by the Bureau of Meteorology.

Project 2 Spatial Patterns of Design Rainfall

Rainfall estimates from IFD relationships are applicable strictly only to a single point and not to a wider area such as a catchment. However, where the catchment area is small, the point IFD relationships is taken to be representative of the areal IFD relationship. For this purpose, a small catchment would be defined as being less than 4km^2. For larger areas, it is unrealistic to assume that the same intensity rainfall will occur over the entire area and reductions in the rainfall intensity are made. Unfortunately, at present there is limited information available regarding values for the Areal Reduction Factor (ARF). Due to this, Canterford et al. (1987) recommended the use of areal reduction factors developed by the US National Weather Service (NOAA, 1980) for the Chicago region as being appropriate for all regions of Australia except for the inland regions where the recommendation is to use areal reduction factors developed for Arizona (NOAA, 1984). Hence the ARFs currently recommended for use in Australia were not defined from Australian data. Since that recommendation, Srikanthan (1995), Siriwardena and Weinmann, (1996) and Catchlove and Ball (2003) have investigated ARFs using Australian data for different regions of the country. These studies have highlighted the inappropriateness of American ARFs for Australian conditions.

As part of the ARR revision SKM (2013, ARR report number P2/S2/012) conducted a review of long duration ARFs. The current study is a pilot study to explore methods of determining short duration ARFs.

M K Bab

Mark Babister
Chair Technical Committee for
ARR Research Projects

James Ball

Assoc Prof James Ball
ARR Editor

AR&R REVISION PROJECTS

The 21 AR&R revision projects are listed below :

ARR Project No.	Project Title	Starting Stage
1	Development of intensity-frequency-duration information across Australia	1
2	**Spatial patterns of rainfall**	**2**
3	Temporal pattern of rainfall	2
4	Continuous rainfall sequences at a point	1
5	Regional flood methods	1
6	Loss models for catchment simulation	2
7	Baseflow for catchment simulation	1
8	Use of continuous simulation for design flow determination	2
9	Urban drainage system hydraulics	1
10	Appropriate safety criteria for people	1
11	Blockage of hydraulic structures	1
12	Selection of an approach	2
13	Rational Method developments	1
14	Large to extreme floods in urban areas	3
15	Two-dimensional (2D) modelling in urban areas.	1
16	Storm patterns for use in design events	2
17	Channel loss models	2
18	Interaction of coastal processes and severe weather events	1
19	Selection of climate change boundary conditions	3
20	Risk assessment and design life	2
21	IT Delivery and Communication Strategies	2

AR&R Technical Committee:

Chair: Mark Babister, WMAwater
Members: Associate Professor James Ball, Editor AR&R, UTS
Professor George Kuczera, University of Newcastle
Professor Martin Lambert, Chair NCWE, University of Adelaide
Dr Rory Nathan, SKM
Dr Bill Weeks, Department of Transport and Main Roads, Qld
Associate Professor Ashish Sharma, UNSW
Dr Bryson Bates, CSIRO
Steve Finlay, Engineers Australia

Related Appointments:

ARR Project Engineer:	Monique Retallick, WMAwater
Assisting TC on Technical Matters:	Dr Michael Leonard, University of Adelaide

PROJECT TEAM

Project Team Members:

- Mark Babister, WMAwater #, ARR TC Project Manager
- Peter Stensmyr, WMAwater
- Monique Retallick, WMAwater

Note: # provided their time as in kind.

This report was independently reviewed by:

- Dr Rory Nathan, SKM

EXECUTIVE SUMMARY

This study is an initial exploration of the viability of using spatial grid interpolation methods to calculate areal rainfall, for use in determining short duration (equal to or less than 18 hours) areal reduction factors. Greater Sydney was chosen as the pilot study area as it has a comparatively high density of pluviographs. A total of 250 pluviograph rainfall gauges were included in the analysis, within an area bounded by Wollongong to the South, Newcastle to the North and Goulburn to the West.

The study uses an empirical approach for interpolating rainfall grids for all historic events in the pluviograph record. These grids were used as a basis for areal rainfall calculations. A number of algorithms for grid interpolation were trialled with natural neighbours chosen as most appropriate for the study.

Areal Reduction Factors (ARFs) are assumed to vary with Annual Exceedance Probability (AEP), duration and catchment area. In the current study, ARFs were calculated by creating an annual maximum series for areal and point rainfall at each gauge, assuming a circular catchment with a gauge located at the centroid. The Generalised Extreme Value distribution (GEV) was fitted to each series, and the design areal and point rainfall extracted for various quantiles. Thus, an ARF was obtained at each pluviograph, for a given catchment area, duration and AEP. Each case creates 250 results, which are filtered for quality and then statistically aggregated. The analysis was run for each combination of:

- 1, 2, 3, 6, 12, 18 hour durations;
- 1 in 2, 5, 10, 20, 50, 100 AEP; and
- 10, 50, 100, 500, 1000, 5000 km^2 circular catchment areas.

The results for the 1 in 2 AEP case are summarised in Figure ES 1 below, which also shows a comparison to the interim equations recommended by the ARR Revision Project 2: Spatial Patterns of Rainfall - Long Duration ARF report (SKM, 2013).

Filtering of results based on data quality was performed. Additionally, the ARF results calculated using natural neighbours grid interpolation were compared to ARF calculated using Thiessen polygons. The results suggest that natural neighbours yield better quality results than other methods trialled as part of the study.

The ARF calculated with this approach are similar to those recommended in the ARR Project 2 Long Duration ARF report (SKM, 2013), indicating that there is validity in the approach and it is recommended the method be explored further.

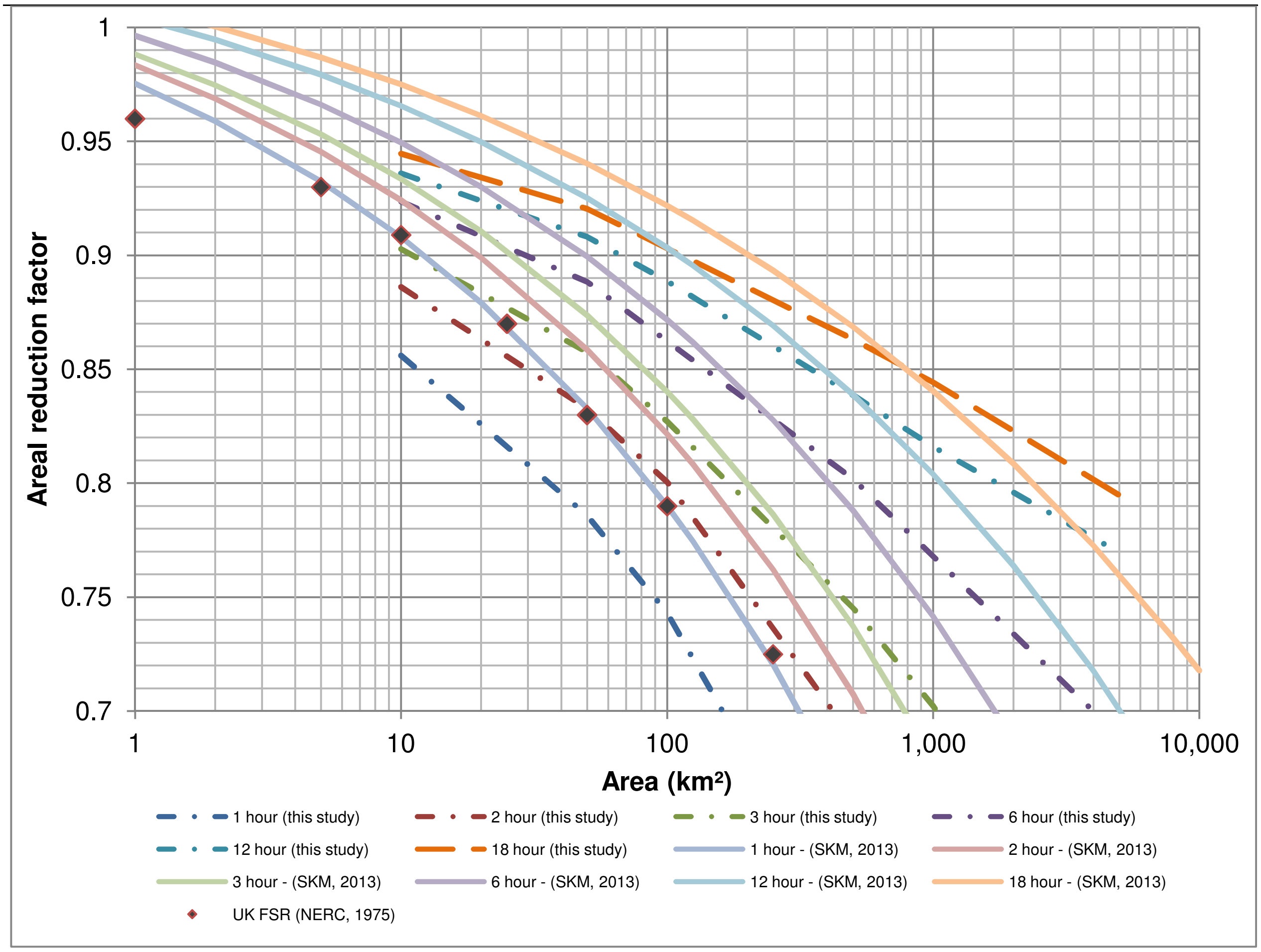


Figure ES 1. ARF Comparison – 1 in 2 AEP

TABLE OF CONTENTS

1. Introduction 1
  1.1. Background 1
2. Method 3
  2.1. Overview 3
  2.2. Short Duration Rainfall Issues and Considerations 3
  2.3. Data 4
  2.4. Software 6
  2.5. Choosing Stations 6
  2.6. Choosing Variables 6
  2.7. Finding Significant Events 7
  2.8. Grid Interpolation of Events 7
  2.9. Calculating Areal Rainfall 12
  2.10. Calculating Areal Reduction Factors 12
3. Results 13
  3.1. Results Filtering 17
  3.2. Bias Checking 19
  3.3. Data Quality 22
  3.4. Interpolation Method Comparison 24
4. Conclusions 27
5. References 28

LIST OF TABLES

Table 1. Distribution of Gauge Ownership 6

LIST OF FIGURES

Figure 1. Distribution of Pluviographs in Greater Sydney Area 4
Figure 2. Ownership of Pluviographs in Greater Sydney Area 5
Figure 3. Number of Stations with Data Used in the Study,1955 – 2012 5
Figure 4. Illustration of Step and Duration for Rainfall Sums 8
Figure 5. Natural Neighbours Interpolation Example 9
Figure 6. Grid Interpolation Method Comparison 11
Figure 7. ARF Comparison – 1 in 2 AEP 13
Figure 8. ARF Comparison – 1 in 100 AEP 14
Figure 9. Box Plots for 60, 180 and 540 minute- 10, 100 and 1000 km^2 Cases 16
Figure 10. Record Length Threshold Comparison 1 in 2 AEP 3 Hour Duration 18
Figure 11. Convex Hull Exclusion Criteria Illustration 19
Figure 12. $100km^2$ catchments with non overlapping sets 20
Figure 13. Results for Non Overlapping catchment set - 100 iterations, 1 in 2 AEP, 120 minute duration 20
Figure 14. Comparison of no overlapping catchment sets 21
Figure 15. Spatial Variability of ARF for the 1 in 2 AEP 120 minute $500km^2$ 22
Figure 16. Poor quality data point example 23
Figure 17. Effect of Raw Rainfall Data on the Final ARF 24
Figure 18. Thiessen Polygon ARF for 60, 180, 540 minute durations and 10, 100, 1000 km^2 cases 25
Figure 19. ARF Comparison – Thiessen Polygons and Natural Neighbours (NN) (1 in 2 AEP) . 26

1. Introduction

When estimating design rainfall for a catchment, it is necessary to account for the fact that point design rainfall intensity, corresponding to a certain AEP, will not occur uniformly over the whole catchment. The ratio of the design average areal rainfall and the point design rainfall is called the Areal Reduction Factor (ARF) and can vary with parameters such as AEP, catchment area, event duration and spatial region. SKM (2013) recommended equations based on these parameters across Australia for determining long duration areal reduction factors. However, some regions of Australia showed no correlation between ARF and AEP.

The majority of studies on ARF in Australia have been based on daily rainfall records (Srikanthan, 1995; Siriwardena and Weinmann, 1996) and are not directly applicable to rainfall durations shorter than 18 hours. As part of the ARR revision, SKM (2013) collated and reviewed the ARF methods recommended in Australia by the CRC-FORGE project. This includes both long and short duration equations. However, the short durations are qualified as being interim with further research recommended. In a review of methods for deriving ARF, Svensson and Jones (2010) stated that the increased amount of sub-daily rainfall data collected since the UK Flood Studies Report (NERC, 1975) would result in improved ARF estimates if new studies were carried out.

The aim of the current study is to explore the viability of deriving short duration ARF using pluviograph records. The Greater Sydney region was selected as for the pilot study as it has a high density of pluviographs. The project team is also familiar with many of the gauges. Other possible candidate locations for testing include Greater Melbourne and the Brisbane/Gold Coast region. The aim of the study is not to determine final ARF values, but rather to provide baseline data to allow an informed decision on the best approach going forward.

1.1. Background

ARR 1987 (Pilgrim, 1987) recommended ARF curves based on work carried out in the United States by Myers & Zehr (1980). Since then, extensive work has been carried out in Australia on ARFs for durations longer than 18 hours, utilising daily rainfall records. Many studies have concluded that the method used in ARR 1987 overestimated ARFs (Nittim, 1989; Porter and Ladson, 1993; Masters and Irish, 1994). Srikanthan (1995) undertook a review of methods for deriving ARF and recommended Bell's method (Bell, 1976) where sufficient data is available. Siriwardena and Weinmann (1996) applied a modified Bell's method for Victoria to derive ARF for durations between 18 and 120 hours, and found that ARF varied with AEP, catchment area and duration for the region. The study concluded that the method is acceptable to be extrapolated outside the 18 to 120 hours duration range; however, no analysis on sub-daily durations was carried out. Catchlove and Ball (2003) derived ARF for short durations for eight gauges in the Upper Parramatta River catchment. While the outcomes of the study was limited by the small number of stations used, the current study draws on the techniques used by Catchlove and Ball (2003). Masters and Irish (1994) applied an analytical method to derive ARF from 1 hour to 72 hours for a 2200 km^2 area, using nine gauges, and found that the method recommended in ARR (1987) overestimated the ARF. Jordan *et al* (2011) applied a modified Bell's method to develop ARF between 1 hour and 120 hours for NSW and ACT, though the sub-daily ARF equations were interpolated between the developed 18 hour curve and the 1 hour ARF curve presented in the UK Flood Studies Report (NERC 1975). These curves have since been incorporated into ARR revision Project 2 Spatial Patterns of Rainfall (SKM, 2013).

A study investigating short duration ARF using a large number of pluviographs over a large area has not previously been conducted in Australia. Due to the increased availability of sub-daily rainfall data, as well as computer processing power, it is now viable to undertake a larger-scale study of short duration ARF and applying methods similar to those developed for longer durations.

Most of the studies in Australia on ARF have used daily rainfall data and need careful translation to short duration events. There are significant differences in the physical behaviour of shorter events when compared to longer duration events. Differences include:

- Storm mechanisms,
- Spatial variability,
- Temporal variability, and
- Storm movement, especially on small catchments.

Using pluviograph data rather than daily gauge data results in technical challenges in terms of data handling, as pluviograph data is orders of magnitude larger than daily rainfall data. This means that many analysis methods used for daily data are not practical with pluviograph data. There are also challenges in terms of quality controlling the data ie. identifying unreliable data points is more difficult for pluviographs, partially due to the amount of data points, but also due to the inherent variability of short duration events.

As part of the CRC-FORGE project (as discussed in SKM, 2013), the various states have individually estimated long duration ARF using a modified Bell's approach (Bell, 1976), which estimates spatial rainfall using Thiessen polygons. Thiessen polygons are more appropriate for larger catchments and longer duration events, where both spatial and temporal variations in rainfall are smaller. The storm mechanism for longer duration events is generally frontal, which exhibits less variation than the convective mechanism responsible for the majority of short duration events. This results in smoother rainfall depth surfaces, which can be adequately represented by Thiessen polygons. It would be possible to use the grid interpolation methods explored in this study for longer durations and to investigate whether the estimates improve.

Current methods for deriving short duration ARF do not utilise the available data to its fullest potential. Current computer power means that it is viable to surface fit all significant historical events using two dimensional (2D) grid interpolation techniques. This can be done on the same temporal resolution as the pluviograph data (5 minutes) and for any rainfall duration of interest.

2. Method

2.1. Overview

The process of deriving the ARF can be split into five main tasks:

- Find all significant historical rainfall (point and areal),
- Spatially interpolating rainfall surfaces for the historical rainfall,
- Calculate point and average areal rainfall for theoretical circular catchments,
- Fitting a Generalised Extreme Value (GEV) distribution to an Annual Maximum Series of point and areal rainfall, and
- Calculating the ARF for a given site, catchment area, AEP and duration by extracting point and areal rainfall quantiles for the gauge and theoretical catchment

To automate this process, a software stack was developed in Python. Each step will be discussed in detail later in this chapter.

2.2. Short Duration Rainfall Issues and Considerations

When using sub-daily rainfall data compared to daily data, a number of complications arise, due to the actual volume of data that has to be processed and from the difference in rainfall behaviour for short duration events compared to long duration events.

Surface fitting short duration events is complicated by difficulty in detecting errors or temporal shifts in pluviograph data. Pluviographs have more sources of error which are harder to detect than for daily gauges. False zero values can significantly distort surfaces, yet are high likely to occur due to gauge failure or simple time shift in the recorder. The data used in the analysis comes from a range of agencies, which can use different treatments of daylight savings which could affect the results. Devising algorithms that correctly identify measurement errors or time shifts in a pluviograph record warrants further investigation.

Short duration events generally display more random behaviour than long duration events. This makes it easier to miss rainfall peaks occurring between gauge locations. It also makes it more difficult to quality control the data, as it is not possible to simply compare suspicious data points against nearby stations to see if the rainfall is plausible. There are also fewer pluviograph stations available to check against. Therefore pluviograph data is more likely to contain bad data points in the record. Due to the volume of data points (orders of magnitude larger than for daily stations), it is often not practical to manually check the data. The rainfall data used for this study has been quality controlled by the Bureau of Meteorology (BoM) as part of the updating of the Intensity Frequency Duration (IFD) data, but there are still many bad data points. The BoM quality checking approach is based on comparing daily totals with daily rainfall grids from the Australian Water Availability Project (AWAP). This approach has been known to flag records from well-known non-BoM gauges as suspicious. In Coffs Harbour, a cluster of gauges investigated by Richards *et al.* (2012) were initially rejected in the IFD analysis because the gauges disagreed with BoM station records, even though all the gauges in the cluster showed similar behaviour.

The developed method is fairly robust to errors in individual gauges. They will increase the error bars on the final ARF calculated, but the mean/median values do not change significantly (see Data Quality for more discussion on this issue).

2.3. Data

The pluviograph data used in this study was provided by the Bureau of Meteorology (BoM) and has been quality checked as part of the ARR Revision Project 1: Development of Intensity Frequency Duration Information across Australia (Green *et al.*, 2012). The data set includes 250 stations operated by 5 agencies including BoM (refer to Table 1 and Figure 2). Figure 1 shows the spatial distribution of the gauges, with the size of each point indicating the record length. Since the quality of the grid interpolation depends on there being a sufficient number of data points available at a given time, the number of stations operating each year is shown in Figure 3. The drop off in the number of gauges around 2008 is likely due to the latest data sets for some locations not being available from some agencies.

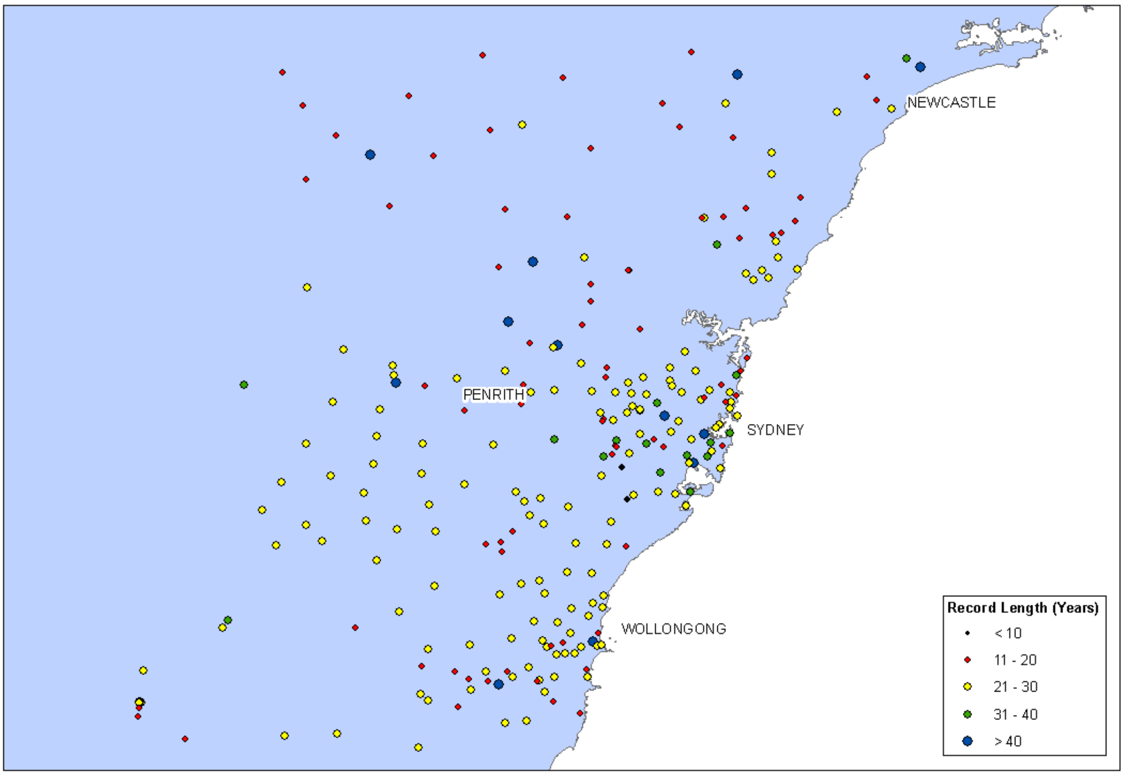


Figure 1. Distribution of Pluviographs in Greater Sydney Area

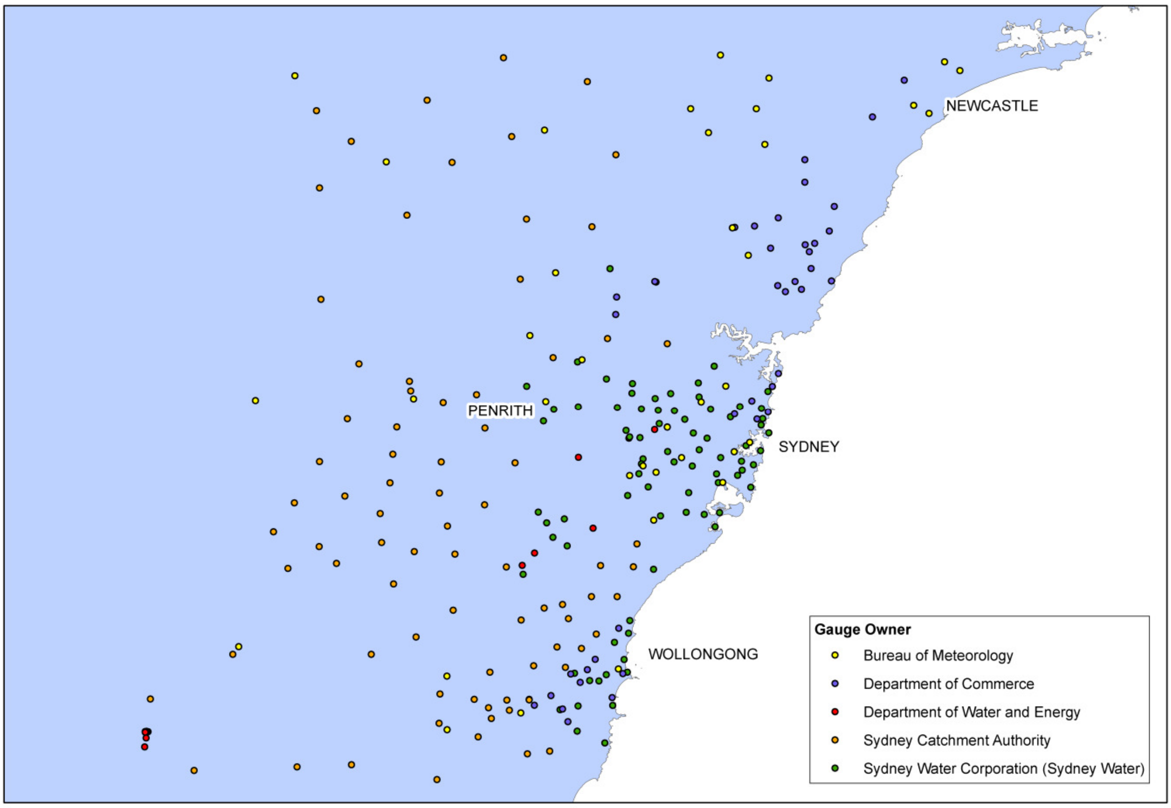


Figure 2. Ownership of Pluviographs in Greater Sydney Area

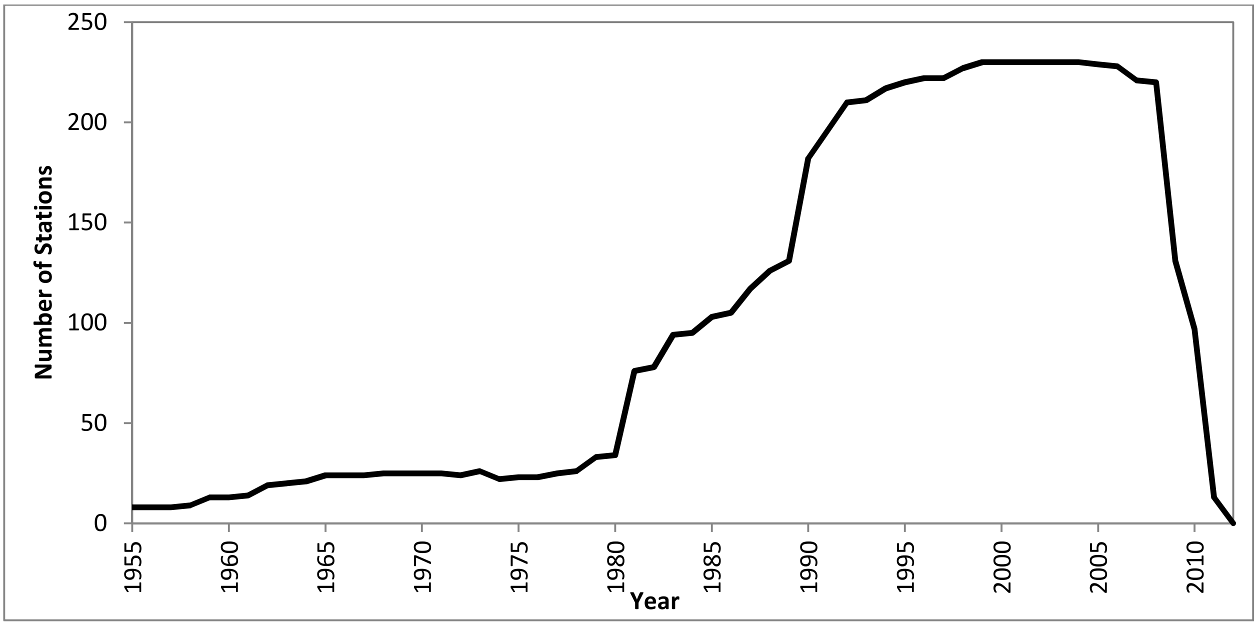


Figure 3. Number of Stations with Data Used in the Study, 1955 – 2012

Table 1. Distribution of Gauge Ownership

Owner	Number of Stations
Bureau of Meteorology	37
Department of Commerce	39
Office of Water	11
Sydney Catchment Authority	83
Sydney Water Corporation (Sydney Water)	80

2.4. Software

To aid with the method for analysis, a software stack for the processing was developed. The software is used to identify the maximum annual rainfall depths and times in the pluviograph record, interpolate grids for point data using a variety of interpolation methods, calculate average areal rainfalls, and fit statistical distributions to data.

2.5. Choosing Stations

Grid interpolation procedures are challenged by the data density. The Greater Sydney area was chosen because it has one of the highest data densities in the country and would therefore reduce gridding errors due to data sparsity. It is possible to use a larger area, but there are diminishing returns when considering the additional processing time and storage space required as a larger grid will be computed. Lower gauge density means that the quality of the spatial rainfall estimates will be lower.

For many of the data points outside the densest areas, the grid interpolations will be of a lower quality due to the lack of nearby stations. Similarly, early in the record when less station records are available the areal estimates will be of lower quality.

2.6. Choosing Variables

ARF can depend on a number of factors. To limit the scope of the study, it was chosen to investigate how ARF varies with duration, AEP and catchment area. Comparisons are also presented of grid interpolation methods and the effect of data quality on ARF.

For each of the 250 pluviograph gauges, an ARF is calculated for each possible combination of the parameters listed below:

- Duration (min): 60, 120, 180, 360, 540, 720, 1080
- AEP (1 in x): 2, 5, 10, 20, 50, 100
- Area (km^2): 10, 50, 100, 500, 1000, 5000

The ARF proposed will then be the median value for each combination, which gives a total of 7 * 6 * 6 = 252 ARF. It might be possible to generalize these values to an equation; however, this was considered beyond the scope of the study.

2.7. Finding Significant Events

To calculate ARF, an annual maximum series of point and areal rainfall is required. The point series can be calculated from the pluviograph data, but the areal annual maximum series is less straight-forward to determine. It is computationally unviable to interpolate a grid for every time 5 minute step between 1955 and 2012, therefore all the significant areal rainfall for each gauge needs to be identified by some means to reduce the number of grids to be interpolated.

To identify all the times of significant areal rainfall, the time of each of the top three independent events per year at each station was extracted. A simplified independence criteria was used, in which events occurring on separate days were considered independent. This criteria should be refined as part of future work. This was done for each of the durations listed in Section 2.6. The start time was then buffered three hours in each direction to identify a window within which grids would be created. Any overlapping events were removed, so that events which occurred at several gauges were only gridded once. Grids for the whole region were interpolated for each event.

For each gauge, any significant spatial rainfall needs to be identified and accounted for, since the spatial rainfall will be calculated whenever there is significant rainfall at that gauge. By gridding the whole area for every significant event at any location, the risk is minimised of missing significant spatial rainfall where there was little point rainfall at a gauge.

Creating the event by buffering the start time three hours either side of the time of interest means that it is possible to miss some spatial rainfall; however, it was necessary to have some limit for computational reasons.

2.8. Grid Interpolation of Events

As discussed in Section 2.7, a series of times were selected for grid interpolation, for each duration. The rainfall sum for the duration with a 5 minute time step over the 6 hour window was then gridded. For example, gridding an 6hr event from 01:30 to 07:30 on 2013/04/18 for the 120 minute duration, 120 minute rainfall sums are calculated every 5 minutes. Hence, this event will generate 12 grids per hour over 6 hours, 72 grids in total. At other gauges the event is likely to have started earlier or finished later, increasing the number of grids required. Figure 4 below illustrates the sums calculated for a single event at a single gauge.

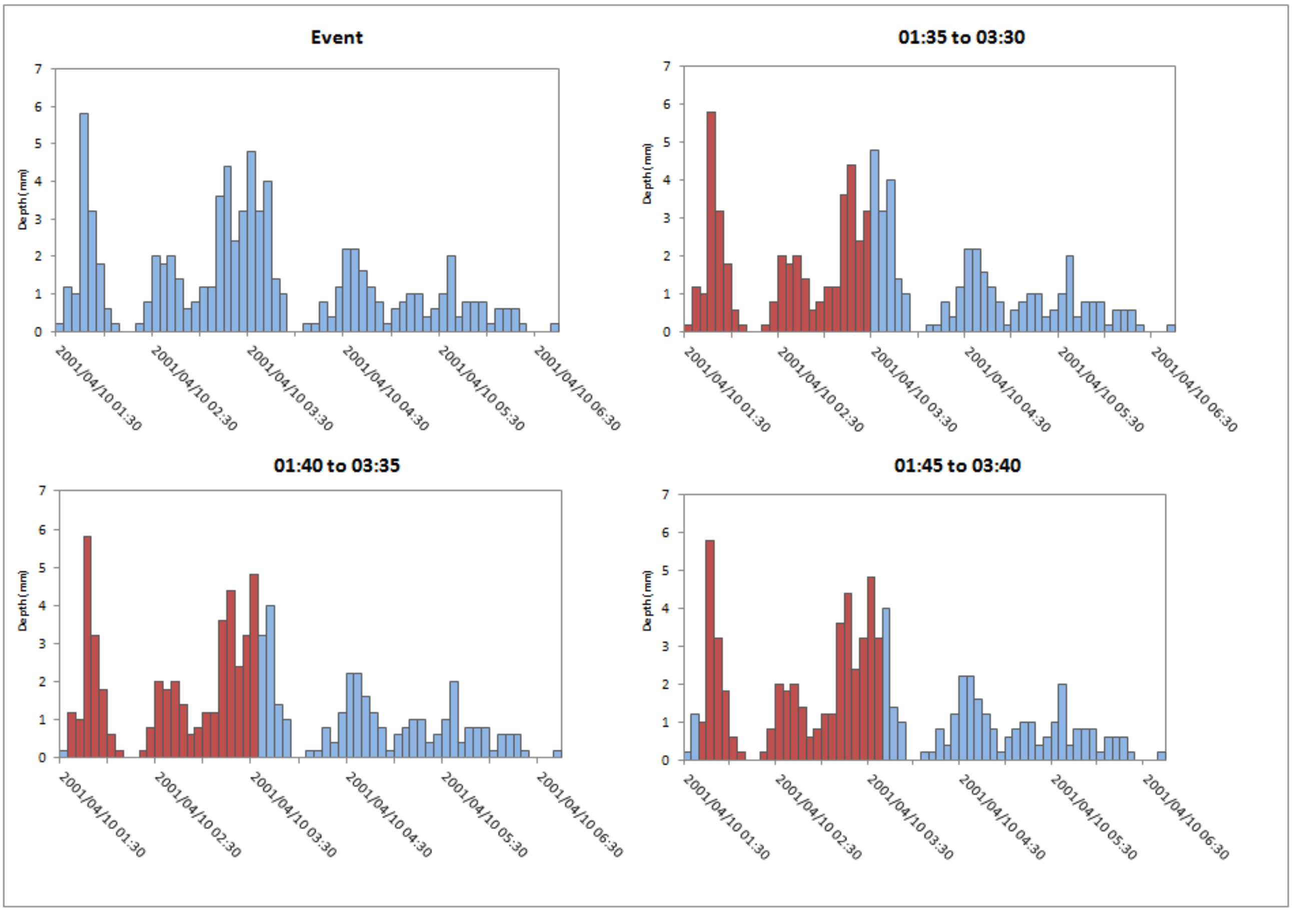


Figure 4. Illustration of Step and Duration for Rainfall Sums

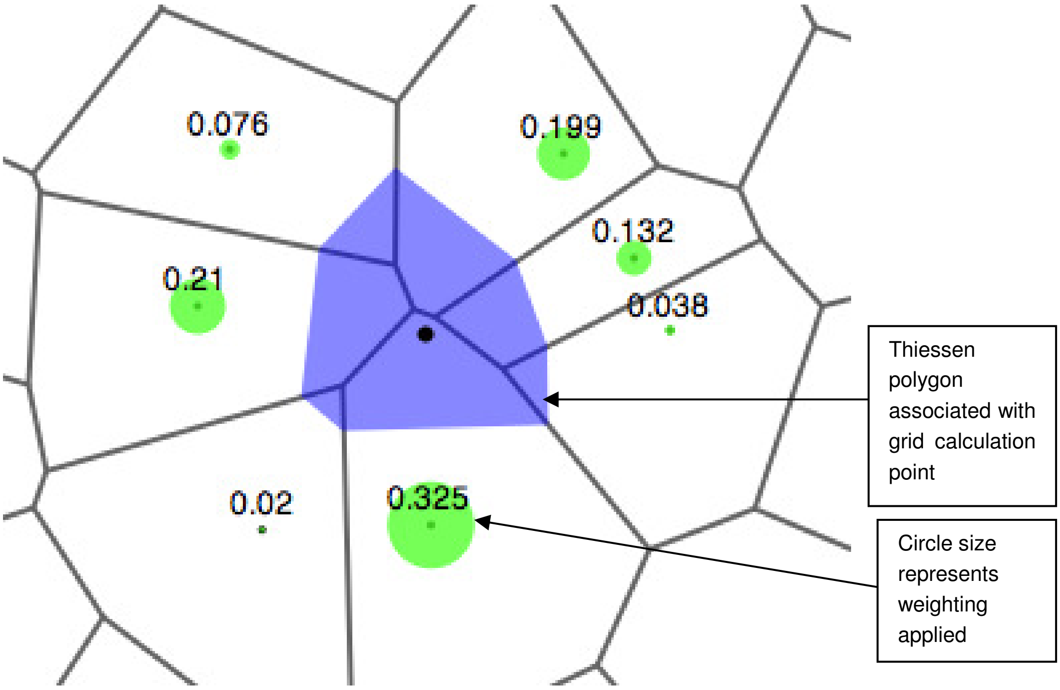
Several interpolation methods were considered for the study, including Thiessen polygons, linear, inverse distance, natural neighbours and cubic spline.

The natural neighbours interpolation method is based on Thiessen polygons. For any given point on a grid, its natural neighbours are the points used to create the neighbouring Thiessen polygons. The algorithm works as follows:

1. Compute the Thiessen polygons for the known data points
2. Insert a new point where a value will be calculated
3. Redraw the Thiessen polygons including the new point
4. Overlay the polygon for the new point with the original Thiessen polygons
5. Each of the original polygons that are within the overlay contributes to the new calculated point value, weighted by area (Sibson, 1981).

Figure 5 below shows the Thiessen polygons before and after the insertion of the new point.

Every calculated grid point is guaranteed to be within the convex hull of the scatter point set. The resulting grid is generally similar to an inverse distance interpolated grid, containing smoothed data with no interpolated values outside the minimum and maximum observed values. However, Figure 5 shows how the interpolated value varies from the inverse distance estimate. All the known points are approximately the same distance from the input point (and hence would have equal impact using inverse distance interpolation), but get weighted differently due to the spatial distribution.



Thiessen polygon associated with grid calculation point

Circle size represents weighting applied

Figure 5. Natural Neighbours Interpolation Example[1]

Natural neighbours was chosen as the method in this study for two main reasons:

- Speed – grid interpolation is reasonably fast – in the order of 4 grids per second with current algorithms and parameters
- Simplicity – With the large volume of grids being generated, it is not possible to quality control individual grids. The natural neighbours algorithm is simple, straight-forward and ensures that no anomalous grid values occur

To keep storage space and interpolation times reasonable, a grid resolution of 2.5 kilometres was adopted, which is similar to the CDIRS resolution of 0.025° used by BoM for the national IFD grids. For the Greater Sydney area, this results in approximately 7500 cells per grid.

The choice of natural neighbours was partially a practical one. Several studies have shown that a spline-based method is potentially superior for the gridding of rainfall surfaces (Luk and Ball, 1996; Umakhanthan, 2002). A simple comparison between interpolation methods is provided in Figure 6. Conducting a comprehensive analysis was considered out of scope for the current project. It is recommended that future work consider a Leave-One-Out (LOO) analysis for comparison of algorithms. There is a also a separate ARR project currently being undertaken by BoM, which uses radar and pluviograph data to generate rainfall surfaces. The outputs of this study could be used to compare different gridding algorithms. The effect of grid size on results should also be investigated in any future studies.

---

[1] Image taken from the Wikipedia Commons [http://commons.wikimedia.org/wiki/File:Natural-neighbors-coefficients-example.png]

Figure 6 depicts interpolated two hour rainfall sums between 17:00 and 19:00 for an event that occurred on 2008-02-06 for four interpolation methods (natural neighbour (top left), Thiessen polygons (top right), linear interpolation (bottom left) and cubic spline (bottom right)). Natural neighbours and the cubic spline grids generate a more realistic grid than the linear interpolation or Thiessen polygons methods. The cubic spline method produces interpolated values that are larger than the maximum observed rainfall. The method does not assume that the maximum rainfall that occurred was captured by a gauge. Hence, the cubic spline method allows for a more sophisticated gridding procedure. However, it also means that cubic spline grids need to be quality checked which is significantly more work. For the purposes of this study, a pragmatic decision was made to use the natural neighbours method.

Section 3.4 contains a comparison between ARFs generated by Thiessen polygons and natural neighbours.

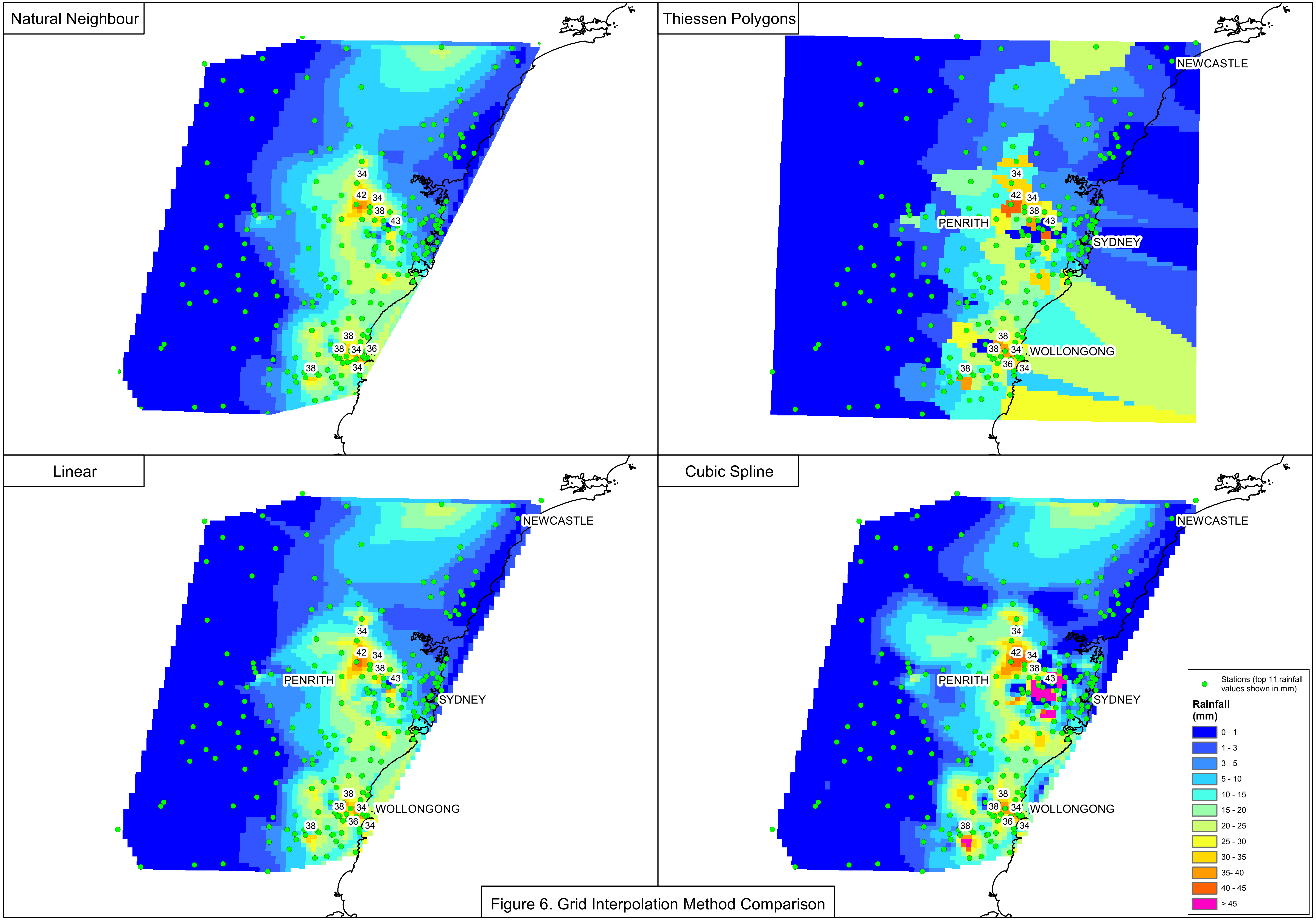


Figure 6. Grid Interpolation Method Comparison

2.9. Calculating Areal Rainfall

An annual maximum series of areal rainfall at each gauge, for each catchment size and rainfall duration was created. Areal rainfall was calculated as the average rainfall within the imaginary circular catchment around each gauge. To avoid introducing a bias in the results, this calculation is only carried out if there is point rainfall recorded at the time for the gauge. Therefore, the annual point series and the annual spatial series will have the same length. This also provides some quality assurance in the areal estimate, as it will not be calculated if there was no data in the area.

Calculating reliable areal estimates is a difficult problem. As part of the process quality indicator parameters are logged, which can be used to exclude certain data points before calculating the ARF. These parameters include the total number of stations used for generating the grid, number of stations within each circular catchment, and number of stations within 25 km of each circular catchment. In addition to this, the percentage "uniqueness" is logged for each areal estimate, therefore the fraction of the circular catchment which does not overlap with a circular catchment from another gauge. These parameters could be used as part of future work, to quantify the impact of different issues with the data set.

2.10. Calculating Areal Reduction Factors

The definition of an ARF is:

$$ARF = \frac{I_{Area}}{I_{Point}}$$

The process for calculating a set of ARF from point and areal data consists of the following steps for a given duration and catchment size:

- Create annual maximum series of point and areal rainfall,
- Fit the Generalised Extreme Value Distribution (GEV) to each series,
- Extract rainfall quantiles at suitable AEP (1 in 2, 5, 10, 20, 50 and 100 AEP),
- The ARF for the gauge at each AEP is then the areal design rainfall divided by the point design rainfall: $d_{AEP, areal} / d_{AEP, point}$

For each combination of AEP, duration and catchment area, point and areal rainfall annual maximum series are constructed. Each series is then fitted to the GEV distribution using the LH-Moments approach described by Wang (1997) with zero shift (H=0). The AEP quantiles are extracted using code based on Hosking and Wallis' LMOMENTS package (Hosking and Wallis, 1990). To remove some bad results before fitting the GEV, records are filtered out where the station record is shorter than three years, and annual records where the point maximum is zero, as the gauge in question was likely not operating over that time even though (zero) rainfall was reported.

With the GEV quantiles calculated for each AEP, duration and catchment area, a series of ARF for each of these cases is calculated. From these series a mean and median ARF for each case is derived.

3. Results

Figure 7 below shows a comparison of the ARF developed as part of this study and the interim short duration ARF recommended by the ARR Project 2 Long Duration ARF report (SKM, 2013) for the 1 in 2 AEP. The ARR Project 2 Long Duration ARF report (SKM, 2013) did not find a correlation between varying AEP and ARF for short durations in the New South Wales GSAM region, therefore there is only a single set of equations presented. Results from the current study are presented for the median ARF for each catchment area and duration. Also presented on Figure 7 are the results from the UK study (NERC, 1975) that the interim equations were based on.

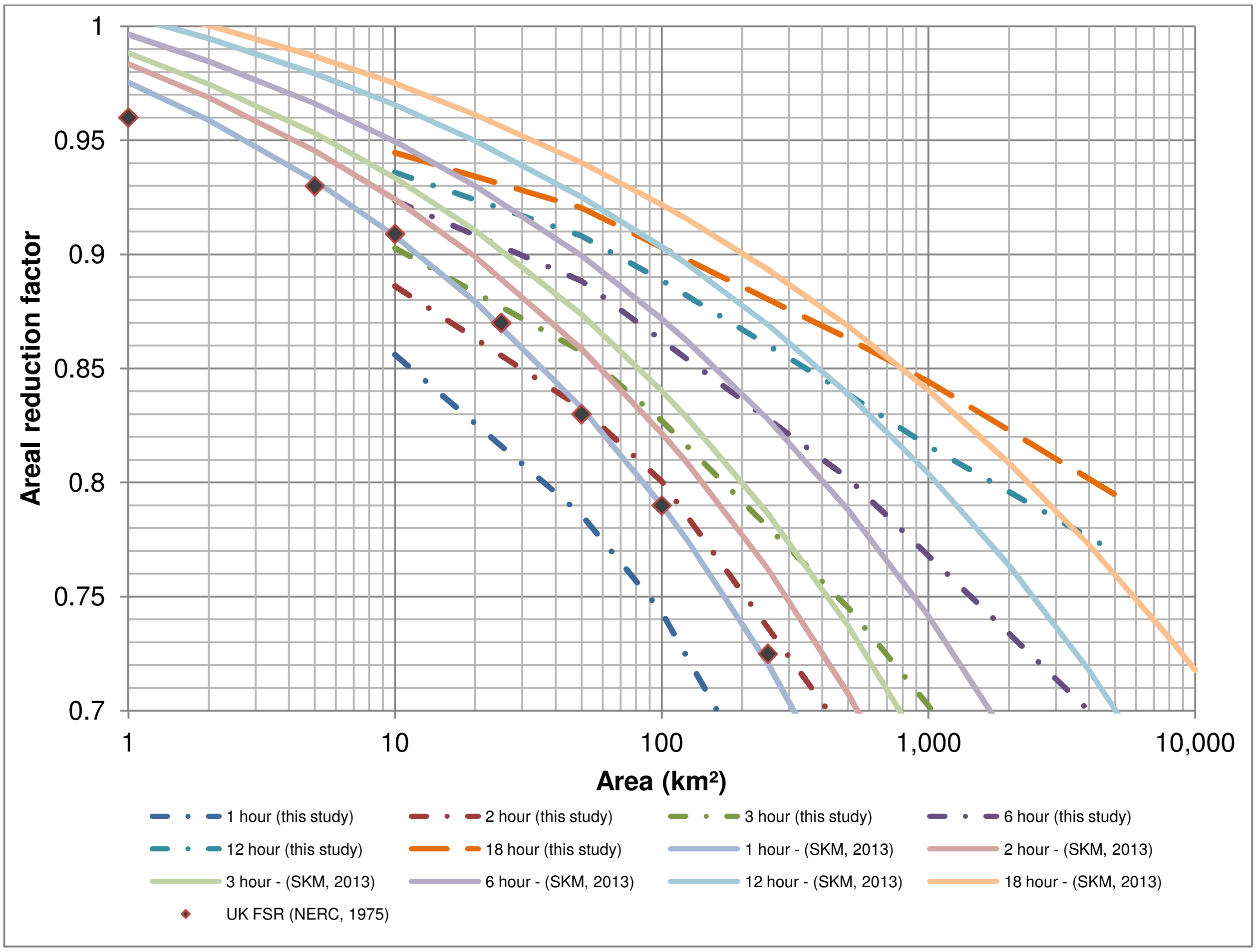


Figure 7. ARF Comparison – 1 in 2 AEP

Figure 8 shows the same comparison for the 1 in 100 AEP case. The ARR Project 2 Long Duration ARF lines are the same as in Figure 7. Note that the vast majority of station records are shorter than 100 years with the average record length in the order of 30 years. The 1 in 100 AEP estimates are included for comparison rather than practical use.

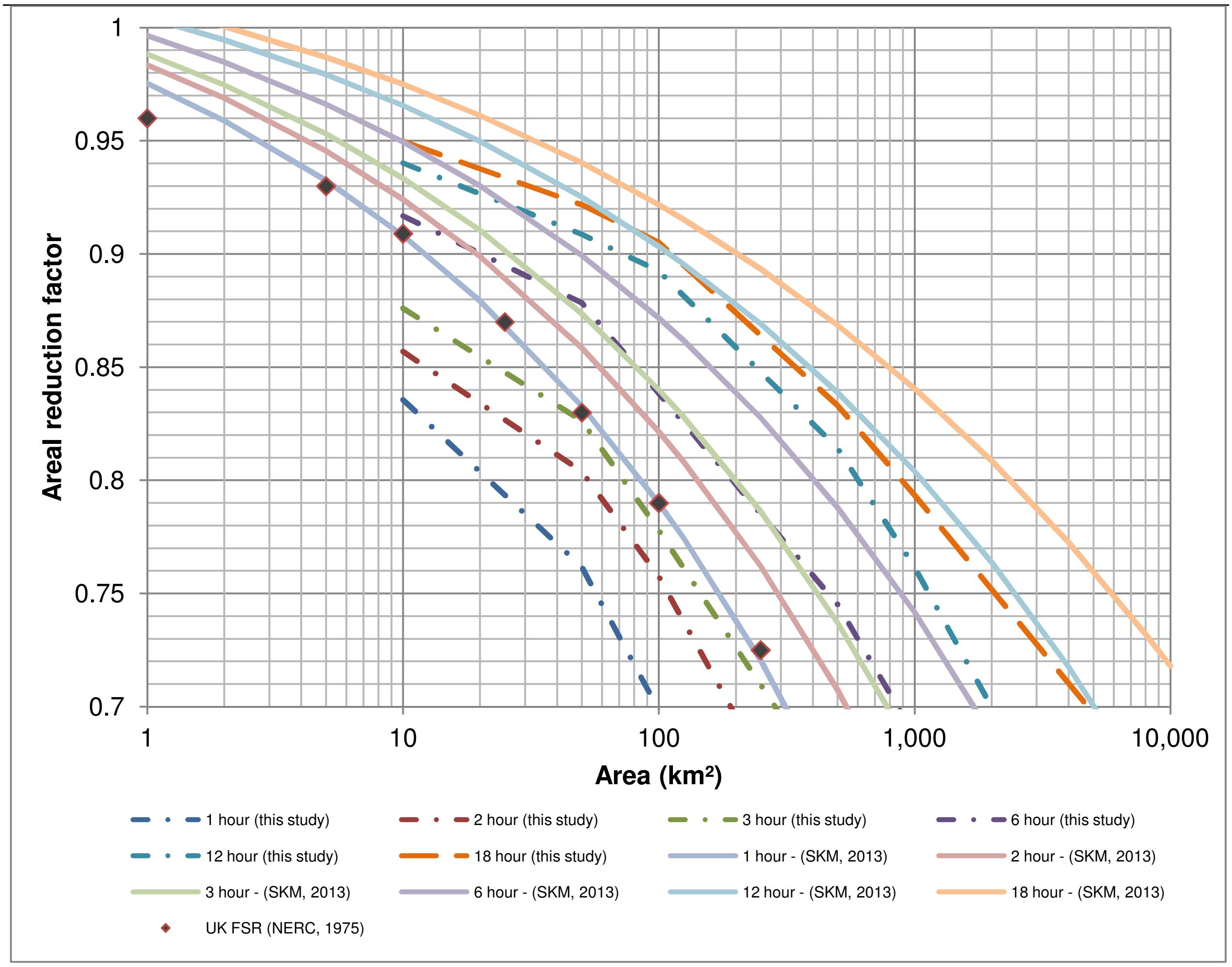


Figure 8. ARF Comparison – 1 in 100 AEP

There are no results for catchment sizes smaller than 10 km^2. The grid size used for the study was 2.5 km, which means that catchments smaller than 10 km^2 are generally contained within a single grid cell. This results in the point rainfall and the areal rainfall being the same value. To allow for ARF calculations for smaller catchment areas, a finer grid size would have to be used.

There is a trend towards lower ARF value with rarer AEP. In Figure 7, the ARF results from the current study cross the ARR Project 2 Long Duration ARF results, whereas in Figure 8, they are consistently lower. These results do not account for any overlap between catchments, which means that there will be an overrepresentation of data from areas with a more dense pluviograph network, such as Central Sydney. The implications of this are discussed below in the Section 3.2.

It would be possible to use the at-site GEV distributions derived by BoM as a source for the point rainfall quantiles, rather than doing a custom analysis for this study. The BoM distributions use regionalised values for the shape and scale parameters in the GEV distribution, which should give more accurate rainfall quantiles, particularly for rarer AEPs. However, the choice was made to use only the at-site GEV fits without regionalised parameters. If the data sources used are not equivalent, there could be a bias introduced which would make the ARF estimates less accurate. Areal rainfall data for a catchment was only logged when there was at-site data available; if BoM estimates were used, there might be a mismatch between the data sets. However, this might explain some of the changes in ARF at rarer AEPs, as the current estimates are of lower quality due to insufficient record lengths. Using regionally estimated parameters might improve these results.

The results presented above are the median ARF of 250 calculated values, with some points filtered out (See Section 3.1 for details). Box plots for a a number of cases are shown in Figure 9. These illustrate the variation between different locations, catchment sizes, AEP and durations.

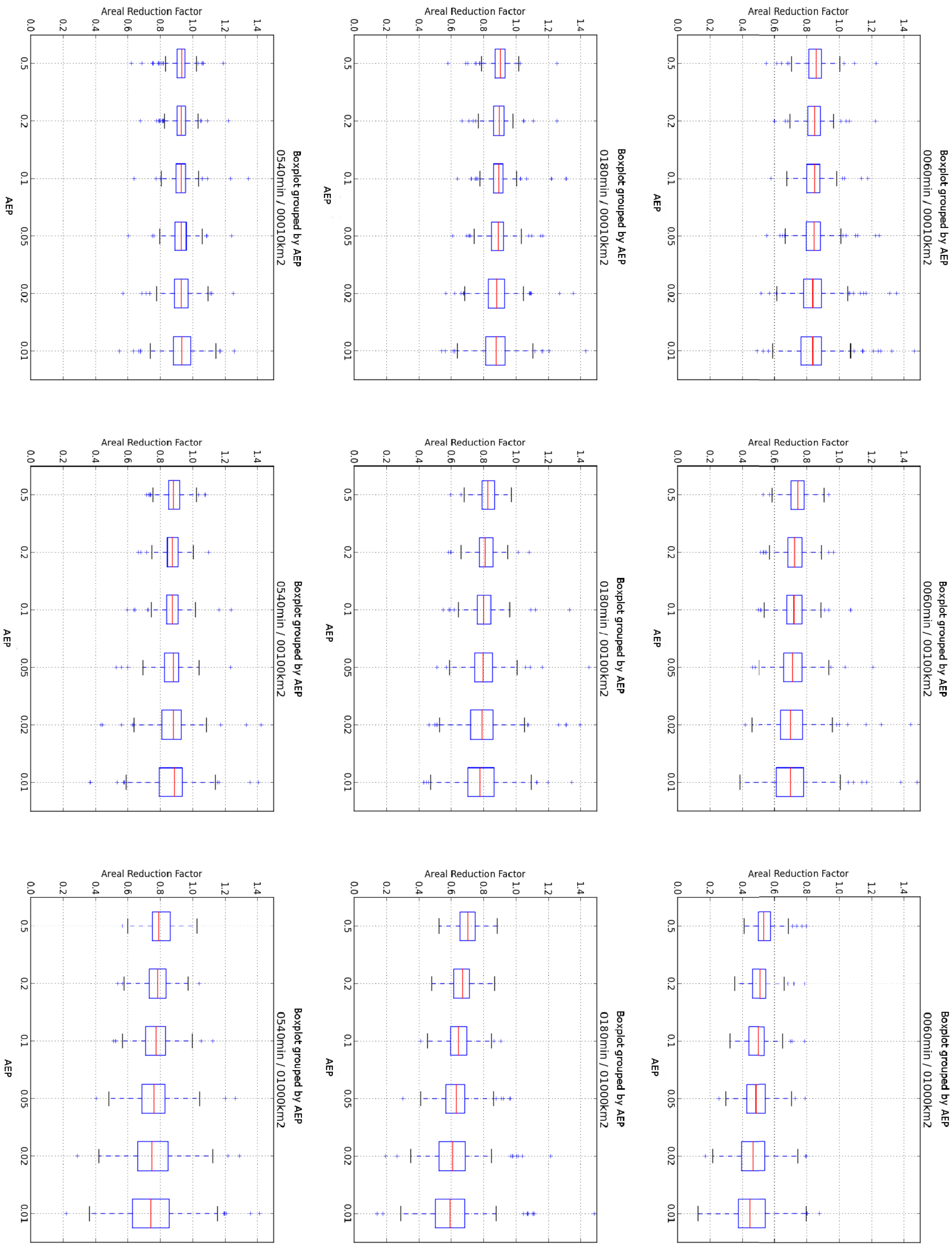


Figure 9. Box Plots for 60, 180 and 540 minute – 10, 100 and 1000 km^2 Cases

From Figure 9 some general observations on trends regarding AEP can be drawn:

- ARF increases with longer duration (less reduction in rainfall)
- ARF decreases with increased catchment area (more reduction in rainfall)
- For longer durations, ARF decreases with rarer AEP (more reduction in rainfall)
- For shorter durations, there is no clear correlation between AEP and ARF

In addition the following general observations can be made:

- There appears to be a bias in the outliers – there are more high outliers than low ones,
- Despite filtering and quality control, there is still a significant number of outliers,
- The 25% and 75% quantiles tend to get wider with larger catchment area, and
- The 25% and 75% quantiles also increase with rarer AEP.

The bias in outliers is most likely caused by the poor quality rainfall data. If an unrealistically high rainfall sum is reported at a gauge, this will result in a higher value in the annual maximum areal rainfall series for the nearby gauges (while not affecting the point series at those gauges). In many cases, the annual maximum areal rainfall for the associated year will be significantly higher than the annual maximum point rainfall. This will impact the final ARF, often resulting in an ARF value greater than 1. The effect becomes more pronounced with rarer AEP. The poor quality rainfall data is discussed in more detail in the Section 3.3. Another factor is that the minimum rainfall is bound at zero, while the maximum rainfall is unbound, therefore there may be a natural trend towards outliers on the high side.

3.1. Results Filtering

Some basic filtering was applied to the results before generating the graphs and box plots. Years in a series where the maximum point rainfall is zero were removed from both point and areal series. Zero maximum point rainfall in a year indicates that the pluviograph was not in operation for the year, but rainfall has still been logged as zero. This appears as though the station is operating, and areal rainfall is logged. A filter is used to remove these records from the analysis.

Stations with annual series shorter than 3 years are also removed. While the data points are valid, fitting a GEV distribution to such a short record is not possible. A record length of 3 years while short, can still be used for 1 in 2 AEP estimates. An indication of the impact of the record length filtering threshold can be seen in Figure 10 which shows ARF for the 1 in 2 AEP, 3 hour duration case using different record length thresholds. The record length appears to have little effect on the final ARF estimate for the 1 in 2 AEP case. However, deciding a cut-off point for station record length requires further investigation. Examining the trade-off between sample size and accuracy should be part of future work.

Grids are only generated for years where there were at least 10 gauges in operation. This is a crude filter for lack of data, and should be investigated further as part of future work.

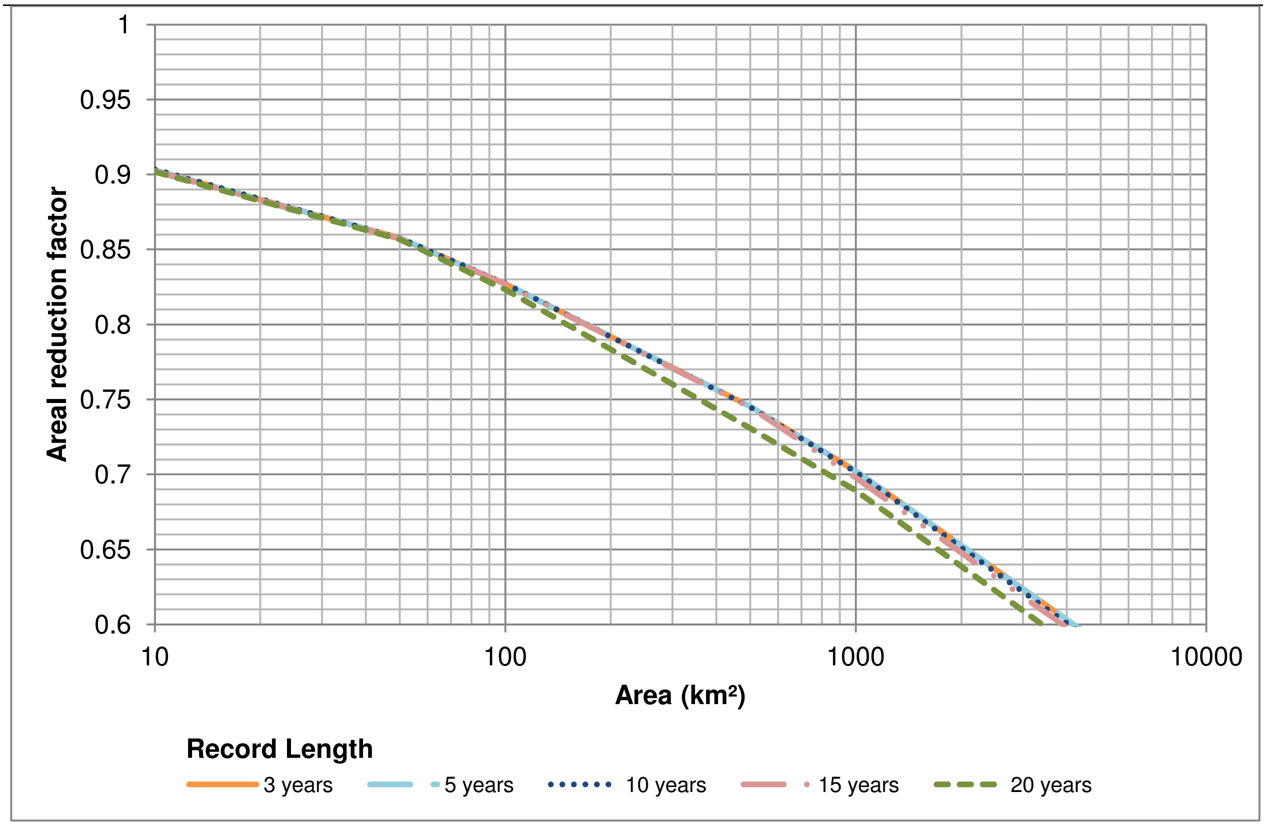


Figure 10. Record Length Threshold Comparison 1 in 2 AEP 3 Hour Duration

Stations where a significant portion (more than 30%) of the catchment lies outside the convex hull of all stations used for the case are also removed from the final analysis. The reason for this is to remove stations on the boundary, where part of the catchment falls outside the interpolated grid surface. As the percentage of the catchment outside the grid boundary increases the closer the areal rainfall will be to the point rainfall, since it will mostly be influenced by the point. A large part of the catchment will be null values, which are ignored when computing areal averages. Figure 11 below shows the convex hull, 100 km^2 area catchments, with the excluded catchments highlighted.

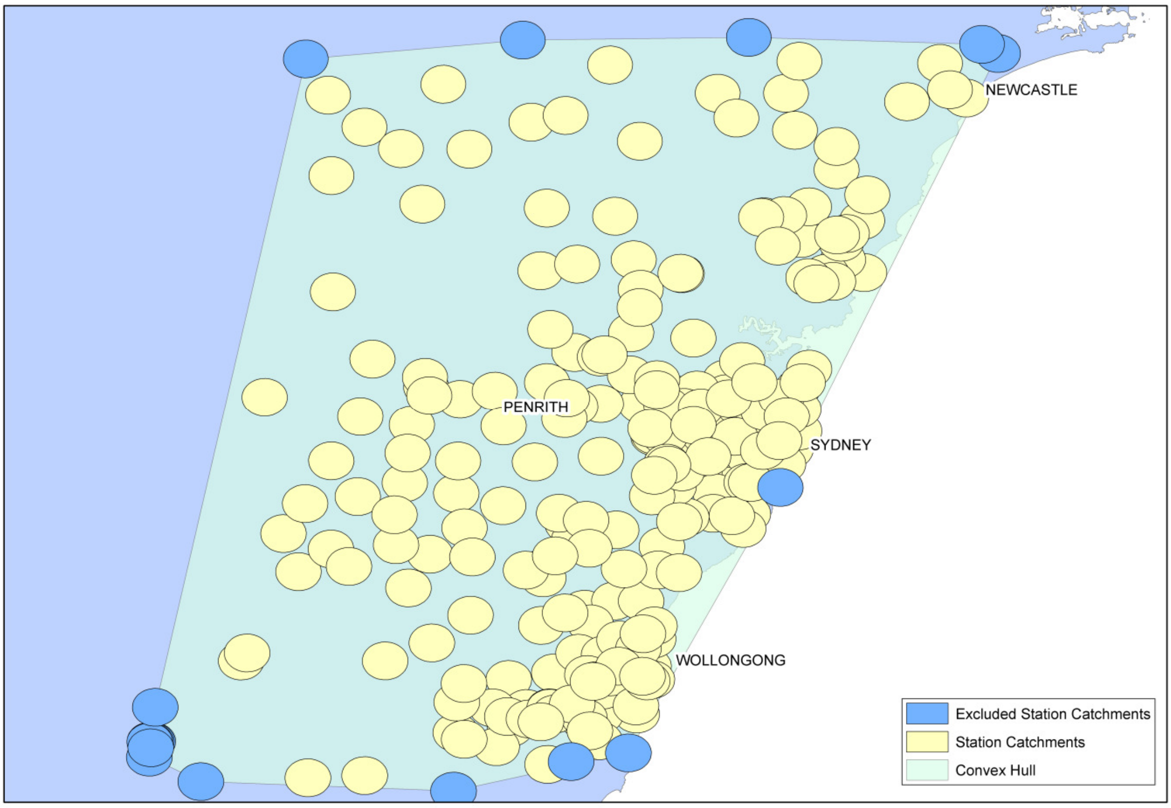


Figure 11. Convex Hull Exclusion Criteria Illustration

Given the known data issues with the data used in this study a reasonably simple and easy to apply filtering method has been applied. Ideally, poor quality data points will not have any significant impact on the final results if the method is robust. This is discussed in more detail in Section 3.3.

3.2. Bias Checking

By controlling the overlap of catchments, the effects of spatial variation of ARF and the bias of counting some areas more than once can be investigated. This allows checking if creating a set of non-overlapping catchments, and using this to calculate ARF, produces different values. If this is the case, the use of adjustment factors or other ways of correcting the bias needs to be considered. Since there are many possibilities to construct a set of non-overlapping catchments from the base set of 250, 100 iterations were run for the 1 in 2 AEP, 120 minute duration case, for each catchment size. The set of non-overlapping catchments was constructed by randomly picking a catchment from the whole set and adding it to the new set if there was no overlap, and discarding it if there was. Figure 12 shows all the 100 km^2 catchments with a non-overlapping set highlighted.

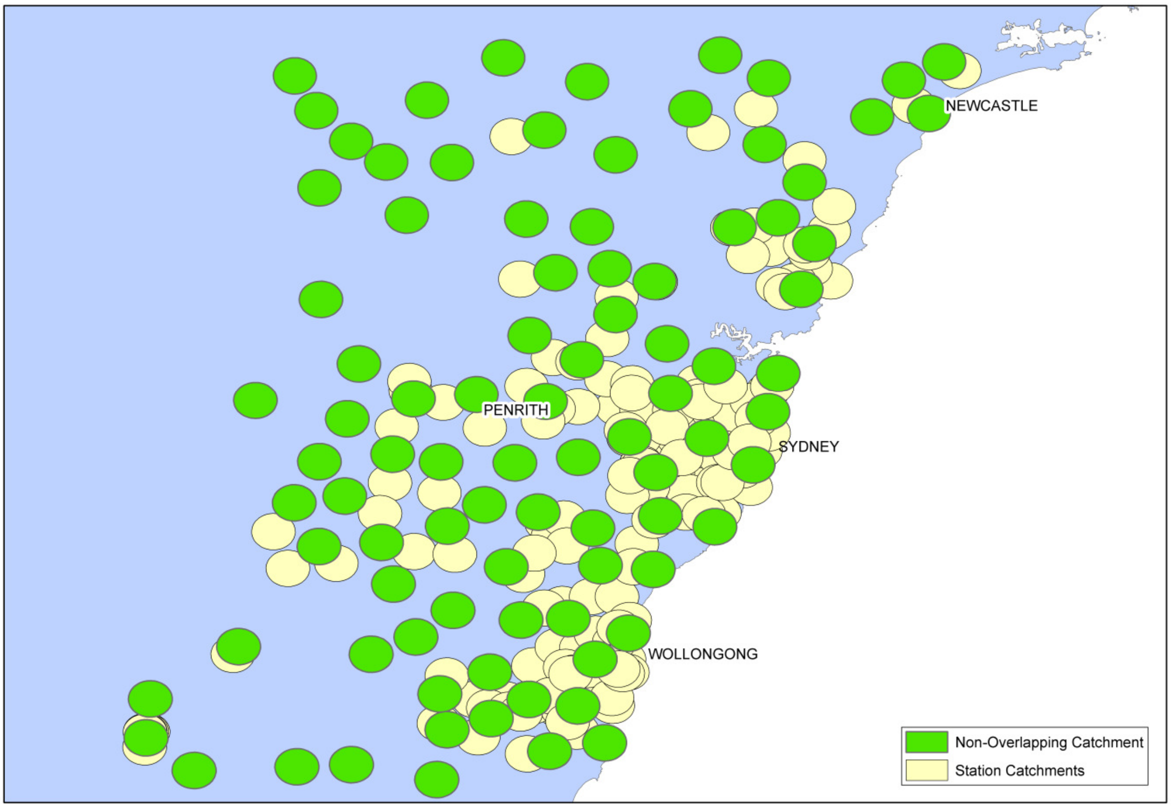


Figure 12. 100km^2 catchments with non overlapping sets

Figure 13 shows the results of the 100 iterations for each catchment area, for the 1 in 2 year AEP, 120 minute duration, as a box plot. For comparison, the original result using the median ARF of all stations is shown as well.

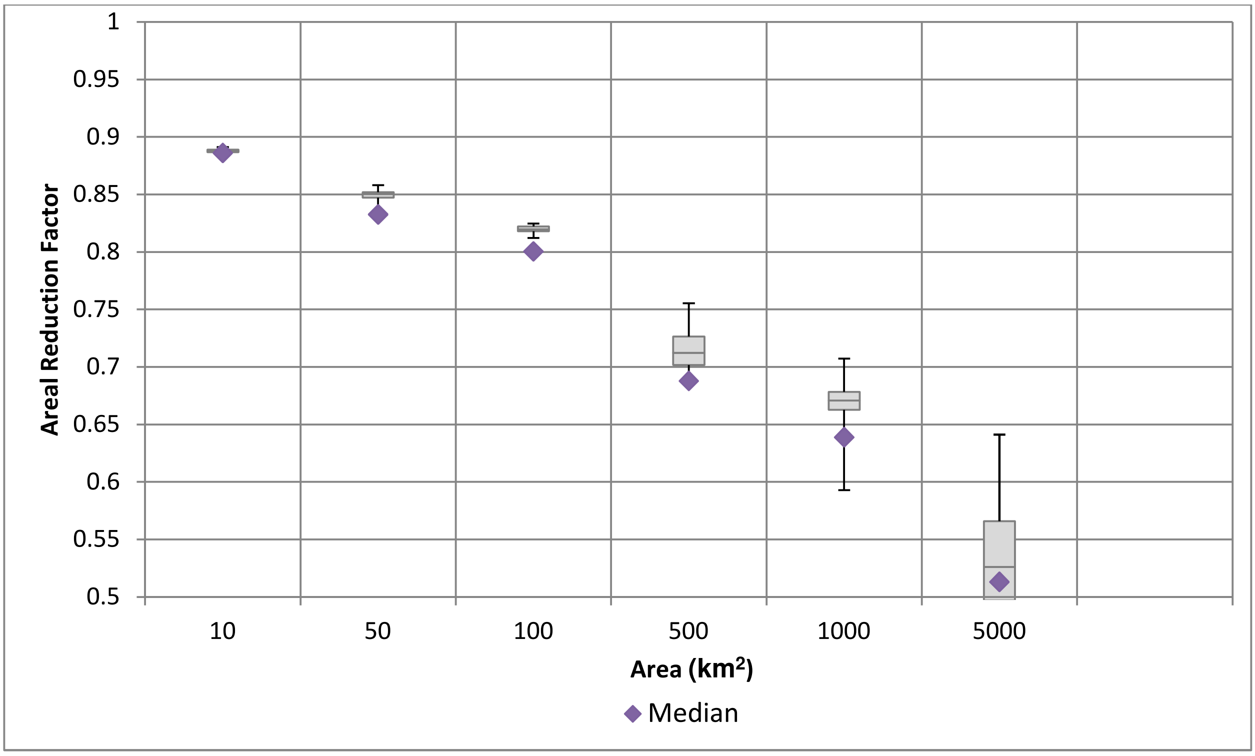


Figure 13. Results for Non Overlapping catchment set - 100 iterations, 1 in 2 AEP, 120 minute duration

The graph for the case is shown on Figure 14 with the result from Figure 7 added for comparison purposes.

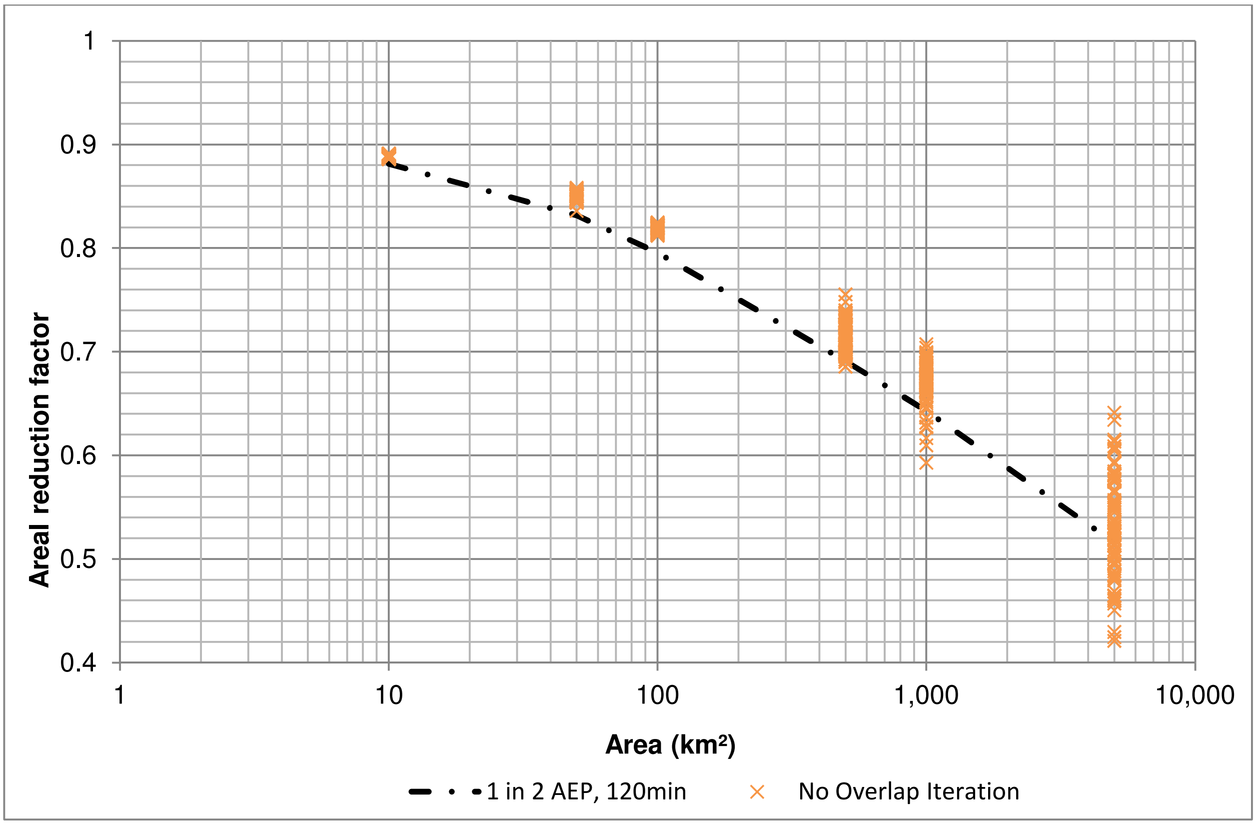


Figure 14. Comparison of no overlapping catchment sets

Using non-overlapping catchment sets produces ARF values that are higher on average, for the 1 in 2 AEP, 120 minute duration case. For small catchments, the effect is very small. The effect is greater with larger catchment areas up to 5 000 km^2. For a catchment of 10 000 km^2 it matches the non-filtered data set for the case. The similarity for small catchments is not surprising, as the catchments are small enough to not have significant overlap; hence, the non-overlapping set will be almost the same as the original set. As the area gets larger, the majority of the catchments filtered out are in the data-dense regions, which are mostly near the coast. It appears as though the ARF varies spatially, which would explain this behaviour. Figure 15 below shows the spatial distribution of the ARF for the 1 in 2 AEP, 120 minute, 500 km^2 case.

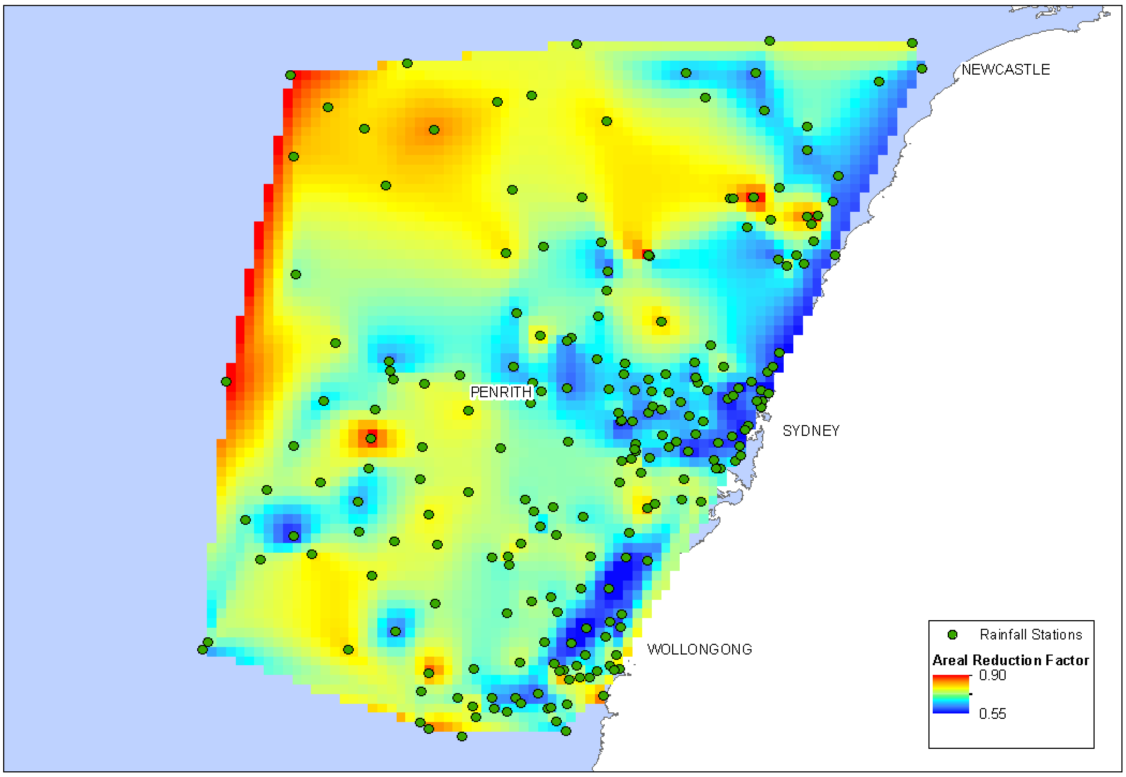


Figure 15. Spatial Variability of ARF for the 1 in 2 AEP 120 minute 500km^2

There is a large concentration of stations near the coast, and the ARF is generally lower in this area. This explains the difference in results and is also something that should be considered for future studies. The lower ARF could be caused by the high IFD gradient near the coast. It is considered worthwhile investigating a correction factor to account for this. Higher rainfall variability across a catchment would generally lead to a lower ARF, as is observed for the coastal catchments.

Using a non-overlapping set avoids some issues and biases, and would make development of generalised equations easier; however, it also means that a large portion of the available data would be discarded. Currently, the fraction of non-overlap for each circular catchment is being logged for each spatial rainfall estimate. In theory, this information could be used to calculate a weighting factor to account for the duplicate information that introduces bias towards data-dense regions. This would make it possible to use more information in the pluviograph records, compared to using a non-overlapping set of catchments. This is recommended for inclusion in future work.

3.3. Data Quality

As mentioned in the Calculating Areal Rainfall section, a few areal rainfall quality indicator parameters are logged for each observation. These are:

- number of gauges within catchment + buffer,
- number of gauges used for grid interpolation, and
- percentage overlap with other catchments.

Currently none of these variables are being utilised in the analysis, but they represent some possibilities to quality filter the output data, should there be a need. For example, using the spatial density of gauges over time, it would be possible to test for stationarity in areal rainfall. Investigating the impact of these quality indicators in more detail would be worthwhile in future work.

Another problem encountered is poor quality rainfall data, which can be very difficult to detect. The pluviograph data has been quality controlled prior to this study by BoM, but there are still poor quality data points that affect the analysis in the dataset. An example is shown in Figure 16.

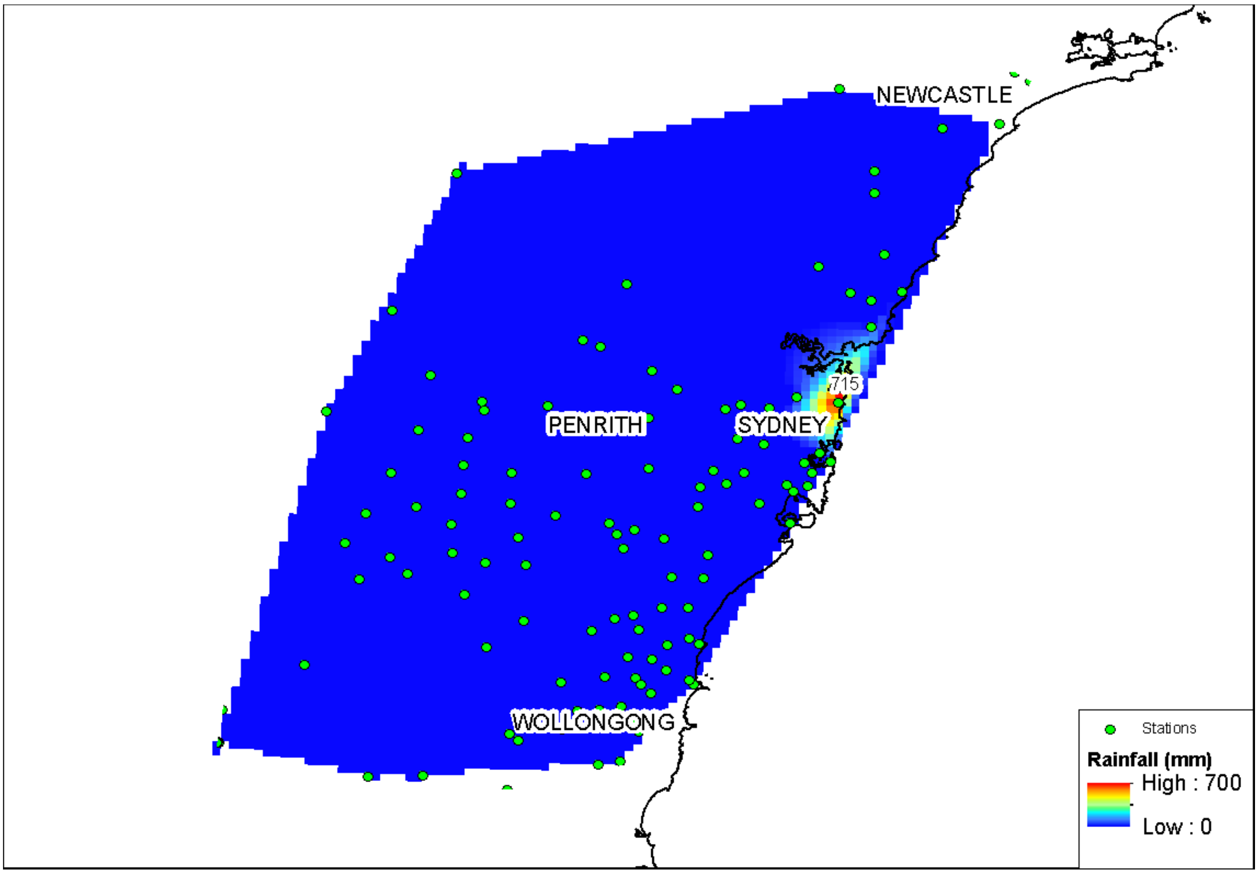


Figure 16. Poor quality data point example

In this example, one station has logged 715 mm of rainfall over a 2 hour period (in a single 5 minute interval), while there was essentially no rainfall occurring anywhere else in Greater Sydney. It seems likely that the gauge was undergoing testing. This data point illustrates the complexity of quality checking – there is no standard testing scheme, and so occasional false data points will get logged because of this. This point skews the annual maximum rainfall for every station in the vicinity of the station. Since the GEV is designed to model extreme values, it will give a large weight to a data point of this nature, potentially skewing the ARF calculation at the nearby gauges.

However, since the number of gauges available is large (250), and the final ARF is calculated from the median of each gauge ARF, the outliers created by poor quality data points will not significantly affect the outcome. It will increase the error bars when plotting the ARF, but the final

ARF used will be similar.

To make a simple attempt at quantifying the effect on final the ARF caused by poor quality rainfall data, the analysis process was run without removing the "-666" quality coded data from the rainfall database. The "-666" flag is applied to gauge data points that appear suspicious and cannot be verified reliably. The majority of the most egregious data points (such as the one shown in Figure 16) have the "-666" flag. Figure 17 shows a comparison for the 1 in 2 AEP, 3 hour duration case.

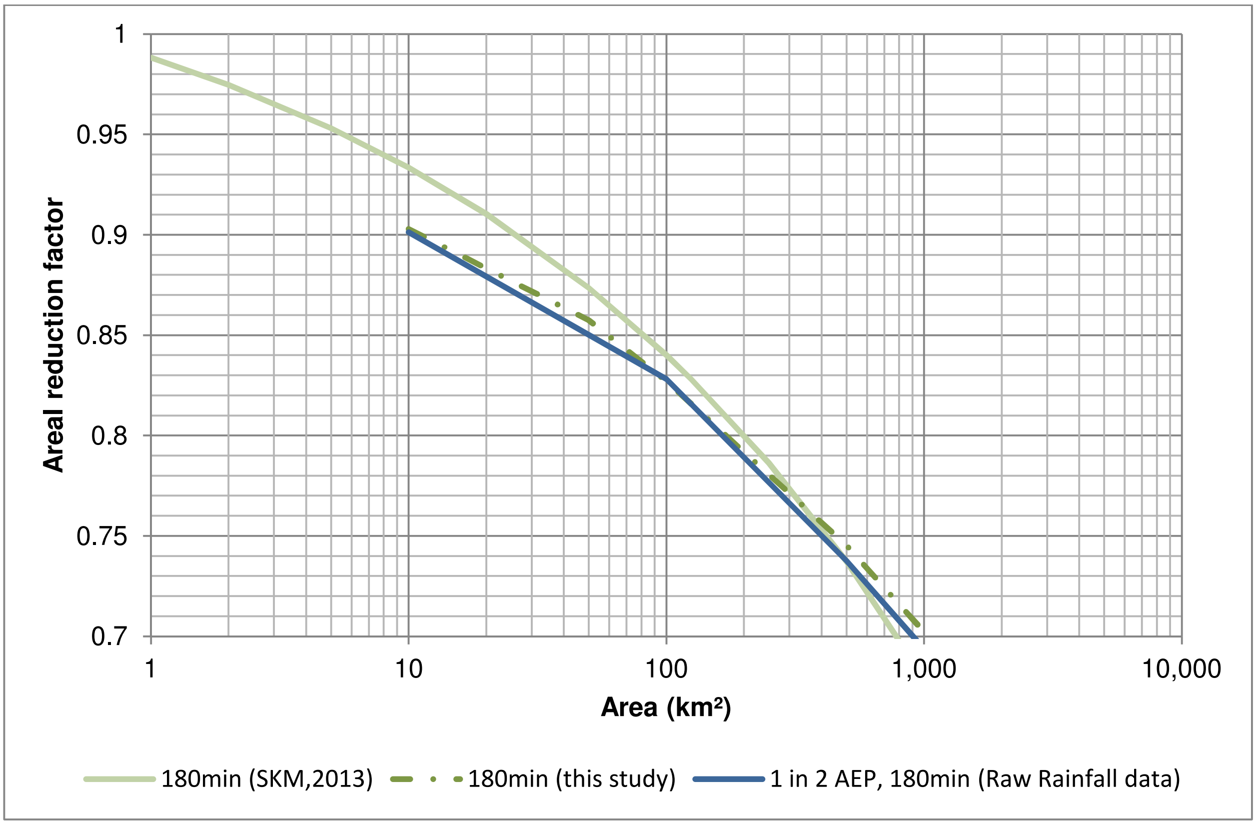


Figure 17. Effect of Raw Rainfall Data on the Final ARF

The poor quality data points do not appear to significantly influence the final median ARF for this case. The difference is less than 0.01 for all catchment sizes. It is possible to do the analysis for all cases, but at this stage only a few comparison runs were carried out, to get a sense of the effect of poor quality input data. Based on the initial results, the method appears robust against these kinds of errors. As part of future work, errors in the input rainfall data should be investigated further.

3.4. Interpolation Method Comparison

To allow for a simple comparison of grid interpolation methods, ARF calculated by using the Thiessen polygon method for interpolating grids are shown for the 60, 180, 540 minute durations and 10, 100, 1000 km^2 area cases are shown on Figure 18.

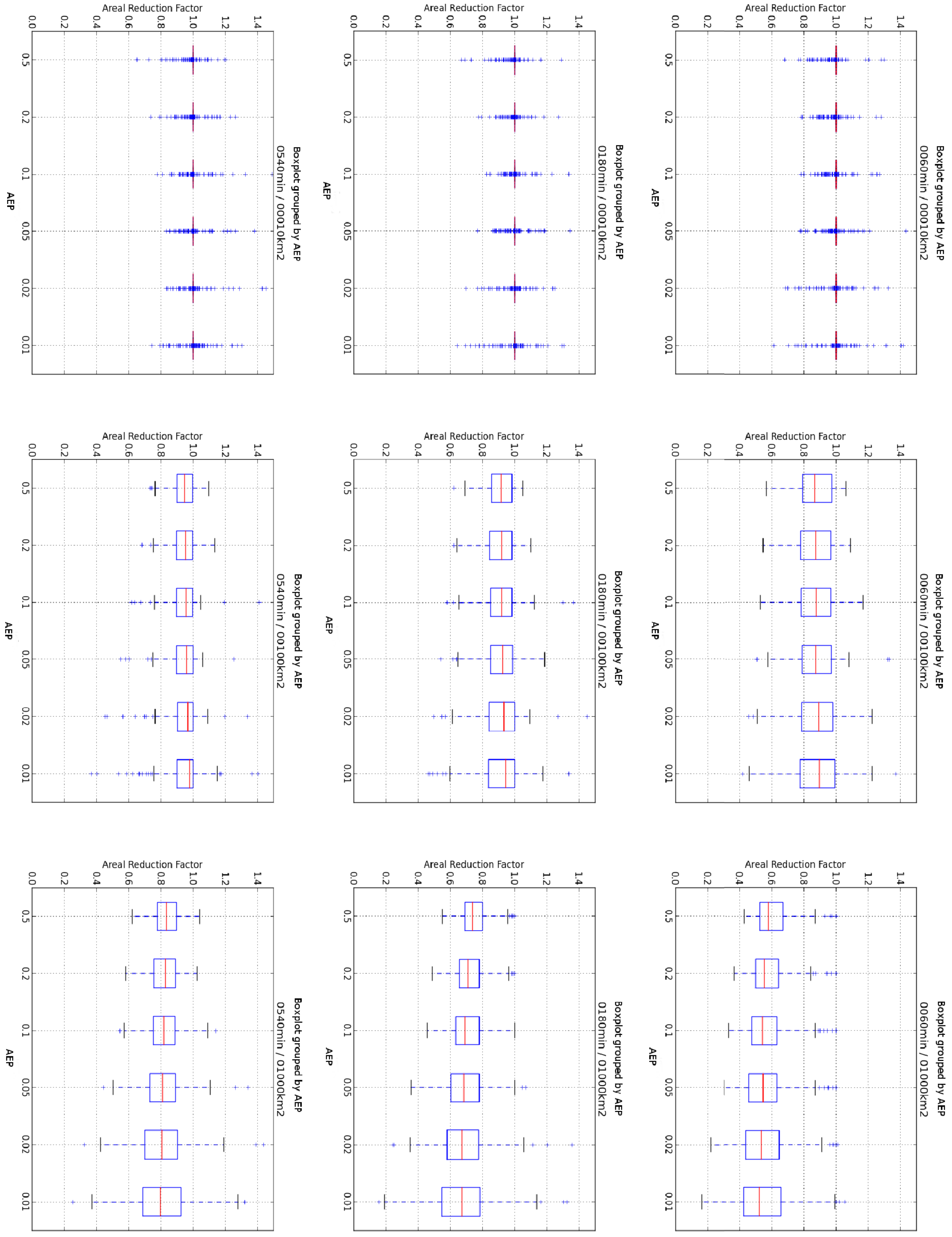


Figure 18. Thiessen Polygon ARF for 60, 180, 540 minute durations and 10, 100, 1000 km^2 cases

Figure 19 below shows the graph with the corresponding results using the natural neighbours interpolation method (presented previously in Figure 7).

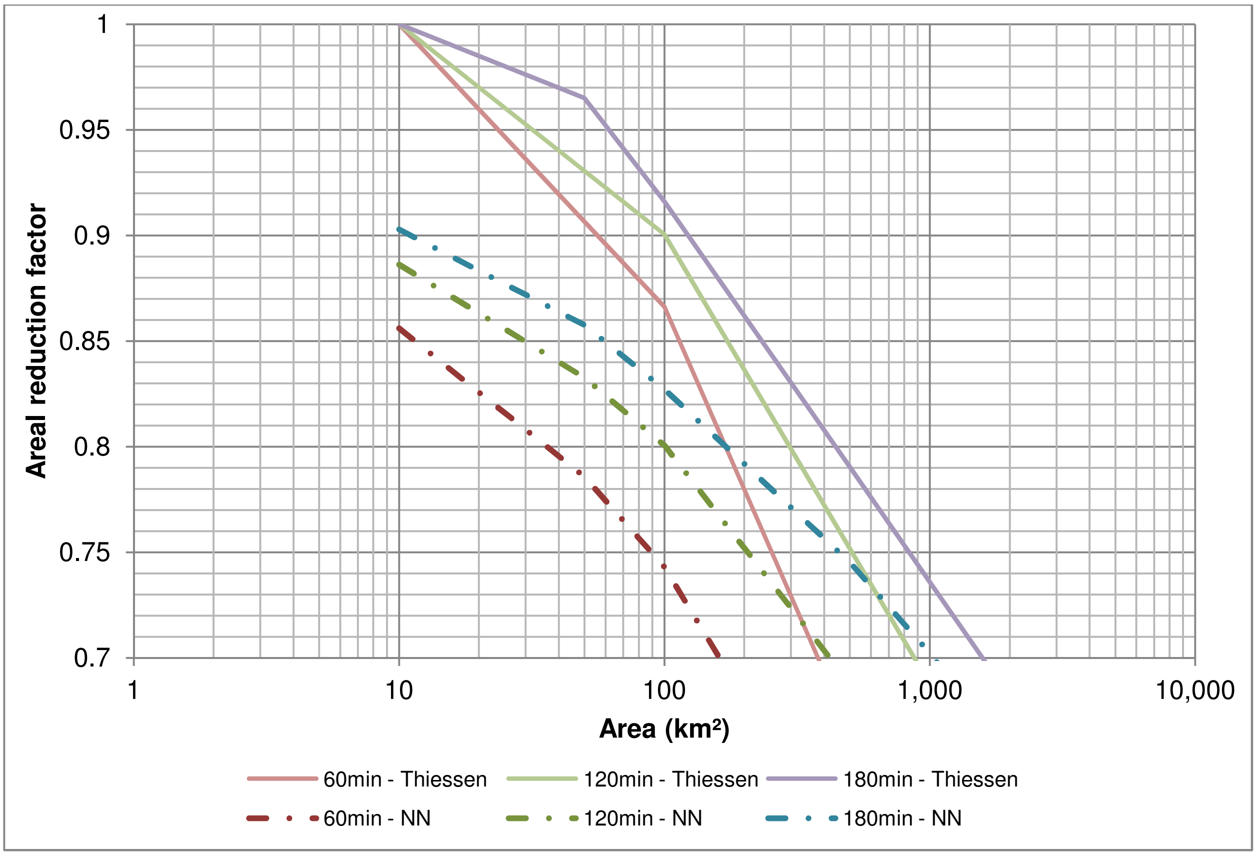


Figure 19. ARF Comparison – Thiessen Polygons and Natural Neighbours (NN) (1 in 2 AEP)

Thiessen polygons are not appropriate for catchments of 10 km^2 (Figure 19). These catchments tend to be influenced only by the nearest station, and so will often have the same areal rainfall as point rainfall. As the area gets larger, the catchment becomes more likely to intersect several Thiessen polygons, which makes for a better estimate. In general, we can see that the Thiessen polygon ARF results are generally higher than the natural neighbours results. The gauge at the centroid of the catchment will have a larger effect on the areal rainfall using Thiessen polygons, hence the areal estimate will generally be closer to the point estimate. As the area increases, the ARF estimates become more similar.

Comparing the box plots in Figure 18 to those in Figure 9 also shows generally wider 25% and 75% quantiles for the Thiessen polygon results. Using Thiessen polygons as the interpolation method, there is more variability in the ARF. In the future, it would be advisable to conduct more detailed validation of both interpolation methods, such as a Leave-One-Out (LOO) analysis. This would provide a better quantification of the quality of the ARF estimates.

4. Conclusions

From this study the following broad conclusions can be drawn:

- It is viable to use spatial interpolation as a basis for calculating ARF,
- Using this method yields similar results to the interim equations recommended by ARR Project 2 Long Duration ARF Report, and
- Using the natural neighbours method for spatial interpolation yields benefits over using Thiessen polygons.

The results of this study indicate the following properties of short duration ARF:

- Longer durations result in higher ARF,
- Larger catchment areas result in lower ARF, and
- Rarer AEP results in lower ARF.

It may be possible to generalise these results to an equation; however, this study is only an initial exploration of the viability of using the method to derive ARF. Generalising the results should be left for a later stage when the method has been more thoroughly vetted.

The next step would be to use the same method for other areas with a dense pluviograph network, such as Brisbane or Melbourne. Additionally it is recommended that parts of the method be explored further, such as the spatial variability of ARF, options for data filtering, pluviograph data quality and handling catchment overlaps.

It would also possible to apply the method on daily rainfall records for the whole country. Since daily records are much faster to process, it would be possible to include a larger area for the analysis.

5. References

Bell, FC, 1976, *The areal reduction factors in rainfall frequency estimation*, Natural Environmental Research Council (NERC), Report No. 35, Institute of Hydrology, Wallingford, UK, 25pp

Catchlove, R. and Ball, J.E. (2003), *A Hydroinformatic Approach to the Development of Areal Reduction Factors*, 28th International Hydrology and Water Resources Symposium, Wollongong, Australia, Vol 1.9-1.15

Green, J., Xuereb, K., Johnson, F., Moore, G., The, C. (2012), *The Revised Intensity-Frequency-Duration (IFD) Design Rainfall Estimates for Australia – An Overview*, 34th Hydrology and Water Resources Symposium, Sydney, Australia, pp. 808-815

Hosking, JRM and Wallis, JR, 1990, *Regional frequency analysis using L-moments,* (RC 15658), IBM Research Division, Yorktown Heights, New York
Institution of Engineers, Australia (1987), *Australian Rainfall and Runoff (ARR87)*, Vol. 1

Jordan, P., Sih, K., Hill, P.I., Nandakumar, N., Weinmann, P.E., Nathan, R.J. (2011), *Areal Reduction Factors for Estimation of Design Rainfall Intensities for New South Wales and the Australian Capital Territory*, 33rd Hydrology & Water Resources Symposium, Brisbane, Australia, pp. 218-225

Luk, K. C. and Ball, J. E., 1996, Application *of GIS modelling of the spatial distribution of rainfall*, Research Report No. 191, Water Research Laboratory, University of New South Wales, Manly Vale NSW, Australia

Natural Environmental Research Council, 1975, *Flood Studies Report, Vol. 2,* Institute of Hydrology, Wallingford, U.K.

Nittim, R., 1989, *Areal Distribution of Design Rainfall*, Hydrology and Water Resource Symposium 1989, Christchurch, New Zealand. Institution of Engineers, Australia, National Conference Publication No. 89/19: 487-488

Masters, C. J. and Irish, J. L. (1994) *Areal Reduction Factors for the Sydney Region Derived from Spatial Characteristics of Heavy Rainfall*, Hydrology and Water Resources Symposium 1994, Adelaide, Australia. Institution of Engineers, Australia, National Conference Publication No. 94/15: 27-32

Myers, V. A. and Zehr, R. M., 1980, *A methodology for point-to-area rainfall frequency ratios*, NOAA Tech, Report NWS 24, Washington D.C.

Porter, J. W. and Ladson, A. R., 1993, *An application of fixed-area rainfall reduction factors*, Hydrology and Water Resources Symposium 1993, Newcastle, Australia. Institution of

Engineers, Australia, National Conference Publication No. 93/14: 273-279

Richards, Z, Sharma, A, Babister, M, 2012, *A Methodology for Incorporating Orographic Information in Deriving Intensity-Frequency-Duration Relationships*, 34th Hydrology and Water Resources Symposium, Sydney, 19 – 22 November, 2012

Sibson, R. (1981), *A Brief Description of Natural Neighbor Interpolation, in "Interpolating Multivariate Data"*, John Wiley & Sons, New York, pp. 21-36

Siriwardena, L. and Weinmann, P. E. (1996) *Derivation of Areal Reduction Factors for Design Rainfalls in Victoria*, Cooperative Research Centre for Catchment Hydrology, Report 94/4

SKM (2013), *Australian Rainfall and Runoff Revision Project 2: Spatial patterns of Rainfall: Collation and Review of Areal Reduction Factors from Applications of the CRC-FORGE Method in Australia*, Institution of Engineers, Australia

Srikanthan, R. (1995) *A review of the methods for estimating areal reduction factors for design rainfalls*, CRC for Catchment Hydrology, Report 93/3

Svensson, C. and Jones, D. A., 2010, *Review of methods for deriving areal reduction factors*, J. Flood Risk Management, 3 (3): 232-245

Umakhanthan (2002), K., *Estimation of the Spatio-Temporal Heterogeneity of Rainfall and its Importance towards Robust Catchment Simulation, within a Hydroinformatic Environment*, Doctoral Dissertation, School of Civil and Environmental Engineering, University of New South Wales, Sydney, Australia
Wang, QJ, 1997, LH-*moments for statistical analysis of extreme events,* Water Resources Research, Vol. 33(12): 2841-48