THE APPLICABILITY OF THE USE OF RADAR DATA TO DEVELOP AREAL REDUCTION FACTORS IN SOUTH AFRICA

Report to the

WATER RESEARCH COMMISSION

by

JA DU PLESSIS, RP BURGER, OJ GERICKE, TJ THOMAS and TS MAREBANE

Stellenbosch University

WRC Report No. 2923/1/20

ISBN 978-0-6392-0195-5

October 2020

Obtainable from

Water Research Commission

Private Bag X03

GEZINA, 0031

orders@wrc.org.z or download from www.wrc.org.za

EXECUTIVE SUMMARY

The main objective of this research is to establish if radar data can be used effectively in the hydrological field and specifically for the development of Areal Reduction Factors (ARFs) in South Africa. Radar data was sourced and analysed to derive ARFs and in doing so, the applicability of the use of the radar data was assessed. It is important to note that the aim is not to derive new ARFs, but the process developed in this research to deriving ARFs from radar data, only serves as a means to assess the applicability of radar data.

An ARF can be defined as a factor that is applied to convert point rainfall depths/intensities to an average rainfall depth/intensity over a specific catchment area. The concept of ARFs provide a powerful mechanism for the analyses of the spatial variability of various hydrological processes. However, a number of empirical methods used to derive ARFs are dependent on the rainfall data collected at rainfall stations. With the decline in the density and number of reliable rainfall stations, radar data has been viewed as an alternative for applications in hydrology and consequently ARFs. Radar data appear to be more efficient than using rainfall station networks, since radar data can capture the internal and spatial distribution of a rainfall event in significant more detail and provide data at a significant larger scale in terms of the spatial distribution of rainfall events.

An extensive review on the 'state' of ARFs was carried out on an international scale. Various factors that influence ARFs, such as catchment and rainfall characteristics, convective and frontal rainfall regions, return period and spatial variability of rainfall, were discussed. Some studies, specifically Bell (1976), suggested that return period also has an influence on ARFs. Various methods of deriving ARFs were researched and it was found that ARF derivations generally either follow a geographically-fixed or a storm-centred approach.

The use of radar data within hydrology was also reviewed. For South Africa, using radar data is considered valuable as insight to rain producing systems, as well as its impact on spatial and temporal resolution when estimating catchment areal rainfall, becomes available. Considering this, AFRs are sensitive to spatial variability and temporal characteristics within its derivation. This research is therefore focused on the storm-centred approach using high spatial and temporal resolution derived rainfall obtained from radar imagery.

The availability of radar data proved to be a critical aspect of this research. While various sources of data have been identified, radar data from only the South African Weather Services (SAWS) and the North West University (NWU) were available for use. SAWS provided composed radar imagery for 1-, 3- and 24-hour storm durations, while NWU provided 6-minute time step data, which was used to compile 1-, 3- and 24-h storm duration rainfall, both in Meteorological Data Volume (MDV) file format. The MDV files were converted to netcdf files for analysis in ArcGIS and the associated software to aid the conversion, were developed in the process. An extensive methodology was put in place, following the storm-centred method, to derive ARFs from these netcdf files using ArcGIS.

To evaluate if the radar data can be used in hydrology, the ARFs, based on the radar data and the methodology developed, were calculated, and compared to the ARFs currently in use in South Africa. It is noted that the existing ARFs in use have been developed in the 1960's with limited data available and that various researchers already highlighted the need for an upgrade

of these ARFs. Due to the limited available radar data set, it was not possible to do any analysis relating to the probability of occurrence (return period) of any event.

The radar data derived ARFs for 1-, 3- and 24-hour storm durations were analysed first. The analysis concluded that:

- An inverse proportional relationship exists between the ARF and storm area, as well as the maximum point rainfall intensity and the maximum point rainfall, as is evident in the existing methods available for use.
- Higher ARFs were obtained for convective rainfall regions than for frontal rainfall regions, which is contradictory to existing results. The limited number of storms (coastal) selected may be partially responsible for this anomaly.
- The influence of storm duration on ARFs did not yield consistent results. It is speculated that the uneven distribution of storm areas within this analysis resulted in possible inconsistencies. An analysis, taking storm area into consideration, proved only partially successful, with storm areas less than 1 000 km^2 following an expected trend of lower ARF's with an increase in duration, but the opposite was observed for storm areas larger than 1 000 km^2.

The derived radar based ARFs were then compared to the results from existing methods in use in South Africa. The analysis concluded that:

- Radar derived ARFs were generally all lower than ARFs derived using the diagrams developed by Van Wyk (1965) or the Op ten Noort and Stephenson's (1982) equation for small catchment areas (< 800 km^2).
- When taking storm area categories into consideration, radar derived ARFs for storm areas less than 800 km^2 for smaller catchments were all, except for a very limited number of storms in the 600-800 km^2 range, also found to be lower than ARFs presently in use in South Africa, as suggested by Van Wyk (1965).
- For large catchment areas (>800 km^2) radar derived ARFs were generally lower than ARFs derived using the diagrams developed by Pullen, Wiederhold and Midgley (1966) as well as the Op ten Noort and Stephenson's (1982) equation.
- When comparing ARFs taking storm duration into consideration for larger catchments areas, radar derived ARFs were lower than ARFs presently in use in South Africa for storm durations of 3- and 24-h. The radar derived ARFs for the 1-hour storm events were found to be overestimated. Using Op ten Noort and Stephenson's (1982) equation produced ARFs that were, on average, significantly higher than the radar derived ARFs.

It can be concluded from the findings highlighted above, that the radar derived ARFs indeed mostly produce findings that is in line with well documented trends, but that ARFs derived from radar data typically provided lower ARF values, indicating that existing methods do provide conservative results.

Radars provide more detailed information for the analyses of storms when compared to using traditional analysis methods. Although radars are known for their large uncertainties, radars expose the potential of significantly improving certain aspects of flood hydrological calculations.

The most significant challenge faced during the research was the accessibility of radar data. The acquisition of quality data for research and analysis is a requirement to understand various phenomena. The channels used to obtain data from credible and reliable institutions/providers to ensure calibration and verification standards are maintained and documented, is of critical importance. It is clear that special attention need to be given to the process to obtain reliable radar data and high level agreements will have to be set in place specifically between the SAWS and research institutions to enable access to radar data, which is considered as a valuable commodity, with a high monetary value attached to it, by the SAWS.

It is also evident that the pool of expertise to be used in the collection, cleaning and post processing of the radar data into a format that can be used for researchers, need to be expand. While various radar data sets might be available, the processing of the data is time consuming which requires some level of skill with the Thunderstorm Identification, Tracking, Analysis and Nowcasting (TITAN) software and netcdf data format to ensure quality data conversion takes place. The process to convert multiple years of MDV data also requires high performance computing power. Some of the practical problems identified during the data processing include:

- Anomalous rainfall reflectivity values that are not consistent with surrounding rainfall values.
- Inconsistent work environments to convert radar data to a format that is compatible in a Geographical Information Systems (GIS) environment can become tedious. For example: Some radar files are processed using a Linux system. The file naming convention on the Linux system is not the same as other operating systems. A Linux system then becomes a crucial part of the process for file name changing.
- Extreme storm events are frequently not recorded by the radar as it is typically non-operational on those days.

Since this research used two different sources of radar data, it was also important to understand the challenges associated with the use of two different data sets. Differences are due to the difference in post-analysis applied by different service providers. For this purpose, ARFs were derived, using the data set from SAWS and NWU separately, for different storms taking maximum point rainfall intensity, storm area, maximum point rainfall and storm duration into consideration. No significant difference could be found when comparing the results of the ARFs as derived based on the SAWS and the NWU radar data sets.

This research proves that radar data can be used successfully for the derivation of ARFs in South Africa. It can be concluded that radar data has the potential to contribute significantly to the field of hydrology and that further research needs to be done to strengthen this potential.

It is recommended that an agreement be drafted between the Water Research Commission and the SAWS to provide resources at a national level to process the available radar data for future hydrological research. Once the availability of data from different sources (specifically the SAWS) has been resolved, it is proposed that further research be done on specifically the selection of appropriate storm duration selection procedures, the selection of more appropriate categories of storms (i.e. coastal vs inland) as well as storms with a higher intensity. Future research also needs to focus on appropriate platforms from where raw radar data (after initial post-analysis by owners) can be accessed as appose to the transfer of huge volumes of data.

ACKNOWLEDGEMENTS

The research in this report emanated from a project funded by the Water Research Commission and entitled: 'The applicability of the use of radar data to develop areal reduction factors in South Africa.'

The Reference Group members are thanked for their input, comments and technical assistance, and these members include:

Mr W Nomquphu (Chair)	Water Research Commission
Dr T Ndarana	Council for Scientific and Industrial Research
Dr CJ Engelbrecht	South African Weather Services
Prof JC Smithers	University of KwaZulu-Natal
Dr WP de Clercq	University of Stellenbosch
Ms KT Chetty	University of KwaZulu-Natal

The researchers responsible for this project consist of the following members:

Prof JA Du Plessis	University of Stellenbosch (Project Leader)
Dr RP Burger	North-West University, Potchefstroom Campus (Researcher)
Dr OJ Gericke	Central University of Technology, Free State (Researcher)
Mr TJ Thomas	University of Stellenbosch (Student)
Mr TS Marebane	North-West University, Potchefstroom Campus (Student)

Special acknowledgements and thanks to:

(a) Mr K de Waal, the original author of the radar file conversion application, for his support;

(b) South African Weather Services for providing radar data, specifically Me Elsa De Jager (now retired); and

(c) Our families and colleagues, for their patience and understanding throughout this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AEP	Annual Exceedance Probability
AMS	Annual Maximum Series
AP	Anomalous Propagation
ARF	Areal Reduction Factor
ARI	Average Recurrence Interval
AWS	Automatic Weather Station
CSIR	Council for Scientific and Industrial Research
DWS	Department of Water and Sanitation
DDF	Depth-duration-frequency
DWAF	Department of Water Affairs and Forestry
DWS	Department of Water and Sanitation
EV1	Extreme Value Type I
GEV	General Extreme Value
GEV/LM	General Extreme Value using Linear Moments
GEV/PWM	General Extreme Value using Probable Weighted Moments
GIS	Geographical Information System
HRU	Hydrological Research Unit
LM	Linear Moments
LN	Log-Normal
LP3	Log-Pearson Type III
LP3/MM	Log-Pearson Type III using Method of Moments
MAP	Mean Annual Precipitation
MDV	Meteorological Data Volume
NEXRAD	Next Generation Radar
NCAR	National Center for Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Services
PDS	Partial Duration Series
PMP	Probable Maximum Precipitation
PWM	Probability Weighted Moments

RAL	Research Application Laboratory
RSA	Republic of South Africa
RDAS	Radar Data Acquisition System
SANCOLD	South African National Committee on Large Dams
SANRAL	South African National Roads Agency Limited
SAWB	South African Weather Bureau
SAWS	South African Weather Services
T_C	Time of concentration
TITAN	Thunderstorm Identification, Tracking, Analysis and Nowcasting
TP	Technical Paper
TP-29	Technical Paper - 29
USA	United States of America
UK FSR	United Kingdom Flood Studies Report
WRC	Water Research Commission

1 INTRODUCTION

1.1 Background

Areal Reduction Factors (ARFs) are used for the development and formulation of design areal rainfalls, which are essential for the design and planning of hydraulic structures. ARFs are essential in the design of hydrological extremes and form part of key functions of storm characteristics such as size and shape. ARFs have been mostly developed in the United States of America, United Kingdom and New Zealand. The use of ARFs are convenient as the networks of rainfall stations with long rainfall records are typically sparse and do not account for appropriate characterisation of associated spatial rainfall patterns (Svensson and Jones, 2010).

An ARF is typically defined as a factor that is applied to convert point rainfall depths/intensities to an average rainfall depth/intensity over a specific catchment area (Gill, 2005). The two main types of ARFs are derived either by considering a specific geographical location (referred to as the "fixed-area" method) or a storm event (referred to as the "storm-centred" method). The fixed-area method consists of analysing rainfall data at a specific geographical location and contains more statistical significance than the storm-centred method. On the other hand, the storm-centred method typically focuses more on analysing aspects of a specific storm event, irrelevant of locality.

Radar data appear to be more efficient than using rainfall station networks as radar data is able to capture the internal and spatial distribution of a rainfall event. Moreover, rainfall station networks produce poor spatial characteristics of a rainfall event, as it is highly dependent on the density of the network. A major challenge with using radar data is using the data to arrive at accurate rainfall values. An abundance of limitations exists in the conversion process and differs from radar to radar data set. These limitations can include radar reflectivity calibration, high estimates due to frozen and wet frozen precipitation and partial beam filling that generally results in signal degradation which leads to rainfall rates being reduced.

The use of radar data in South Africa has only been used to a limited extent. Currently, there is a significant amount of uncertainty with the use of radar data in the hydrological research field, not only due the problems associated with the production of many inaccuracies associated with the process, but also due to cost, technical infrastructure and topography.

1.2 Problem Statement

Currently, the ARFs used in engineering applications in South Africa to arrive at design rainfalls were mainly produced in 1965 (ARFs for small areas) and 1969 (AFRs for larger areas). Although these ARFs were used to calculate many areal design rainfalls, they are now considered outdated as a significant amount of geomorphology, historical rainfall patterns and behavioural changes have since occurred. A significant amount of additional information is also now available. With this, an evolution of technology has occurred, and weather data capturing is currently more convenient and accessible than before. With traditional rainfall station data under pressure due to a lack of maintenance, radar data potentially could act as a suitable means to obtain areal rainfall data to estimate ARFs. Furthermore, of the two methods to obtain ARFs currently, a storm-centred approach would prevail as rainfall stations tend to exclude the spatial characteristics of a specific storm event.

1.3 Research Objectives

The main objective of this research was to investigate the potential to develop ARFs using radar data provided by the SAWS. These objectives were achieved by addressing the following sub-objectives:

a) Obtain estimates of rainfall values derived from raw radar data;
b) Identify potential obstacles that could hinder the use of radar data;
c) Use radar rainfall data to establish storm-centred depth area relationships for convective and frontal rainfall regions in South Africa;
d) Calculate ARFs using radar data from a reliable radar(s) in South Africa;
e) Compare the ARFs calculated from radar data to the ARFs currently used in South Africa;
f) Explore various aspects that could affect ARFs, namely storm size and shape and climate change, with radar data;
g) Expose the potential of the use of radar data, specifically for the hydrological environment.

1.4 Limitations of Research

The scope of this research was limited by the following:

- The availability of radars;
- The historical record of radar data available for analysis;
- The location of the radars; convective region and frontal regions;
- Type of storm events captured by radars for analysis;
- The available different storm duration (1-h, 3-h and 24-h) data as captured by the SAWS.

2 LITERATURE REVIEW

2.1 Introduction

This section discusses design rainfall as the main input parameter towards the calculation of flood events for the design of associated hydraulic structures. In deriving an appropriate design rainfall, the development of ARFs and various elements and factors that influence ARFs, are discussed. The methods of deriving ARFs in South Africa are presented in detail. Thereafter, the use of radar data in hydrology in general and specifically in relation to storm rainfall, is discussed. Finally, the international use of radar imagery, with multiple considerations, is explained, specifically for the derivation of ARFs.

2.2 Design Rainfall Estimation

Design rainfall comprises of a depth and duration associated with a given return period or annual exceedance probability (AEP) (Smithers and Schulze, 2004). Short and long duration design rainfall estimations can either be based on point or regionalised data. Rainfall durations less than 24 hours are generally classified as short, while long durations typically range from 1 to 7 days (Smithers and Schulze, 2004). Several regional and national scale studies in South Africa based on short durations and point data were conducted between 1945 and 2001. Studies focusing on long durations based on daily point rainfall data included studies done by the SAWB (South African Weather Bureau), Schulze (1980), Adamson (1981), Pegram and Adamson (1988) and Smithers and Schulze (2000b). Smithers and Schulze (2000a; 2000b) also used a regionalised approach to increase the reliability of the design values at gauged sites, as well as for the estimation of design values at ungauged sites (Smithers and Schulze, 2003).

2.2.1 Single site approach

A single site approach requires that each rainfall station within the relevant catchment be investigated to determine the record length, data quality (errors, missing data and outliers) and topographical position (Smithers and Schulze, 2000a). In order to develop the depth-duration-frequency (DDF) relationship at every single site, the following steps are of importance (Smithers and Schulze, 2000a):

(a) Selection of the most appropriate data set. This may either be the annual maximum series (AMS) or partial duration series (PDS) with a sufficient record length;
(b) Selection of the most appropriate probability distribution; and
(c) Selection of a suitable parameter and quantile method.

A probabilistic analysis needs to be conducted at each rainfall station and it is thus advisable not to use rainfall stations with short record lengths. Furthermore, it is impossible to conclusively select a distribution that could consistently provide adequate rainfall frequency estimates for return periods greater than the period of record. On the other hand, small samples may define a distribution which is markedly different from the parent population (Smithers and Schulze, 2000a).

According to Viessman, Lewis and Knapp (1989), a minimum record length of 10 years is required, while Schulze (1984) questioned the significance of the record length for extreme events recorded and hence the design values. Hogg (1992) demonstrated that even 20 years of data are not stable enough to estimate the 10-year return period event. Hogg (1992) indicated that the assumptions of stationarity and homogeneity of the AMS of rainfall are seldom valid. It is suggested that a regional approach be used to improve the frequency analysis of extreme rainfall events.

According to Weddepohl (1988), the malfunctioning of rainfall gauges and processing errors are inherent in rainfall data. The spatial density and distribution of rainfall gauges, sporadic rainfall events as opposed to the continuous digitised data in use, the length of available records and the presence of outliers are all problems associated with these errors (Weddepohl, 1988).

The selection of the most suitable probability distribution resembling the probability distribution of the population must be made according to the theoretical basis, consistency, acceptance, user-friendliness and applicability thereof (Cunnane, 1989; cited by Smithers and Schulze, 2000a). This selection is particularly important when estimating extreme events with return periods greater than the length of record. Equally important are that, factors such as the type of data in use, data stationarity and the method of fitting the distribution, should also be considered (Cunnane, 1989; cited by Smithers and Schulze, 2000a).

The Extreme Value Type I (EV1) distribution has been extensively used in rainfall DDF studies in South Africa since 1963, while the use of the integrated General Extreme Value (GEV) distribution is growing in the application of frequency analysis. Van der Spuy and Rademeyer (2018) propose the use of the Log-Normal (LN), Log-Pearson Type III (LP3) as well as GEV using the Method of Moments (MM), Probable Weighted Moments (PWM) or Linear Moments (LM) to estimate the required design rainfall depths in South Africa.

The Technical Report 102 (TR102; Adamson, 1981) is an example of a design point rainfall database based on a single-site approach and is commonly used in South Africa. Adamson (1981) estimated the 1, 2, 3 and 7-day extreme design point rainfall depths for return periods of 2, 5, 10, 20, 50, 100 and 200 years using approximately 1 946 rainfall stations. A censored LN distribution based on the PDS was used to estimate the design point rainfall depths at a single site.

2.2.2 Regional approach

Regional frequency analysis is based on the assumption that the standardised variate distributions of rainfall data are similar at every single site in a region and that the data from various single sites in a region can thus be combined to generate a single regional rainfall frequency curve representative of any site in the specific region with appropriate site-specific scaling. An advantage of this approach is that it can be used to estimate events at ungauged sites where no rainfall data exists (Alexander, 2001; Cunnane, 1989; cited by Smithers and Schulze, 2003). In nearly all practical situations, a regional approach is preferred to a single site approach primarily based on the efficiency and accuracy of the rainfall quantile estimation and where statistical homogeneity or heterogeneity might exist (Hosking and Wallis, 1997; cited by

Smithers and Schulze, 2003). The large degree of uncertainty introduced in the extrapolation of AEPs beyond the record length of data can also be reduced by regionalisation, since the observed rainfall at a single site is then related to the hydrological response at a regional scale by making use of an extended or combined record length of data (Smithers and Schulze, 2003).

In considering the limitations of a single-site approach and the paucity of sub-daily rainfall data in South Africa, *i.e.*, 412 sub-daily rainfall stations and only 49 of these rainfall stations having record lengths exceeding 30 years, Smithers and Schulze (2000a; 2000b; 2003; 2004) developed a regional scale invariance approach to estimate the mean point rainfall AMS for any duration and associated 'scaling factors' as an alternative for the 'conversion factors' proposed by Adamson (1981). These 24-hour to 1-day continuous rainfall measurement 'scaling factors' range between 1.14 and 1.30 in South Africa (Smithers and Schulze, 2003).

Smithers and Schulze (2003; 2004) established 78 homogeneous long duration rainfall clusters, 15 short duration rainfall clusters, and estimated index values (mean *n*-hour AMS values) derived from at-site data. Cluster analysis of site characteristics was used to group the 78 long duration rainfall clusters into 7 regions with 6 associated region-specific regression parameters. Firstly, the mean of the 1-day fixed time interval point rainfall AMS was estimated using regional regression relationships. Thereafter, the mean of the 24-hour continuously recorded point rainfall AMS was estimated directly from the 1-day value for the specific site under consideration. Lastly, the mean of the point rainfall AMS values for durations shorter and longer than 1 day were scaled directly from the mean of the continuous 24-hour and 1-day values, respectively, using the established regression parameters. The up- and downscaling were found to scale linearly as a function of the mean 1-day and continuous 24-hour values, respectively. In the application of the regression relationships to estimate the mean of the AMS for durations shorter and longer than 1 day, inconsistencies in the growth curves derived from the 24-hour continuously recorded and daily rainfall data were evident due to the quality and non-concurrent periods of the digitised rainfall data, as well the differences in the AMS extracted from: (i) continuously recorded data using a sliding window, and (ii) daily rainfall data using a fixed period window.

As a result, a scale invariance approach was introduced to the Regional Linear Moment Algorithm and termed the RLMA&SI approach to address the inconsistencies evident in the above-mentioned growth curves (Smithers and Schulze, 2003). In South Africa, the RLMA&SI approach is the preferred method for design rainfall estimation and is automated and included in the software program, *Design Rainfall Estimation in South Africa* (Smithers and Schulze, 2003; 2004). The latter software facilitates the estimation of design rainfall depths at a spatial resolution of 1-arc minute, for any location in South Africa, for durations ranging from 5 minutes to 7 days and for return periods of 2 to 200 years.

Irrespective of whether a single site or regional approach is adopted, the design rainfall depth to be used in design flood estimation, especially in the deterministic methods, must be based on the critical storm duration or time of concentration (T_C) of a catchment. Thus, depending on the T_C, the daily design rainfall depth used in flood estimations must either be increased or decreased. In order to convert the daily design rainfall depth values to independent durations of the same length, conversion and/or scaling factors have to be used. The conversion factors are dependent on the duration in question and various values have been proposed.

The use of conversion factors (Adamson, 1981) is generally accepted in South Africa to convert 1-day fixed time interval rainfall (08h00 to 08h00) to continuous measures of *n*-hour rainfall associated with T_C. Adamson (1981) proposed the use of a conversion factor of 1.11 to convert daily rainfall depths recorded at fixed 1-day intervals to continuous 24-hour rainfall depths. At an international level, similar conversion factors have been proposed to convert daily fixed time interval rainfall depths to continuous 24-hour maxima, *e.g.* 1.13 in the USA (Hershfield, 1962), 1.06 in the UK (NERC, 1975), and 1.13 in South Africa (Alexander, 1978). In order to convert continuous 24-hour rainfall series to critical storm or T_C durations ranging between 0.10 hour and 24 hours, Adamson (1981) proposed the use of the conversion factors as listed in Table 2.1. The conversion factors listed in Table 2.1 are considered to be independent of return period, but are influenced by regional climatological differences as evident in the summer rainfall/inland and winter rainfall/coastal regions of South Africa (Midgley and Pitman, 1978).

Converting daily rainfall depths to durations longer than 1-day simply entails the conversion of fixed time interval rainfall to continuous measures of rainfall (e.g., 2-days to 48-hour, 3 days to 72-hour, etc.), and interpolating between the different T_C durations as listed in Table 2.2.

The conversion factors listed in Table 2.2 are normally used in practice (Van der Spuy and Rademeyer, 2018); however, no literature is available as to how these conversion factors were derived.

Table 2.1: Conversion of continuous 24-hour rainfall depths to T_C-hour rainfall depths (Adamson, 1981)

T_C (hours)	Conversion factor Summer/inland region	Conversion factor Winter/coastal region
0.10	0.17	0.14
0.25	0.32	0.23
0.50	0.46	0.32
1	0.60	0.41
2	0.72	0.53
3	0.78	0.60
4	0.82	0.67
5	0.84	0.71
6	0.87	0.75
8	0.90	0.81
10	0.92	0.85
12	0.94	0.89
18	0.98	0.96
24	1.00	1.00

Table 2.2: Conversion of fixed time interval rainfall to continuous estimates of *n*-hour rainfall (Van der Spuy and Rademeyer, 2018)

Duration		Conversion factor
From (days)	To (hours)	
1	24	1.11
2	48	1.07
3	72	1.05
4	96	1.04
5	120	1.03
7	168	1.02
> 7	> 168	1

However, the latter South African approaches as listed in Table 2.1 and Table 2.2 are regarded as outdated and Smithers and Schulze (2000a) developed regionalised relationships for 15 relatively homogeneous short duration rainfall clusters in South Africa, with a national average of 1.21.

2.3 Areal Reduction Factors

ARFs have many definitions that allude to calculating the same intrinsic value. This results from the calculation of the ARF incorporating either the return period, catchment area or storm duration of a particular rainfall event; or a combination of these three characteristics. Thus, a basic definition of an ARF can be defined as a factor that is applied to convert point rainfall depths to an average rainfall depth over a specific catchment area. The concept of ARFs provides a powerful mechanism for the analyses of the spatial variability of the various hydrological processes (Gill, 2005)

The design of hydraulic structures requires the knowledge of the amount of rainfall that is likely to fall within a certain amount of time, over a specific area. Furthermore, ARFs are central to conventional flood risk assessment. Errors and inaccurate estimations of ARFs has the ability to produce large errors in subsequent estimates of design rainfall and discharge, specifically for flood frequency analyses. Point rainfall is only representative for a limited area, whereas for larger areas, the average areal rainfall depth is highly probable to be smaller than the maximum observed point rainfall depth (Svensson and Jones, 2010). ARFs has been found to vary with predominant weather types, season and return period and thus, provide guidance for catchments with minimal or insufficient spatial rainfall station density or inadequate historical rainfall data. Moreover, ARFs allow for spatial smoothing of sampling variations and facilitate the development of regional engineering guidelines.

There are several methods for deriving ARFs which can be categorised into two groups: empirical and statistical methods. Traditionally, ARF estimates are based on empirical methods; however, a range of analytical methods have been researched and applied more recently (Svensson and Jones, 2010). Empirical methods have the tendency to disregard the influence of return period on the value of an ARF. That is, until Svensson and Jones (2010) conducted a review study that clearly show the influence of return period on the derivation of ARFs. In addition, Bell (1976) derived ARFs from rainfall frequency curves and found that ARFs do in fact vary with return period. For South Africa, Du Plessis and Loots (2019) re-evaluated the ARFs in use in South

Africa, using 19 test sites well spread over South Africa, adding approximately 20 years of data to the previous set used during the development of ARF for South Africa by various researchers. They concluded, based on the test sites that return period indeed also have an impact on ARFs in South Africa.

Incorrect ARFs have the possibility of resulting in major errors in estimates of the intensity of areal rainfall, including the development of design storms. This directly impacts on subsequent flood risk estimates (Wright, Smith and Baeck, 2014). ARFs should at least represent the rudimentary properties of an observed storm structure, along with its variability, based on the premise that flooding is not the end product of idealised design storms, but instead, highly complex meteorological systems. Thus, ARFs should at least address the basic properties of observed storm structure and variability.

2.3.1 Factors influencing Areal Reduction Factors

Various factors affect the ratio between the point rainfall in a catchment area and the spatial areal rainfall over that area. These issues include factors that relate to the characteristics of rainfall itself as well as to the physical characteristics of the catchment area, including the data and methods used to arrive at an ARF.

2.3.1.1 Catchment characteristics

Most research conducted on the estimation of ARFs concluded that catchment geomorphology (e.g., area, shape and topography) has an insignificant influence on ARFs (Svensson and Jones, 2010). In catchments with areas less than 800 km², ARFs are mainly a function of the area and point intensity, since the relationship between rainfall intensity and the infiltration rate of the soil is predominant. In catchments with areas of up to 30 000 km², ARFs are mainly a function of the area and storm duration (Alexander, 2001; SANRAL, 2013). Lambourne and Stephenson (1986) demonstrated that the ARF will decrease from unity with an increasing catchment area.

Elongated catchment shapes tend to result in variable ARFs, with a dependency on typical rainfall isohyets that are aligned along the catchment or perpendicular to the catchment. Veneziano and Langousis (2005) investigated rainfall fields and deriving ARFs from a theoretical multifractal perspective and concluded that the influence of catchment shape was generally small. In addition, they noted that highly elongated catchments are rare in hydrology.

Windward and leeward effects of mountainous and hilly regions potentially have an influence on ARFs. For the calculation of areal rainfall, which is essential for the derivation of ARFs, using Thiessen polygons with inverse distance weighting methods are not true representatives of the topography being examined. This comes as a potential challenge at higher elevations where the network of rainfall stations tend to be less dense (Prudhomme, 1999). With this new knowledge at the time, Allen and DeGaetano (2005) developed a topographical bias adjustment factor to combat the inaccuracy that comes with rainfall interpolation procedures in mountainous regions. This bias acted as a means of modifying the areal rainfall values that were given by the interpolation procedures. Their research concluded that the biases appeared to be insignificant for the derivation of ARFs.

A study conducted by Huff (1995), noted that there was a potential difference in ARFs between urban areas and its surrounding rural areas. A number of storms in Chicago, the study area, were found to have a slower rate of decrease in the derivation of ARFs within a 500 km^2 radius of the urban storm centre in comparison with several rural storms. For larger areas, it was found that the rate of decrease for urban storms exceeded that for rural storms. It is important to note that the sample size of storms for the study was relatively small, and the natural variability in spatial rainfall characteristics is generally larger. Thus, Huff (1995) concluded that this anomaly could potentially be a result of the natural variability, instead of the urban rainfall effect.

2.3.1.2 Rainfall characteristics

According to Huff and Shipp (2002) various synoptic weather types produce different spatial rainfall patterns. Skaugen (1997) conducted research on scale properties of daily areal rainfall in Norway, with the rainfall events being classified into small- and large-scale events, based on their statistical pattern recognition. In the research, the rainfall events are also classified into convective showers and frontal rainfall events. It was concluded that the spatial averages for large scale frontal events had not reduced much in magnitude with increasing area as opposed to the small scale convective rainfall events. Huff and Shipp (2002) used a detailed classification method and found that the decay in spatial correlation in smaller storms that occurred in lower pressure centres as opposed to the fronts associated with mid-latitude cyclones, and that it is greatest in air mass storms.

2.3.1.3 Climate and rainfall types

The climate is highly variable in South Africa. Hence, hydrological and climatological information were used by Alexander (2010) to define nine distinctive climatological regions in South Africa as illustrated in Figure 2.1. Typically, apart from climate, other factors such as geographical location, altitude above mean sea level, rainfall type (convective, frontal and/or orographic), rainfall seasonality (summer, winter and/or all year) and average catchment slope classes (flat, moderate or steep) were also considered to define the various regions as shown in Figure 2.1.

Typically, in the south-western Cape (Mediterranean, and Southern Coastal regions), the climate is characterised by winter rainfall and warm windy summers, while highly variable, non-seasonal rainfall and extreme temperatures occur in the Karoo (KAR) region. Hot summers with convective thunderstorms and cold winters are typical on the Highveld, while mesic-subtropical conditions dominate on the KwaZulu-Natal coast of the Escarpment region (Davies and Day, 1998; Alexander, 2010). The MAP decreases, while potential evaporation increases westwards and northwards across South Africa. The overall MAP is 452 mm, but in many parts of the country, the MAP is much less. Evaporation exceeds rainfall throughout the country, except in the mountainous Escarpment and Mediterranean regions. In the central parts of South Africa, evaporation is approximately twice the rainfall, while in the western parts of the country evaporation exceeds the rainfall by a factor of ten (Davies and Day, 1998).

The temporal and spatial distribution of rainfall is highly variable on a seasonal and annual basis, since the rainfall is produced by different weather systems in different regions and at different times of the year (Davies and Day, 1998). In winter, the prevailing north-westerly winds result in high rainfall in the western part of the country, while the southern interior and Karoo remain dry. Summer rainfall is normally higher in the north and east, but due to dry high-pressure air masses that persist for long periods, the rainfall is low in the western parts of the country (Davies and Day, 1998).

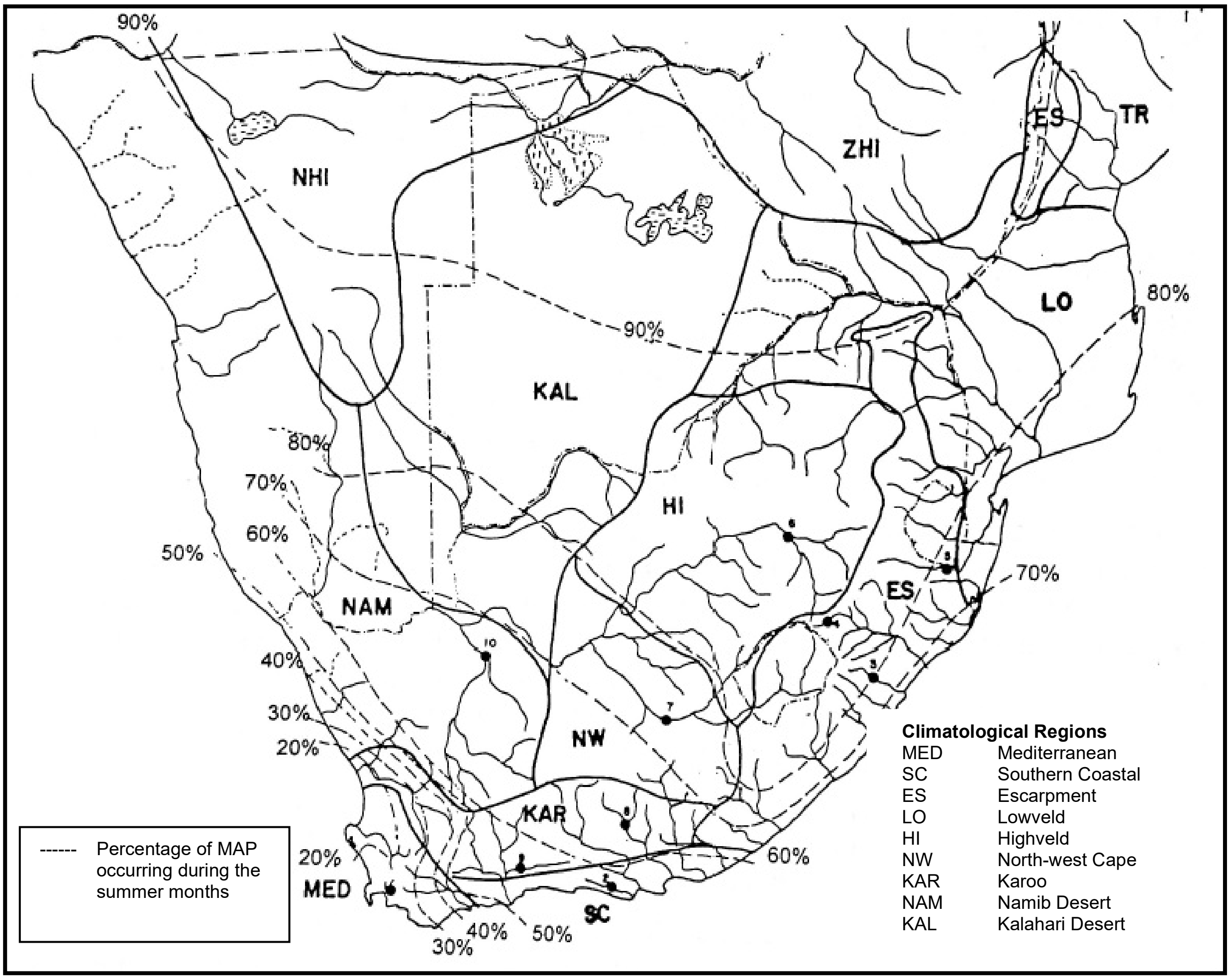


Figure 2.1: Climatological regions for South Africa (Alexander, 2010)

Bárdossy and Pegram (2018) highlighted that a 10-year return period rainfall event somewhere in London (UK) will be greater that the same return period event occurring somewhere in Munich, due to their differences in area, i.e. the greater the area, the higher the probability of occurrence of a storm of a given magnitude. Climate does not only affect rainfall distribution, but also rainfall intensity, duration and variability, which are all interdependent. However, the four major rainfall processes occurring in South Africa will also affect this interdependency and are most likely to have different influences on the estimation of ARFs. The four major rainfall processes occurring in South Africa can be summarised as follows (Haarhoff and Cassa, 2009; Van der Spuy and Rademeyer, 2018):

a) Convective rainfall: This process typically occurs during the summer season when air layers (closest to the earth's surface) saturated with water vapour are heated and subsequently tend to rise and cool down, resulting in cloud formation and rainfall. The rainfall intensity is normally high to very high with associated thunder activity. Convective rainfall is characteristic of the Highveld region which covers the Free State, Gauteng and Mpumalanga provinces.

b) Cyclonic rainfall: This rare process typically occurs over the open sea and is formed when cyclones (large circular patterns) are growing in size, allowing moist air to be drawn into the cyclone vortex and allowing mist to be lifted up into the centre, resulting in very strong winds and extremely high rainfall intensities.

c) Frontal rainfall: This inland process typically occurs when cold or warm fronts are moving across the country and interact with one another. The cold air has the tendency to move underneath the warm air, and the warm air is deflected upwards by the trailing edge of the cold air. In both cases, the warm air is lifted up into the colder region, resulting in rainfall.

d) Orographic rainfall: This process usually occurs near coast lines and typically develops when wind blows over the open sea towards land carrying air saturated with water vapour until it reaches a mountain range. At these geographical barriers, the saturated air is forced upwards to result in condensation and rainfall. The rainfall intensity is normally regarded as moderate and dependent on wind blowing towards the inland areas. Orographic rainfall is characteristic of the coast lines of KwaZulu-Natal and the Western Cape provinces.

The rainfall types listed in (a) to (d) were carefully considered to highlight and describe the direct influence of these on the estimation of ARFs. The magnitude of ARFs is highly dependent on the different storm mechanisms associated with different rainfall types. In a specific region with more frequent thunderstorms (convective rainfall) occurring than frontal storms (wide spread rainfall), the typical observed point rainfall AMS for that specific region would likely consist of rainfall values associated with convective activity (rainfall with rapidly changing intensity); whereas, the frontal rainfall values could have been more representative of the actual rainfall process in that particular catchment or region. This may result in much lower probabilistically correct ARFs (thunderstorms with high intensities), as opposed to the probabilistically higher ARFs represented by the frontal activity (Siriwardena and Weinmann, 1996).

In recognition of the above-mentioned interdependencies, Weddepohl (1988; cited by Schulze et al., 1992) demarcated South Africa into four distinctive daily rainfall intensity distribution regions. Typically, Region 1 is associated with a Type 1 design rainfall intensity distribution which is regarded as the lowest, while Type 4 is associated with the highest rainfall intensity. The spatial distribution of these regions can be summarised as follows: (i) Region 1: Eastern Cape, namely, East London and Port Elizabeth, (ii) Region 2: Western Cape (Karoo) and Free State, (iii) Region 3: Northern Cape, namely, Upington and Kimberley, as well the Highveld, including Gauteng and Mpumalanga, and (iv) Region 4: the remainder of the country.

Considering ARFs with rainfall types (a) and (c), Skaugen's (1997) research in Norway (as mentioned in Section 2.3.1.2) found that the difference in ARF curves between convective and frontal rainfall events had become more pronounced for longer return periods. There is a considerably higher rate of decrease in ARFs for convective rainfall events than for frontal evens; both convective and frontal events' ARFs decrease with an increase in return period. Allen and DeGaetano (2005) conducted research on considerations for the use of radar-derived rainfall estimates in determining return intervals for extreme areal rainfall amounts in USA. This research concluded that ARFs are smaller in warmer seasons than in the colder seasons presumed to be responsive to the increased convection over the summer season. Huff and Shipp (2002) found a comparative seasonal difference: the decay with distance of spatial correlation pattern of rainfall was greater in the warmer seasons than the colder seasons. The decrease in ARFs with the increase in return period could potentially reflect the significance of convection in producing heavy point rainfalls.

Skaugen (1997) pointed out that point rainfall extremes associated with convective rainfall events tend to occur inland, as opposed to the maxima of the large-scale events that usually occur along the coast.

Allen and DeGaetano (2005a) found that the warmer seasons in the Eastern United States regions (April-September) produced ARF values that decay at a faster rate when compared to the colder seasons (October-March). This was said to be attributed to the season dependent rainfall mechanisms and the associated spatial variability of rainfall.

2.3.1.4 Return period

Bell (1976) conducted a study that specifically focused on ARFs in rainfall frequency estimations. Bell (1976) re-examined the ARFs produced by the Flood Studies Report (FSR) (NERC, 1975), with a focus on the influence of return period. Two methods were used to test the significance of the influence of return period: a non-parametric sign test and an adaption of the *t*-test for comparing means of samples from populations with different variances.

The non-parametric sign test emphasised the differences between group values for any pair of return periods to be significant at 95% level, for the 24-hour ARFs. Similarly, its application to the short duration ARFs showed increasing significance at 99% levels. With the *t*-test significant differences were also found between the 2 year and 20-year 24-hour ARFs at 95% level. This suggests that the data provided reasonable evidence that ARFs decrease with increasing return periods. Bell (1976) pointed out that the values of ARFs in the FSR correspond to return periods of 5 to 10 years and with a tendency to be conservative for longer return periods.

Stewart (1989) carried out a study on ARFs that were used for design storm areal rainfalls by using rainfall station and radar data. Stewart (1989) followed Bell's (1976) methodology which consisted of relying on obtaining ARFs directly from frequency curves. The study introduced a standardisation of the rainfall data through division by the mean annual maximum rainfall. In this way, the ARFs were derived using rainfall growth curves instead of rainfall frequency curves. This enhances the effect of locational variations to be represented by differences in the mean annual maximum values, while return period effects are represented by the growth curves. The findings were concurrent to Bell's (1976) findings in that the ARF had the tendency to decrease for longer return periods. Allen and DeGaetano (2005a) estimated ARFs with a high rainfall station density, following the fixed-area approach and found that ARF values for a 2-year return period decay exponentially. It was found that higher return periods were associated with lower ARF values.

In terms of the methods used to derive ARFs, Omolayo (1993) suggested that storm-centred ARFs tend to be incorrect for estimating areal rainfall of a particular frequency from point rainfalls due to its use being mainly suited for probable maximum precipitation (PMP) studies.

2.3.1.5 *Spatial variability of rainfall*

Kim J, Lee, Kim D, and Kang (2019) studied the role of rainfall spatial variability in deriving ARFs. The research highlighted the influence of internal spatial variability of storms on ARFs. A storm identification algorithm on composite radar data was employed to identify some 55 000 elliptically shaped extreme storms over a six-year period. Thereafter, an investigation on various storm characteristics with the corresponding ARF values were carried out. The main assumption was accepting that the ARF generally increases with duration, with an inverse relationship to storm area. With this, the spatial variability within a storm is found by calculating a coefficient of variation of radar image pixel rainfall values, as this proved to be a strong predictor of the ARF value, along with area and duration. The difference of ARF values between storms that have elliptical shapes and those that are circular over the same area was found to be approximately 20% on average. With these findings, the current design framework of areal rainfall estimation is said to be improved by integrating the information of the rainfall spatial variability and storm shape.

2.3.2 Methods of deriving Areal Reduction Factors

Theoretical approaches for deriving ARFs were developed based on the relationship characteristics of extreme storms and the extent to which their characteristics associate with one another. The earliest attempt at deriving an ARF followed an empirical approach that focused on single storm events, dating back to 1957 (US Weather Bureau, 1957). These approaches had the tendency to disregard the influence of return period and were pioneering for countries like Italy where they are currently frequently used for defining design storms for urban drainage systems (Supino, 1964). The theoretical approach was furthered with the inclusion of variance functions and reduction parameters (Rodriguez-Iturbe and Mejía, 1974). By 1984, Waymire *et al.* (1984) carried-out analyses using a stochastic derivation of ARFs for rainfall events, which included the spatial and temporal characteristics of the rainfall events.

Unfortunately, the two recognised approaches, namely, the storm-centred and geographically-centred approaches, used to estimate ARFs generally provide inconsistent results. In using a storm-centred approach, the isohyets of a complete storm are analysed without considering the geographical location thereof (Alexander, 2001). In the case of a geographically-centred approach, storms occurring over a fixed area or collection of rainfall stations on the catchment's surface are considered (Alexander, 2001). Bell (1976) highlighted that the theoretical significance of the geographically-centred approach is more statistical than physical and is therefore best interpreted in terms of average areal point rainfall frequency curves, which simply provides the ratios of areal to point rainfall with the same AEP.

Thus, it is quite evident that the use of different methodologies to estimate ARFs is likely to result in different ARF estimates.

2.3.2.1 *Storm-centred ARFs*

Stormed-centred ARFs are based on particular individual storm event(s). Storm-centred ARFs are typically computed by dividing an observed area-averaged accumulation by the maximum observed point accumulation from that particular storm. With the exception of forecasting and nowcasting, the storm centre is typically calculated after the storm event, since there is no complete accurate method of predicting where the storm centre is located within a storm. Storm-centred ARFs are often used to develop estimates of PMP. The storm-centred ARF approach is rarely used outside of PMP, partially due to the fact that they are inextricably linked to storm classification and complications associated with multicellular storms (Wright *et al.* 2014).

Pavlovic *et al*. (2016) defined storm-centred ARFs as ratios formulated and are based on the analysis of individual storms, which are used to convert point estimates of PMP to areal rainfall estimates. This is in line with the definitions of storm-centred ARFs from Wright *et al.* (2014). Although there are various challenges associated with deriving storm-centred ARFs, they can be considered as an authenticity check to examine the validity of their fixed area counterparts.

Storm-centred ARFs are affiliated with the calculation of the effective depth for discrete storms as well as signify profiles of individual storms, with data that are usually provided by a reliable weather service. In reality, the area in which the rain falls is not predetermined as it changes with each individual storm. With storm-centred ARFs, the maximum point rainfall from a rainfall event is considered the centre of the storm which is crucial for the calculation of ARFs. The ratio of the average areal storm rainfall depth and the maximum storm point rainfall is characterised with the aid of these values. Generally, rainfall depths, obtained from isohyets, are divided by the maximum point rainfall of the same storm; they are integrated to obtain the average storm rainfall depth (Gill, 2005). Storm-centred ARFs are usually calculated using Equation **1**:

$$\boldsymbol{ARF} = \frac{\boldsymbol{R}}{\boldsymbol{P}} \qquad \text{Equation 1}$$

Where:

R is the areal storm rainfall enclosed by a specific isohyet, within which the secluded rainfall is greater than or equal to the value of the isohyet, and

P is the maximum point rainfall at the storm centre.

<u>This research focuses on incorporating a storm-centred approach</u> due to the spatial and temporal characteristics of a storm being factored into the derivation of the ARF. A geographically-fixed approach typically implies using rainfall data, which is increasingly problematic in the hydrological field due to the declining number of functional rainfall stations and unreliability that is associated with badly maintained rainfall stations and its data.

2.3.2.2 Geographically-fixed ARFs

Geographically-fixed derived ARFs describe the relationship between the areal average design rainfall over a geographically-fixed area with a corresponding design point rainfall value that is representative of the area considered. In other words, the ARF is used for percentage reduction which directly relates to the statistics of areal and point rainfall. As a result, this considers the uniform temporal and spatial distribution of rainfall over a specified area (Pietersen, 2016). Of the two types of ARFs, geographically-fixed derived ARFs are more frequently used due to the degree of difficulty that comes along with the storm-centred derived ARFs. This type of ARF considers different parts of different storms, instead of considering the highest point values at respective storm centres. Thus, these ARFs are not necessarily related to any individually recorded storms but do originate from rainfall statistics (Omolayo, 1993).

Wright *et al.* (2014) defined geographically-fixed ARFs values that are computed by dividing the extreme value of area-averaged rainfall through an extreme-point rainfall value of the same time period that is typical for that specific area. This is in line with the definitions defined by Pietersen *et al*. (2015) and Pavlovic *et al.* (2016).

Geographically-fixed ARFs are conventionally estimated from the average of frequency based quantile estimates, with using annual maxima rainfall series observed at a specific fixed location. These correlate the point rainfall depth of a specific area to the average rainfall depth for that area. Particularly, the representative point is an average point having the mean of the collective point rainfalls in the specified area. The area that is observed is dually fixed in time and space (Gill, 2005). In this case, the centre of the storm does not need to coincide with the centre of catchment area. Thus, the values of the ARFs are naturally based on different parts of different storms rather than at the highest point rainfall values at the respective storm centres. Geographically-fixed ARFs are thus derived from rainfall data accumulations instead of individual storms events. Geographically-fixed ARFs can be represented by Equation 1 in Section 2.3.2.1, with the exception that R is the mean of annual maximum values and P is generally the weighted mean (due to uneven spatial distribution of rainfall stations) of annual maximum point rainfall values at gauged locations, within the specified area (Bell, 1976).

A number of researchers noted that ARFs vary with geographical location and climate due to a difference in the predominant rainfall generating mechanisms (Svensson and Jones, 2010). On a global scale, a study by Omolayo (1993) suggested that 1-day ARFs are generally larger using USA's method, as opposed to Australia's method. Furthermore, Asquith and Famiglietti (2000) found that ARFs decline more rapidly in semi-arid south-western USA (specifically Texas) than elsewhere in the country.

Although geographically-fixed ARFs are generally used in the hydrology field, it lacks the spatial variability that comes with all storm events. This ultimately leads to a conservative approach, as certain characteristics of storm events are completely neglected.

2.3.2.3 Annual maxima centred ARFs

The annual maxima centred ARFs were proposed by Asquith and Famiglietti (2000). This approach considers the distribution of concurrent rainfall surrounding annual maxima. The approach is not dependent on the prior spatial averaging of rainfall, explicit determination of spatial correlation coefficients or the explicit definition of a representative area of a particular storm event for analysis. Instead, the approach is designed to make extensive use of the wide availability of dense rainfall station data in various region across the world. It considers the spatial distribution of rainfall occurring concurrently with and surrounding an annual maximum value at a particular point within a catchment area (Asquith and Famiglietti, 2000).

In order to arrive at an annual maxima centred ARF, the following steps were implemented by Asquith and Famiglietti (2000). Firstly, the ratio of the annual maxima depth to the concurrent rainfall was computed for every annual maxima within a specific database, as well as the separation distances between the two respective rainfall stations. Thereafter, these ratios produced insight on the description of the relation between criteria conditioned sample ratio values and separation distances. These relations are defined by specific functions fitted to an empirical ratio relation. Thus, a best fit line that produces the expected ratio is achieved. Lastly, the ARFs are computed for a user defined area with specific design criteria. This explains why ARFs are functions of catchment area, geographical location, shape and return period.

2.3.2.4 United States of America

Point rainfall frequency estimates are used in many infrastructure designs in the United States of America (USA). These estimates are only representative for a limited proximity of area around the rainfall point and as such reduces its usefulness in many applications that demand areal rainfall frequency estimates. ARFs are highly sensitive for the method by which it is derived from.

Research carried out by Pavlovic *et al.* (2016), in the USA dealt with the analysis of differences amongst ARFs from various geographically-fixed ARF methods. The researchers defined an ARF to be a concept used in engineering design that converts point rainfall into areal rainfall, for specified durations and frequencies. For the comparison of ARFs, one representative method from each of the four main categories: empirical (M1), spatial correlation structure of rainfall (M2), temporal and spatial scaling properties of rainfall (M3), and extreme value theory (M4). Next Generation Radar (NEXRAD) rainfall data were used in the latter cases. Pavlovic *et al.* (2016) highlighted that the predominant source of uncertainty in raw radar precipitation estimates, i.e. the assumed relationship between the radar reflectivity and precipitation amount.

The M1 method is based on an original empirical ARF method. The United States Weather Bureau (USWB) formulated several curves to transform point rainfall to areal rainfall for storm durations between 1 and 24 hours, for areas less than 1 000 km^2. As most methods were empirically-based over 50 years ago, this method disregards the influence of return period. The M2 method incorporates the spatial correlation structure of rainfall within the derivation of ARFs. With this, an exponential distribution for the point parent rainfall and a Gamma distribution for the areal average parent rainfall are assumed. The M3 method uses concepts of dynamic scaling and statistical self-affinity in order to arrive at a general expression for the mean annual maxima as a function of rainfall and area. Essentially, ARFs are derived by fitting a general expression to the ARF estimates calculated from M1. In addition, the M3 method does not account for the influence of the return period on ARFs. Finally, for the M4 method, ARFs are derived by estimating the precipitation frequency estimates that are obtained by fitting the GEV to the mean regional annual maxima series data for each combination of chosen storm durations and catchment sizes (Pavlovic *et al.*, 2016). It is evident that the USA has an extensive analysis of deriving ARFs by using the fixed-area method. The methods used are highly dependent on rainfall stations, but included spatial correlation structure to subsequently combat the problem of defining the spatial behaviour of a storm event.

Wright *et al.* (2014) argued that insufficient attention has been given to frequently used ARFs, with the formulations used to derive them. The results portray that there exist large discrepancies between the frequently used ARFs and the true characteristics and properties of extreme rainfall events, for a specific study region. The researchers mention that using a storm-centred approach to derive ARFs can serve as a reality check for the examination of their geographically-fixed area counterparts.

By using the storm-centred approach, Wright *et al.* (2014) found that the frequently used ARFs failed this reality check, and this consequently suggested important implications for flood risk estimations. Specifically, the latter research analysed storm-centred derived ARFs for five storms in order to compare them to storm-centred derived ARFs from a larger population of storms. This was done to ultimately determine the difference between ARFs from more extreme storms (larger return periods) and more frequent storms (lower return periods). The most widely used source of ARFs in the USA originates from Technical Paper 29 (TP-29), by the US Army Corps of Engineers (Beard, 1967). Wright *et al.* (2014) pointed out that a principle weakness in the method TP-29 uses to derive ARFs is that the maximum areal rainfall for a specified area and duration and the maximum point rainfall for a specified area and duration are not from the same storm event. Furthermore, the TP-29 states that its ARF estimates are not dependent on storm magnitude, and in turn, disregards the influence of return period (Wright *et al.* 2014). This shows that the storm-centred ARFs derived from TP-29 lacks cohesion and are subjected to a high degree of uncertainty. The uncertainty results from using storm characteristics from different storms for the formulation of ARFs. The ARFs tend to be higher, as the maximum areal rainfall would generally be greater than the averaged areal rainfall, while the maximum point rainfall always stays the same for the data set.

2.3.2.5 United Kingdom

The geographically-fixed area method used in the UK (NERC, 1975), does not consider the influence of return period on ARFs, as its influence was regarded negligible. For this method, annual maximum areal rainfall over a particular region is calculated. With this, the point rainfall measurements for a particular station and year are recorded. Independently, the annual maximum point rainfalls at each station for each year in the area of interest is recorded. The ARF is then calculated for a specific area and duration. This method is considered to be a simplification of the US Weather Bureau method. It was purely adopted for computational convenience. Svensson and Jones (2010) considers this method to be unorthodox, on the basis that it is an average of ratios that is safely approximating a ratio of averages.

Following Section 2.3.1.4 and 2.3.2.4, Bell's (1976) method of obtaining ARFs involve ranking a PDS using Thiessen's Polygon as well as ranking the AMS per rainfall station in the area under consideration. In order to obtain a single point rainfall frequency curve that is representative for a specific area, the Thiessen-weighted mean of the annual maximum point rainfalls of the same rank was computed (Svensson and Jones, 2010). Thereafter, frequency distributions were fitted to the areal and point rainfall series to obtain ARFs calculated for the different return periods.

2.3.2.6 National Weather Service method

This method falls under the M1 method of deriving ARFs in Section 2.3.2.4. It is outlined in the National Oceanic and Atmospheric Administration (NOAA) TP National Weather Services (NWS) 24. The method was mainly developed to take advantage of the AMS based on the two-station average rainfall statistics. The mean and standard deviation from the same points are estimated for station pairs at arbitrary distances for various durations. A smoothing surface is then fitted to each parameter in the distance-duration-space.

The method is based on the frequency analysis of annual maximum pairs of stations and the distance between them. This way of deriving ARFs explicitly takes the variation of ARF with return period into account, and this reduces the need for large, dense networks with concurrent data observations by using statistics of station-pairs and small five-station networks (Bell, 1976). Isotropy in the spatial rainfall area was assumed due to the random locations of the station pairs and five-station network. Thus, elongated catchments with rainfalls generally aligned in one direction is not considered. Svensson and Jones (2010) mentioned that it is questionable as to whether this complex methodology is justified as precipitation observations become more over time.

2.3.2.7 Storm movement

Bengtsson and Niemczynowicz (1986) conducted research on deriving ARFs from rain movement. The researchers took a simplified conceptual physics approach to deriving ARFs by specifically moving an idealised storm across an area; this method closely maps the storm-centred method. Data requirements for this study are limited due to the intention of the research focusing on urban catchments up to 30 km^2 with relatively short durations up to 40 minutes. This method of deriving ARFs was referred to as: the moving storm derived ARF (M-ARF).

The derivation of M-ARFs were based on using 12 recording rainfall stations in Sweden and focussed on the movement of convective storms. The M-ARFs were calculated from rainfall observations at a fixed point and its storm speed. The main assumption made was that the shape of the hyetograph and the velocity of movement does not change during the storm passage over the area. Since urban areas are limited in areal extent, rainfall intensities were assumed not to change drastically. The method assumes a lateral decay in rainfall intensity that is applicable for small convective storms, with its main area of application for urban hydrology. According to Svensson and Jones (2010), this method is not applicable to larger catchment areas as well as for longer rainfall durations.

M-ARFs was found not to have significant dependence on a particular rain gauge hyetograph. Instead, average M-ARFs derived from hyetographs at any of the rainfall stations used in the study, produced stable estimates. Bengtsson and Niemczynowicz (1986) found that the M-ARFs "…agree well with true areal reduction factors…" (Bengtsson and Niemczynowicz, 1986), which have similar values all over Sweden. Where hyetographs were not available, synthetic storms were simulated to move across the catchment (Svensson and Jones, 2010).

2.3.2.8 Review of methods for deriving ARFs

The relationship between point rainfalls and areal rainfalls has been found to mainly vary with predominant weather type, season, return period and estimation method. Svensson and Jones (2010) point out that analytical methods has the tendency to categorise ARF estimations on a sound scientific basis. On the contrary, the analytical methods are typically based on assumptions that are not entirely considered as ground truth descriptions of the real rainfall process, which Svensson and Jones (2010) classifies as "… cause for concern and uncertainty regarding the results.". This concern is compounded by the limited amount of rainfall data that is used to verify results. On the other hand, with a smaller amount of computational effort and data requirements, some ARF estimation methods prove to produce reasonably acceptable results in comparison to traditional ARF estimation methods.

Empirical and analytical methods may not produce areal rainfall estimates that are probabilistically correct. Applying an ARF to a *T*-year rainfall to obtain areal rainfall may not produce an areal rainfall of the same return period. Hence, more focus is put on the measurement of the discrepancy; a small difference in results may be acceptable when it is considered in conjunction with its advantages. After these discrepancies are addressed and assessed, it seems prudent to recommend these methods for use with rainfall frequency estimates. For any method of deriving ARFs, the underlying data has more importance than any results obtained.

2.3.3 Areal Reduction Factors in South Africa

2.3.3.1 Current ARFs

The main ARFs used for South Africa is found in the SANRAL Drainage Manual (2013). Van Wyk (1965) was the first South African to analyse ARFs based on a storm-centred approach, which was conducted on a small-scale (catchment areas ≤ 800 km²) in Pretoria, Gauteng. In addition, a few rainfall storm areas from the USA and Canada were analysed for comparison purposes. Intensity duration frequency curves were based on Gumbel's theory of extremes using short duration rainfall data. By focussing on smaller catchments, storms were predicted to be intense and short, producing maximum response or peak discharge. The data used was the yearly records of 20 rainfall stations, typically inclusive of in the Pretoria region in South Africa that used maximum rainfalls from 15 minutes up to 1 440 minutes. Hershfield (1961) suggested that a factor of 1.13 be applied to the output of the Gumbel values which were applicable in the USA based on the assumption that rainfall was the same all over the world. It was found that intensities for winter were lower than those for summer, with year-round rainfall lying in between. The areal distribution of rainfall for small-area storms were defined as major percentage of total volume in area to an area less than 800 km^2. Isohyetal maps of several storms were plotted based on the average areal rainfall depths in catchments ranging from 10 km² to 800 km² centred on the maximum point rainfall and expressed as a percentage of point rainfall at the storm centre (Van Wyk, 1965, cited by Lambourne and Stephenson, 1986).

The ARFs were also expressed as a function of the point source rainfall intensity, particularly an average intensity over the storm duration at the storm centre (Van Wyk, 1965, cited by Lambourne and Stephenson, 1986). As a result, depth-intensity-area envelope diagrams were developed (Figure 2.2). From this, it is evident that the ARFs are mainly a function of area and design point rainfall intensity, since the relationship between rainfall intensity and infiltration rate of the soil is predominant (Pietersen *et al*., 2015). It was recommended by Van Wyk (1965) that radar data be used for a suitable depth-area analysis, which is the focus of this research.

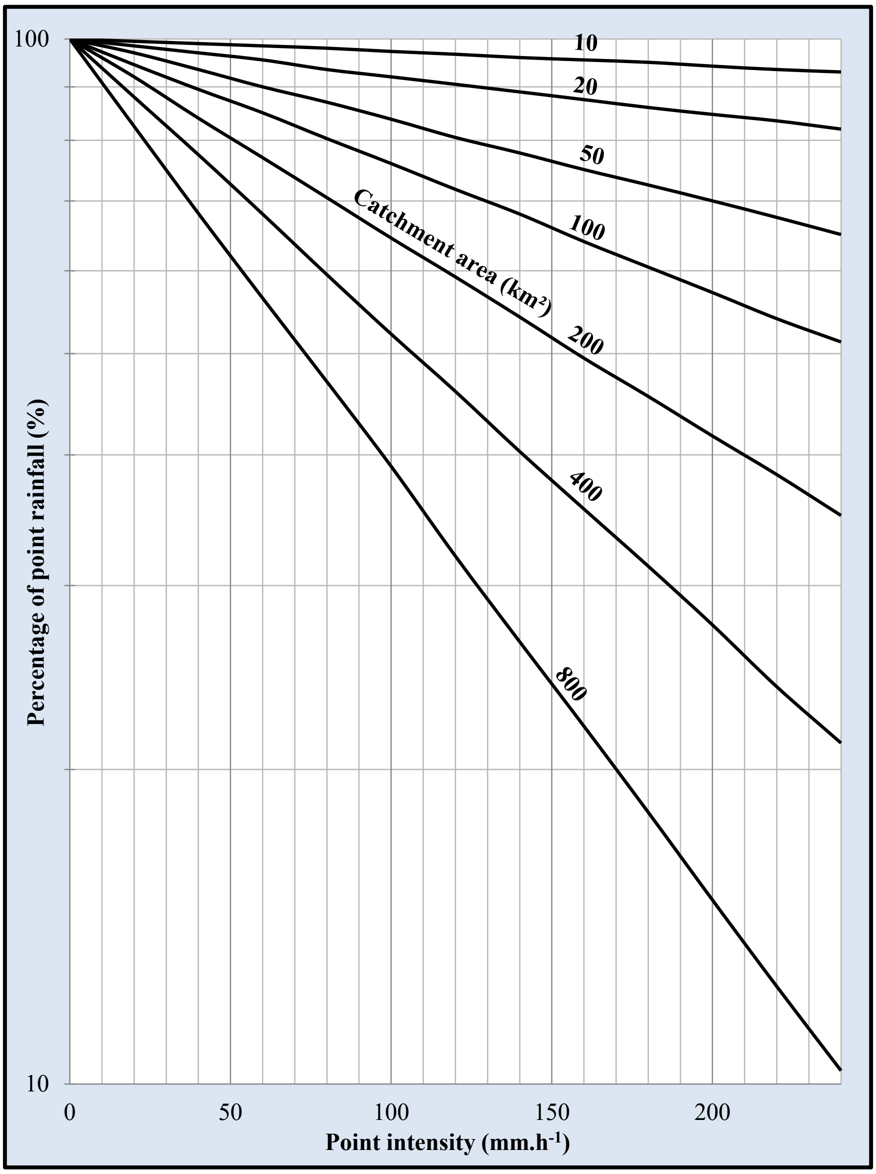


Figure 2.2: Expected percentage of runoff as a function of point rainfall intensity (SANRAL, 2013)

Op ten Noort and Stephenson (1982) converted Figure 2.2 into a mathematical expression using regression analysis as presented in Equation **2**:

ARF $= \boldsymbol{Exp}(-\mathbf{0.000068}\boldsymbol{iA}^{\mathbf{0.77}})$ Equation 2

where:

ARF = areal reduction factor for point rainfall (fraction),

A = catchment area (km²), and

i = point rainfall intensity at the storm centre (mm.h^{-1}).

Pullen, Wiederhold and Midgley (1966) produced an alternate study that focused on large-area storms using depth-area-duration analyses. The study analysed some 170 storms that covered catchment areas between 500 km² and 30 000 km² within 18 regions delineated for South Africa.

A formulation of isopercentile maps were used, which are smooth patterns of isometric lines that are drawn amongst sample observations when expressed as percentages of local mean annual rainfalls. The methodology for deriving depth-area-duration curves involved the determination of maximum average rainfall depths occurring within selected time intervals throughout the total storm period, on areas encompassed by each isohyet of the total storm isohyet map. The large area storms were delineated while the point rainfall depths at each rainfall station were used to fit a 6th-degree polynomial surface to enable the plotting of isohyets. Regionalised depth-area curves were produced for each storm at a daily interval resulting in co-axial diagrams to estimate the rainfall equalled or exceeded for storm durations of one day or longer. The developed depth-duration-area envelope diagram is shown in Figure 2.3.

In the case of large area storms with associated storm durations less than 24 hours, the areal average rainfall over increasing areas (durations of 1 to 6 days) within each of the 18 regions were expressed as percentages of the maximum observed point rainfall. Depth-area diagrams were produced for durations of 1 to 6 days. The upper envelope diagrams (of individual durations) were then re-plotted to produce depth-duration-area diagrams. Thereafter, the 24-hour to 1-hour durations were linearly extrapolated to express the rainfall associated with a given area as a proportion of the point rainfall between one and 72 hours (Lambourne and Stephenson, 1986).

ARFs were mainly functions of area and storm durations, since the quantity of the rainfall relative to the number of storage areas is of great significance (Pietersen *et al.*, 2015).

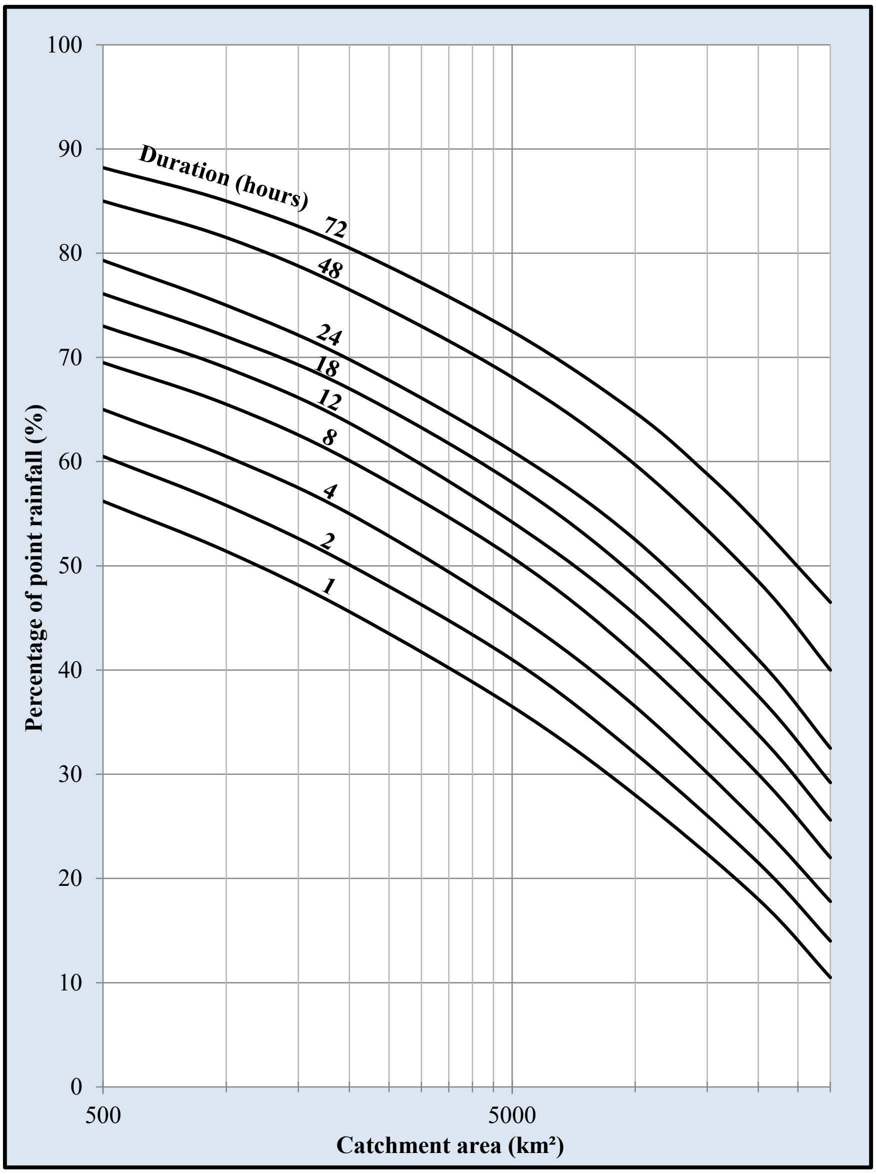


Figure 2.3: Expected percentage of runoff as a function of storm duration (SANRAL, 2013)

Op ten Noort and Stephenson (1982) converted Figure 2.3 to the mathematical expression using regression analysis as shown in $[1.343\text{-}0.09\ln(\mathrm{A})]\mathrm{T}_{\mathrm{d}}^{0.03\mathrm{A}^{0.19}}$
Equation 3.

ARF = $[1.343 - 0.09\ln(A)]T_d^{0.03A^{0.19}}$ Equation 3

where:

ARF = areal reduction factor for point intensity (fraction),

A = catchment area (km²), and

T_d = storm duration (hours).

Op ten Noort and Stephenson (1982) compared Equation 2 and Equation 3 and established that the use thereof could cause a discontinuity in storm runoff estimation. Consequently, Figure 2.3

was extrapolated such that the ARFs approach unity at short durations. This relationship is expressed by Equation **4**.

ARF = $[1.04 - 0.08\ln(A)]T_d^{0.02A^{0.28}}$ Equation 4

where:

ARF = areal reduction factor for point intensity (fraction),

A = catchment area (km²), and

T_d = storm duration (hours).

Apart from these methods discussed in the sections above, Table in Appendix B contain a summary of additional storm-centred ARF estimation methods used internationally.

2.3.3.2 UK FSR

The United Kingdom Flood Studies Report (NERC, 1975) contains extensive research of the analysis of available hydrometrical and rainfall records, carried out in Britain and Ireland. Volume II of the FSR contains a copious amount of valuable information and knowledge in the field of hydrology by focusing on the estimation of rainfall depths corresponding to rainfall durations and return periods. Furthermore, the FSR provides insightful knowledge on ARFs. Cunnane and Lynn (1975) produced a review of the FSR with regards to flood estimation.

The FSR produced research on rainfall studies that investigated point rainfall estimates, ARFs, storm profiles and estimated maximum rainfall. Point rainfall estimates were obtained by analysing rainfalls of 5-year return periods with durations of 60 minutes and 2 days. Growth factors along with long term rainfall and the rainfall for a 5-year return period were used to arrive at the point rainfall estimates (Cunnane and Lynn, 1975).

The ARF is applied to the point rainfall of given frequency to arrive at the corresponding areal rainfall. The ARFs derived in the UK FSR is provided in Table 2.3, with recommended ARF values for areas up to of 30 000 km^2 and durations of up to 25 days. The ARF was found to not vary significantly with location or return period. On the other hand, the FSR shows that the ARF does, in fact, increase for a specified area with an increasing in duration; the ARF diminishes with an increase in area, for a specific duration. The results displayed in Table 2.3 are in accordance with experience from which it would be expected that persistent rain (longer duration) would have greater areal uniformity, while short duration rainfall might not be uniform, as reflected in the smaller ARFs.

Table 2.3: Relation of ARF with duration (D) and area (A) (Cunnane and Lynn, 1975)

Duration, D	Area A (km²)									
	1	5	10	30	100	300	1000	3000	10000	30000
1 min	0.76	0.61	0.52	0.4	0.27	-	-	-	-	-
2 min	0.84	0.72	0.65	0.53	0.39	-	-	-	-	-
5 min	0.9	0.82	0.76	0.63	0.51	0.38	-	-	-	-
10 min	0.93	0.87	0.83	0.73	0.59	0.47	0.32	-	-	-
15 min	0.94	0.89	0.83	0.77	0.64	0.53	0.39	0.29	-	-
30 min	0.93	0.91	0.89	0.82	0.72	0.62	0.51	0.41	0.31	-
60 min	0.96	0.93	0.91	0.86	0.79	0.71	0.62	0.53	0.44	0.35
2h	0.97	0.95	0.93	0.9	0.84	0.79	0.73	0.65	0.55	0.47
3h	0.97	0.96	0.94	0.91	0.87	0.83	0.78	0.71	0.62	0.54
6h	0.98	0.97	0.96	0.93	0.9	0.87	0.83	0.79	0.73	0.67
24 h	0.99	0.98	0.97	0.96	0.94	0.92	0.89	0.86	0.83	0.8
48 h	-	0.99	0.98	0.97	0.96	0.94	0.91	0.88	0.86	0.82
96 h	-	-	0.99	0.98	0.97	0.96	0.93	0.91	0.88	0.85
192 h	-	-	-	0.99	0.98	0.97	0.95	0.92	0.9	0.87
25 days	-	-	-	-	0.99	0.98	0.97	0.95	0.93	0.91

Following Section 2.3.1.4, Bell (1976) found that ARFs for 1- and 2-hour durations both decreased with longer return periods, which contradicts what the FSR provides.

2.3.3.3 Review of ARFs for South Africa

Recently, Pietersen *et al.* (2015) carried out a review of the current methods for estimating ARFs used in South Africa, with a preliminary focus on the identification of new methods. The main objectives of this study included a national and international comparison of ARF estimation methods, with an emphasis on the differences in these methods, along with the assessment of graphical and numerical ARF estimation methods, using standard input variables (such as catchment area, time of concentration, duration and rainfall intensity). The overarching theme was that the ARFs currently implemented in South Africa are outdated with a need for renewal. The review study looked at the ARFs currently presented in the SANRAL Drainage Manual (2013). The two storm-centred methods, as discussed in Section 2.3.1.1, were compared to their numerical estimated counterparts, as presented by Op ten Noort and Stephenson (1982). For the geographically-fixed area method, Alexander (1980), who produced ARFs for South Africa based on the UK FSR (NERC, 1975), was reviewed. However, these ARFs were only applicable when assuming uniform temporal and spatial rainfall distribution over a catchment. In addition, Alexander (1980) had produced a numerical relationship for his graphical results, which included the critical duration of a storm event.

The research continued to produce a case study to apply the review ARF estimation methods; firstly, by using standard input variables, and secondly, by applying these ARF estimations to a pilot study area. For the standard input variable phase, it was found that the ARFs decreased for an increase in catchment area with significant differences that presented the presence of inconsistencies between results from the numerical and graphical methods. Pietersen *et al.* (2015) pointed out that Van Wyk (1965) and Pullen, Wiederhold and Midgley (1966) methods are not suitable for estimating catchment areal design rainfall from design point rainfalls. In doing so, the incorrect assumption is made that extreme design point rainfall and extreme areal design rainfall are produced by the same rainfall event or rainfall type. In the case of the pilot study area, it was found that the geographically-fixed numerical ARF estimation methods were more consistent, with the exclusion of the influence of return period. Pietersen *et al.* (2015) points out that ARFs should be derived from local rainfall data as opposed to the UK FSR (NERC, 1975) transposed data, due to the variation observed from areal rainfall characteristics in South African catchments. Furthermore, it was suggested that the current ARFs for South Africa need to be updated by utilising longer rainfall records. The variation of ARF with return period and rainfall-producing mechanisms was suggested for further research. Finally, Bell's (1976) approach was recommended due to its geographically-fixed approach which encompasses an unofficial national conventional method, considering that Van Wyk (1965) and Pullen, Wiederhold and Midgley (1966) has a major discrepancy methodology in its process to derive ARFs.

2.4 The Use of Radar Data in Hydrology

2.4.1 Radar data for research in hydrology

Weather radars provide quantitative rainfall estimates with a high spatial and temporal resolution. With this, adjustments need to be applied due to gross errors like beam attenuation caused by strong rainfall and those caused by variability of the drop size distribution and a non-uniform profile of reflectivity.

South Africa is a semi-arid country with large portions of its surface area receiving annually on average less than 500 mm of rainfall. Large-scale flood events are often brought on by prolonged periods of drought. In addition, convective storms produce a large percentage of the annual rainfall and this adds to the potential to cause local flash floods and other severe storm related damages.

Terblanche, Pegram and Mittermaier (2001) point out that SAWS is the custodian for the collection of rainfall data as part of its climate database. However, daily records of rainfall data from rainfall stations has decreased from 4 500 to 1 750 active gauges in about 55 years, with about 600 gauges producing daily recordings. This add to the urgency to improve the country's ability to monitor rainfall over large areas in close to real-time during flood events. In addition, there is a need to improve the spatial resolution of rainfall measurements that can determine more accurate catchment rainfall estimates under convective conditions. This points to radar data as a possible alternative. The UK, USA and several other European countries have already successfully integrated weather radar data for hydrological applications such as flood warning and water resources management systems. In South Africa, limited resources are available as it is a large developing country.

From 1970 to 1990, radar meteorological research in South Africa was limited to studies and activities that focussed on improving the understanding of natural rainfall processes, severe storms and the possibility of developing a viable rainfall enhancement technique. Storm dynamics, cloud microphysics and hail for severe storm studies were the priority research at the Council for Scientific and Industrial Research (CSIR). In the early 1990s, this programme was terminated, and one S-band and two C-band Doppler radars were constructed and operated in the Pretoria and Johannesburg regions, which resulted in multiple Doppler studies. Simultaneously, rainfall enhancement research was carried out in Nelspruit. Thereafter, rapid progress was made towards developing 'new' cloud seeding technology for the rainfall enhancement. At this point, the emphasis on radar research moved away from the radar-based comparison between seeded and natural storms in randomised seeding experiments, towards quantitative measurements of areal rainfall (Terblanche, Pegram and Mittermaier, 2001).

Terblanche, Pegram and Mittermaier (2001) used weather radars as a research and operational tool for hydrology in South Africa. A Radar Data Acquisition System (RDAS) for control of radar antennae, digitising and processing outputs from the radars receiver into useful reflectivity, was developed. A 'displace' processing method for the elimination of digitised receiver averaging errors in areas of steep reflectivity gradients was developed. Furthermore, development for the performance of testing procedures and upgrades to ensure sustained radar operations and high-quality radar data was carried out. Finally, the introduction of Thunderstorm Identification Tracking and Nowcasting (TITAN) real-time storm tracking and analysis system was presented for use in South Africa, which is detailed in Chapter 3.

2.4.2 Radar data for the derivation of ARFs

The derivation of ARFs have been described in Section 2.3, with nearly every empirical derivation making use of rainfall station data, with limited methods using radar data. Radar data has become available and numerous researchers have investigated its use in hydrology as an alternative to using ground observation based data (Stewart, 1989; Terblanche, Pegram and Mittermaier, 2001; Allen and DeGaetano, 2005b; Sinclair and Pegram, 2005; Wright *et al.* 2014).

Svensson and Jones (2010) mentioned that radar data provides an improved spatial coverage of rainfall events, in comparison with dense rainfall station networks. On the other hand, records obtained from radars tend to be short, particularly for fine spatial resolution. In addition, quantitative measurements are poor in comparison with rainfall station data, which can be resolved by incorporating rainfall station data with radar data.

2.4.2.1 Rainfall stations

Rainfall stations are the most widely used data source when it comes to deriving ARFs. Rainfall stations are generally considered to provide the most accurate rainfall information at any given location, with certain limitations. These include instrumental errors and under-catch due to wind and erratic behaviour of the mechanical aspects of the rainfall station during intense rainfall. Furthermore, conversion of point rainfall to areal rainfall is sensitive to the interpolation technique employed. This is due to the high spatial variability of rainfall and presents a significant challenge for ARF calculations, particularly at sub-daily time-steps (Pavlovic *et al.*, 2016).

Multiple studies have been carried out on the joint use of rainfall station and radar data. Stewart (1989) described the analysis of spatial variability of rainfall, considering the joint use of rainfall station and radar data for North West England, UK. A fixed-area ARF derivation approach was implemented for extreme rainfall events with durations of 1 hour to 8 days, within an area of 100 000 km^2. A similar methodology to Bell (1976) was put in place, where ARFs were derived directly from frequency curves. Due to the limitations of the record of radar data, the calculation of ARFs were carried out directly for 1 to 8-day durations, ranging in areas of 25 to 10 000 km^2, from which average areal values were computed. It was found that ARFs increased with duration for a given area, and decreased with area for a given duration, in line with findings from several other studies, e.g. Bell (1976). These ARFs were compared to ARFs produced in the FSR and showed that the FSR contained conservative ARF values. Furthermore, ARFs decreased with return period; there was no strong tendency for ARFs to vary with location for the study area. The study concluded that more confidence is placed on ARFs obtained for smaller areas than for larger areas.

2.4.2.2 Storm centres

Convective cells, from a storm event, are dynamic objects that comes with the difficulty of tracking with conventional approaches due to their rapid change in shape. Algorithms for the identification of storm centres in radar data have been developed over time, following two approaches. The first being correlation tracking algorithms that provide velocity and direction information for larger area events and the second being cell identification and tracking algorithms that provide location information on isolated cells. A study carried out by Picus *et al.* (2008) focused on tracking storm centres in radar data for short term weather prediction, using a mean-shift method with integral image computation. The mean-shift algorithm identifies high density of modes within a complex feature. For the identification of each mode, the mean-shift algorithm requires an iterative computation of area integrals. It was concluded that the mean-shift algorithm was able to correctly track small storms.

Following Section 2.3.1.5, Kim *et al.* (2019) identified that storm centres were based on the assumption that extreme events could be captured by referring to the pixels of a radar image, on the basis that the accumulated rainfall volume with a given duration exceeds a specific threshold over the entire record of the six-year period used in their research. Storm centres that corresponded spatially and temporally with extreme storms types, were identified. It was pointed out that the spatial and temporal correspondence of each storm centre are of great importance, since ARFs are primarily used to reflect the characteristics of extreme storms in designing flood defence systems. An assumption was made that the extreme storms contained at least one of the extreme pixels that have a rainfall depth greater than the rainfall associated with the10-year return period, and less than a 200-year return period. With the storm centre being identified, an analysis was carried out on each storm, to eliminate instances of overlap in time.

2.4.2.3 Other countries

Following Section 2.3.2.4, Wright *et al.* (2014) derived ARFs from a 10-year, high resolution radar with a bias corrected, data record from the Hydro-NEXRAD system. The research made use of a mean-field bias correction of the 10-year record at a daily scale using rainfall stations within the study region, Charlotte, USA. ARFs were calculated for 1, 3, 6 and 12-hour durations for a threshold area of 3 600 km^2, using the storm-centred approach. It was established that the calculated ARFs are lower than the ARFs in TP-29. With longer durations, the ARFs tend to approach ARF values presented in TP-29. It was highlighted that longer duration ARFs compared well with the ARFs found in TP-29. This resulted from the storms that produced high long-duration point accumulations which have the tendency to additionally produce high long-duration areal rainfall accumulations. However, storms that produce high short-duration point accumulations, did not necessarily produce high short-duration accumulations over larger areas.

The mean ARFs for tropical and non-tropical storms are significantly less than the ARFs presented in TP-29 for all durations. The researchers highlighted that there is a tendency towards multicellular storm structure for longer-duration storms that are based on the number of ARFs that do not monotonically decrease with area. In addition, ARFs were found to decay more rapidly with area for the storms that were selected based on the size and shape of the catchment used for the study (Wright *et al.* 2014).

Wright *et al.* (2014) concluded that storm type has a significant influence on the derivation of storm-centred ARFs, especially within the study area of Charlotte, North Carolina. The rainfall from tropical storms had a tendency to be spatially larger and of longer duration than rainfall from organised storm systems. Therefore, tropical storms decay less rapidly with increasing area. The study used radar rainfall data as an alternative for improving ARF estimates, based on the principle that radar data capture a wide range of storm behaviour, which can then be readily used to characterise rainfall spatial variability. Furthermore, the variability has the potential to be incorporated into design storms and flood risk estimates.

Pavlovic *et al.* (2016) used NEXRAD gridded rainfall data with rainfall station data to determine ARFs for Oklahoma, USA. It was pointed out that the predominant source of uncertainty with raw radar rainfall estimates is the assumed relationship between the reflectivity and rainfall amount. This varies based on rainfall type. Since rainfall is sensed well above ground surface, the rainfall detected by the radar could potentially move large distances downwind or evaporate before reaching the ground. In addition, uncertainty arises from radar technology itself. Conversely, radar data has the potential to be useful for calculating the statistics of extreme events, and spatial pooling can typically be employed to compensate for short radar records.

Following the study carried out by Pavlovic *et al.* (2016), ARFs estimates were compared for the M1 through M4 methods, as discussed in Section 2.3.2.4. ARFs were calculated from the averages of the AMS data and compared to the US Weather Bureau ARF estimates. It was found that for longer durations, the ARFs were similar for all the area sizes, but the differences became more pronounced as the duration decreased. All four methods were classified as conservative with regards to the US Weather Bureau across all durations and all areas considered. This notion lies more in line with the TP-29 estimates, and consequently, contradicts the conclusions stated by Wright *et al.* (2014). This could be due to the methodologies (storm-centred vs geographically-fixed) to arrive at these ARF estimates. Furthermore, the average recurrence interval's (ARI's) influence was tested on the ARF estimates and it was reported that there was a clear separation of ARFs with longer ARIs, in line with the findings of Bell (1976). The dependency of ARFs on ARI was found to be more pronounced for shorter durations.

Bacchi and Ranzi (1996) obtained ARFs using a stochastic derivation of storm intensity, based on the analysis of crossing properties of the rainfall process aggregated in time and space for Italy. Thus, a storm-centred approach was mapped. The data used included radar images with a time resolution of 15 minutes. This data was verified with corresponding rainfall station data for 17 stations, and produced satisfactory results. It was established that radar data is more efficient than common rainfall station networks as radars host the ability to capture internal structure and spatial distributions of storms. The main findings of the research included an observation that the

ARFs decay according to a power function with respect to area, as well as a weak decay of the ARFs with respect to the return period in urban areas.

Lombardo, Napolitano and Russo (2006) used radar reflectivity for the estimation of ARFs. The study consisted of analysing ARFs using radar reflectivity maps that were collected with a Polar 55C weather radar, which is a C-band Doppler dual polarised coherent weather radar with polarisation agility. This radar measures the most used horizontally reflectivity factor (Z_h), the differential reflectivity and the differential phase shift. These measurements are obtained by averaging 64 pulses with a range-bin resolution of 75 m, and threshold of 120 km away from the radar location. The conversion of reflectivity to rainfall intensity (*R*) was based on a non-linear regression analysis (Equation 5):

$$R = 7.27 \cdot 10^{-2} Z_h^{0.62} \qquad \text{Equation 5}$$

An approach different to the geographically-fixed or storm-centred method was followed; a scaling law was used and obtained by the ratio between the radar rain rate estimates over an area ranging from 1 km^2 to 900 km^2 and the radar rain rate estimates over 1 km^2. Similar results as reported by Bell (1976) and Wright *et al.* (2014) were reported; the longer the duration, the higher the ARFs for selected areas as well as the larger the area, the higher the ARFs for selected durations. These results were found to be more applicable to floodplain management as well as in the design of urban drainage systems, for basins 200-900 km^2, with storm rainfalls that are typically associated with 25 to 50 year return periods. This includes 1 to 2-hour storm concentration times: where estimated ARF values range from 0.1 to 0.3.

Extreme areal rainfall depths are typically obtained by spatial interpolation of rainfall station data. A reliable estimation of these depths is often hampered by the low spatial density of rainfall station networks. Overeem, Buishand, Holleman and Uijlenhoet (2010) investigated extreme value modelling of areal rainfall obtained using weather radar. An 11-year radar rainfall data set was used to abstract annual maximum rainfall depths for durations of 15 minutes to 24 hours, with area sizes of 6 to 1 700 km^2 for the Netherlands. A GEV distribution was fitted to the annual maxima for each area size and duration, separately. With this, areal rainfall depth-duration-frequency curves were derived and ARFs calculated. The ARFs in this study were compared to ARFs produced in the FSR; it was concluded that the difference in ARFs of rainfall station data and rain data estimated through radar pixels, is small for a duration of 24 hours. Adjustment factors were applied to the ARFs in the FSR which produced ARFs for short durations, which had strong correlations with ARFs produced in the Overeem *et al.* (2010) study. Figure 2.4 displays the radar-based ARFs and highlights the influence thereof on the return period. For areas up to 500 km^2, ARFs decrease significantly with area size, for both long and short durations. These results are indicative of rare events that have relatively high spatial gradients.

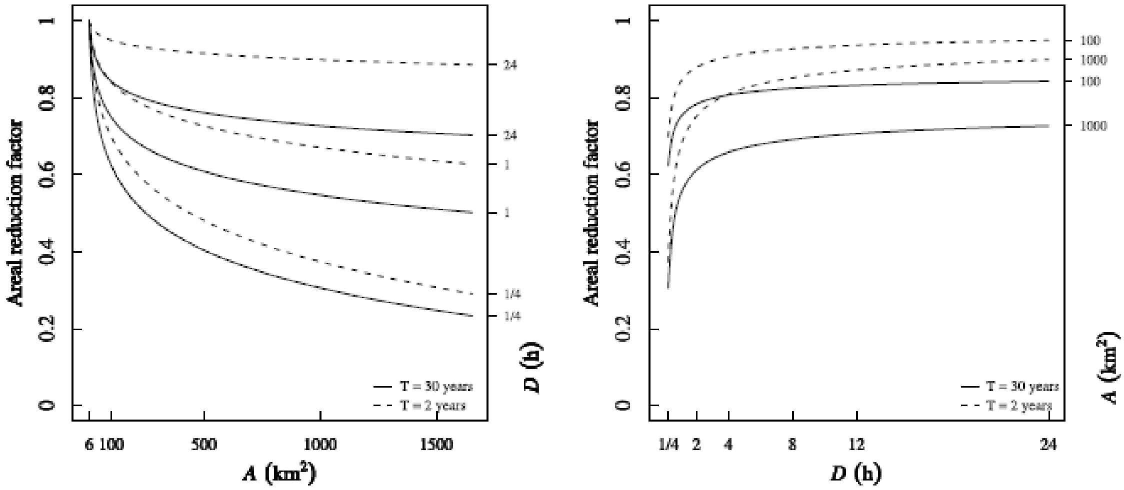


Figure 2.4: Radar-based ARFs plotted against area size (left) and duration (right) (Overeem *et al.,* 2010)

Following Section 2.3.1.5, Kim *et al.* (2019) analysed the influence of various storm characteristics on ARFs. The characteristics considered were the maximum, mean, variance and coefficient of variation of the pixel radar rainfall values within a storm. It was found that an increase of the mean storm rainfall was associated with an increase in ARF value, for most areas and durations. This results from the ARF being a direct function of mean storm rainfall. The increase of the maximum storm rainfall was associated with an increase in ARF values for larger areas and shorter durations, with an opposing trend for smaller areas and longer durations. An increase in rainfall variance generally lead to the decrease of ARF values due to spatially varying storms (such as convective storms) having a tendency to produce smaller ARFs in comparison to storms that have little spatial variability (such as frontal storms). The coefficient of variation was used to test the influence of rainfall spatial variability, regardless of the mean storm rainfall. It was found that the increase in the coefficient of variance strongly correlates with a decrease in ARF values. This signified the most consistent factor affecting the ARF amongst the other storm characteristics.

3 RADAR DATA SOFTWARE APPLICATION

3.1 Weather Radars in South Africa

Following Section 2.4.1, Terblanche, Pegram and Mittermaier (2001) conducted research on the development of the weather radar for South Africa. There are various types of radars that exist in the world, each suited for its specialising functions. Figure 3.1 displays the waves and frequency ranges that radars operate in and are typically named after.

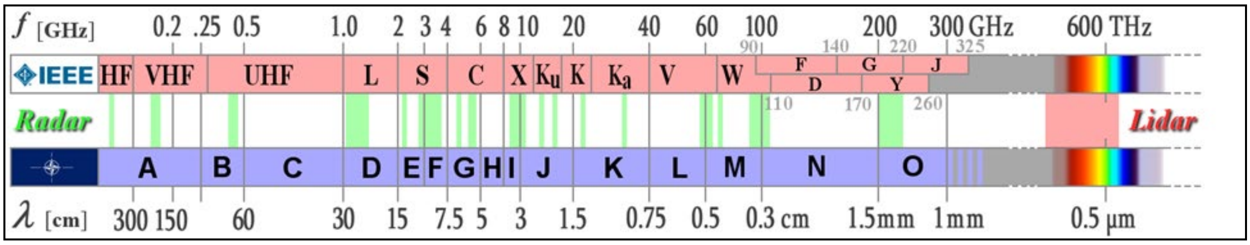


Figure 3.1: Waves and frequency ranges that are used by radars (Wolff, 2002)

In 2002, the South African Weather Bureau (SAWB) owned ten C-band Enterprise radars of various ages. At this point in time, the newer radar systems made use of Doppler facilities. The radars were installed in Cape Town, Port Elizabeth, East London, Durban, De Aar, Bloemfontein, Bethlehem, Irene, Ermelo and Pietersburg, as shown in Figure 3.2. The Water Research Commission (WRC) owned an MRL-5 dual wavelength (S-and X-band) research radar near Bethlehem and a Pacer C-band radar situated near Tzaneen. The Pacer radar was used to exclusively support the rainfall enhancement programme in the Northern Province. The spacing of the radars shown in Figure 3.2 resulted from the need for storm and rainfall intensity surveillance at the regional offices of SAWB, with no intention of hydrological application in mind. Terblanche, Pegram and Mittermaier (2001) point out that the spacing of the radars are not ideal for observing stratiform rainfall, as these systems are relatively shallow, generally resulting in the radar beam overshooting the echo top at long ranges. Furthermore, convective rainfall systems have deep vertical dimensions that allows them to be observed at larger ranges. For the radar horizon that determines blockage, the radar sites are said to be well selected. However, some beam blockage occurs at the Bethlehem C-band and Tzaneen radars, at low elevations.

A PC-based radar data acquisition and antenna control system (RDAS) was developed to ensure flexibility in terms of data collection, from the various types and models of radars used. The RDAS software was expanded to include routines to assist during the calibration of a radar as well as to ensure uniformity in the way any calculations and procedures were carried out. This relieved the hydrological readers with a copious amount of technical detail. The procedure eliminated manual and error prone calculations. The result was a receiver slope with an electronic file that contained calibration information used in the operation of the radar, and consequently formed part of the calibration history for a specific radar. Calibration checks were carried out consistently and was found to be "…remarkable stable" (Terblanche, Pegram and Mittermaier, 2001).

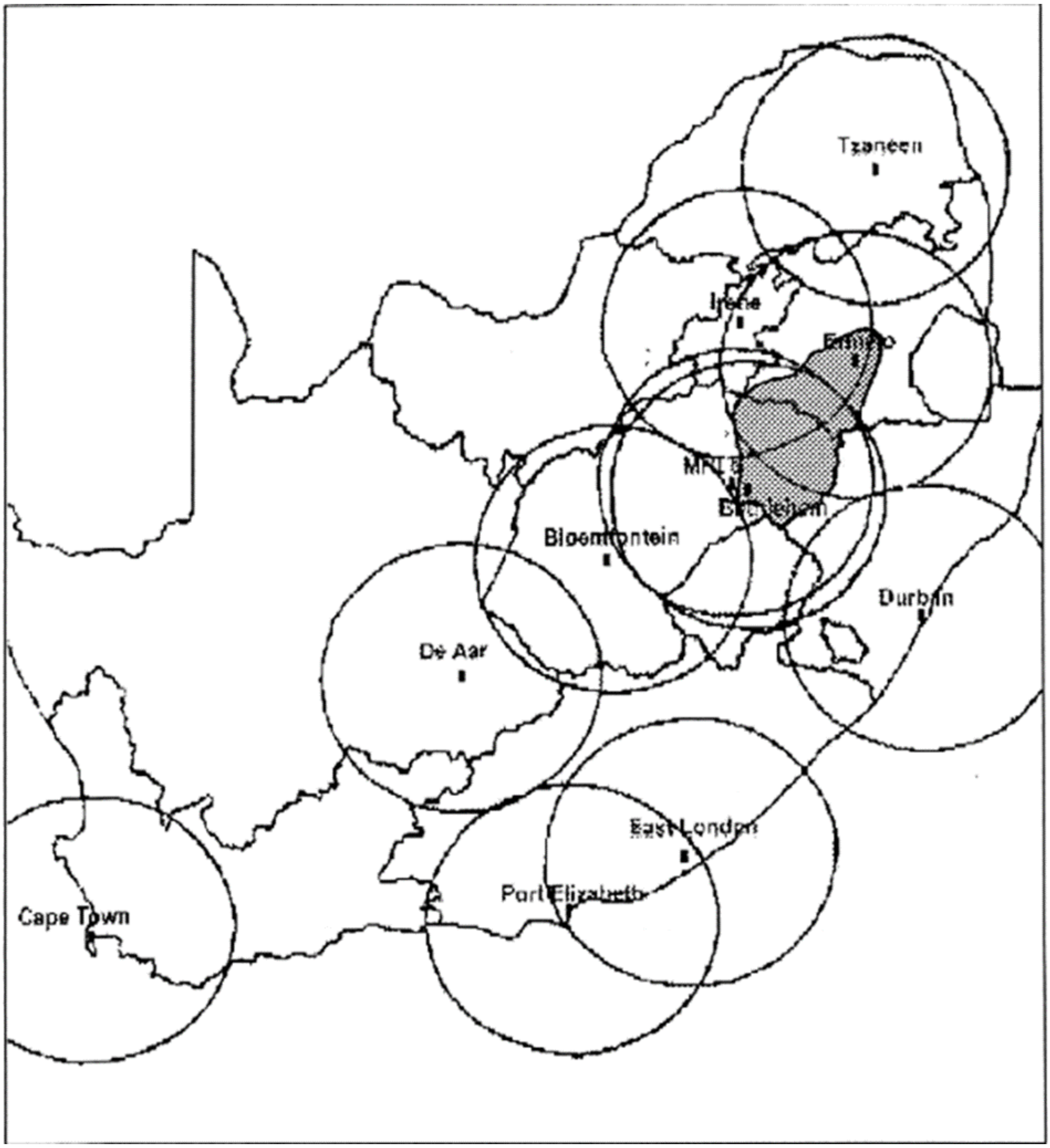


Figure 3.2: The South African Weather Radar Infrastructure, with circles that represent 200 km data collection range (Terblanche, Pegram and Mittermaier, 2001)

After the arrival of the MRL-5 in 1994, RDAS was upgraded to facilitate the calibration and data processing from the S- and X-band radars. A continuous power supply was generated to the system to facilitate 24-hour operations, with a microwave link installed for real-time data transfer to Bethlehem. Thus, the MRL-5 was classified as a reliable system, for 24-hour volume scan operations.

An efficient processing algorithm was developed to replace the customary averaging of the logarithmic output and correcting it with a 2.5 dB averaging bias. The algorithm, created by Terblanche (1996a, b), was called ‘displace’, which carries out pair-wise averaging on digitised logarithmic receiver samples using averaging lookup values that were functionally dependent on the difference between the pair and the receiver transfer function to be simulated. In this way, a true unbiased average of received power was obtained when simulating a quadratic receiver. This eliminates the underestimation that occurs in areas of steep reflectivity gradients when using the customary averaging technique.

The algorithm was further developed to achieve more accurate interpolations when converting spherical coordinate volume scanned radar data to Cartesian coordinates. This technique was found to be twice as computationally efficient as conventional interpolation techniques. Constant Altitude Plan Position Indicator (CAPPI), a coordinate file type that is an output from the 'displace' algorithm, simplified the merging of information from several radars. This is a requirement for the TITAN software, which plays a pivotal role in the South African weather radar data manipulation and representation system.

The acquisition of quality data for research and analysis is an imperative process in the quest to understand various phenomena. The channels used to obtain data should be from credible and reliable institutions/providers to ensure calibration and verification standards are maintained and documented. The data required for this project are primary focused on rainfall estimates based on radar derived data. The rainfall estimates are essential for the derivation of ARFs as they are used to calculate areal rainfall, and the maximum point rainfall for each storm event. The rainfall estimates are the main component for each file, within each data set, and these estimates are dependent on the storm duration. Since the storm durations are controlled (1, 3 and 24 hours), the rainfall estimates are now considered accumulated rainfall estimates.

The next section provides an overview of the software used throughout the data acquisition process to convert and analyse the radar data. This includes the gathering and viewing of the raw data files, conversion of the raw data to commonly used data formats, and methods to quickly view and analyse the data. Scripts to view the raw data on various platforms are presented in this document to show that the data contain no major spatial-temporal discrepancies.

3.2 Radar Software

Radar data can be analysed in multiple software applications. However, emphasis is put on a file that is compatible in a Geographical Information System (GIS) environment to allow for further spatial-temporal analysis.

The TITAN software package is used in the first stages of data acquisition. Here raw radar data is visualised and algorithms interpret and produce products for use in operational environments. For geospatial research purposes, the data is further processed to be analysed using software packages such as the open source visualisation package Ncview, Panoply, and GIS software, QGIS, and/or proprietary GIS software, ArcGIS. Scripting languages provide a powerful tool to automate various tasks, scripting also allows users to easily do reproducible analyses on the data without altering the structure of the raw data files. For this reason, statistical programming languages such as *R* and Python will be discussed. Scripts that were used to inspect the data is also presented in this report.

3.2.1 TITAN

The TITAN software system developed at the National Centre for Atmospheric Research (NCAR) in the USA by Dixon and Wiener (1993) and based on previous software developed in South Africa, was introduced in 1995. TITAN is a real-time storm tracking system that allows its users to analyse the space-time evolution of various storms, and unique physical characteristics of these events. TITAN produces additional fields including vertically integrated liquid, accumulated rainfall estimates and projected storm movement and size changes. These fields are often used in severe storm identification and research. The combination of the RDAS and TITAN is successfully used in operational weather forecasting and atmospheric research settings around the world. TITAN uses a MDV data format for all its input and generate user defined outputs specifically designed for gridded two- and three-dimensional meteorological data. From 2001, this data format has become the common denominator between radar, satellite and other data and their derived products in South Africa.

TITAN identifies a storm as a region where the decibel relative to Z (DbZ) and volume exceeds a predefined threshold. Once a storm is identified, TITAN can spatially define and track the storm as it evolves and dissipates. According to the TITAN manual the following precipitation parameters can be derived from TITAN package.

- Precipitation area (km^2);
- Precipitation area centroid x;
- Precipitation area centroid y;
- Precipitation area ellipse orientation;
- Precipitation area ellipse minor radius; and
- Precipitation area ellipse major radius.

TITAN has two methods of defining a storm area, i.e. exact grid points on the edge of the identified storm boundary to form a polygon, or a simplified ellipse structure encompassing the identified storm area. The polygon is considered a complex shape and might distort some features on the edges of the storm, in contrast the ellipse is a simplified representation of the storm area, but as noted in TITAN manual, single cells within a squall line type storm might not be represented as an individual storm type. Figure 3.3 (from the TITAN manual) indicates the two methods of storm identification in TITAN.

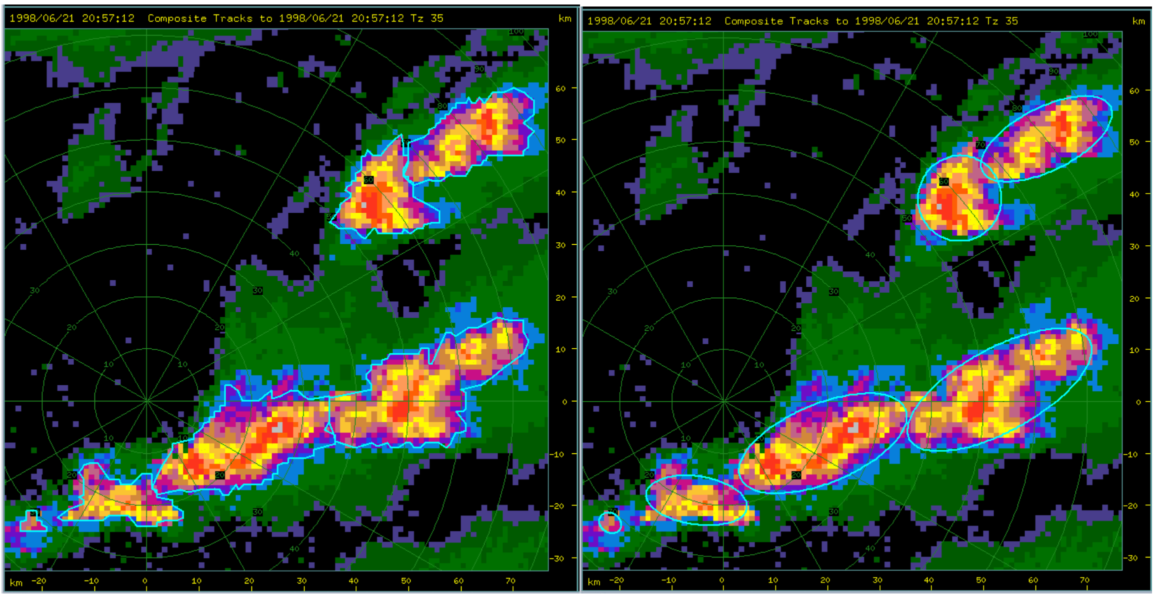


Figure 3.3: Storm identification in TITAN

3.2.2 Ncview

Ncview is a free and open source software to quickly inspect and visualise multidimensional netcdf files. The software is available as a precompiled package on all major Linux distributions. The software can also be used in Windows 10 systems with the ability to run the Windows Subsystem for Linux (WSL). Figure 3.4 shows an example of Ncview used to view the Maximum DbZ from one time step during radar scan that converted to a netcdf file as discussed later in this section. Note that the product shown here was not from the Irene radar, but the NWU Lekwena Radar, shown only as an example.

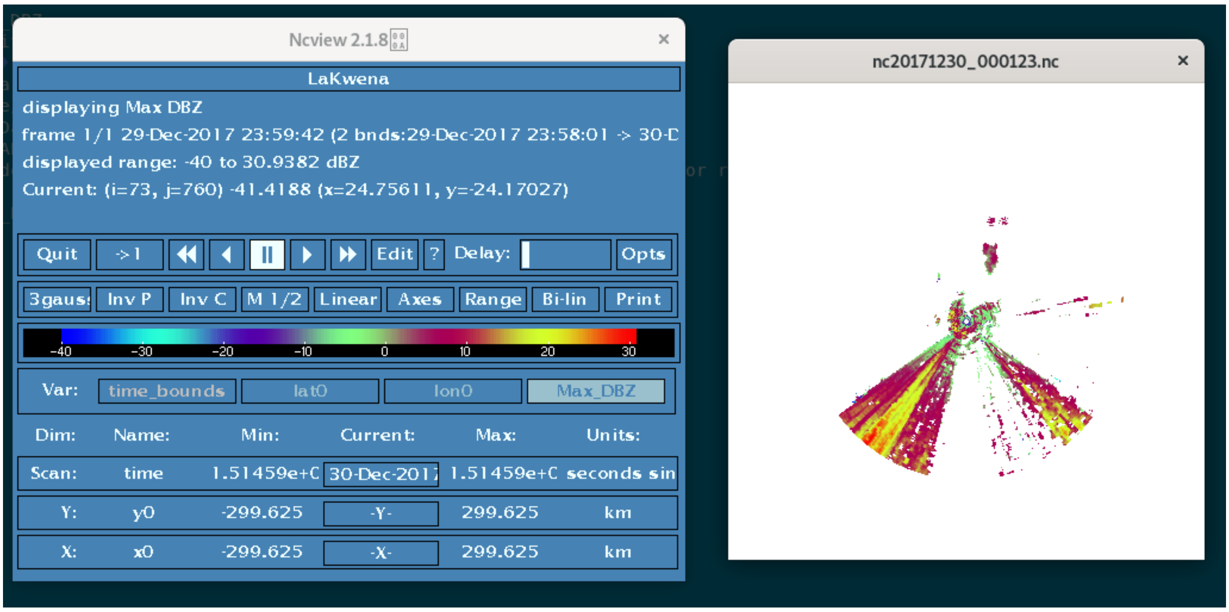


Figure 3.4 : An example of Ncview

3.2.3 *Climate Data Operators*

Climate Data Operators (CDO) developed by the Max Planck Institute for Meteorology in Germany, provides a command line interface to extract information, manipulate, inspect, quality control, perform calculations, and convert the various data formats commonly used in atmospheric sciences, including netcdf, and GRIB files. CDO is open source software and can be used on all major Linux operating systems and newer Windows 10 systems with WSL. CDO has a small memory footprint and can process large quantity of data sets, larger than the physical memory of the systems, and can be used in a scripting language such as Bash to analyse large numbers of files efficiently.

Figure 3.5 shows the interface to inspect netcdf files, in this case a radar derived file containing the maximum DbZ and it illustrates the details around the coordinate system and projections used during radar runtime. The netcdf file was created using a TITAN utility called MdvConvert as discussed later in this section.

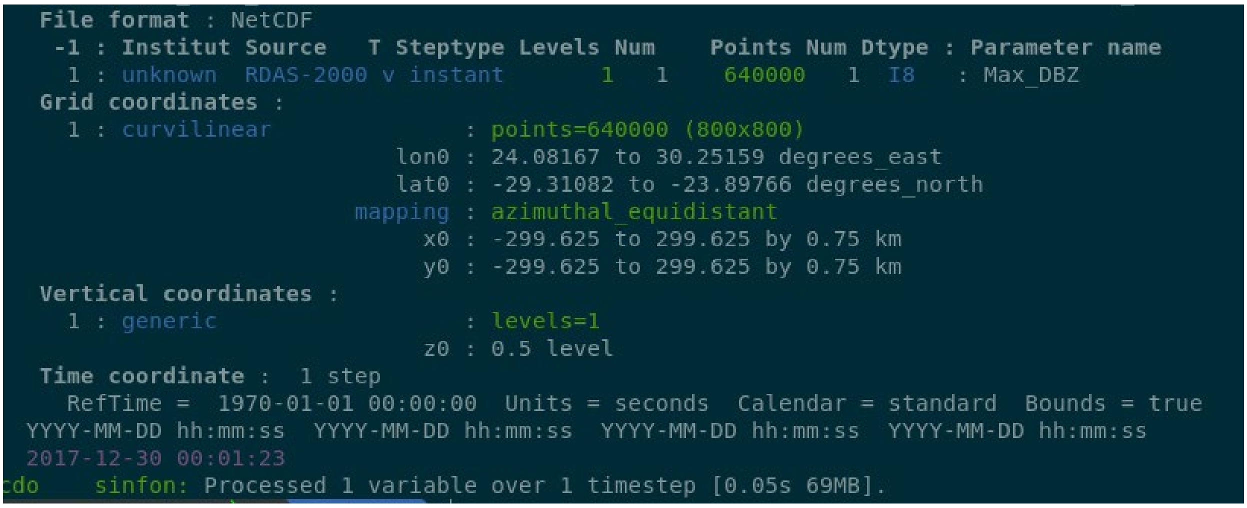


Figure 3.5: Interface to inspect netcdf files

3.2.4 *Panoply*

Panoply is used to view the contents of netcdf, HDF, and GRIB files. These file formats are commonly used in atmospheric research. Panoply is considered a cross-platform application that runs on the three major operating systems Macintosh, Windows and Linux, and can be compiled on other desktop computers. For this project, it is necessary to be able to open, view and separate sub data sets within a netcdf file environment. MDV produced by TITAN can be converted to netcdf file format using TITAN module MdvConvert and Mdv2netcdf. These files can then be accessed through Panoply software on any operating system and allows for the user to obtain the rainfall accumulation estimates.

Panoply is open source and free software, and it has the functionality to:

- Slice and plot geo-referenced latitude-longitude, latitude-vertical, longitude-vertical, time-latitude or time-vertical arrays from larger multidimensional variables.
- Slice and plot ‘generic’ 2D arrays from larger multidimensional variables.
- Slice 1D arrays from larger multi-dimensional variables and create line plots.

- Combine two geo-referenced arrays in one plot by differencing, summing or averaging.
- Plot long-lat data on a global or regional map using any of over 100 map projections or make a zonal average line plot.
- Overlay continent outlines or masks on long-lat map plots.
- Use any of numerous colour tables for the scale colour bar, or apply your own custom ACT, CPT, or RGB colour table.
- Save plots to disk GIF, JPEG, PNG or TIFF bitmap images or as PDF or PostScript graphics files.
- Export long-lat map plots in KMZ format.
- Export animations as MP4 video or as a collection of individual frame images.
- Explore remote THREDDS and OpenDAP catalogues and open datasets served from them.

Panoply's features can be viewed and the software can be downloaded from the National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS) website (see https://www.giss.nasa.gov/tools/panoply/). Figure 3.6 shows the typical Panoply interface, where the user can easily select the variables to plot, while also having control over the colour map selection and various other visual controls for publication quality graphics.

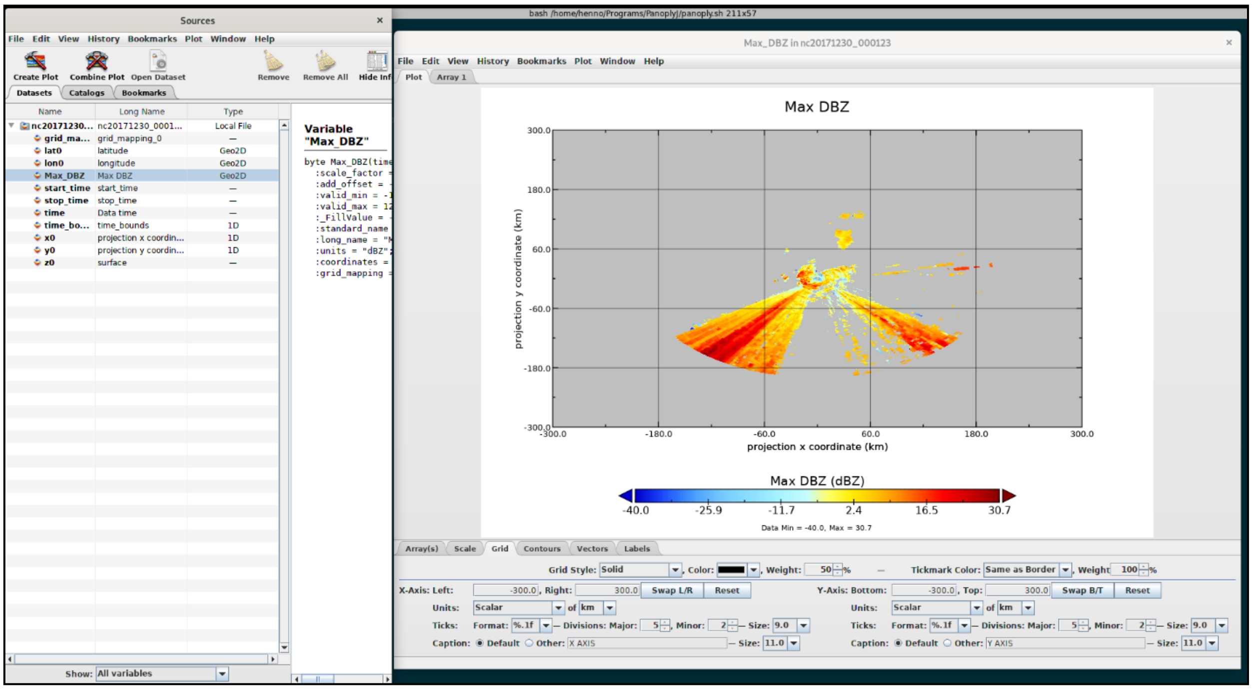


Figure 3.6: Typical Panoply interface

3.2.5 ArcGIS

ArcGIS is a widely used and commercially supported GIS software program. It is known for its ability to map and visualize most georeferenced data sets and comes with a set of tools for a wide variety of geo-spatial analyses. ArcGIS plays an integral role in being used to visualise the geo-referenced netcdf files, and to calculate the isohyetal interval storm area, the total storm area and to implement storm threshold values.

3.2.6 Statistical programming and scripting languages: Bash, R, and Python

The ***R*** statistical programming language is a popular open source alternative to proprietary statistical software. *R* is powerful and modular, allowing the user to do a variety of statistical analysis on various data types. R also contains a large number of so-called "packages" which allows the user to easily manipulate and work on various data types, and this includes netcdf files and geospatial visualisation. Where R is particularly useful, is in its ability to process large data sets, and *R* can also be used on servers where it is able to utilise the full capability of a system by means of parallelisation. *R* can also be used as a scripting language allowing users to automate tasks. Powerful *R* packages include *ggplot2* which creates publication quality graphics and also allows the user to do spatial visualisation of data. The *netcdf* and *ncdf4* packages allows users to view and manipulate netcdf data. *R* can be installed on all operating systems and also allows for the use of user friendly Graphical User Interface (GUI's) in the form of R-Studio

Similar to *R*, **Python** is a modular and powerful programming language that can be used to inspect, visualise and perform statistical analysis on various data sources including netcdf data and other meteorological data sources. Python includes a variety of libraries to inspect and perform quality control on the large data sources. Python can be used on a cloud platform such as Jupyter Notebook allowing for collaborative work on the same project. As with *R*, Python can be used as a scripting language allowing for the automation of tasks and allowing the user to run tasks using parallel capabilities of a computer and/or server.

Bash is a UNIX command language that allows the user to automate and script tasks. Bash (and other command languages) includes various utilities that can be used to quickly inspect data and also perform data manipulation in a controlled and scripted fashion using regular expressions. New Windows 10 systems with WSL allows users to run Bash on Windows, while the language is standard on Mac and Linux systems. Bash is particularly powerful for task automation and a user can call programs within a script to perform multiple analysis. An example could be to automate the conversion of MDV files to netcdf and then automatically perform prescribed statistics on the data set.

3.3 File Types

3.3.1 Meteorological Data Volume

The Meteorological Data Volume (MDV) format for gridded data was developed by the Research Applications Laboratory (RAL) at NCAR in the 1990s. At the time, several gridded data formats were used at RAL. For the simplification of these data systems, it was decided to standardise on a single gridded data type for internal use. MDV evolved as a data format unique to RAL and NCAR. It is described as an effective format for gridded data, with suffice meta-data support and efficient internal compression capability that allows for selected decompression of single planes from single data fields (Dixon, 2006).

In essence, MDV is a general-purpose data file format for storing two- and three-dimensional gridded data. It is a single-time format; each MDV data set contains data for a single time period. Time searching and retrieval are handled by a time-based file naming convention. MDV provides the capabilities for managing multiple data fields in a single file. Furthermore, the MDV format is extensible in that it provides space and access capabilities for optional generic “chunk” data defined by the MDV user. The chunk data allows MDV users to attach additional information that is not suitable for storage in the MDV headers or data fields to the data set (Dixon, 2006).

Figure 3.7 displays the layout of the data set for the MDV format, with the header lengths given in bytes. All the MDV header information appears at the beginning of the file. Thereafter, field and chunk data appear. The master header contains file offsets to the field header array, ‘Vlevel’ header array and the optional chunk header array. The field here contains file offsets of the field data. Vlevel headers are used to store the details of third dimensional data (such as radar elevation angles and Cartesian plane heights). The chunk headers contain file offsets to the chunk data.

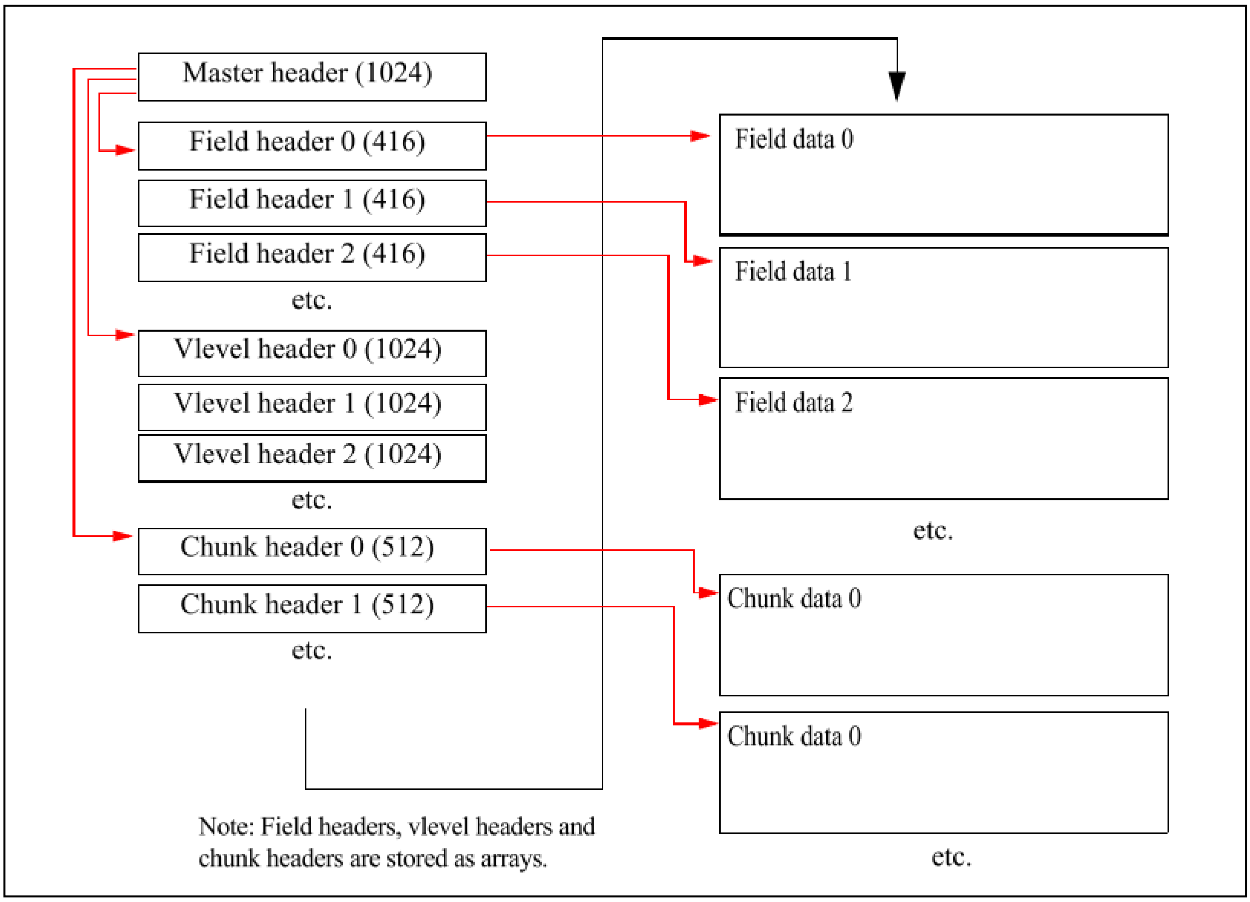


Figure 3.7: MDV data set structure organisation (Dixon, 2006)

The MDV file naming convention is essential for the retrieval of specific MDV files. Specifically, time is considered an important attribute for meteorological data. Thus, each MDV file contains data for a single time period. MDV files are named according to the time of data stored within the file. Internationally, UTC times are used. However, since the use of TITAN is well integrated with the SAWS, GMT (+2) is used. Times applicable for the naming convention include:

- Valid time – the time at which the radar observation occurred;
- Generate time – the time at which a model was executed, or a forecast generated;
- Forecast time – the time at which a forecast is considered valid; and
- Lead time – time difference between forecast and generated time for a forecast time.

For this project, MDV valid times were used, since storms that already occurred, have been analysed. Thus, the file name is as follows:

"data_dir/yyyymmdd/hhmmss.mdv"

The "yyyymmdd" implies the folder in which the data are stored, with "data_dir" being the directory for the storage of all MDV files. The "yyyy", indicates the year, "mm", the month and "dd", the date. The actual file is stored as a timestamp "hhmmss.mdv", with "hh", indicating the hour of 24-hour digital time of the observation, "mm" the minutes and "ss", the seconds. As an example, a storm captured on the 25 November 1995 at 06:30 AM would be in the following location, with the following file name:

"data_dir/19951125/063000.mdv"

More of the MDV Interface Control Document could be found on: https://www.eol.ucar.edu/system/files/MDV_format_ICD.pdf.

3.3.2 Network Common Data Form

The netcdf file is an interface of data access functions for storing and retrieving data in the form of arrays. An array can be considered as an *n*-dimensional rectangular structure (i.e. rows and columns of data) containing information which all has the same data type.

An excerpt from https://www.unidata.ucar.edu/software/netcdf/docs/netcdf_introduction.html describes netcdf files as follows:

"*Netcdf is an abstraction that supports a view of data as a collection of self-describing, portable objects that can be accessed through a simple interface. Array values may be accessed directly, without knowing details of how the data are stored. Auxiliary information about the data, such as what units are used, may be stored with the data. Generic utilities and application programs can access netcdf datasets and transform, combine, analyze, or display specified fields of the data. The development of such applications has led to improved accessibility of data and improved re-usability of software for array-oriented data management, analysis, and display.*"

The netcdf file format is compatible with various GIS environments, and commonly used as a global data format in atmospheric research, and the data can be curtailed to adhere to the project objectives.

The process to convert MDV data to netcdf was necessary and a MDV conversion function was formulated in order to convert the files to netcdf files. TITAN, the programme that generates the MDV files, was developed for operational settings, and is less suited for research purposes compared to typical GIS software (e.g. ArcGIS, QGIS, etc.). By converting the MDV files to a netcdf file type, it becomes more accessible and convenient for data analysis.

The MdvConvert function acts as a key component for the conversion from MDV to netcdf. The conversion process needs to be controlled carefully not to change valuable information when the file is converted. The gridded rainfall values within the MDV file are particularly important as they are essential to the derivation of ARFs. The following parameters needs to be set appropriately for the MdvConvert function to provide accurate data.

- "`apply_threshold_to_field_values`" limits the values (in this case rainfall) in specified fields between the "`min_threshold`" and "`max_threshold`" values.
- "`remap_pole_at_north`" is a flag indicating that stereographic occurs over the North Pole.
- "`ncf_set_global_attributes`" an option to set specific global attributes to the output netcdf file. Mainly includes strings as attributes.
- "`ncf_compress_data`" acts as an option to compress field data.
- "`ncf_output_latlon_arrays`" the latitude and longitude arrays of grid are in the output file.

Thus, after applying the conversion tool, a netcdf file was readily available to be used in the GIS environment.

3.4 Data Providers

One of the main objectives in the process of verifying the applicability of radar data for use in the evaluation of ARFs is to identify and contact the custodians of South Africa's Radar network, evaluate the current archival format of the radar data, and the availability and usability thereof. While various radar data sets are available, the retrieval of the data is a time-consuming process which requires some level of skill with the TITAN software and netcdf data format to ensure quality data conversion takes place. The process to convert multiple years of MDV data also requires high performance computing power. The capacity to execute the conversion and verification of the data in a format that is accessible to researchers not familiar with MDV data, as discussed above, is currently a time-consuming process.

Three main data providers have been identified for use in this project:

- SAWS;
- North-West University (NWU); and
- eThekwini Municipality.

While it is also known that the South African Airforce has several operational radars, all attempts to access their information failed in the timeframe of this project.

3.4.1 South African Weather Services

3.4.1.1 Radar data received in 2018

The SAWS originally provided data to the project team in MDV format. The following procedure was implemented to procure the data:

1. Hourly rainfall data (not radar data!) from automatic weather stations (AWS) were requested, for a two-year period from 1 Jan 2016 at 00:00 to 31 December 2017 23:00. This was required to identify major storm events to be investigated.
2. The hourly data was used to gather 3- and 24-hourly rainfall accumulations from the automatic weather stations. The top fifty storms were selected, in terms of rainfall accumulation, for each accumulation period. These storms were dated.
3. A request was sent to SAWS for radar images of the top 50 storms, referenced from the SAWS readings.

A detailed description of the data originally provided by SAWS is shown in Table 3.1.

Table 3.1: Data obtained from SAWS

Folder	Number of storms	Date range	Size	File type
1 hour	45	8/1/2016-30 /12/2017	7.38 MB	MDV
3 hours	41	8/3/2017-29/12/2017	36.2 MB	MDV
24 hours	42	9/3/2017-29/12/2017	100 MB	MDV

Individual radar images are merged into one field to produce a mosaic field for the whole of South Africa. This is done by means of TITAN composing scripts. The standard radar reflectivity measurements from the mosaic field and for each 6-minute radar scan is used in the standard Marshall-Palmer relationship to calculate the rainfall. Thereafter, TITAN scripts are used to accumulate the 6-minute rainfall values, to produce an hourly rainfall. This was carried out for the 3- and 24-hour periods as well.

Minor ground clutter, attenuation and occultation corrections were executed on the raw radar reflectivity data that was used to arrive at rainfall data. Many influencing factors on the radar is challenging to remove. However, there are possible corrections steps built into TITAN to deal with these influences. SAWS radars are S-band, which implies little attenuation and RLAN interference during operational use.

No bias corrections were performed on the radar derived rainfall data. Since the data is estimates of rainfall based on radar reflectivity, it is possible to overestimate or underestimate rainfall values. The rainfall data received directly from SAWS are assumed to be satisfactory estimates for this project as it makes use of the storm-centred ARFs. Thus, the only errors that may produce incorrect ARFs are those that affect both the calculation of the areal rainfall and the maximum point rainfall. From a theoretical point of view, if the reflectivity is overestimated, the areal rainfall is consequently overestimated, as well as the maximum point rainfall. Once the areal rainfall is divided by the point rainfall, the assumption was made that it nullifies out any errors that are brought upon by the reflectivity errors. After the data was obtained, it was converted to netcdf using the conversion software (MdvConvert). Once all the MDV files were converted to netcdf, Panoply was used to view the data which facilitated an overall understanding of the information contained in the file.

Table 3.2 provides the date that each radar (operated by SAWS) was commissioned, along with its operational frequency and capabilities.

Table 3.2: Radar network from SAWS (Becker and Pegram, 2014)

Radar	Date Commissioned	Frequency Band	Capabilities
Bethlehem	March 2010	S-band (10cm)	Doppler & Dual-Pol
Bloemfontein	July 2011	S-band (10cm)	Doppler
Cape Town	< 1998	C-band (5cm)	Reflectivity Only
Cape Town Int. Airport	Not yet active	X-band (3cm)	Doppler & Dual-Pol
De Aar	< 1998	C-band (5cm)	Reflectivity Only
Durban	May 2011	S-band (10cm)	Doppler
East-London	May 2011	S-band (10cm)	Doppler
Ermelo	November 2010	S-band (10cm)	Doppler
George	January 2012	S-band (10cm)	Doppler
Irene	January 2010	S-band (10cm)	Doppler
OR Tambo Int. Airport	Not yet active	X-band (3cm)	Doppler & Dual-Pol
Ottosdal	November 2010	S-band (10cm)	Doppler
Mthatha	March 2010	S-band (10cm)	Doppler
Polokwane	November 2010	S-band (10cm)	Doppler
Port-Elizabeth	< 1998	C-band (5cm)	Reflectivity Only
Skukuza	February 2007	S-band (10cm)	Reflectivity Only

Figure 3.8 displays the radar network in South Africa. The white rings represent S-band frequency radars, and the green, the C-band frequency radars. The smaller rings represent a range of 200 km, with the larger rings covering a range of 300 km. Blue dots indicate radars that make use of Doppler capabilities, with the red dots showing the lack thereof. The turquoise dot represents Doppler and Dual-Polarisation capabilities.

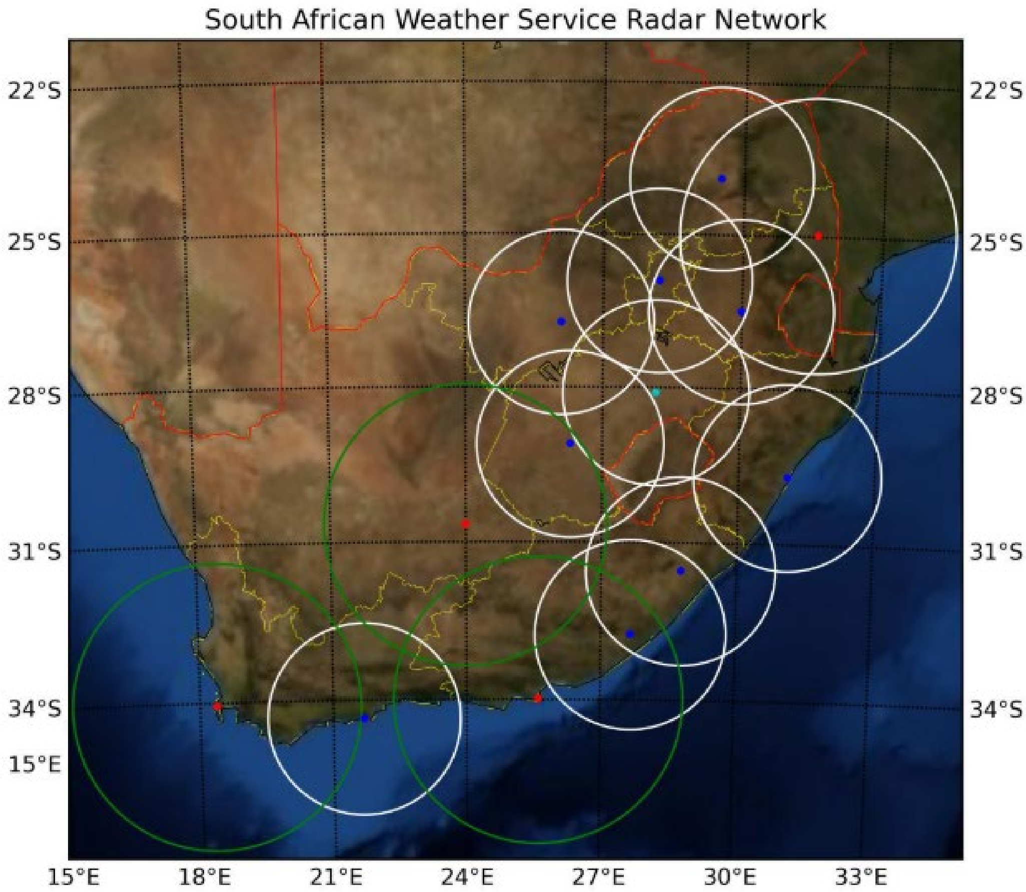


Figure 3.8: Current radar network employed by SAWS (Becker and Pegram, 2014)

3.4.1.2 Radar data

Within the limited capacity of the SAWS, additional data was made available during a second phase. The data set comprises of files for 3 hour intervals, which, for any particular day, starting at 0:00, 3:00, 6:00, 9:00 …till 24:00. As an example, one particular file would contain the georeferenced rainfall accumulation starting at 0:00 and ending at 03:00. Table 3.3 displays a detailed summary of the data.

Table 3.3: Radar data obtained from SAWS in 2019

Folder	Number of storms	Date range	Size	File type
3 hours	Varying	2/3/2016-13/10/2019	680 MB	MDV
6-minute rainfall	Varying	2/3/2016-14/10/2016	13.6 GB	MDV

3.4.2 North-West University

North-West University (NWU) has supplied MDV files from the Irene radar. The data contain the 6-minute rainfall readings for an extensive record, hourly rainfall data and raw radar data. This is displayed in Table 3.4. The MDV data must be converted to netcdf in order to be used in a GIS environment. The 'vol' files contained radar information and is obtained directly from the radar itself and is not needed for this project. The resolution of the rainfall data is 0.5 km x 0.5 km.

Table 3.4: Radar Data obtained from NWU

Folder	Date range	Size	File type
1-Hour data	14/1/2013-18/4/2017 (not usable file info)	796.7 GB	netcdf
3-Hour data	14/1/2013-18/4/2017	56.1 GB	netcdf
24-Hour data	14/1/2013-18/4/2017	3.9 GB	netcdf

One of the challenges with the use of two different data sets, is the difference in post analysis, resulting in possible differences in the accuracy thereof. Radar data from NWU does not have the same level of post analysis to convert data accurately, as what was done (received from) by the SAWS. The MDV files obtained from SAWS have some correction calculations applied to the data itself. This presents a potential problem of inconsistency in the data being analysed, which could lead to varying results.

3.4.3 eThekwini Municipality

The eThekwini Municipality (Durban) recently commissioned their own radar to assist them in flood forecasting. Data is available from them for use, but were not received, despite various attempts, for use in this project.

3.5 Derived Data

The derived products from the TITAN software contains various geo-referenced atmospheric variables that the user can view, analyses, and do geo-spatial statistics on. These were discussed in Section 3.2.1. Other derived fields that allow the user to understand the type, intensity, and characteristics of the specific event are also available, which includes well known parameters such as:

- Cloud Top (km MSL);
- Cloud Base (km MSL);
- Volume (km^3);
- Mean area (km^2);
- Precipitation flux (m^3/s); and
- Mass (ktons).

Severe storm indicators, which can be useful to identify cloud types and extreme events, are also available:

- FOKR category: 0-4;
- Waldvogel probability: 0-1;
- Hail mass aloft: ktons; and

● Vertically Integrated hail mass, from max Z: kg/m^2.

A raw derived data file structure is presented as an example in Figure 3.9.

```
NSimpleTracks ComplexNum SimpleNum Year Month Day Hour Min Sec dBZThreshold RemainingDuration(hrs) VolCentroidLat(deg) VolCentroidLon(deg) VolCentroidX(km) VolCentroidY(km) VolCentroidZ(km)
ReflCentroidLat(deg) ReflCentroidLon(deg) ReflCentroidX(km) ReflCentroidY(km) ReflCentroidZ(km) PrecipCentroidLat(deg) PrecipCentroidLon(deg) PrecipCentroidX(km) PrecipCentroidY(km) PrecipArea(km2)
PrecipOrientation(degT) PrecipMajorRadius(km) PrecipMinorRadius(km) PrecipFlux(m3/s) PrecipRate(mm/hr) EnvelopeCentroidLat(deg) EnvelopeCentroidLon(deg) EnvelopeCentroidX(km) EnvelopeCentroidY(km)
EnvelopeArea(km2) EnvelopeOrientation(degT) EnvelopeMajorRadius(km) EnvelopeMinorRadius(km) Top(km) Base(km) Volume(km3) MeanArea(km2) Mass(ktons) TiltAngle(deg) TiltOrientation(degT) MaxDBZ(dBZ)
MeanDbz(dBZ) MaxDBZGradient(dBZ/km) MeanDbzGradient(dBZ/km) HtOfMaxDBZ(km) Vorticity(/s) VilFromMaxZ(kg/m2) U(km/hr) V(km/hr) Dtop/Dt(km/hr) Dvolume/Dt(km3/hr) DprecipFlux/Dt(m3/s2) Dmass/Dt(ktons/hr)
DdBzMax/Dt(dBZ/hr) Speed(km/hr) Dirn(DegT) %Vol>40dBZ %Area>40dBZ %Vol>50dBZ %Area>50dBZ %Vol>60dBZ %Area>60dBZ %Vol>70dBZ %Area>70dBZ HailFOKRCat0-4 HailWaldvogelProb HailMassAloft(ktons) HailVihm(kg/m2)
parents children
3 3 3 2013 11 14 15 04 50 35 -384567 -27.4715 28.2702 5.86698 -173.43 4.88659 -27.4703 28.2705 5.89557 -173.304 4.78173 -27.4731 28.2726 6.10241 -173.608 41.5 58.8952 4.95578 2.66555 96.9757 8.41235
-27.4719 28.2719 6.03613 -173.481 43.25 58.5934 4.82061 2.85584 7.625 3.125 131.438 29.2083 61.3374 32.5074 324.905 40.5 37.4373 -0.0952381 -0.335025 6.5 0 3.09034 50.7729 -15.2988 -7.43252 -212.428
-333.564 -99.1331 -26.415 53.0277 106.769 21.8735 29.3173 0 0 0 0 0 0 0 0 0 0 - 15,16
1 6 6 2013 11 14 15 04 50 35 -384567 -24.2794 28.2548 4.46197 181.527 4.85194 -24.2801 28.2506 4.03913 181.447 4.80118 -24.2787 28.2556 4.54762 181.599 31.5 72.7716 5.33083 1.8809 66.5596 7.60682 -24.2787
28.2556 4.54762 181.599 31.5 72.7716 5.33083 1.8809 7.625 3.875 57.9375 15.45 25.4995 52.1066 207.187 40.5 36.9902 -1.66666 -0.580099 5 0 2.0768 0.0754605 -11.9598 -1.57384 -133.682 -240.223 -58.8363
-17.501 11.9601 179.638 17.26 21.164 0 0 0 0 0 0 0 0 0 0 - -
1 8 8 2013 11 14 15 04 50 35 -384567 -24.7999 28.6399 43.3168 123.579 5.78415 -24.8083 28.6315 42.475 122.643 4.69012 -24.7993 28.6383 43.1639 123.64 52.25 48.3253 4.80792 3.45923 383.713 26.4377 -24.8
28.642 43.5338 123.562 57.25 64.4247 4.91554 3.70727 10.625 2.375 312.938 37.9318 297.954 9.68864 110.72 56.5 44.0609 -2.27272 -1.30293 3.5 0 24.0283 11.135 -1.66184 3.25243 925.236 1413.03 792.842 52.0389
11.2583 98.4885 62.9119 83.8915 13.2415 27.2727 0 0 0 0 2 0.4 0 0.785394 - -
1 10 10 2013 11 14 15 04 50 35 -384567 -27.0153 27.9597 -24.8757 -122.727 4.33716 -27.0156 27.9592 -24.9191 -122.765 4.23331 -27.0151 27.9605 -24.7955 -122.71 24.75 33.4546 3.5931 2.19258 62.1188 9.03547
-27.0151 27.9605 -24.7955 -122.71 24.75 33.4546 3.5931 2.19258 6.875 2.375 69.375 15.4167 33.8232 28.3907 15.6533 41.5 37.794 -1.12381 -0.539231 3.5 0 3.02888 0 0 0 0 0 0 0 0 0 28.4685 38.0471 0 0 0 0 0 0 0
0 0 0 - -
```

Figure 3.9: Raw derived data file structure

3.5.1 Quality control

The first quality control is to check the data integrity. To enable this, the data is plotted on a polar grid to monitor if any data fall outside the expected radar range. The following *R* script is used to load and visualise the data.

```
#Load Libraries
library(ggplot2)
library(tidyverse)
library(reshape2)
library(viridis)
library(grid)
library(readr)
library(gridExtra)
library(raster)
library(GSODR)
library(rgeos)
library(sp)
library(data.table)
library(magrittr)
library(data.table)
library(readr)
theme_set(theme_classic(base_size=22))

# Load data
df <- fread("StormData.csv", header=TRUE, sep=" ")

# We can quickly inspect the data (df)
head(df)

# To inspect individual values we can use the summary function
summary(df$PrecipArea.km2.)
     Min.  1st Qu.   Median     Mean  3rd Qu.    Max.
     0.00  26.50 47.25   131.20   103.50 24143.50

# Now we can quickly define our basemap for South-Africa
basemap <- subset(map_data("world"), region %in% c("South Africa"))
```

```
# To visualize the data and see if any points fall outside the
# bounds
ggplot() +
      geom_path(data=basemap, aes(x=long, y=lat, group=group)) +
      geom_jitter(data=df, size=0.05, aes(x=ReflCentroidLon.deg.,
      y=ReflCentroidLat.deg., color=MaxDBZ.dBZ.)) +
      coord_map(ylim=c(max(df$ReflCentroidLat.deg.) + 0.5,
      min(df$ReflCentroidLat.deg.) - 0.5),
      xlim=c(min(df$ReflCentroidLon.deg.) - 0.5,
      max(df$ReflCentroidLon.deg.) + 0.5)) +
      scale_colour_viridis_c(direction=-1) +
      ylab("Latitude") + xlab("Longitude")

# This script can easily be modified to visualize the
# spatio-temporal integrity of the data, eg:
ggplot() +
      geom_path(data=basemap, aes(x=long, y=lat, group=group)) +
      geom_jitter(data=df, size=0.05, aes(x=ReflCentroidLon.deg.,
      y=ReflCentroidLat.deg., color=MaxDBZ.dBZ.)) +
      coord_map(ylim=c(max(df$ReflCentroidLat.deg.) + 0.5,
      min(df$ReflCentroidLat.deg.) - 0.5),
      xlim=c(min(df$ReflCentroidLon.deg.) - 0.5,
      max(df$ReflCentroidLon.deg.) + 0.5)) +
      scale_colour_viridis_c(direction=-1) +
      facet_wrap(Month ~ .) +
      ylab("Latitude") + xlab("Longitude")
```

The first ***ggplot*** call generates a figure to quickly inspect the spatial integrity of all the data points after MDV, and netcdf conversion have taken place. From Figure 3.10, it is clear that the data is structurally intact and no data fall outside the bounds. If there was any spatial error, the data processing scripts would have needed to undergo additional checks to establish whether the processing was erroneous or if the data was corrupt.

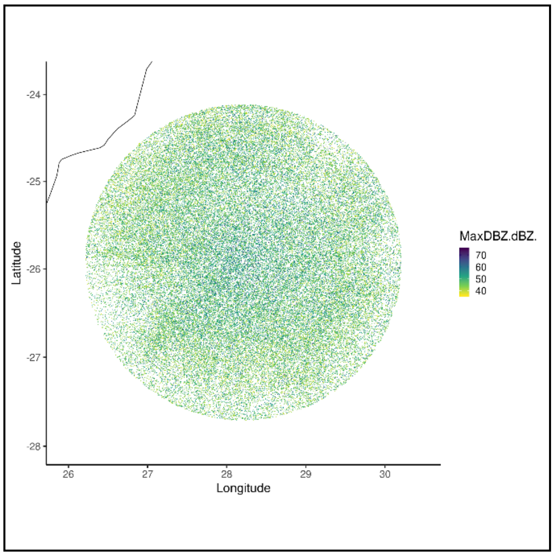


Figure 3.10: Spatial integrity of all the data points

The second **ggplot** call is an example of a spatial-temporal visualisation of the data, which allows for the assessing of storms to see if they are captured correct in time, using current knowledge on the seasonal distribution of thunderstorm events. This is shown in the Figure 3.11.

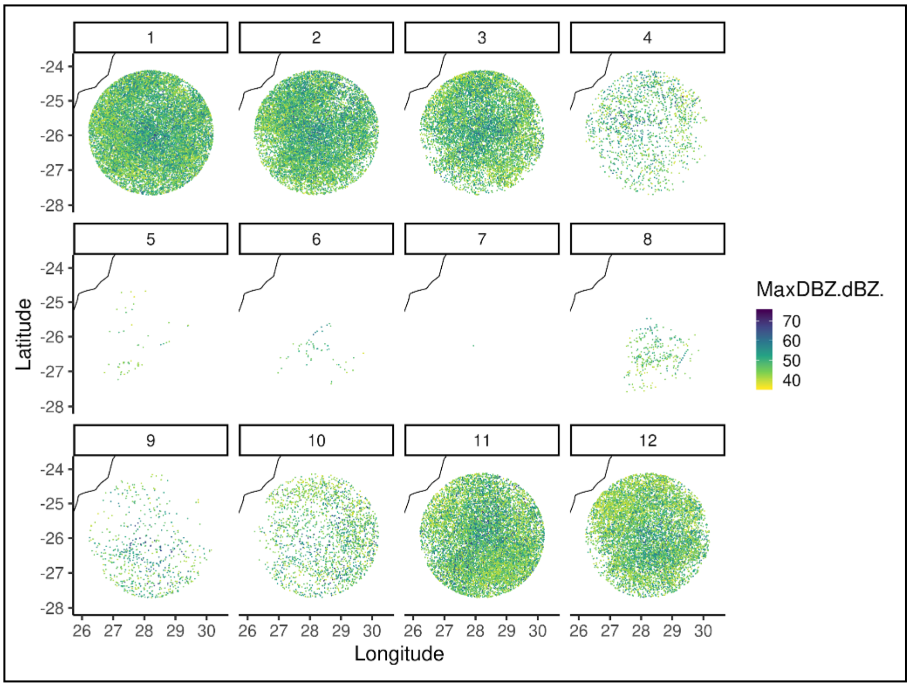


Figure 3.11: Spatial-temporal visualisation of the data

The data processed from the Irene radar as illustrated in Figure 3.11 provide confirmation that the spatial and temporal extend agree with the expected results. The data fall within the 200 km radius and no data points are out of bounds. This is confirmed as the function in the *R* script plots the data based on the minimum and maximum Lat/Long limits and not a user defined minimum or maximum coordinate system. Temporally most events start occurring in October through to March. This was expected as summer is when the Highveld Region receives most of its rainfall, mostly in the form of convective events.

More thorough statistical analysis can be done on the data set by means of correlation statistics. *R* and it's parallelisation capabilities can be used to quickly process and correlate the data:

```
#Load Libraries
library(parallel)
library(MASS)
library(foreach)
library(doParallel)

# Let R get the specs of your system, in this case we have 48
# processors. We leave 5 open to avoid a system crash
numCores <- detectCores() - 5
numCores
registerDoParallel(numCores)

# For correlation statistics we use only variables that are
# logically sensible to use. We select the column names of these
```

```
# variables we want to use
colnames <- c(names(df)[10:62], names(df)[71:76])
plot_list = list()

# Here we run our first loop, this is not so computationally
# expensive so we do not parallelize yet. In the plots that we do
# here x= the maximum reflectivity while y = all the variables
# defined in colnames. So we'll have correlation statistics on the
# maximum dbz and all other variables. It is easy enough to change x
# to whichever variable needed
for (i in colnames)
{
	plt <- ggplot(df, aes_string(x="MaxDBZ.dBZ.", y=i,
color="MaxDBZ.dBZ.")) +
	geom_jitter(alpha=0.2) +
	ggtitle(paste("Brazil - Max DbZ", i)) +
	ylab(paste(i)) + xlab("Max DbZ") +
	scale_fill_viridis_c() +
	labs(fill="Max DbZ") + guides(color=FALSE)
	plot_list[[i]] = plt
}

# Here we use the full power of parallelization to quickly plot all #
the data correlation graphs and save them. We can change the
foreach(i = colnames) %dopar% {
	file_name = paste("Irene_Maxdbz_CorrPlots", i, ".png", sep="")
	png(file_name, width = 8 * 500, height = 4 * 500, res = 300)
	print(plot_list[[i]])
	dev.off()
}
```

After running the above script, the correlations graphics can be viewed for all the variables (with Max DbZ being the dependent variable). Figure 3.12 is an example of the output from the script provided above and clearly illustrates the relationship between the Maximum Reflectivity and cloud top height. Both are important to distinguish between various rainfall and cloud types. Typically, high reflectivity and a high cloud top indicates the presence of cumulonimbus clouds which can lead to extreme rainfall and hail events. The correlation scripts allow the user to quickly inspect various relationships among the data points.

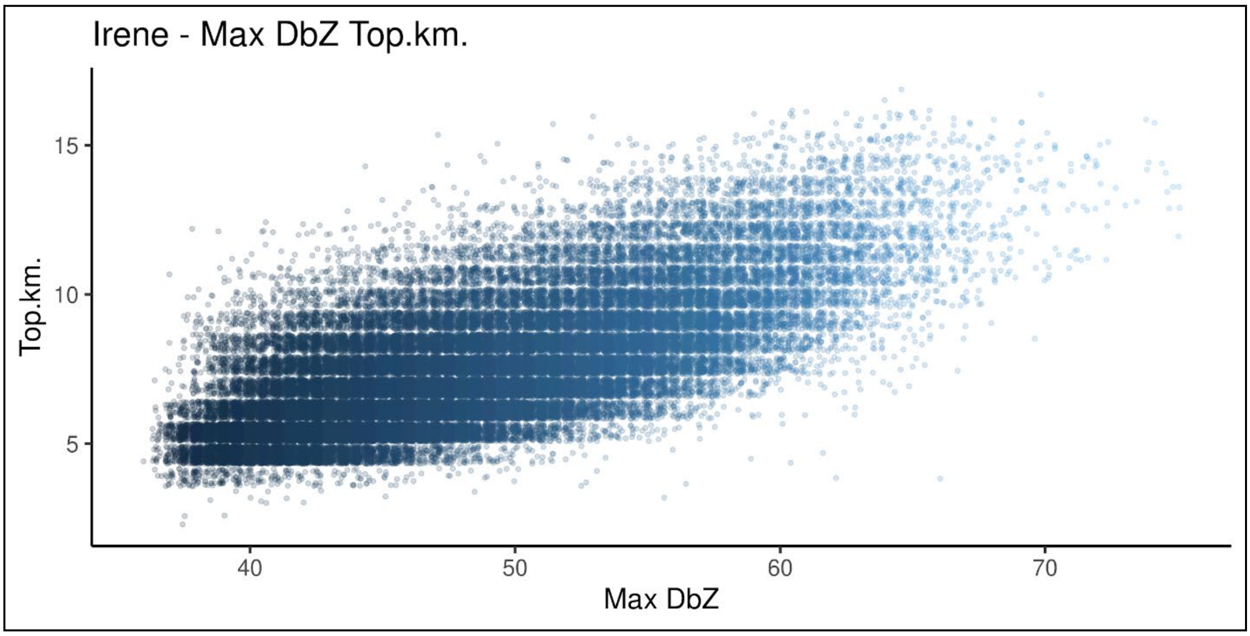


Figure 3.12: Relationship between the maximum reflectivity and cloud top height

3.6 Radar Reflectivity Problems

Reflectivity measurements must be carefully corrected and extracted before rainfall can be estimated. Beam blocking, anomalous propagation hail and bright band are a few of the errors that contribute to incorrect rainfall estimates. These will be discussed in more detail in the following sections.

3.6.1 Beam blocking

When a radar beam meets a fixed object (for example, a mountain or skyscraper), it constitutes to beam blocking. Beam blocking tends to be a significant source of error within rainfall estimation. Generally, two types of beam blocking occur. Partial beam blocking is when the apex or upper segment of a fix object is within range of the beam, causing power losses and severe underestimation of rainfall estimations. The second type is called total beam blocking, in which a fixed object comes into the full range of the beam. This causes full power blocking with no rainfall estimations to be read beyond the range. Thus, this makes beam blocking extremely dependent on the surrounding topography of its location (Becker and Pegram, 2014).

3.6.2 Bright band

Bright band occurs when observations of a uniform band of higher reflectivity are captured just beneath the zero-degree isotherm (at freezing level). Higher reflectivity is a result of several characteristics of electromagnetic waves and hydrometers within the atmosphere. These include the differences of ice and water with regards to their reflective properties, density above and below the melting level and terminal velocity. This phenomenon is observed in more organised stratiform rainfall, where a clear distinction in particles between different layers in the atmosphere is present. When particles precipitate from a cloud and the ice particles move through the melting layer, they start to melt from the outside inward (Becker and Pegram, 2014).

3.6.3 Ground clutter

Ground clutter appears when the main beam intersects the ground. This results into an echo returned to the radar. The targets typically consist of mountains and tall structures in the immediate area close in proximity to the radar. Most times, the object causing ground clutter is easily identified. Usually, a radar clutter map is used to identify and delete ground clutter from a weather radar display (Becker and Pegram, 2014).

3.6.4 Anomalous Propagation

Anomalous Propagation (AP) is defined as the extended detection of ground targets. It usually occurs in clear weather conditions with the presence of a temperature inversion, or whenever the water vapour content of the atmosphere is at relatively high levels. This results in a more refracted radar beam, as opposed to one obtained in normal atmospheric conditions. In extreme cases, this can cause the beam to curve toward the surface of the Earth (Becker and Pegram, 2014).

3.7 Converting Radar Reflectivity to Rainfall

After corrections for reflectivity errors are applied, the conversion of the reflectivity to rainfall values can be done. The Marshall-Palmer relationship is one of the most commonly used relationships that is applicable to stratiform rainfall as well as radar derived rainfall. The relationship was derived by comparing Drop Size Distributions and radar reflectivity measurement from numerous rainfall events. The relationship is expressed in

Equation 6:

$$Z = 200 \cdot R^{1.6} \qquad \text{Equation 6}$$

Where:

Z in mm^6 per m^3 and

R in mm per h.

This project makes use of this relationship by default, as it is received from SAWS with the relationship already being applied.

3.8 Radar Data Challenges

Radar data has a reputation of producing large amounts of errors and the accuracy of radar data is highly dependent on a multitude of factors and parameters. Thus, the following challenges were encountered when analysing and using the radar data, particularly within a GIS environment:

- Anomalous rainfall reflectivity values that are not consistent with surrounding rainfall values.
- Inconsistent work environments to arrive at a file that is compatible in a GIS environment can become an inconvenience. For example: Some radar files are processed using a Linux system. The file naming convention on the Linux system is not the same as other operating systems. A Linux system then becomes a crucial part of the process for file name changing.
- The data within some radar files are inconsistent and noticeably erroneous.

- MDV files have to be converted to netcdf files to be used in a GIS environment.
- Extreme storm events are many times not recorded by the radar as it is frequently non-operational on that particular day.
- Some radar files, as received from the data providers, contain no useful information, except erroneous values that are unrealistic.

3.9 Summary of Radar Information for this Study

Individual radar images are merged into one field to produce a mosaic field for the whole of South Africa. This is done by means of TITAN composing scripts. The standard radar reflectivity measurements from the mosaic field and for each 6-minute radar scan is used in the standard Marshall-Palmer relationship, described in Section 3.7, to calculate the rainfall rate. Thereafter, TITAN scripts are used to accumulate the 6-minute rainfall rates, to produce an hourly rainfall. This is carried out for 3- and 24-hour periods.

Minor ground clutter, attenuation and occultation corrections were carried out on the raw radar reflectivity data that was used to arrive at rainfall data. While many artefacts on the radar is said to be a challenge to remove, TITAN does allow for corrections for scenarios like this. Bright band and anomalous propagation corrections were needed to be calculated separately. Most of SAWS radars are S-band, which implies little attenuation and RLAN interference.

No bias corrections were performed on the radar rainfall data. Since the data is estimates of rainfall based on radar reflectivity, it is possible to overestimate or underestimate rainfall values. The rainfall data received directly from SAWS are assumed to be satisfactory estimates for this research as it makes use of the storm-centred method of deriving ARFs. It was assumed in this project that possible errors in the conversion of radar data to actual rainfall data will affect both areal rainfall and the maximum point rainfall, which will be eliminated during the calculation process of ARFs.

Becker and Pegram (2014) investigated the accuracy of rainfall estimates, comparing TITAN rainfall estimates with rainfall station data. The results are presented in Table 3.5. It shows a slight tendency of overestimation of rainfall, due to the bias being greater than 1. Their research mentions the following:

"The S-band radar at Irene produces high quality reflectivity data that is free from RLAN interference and less effected by attenuation, particularly over short ranges from the radar. This is favourable for rainfall estimates. Ground clutter from the reflectivity data has been removed as well as possible by the Doppler filter." – (Becker and Pegram, 2014).

Table 3.5: Scores calculated from contingency tables using various thresholds, for hourly measurements (Becker and Pegram, 2014)

	Scores (1mm)	Scores (5mm)	Scores (10mm)
Accuracy	0.973	0.99	0.995
Bias	1.6	1.719	1.766
Probability of Detection	0.911	0.848	0.728
False Alarm Ratio	0.431	0.507	0.588
Critical Succes Index	0.539	0.453	0.357
Equitable Thread Score	0.524	0.449	0.355
Heidke Skill Score	0.687	0.619	0.524

Despite the problems expected to be encountered in the different steps of radar data collection and processing, the data was deemed appropriate for use in this research to evaluate the applicability of the radar data in further research into the calculation of ARFs. In the next section, the methodology used to derive ARF with the radar data will be discussed in detail.

4 METHODOLOGY

4.1 Introduction

The methodology adopted using radar imagery to derive ARFs, is discussed in this Chapter. An overview of the research plan, including the methods of obtaining radar data, is presented. Thereafter, the execution of 'converting' radar imagery to ARFs is explained by utilising ArcGIS and Excel.

4.2 Research Plan

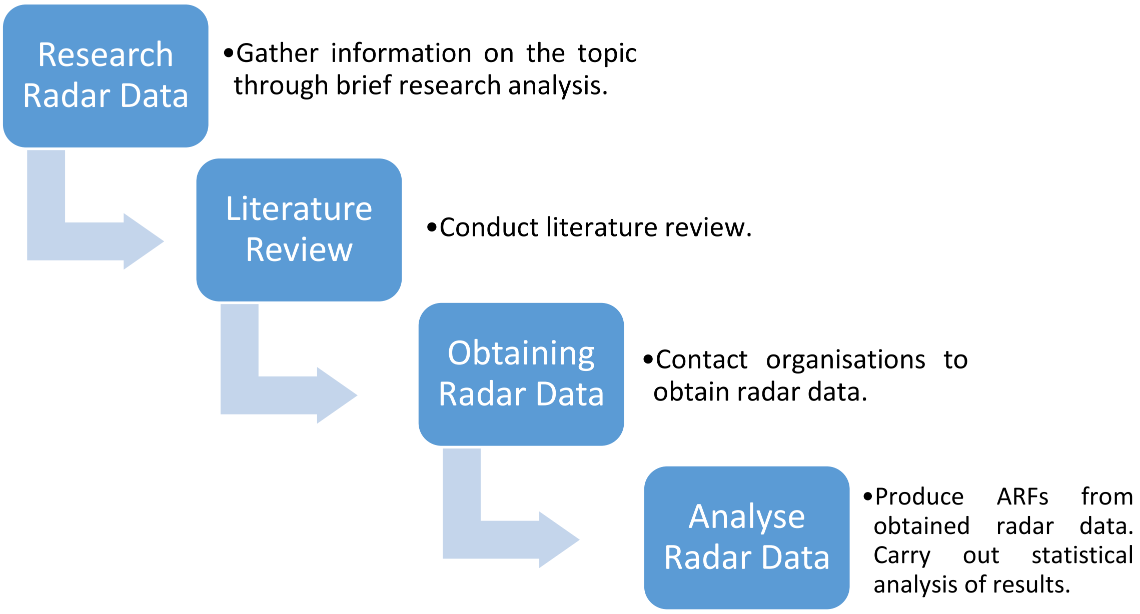


Figure 4.1: Broad research plan to arrive at end results

Figure 4.1 displays the steps followed in this project. First, brief research was carried out to establish the need for the research in the hydrological field on a global and national scale. Thereafter, a literature review was conducted, to gain perspective on the progress of the research field of radar data and ARFs at a global and national scale. Organisations and companies that could produce radar data were approached. After the radar data was obtained, a methodology was put in place to arrive at ARFs.

The topic of radar data has become more and more popular over the years, especially for the hydrological fields. From flash flood warning systems to ARFs, radars are proving their true potential in the ability to add knowledge and significant research to the hydrological field. On a global scale, radars have been extensively used in research, and radars themselves, have become more refined and streamlined to fit certain functions in order to obtain knowledge that is useful. On a national scale, radars are proving to be more and more useful towards South African hydrology as highlighted by researchers such as Terblanche, Pegram and Mittermaier (2001). The SAWS forms an integral part of the process to use radar data in hydrology.

4.2.1 Obtaining storms for analysis

Only a selected number of storms was obtained based on the rainfall record that was used for storm dating. SAWS was not able to provide the radar data for the missing storms due to the radar being non-operational, for said storm dates. Thus, with the given radar data from SAWS, storms were selected based on radar image quality, location (convective or frontal region) and size.

4.2.2 Analysis of radar data to arrive at an ARF

Approximately 100 storm files, in MDV format, were received from SAWS. An extensive process was carried out to produce ARFs from the MDV file format. A broad four step methodology, as presented in Figure 4.2, was followed.

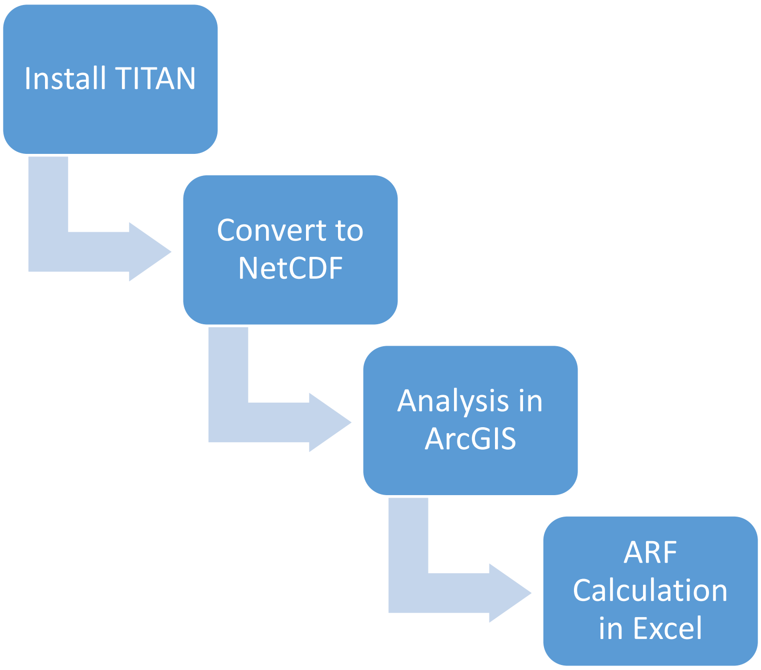


Figure 4.2: Broad methodology to arrive at ARFs

4.2.2.1 TITAN

The data provided from SAWS was in MDV format, which is extensively described in Section 3.3. TITAN was installed following the installation guide found on https://ral.ucar.edu/projects/titan/docs/ . TITAN is known for being a specialised software tool that is used by experts. This research aims to make radar data more usable for hydrologists in South Africa. A conversion tool developed by De Waal (2018) for TITAN, creates a netcdf that is compatible in GIS systems. The software code is presented in APPENDIX A. Figure 4.3 shows TITAN displaying a 1-hour storm event that occurred on 2018/04/11 at 16:00. This resulted from the execution of the "`start_all`" function.

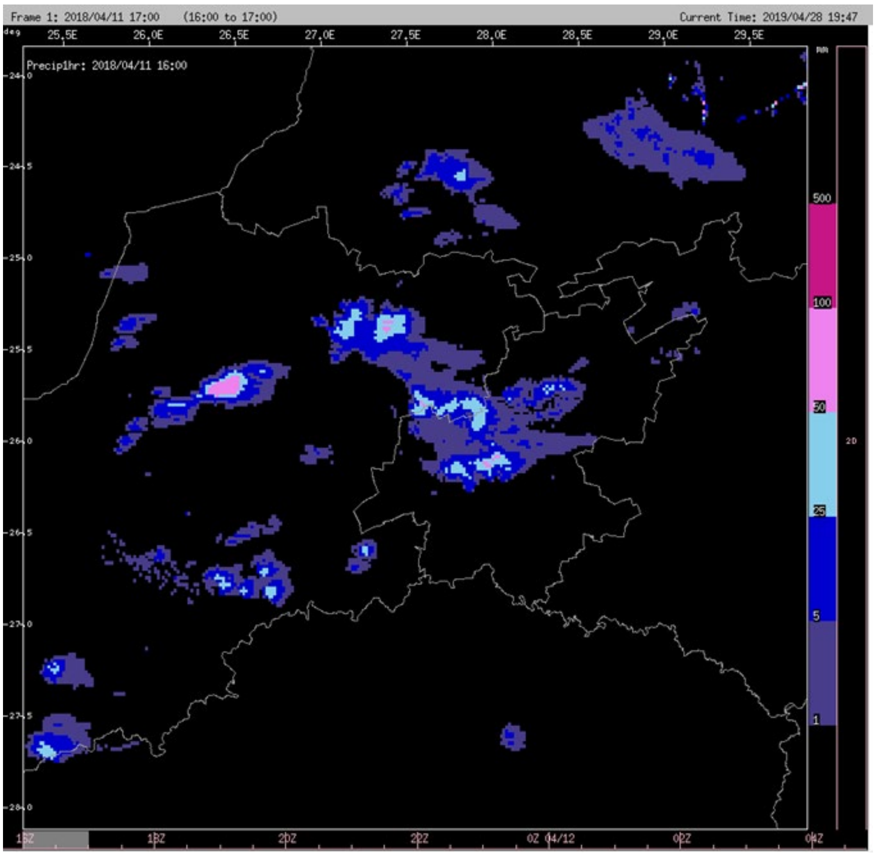


Figure 4.3: One-hour storm rainfall event in TITAN (MDV format)

4.2.2.2 Converting to NETCDF

A MDV conversion function was formulated in order to convert the MDV files to netcdf, which is compatible in most GIS systems. This was done as TITAN, the programme that hosts MDV files, is less suited for the research than typical GIS software (e.g. ArcGIS, QGIS, etc.). By converting this to a netcdf file type, it becomes more accessible and convenient for data interpretation.

The MdvConvert function acts as a key component for the conversion from MDV to netcdf. Careful factors need to be considered when it comes to file conversion. The gridded rainfall values within the MDV file are particularly important as they are the essential to the derivation of ARFs. The following is considered and integrated into the MdvConvert function as “`TRUE`”, found in APPENDIX A:

- “`_latest_data_info_`” was written for the output files
- “`apply_threshold_to_field_values`” limits the values (in this case rainfall) in specified fields between the “`min_threshold`” and “`max_threshold`” values
- “`remap_pole_at_north`” is flag indicating stereographic occurs over the North Pole
- “`ncf_set_global_attributes`” an option to set specific global attributes to the output netcdf file. Mainly includes strings as attributes
- “`ncf_compress_data`” acts as an option to compress field data
- “`ncf_output_latlon_arrays`” the latitude and longitude arrays of grid are in the output file

The storm event displayed in Figure 4.3 is in MDV format. Thus, after applying the conversion tool, a netcdf file was readily available to be used in ArcGIS. The following code was executed:

“`MdvConvert -params MdvConvert.1hr -f /home/titan5/projDir/data/mdv/precip/1hr/20180411/160000`”

4.2.2.3 Analysis in ArcGIS

After conversion of the MDV to netcdf, the file was imported into ArcGIS. An example for one file, specifically a 1-hour storm occurring on 9/03/2016 at 16:00, is thoroughly explained in this section and this methodology is applied for every netcdf storm file. The following methodology was put in place to arrive at Excel input values:

1. ArcGIS was setup to have a Projected Coordinate System, with the characteristics displayed in Figure 4.4.

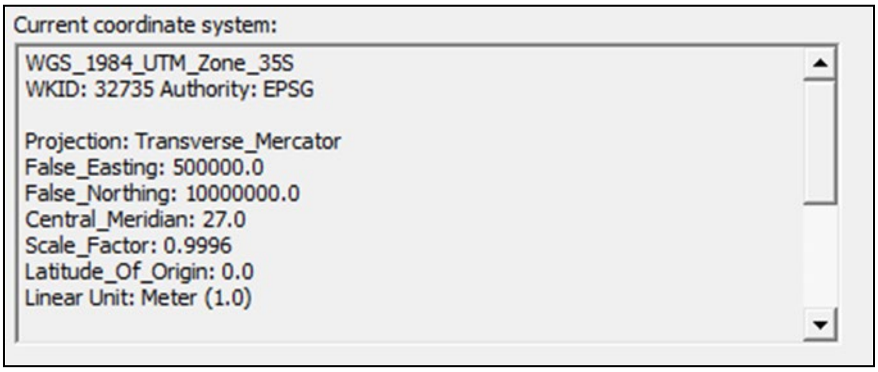


Figure 4.4: Projection coordinate system in ArcGIS

2. The netcdf file was imported in ArcGIS as a raster layer with the setting displayed in Figure 4.5. The netcdf layer uses the same projected coordinate system as in Point 1.

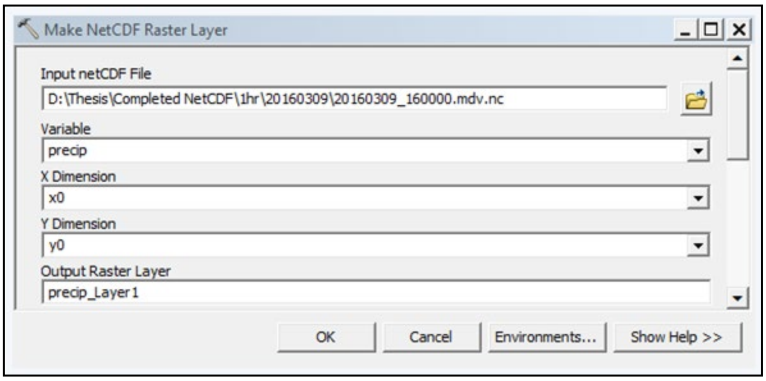


Figure 4.5: Importation of a netcdf file into ArcGIS

3. The netcdf, now considered as a raster file, need to be re-classified for the rainfall to be represented by isohyets. The ‘Reclassify’ function in ArcGIS can do this with the settings displayed in Figure 4.6. Isohyet intervals of 5 mm were chosen, for accuracy purposes.

The ‘Old values’ in Figure 4.6 represent the grouping of gridded values found in the rainfall data. The ‘New values’ represent the label value of each isohyet.

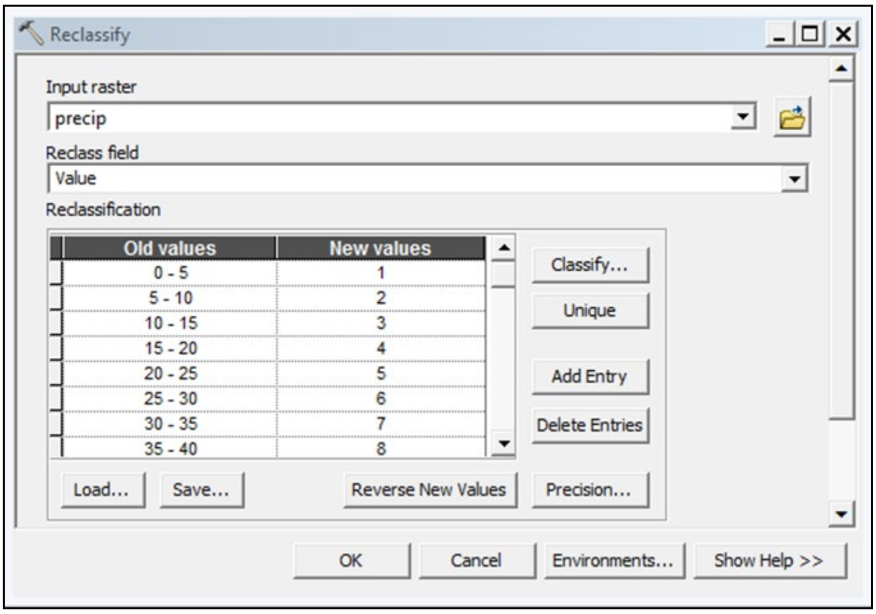


Figure 4.6: Reclassify parameters used in ArcGIS

4. From the reclassified raster layer, a single cell storm is selected. This is done by drawing a polygon around the storm. The single cell storm is extracted using the raster processing function called ‘Clip’, which uses a polygon as the clipping extent. The single sell storm is displayed in Figure 4.7.

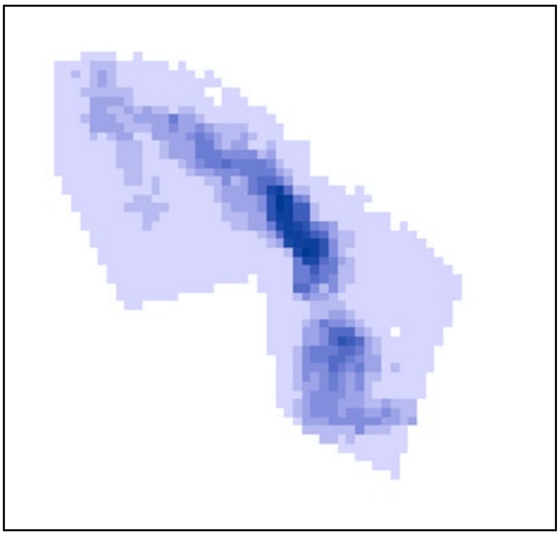


Figure 4.7: Single cell storm extracted from the reclassified layer using a polygon as the extracting Extent

5. The 'Zonal Statistics as Table' function was used to gather information of the single cell storm in Point 4. Figure 4.8 illustrates the settings used to arrive at a table of information of the single cell storm.

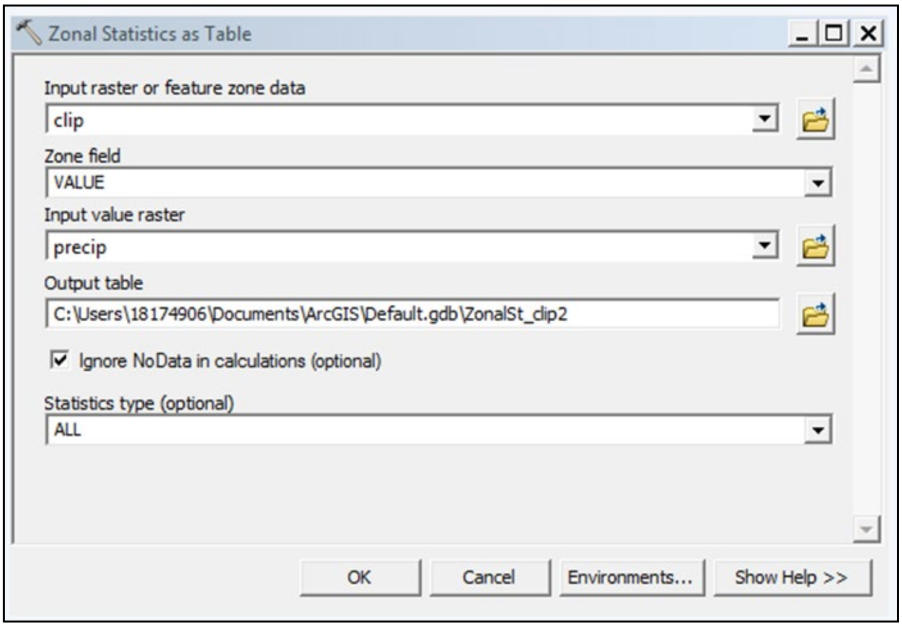
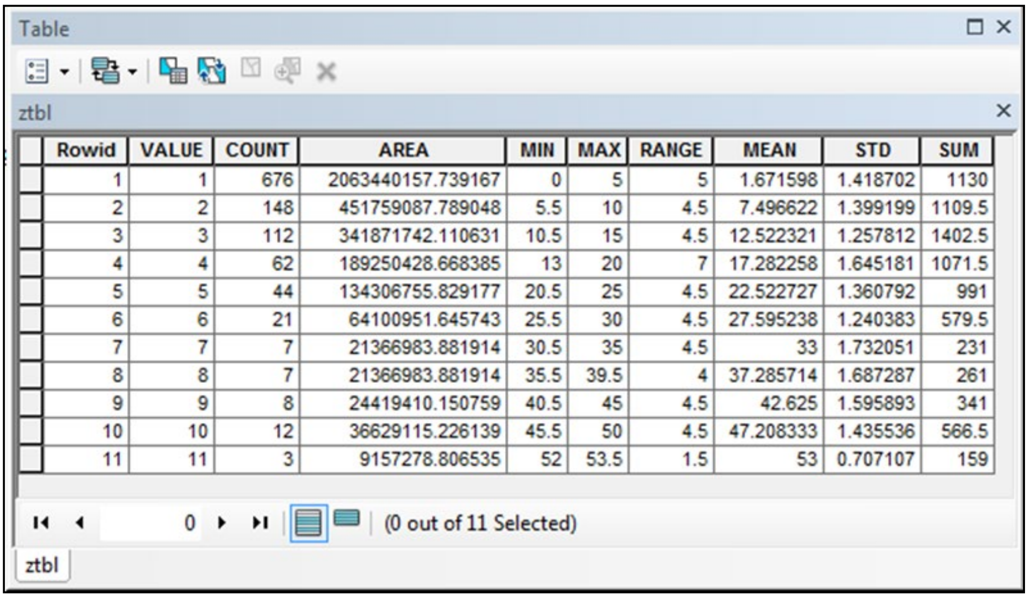


Figure 4.8: Zonal Statistics as table function settings in ArcGIS

6. The Zonal Statistics Table output is shown in Figure 4.9. The heading 'VALUE' represents the re-classed interval, 'COUNT' represents the number of cells within the re-classed interval, 'AREA' represents the area of cells for each interval and the rest of the column headings are descriptive statistics for each interval. The maximum point rainfall is found by the highest value in the 'MAX' column.



Rowid	VALUE	COUNT	AREA	MIN	MAX	RANGE	MEAN	STD	SUM
1	1	676	2063440157.739167	0	5	5	1.671598	1.418702	1130
2	2	148	451759087.789048	5.5	10	4.5	7.496622	1.399199	1109.5
3	3	112	341871742.110631	10.5	15	4.5	12.522321	1.257812	1402.5
4	4	62	189250428.668385	13	20	7	17.282258	1.645181	1071.5
5	5	44	134306755.829177	20.5	25	4.5	22.522727	1.360792	991
6	6	21	64100951.645743	25.5	30	4.5	27.595238	1.240383	579.5
7	7	7	21366983.881914	30.5	35	4.5	33	1.732051	231
8	8	7	21366983.881914	35.5	39.5	4	37.285714	1.687287	261
9	9	8	24419410.150759	40.5	45	4.5	42.625	1.595893	341
10	10	12	36629115.226139	45.5	50	4.5	47.208333	1.435536	566.5
11	11	3	9157278.806535	52	53.5	1.5	53	0.707107	159

Figure 4.9: Zonal Statistics as table output

7. The values in Figure 4.9 are exported to Excel, where the data is extracted and the ARFs calculated.

Table 4.1: Calculation of ARFs from Zonal Statistics

Interval (mm)		VALUE ID	COUNT	Area km^2	Rainfall (mm)	ARF	Max Point Rainfall (mm)
0	5	1	676	2063.44	16.57	Ignore	
5	10	2	148	451.76	10.88		
10	15	3	112	341.87	13.73		
15	20	4	62	189.25	10.64		
20	25	5	44	134.31	9.71	**57.27%**	**53.5**
25	30	6	21	64.10	5.66		
30	35	7	7	21.37	2.23		
35	40	8	7	21.37	2.57		
40	45	9	8	24.42	3.33		
45	50	10	12	36.63	5.59		
50	55	11	3	9.16	1.54		
			Total	311.35	30.64		

4.2.2.4 ARF calculation in Excel

For the Excel calculations, the first step was to calculate the areal rainfall. The areal rainfall, in Table 4.1, was obtained following the isohyetal method, where

Equation **7** is valid:

$$R_i = \frac{(I_{i-1}+I_i)}{2} \cdot \frac{A_i}{A_T} \qquad \text{Equation 7}$$

Where:

R_i = Rainfall for current interval

I_i = Current rainfall interval boundary

I_{i-1} = Previous rainfall interval boundary

A_i = Current interval area

A_T = Total area

A threshold of 20 mm for 1-h, 25 mm for 3-h and 40 mm for 24-h storm durations were suggested and implemented by Olivera *et al.* (2006). Thus, 'VALUE' ID 1-4 (0-20 mm) is ignored, as this is a 1-hour storm. This is also convenient as it assists in classing single cellular storms. Figure 4.10 displays the threshold value of the rainfall array in Excel.

The ARF is calculated using Equation 1, with the areal rainfall as 30.64 mm and a maximum point rainfall of 53.5 mm from Table 4.1. Thus, the ARF equals 57.27%.

												25.5	28.5			
												21	24	29	22	20
												21	22	34.5	23	20.5
												24.5	41	45	30	21.5
												43	52	45.5	35.5	
										25.5	41.5	48	50	46.5	38.5	
									26	37	46	47	45	35	28	
								23	40.5	45	46	43	35.5	22.5		
								36	46.5	47	46.5	33	30.5			
								39.5	53.5	46	48	24				
						24	27.5	42	50	46.5	27					
						24	27.5	30	24	21.5						
			20	20	21.5	24.5	22	22								
	23	24	26.5	22.5	25	21	23.5									
22.5	25	24		20	20											
22.5	21															
24																
20																

Figure 4.10: Rainfall array in Excel

5 RESULTS

This chapter presents the results of the calculated ARFs obtained from the analyses of multiple storm events. The storm data used was obtained from SAWS and NWU, as discussed in Chapter 3, Section 3.4. Data used from SAWS consisted of two sets: A set provided in 2018 as a mosaic for South Africa that contained 1, 3 and 24-hour data and the second, provided in 2019, containing only 3-hour data for the Irene area. Only a selected number of storms was obtained and used from the rainfall record originally used for storm dating. SAWS was not able to provide the radar data for all the storms due to the radar being non-operational at specific times. From the SAWS radar data provided, storms were selected based on radar image quality, location (convective or frontal region) and size. The data obtained from NWU consisted of 3 and 24-hour data that was based in the Irene area. This data set ranged from 2014 to 2017, as elaborated on in Chapter 3, Section 3.4.

The influence of storm area, maximum point rainfall intensity, rainfall process and storm duration on ARFs, using radar-based data, are presented. Thereafter, the ArcGIS (radar-based) derived ARFs were compared with the current ARFs used in South Africa. All the results and data used in this chapter is presented in Appendix C.

5.1 Factors Influencing ARFs

The results, using statistical correlation (curve fitting) between various factors that could affect the radar derived ARFs, are presented in this section. The ArcGIS isohyetal areal derived rainfall (Table C.1 (column 2), Appendix C), were used in the analysis.

ARFs produced in South Africa by Van Wyk (1965) and Pullen, Wiederhold and Midgley (1966) considers ARFs to be estimated based on point rainfall, storm area and duration. In Chapter 2, Section 2.3.1, the influences affecting ARFs were discussed. As highlighted before, storm durations of 1-h, 3-h and 24-h were used as a means of categorising the ARFs for each statistical analysis.

5.1.1 ARFs and storm area per duration category

Storm area is rarely used as the only variable to determine an ARF. The literature from Chapter 2, specifically with regards to the ARFs derived in South Africa by Van Wyk (1965) and Pullen, Wiederhold and Midgley (1966), shows however that the storm area is a significant variable to consider when deriving ARFs. The data used for the analysis in this section is presented in Table C.1 (columns 3 and 6).

Figure 5.2 illustrates the relationship between the storm area and radar derived ARFs for **1-hour** storm duration. From the regression analysis, it is clearly evident that a logarithmic trend line provides the optimum fit. Figure 5.2 also shows that for storm areas less than 1 000 km^2, ARFs appear to be more scattered, and could be ascribed to the variability of smaller storms (area) being mainly a function of storm area and point rainfall intensity (Alexander, 2001).

For storm areas greater than 1 000 km^2, the ARFs resulting from larger storm areas tend to follow the trend line. This could possibly result from larger storms being functions of storm area and duration (Alexander, 2001). An R^2 and r - value of 0.38 and -0.62 was obtained, respectively, with a *p*-value approximately close to 0. The R^2 value is indicative of a low degree of association, especially for storm areas less than 1 000 km². The *r* - value and *p*-value indicate weak, but satisfactory inverse correlation and a high significance.

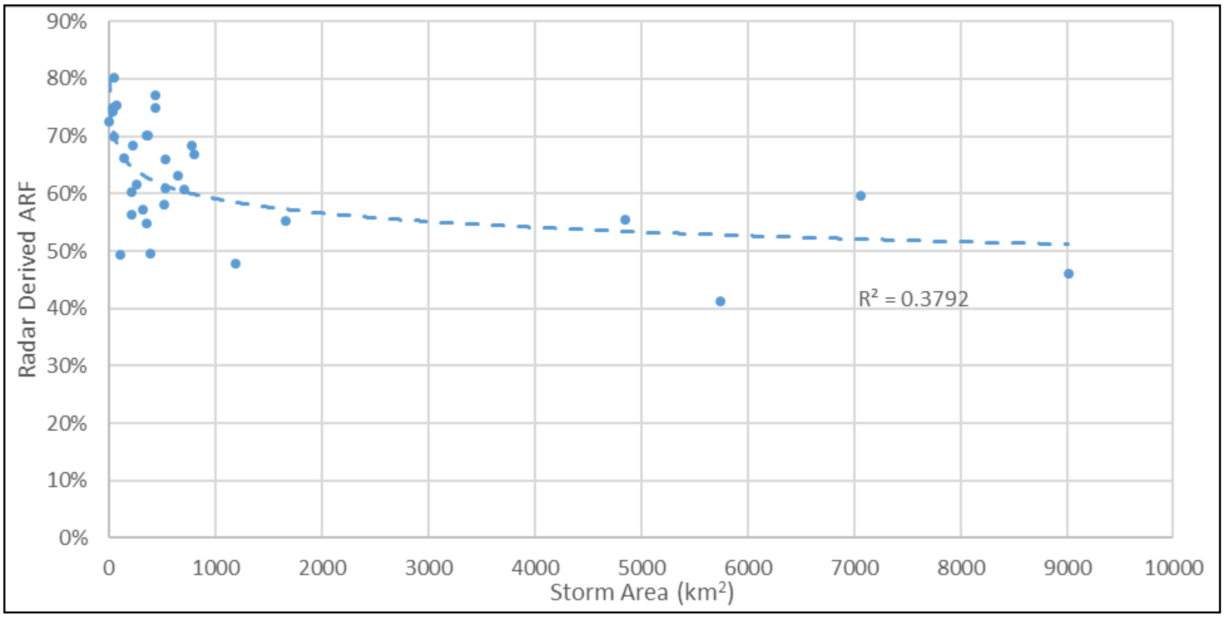


Figure 5.1: Association between ARFs and storm area for the 1-hour storm duration

Figure 5.2 displays the relationship between the storm area and the radar derived ARFs for the **3-hour** storm duration. As in the case of the 1-hour storm duration, logarithmic trend line provides an optimum fit with a higher R^2 value of 0.39. It is evident that the storms with a storm area of less than 1 000 km² produce ARFs that are more scattered, which is in line with the 1-hour storm duration ARFs. From Figure 5.2, there is a tendency for the ARFs to fit the regression line slightly better as the storm area increases. Thus, also attaining to the notion that ARFs for storm areas greater than 1 000 km², are mainly functions of storm area and storm duration (Alexander, 2001). The *r* - value of -0.62 and a *p*-value of approximately zero, highlight a reasonable inverse correlation and high significance.

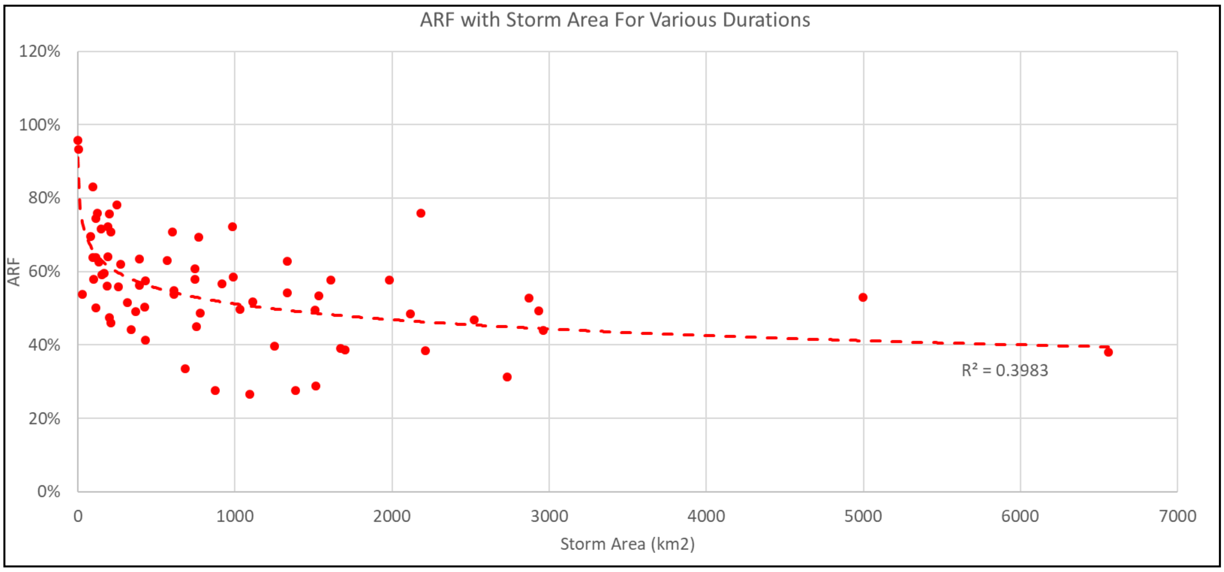


Figure 5.2: Association between ARFs and storm area for the 3-hour storm duration

Figure 5.3 illustrates the relationship between the storm area and the radar derived ARFs for the **24-hour** storm duration. In considering the regression analysis, it is clearly evident that the degree of association (R^2 value) is significantly lower; however, for storms areas less than 1 000 km^2, similar trends as for the 1- and 3-hour storm duration ARFs are evident. Fewer storms exceeding 3000 km^2 were analysed in the case of the 24-hour storm events; however, a better fit is witnessed in this area range.

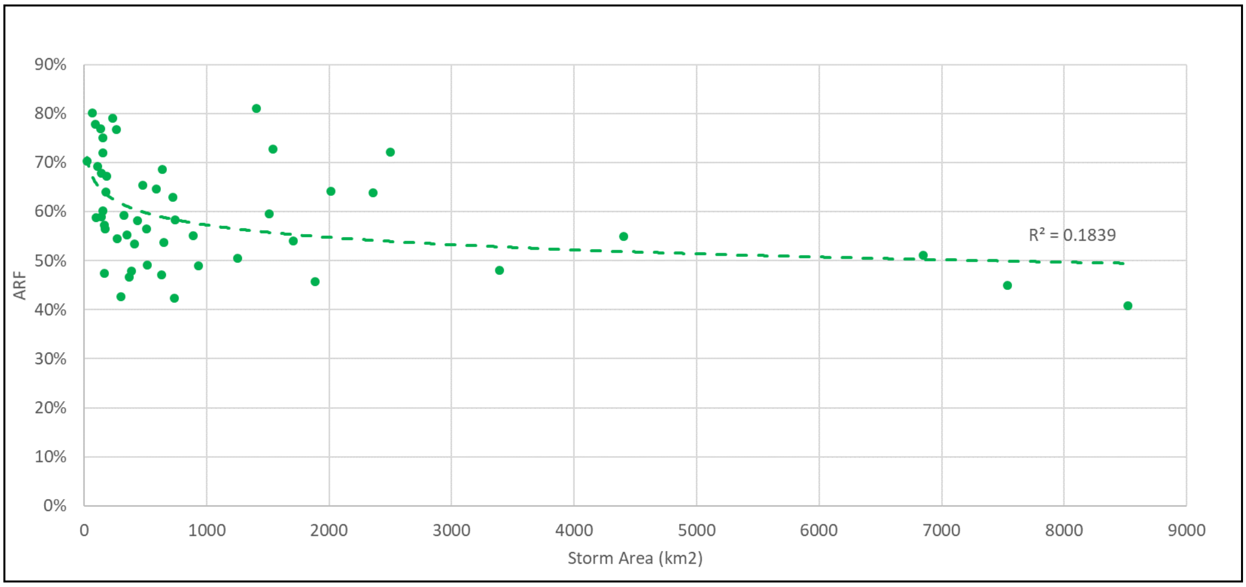


Figure 5.3: Association between ARFs and storm area for the 24-hour storm duration

Figure 5.4 summarises the correlation between ARFs and storm areas for all the storm durations (e.g. 1, 3 and 24-hour) under consideration. The R^2 value of 0.32 confirms the overall low degree of association and also highlights that as the storm area increases, the ARFs decrease. Storm areas less than 2 000 km^2 tend to produce more scattered ARFs.

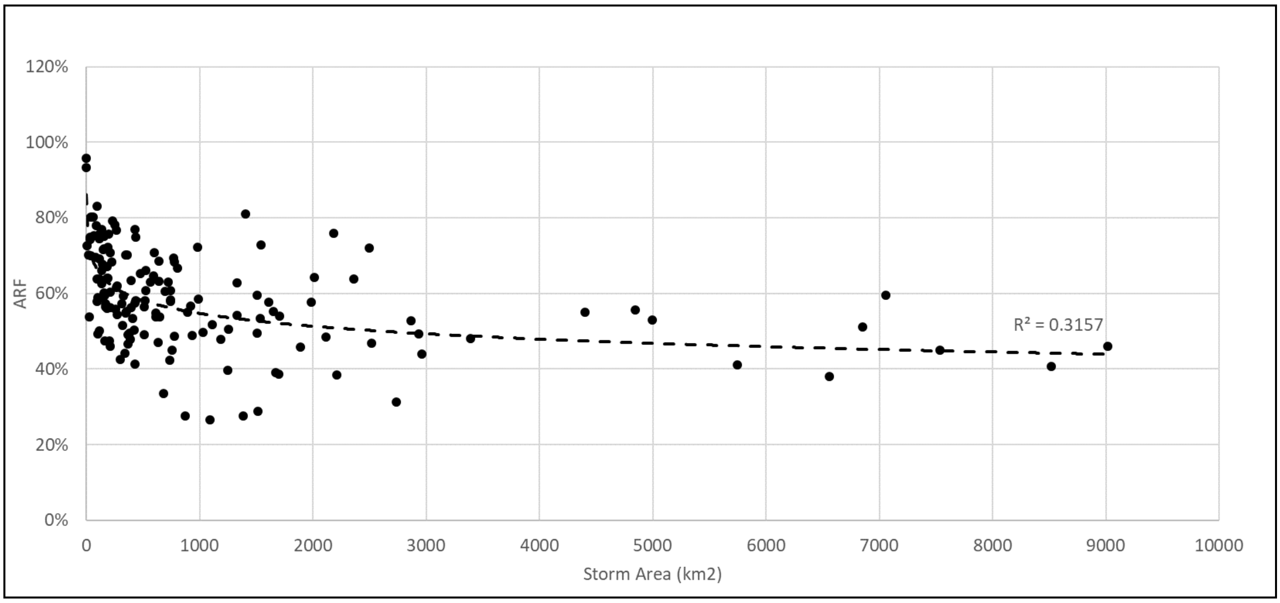


Figure 5.4: Association between ARFs and storm area for all the storm durations

Considering all the statistics obtained from the regression analyses, it was found that storm area does have a significant influence on radar derived ARFs. When compared separately for each storm duration, it was found that the 3-hour storm durations produced the best relationship between storm area and radar derived ARFs. The results are characterised by a weak, but better relationship between storm duration and storm area (assuming a logarithmic relationship), for the 1- and 3-hour storm durations, than for the 24-hour storm durations.

5.1.2 ARFs and maximum point intensity

In this section, similar to Van Wyk (1965), ARFs based on storm area and point rainfall intensity are discussed and presented in Table C.1 (columns 5 and 6).

Figure 5.5 displays the relationship between the maximum point rainfall intensity and ARFs using radar derived data. As in Section 5.1.1, the analysis is presented in the three storm duration categories, e.g. 1, 3 and 24-hour. Overall, the logarithmic transformation proved to be appropriate, with R^2 values ranging between 0.79 (1-hour) and 0.87 (3-hour). Similarly, for the 3-hour and 24-hour storm durations, a strong inverse correlation (r - values ranging between -0.94 and 0.91), is also evident. In considering all the storm durations together (black dotted line in Figure 5.5), i.e. no storm duration categories, a much lower degree of association (R^2 = 0.11; r = -0.33) is evident. Hence, categorisation using storm durations is required to optimise the regression curve fitting and also highlights the impact of storm duration on ARFs.

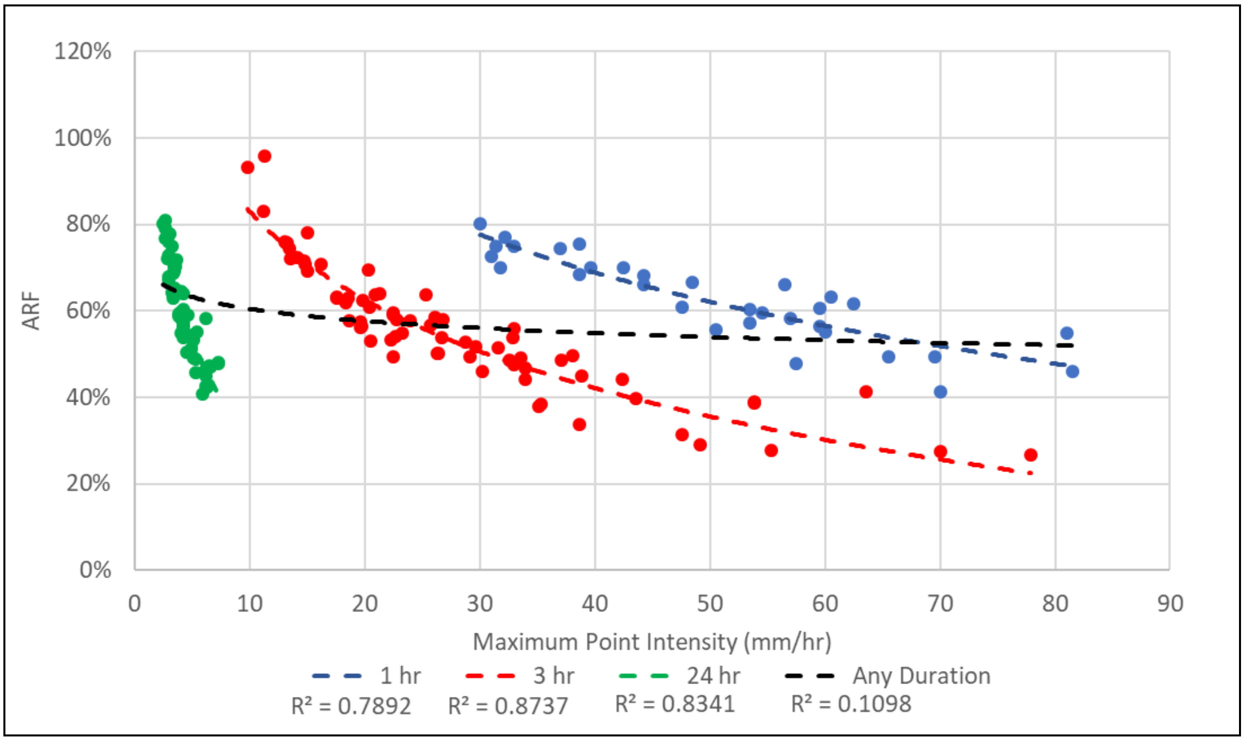


Figure 5.5: Association between ARFs and maximum point rainfall intensity

Figure 5.6 displays the relationship between the maximum point rainfall (and not maximum point rainfall intensity) with the corresponding ARFs. The maximum point rainfalls were not categorised based on storm durations as the relationship between the maximum point rainfalls and ARFs and the relationship between the maximum point rainfall intensities and ARFs are equivalent, for the same storm durations.

The following relationship between the ARF and maximum point rainfall for all storm duration was observed: As the maximum point rainfall increases logarithmically, the ARF decreases linearly. When considering the relationship between maximum point rainfall and ARFs, disregarding the influence of storm duration, a R^2 value of 0.63, r - value of -0.79 and *p*-value of approximately zero were obtained. This shows that when disregarding the influence of storm duration, an improved, although still low correlation exists between maximum point rainfall and ARFs.

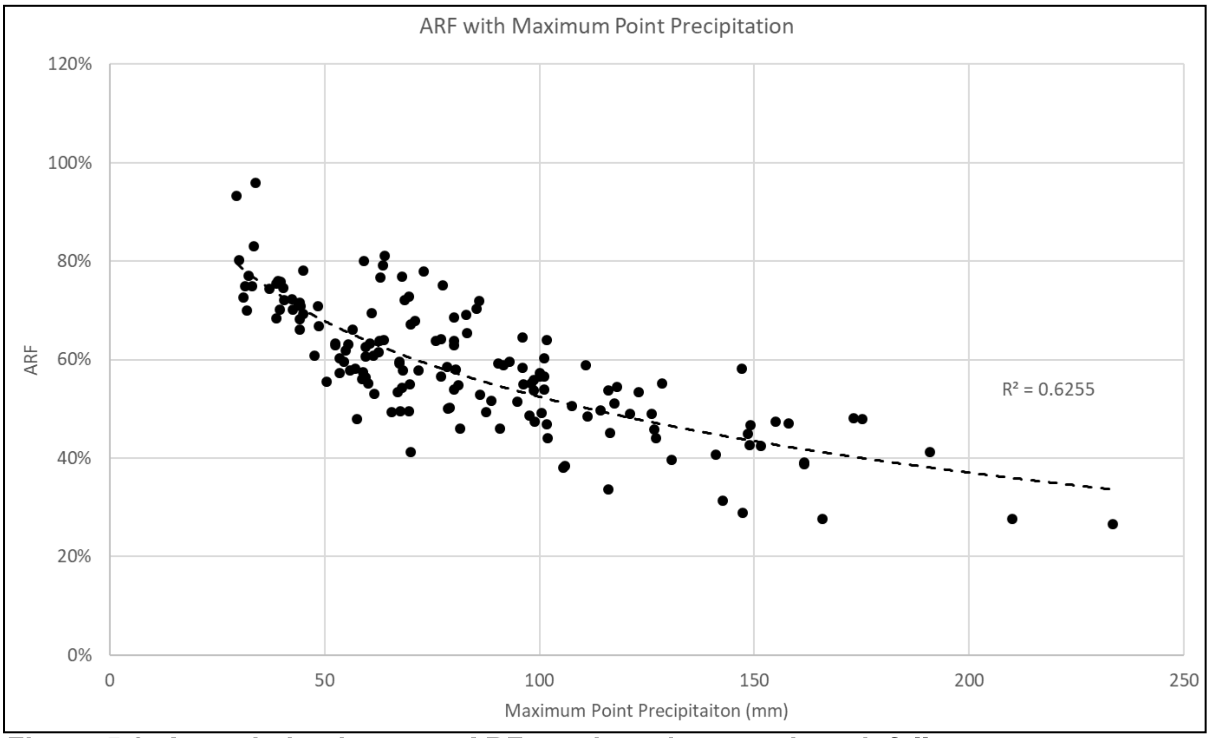


Figure 5.6: Association between ARFs and maximum point rainfall

5.1.3 ARFs and different rainfall processes

The rainfall processes considered in this project are representative of convective and frontal rainfall. Convective regions were defined to be located in the vicinity of Pretoria, and frontal regions were defined to occur along the coast. In Chapter 2, Section 2.3.1.3, the various rainfall processes were described with their influences on the derivation of ARFs. It was stated that in convective regions, ARFs may be lower due to thunderstorms with rapid changing rainfall intensities. On the other hand, frontal rainfall regions may produce higher ARFs as these ARFs are considered to be more representative of the actual rainfall process.

The results displayed in Figure 5.7, highlight that convective rainfall regions produced slightly higher ARFs than those in frontal rainfall regions. It is important to note that the number of convective storms analysed is not the same as the number of frontal storms analysed due to data availability. For the 1-h storm duration, (12 frontal and 20 convective), the 3-h storm duration (4 frontal and 67 convective), and the 24-h duration (13 frontal and 39 convective) storms were analysed. Thus, the results obtained does not truly contradict the notion that ARFs for frontal regions are lower than ARFs for convective regions, as many more factors need to be evaluated.

However, this conforms to the limitations presented in Chapter 1, Section 1.4, which do not accommodate various controlled factors. These controlled factors could include constant storm areas, duration and maximum point intensity.

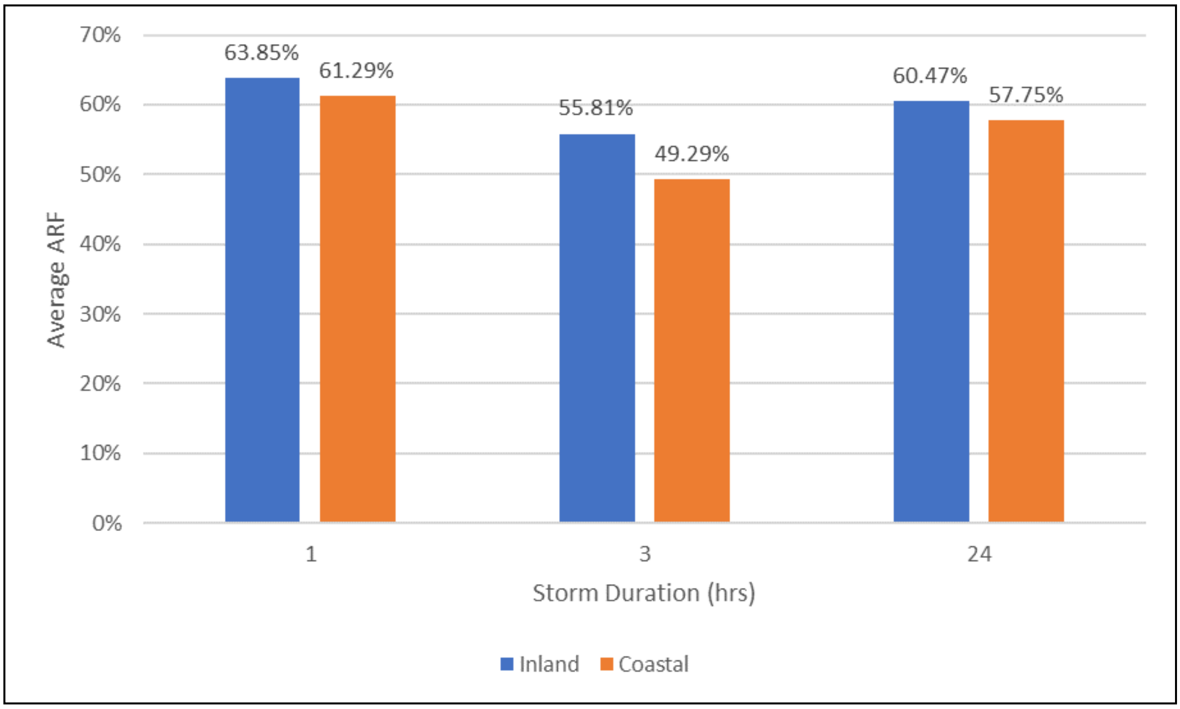


Figure 5.7: ARFs for convective and frontal rainfall regions

5.1.4 ARFs and storm duration

ARFs were analysed for each storm duration, which was bound by the limitations mentioned in Chapter 1, Section 1.4. Figure 5.8 displays the ARFs obtained for all storm durations with the respective average ARFs. The average of the ARFs decrease from the 1-hour to the 3-hour storm durations. This is expected as ARFs typically decrease with an increase in storm duration (Kim *et al.,* 2019). Contrary to this observation, a slight increase in terms of the average ARFs from the 3-hour to the 24-hour storm duration was observed.

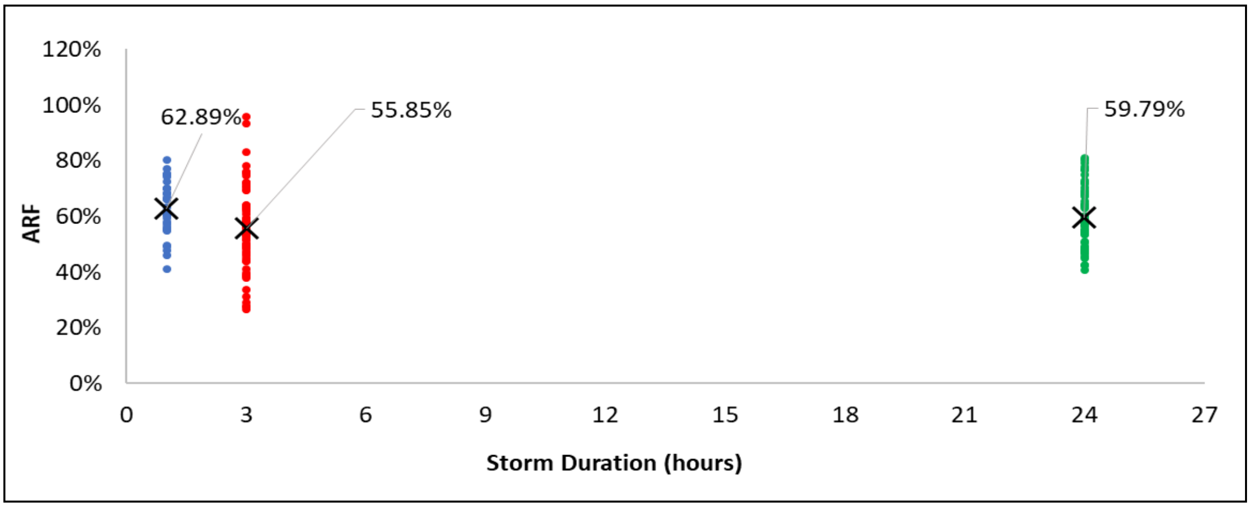


Figure 5.8: ARFs for various storm durations

In acknowledging that storm area also plays an important role, a further evaluation of the available information is presented in Table 5.1. Table 5.1 displays the average of the storm areas, radar derived ARFs and standard deviation of the ARFs for each storm duration. The average storm area for the 1-hour storm events are the highest and is not in agreement with the conclusion made in Section 5.1.1, as well as shown in Figures 2.2. and 2.3, respectively. In other words, the largest storm area did not yield the lowest ARFs. The 3-hour storm events actually host the smallest storm area with the lowest average ARF.

Table 5.1: Storm averages

Storm duration (hours)	Average storm area (km^2)	Average ARF	Standard deviation of ARFs	Number of storms with storm areas < 1000 km^2	Number of storms with storm areas > 1000 km^2
1	1186.26	62.89%	9.94%	26	6
3	979.14	55.85%	14.57%	47	24
24	1162.12	59.79%	10.96%	38	14
Any Duration	**1083.29**	**58.63%**	**12.81%**	**111**	**44**

The standard deviation for the 1-hour storms is the lowest, indicating that the data is the least scattered. The 24-hour storm events had a lower standard deviation than the 3-hour storm events, indicating that the ARFs from the 24-hour storm events are more scattered than that of the 3-hour storm events. The increase in the average ARFs from 3 hours to 24-hour storm events could be explained by the number of storms with corresponding ARFs of storm areas less than 1 000 km^2.

It can be speculated that, since more storms with storm areas exceeding a 1 000 km^2 were produced by the 3-hour storm events (24 storms), than the number of 24-hour storm events (14), a lower average ARF was produced for the 3-h storm duration.

However, since these findings are not very convincing, further analyses were required. It was found necessary to group the ARFs for each storm duration, using storm areas of 200, 500, 1 000 and 10 000 km^2, as shown in Figure 5.9. Figure 5.9 must be viewed in conjunction with the results listed in Table 5.2. From the literature (Chapter 2, Figure 2.3), an increase in ARFs are associated with an increase in storm duration and decrease in storm area, respectively. Hence, the predicted results should therefore portray an increase in ARFs for an increase in storm duration on each graph ((a) to (d)) in Figure 5.9, as these graphs already represent the different storm area categories.

Figure 5.9 (a) shows a decline in the average ARF for storm areas less than 200 km^2 associated with 1-hour to 24-hour storm events, respectively as expected. In Figure 5.9 (b), (c) and (d), a decrease in the average ARFs from the 1-hour to 3-hour storm events is witnessed (as expected), while an increase in average ARFs from the 3 to 24-hour storm events are observed.

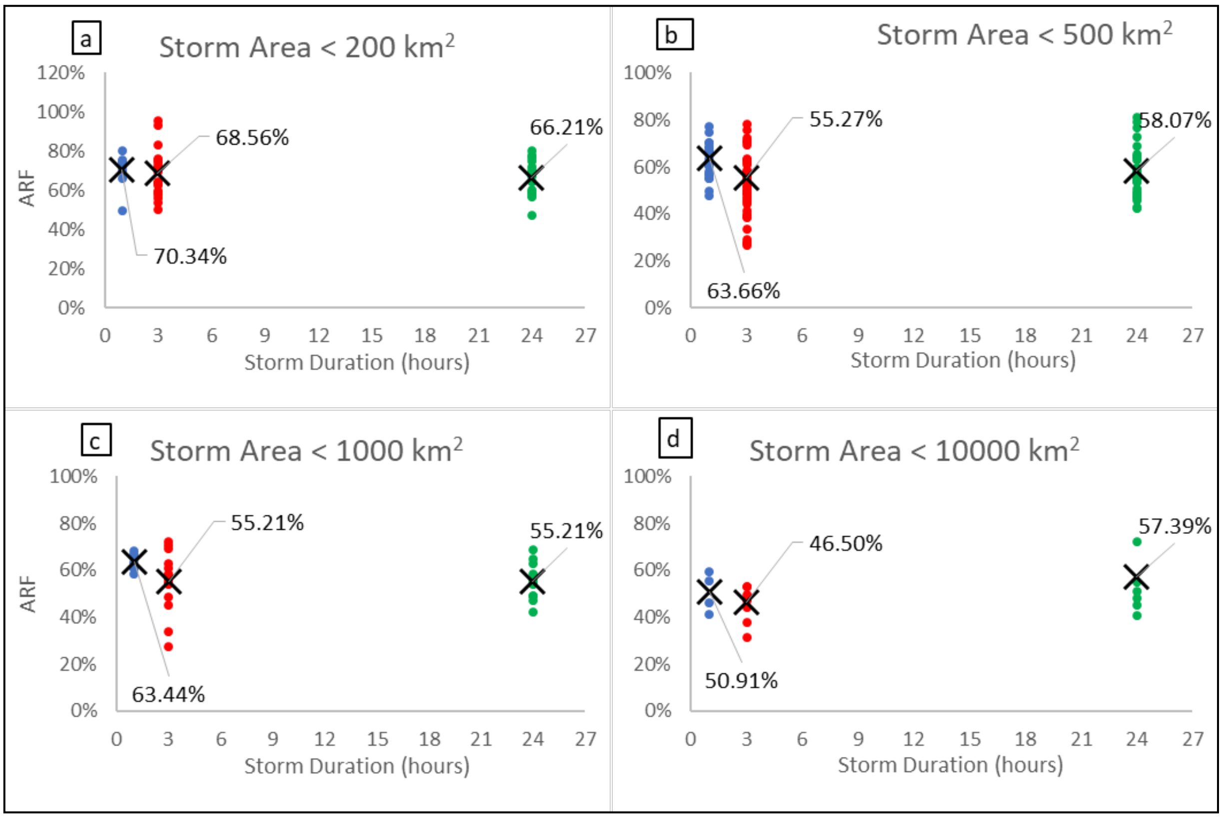


Figure 5.9: Average ARFs per storm duration in different storm area categories

The latter deviation from the norm, i.e. ARFs increase as the storm duration increases and the storm area decreases, could partially be ascribed to the limited number of storms considered and listed in Table 5.2.

Table 5.2: Average storm area for each storm duration

Storm duration (hours)	< 200 km^2	< 500 km^2	< 1 000 km^2	< 10 000 km^2
1	56.47	320.27	640.14	4917.46
3	117.76	321.07	759.58	2173.03
24	133.90	347.84	688.31	3349.28

5.2 Comparison of ARFs

Van Wyk (1965) suggested that the ARFs produced from his study be validated using radar technology. This section compares the radar-based ARFs that were derived from ArcGIS with ARFs that are currently implemented in South Africa, i.e. ARFs derived by Van Wyk (1965) and Pullen, Wiederhold and Midgley (1966) using a storm-centred approach, and Op ten Noort and Stephenson (1982).

The storm-centred ARFs currently implemented in South Africa are separated into two categories based on the storm criteria: catchments less than 800 km^2; and catchments greater than 800 km^2 but less than 30 000 km^2. Thus, for the comparison of ARFs, the radar-based ARFs were also separated into the two categories using the same criteria.

5.2.1 Current ARFs for small catchments

The radar derived ARFs were compared to Van Wyk's (1965) ARFs for catchments less than 800 km^2, following Section 2.3.3, Chapter 2. The data set for comparison with Van Wyk (1965) was obtained using the radar derived maximum point intensity and storm area to arrive at each ARF, for each storm. These ARFs were also compared to the Op ten Noort and Stephenson's (1982) (OTN) Equation 2. A graphical illustration of the results is presented in Figure 5.10. All the data used in this section is contained in Table C.2 in Appendix C.

Table 5.3: Comparison between radar-based ARFs and the current storm-centred approaches used in South Africa for areas less than 800 km^2

Method	Storm area range	Difference Equation	Minimum difference	Maximum difference	Average difference	No of storms with difference < 0%
Van Wyk	< 800 km^2	(Van Wyk ARF) - (Radar Derived ARF)	-2.55%	54.43%	26.90%	3
OTN Equation 2	< 800 km^2	(OTN ARF) - (Radar Derived ARF)	-56.13%	45.65%	6.22%	34

A regression analysis was carried out between the Van Wyk (1965) ARFs and the radar derived ARFs. A R^2 value of 0.11 and r - value of 0.33 was respectively obtained; hence, confirming the low degree of association between the two methods. Upon further analyses, it was found that Van Wyk's (1965) storm-centred derived ARFs were, on average, 26.90% higher than the radar derived ARFs, as listed in Table 5.3. This was expected as Van Wyk's (1965) method of deriving storm-centred ARFs included rainfall stations' data, which could, due to the latter stations poor spatial distribution, produce less representative ARFs (Sinclair and Pegram, 2005).

The graph presented in Figure 5.10 contains a reference line called "Perfect Correlation". The "Perfect Correlation"- line can be used to graphically illustrates the deviation of the comparison between the radar derived ARFs with the ARFs currently in use in South Africa. In other words, the "perfect correlation"- line would exist if all the radar derived ARFs were the same as the Van Wyk's (1965) ARFs, using storm area and maximum point intensity as input parameters.

Regarding Op ten Noort and Stephenson's (1982) Equation 2, it was found that using the equation to calculate ARFs, produced ARFs that were, on average, only 6.22% higher than the radar derived ARFs. Figure 5.10 shows that the trend line of the Op ten Noort and Stephenson's (1982) ARFs almost identically matches the "Perfect Correlation"- line. Although, the trend lines are almost identical it still represents a poor correlation, i.e. R^2 value of 0.14 and r - value of 0.37. In other words, both data sets contain trend lines that fits their respective data sets similarly; however, there still exists a poor correlation between the two data sets.

Comparing the average differences in ARFs from Van Wyk (1965) and Op ten Noort and Stephenson's (1982), it is clear that using Op ten Noort and Stephenson's (1982) Equation 2, produces ARFs closer to the radar derived ARFs. Figure 5.10 also shows the deviation of the results from the perfect correlation with Van Wyk's (1965) ARFs.

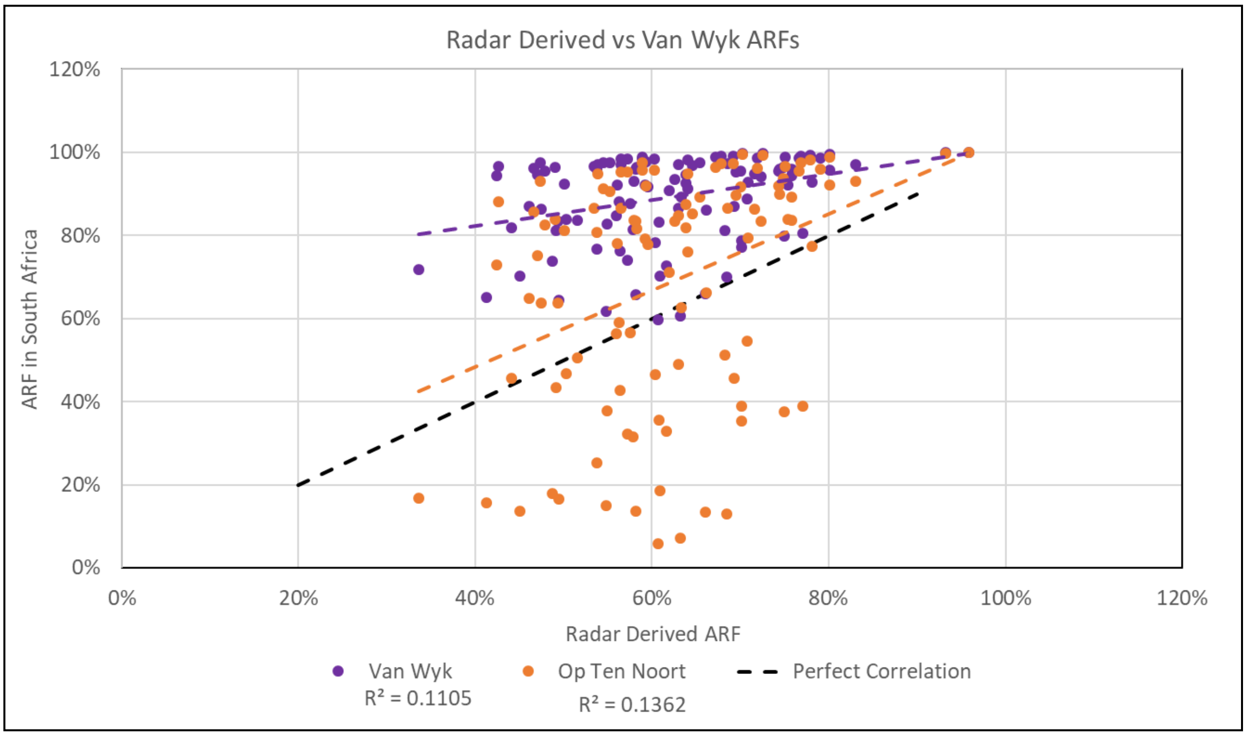


Figure 5.10: Statistical Comparison of ARFs for storm areas ≤ 800 km^2

However, Op ten Noort and Stephenson's (1982) stated that their derived ARFs compare very well with those produced (Figure 5.11) by Van Wyk (1965). By investigating the latter statement, R^2 values > 0.9 were obtained.

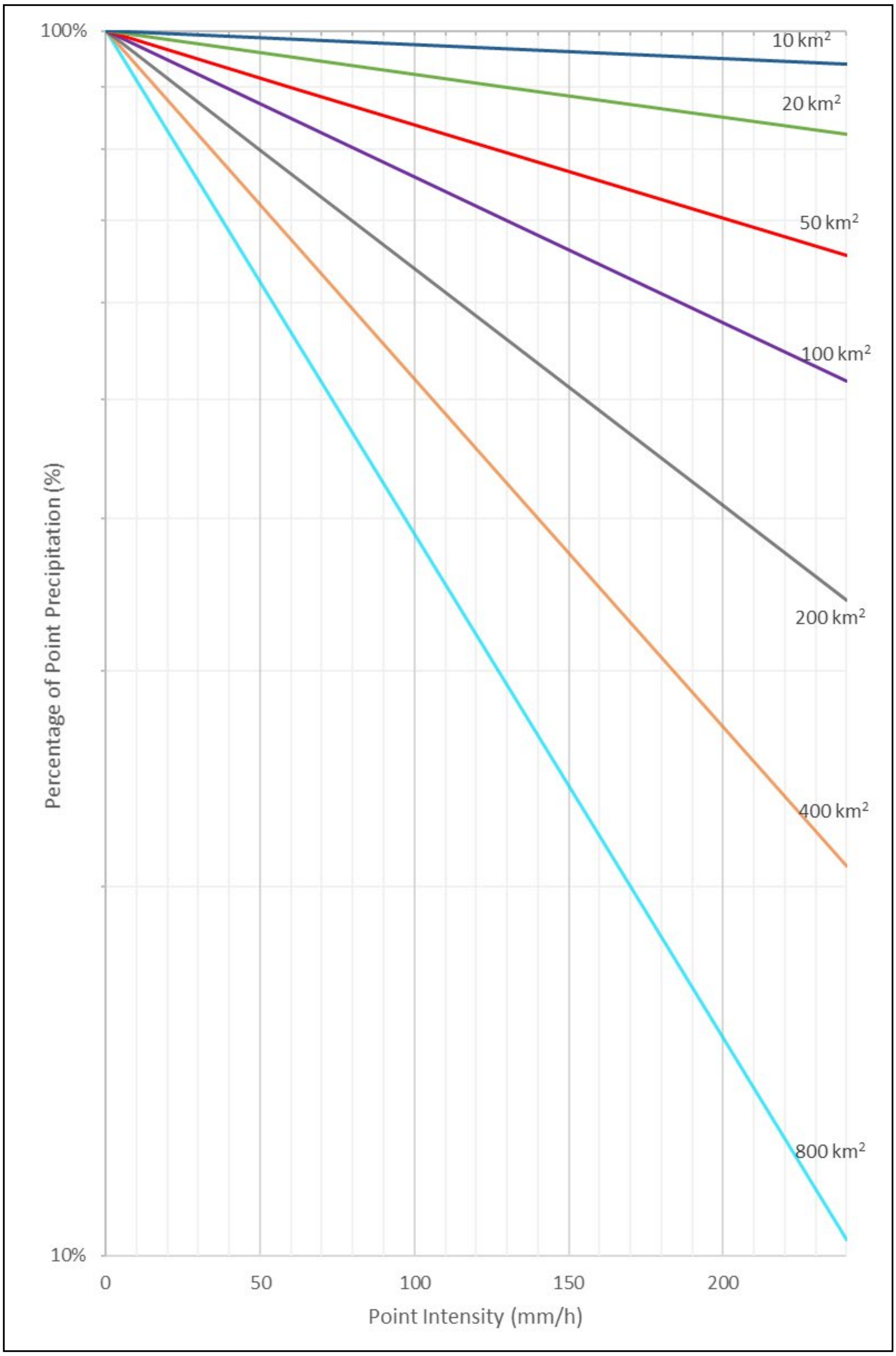


Figure 5.11: Adapted from Van Wyk's (1965) ARFs in South Africa

Figure 5.12 illustrates the comparison between the radar derived ARFs (or reduction in storm point rainfall), and ARFs currently used in South Africa (Van Wyk's graph). The radar data available only reflect a maximum point rainfall intensity of 81 mm/h. Thus, Figure 5.12 only compares the radar derived ARFs with ARFs in South Africa up to 81 mm/h. These can be considered as a "low point intensity storms", for comparison purposes. The storms were grouped, for comparison purposes, into the following seven storm area boundaries: 0-15, 15 -35, 35-75, 75-150, 150-300, 300-600 and 600-800 km^2, which matched the storm areas used by Van Wyk (1965): 10, 20, 50, 100, 200, 400 and 800 km^2 in Figure 5.11.

Trend lines were then established for each data set. Only three storm events in the research data set conformed to the 0-15 km^2 boundary. The ARFs for those storms were found to be generally in agreement with Van Wyk's (1965) ARFs. The radar-based storm events in the storm categories of 20, 50 and 100 km^2, in Figure 5.12 (b), (c) and (d), were found to be all lower than the Van Wyk's ARFs, with similar gradients (increase in rainfall intensity is associated with a decrease in ARFs).

For the radar derived ARFs in the 20 km^2 category (Figure 5.12 (b)), the ARFs are all lower than Van Wyk's, with an insignificant tendency for an increase in ARFs as the rainfall intensity increases. This is inconclusive based on the amount of data points for the 20 km^2 category. For radar derived ARFs in the 200 km^2 category, Figure 5.12 (e), lower ARFs than suggested by Van Wyk, are evident. However, as the rainfall intensity increases, the differences between the ARFs decreased slightly. A similar trend was evident for the radar derived ARFs in the 400 and 800 km^2 categories (Figure 5.12 (f) and (g)). In these cases, lower radar derived ARFs were obtained when compared with the Van Wyk's ARFs up until a rainfall intensity of 60 and 80 mm/h, respectively. For rainfall intensities of 60 mm/h or more, the 800 km^2 category produced higher radar-based ARFs than those suggested by Van Wyk.

From all the catchment area categories presented in Figure 5.12, except for the 800 km^2 category, all radar derived ARFs were lower than those suggested by Van Wyk, with some trend lines highlighting that a threshold value do exist where, for higher rainfall intensities, the radar-based ARFs will exceed those suggested by Van Wyk.

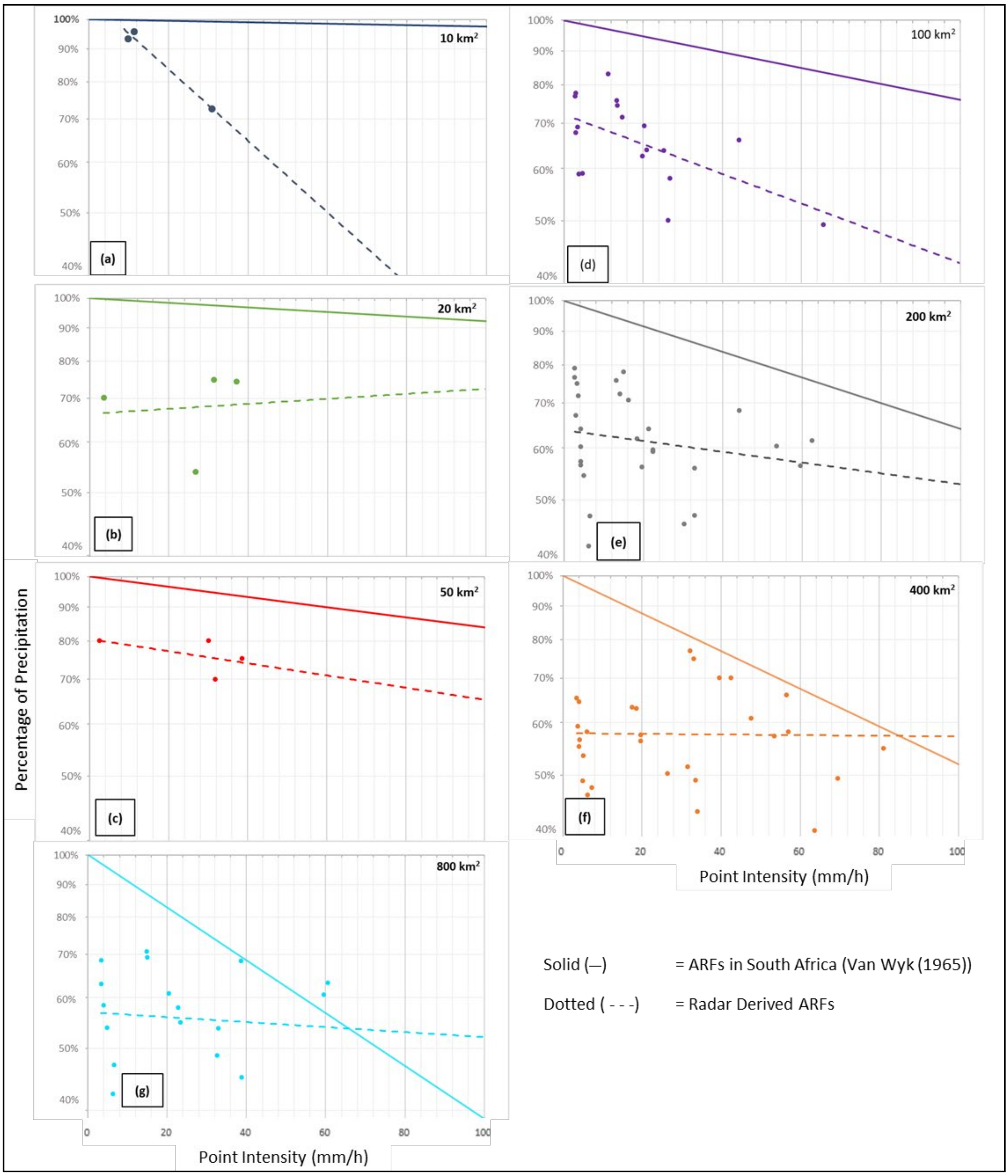


Figure 5.12: Comparison of ARFs in different storm area categories ≤ 800 km²

5.2.2 Current ARFs for larger catchments

The radar derived ARFs were also compared to Pullen, Wiederhold and Midgley's (1966) ARFs for catchments larger than 800 km^2, but less than 30 000 km^2. The ARFs as derived using the latter methodology, were obtained using the radar derived storm area and durations. Similarly, ARFs were also calculated using the Op ten Noort and Stephenson's (1982) Equation 3 and the same radar derived input parameters, A graphical illustration of the results is presented in Figure 5.13. The data used in this section are also contained in Appendix C (Table C.3).

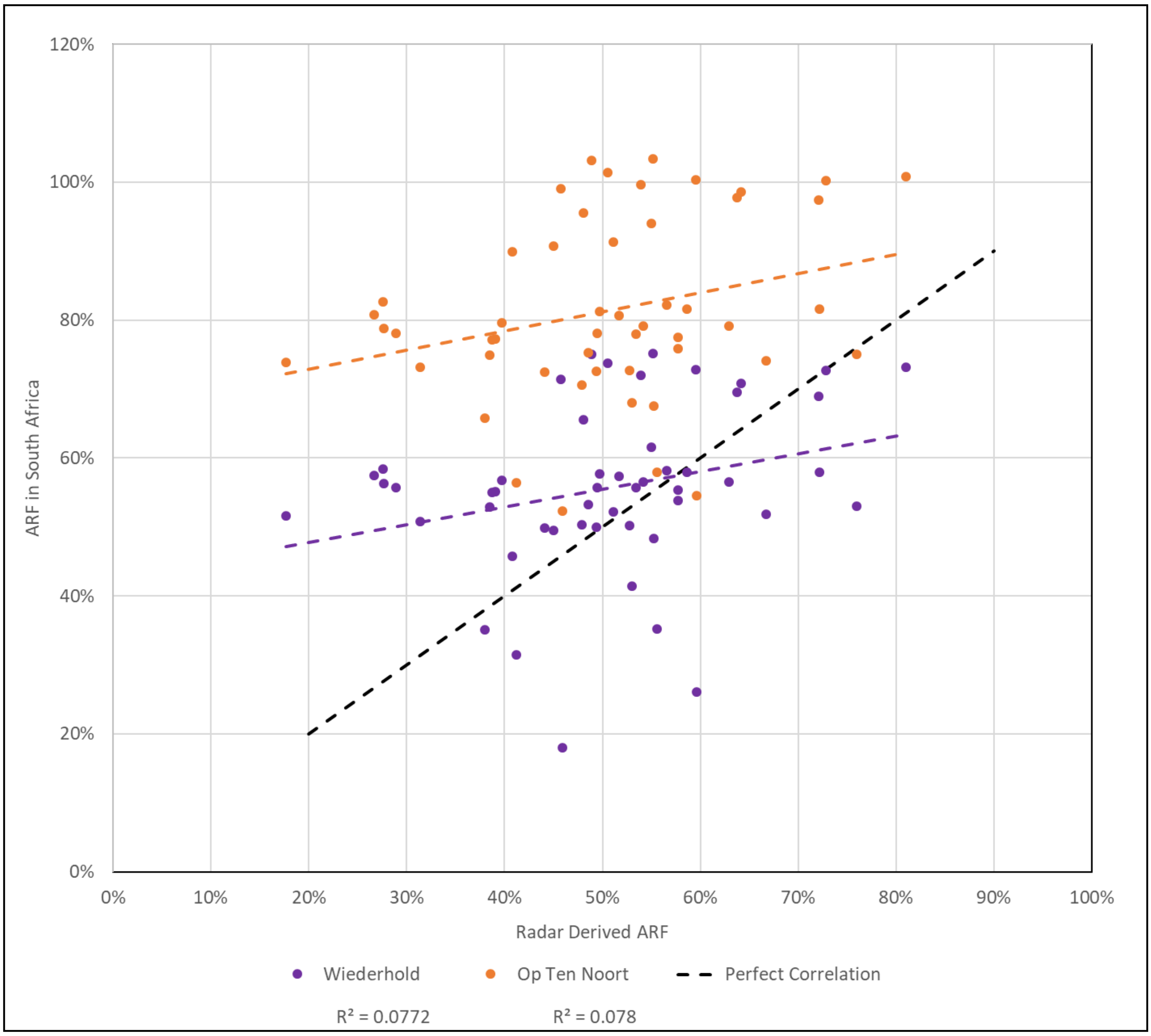


Figure 5.13: Statistical ARF comparison for larger areas

By considering the regression analysis between these methods, it is clearly evident that almost no correlation exists, i.e. R^2 = 0.08 and r = 0.28. Although, a slightly higher correlation exists between the radar-based ARFs and Equation 3. Upon further analysis, Pullen, Wiederhold and Midgley's (1966) storm-centred derived ARFs were, on average, only 5.01% higher than the radar derived ARFs. The latter results are presented in Table 5.4.

Table 5.4: Comparison between radar-based ARFs and the current storm-centred approaches used in South Africa for areas exceeding 800 km²

Comparison	Storm area range		Minimum difference	Maximum difference	Average difference	No of storms with difference < 0%
Wiederhold	> 800 km²		-33.55%	33.94%	5.01%	18
OTN Equation 3	> 800 km²		-5.05%	56.15%	30.77%	2

Figure 5.13 also contains a reference line called "Perfect Correlation" as explained in section 5.2.1. The "perfect correlation" - line would exist if all the radar derived ARFs were the same as ARFs from Pullen, Wiederhold and Midgley's (1966), using radar derived storm area and duration as input parameters.

The results indicate that using Op ten Noort and Stephenson's (1982) Equation 3 to estimate ARFs, produced values that were on average 30.8% higher than the radar derived ARFs. This is displayed in Table 5.4. Comparing the average difference between radar derived ARFs and ARFs from Pullen, Wiederhold and Midgley (1966) and Op ten Noort and Stephenson's (1982) as presented in Table 5.4, it is clear that the radar derived ARFs provided results closer to those suggested by Pullen, Wiederhold and Midgley's (1966) method in comparison to those suggested by Op ten Noort and Stephenson (1982).

A regression analysis was carried out on the ARFs produced by Op ten Noort and Stephenson's (1982) Equation 3 and Pullen, Wiederhold and Midgley's (1966), using the radar derived storm inputs parameters, presented in Appendix C (Table C.3 (columns 5 and 6)). This was carried out in order to validate the high correlation present between the two existing ARF data sets used in South Africa, as mentioned by Op ten Noort and Stephenson's (1982). A R^2 value of more than 0.85 was obtained, indicating and confirming that there exists indeed a high correlation between the two ARF data sets, as expected.

Figure 5.14 displays the comparison of the radar derived ARFs with the ARFs suggested (and presently in use in South Africa) by Pullen, Wiederhold and Midgley's (1966). The radar derived ARFs were categorised based on the storm durations 1, 3 and 24 hours. Trend lines were then produced for each storm duration. Figure 5.14 illustrates that for 1-hour radar-based storm events, all the ARFs are higher than those suggested by Pullen, Wiederhold and Midgley for all storm areas. The radar-based ARFs for the 3- and 24-hour storm events are all less than the ARFs suggested by Pullen, Wiederhold and Midgley for storm areas up to 7 500 km². The results therefore suggest that radar derived ARFs overestimate ARFs for storm durations less than 1-hour, and underestimates ARFs for storms exceeding 1 hour, up until a certain storm area threshold.

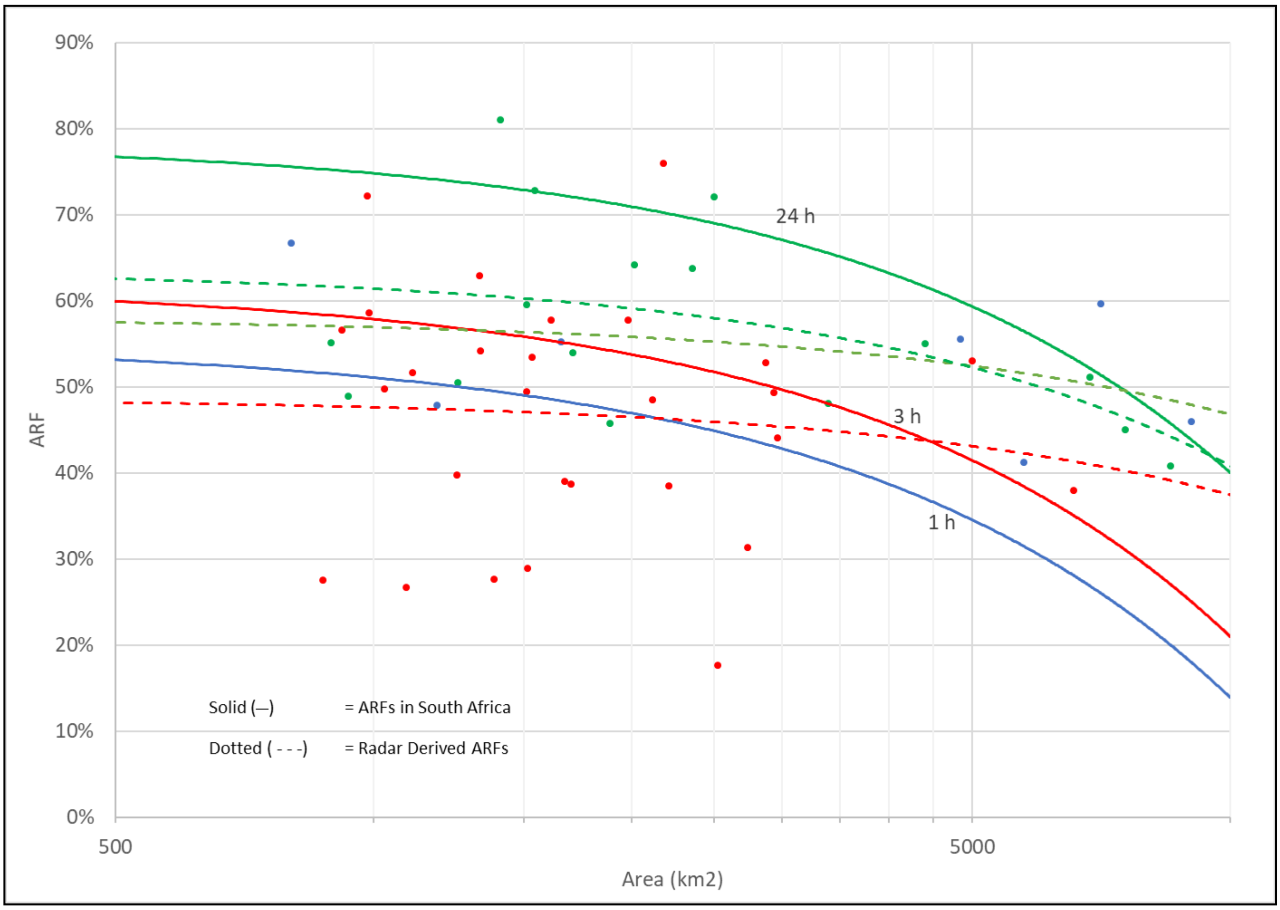


Figure 5.14: Comparison of ARFs for larger catchments (800 km^2 < A ≤ 30 000 km^2)

5.3 Data Provider Comparison

In Chapter 3, Section 3.4, the sources of data used were discussed. Radar data was provided from SAWS and the NWU. While it was previously stated that different processing procedures might result in different results, a comparison between the results based on the different data sources, were executed. The maximum point rainfall intensity, storm area, maximum point rainfall depth and storm duration were used as criteria for the comparison between the data sets. In all cases the ARFs were calculated using the two data sets separately and the values compared for those storms where the different criteria listed above, were the same.

5.3.1 Maximum point rainfall intensity

Figure 5.15 displays the comparison between the SAWS and NWU 3-hour data sets using the maximum point rainfall intensity as criteria. The high R^2 values are indicative of a favourable correlation with similar trends. Similar results are evident for the 24-hour data set in Figure 5.16.

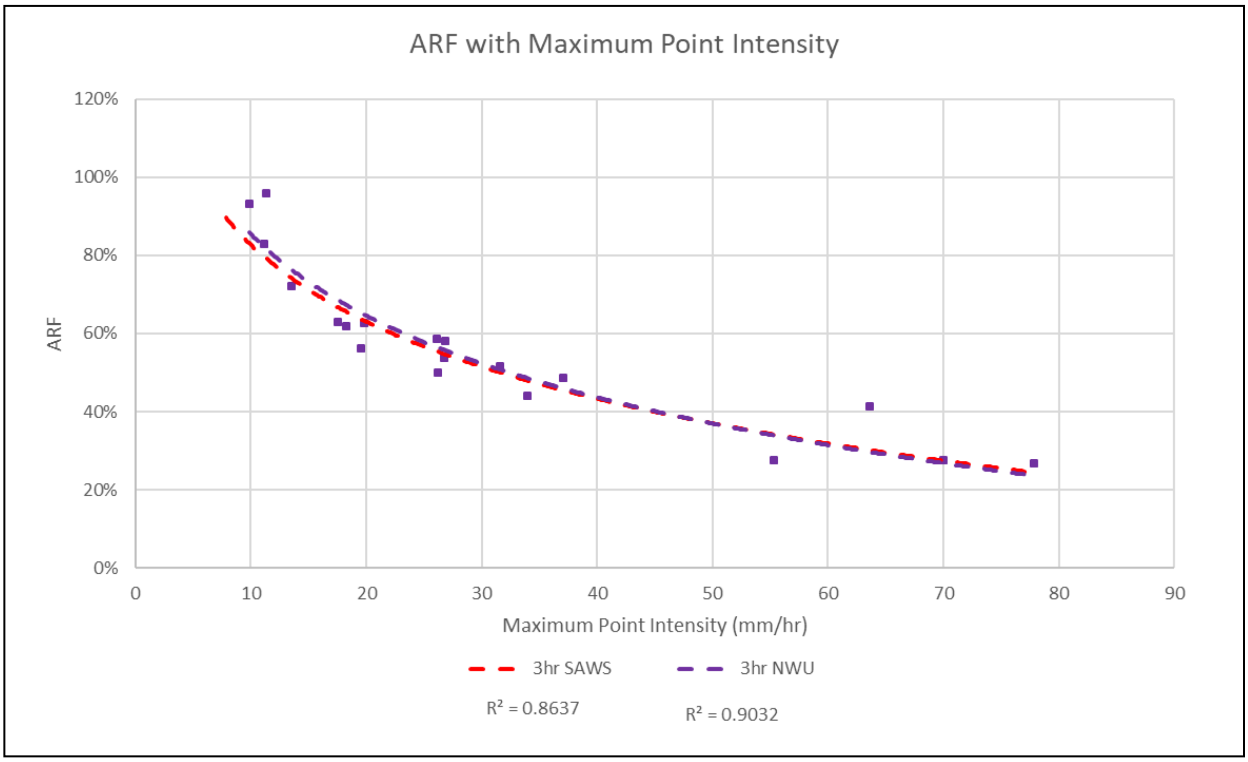


Figure 5.15: Comparison of the 3-h SAWS and NWU maximum point intensity data sets

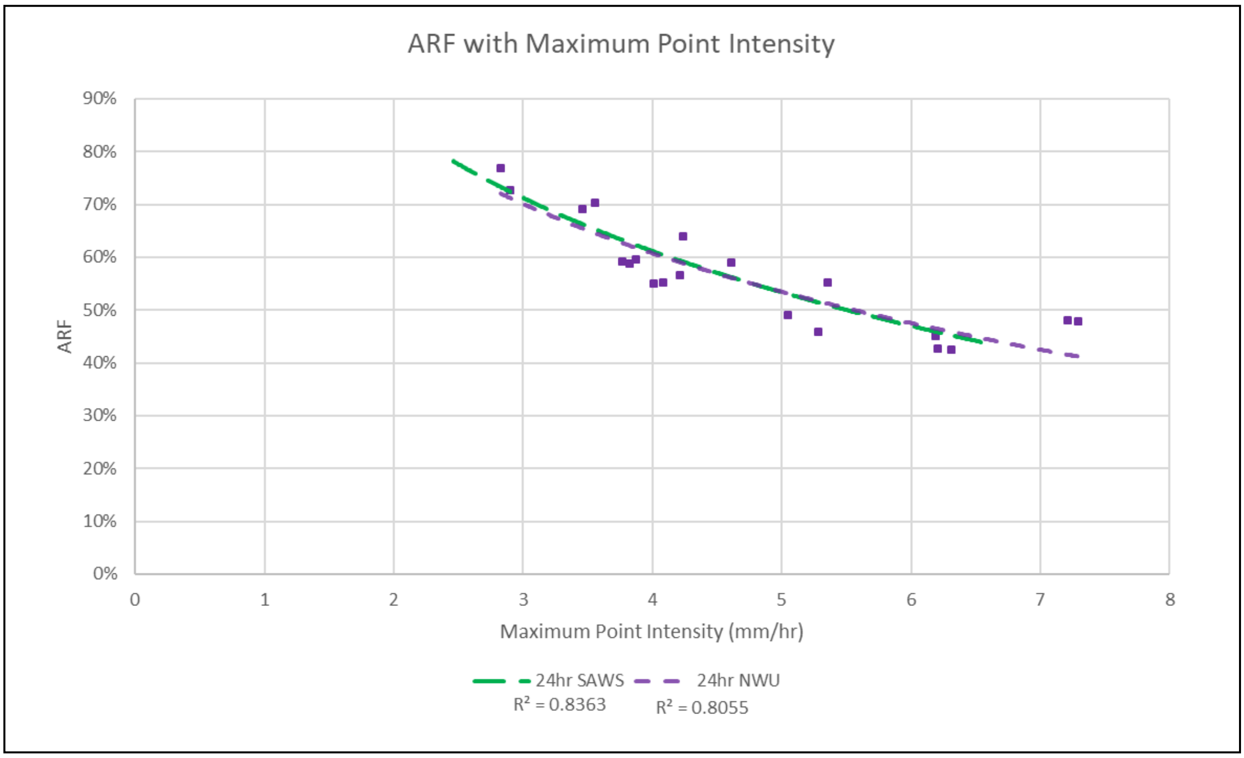


Figure 5.16: Comparison of the 24-h SAWS and NWU maximum point intensity data sets

It is clear from both Figure 5.15 and Figure 5.16, that for storms with a similar maximum point intensity, the two data sets provided acceptable results.

5.3.2 Storm area

Figure 5.15 displays the comparison between the SAWS and NWU 3-hour data sets using storm area as criteria. The low R^2 values are indicative of a low degree of association, although both data sets follow a logarithmic trend. Similar results are evident for the 24-hour data set in Figure 5.18.

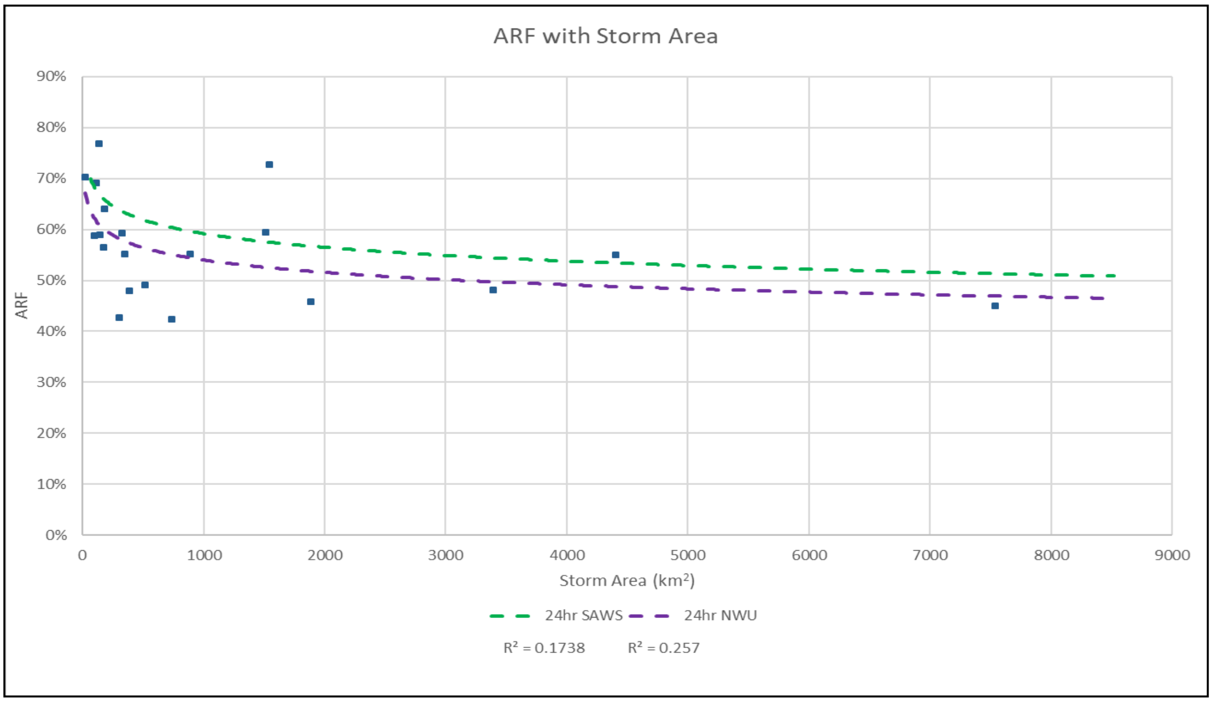


Figure 5.17: Comparison of the 3-h SAWS and NWU storm area data sets

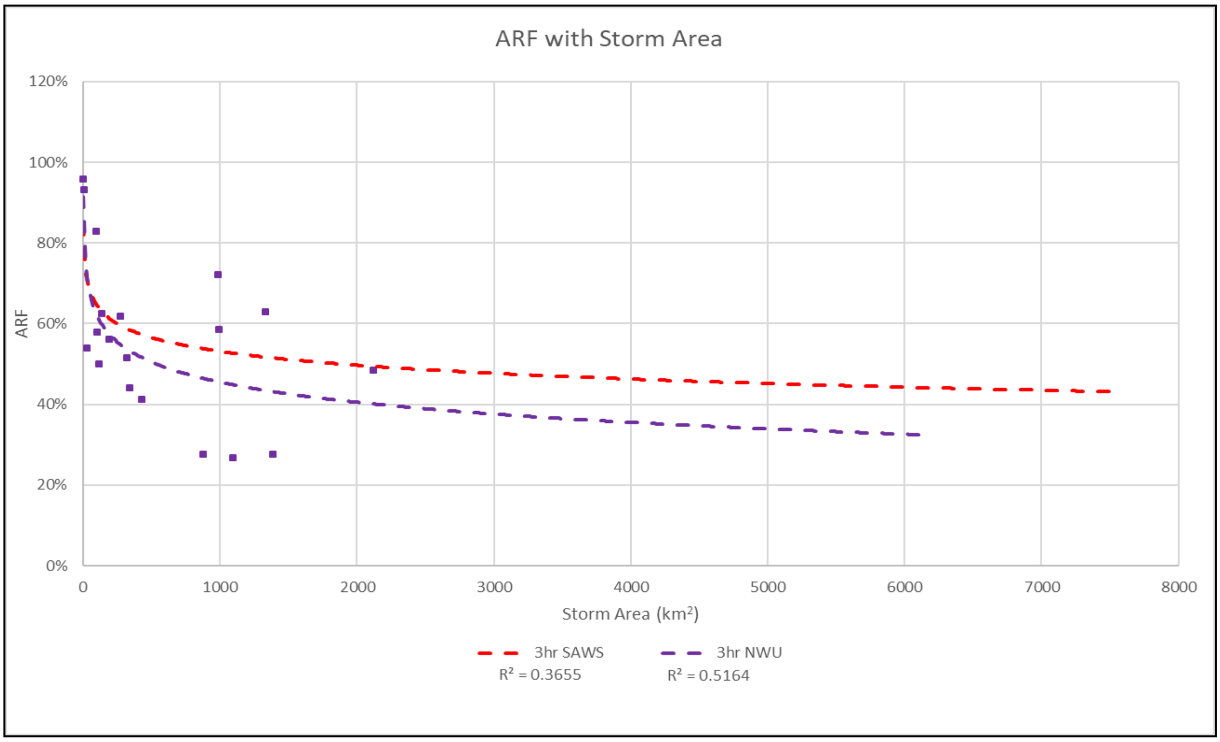


Figure 5.18: Comparison of the 24-h SAWS and NWU storm area data sets

5.3.3 Maximum point rainfall depth

Figure 5.15 displays the comparison between the SAWS and NWU data sets for all durations using the maximum point rainfall depth as criteria. The high R^2 values are indicative of a favorable correlation with similar trends.

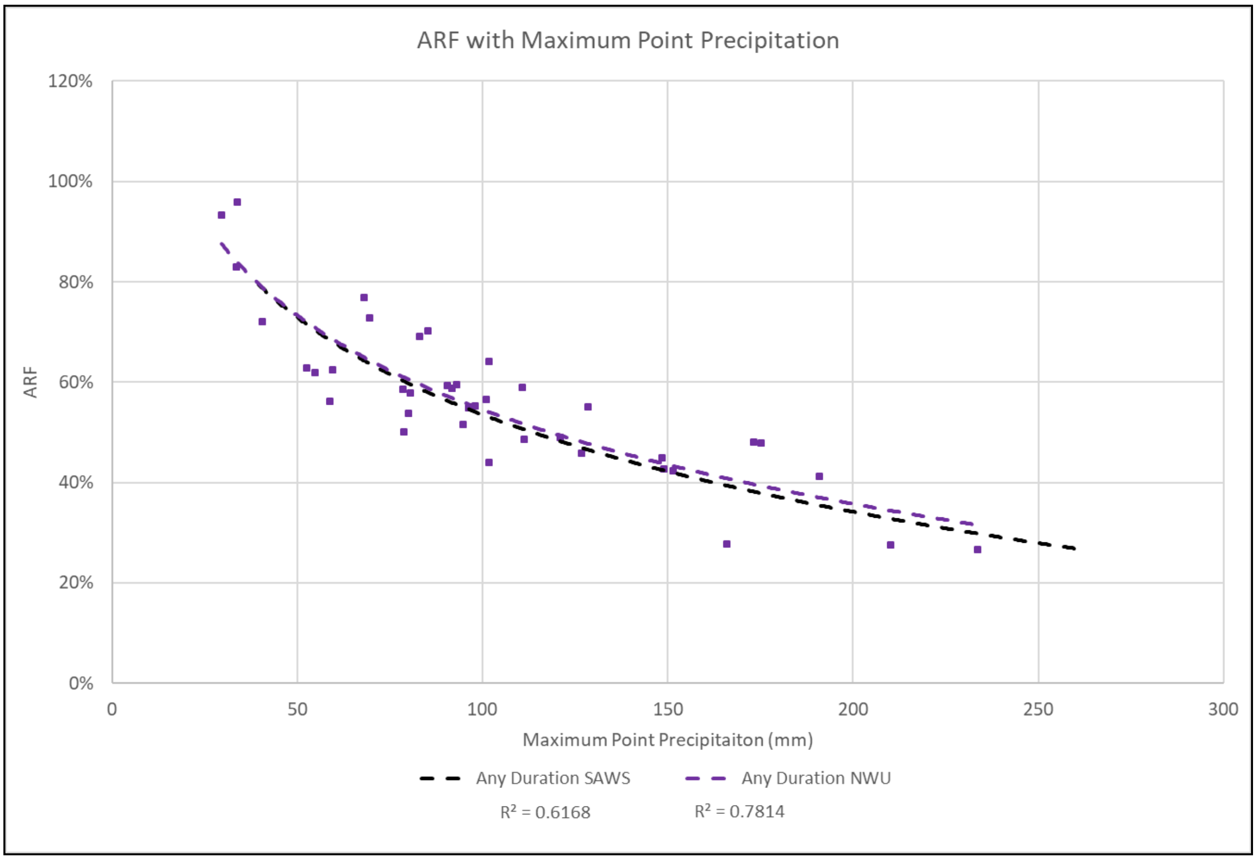


Figure 5.19: Comparison of the SAWS and NWU maximum Point rainfall depth data sets

5.3.4 Storm duration

Figure 5.20 displays the <u>average</u> ARFs for the 3 and 24-hour data sets (the 1-h radar data from NWU proved to be unusable), for each storm area category. The SAWS data sets yielded higher average ARFs for all storm area and duration categories. The lowest average ARF difference is 0.54% for the 3-hour data set associated with storm areas less than 200 km^2. The maximum average ARF difference is 10.72% for the 24-hour data set associated with storm areas less than 500 km^2. The differences between the average ARFs for different storm durations associated with different storm areas are regarded as negligible, especially if the size of the data sets used and the approaches followed to provide one consolidate data set, are considered.

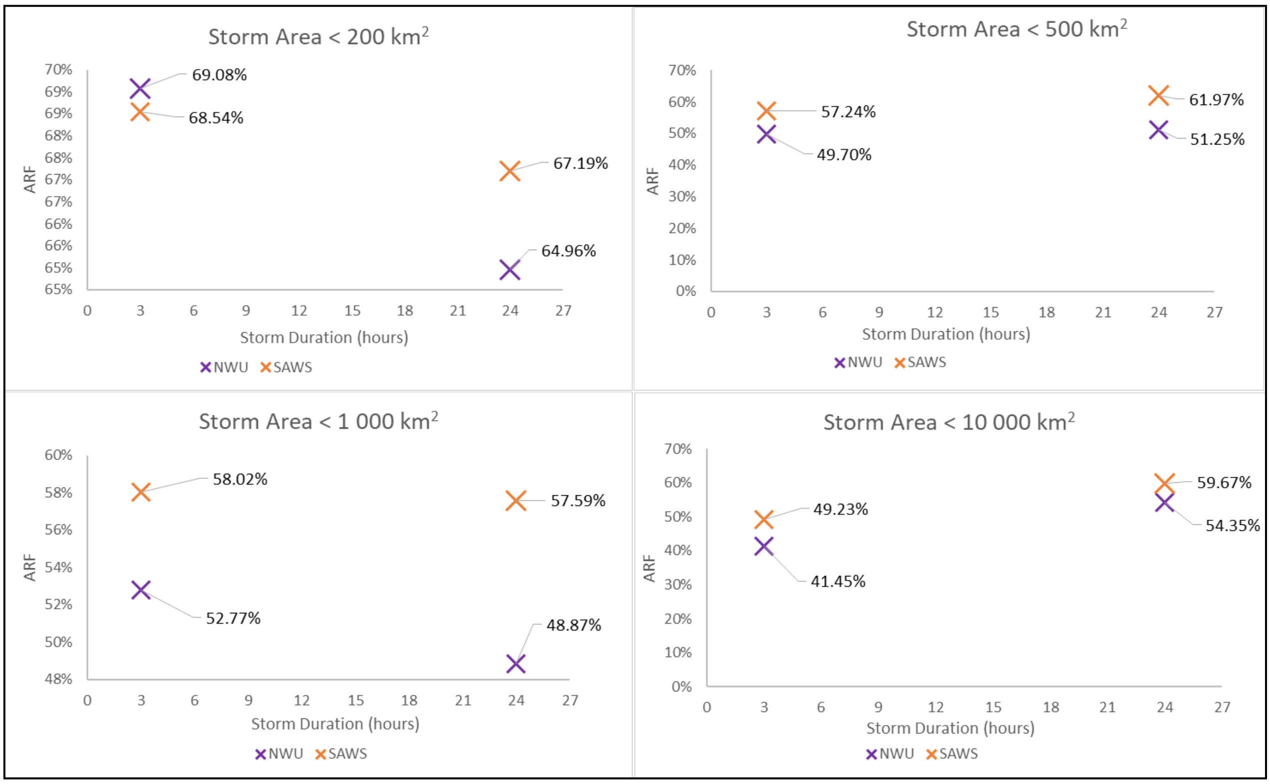


Figure 5.20: Comparison of the SAWS and NWU storm duration data sets

6 CONCLUSIONS

The main objective of this research was to establish if radar data can be used to derive ARFs in South Africa. In addressing this question, radar data was used to derive ARFs and in doing so, the applicability of the use of the radar data was assessed. It is important to note that the aim is not to derive new ARFs, but the process of deriving ARFs only serves as a means to assess the applicability of radar data.

An areal reduction factor is defined as a factor that is applied to convert point rainfall depths to an average rainfall depth over a specific catchment area. There are two main methods currently used to derive ARFs: the storm-centred method, which considers the area of a specific storm, and the geographically-fixed method, which considers rainfall occurring over a specific, fixed location. For South Africa, the ARFs currently implemented were derived in the 1960's and estimated by Van Wyk (1965) and Pullen, Wiederhold and Midgley (1966), using the storm-centred method. The storms used in these analyses lack adequate spatial resolution, which could significantly affect the derivation of ARFs. Thus, radar data was tested as an alternative, as it hosts the ability to produce high spatial and temporal resolution storms for analysis and it considers the internal structure and behaviour of storms.

The availability of data proves to be a critical aspect of this research. While various sources of data have been identified, various challenges resulted into data being made available only from the SAWS and the NWU and in both cases, in limited numbers mainly due to the processing required before the data can be used. SAWS provided radar imagery for 1-, 3- and 24-hour storm durations, in MDV file format that was used for the analyses. MDV files were converted to netcdf files for analysis in ArcGIS. The NWU provided 6-minute time step MDV files for the radar data, which was used to compile 1-, 3- and 24-h storm duration netcdf data files. An extensive methodology was put in place, following the storm-centred approach, to derive ARFs from these netcdf files using ArcGIS.

The radar data derived ARFs for 1-, 3- and 24-hour storm durations were statistically analysed using regression analyses. The analysis concluded that:

- An inverse proportional relationship exists between ARFs and the respective storm areas.
- An inverse proportional relationship exists between ARFs and the respective maximum point rainfall intensity associated with different storm durations. When disregarding the influence of storm duration, little correlation was found.
- An inverse proportional relationship exists between ARFs and the maximum point rainfall depths.
- Higher ARFs were associated with convective rainfall as opposed to frontal rainfall. The latter is in contradiction with existing literature. The limited number of storms (coastal) selected may be partially responsible for this anomaly.
- The influence of storm duration on ARFs did not yield consistent results. It is speculated that the uneven distribution of storm areas within these analyses resulted in possible inconsistencies. An analysis, taking storm area into consideration, proved only partially successful, with storm areas less than 1 000 km^2 following an expected trend of lower ARF's with an increase in duration. The opposite was observed for storm areas larger than 1 000 km^2.

- Radar derived ARFs were generally all lower than the ARFs derived using the diagrams developed by Van Wyk (1965) or the Op ten Noort and Stephenson's (1982) Equation 2 for small catchment areas (< 800 km^2).
- By considering storm area categories, the radar derived ARFs for storm areas less than 800 km^2 for smaller catchments were all, except for a very limited number of storms in the 600-800 km^2 category, found to be lower than the ARFs currently in use in South Africa. . A tendency was observed that a threshold value might exist where, for higher rainfall intensities, the radar-based ARFs might exceed those suggested by Van Wyk (1965).
- For large catchment areas (>800 km^2) radar derived ARFs were generally lower than those ARFs derived using the diagrams developed by Pullen, Wiederhold and Midgley (1966), as well as the Op ten Noort and Stephenson's (1982) Equation 3, with the latter equation being more conservative (higher).
- When comparing ARFs taking storm duration into consideration for larger catchments areas, radar derived ARFs were lower than the ARFs currently in use in South Africa for storm durations of 3- and 24-h. The radar derived ARFs for the 1-hour storm events were found to be overestimated. In using the Op ten Noort and Stephenson's (1982) Equation 3, ARFs were produced that were, on average, significantly higher than the radar derived ARFs.
- No significant difference could be found when comparing the results of the ARFs as derived based on the SAWS and the NWU radar data sets. These comparisons took maximum point rainfall intensity, storm area, maximum point rainfall depth and storm duration into consideration.

It is clear from these findings that radar data can be used to derive ARFs in South Africa. Most of the findings were in agreement with the expected outcomes, based on methodologies used in South Africa since the 1960's. The comparisons of radar derived ARFs and the ARFs currently used in South Africa, suggest that there could be value in reproducing ARFs for South Africa, considering the significant evolution of data capturing since 1965, using advanced technology (such as radars).

The radar-based ARFs are generally lower than those estimated using conventional methods. This suggests that the current methods may yield ARFs that are conservative. Radars provide more detailed information for the analyses of storms when compared to using traditional methods. Although radars are known for their large uncertainties, radars expose the potential of significantly improving certain aspects of hydrological calculations.

Perhaps the most challenging finding of the research is the accessibility of data. The methodology has been developed and tested on the limited available data sets, but the retrieval of data from the different data suppliers proves to be challenging, mainly due to available resources at the different institutions. It was also clear during the process, that high level agreements will have to be put in place; specifically between the SAWS and research institutions to enable access to radar data, which is considered as a valuable commodity, with a high monetary value attached to it, by the SAWS.

It is also evident that the pool of expertise to be used in the collection, cleaning and post processing of the radar data into a format that can be used for researchers, need to be expanded.

Once the availability of data from different sources (specifically the SAWS) has been resolved, it is proposed that further research be done on specifically the selection of appropriate storm duration selection procedures, the selection of more appropriate categories of storms (i.e. coastal vs inland) as well as storms with a higher intensity. Future research also needs to focus on appropriate platforms from where raw radar data (after initial post-analysis by owners) can be accessed as appose to the transfer of huge volumes of data.

****

7 REFERENCES

Adamson, PT. (1981). Southern African Storm Rainfall. Technical Report TR102. Department of Environmental Affairs and Tourism. Pretoria, RSA.

Alexander, WJR. (1978). Depth-Area-Duration-Frequency Properties of Storm Rainfall in South Africa. Technical Report TR83. Department of Water Affairs, Pretoria, RSA.

Alexander, WJR. (2001). Flood Risk Reduction Measures: Incorporating Flood Hydrology for Southern Africa. Department of Civil and Biosystems Engineering, University of Pretoria, Pretoria, RSA.

Alexander, WJR. (2010). Analytical Methods for Water Resource Development and Management: Handbook for Practitioners and Decision Makers, [online]. Department of Civil and Biosystems Engineering, University of Pretoria, Pretoria, RSA. Available from: http://www.droughtsandfloods.com [5 February 2013].

Allen, RJ and DeGaetano, AT. (2005a) 'Areal Reduction Factors for Two Eastern United States Regions with High Rain-Gauge Density', *Journal of Hydrologic Engineering*, 10(4), pp. 327-335. doi: 10.1061/(asce)1084-0699(2005)10:4(327).

Allen, RJ and DeGaetano, AT. (2005b) 'Considerations for the use of radar-derived precipitation estimates in determining return intervals for extreme areal precipitation amounts', *Journal of Hydrology*, 315(1-4), pp. 203-219. doi: 10.1016/j.jhydrol.2005.03.028.

Asquith, WH and Famiglietti, JS. (2000) 'Precipitation areal-reduction factor estimation using an annual-maxima centered approach', *Journal of Hydrology*, 230(1-2), pp. 55-69. doi: 10.1016/S0022-1694(00)00170-0.

Bacchi, B and Ranzi, R. (1996) 'On the derivation of the areal reduction factor of storms', *Atmospheric Research*, 42(1-4), pp. 123-135. doi: 10.1016/0169-8095(95)00058-5.

Beard, LR. (1967) 'Optimization techniques for hydrologic engineering', *Water Resources Research*, 3(3), pp. 809-815. doi: 10.1029/WR003i003p00809.

Becker, EH and Pegram, G. (2014) *College of Agriculture , Engineering and Science Handbook for 2013*.

Bárdossy, A and Pegram, G. (2018). Intensity-duration-frequency curves exploiting neighbouring extreme precipitation data. Hydrological Sciences Journal 63 (11): 1593-1604.

Bell, F. (1976) 'The Areal Reduction Factor in Rainfall Frequency Estimation', *Report No. 35*.

Bengtsson, L and Niemczynowicz, J. (1986) 'Areal Reduction Factors from Rain Movement', *Hydrology Research*, 17(2), pp. 65-82. doi: 10.2166/nh.1986.0005.

Bureau, UW. (1957) '1957. Rainfall intensity-frequency regime', *Technical Paper 29*.

Cunnane, C and Lynn, MA. (1975) 'Flood Estimation following the Flood Studies Report', *The Institution of Engineers of Ireland*, (February), p. 39.

Cunnane, C. (1989). Statistical Distributions for Flood Frequency Analysis. WMO Report 718. World Meteorological Organisation, Geneva, Switzerland.

Davies, B and Day, J. (1998). Vanishing Waters. 1st ed. University of Cape Town Press, Cape Town, RSA.

De Michéle, C, Kottegoda, NT and Rosso, R. (2001). The derivation of areal reduction factor of storm rainfall from its scaling properties. Water Resources Research 37 (12): 3247-3252.

De Waal, K. (2018). Email Correspondence. 4 September. Stellenbosch.

Dixon, M. (2006) 'MDV FORMAT Interface Control Document (ICD)', (November), pp. 1-30.

Du Plessis, JA and Loots, W. (2019) 'Re-evaluation of Areal Reduction Factors and their impact on floods in South Africa', *Arabian Journal of Geosciences, in press*.

Gill, TD. (2005) 'Transformation of Point Rainfall To Areal Rainfall By', (May).

Grebner, G and Roesch, T. (1997). Regional dependence and application of DAD relationships. International Association of Hydrological Sciences 246: 223-230.

Haarhoff, J and Cassa, AM. (2009). Introduction to Flood Hydrology. Juta and Company Limited, Cape Town, RSA.

Hershfield, DM. (1962). Extreme rainfall relationships. Proceedings, American Society of Civil Engineers HY6 (11):73-92.

Hogg, WD. (1992). Inhomogeneities in Time Series of Extreme Rainfall. 5th International Meeting on Statistical Climatology. The Steering Committee for International Meetings on Statistical Climatology, Toronto, Canada 481-484.

Hosking, JRM and Wallis, JR. (1997). Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press. Cambridge, UK.

Huff, FA and Shipp, WL. (1969). Spatial correlations of storm, monthly, and seasonal rainfall. Journal of Applied Meteorology 8 (4): 542-550.

Huff, FA and Shipp, W. L. (2002) 'Spatial Correlations of Storm, Monthly and Seasonal Precipitation', *Journal of Applied Meteorology*, pp. 542-550. doi: 10.1175/1520-0450(1969)008<0542:scosma>2.0.co;2.

Kim, J, Lee, J, Kim, D and Kang, B. (2019) 'The role of rainfall spatial variability in estimating areal reduction factors', *Journal of Hydrology*. Elsevier, 568(June 2018), pp. 416-426. doi: 10.1016/j.jhydrol.2018.11.014.

Lambourne, JJ and Stephenson, D. (1986). Research in Urban Hydrology and Drainage: Factors affecting Storm Runoff in South Africa. Water Systems Research Programme Report No. 2/86. University of Witwatersrand. Johannesburg, RSA.

Lombardo, F, Napolitano, F and Russo, F. (2006) 'On the use of radar reflectivity for estimation of the areal reduction factor', *Natural Hazards and Earth System Science*, 6(3), pp. 377-386. doi: 10.5194/nhess-6-377-2006.

Midgley, DC and Pitman, WV. (1978). A Depth-Duration-Frequency Diagram for Point Rainfall in Southern Africa. HRU Report 2/78. University of Witwatersrand. Johannesburg, RSA.

Mineo, C, Ridolfi, E, Napolitano, F and Russo, F. (2018). The areal reduction factor: A new analytical expression for the Lazio Region in central Italy. Journal of Hydrology 560: 471-479.

NERC. (1975). Flood Studies Report. Natural Environment Research Council, London, UK.

Omolayo, AS. (1993) 'On the transposition of areal reduction factors for rainfall frequency estimation', *Journal of Hydrology*, 145(1-2), pp. 191-205. doi: 10.1016/0022-1694(93)90227-Z.

Op ten Noort, TH and Stephenson, D. (1982). Flood Peak Calculation in South Africa. Water System Research Programme Report No. 2/1982. University of the Witwatersrand. Johannesburg, RSA.

Overeem, A, Buishand, T, Holleman, I and Uijlenhoet, R. (2010) 'Extreme value modeling of areal rainfall from weather radar', *Water Resources Research*, 46(9), pp. 1-13. doi: 10.1029/2009WR008517.

Pavlovic, S, Perica, S, St Laurent, M and Mejía, A. (2016) 'Intercomparison of selected fixed-area areal reduction factor methods', *Journal of Hydrology*. Elsevier B.V., 537, pp. 419-430. doi: 10.1016/j.jhydrol.2016.03.027.

Peleg, N, Marra, F, Fatichi, S, Paschalis, A, Molnar, P and Burlando, P. (2018). Spatial variability of extreme rainfall at radar subpixel scale. Journal of Hydrology 556: 922-933.

Pegram, GGS and Adamson, PT. (1988). Revised risk analysis for extreme storms and floods in

Natal/KwaZulu. The Civil Engineer in South Africa: 331-336.

Picus, C, Beleznai, C, Nowak, C, Ramoser, H and Mitterhubber, S. (2008) 'Kernel-based storm tracking in radar data', *2008 IEEE Radar Conference, RADAR 2008*. doi: 10.1109/RADAR.2008.4721112.

Pietersen, JPJ, Gericke, OJ, Smithers, JC and Woyessa, YE. (2015). Review of current methods for estimating areal reduction factors applied to South African design point rainfall and preliminary identification of new methods. Journal of the South African Institution of Civil Engineering 57 (1): 16-30.

Podger, S, Green, J, Jolly, C and Beesley, C. (2015a). Creating long duration areal reduction factors. In: Proceedings, 36th Hydrology and Water Resources Symposium: The Art and Science of Water. Institution of Engineers, Australia. p. 39.

Podger, S, Green, J, Steynsmyr, P and Babister, M. (2015b). Combining long and short duration areal reduction factors. In: Proceedings, 36th Hydrology and Water Resources Symposium: The Art and Science of Water. Institution of Engineers, Australia. p. 210.

Prudhomme, C. (1999) 'Mapping a statistic of extreme rainfall in a mountainous region', *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 24(1-2), pp. 79-84. doi: 10.1016/S1464-1909(98)00015-X.

Pullen, R., Wiederhold, J. and Midgley, D. (1966) 'STORM STUDIES IN SOUTH AFRICA : LARGE-AREA STORMS : DEPTH-AREA-DU RATION ANALY-SIS BY DIGITAL COMPUTER', *The Civil Engineer in South Africa*, pp. 173-187.

Rodriguez-Iturbe, I and Mejía, J. (1974) 'On the transformation of point rainfall to areal rainfall', *Water Resour. Res., 10 (4)*, pp. 729-735.

SANRAL. (2013). Drainage Manual. 6th ed. South African National Roads Agency Limited, Pretoria, RSA.

Schulze, RE. (1980). Potential Flood Producing Rainfall for Medium and Long Duration in Southern Africa. Report to Water Research Commission, Pretoria, RSA.

Schulze, RE. (1984). Depth-duration-frequency Studies in Natal based on Digitised Data. South African National Hydrology Symposium, Technical Report TR119. Department of Environment Affairs. Pretoria, RSA.

Schulze, RE, Schmidt, EJ, and Smithers, JC. (1992). SCS-SA User Manual: PC-Based SCS Design Flood Estimates for Small Catchments in Southern Africa. ACRU Report No. 40. Department of Agricultural Engineering, University of Natal, Pietermaritzburg, RSA.

Sinclair, S and Pegram, G. (2005) 'Combining radar and rain gauge rainfall estimates using conditional merging', *Atmospheric Science Letters*, 6(1), pp. 19-22. doi: 10.1002/asl.85.

Siriwardena, L and Weinmann, PE. (1996). Derivation of Areal Reduction Factors for Design Rainfalls in Victoria. Report No. 96/4. Cooperative Research Centre for Catchment Hydrology. Victoria, Australia.

Sivapalan, M and Blöschl, G. (1998). Transformation of point rainfall to areal rainfall: Intensity-duration-frequency curves. Journal of Hydrology 204 (1-4): 150-167.

Skaugen, T. (1997) 'Classification of rainfall into small- and large-scale events by statistical pattern recognition', *Journal of Hydrology*, 200(1-4), pp. 40-57. doi: 10.1016/S0022-1694(97)00003-6.

Smithers, JC and Schulze, RE. (2000a). Development and Evaluation of Techniques for Estimating Short Duration Design Rainfall in South Africa. WRC Report No. 681/1/00. Water Research Commission, Pretoria, RSA.

Smithers, JC and Schulze, RE. (2000b). Long Duration Design Rainfall Estimates for South Africa. WRC Report No. 811/1/00. Water Research Commission, Pretoria, RSA.

Smithers, JC and Schulze, RE. (2003). Design Rainfall and Flood Estimation in South Africa. WRC Report 1060/01/03. Water Research Commission, Pretoria, RSA.

Smithers, JC and Schulze, RE. (2004). The estimation of design rainfall for South Africa using a regional scale invariant approach. Water SA 30 (4):435-444.

Stewart, EJ. (1989) 'Areal reduction factors for design storm construction: Joint use of raingauge and radar data', *IAHS Publ.*, 181(181), pp. 31-49.

Supino, G. (1964) 'Le reti idrauliche'.

Svensson, C and Jones, D. A. (2010) 'Review of methods for deriving areal reduction factors', *Journal of Flood Risk Management*, 3(3), pp. 232-245. doi: 10.1111/j.1753-318X.2010.01075.x.

Terblanche, DE, Pegram, GGS and Mittermaier, MP. (2001) 'The development of weather radar as a research and operational tool for hydrology in south africa', *Journal of Hydrology*, 241(1-2), pp. 3-25. doi: 10.1016/S0022-1694(00)00372-3.

USWB. (1957). Rainfall Intensity-Frequency Regime, The Ohio Valley. Technical Paper 29. United States Weather Bureau, Washington, DC, USA.

Van der Spuy, D and Rademeyer, PF. (2018). Flood Frequency Estimation Methods as Applied in the Department of Water and Sanitation. DWS, Pretoria, RSA.

Van Wyk, W. (1965). Aids to the Prediction of Extreme Floods from Small Watersheds. Department of Civil Engineering, University of the Witwatersrand, Johannesburg, RSA.

Veneziano, D and Langousis, A. (2005) 'The areal reduction factor : A multifractal analysis', 41, pp. 1-15. doi: 10.1029/2004WR003765.

Viessman, W, Lewis, GL, and Knapp, JW. (1989). Introduction to Hydrology. 3rd ed. Harper and Row Publishers Incorporated, New York, USA.

Weddepohl, JP. (1988). Design Rainfall Distributions for Southern Africa. Unpublished MSc dissertation, Department of Agricultural Engineering, University of Natal. Pietermaritzburg, RSA.

Wiederhold, JFA. (1969). Design Storm Determination in South Africa. HRU Report No. 1/1969. Hydrological Research Unit, University of the Witwatersrand. Johannesburg, RSA.

Wolff, C. (2002) *Radar Basics*. Available at: http://www.radartutorial.eu/07.waves/Waves and Frequency Ranges.en.html.

Wright, DB, Smith, JA, Villarini, G. and Baek, M. L. (2013) 'Estimating the frequency of extreme rainfall using weather radar and stochastic storm transposition', *Journal of Hydrology*. Elsevier B.V., 488, pp. 150-165. doi: 10.1016/j.jhydrol.2013.03.003.

Wright, DB, Smith, JA and Baeck, ML. (2014) 'Critical Examination of Area Reduction Factors', *Journal of Hydrologic Engineering*, 19(4), pp. 769-776. doi: 10.1061/(ASCE)HE.1943-5584.0000855.

Yoo, C, Kim, K, Kim, HS and Park, MJ. (2007). Estimation of areal reduction factors using a mixed gamma distribution. Journal of Hydrology 335 (3-4): 271-284.

APPENDIX A

```
/**********************************************************************
 * TDRP params for MdvConvert
 **********************************************************************/

//======================================================================
//
// Program name: MdvConvert.
//
// MdvConvert reads mdv data, converts it in various ways, and writes it
//   out. The usage is 'MdvConvert -params params_file'.
//
//======================================================================

//======================================================================
// Convert UAE Merge for GCC use to NetCDF.
//======================================================================

//======================================================================
//
// DEBUGGING AND PROCESS CONTROL.
//
//======================================================================

///////////// debug ///////////////////////////////////
//
// Debug option.
// If set, debug messages will be printed appropriately.
//
// Type: enum
// Options:
//     DEBUG_OFF
//     DEBUG_NORM
//     DEBUG_VERBOSE
//

debug = DEBUG_OFF;

///////////// instance ////////////////////////////////
//
// Process instance.
// Used for registration with procmap.
// Type: string
//

instance = "precip";

///////////// reg_interval ////////////////////////////
//
// Registration interval.
// The number of seconds between expected procmap registrations.
// Type: int
//

reg_interval = 60;

//======================================================================
//
// DATA INPUT.
//
//======================================================================

///////////// mode ////////////////////////////////////
//
// Operating mode.
// In REALTIME mode, the program waits for a new input file.
```

```
//   LOCAL_FILEPATH_REALTIME is a realtime mode used when data resides on
//   the host where the application is running. This was added due to the
//   problems that the data server layer had distinguishing data times of
//   data written very close together. NOTE that in this mode the
//   input_url parameter should be set to an input directory, not an input
//   url. In ARCHIVE mode, it moves through the data between the start and
//   end times set on the command line. In SPEC_FCAST_REALTIME mode, the
//   program waits for a new input file that is a forecast file with the
//   specified forecast lead time. The forecast lead time is specified in
//   the fcast_lead_time parameter. In FILELIST mode, it moves through the
//   list of file names specified on the command line. Paths (in FILELIST
//   mode, at least) MUST contain a day-directory below the data file --
//   ./data_file.mdv will not work as a file path. In ARCHIVE_FCST mode,
//   it moves through the data between the start and end times set on the
//   comand line, and processes all lead times found in the forecast
//   files. In SPEC_FCST_ARCHIVE mode, it moves through the data between
//   the start and end times set on the comand line, and processes
//   forecast files with the specified forecast lead time.
//
// Type: enum
// Options:
//     ARCHIVE
//     REALTIME
//     FILELIST
//     SPEC_FCAST_REALTIME
//     REALTIME_FCST_DATA
//     LOCAL_FILEPATH_REALTIME
//     ARCHIVE_FCST
//     SPEC_FCST_ARCHIVE
//

mode = FILELIST;

///////////// local ///////////////////////////////////
//
// Setup for LOCAL_FILEPATH_REALTIME mode ONLY. Max age of input,
//   seconds, if we use latest data info to trigger, and if we should only
//   process the latest file.
// Defaults should generally be fine.
//
// Type: struct
//   typedef struct {
//      int lookback;
//      boolean use_ldata_info;
//      boolean latest_file_only;
//   }
//
//

local = {
    lookback = 1200,
    use_ldata_info = TRUE,
    latest_file_only = FALSE
};

///////////// input_url ///////////////////////////////
//
// URL for input data.
// This is used in REALTIME and ARCHIVE modes only. In FILELIST mode,
//   the file paths are specified on the command line. In
//   LOCAL_FILEPATH_REALTIME mode, set this to a directory, not a URL.
// Type: string
//

input_url = "$(DATA_DIR)/mdv/precip/1hr";

///////////// writeLdataInfo //////////////////////////
//
// Write _latest_data_info files for output files.
```

```
// If false, will suppress writing of _latest_data_info files.
// Type: boolean
//

writeLdataInfo = TRUE;

///////////// fcast_lead_time /////////////////////////
//
// Forecast lead time information for forecast files to be processed.
// Used only if mode is set to SPEC_FCAST_REALTIME or SPEC_FCST_ARCHIVE.
//
// Type: struct
//   typedef struct {
//      int lead_time_secs;
//      boolean use_gen_time;
//   }
//
//

fcast_lead_time = {
    lead_time_secs = 0,
    use_gen_time = FALSE
};

///////////// do_lead_time_subsampling ////////////////
//
// Lead time subsampling flag.
// Set to true to enable lead time subsampling.  Used only if mode is
//   REALTIME_FCST_DATA or ARCHIVE_FCST.
// Type: boolean
//

do_lead_time_subsampling = FALSE;

///////////// subsample_lead_time_hour ////////////////
//
// The subsampled lead times to process.
// Type: double
// 1D array - variable length.
//

subsample_lead_time_hour = { 0 };

///////////// set_field_nums /////////////////////////
//
// Option to set field numbers.
// Type: boolean
//

set_field_nums = FALSE;

///////////// field_nums /////////////////////////////
//
// Field number list.
// Type: int
// 1D array - variable length.
//

field_nums = {
 0
};

///////////// set_field_names ////////////////////////
//
// Option to set field names.
// Type: boolean
//

set_field_names = FALSE;
```

```
///////////// field_names /////////////////////////////
//
// Field name list.
// Type: string
// 1D array - variable length.
//

field_names = {
 "0"
};

///////////// rename_fields ///////////////////////////
//
// Option to set field names.
// Type: boolean
//

rename_fields = FALSE;

///////////// new_names ///////////////////////////////
//
// Provides a map from old field name to new field name. Note that
//   either the filed name or the long field name must match the
//   old_field_name specified for the renaming to take effect, and that if
//   the renaming happens then both the field name and the long field name
//   are renamed to the new_field_name.
//
// Type: struct
//   typedef struct {
//      string old_field_name;
//      string new_field_name;
//   }
//
// 1D array - variable length.
//

new_names = {
  {
    old_field_name = "",
    new_field_name = ""
  }
};

///////////// apply_thresholds_to_field_values ////////
//
// Option to threshold field values. Points with values outside the
//   specified limits will be set to missing.
// NOTE: this works on the output field names. If rename_fields is
//   false, then the input and output field names are the same. If rename
//   fields is true, the field name change is performed first, before the
//   field values are thresholded.
// Type: boolean
//

apply_thresholds_to_field_values = TRUE;

///////////// thresholded_fields //////////////////////
//
// Limit the values in specified fields to between min_threshold and
//   max_threshold. Values outside this range will be set to missing.
//
// Type: struct
//   typedef struct {
//      string output_field_name;
//      double threshold_min;
//      double threshold_max;
//   }
//
```

```
// 1D array - variable length.
//

thresholded_fields = {
  {
    output_field_name = "precip",
    threshold_min = 0,
    threshold_max = 250
  }
};

//======================================================================
//
// DATA OUTPUT.
//
//======================================================================

///////////// output_url //////////////////////////////
//
// Output URL.
// Output data is written to this URL.
// Type: string
//

output_url = "$(DATA_DIR)/netcdf/precip1hr";

///////////// output_as_forecast //////////////////////
//
// Set to output the data as forecast in mdv format.
// This forces a forecast-style output, whether the data is of forecast
//   type or not.
// Type: boolean
//

output_as_forecast = FALSE;

///////////// if_forecast_output_as_forecast //////////
//
// Set to output the data as forecast, if the data is of a forecast
//   type.
// This only writes out in forecast-style output if the
//   data_collection_type in the master header is of type FORECAST or
//   EXTRAPOLATED.
// Type: boolean
//

if_forecast_output_as_forecast = FALSE;

///////////// output_format ///////////////////////////
//
// Specify format of file on output.
// FORMAT_MDV: normal MDV formal. FORMAT_XML: XML format. XML format
//   writes out 2 files: *.mdv.xml and *.mdv.buf. The xml file contains
//   the meta-data. The buf file contains the binary fields.
// NOTE: only COMPRESSION_NONE and COMPRESSION_GZIP_VOL are supported in
//   XML. FORMAT_NCF: write file in netCDF CF format. Extension will be
//   .nc.
//
// Type: enum
// Options:
//     OUTPUT_FORMAT_MDV
//     OUTPUT_FORMAT_XML
//     OUTPUT_FORMAT_NCF
//

output_format = OUTPUT_FORMAT_NCF;

///////////// write_to_path ///////////////////////////
//
```

```
// Write the file to a specified path.
// This overrides output_url.
// Type: boolean
//

write_to_path = FALSE;

///////////// output_path ////////////////////////////
//
// Output path.
// See 'write_to_path'.
// Type: string
//

output_path = "./output/test.mdv";

//======================================================================
//
// GEOMETRY CONVERSION.
//
//======================================================================

///////////// set_horiz_limits ///////////////////////
//
// Option to set horizontal limits.
// Type: boolean
//

set_horiz_limits = FALSE;

///////////// horiz_limits ///////////////////////////
//
// Set horizontal limits.
//
// Type: struct
//   typedef struct {
//      float min_lat;
//      float min_lon;
//      float max_lat;
//      float max_lon;
//   }
//
//

horiz_limits = {
    min_lat = -90,
    min_lon = -180,
    max_lat = 90,
    max_lon = 180
};

///////////// set_vlevel_limits //////////////////////
//
// Option to set plane vlevel limits.
// Mutually exclusive with set_plane_num_limits.
// Type: boolean
//

set_vlevel_limits = FALSE;

///////////// lower_vlevel ///////////////////////////
//
// Lower plane vlevel limit.
// Type: float
//

lower_vlevel = 0;

///////////// upper_vlevel ///////////////////////////
```

```
//
// Upper plane vlevel limit.
// Type: float
//

upper_vlevel = 0;

///////////// override_vlevels ///////////////////////
//
// Option to override the vlevels in the vlevel header.
// If true, will replace the vlevels in the header with those specified
//   in 'vlevel_array'. This does not affect the actual data in the file.
// Type: boolean
//

override_vlevels = FALSE;

///////////// vlevel_array ///////////////////////////
//
// vlevel values to override what is already in the file.
// See 'override_vlevels'.
// Type: double
// 1D array - variable length.
//

vlevel_array = {
 0
};

///////////// set_plane_num_limits ///////////////////
//
// Option to set plane number limits.
// Mutually exclusive with set_vlevel_limits.
// Type: boolean
//

set_plane_num_limits = FALSE;

///////////// lower_plane_num ////////////////////////
//
// Lower plane num limit.
// Type: int
//

lower_plane_num = 0;

///////////// upper_plane_num ////////////////////////
//
// Upper plane num limit.
// Type: int
//

upper_plane_num = 0;

///////////// composite //////////////////////////////
//
// Option for creating composite.
// Composite is a plane in which each grid location contains the maximum
//   value at any height.
// Type: boolean
//

composite = FALSE;

///////////// remap_z_to_constant_grid ///////////////
//
// Option to remap the Z levels onto a grid with constant dz.
// Field data will be remapped onto the specified Z levels using the
//   nearest neighbor method. See 'remap_z_grid'. Note that this actually
```

```
//   changes the data. Whereas 'override_vlevels' only changes the vlevels
//   in the headers, and does not change the data.
// Type: boolean
//

remap_z_to_constant_grid = FALSE;

///////////// remap_z_grid ///////////////////////////
//
// Specified Z levels for remapping.
//
// Type: struct
//   typedef struct {
//      int nz;
//      double minz;
//      double dz;
//   }
//
//

remap_z_grid = {
    nz = 18,
    minz = 0,
    dz = 1
};

///////////// remap_xy ///////////////////////////////
//
// Option to remap grid in x,y.
// If true, set the remap parameters below.
// Type: boolean
//

remap_xy = FALSE;

///////////// auto_remap_to_latlon ///////////////////
//
// Option to automatically remap the grid to a lat-lon projection.
// If true, the data in the file will be remapped to a latlon grid which
//   matches the existing grid in resolution and extent. Other remap
//   parameters will be ignored.
// Type: boolean
//

auto_remap_to_latlon = FALSE;

///////////// remap_at_source ////////////////////////
//
// Flag indicating where to do the remapping.
// If set to true, the remapping is done on the source machine by
//   setting the remapping in the MDV read request. This is the default.If
//   set to false, the remapping is done on the destination machine by
//   doing a remap command after the read is done. This is useful if you
//   are reading the data from a machine that is overloaded.
// Type: boolean
//

remap_at_source = FALSE;

///////////// remap_projection ///////////////////////
//
// Projection for remapping in x,y. See projection param below.
//     PROJ_LATLON: simple lat/lon grid (Equidistant Cylindrical)
//     PROJ_FLAT: Azimuthal Equidistant (Radar)
//     PROJ_LAMBERT_CONF: Lambert Conformal Conic
//     PROJ_LAMBERT_AZIM: Lambert Azimuthal Equal Area
//     PROJ_MERCATOR: Mercator - EW orientation
//     PROJ_TRANS_MERCATOR: Tranverse Mercator - NS orientation
//     PROJ_POLAR_STEREO: Stereographic- polar aspect
```

```
//     PROJ_OBLIQUE_STEREO: Stereographic - oblique aspect
//     PROJ_ALBERS: Albers Equal Area Conic
//     PROJ_VERT_PERSP: Vertical Perspective (satellite view).
//
// Type: enum
// Options:
//     PROJ_LATLON
//     PROJ_LAMBERT_CONF
//     PROJ_MERCATOR
//     PROJ_POLAR_STEREO
//     PROJ_FLAT
//     PROJ_OBLIQUE_STEREO
//     PROJ_TRANS_MERCATOR
//     PROJ_ALBERS
//     PROJ_LAMBERT_AZIM
//     PROJ_VERT_PERSP
//

remap_projection = PROJ_LATLON;

///////////// remap_grid //////////////////////////////
//
// Grid parameters for remapping in x,y.
// Units in km, except for LATLON, which is in degrees.
//
// Type: struct
//   typedef struct {
//      int nx;
//      int ny;
//      double minx;
//      double miny;
//      double dx;
//      double dy;
//   }
//
//

remap_grid = {
    nx = 1,
    ny = 1,
    minx = 0,
    miny = 0,
    dx = 1,
    dy = 1
};

///////////// remap_rotation //////////////////////////
//
// Remapped grid rotation.
// This applies only to PROJ_FLAT projections.
// Type: double
//

remap_rotation = 0;

///////////// remap_origin_lat ////////////////////////
//
// Remapped grid origin latitude.
// This applies to all projections except LATLON.
// Type: double
//

remap_origin_lat = 0;

///////////// remap_origin_lon ////////////////////////
//
// Remapped grid origin longitude.
// This applies to all projections except LATLON.
// Type: double
```

```
//

remap_origin_lon = 0;

///////////// remap_lat1 //////////////////////////////
//
// Remapped grid reference latitude 1.
// This applies to LAMBERT_CONF and ALBERS projections.
// Type: double
//

remap_lat1 = 0;

///////////// remap_lat2 //////////////////////////////
//
// Remapped grid reference latitude 2.
// This applies to LAMBERT_CONF and ALBERS projections.
// Type: double
//

remap_lat2 = 0;

///////////// remap_central_scale ////////////////////
//
// Central scale for remapped projections.
// This applies to POLAR_STEREO, OBLIQUE_STEREO and TRANSVERSE_MERCATOR
//   projections.
// Type: double
//

remap_central_scale = 1;

///////////// remap_tangent_lat //////////////////////
//
// Remapped tangent latitude (deg).
// This applies to OBLIQUE_STEREO only.
// Type: double
//

remap_tangent_lat = 0;

///////////// remap_tangent_lon //////////////////////
//
// Remapped tangent longitude (deg).
// This applies to OBLIQUE_STEREO and POLAR_STEREO.
// Type: double
//

remap_tangent_lon = 0;

///////////// remap_pole_is_north ////////////////////
//
// Flag indicating stereogtraphic is over the NORTH pole.
// This applies to POLAR_STEREO. If false, the projection is over the
//   south pole.
// Type: boolean
//

remap_pole_is_north = TRUE;

///////////// remap_persp_radius /////////////////////
//
// Radius of perspective point (km).
// This applies to VERT_PERSP.
// Type: double
//

remap_persp_radius = 35786;
```

```
///////////// remap_false_northing ////////////////////
//
// Remapped false northing correction.
// Occasionally, this is added to the Y coordinate so that all
//   coordinates are positive. Normally 0. As an alternative to
//   false_northing and false_easting, you can set the offset_latitude and
//   offset_longitude.
// Type: double
//

remap_false_northing = 0;

///////////// remap_false_easting /////////////////////
//
// Remapped false easting correction.
// Occasionally, this is added to the X coordinate so that all
//   coordinates are positive. Normally 0.
// Type: double
//

remap_false_easting = 0;

///////////// remap_set_offset_origin /////////////////
//
// Do you want to specify an offset origin using lat/lon instead of
//   false_northing and false_easting?.
// If true, set remap_offset_origin_latitude and
//   remap_offset_origin_longitude.
// Type: boolean
//

remap_set_offset_origin = FALSE;

///////////// remap_offset_origin_latitude ////////////
//
// Latitude of offset origin.
// See remap_set_offset_origin.
// Type: double
//

remap_offset_origin_latitude = 0;

///////////// remap_offset_origin_longitude ///////////
//
// Longitude of offset origin.
// See remap_set_offset_origin.
// Type: double
//

remap_offset_origin_longitude = 0;

//======================================================================
//
// ENCODING AND COMPRESSION CONVERSION.
//
//======================================================================

///////////// encoding_type ///////////////////////////
//
// Set encoding type.
//
// Type: enum
// Options:
//     ENCODING_ASIS
//     ENCODING_INT8
//     ENCODING_INT16
//     ENCODING_FLOAT32
//
```

```
encoding_type = ENCODING_ASIS;

///////////// compression_type ///////////////////////
//
// Set compression type.
// See <toolsa/compress> for details on the compression types.
//
// Type: enum
// Options:
//     COMPRESSION_ASIS
//     COMPRESSION_NONE
//     COMPRESSION_RLE
//     COMPRESSION_LZO
//     COMPRESSION_ZLIB
//     COMPRESSION_BZIP
//     COMPRESSION_GZIP
//     COMPRESSION_GZIP_VOL
//     COMPRESSION_TYPES_N
//

compression_type = COMPRESSION_GZIP;

///////////// force_scale_change /////////////////////
//
// Option to force a scaling change in the data.
// If this option is chosen, the data is read in as float data and then
//   is converted to the chosen output encoding type using the scaling
//   options specified below.
// NOTE: When using this option, if you set the encoding_type option to
//   ENCODING_ASIS, the output will use FLOAT32 encoding.
// Type: boolean
//

force_scale_change = FALSE;

///////////// scaling_type ///////////////////////////
//
// Set scaling type.
// This is only relevant when converting from float32 to int8 or int16
//   or if force_scale_change is set.
//
// Type: enum
// Options:
//     SCALING_ASIS
//     SCALING_NONE
//     SCALING_ROUNDED
//     SCALING_INTEGRAL
//     SCALING_DYNAMIC
//     SCALING_SPECIFIED
//

scaling_type = SCALING_ROUNDED;

///////////// scale //////////////////////////////////
//
// Input scaling scale.
// For SCALING_SPECIFIED only.
// Type: float
//

scale = 1;

///////////// bias ///////////////////////////////////
//
// Input scaling bias.
// For SCALING_SPECIFIED only.
// Type: float
//
```

```
bias = 0;

//======================================================================
//
// DECIMATION.
//
//======================================================================

///////////// decimate ////////////////////////////////
//
// Option to decimate in x,y.
// If true, each plane is decimated to force the number of grid points
//   to be less than 'decimate_max_nxy'.
// Type: boolean
//

decimate = FALSE;

///////////// decimate_max_nxy ////////////////////////
//
// Max number of xy grid points in decimation.
// See 'decimate'.
// Type: int
//

decimate_max_nxy = 1000000;

//======================================================================
//
// INVERT PLANES IN THE VERTICAL SENSE.
//
//======================================================================

///////////// invert_vertically ///////////////////////
//
// Invert the vertical levels in all fields.
// This inversion is applied after the remap, forced scale change,
//   overriding of V levels, and linear transformations.
// Type: boolean
//

invert_vertically = FALSE;

//======================================================================
//
// BYTE ORDERING.
//
//======================================================================

///////////// input_be ////////////////////////////////
//
// Are input files big-endian.
// Type: boolean
//

input_be = TRUE;

///////////// output_be ///////////////////////////////
//
// Are output files big-endian.
// Type: boolean
//

output_be = TRUE;

//======================================================================
//
// APPLY LINEAR TRANSFORM FUNCTION TO SELECTED FIELDS.
//
```

```
//======================================================================

///////////// apply_linear_transform //////////////////
//
// Option to apply a linear transform function to the data in selected
//   data fields.
// Field names and transform parameters are given in linear_transforms
//   parameter.
// Type: boolean
//

apply_linear_transform = FALSE;

///////////// linear_transforms ////////////////////////
//
// Array specifying the transform functions and the field names to which
//   they apply.
// The transform will only be applied to the specified fields. If a
//   field which is specified does not exist, a warning will be issued.
//
// Type: struct
//   typedef struct {
//      string field_name;
//      double scale;
//      double bias;
//   }
//
// 1D array - variable length.
//

linear_transforms = {
  {
    field_name = "DBZ",
    scale = 1,
    bias = 0
  }
};

//======================================================================
//
// WRITE USING EXTENDED PATHS.
//
// This will be overridden if the environment variable
//   MDV_WRITE_USING_EXTENDED_PATHS exists and is set to TRUE.
//
//======================================================================

///////////// write_using_extended_paths //////////////
//
// Option to write files with extended paths.
// If specified, this will override that specified by the client.
//   Default is FALSE.

// If set, paths will include a separate year subdirectory, and the file
//   name will include date and time.

// Non-forecast path:
//   dir/yyyy/yyyymmdd/yyyymmdd_hhmmss.mdv.

// Forecast path:
//   dir/yyyy/yyyymmdd/yyyymmdd_g_hhmmss_f_llllllll.mdv.
// Type: boolean
//

write_using_extended_paths = TRUE;

//======================================================================
//
// CONTROL OF CONVERSION TO NETCDF.
```

```
//
// The following parameters control conversion of MDV files to NetCDF
//   CF-compliant files.
//
//======================================================================

///////////// ncf_set_global_attributes ///////////////
//
// Option to set specify global attributes in the NCF file.
// The global attributes are 'institution', 'references' and 'comment'.
// Type: boolean
//

ncf_set_global_attributes = TRUE;

///////////// ncf_global_attributes ///////////////////
//
// Global attributes for netCDF file.
// These strings will be included as global attributes in the NetCDF
//   file. Other global attributes will be determined from the MDV
//   headers.
//
// Type: struct
//   typedef struct {
//      string institution;
//      string references;
//      string comment;
//   }
//
//

ncf_global_attributes = {
    institution = "SAWS",
    references = "SA Radar Merge Precip 1hr",
    comment = "Converted by MdvConvert"
};

///////////// ncf_transform_fields ////////////////////
//
// Option to tranform field names, units and values when converting MDV
//   to NCF.
// Type: boolean
//

ncf_transform_fields = FALSE;

///////////// ncf_field_transforms ////////////////////
//
// List of transforms. If mdv_field_name is found in the MDV data, these
//   other parameters will be used to set the field variable in the netCDF
//   file.
// See mdv2ncf_transform_fields.
//
// Type: struct
//   typedef struct {
//      string mdv_field_name;
//      string ncf_field_name;
//      string ncf_standard_name;
//      string ncf_long_name;
//      string ncf_units;
//      boolean do_linear_transform;
//      float linear_multiplier;
//      float linear_const;
//      data_pack_t packed_data_type;
//        Options:
//          DATA_PACK_FLOAT
//          DATA_PACK_SHORT
//          DATA_PACK_BYTE
//          DATA_PACK_ASIS
```

```
//   }
//
// 1D array - variable length.
//

ncf_field_transforms = {
  {
    mdv_field_name = "mdv_field_name",
    ncf_field_name = "ncf_field_name",
    ncf_standard_name = "ncf_standard_name",
    ncf_long_name = "ncf_long_name",
    ncf_units = "ncf_units",
    do_linear_transform = FALSE,
    linear_multiplier = 1,
    linear_const = 0,
    packed_data_type = DATA_PACK_ASIS
  }
};

///////////// ncf_compress_data //////////////////////
//
// Option to compress field data.
// Only applies to NETCDF4 and  NETCDF4_CLASSIC files.
// Type: boolean
//

ncf_compress_data = TRUE;

///////////// ncf_compression_level //////////////////
//
// Compression level from 1 to 9 with 9 being the greatest compression.
//   Default is 9.
// Only applies to NETCDF4 and  NETCDF4_CLASSIC files.
// Type: int
//

ncf_compression_level = 9;

///////////// ncf_filename_suffix ////////////////////
//
// Suffix of netCDF files.
// File extension is always .nc. File name will end with mdv.suffix.nc.
//   Set to the empty string for no suffix, in which case file name will
//   end with .mdv.nc.
// Type: string
//

ncf_filename_suffix = "";

///////////// ncf_file_format ////////////////////////
//
// NetCDF file format.
// netCDF classic format, netCDF 64-bit offset format, netCDF4 using
//   HDF5 format, netCDF4 using HDF5 format but only netCDF3 calls.
//
// Type: enum
// Options:
//     CLASSIC
//     NC64BIT
//     NETCDF4
//     NETCDF4_CLASSIC
//

ncf_file_format = NETCDF4;

///////////// ncf_polar_radar_file_type //////////////
//
// Output format for polar radar data.
//
```

```
// Type: enum
// Options:
//     FILE_TYPE_CF
//     FILE_TYPE_CF_RADIAL
//     FILE_TYPE_DORADE
//     FILE_TYPE_UF
//

ncf_polar_radar_file_type = FILE_TYPE_CF;

///////////// ncf_output_latlon_arrays ////////////////
//
// If true latitude and longitude arrays of each grid point are output.
// The CF convention requires that these arrays are present in the
//   netCDF file; however, the information is redundant since the lat and
//   lon arrays could be constructed using the other projection and grid
//   information required with a gridded data field.
// Type: boolean
//

ncf_output_latlon_arrays = TRUE;

///////////// ncf_output_mdv_attributes ///////////////
//
// Option to output non-CF compliant MDV attributes.
// If true, MDV attributes which are not CF compliant will be output.
//   This will facilitate the translation of the data back into MDV with
//   the minimal loss of information.
// Type: boolean
//

ncf_output_mdv_attributes = TRUE;

///////////// ncf_output_mdv_chunks ///////////////////
//
// Option to output non-CF compliant MDV chunks.
// If true, MDV chunks will be included as byte binary variables.
// Type: boolean
//

ncf_output_mdv_chunks = TRUE;
```

APPENDIX B

Table B.1

Method	Mathematical algorithm	Origin	Comments
Annual maxima-centered method (Asquith & Famiglietti, 2000)	*ARF* = $\frac{\int_0^R 2rS_{T(r)}\Delta r}{R^2}$ where: *ARF* = areal reduction factor, *A* = rainfall storm areas (km²), *R* = maximum radius of circular catchment or integration limit (km), *r* = radius of concentric circle within the catchment (km), and $S_{T(r)}$ = ratio between rainfall depth at a specific location, distance *r* from the point of the design storm and the annual maxima rainfall.	USA	• Method developed for the Austin, Dallas, and Houston regions, USA with a dense rainfall-monitoring network. • The Austin region (15 600 km²) had 108 daily rainfall stations, Dallas region (21 000 km²) had 103 daily rainfall stations and Houston region (35 800 km²) had 193 daily rainfall stations. • Several record lengths exceeded 80 years. • Method focuses on the analysis of the areal rainfall distribution to estimate ARFs for design storms. • ARFs decrease rapidly with increasing AEPs.
Rodriquez-Iturbe-Mejia method (Rodriquez-Iturbe & Mejia, 1974; cited by Svensson & Jones, 2010)	*ARF* = $\sqrt{E(\rho(d))}$ where: *ARF* = areal reduction factor, and *E* (ρ(*d*)) = expected correlation coefficient for the characteristic correlation distance.	Various	• Simple ARF estimation approach used in various areas. • Based on a spatial correlation structure using either an exponentially decaying function or a Bessel-type correlation structure. • Dependent on all observed rainfall data, *i.e.*, the primary data and not only the AMS. • 'Design storm' areal rainfall distributions are not included.

Method	Mathematical algorithm	Origin	Comments
Storm movement method (Bengtsson & Niemczynowicz, 1986)	$ARF = \frac{L_P}{L} = \frac{vT_d}{L}$ if $L_P < 0.5$ $ARF = 1 - \frac{0.25L}{L_P} = 1 - \left(\frac{0.25L}{vT_d}\right)$ if $L_P \geq 0.5$ where: *ARF* = areal reduction factor, *L* = catchment length (km), L_P = extension of block rain cell (km), T_d = storm duration (hours), and *v* = storm speed ($m.s^{-1}$).	Sweden	• Represents the relationship between rainfall movement and ARFs. • ARFs are based on the limited extension of rain cells, movement and spacing between rain cells and the effect of rain cells on each other. • ARFs were obtained from point rainfall hyetographs and storm speeds. • Relations were established between moving storm-derived ARFs and ARFs estimated by a dense rainfall-monitoring network. • ARFs proved to be constant in Norway.

Method	Mathematical algorithm	Origin	Comments
Bacchi-Ranzi method (Bacchi & Ranzi, 1996)	$ARF_{(A,\ Td,\ F)} = \dfrac{T_{A,Td}(F')}{T_A(F')}$ where: *ARF* = areal reduction factor, *A* = area under consideration (km²), *F'* = *F*-quantile of the corresponding probability distribution, T_d = duration within the space-time domain where the rainfall process can be assumed uniform (hours), and *T* = return period (years).	Italy	• Sixteen Constant Altitude Plan Position Indicator (CAPPI) maps were recorded and analysed from the C-band weather radar to be compared with the corresponding rainfall data from 17 rainfall stations. • Based on the analysis of the crossing properties of the spatial and temporal rainfall process. • High rainfall intensity processes were assumed to be Poisson distributed. • ARF expressed as the ratio of areal and point rainfall intensity values associated with the same duration and frequency. • ARFs are dependent on the return period and catchment area.

Method	Mathematical algorithm	Origin	Comments
Sivapalan-Blöschl method (Sivapalan & Blöschl, 1998)	$ARF\left[k^2\left(\frac{A}{\lambda^2}\right),T_d,T\right]=\frac{b(T_d)c(T_d)k^2F_2(k^{-2})-\frac{k^2}{F_1(k^{-2})}\ln\left[\ln\left(\frac{T}{T-1}\right)\right]}{b(T_d)c(T_d)-\ln\left[\ln\left(\frac{T}{T-1}\right)\right]}$ where: *ARF* = areal reduction factor, *A* = catchment area (km²), *b* = function of duration, where $b(T_d)=-0{,}05+0{,}25T_d^{\;0{,}49}$ *c* = function of duration, where $c(T_d)=0{,}2+20T_d^{\;-0{,}7}$ $F_1(k^{-2})$ = generic properties of the gamma distribution, $F_2(k^{-2})$ = generic properties of the gamma distribution, k^2 = rainfall correlation structure, *T* = return period (years), T_d = storm duration (hours), and λ = spatial correlation length (km).	Austria	• Based on a spatial correlation structure using both extreme value and/or parent distributions. • ARF values are dependent on the catchment area, storm duration (spatial correlation structure) and return period. • The ARF values are independent of the rainfall regime. • ARF values decrease with an increasing catchment area and return period. • Method is rather regarded as a 'geographically-centred' method as opposed to 'storm-centred'. • The final ARF expression is regarded as complex and not user-friendly.
Polar 55C method (Lombardo *et al.*, 2006)	$ARF_{(Td,\,T)}=\frac{i_A(T_d,T)}{i_{A=1}(T_d,T)}$ where: *ARF* = areal reduction factor, *A* = area under consideration (km²), *i* = rainfall intensity (mm.h^{-1}), *T* = return period (years), and T_d = storm duration (hours).	Italy	• The ARF values were estimated by using radar reflectivity maps collected with Polar 55C. • Rainfall intensities over the radar scanning region (allowing a single radar image to last for one minute) were estimated for durations (1, 5, 10, 60 and 120 minutes) and return periods (2, 10, 25 and 50 years) by using the Arithmetic mean and Thiessen polygon methods.

			• The radar rainfall estimates were integrated for heavy rainfall data over an area of 900 km². • Radar located 15 km south-east of Rome. • Study focussed of the influences of area, storm duration, intensity and return period. • The ARFs exceeded unity in small areas characterised by relative longer storm durations.

Method	Mathematical algorithm	Origin	Comments
Kim *et al.* method (Kim *et al*., 2019)	*ARF* $= \frac{\frac{\Delta A}{A_0}\sum R_i}{R_0}$ where: *ARF* = areal reduction factor, A_0 = area of an elliptical storm (m²), ΔA = area of radar cells (i.e. 1 km x 1 km), R_0 = maximum rainfall depth of an ellipse (mm), and R_i = rainfall volume of the *i*-th cell within the storm (mm).	South Korea	• Build a time series at every radar pixel for six (6) durations. • Used storm centres that contained specific time and location information. • Identified specific elliptical shapes (aspect ratio, orientation and sizes) with a maximizing rainfall volume of 13 areas. • Performed an additional filtering process to verify spatiotemporal proximity of storms. • Storms were reduced from 2218 905 to 54 758 as part of the filtering process.

APPENDIX C

Table C1: Results

	Date	Rainfall ArcGIS (mm)	Area of Storm (km^2)	Max Point Precipitation (mm)	Max Point Intensity (mm/h)	ARF ArcGIS (%)	Region (Inland/ Coastal)
Column Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1 hr	24/01/2016	27.54	1187.39	57.50	57.50	47.89%	Coastal
1 hr	06/02/2016	24.80	430.39	32.20	32.20	77.03%	Inland
1 hr	25/02/2016	38.49	262.51	62.50	62.50	61.58%	Costal
1 hr	09/03/2016	30.64	311.35	53.50	53.50	57.27%	Inland
1 hr	10/03/2016	27.50	33.58	37.00	37.00	74.32%	Coastal
1 hr	14/03/2016	26.42	775.32	38.60	38.60	68.44%	Inland
1 hr	17/03/2016	37.32	521.96	56.50	56.50	66.06%	Coastal
1 hr	22/10/2016	28.85	5744.67	70.00	70.00	41.21%	Coastal
1 hr	23/10/2016	28.05	4847.25	50.50	50.50	55.55%	Coastal
1 hr	26/10/2016	38.23	644.06	60.50	60.50	63.20%	Inland
1 hr	05/11/2016	24.04	39.68	30.00	30.00	80.13%	Coastal
1 hr	10/11/2016	37.44	9013.81	81.50	81.50	45.94%	Inland
1 hr	11/11/2016	32.48	7057.21	54.50	54.50	59.60%	Inland
1 hr	30/11/2016	32.37	802.79	48.50	48.50	66.74%	Coastal
1 hr	13/12/2016	28.85	521.96	47.60	47.60	60.85%	Inland
1 hr	19/12/2016	33.15	1654.42	60.00	60.00	55.24%	Coastal
1 hr	03/01/2017	44.42	345.42	81.00	81.00	54.84%	Inland
1 hr	04/01/2017	36.08	699.01	59.50	59.50	60.64%	Inland
1 hr	06/01/2017	24.74	436.50	33.00	33.00	74.96%	Inland
1 hr	07/01/2017	29.61	137.36	44.20	44.20	66.10%	Inland

	Date	Rainfall ArcGIS (mm)	Area of Storm (km^2)	Max Point Precipitation (mm)	Max Point Intensity (mm/h)	ARF ArcGIS (%)	Region (Inland/ Coastal)
Column Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1 hr	26/01/2017	34.38	381.55	69.50	69.50	49.47%	Inland
1 hr	27/01/2017	33.51	210.62	59.50	59.50	56.33%	Coastal
1 hr	30/01/2017	33.15	515.86	57.00	57.00	58.16%	Inland
1 hr	09/02/2017	29.09	67.15	38.60	38.60	75.37%	Inland
1 hr	20/02/2017	22.50	3.05	31.00	31.00	72.58%	Inland
1 hr	21/02/2017	23.50	30.52	31.40	31.40	74.84%	Inland
1 hr	06/04/2017	29.79	360.19	42.50	42.50	70.09%	Coastal
1 hr	30/09/2017	30.17	222.83	44.20	44.20	68.26%	Inland
1 hr	23/11/2017	27.76	351.03	39.60	39.60	70.10%	Inland
1 hr	27/11/2017	26.73	39.68	31.80	31.80	69.98%	Inland
1 hr	06/12/2017	32.35	100.73	65.50	65.50	49.39%	Inland
1 hr	29/12/2017	32.28	210.62	53.50	53.50	60.34%	Coastal
3 hr	08/03/2016	34.31	210.62	48.44	16.15	70.83%	Inland
3 hr	09/03/2016	35.00	567.75	55.52	18.51	63.04%	Inland
3 hr	16/03/2016	51.91	1251.49	130.71	43.57	39.72%	Coastal
3 hr	26/04/2016	47.54	778.37	97.60	32.53	48.71%	Inland
3 hr	26/04/2016	47.73	2521.30	101.76	33.92	46.91%	Inland
3 hr	14/05/2016	56.78	1031.72	114.15	38.05	49.74%	Inland
3 hr	26/07/2016	44.74	2734.97	142.71	47.57	31.35%	Coastal
3 hr	22/10/2016	41.52	1611.68	71.90	23.97	57.76%	Inland
3 hr	22/10/2016	40.05	6559.66	105.43	35.14	37.99%	Inland
3 hr	26/10/2016	53.02	613.54	98.60	32.87	53.77%	Inland
3 hr	11/11/2016	55.99	2960.85	127.02	42.34	44.08%	Inland
3 hr	20/11/2016	45.46	2869.28	86.10	28.70	52.80%	Inland

	Date	Rainfall ArcGIS (mm)	Area of Storm (km^2)	Max Point Precipitation (mm)	Max Point Intensity (mm/h)	ARF ArcGIS (%)	Region (Inland/ Coastal)
Column Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
3 hr	26/11/2016	41.78	210.62	90.70	30.23	46.06%	Inland
3 hr	30/11/2016	29.96	198.41	39.55	13.18	75.75%	Inland
3 hr	04/12/2016	48.42	115.99	75.86	25.29	63.83%	Inland
3 hr	07/12/2016	52.48	753.95	116.43	38.81	45.07%	Inland
3 hr	19/12/2016	39.47	744.79	68.17	22.72	57.90%	Inland
3 hr	03/01/2017	38.37	613.54	69.87	23.29	54.92%	Inland
3 hr	04/01/2017	63.09	1672.73	161.53	53.84	39.06%	Inland
3 hr	07/01/2017	33.35	1510.95	67.46	22.49	49.43%	Inland
3 hr	07/01/2017	39.00	680.69	115.94	38.65	33.64%	Inland
3 hr	20/01/2017	46.89	201.46	98.82	32.94	47.45%	Inland
3 hr	26/01/2017	49.38	366.29	100.53	33.51	49.11%	Inland
3 hr	30/01/2017	42.63	1514.00	147.37	49.12	28.93%	Inland
3 hr	21/02/2017	31.15	769.21	44.95	14.98	69.31%	Inland
3 hr	09/03/2017	40.70	2213.01	105.84	35.28	38.45%	Inland
3 hr	10/04/2017	39.77	424.29	79.11	26.37	50.27%	Coastal
3 hr	14/04/2017	33.28	393.76	52.52	17.51	63.35%	Inland
3 hr	12/05/2017	30.18	125.15	39.80	13.27	75.84%	Coastal
3 hr	13/05/2017	30.00	115.99	40.28	13.43	74.49%	Inland
3 hr	02/10/2017	30.56	189.25	42.26	14.09	72.32%	Inland
3 hr	23/11/2017	43.18	2933.38	87.47	29.16	49.37%	Inland
3 hr	28/11/2017	35.18	250.30	45.03	15.01	78.13%	Inland
3 hr	23/11/2014	27.85	95.63	33.55	11.18	83.01%	inland
3 hr	27/11/2014	33.04	1330.88	52.52	17.51	62.91%	inland
3 hr	22/12/2014	27.50	2.81	29.49	9.83	93.25%	inland

	Date	Rainfall ArcGIS (mm)	Area of Storm (km^2)	Max Point Precipitation (mm)	Max Point Intensity (mm/h)	ARF ArcGIS (%)	Region (Inland/ Coastal)
Column Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
3 hr	23/12/2014	37.23	135.00	59.52	19.84	62.55%	inland
3 hr	24/12/2014	48.81	317.25	94.76	31.59	51.51%	inland
3 hr	01/01/2015	39.43	116.44	78.73	26.24	50.08%	inland
3 hr	16/10/2016	32.50	1.69	33.92	11.31	95.82%	inland
3 hr	23/12/2015	43.16	29.81	80.12	26.71	53.87%	inland
3 hr	10/01/2016	53.99	2118.69	111.25	37.08	48.53%	inland
3 hr	02/03/2016	45.90	1384.88	165.90	55.30	27.66%	inland
3 hr	02/03/2016	44.88	340.88	101.79	33.93	44.09%	inland
3 hr	01/11/2016	45.94	988.88	78.40	26.13	58.60%	inland
3 hr	07/11/2016	78.73	428.63	190.81	63.60	41.26%	inland
3 hr	12/11/2016	34.00	273.38	54.90	18.30	61.94%	inland
3 hr	23/11/2016	57.92	873.56	210.02	70.01	27.58%	inland
3 hr	23/11/2016	62.30	1092.38	233.53	77.84	26.68%	inland
3 hr	09/12/2016	46.58	97.88	80.39	26.80	57.94%	inland
3 hr	28/12/2016	29.26	984.94	40.56	13.52	72.14%	inland
3 hr	04/01/2017	32.94	186.19	58.71	19.57	56.11%	inland
3 hr	09/03/2016	39.80	152.62	67.27	22.42	59.16%	inland
3 hr	19/10/2016	42.31	79.36	60.85	20.28	69.53%	inland
3 hr	01/11/2016	40.81	189.25	63.70	21.23	64.06%	inland
3 hr	07/11/2016	55.24	256.40	98.73	32.91	55.95%	inland
3 hr	09/12/2016	43.60	918.78	77.03	25.68	56.59%	inland
3 hr	04/01/2016	62.55	1700.20	161.53	53.84	38.73%	inland
3 hr	11/01/2017	40.08	94.63	62.76	20.92	63.87%	inland
3 hr	02/03/2017	40.09	164.83	67.30	22.43	59.57%	inland

	Date	Rainfall ArcGIS (mm)	Area of Storm (km^2)	Max Point Precipitation (mm)	Max Point Intensity (mm/h)	ARF ArcGIS (%)	Region (Inland/ Coastal)
Column Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
3 hr	21/10/2017	45.92	1111.08	88.82	29.61	51.70%	inland
3 hr	25/11/2017	32.29	1984.08	55.92	18.64	57.74%	inland
3 hr	02/12/2017	35.74	1532.32	66.91	22.30	53.41%	inland
3 hr	15/01/2018	31.67	146.52	44.21	14.74	71.64%	inland
3 hr	22/03/2018	32.57	4996.82	61.45	20.48	53.00%	inland
3 hr	13/10/2018	37.23	744.79	61.22	20.41	60.81%	inland
3 hr	20/11/2018	29.77	2182.48	39.20	13.07	75.93%	inland
3 hr	03/01/2019	33.93	427.34	58.97	19.66	57.54%	inland
3 hr	15/01/2019	36.81	1333.91	67.92	22.64	54.20%	inland
3 hr	28/01/2019	31.46	601.33	44.43	14.81	70.81%	inland
3 hr	10/10/2019	33.28	393.76	59.08	19.69	56.33%	inland
24 hr	13/03/2016	54.48	1709.36	101.00	4.21	53.94%	Inland
24 hr	15/03/2016	50.23	235.04	63.50	2.65	79.10%	Inland
24 hr	12/06/2016	48.15	140.41	71.00	2.96	67.82%	Coastal
24 hr	25/07/2016	51.01	2359.53	80.00	3.33	63.77%	Coastal
24 hr	18/10/2016	47.26	64.10	59.00	2.46	80.10%	Inland
24 hr	20/10/2016	54.29	1254.55	107.50	4.48	50.50%	Coastal
24 hr	25/10/2016	61.98	592.17	96.00	4.00	64.57%	Coastal
24 hr	01/11/2016	56.51	509.76	100.00	4.17	56.51%	Inland
24 hr	04/11/2016	61.63	934.04	126.00	5.25	48.92%	Coastal
24 hr	09/11/2016	64.24	271.67	118.00	4.92	54.44%	Inland
24 hr	29/11/2016	48.31	262.51	63.00	2.63	76.69%	Inland
24 hr	04/12/2016	47.00	183.15	70.00	2.92	67.14%	Coastal
24 hr	06/12/2016	56.81	155.67	86.00	3.58	71.92%	Inland

	Date	Rainfall ArcGIS (mm)	Area of Storm (km^2)	Max Point Precipitation (mm)	Max Point Intensity (mm/h)	ARF ArcGIS (%)	Region (Inland/ Coastal)
Column Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
24 hr	09/12/2016	72.50	634.90	158.00	6.58	47.08%	Inland
24 hr	18/12/2016	57.23	167.88	100.00	4.17	57.23%	Coastal
24 hr	27/12/2016	60.08	6849.64	117.50	4.90	51.14%	Coastal
24 hr	03/01/2017	55.98	741.74	96.00	4.00	58.31%	Coastal
24 hr	06/01/2017	48.66	2499.94	68.50	2.85	72.09%	Inland
24 hr	12/01/2017	65.72	412.08	123.00	5.13	53.43%	Inland
24 hr	25/01/2017	73.43	164.83	155.00	6.46	47.37%	Coastal
24 hr	29/01/2017	56.83	91.57	73.00	3.04	77.85%	Inland
24 hr	09/02/2017	69.46	366.29	149.00	6.21	46.62%	Inland
24 hr	19/02/2017	85.51	436.50	147.00	6.13	58.17%	Coastal
24 hr	24/02/2017	51.85	1407.17	64.00	2.67	81.01%	Inland
24 hr	09/04/2017	62.38	650.17	116.00	4.83	53.78%	Inland
24 hr	12/05/2017	57.48	8519.32	141.00	5.88	40.77%	Coastal
24 hr	14/05/2017	54.25	479.23	83.00	3.46	65.36%	Inland
24 hr	30/09/2017	50.38	726.48	80.00	3.33	62.97%	Inland
24 hr	02/10/2017	60.83	155.67	101.00	4.21	60.23%	Inland
24 hr	04/10/2017	60.83	155.67	77.50	3.23	75.05%	Coastal
24 hr	26/11/2017	49.42	2014.60	77.00	3.21	64.18%	Inland
24 hr	29/12/2017	54.87	637.96	80.00	3.33	68.59%	Inland
24 hr	11/01/2014	65.10	180.56	101.64	4.23	64.05%	inland
24 hr	27/11/2014	52.90	4405.50	96.24	4.01	54.97%	inland
24 hr	23/12/2014	65.26	140.63	110.71	4.61	58.95%	inland
24 hr	01/01/2015	53.95	99.00	91.67	3.82	58.85%	inland
24 hr	16/10/2015	66.85	7537.50	148.51	6.19	45.01%	inland

	Date	Rainfall ArcGIS (mm)	Area of Storm (km^2)	Max Point Precipitation (mm)	Max Point Intensity (mm/h)	ARF ArcGIS (%)	Region (Inland/ Coastal)
Column Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**
24 hr	14/12/2015	70.87	892.69	128.56	5.36	55.13%	inland
24 hr	15/12/2015	54.11	352.13	97.99	4.08	55.21%	inland
24 hr	20/12/2015	57.40	112.50	82.97	3.46	69.18%	inland
24 hr	22/12/2015	59.94	23.06	85.31	3.55	70.26%	inland
24 hr	10/01/2016	59.38	514.69	121.06	5.04	49.05%	inland
24 hr	09/03/2016	55.33	1510.31	92.99	3.87	59.50%	inland
24 hr	01/11/2016	83.15	3392.44	173.05	7.21	48.05%	inland
24 hr	13/11/2016	52.21	135.00	67.91	2.83	76.88%	inland
24 hr	24/11/2016	57.14	172.69	101.09	4.21	56.53%	inland
24 hr	09/12/2016	64.25	736.88	151.45	6.31	42.42%	inland
24 hr	13/12/2016	83.86	385.31	175.09	7.30	47.89%	inland
24 hr	21/12/2016	50.68	1542.94	69.63	2.90	72.79%	inland
24 hr	01/01/2017	53.60	325.69	90.41	3.77	59.28%	inland
24 hr	13/01/2017	58.02	1887.19	126.75	5.28	45.78%	inland
24 hr	30/01/2017	63.47	299.81	148.91	6.20	42.62%	inland

Table C.2: Comparison of Radar Derived ARFs with Smaller Catchment ARFs

	Duration	ArcGIS area	Max Point Precipitation (mm)	ARCGIS Point Intensity	ARF ArcGIS (%)	Van Wyk ARF	Op Ten Noort
Column Number ➔	1	2	3	4	5	6	7
	3	1.69	33.92	11.31	95.82%	99.95%	99.87%
	3	2.81	29.49	9.83	93.25%	99.93%	99.81%
	1	3.05	31	31.00	72.58%	99.76%	99.36%
	24	23.06	85.31	3.55	70.26%	99.68%	99.44%
	3	29.81	80.12	26.71	53.87%	97.06%	94.73%
	1	30.52	31	31.40	74.84%	96.49%	93.69%
	1	33.58	37	37.00	74.32%	95.54%	91.90%
	1	39.68	30	30.00	80.13%	95.81%	92.22%
	1	39.68	32	31.80	69.98%	95.56%	91.78%
	24	64.10	59	2.46	80.10%	99.50%	98.93%
	1	67.15	39	38.60	75.37%	92.25%	83.84%
	3	79.36	60.85	20.28	69.53%	95.38%	89.63%
	24	91.57	73	3.04	77.85%	99.22%	98.12%
	3	94.63	62.76	20.92	63.87%	94.63%	87.41%
	3	95.63	33.55	11.18	83.01%	97.07%	92.99%
	3	97.88	80.39	26.80	57.94%	93.02%	83.67%
	24	99.00	91.67	3.82	58.85%	98.96%	97.46%
	1	100.73	66	65.50	49.39%	83.49%	63.85%
	24	112.50	82.97	3.46	69.18%	98.98%	97.39%
	3	115.99	76	25.29	63.83%	92.66%	81.92%
	3	115.99	40	13.43	74.49%	96.03%	89.95%

Duration	ArcGIS area	Max Point Precipitation (mm)	ARCGIS Point Intensity	ARF ArcGIS (%)	Van Wyk ARF	Op Ten Noort
1	**2**	**3**	**4**	**5**	**6**	**7**
3	116.44	78.73	26.24	50.08%	92.38%	81.24%
3	125.15	40	13.27	75.84%	95.88%	89.32%
3	135.00	59.52	19.84	62.55%	93.59%	83.35%
24	135.00	67.91	2.83	76.88%	99.06%	97.44%
1	137.36	44	44.20	66.10%	86.16%	66.18%
24	140.41	71	2.96	67.82%	98.99%	97.21%
24	140.63	110.71	4.61	58.95%	98.43%	95.68%
3	146.52	44.21	14.74	71.64%	94.92%	86.35%
3	152.62	67.27	22.42	59.16%	92.16%	79.24%
24	155.67	86	3.58	71.92%	98.68%	96.28%
24	155.67	101	4.21	60.23%	98.46%	95.64%
24	155.67	78	3.23	75.05%	98.81%	96.64%
24	164.83	155	6.46	47.37%	97.54%	93.02%
3	164.83	67.30	22.43	59.57%	91.72%	77.77%
24	167.88	100	4.17	57.23%	98.38%	95.35%
24	172.69	101.09	4.21	56.53%	98.33%	95.17%
24	180.56	101.64	4.23	64.05%	98.27%	94.93%
24	183.15	70	2.92	67.14%	98.79%	96.43%
3	186.19	58.71	19.57	56.11%	92.07%	78.06%
3	189.25	42	14.09	72.32%	94.15%	83.42%
3	189.25	63.70	21.23	64.06%	91.32%	76.09%
3	198.41	40	13.18	75.75%	94.32%	83.70%
3	201.46	99	32.94	47.45%	86.29%	63.68%

Duration	ArcGIS area	Max Point Precipitation (mm)	ARCGIS Point Intensity	ARF ArcGIS (%)	Van Wyk ARF	Op Ten Noort
1	**2**	**3**	**4**	**5**	**6**	**7**
1	210.62	54	53.50	60.34%	78.32%	46.48%
3	210.62	91	30.23	46.06%	87.10%	64.86%
1	210.62	60	59.50	56.33%	76.20%	42.65%
3	210.62	48	16.15	70.83%	92.88%	79.35%
1	222.83	44	44.20	68.26%	81.27%	51.18%
24	235.04	64	2.65	79.10%	98.73%	95.86%
3	250.30	45	15.01	78.13%	92.80%	77.45%
3	256.40	98.73	32.91	55.95%	84.73%	56.34%
24	262.51	63	2.63	76.69%	98.67%	95.42%
1	262.51	63	62.50	61.58%	72.77%	32.77%
24	271.67	118	4.92	54.44%	97.47%	91.32%
3	273.38	54.90	18.30	61.94%	90.89%	71.17%
24	299.81	148.91	6.20	42.62%	96.65%	88.12%
1	311.35	54	53.50	57.27%	74.14%	32.22%
3	317.25	94.76	31.59	51.51%	83.61%	50.59%
24	325.69	90.41	3.77	59.28%	97.85%	92.00%
3	340.88	101.79	33.93	44.09%	81.82%	45.54%
1	345.42	81	81.00	54.84%	61.80%	14.92%
1	351.03	40	39.60	70.10%	78.79%	38.86%
24	352.13	97.99	4.08	55.21%	97.56%	90.69%
1	360.19	43	42.50	70.09%	77.11%	35.31%
24	366.29	149	6.21	46.62%	96.23%	85.67%
3	366.29	101	33.51	49.11%	81.28%	43.40%

Duration	ArcGIS area	Max Point Precipitation (mm)	ARCGIS Point Intensity	ARF ArcGIS (%)	Van Wyk ARF	Op Ten Noort
1	**2**	**3**	**4**	**5**	**6**	**7**
1	381.55	70	69.50	49.47%	64.36%	16.48%
24	385.31	175.09	7.30	47.89%	95.44%	82.60%
3	393.76	59.08	19.69	56.33%	88.02%	59.02%
3	393.76	53	17.51	63.35%	89.28%	62.58%
24	412.08	123	5.13	53.43%	96.66%	86.62%
3	424.29	79	26.37	50.27%	83.77%	46.73%
3	427.34	58.97	19.66	57.54%	87.59%	56.48%
3	428.63	190.81	63.60	41.26%	65.15%	15.66%
1	430.39	32	32.20	77.03%	80.45%	38.97%
24	436.50	147	6.13	58.17%	95.92%	83.38%
1	436.50	33	33.00	74.96%	79.90%	37.55%
24	479.23	83	3.46	65.36%	97.57%	89.34%
24	509.76	100	4.17	56.51%	96.99%	86.55%
24	514.69	121.06	5.04	49.05%	96.35%	83.82%
1	515.86	57	57.00	58.16%	65.80%	13.54%
1	521.96	48	47.60	60.85%	70.33%	18.46%
1	521.96	57	56.50	66.06%	65.88%	13.46%
3	567.75	56	18.51	63.04%	86.64%	48.95%
24	592.17	96	4.00	64.57%	96.87%	85.12%
3	601.33	44.43	14.81	70.81%	88.83%	54.58%
3	613.54	70	23.29	54.92%	82.85%	37.85%
3	613.54	99	32.87	53.77%	76.70%	25.38%
24	634.90	158	6.58	47.08%	94.71%	75.26%

Duration	ArcGIS area	Max Point Precipitation (mm)	ARCGIS Point Intensity	ARF ArcGIS (%)	Van Wyk ARF	Op Ten Noort
1	**2**	**3**	**4**	**5**	**6**	**7**
24	637.96	80	3.33	68.59%	97.28%	86.54%
1	644.06	61	60.50	63.20%	60.65%	7.07%
24	650.17	116	4.83	53.78%	96.04%	80.76%
3	680.69	116	38.65	33.64%	71.83%	16.72%
1	699.01	60	59.50	60.64%	59.66%	5.91%
24	726.48	80	3.33	62.97%	97.07%	84.82%
24	736.88	151.45	6.31	42.42%	94.48%	72.89%
24	741.74	96	4.00	58.31%	96.45%	81.73%
3	744.79	61.22	20.41	60.81%	83.13%	35.57%
3	744.79	68	22.72	57.90%	81.41%	31.64%
3	753.95	116	38.81	45.07%	70.21%	13.67%
3	769.21	45	14.98	69.31%	87.08%	45.67%
1	775.32	39	38.60	68.44%	69.90%	13.07%
3	778.37	98	32.53	48.71%	73.89%	17.87%

Table C.3: Comparison of Radar Derived ARFs with Larger Catchment ARFs

Duration	ArcGIS area	Max Point Precipitation (mm)	ARF ArcGIS (%)	ARF Drainage	Op Ten Noort	Difference Drainage	Difference OTN
1	2	3	4	5	6	7	8
1	802.79	49	66.74%	51.89%	74.11%	-14.85%	7.37%
1	1187.39	58	47.89%	50.30%	70.58%	2.41%	22.69%
1	1654.42	60	55.24%	48.37%	67.60%	-6.87%	12.36%
1	4847.25	51	55.55%	35.18%	57.92%	-20.37%	2.37%
1	5744.67	70	41.21%	31.47%	56.40%	-9.74%	15.18%
1	7057.21	55	59.60%	26.05%	54.54%	-33.55%	-5.05%
1	9013.81	82	45.94%	17.96%	52.34%	-27.98%	6.40%
3	1031.72	114	49.74%	57.73%	81.27%	7.99%	31.53%
3	873.56	210.02	27.58%	58.38%	82.64%	30.80%	55.06%
3	918.78	77.03	56.59%	58.20%	82.23%	1.60%	25.63%
3	984.94	40.56	72.14%	57.93%	81.65%	-14.21%	9.52%
3	988.88	78.40	58.60%	57.91%	81.62%	-0.69%	23.02%
3	1092.38	233.53	26.68%	57.48%	80.79%	30.81%	54.12%
3	1111.08	88.82	51.70%	57.41%	80.65%	5.71%	28.95%
3	1330.88	52.52	62.91%	56.51%	79.16%	-6.40%	16.25%
3	1333.91	67.92	54.20%	56.49%	79.14%	2.30%	24.94%
3	1384.88	165.90	27.66%	56.29%	78.83%	28.62%	51.16%
3	1532.32	66.91	53.41%	55.68%	77.99%	2.27%	24.58%
3	1700.20	161.53	38.73%	54.99%	77.12%	16.27%	38.40%
3	1984.08	55.92	57.74%	53.83%	75.84%	-3.91%	18.09%
3	2118.69	111.25	48.53%	53.28%	75.29%	4.74%	26.76%

Column Number →

Duration	ArcGIS area	Max Point Precipitation (mm)	ARF ArcGIS (%)	ARF Drainage	Op Ten Noort	Difference Drainage	Difference OTN
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
3	2182.48	39.20	75.93%	53.02%	75.04%	-22.92%	-0.89%
3	4996.82	61.45	53.00%	41.48%	68.08%	-11.52%	15.08%
3	1251.49	131	39.72%	56.83%	79.67%	17.11%	39.95%
3	1510.95	67	49.43%	55.77%	78.10%	6.33%	28.67%
3	1514.00	147	28.93%	55.76%	78.09%	26.83%	49.16%
3	1611.68	72	57.76%	55.36%	77.57%	-2.40%	19.81%
3	1672.73	162	39.06%	55.11%	77.26%	16.05%	38.20%
3	2213.01	106	38.45%	52.89%	74.93%	14.44%	36.47%
3	2521.30	102	17.69%	51.63%	73.84%	33.94%	56.15%
3	2734.97	143	31.35%	50.75%	73.15%	19.40%	41.81%
3	2869.28	86	52.80%	50.20%	72.75%	-2.60%	19.95%
3	2933.38	87	49.37%	49.94%	72.57%	0.57%	23.20%
3	2960.85	127	44.08%	49.83%	72.49%	5.75%	28.41%
3	6559.66	105	37.99%	35.07%	65.76%	-2.91%	27.78%
24	934.04	126	48.92%	75.05%	103.19%	26.13%	54.28%
24	1254.55	108	50.50%	73.81%	101.45%	23.31%	50.95%
24	892.69	128.56	55.13%	75.21%	103.46%	20.08%	48.34%
24	1510.31	92.99	59.50%	72.82%	100.36%	13.32%	40.86%
24	1542.94	69.63	72.79%	72.70%	100.23%	-0.09%	27.44%
24	1887.19	126.75	45.78%	71.37%	99.04%	25.59%	53.27%
24	3392.44	173.05	48.05%	65.55%	95.57%	17.50%	47.52%
24	4405.50	96.24	54.97%	61.63%	94.01%	6.66%	39.04%
24	7537.50	148.51	45.01%	49.53%	90.74%	4.52%	45.73%
24	1407.17	64	81.01%	73.22%	100.78%	-7.79%	19.76%

Duration	ArcGIS area	Max Point Precipitation (mm)	ARF ArcGIS (%)	ARF Drainage	Op Ten Noort	Difference Drainage	Difference OTN
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
24	1709.36	101	53.94%	72.05%	99.63%	18.11%	45.69%
24	2014.60	77	64.18%	70.87%	98.66%	6.70%	34.48%
24	2359.53	80	63.77%	69.54%	97.73%	5.78%	33.96%
24	2499.94	69	72.09%	69.00%	97.38%	-3.09%	25.29%
24	6849.64	118	51.14%	52.18%	91.33%	1.05%	40.20%
24	8519.32	141	40.77%	45.73%	89.98%	4.96%	49.22%

Table C.4: Statistical Results

Durations	X variable	Y variable	R^2 value	r-value	p-value
3hr	Storm Area	Radar Derived ARF	0.3983	-0.6311	3.64E-09
24 hr	Storm Area	Radar Derived ARF	0.1839	-0.4289	0.001513
All duration	Storm Area	Radar Derived ARF	0.3157	-0.5618	2.83E-14
3hr	Maximum Point Intensity	Radar Derived ARF	0.8737	-0.9347	1.03E-32
24 hr	Maximum Point Intensity	Radar Derived ARF	0.8341	-0.9133	3.82E-21
All duration	Maximum Point Intensity	Radar Derived ARF	0.1098	-0.3314	2.54E-05
All duration	Maximum Point Precipitation	Radar Derived ARF	0.6255	-0.7909	1.88E-34
	Radar Derived ARF	Van Wyk ARF	0.1105	0.3324	0.000567
	Radar Derived ARF	Op ten Noort Equation 2 ARF	0.1362	0.3691	0.000116
	Van Wyk ARF	Op ten Noort Equation 2 ARF	0.9475	0.9734	4.38E-67
	Radar Derived ARF	Wiederhold	0.0772	0.2779	0.048351
	Radar Derived ARF	Op ten Noort Equation 3 ARF	0.0780	0.2792	0.047247
	Wiederhold	Op ten Noort Equation 3 ARF	0.8475	0.9206	1.21E-21