Atmospheric thermodynamics and circulation associated with heavy rainfall over the Gauteng Province, South Africa

by

Liesl Letitia Dyson

Submitted in partial fulfilment of the requirement for the degree of

Doctor of Philosophy (Meteorology)

In the Faculty of Natural & Agricultural Sciences

University of Pretoria

May 2013

Atmospheric thermodynamics and circulation associated with heavy rainfall over the Gauteng Province, South Africa

Student: Liesl Letitia Dyson

Supervisor: Prof. Johan van Heerden

Co-supervisor: Prof. Paul Sumner

Department: Geography Geoinformatics and Meteorology

Degree: Doctor of Philosophy (Meteorology)

ABSTRACT

The primary focus of this thesis is to describe the prevailing atmospheric conditions when heavy rainfall occurs over the Gauteng Province in South Africa. This thesis first describes the characteristics of daily heavy rainfall over Gauteng by defining different heavy rainfall classes and considering the seasonal distribution of these events. Late summer (January, February and March) has considerably more heavy rainfall days than early summer. The change of the character of the atmosphere as the summer season progresses is highlighted by the investigation into the monthly average synoptic circulation patterns when heavy rainfall occurs. The weather systems change from extra-tropical in the first few months of the summer rainfall season to tropical in February months. It is also shown how cyclonic vorticity advection occurs in the upper troposphere whenever heavy rainfall occurs, irrespective of the time of the season. A deep layer of horizontal wind convergence is also present when heavy rainfall occurs and this is replaced by horizontal wind divergence above that. A monthly climatology of sounding-derived parameters associated with heavy rainfall is constructed and it is again apparent how the atmosphere changes from one where conditional instability dominates the production of heavy rainfall in early summer to one where convective instability plays a dominant role in late summer. Twelve sounding-derived variables are identified to describe the thermodynamical profile of the atmosphere when heavy rainfall occurs over Gauteng. They include variables not previously used such as the Elevated *K*-Index and the meridionial wind component near the surface. Self-organizing maps are used to create a climatology of the vertical profile of the

atmosphere during heavy rainfall and this methods captures the changes to the atmospheric state during the progression of the summer season. Favourable sounding-derived parameters and circulation criteria are combined in a self-organizing map to predict daily rainfall frequencies. This method produces encouraging results and methods should be explored to create probabilistic daily rainfall forecast for Gauteng in an operational environment.

Acknowledgements

"I believe in a magnificent God."

I would like to extend my sincerest appreciation to

- *Professors Johan van Heerden* and *Paul Sumner* for their support, encouragement, guidance and expert advice during the course of this study. *Prof van Heerden* has mentored me through my entire career and I truly appreciate his patience and support.
- The Head *Professor Hannes Rautenbach* and all the staff in the Department of Geography Geoinformatics and Meteorology. I would especially like to acknowledge my second floor neighbour *Dr Dan Darkey* for listening to all my stories and *Ms Nerhene Davis* for sharing this experience. A special word of thanks goes to *Ms Ingrid Booysen* for help with the graphics and general interest in my progress as well as *Ms Corné van Aardt* and *Mr André Daniels* for all the logistical help while I was on study leave.
- The management of *Faculty of Natural and Agricultural Science* for granting of study leave to finalize this thesis.
- The *South African Weather Service* and especially *Ms Colleen de Villiers* for the friendly and efficient supply of rainfall data.
- The *Water Research Commission* for supporting research into heavy rainfall over Gauteng.
- The external examiners whose expert comments helped to improve this thesis.
- My family for their backing, encouragement and interest and my dad *Dr Squire Benjamin Dyson* who would have been delighted.
- My meteorology friends and especially to *Ms Christien Engelbrecht* for the many coffees and discussions about the weather and intricacies of SOMs.
- A mentor, friend and inspiration *Mr Phillip Strydom* who has kept me on my meteorological toes for more than 20 years.

Many meteorological students have a tale about a Geography school teacher who inspired them to become meteorologist. In my case it was *Mrs Bezuidenthout* at the Hoërskool Brandwag in Benoni in the 1980's. Her passion for Geography lighted a fire in me, she was slightly eccentric and I responded to it, she cared and I appreciated it. *You changed my life. Thank you*.

DECLARATION

I, L.L. Dyson declare that the thesis, which I hereby submit for the degree PhD in Meteorology at the University of Pretoria, is my own work and has not been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE:.....

DATE:.....

Contents

List of Symbols and Abbreviations

AGL	Above ground level
AHR	Area-averaged heavy rainfall
a. m. s. l.	Above mean sea level
AOH	Atlantic Ocean High
CAPE	Mean Layer Convective Available Potential Energy
COL	Cut-off Low
EKI	Elevated *K*-Index
ETTI	Elevated Total Totals Index
FL	Freezing level height
ITCZ	Inter Tropical Convergence Zone
IOH	Indian Ocean High
KI	*K* -Index
M	Average mixing ratio
MCS	Mesoscale Convective System
MWS	Magnitude of wind shear
NCEP	National Centre for Environmental Prediction
NWP	Numerical Weather Prediction
p	Pressure
PFC	Pressure of Free Convection
PW	Precipitable water
q	Mixing ratio
QC	Quality Control

RMSE	Root mean square error
SAWS	South African Weather Service
SHR	Single station heavy rainfall
SI	Showalter index
SOM	Self-organizing map
T	Temperature
T_{53}	Average 500-300 hPa temperatures
TC	Tropical Cyclone
TH	Geopotential thickness
TL	Temperature lapse rate
TTI	Total Totals Index
TTT	Tropical temperate trough
u	Zonal wind speed
v	Meridionial wind speed
VTD	Vertical temperature difference
W64	Average wind speed in the 600-400 hPa pressure levels
WV86	Average meridionial wind speed in the 800-600 hPa pressure levels
WCD	Mean Layer warm cloud depth
WS	Bulk wind shear
Z	Geopotential height
Θe	Equivalent potential temperature
Θe_{ave}	Average tropospheric equivalent potential temperature
ΔΘe	Equivalent Potential Temperature lapse rate

List of Figures

List of Tables

1 INTRODUCTION

1.1 Background

Predicting rainfall accurately remains one of the major challenges for meteorologists. Heavy rainfall may lead to floods which in turn may lead to loss of life and damage to infrastructure. An absence of rainfall can also be quite devastating when droughts occur which negatively impact on the economy and environment. Tools available to the weather forecaster are at present still unable to provide accurately predicted rainfall amounts and pinpointing the exact area where the heavy rainfall will be concentrated is very difficult. This is so in spite of all the recent advances in Numerical Weather Prediction (NWP) models (Landman et al., 2012)

Rainfall over southern Africa varies considerably from year to year (Kruger, 2004). Below normal rainfall and droughts occur frequently. From time to time heavy rainfall and floods occur over southern Africa and are often associated with large financial losses, damage to infrastructure and loss of life. SAWB (1991), cited in Alexander and van Heerden (1991), lists 184 noteworthy flood events during the years from 1911 to 1988. Close to 30 % of these flood events occurred over Kwazulu-Natal and 42 % over the four northern provinces and the central interior of South Africa. Between 1990 and 2012 several remarkable flood events occurred over South Africa. A few examples include the floods in February 1996 over the central interior (De Coning et al., 1998) and the widespread heavy rainfall over the northern provinces of South Africa in February 2000 (Dyson and van Heerden, 2001). During the summer rainfall seasons (October-March) of 2009/2010 and 2010/2011 above normal rainfall occurred over much of the summer rainfall area (DWA, 2012). The summer of 2009/2010 was particularly wet as floods occurred in 7 of the 9 provinces in South Africa forcing the South African government to declare 33 disaster zones (Guardian, 2011). In the winter of 2012 heavy rainfall and flooding occurred over the Western and Eastern Cape Provinces as a progression of cold fronts brought an anomalously cold and wet winter over these parts (News24, 2012; DWA, 2012).

This focus of this thesis is heavy rainfall during summer (October to March). Taljaard (1996) identified a few synoptic-scale weather systems over South Africa which may lead to heavy rainfall. These weather systems can broadly be classed into tropical weather systems which occur in the late summer and extra-tropical weather systems which occur in early summer. An important mechanism in the production of rainfall over South Africa is the transport of moist tropical air originating at the Inter Tropical Convergence Zone (ITCZ) over Angola and Zambia to South Africa. This transport of moisture often takes place by means of the tropical temperate trough (TTT) which is clearly visible on

satellite imagery as long cloud bands linking tropical convective systems to the mid-latitude (temperate) frontal systems. Hart et al. (2010) described three heavy rainfall events over South Africa associated with the TTT, all three events occurred in December and January months. Many of the heavy rainfall events in January 2010 were caused by a tropical weather system moving slowly from over Botswana to the eastern parts of South Africa. The heavy rainfall associated with tropical weather systems often take place over large areas and lasts for days, resulting in widespread flooding and disruption of infrastructure and even loss of life (Dyson and van Heerden 2002). It occurs predominately from December to February months. During spring and autumn cut-off low (COL) pressure systems may be responsible for widespread and heavy rainfall over South Africa (Taljaard, 1985).

Heavy rainfall is not always associated with large synoptic-scale weather systems as so called mesoscale weather systems may also be responsible for extreme rainfall events. Synoptic-scale weather systems such as fronts, highs, and lows are associated with wavelengths in the order of 1000 km (Holton, 1992) and they can persist for days to weeks. Mesoscale weather systems range from near synoptic-scales (mesoscale-alpha) down to individual cloud cells with dimensions of 1–20 km and life spans less than one hour (mesoscale-gamma) (Orlanski, 1975). On the synoptic-scale the hydrostatic balance holds and vertical motion are in the order of cm s^{-1} with negligible vertical acceleration of air. In mesoscale weather systems, this assumption cannot be made as the vertical velocity may reach values in the order of 10 ms^{-1} (Holton, 1992). Hydrostatic NWP models are therefore not capable of explicitly resolving convective weather systems and these models employ numerical parameterization schemes to deal with convection (Pielke, 2002). An example of a mesoscale weather system is the Mesoscale Convective Complex (MCC) which occurs predominantly from November to February over South Africa but with the highest frequency in November and December months (Blamey and Reason, 2012). Mesoscale weather systems do not necessarily cause heavy rainfall but when it does occur the heavy rainfall may be of a short duration (but intense) and is often associated with strong winds and hail (Viviers and Chapman, 2008; Ngubo et al., 2008).

Rainfall over the summer rainfall area of South Africa is essentially convective in nature. This fact has been highlighted by a 6-year lightning climatology over South Africa (Gijben, 2012). Most of the summer rainfall area of South Africa receives on average between 5-10 lightning ground strokes per square kilometre per year. Over the escarpment of Mpumalanga there are 20-50 lighting strokes per annum in a square kilometre while the Witwatersrand in the southern parts of Gauteng receive 10-15 ground strokes per square kilometre per year. It is therefore important to note that

irrespective of the type of weather system responsible for rainfall over the summer rainfall area, convection and the accompanying atmospheric instability plays an important role in its development.

This thesis focuses on the Gauteng Province (here after Gauteng) of South Africa (Fig. 1-1). Gauteng falls into the summer rainfall area of South Africa and is situated over the northeastern parts of the country on the interior plateau. The Witwatersrand is located in south Gauteng (approximately 1 700 m a. m. s. l.) and receives roughly 700 mm of rain per annum. The northern parts of the province receive about 600 mm rainfall per annum and north of the Magaliesberg the height a. m. s. l. decreases to around 1 100 m. Gauteng is cool with the average annual maximum temperatures of 22 °C in the south but 25 °C in the north. There are about 100 days with rain in Johannesburg (South) and 85 in Pretoria (North) (Kruger, 2004). Even though there are some variation in rainfall, topography and temperatures over Gauteng the area has similar climatological characteristics. Further detail about Gauteng is provided in Section 2.2. There are ample meteorological data available in the province. There are two main weather offices in Gauteng, at OR Tambo International Airport and the head office of the South African Weather Service (SAWS) in Pretoria. Upper air soundings are done daily at the Irene Weather Office, located in the centre of the province, and there are about 50 rainfall stations in Gauteng (See Fig. 2-1).

One of the forecasting challenges for Gauteng is that the type of weather system responsible for precipitation, and heavy rainfall events, differs considerably from early to late summer. During early summer the atmosphere has a distinct extra-tropical nature when weather systems such as COLs are frequent (Singleton and Reason, 2007) but late summer (January and February) tropical circulation systems are much more prevalent over South Africa (Dyson and Van Heerden, 2002). Heavy rainfall from the different types of weather systems are produced by different mechanisms and thus also require different forecasting techniques. The rainfall prognosis from NWP models plays an increasingly important role in the weather forecasting process. However, in a recent study Landman (2012) showed that the bias of the rainfall forecast in a multi-model ensemble system increases from early to late summer as the tropical weather systems dominate the circulation over South Africa.

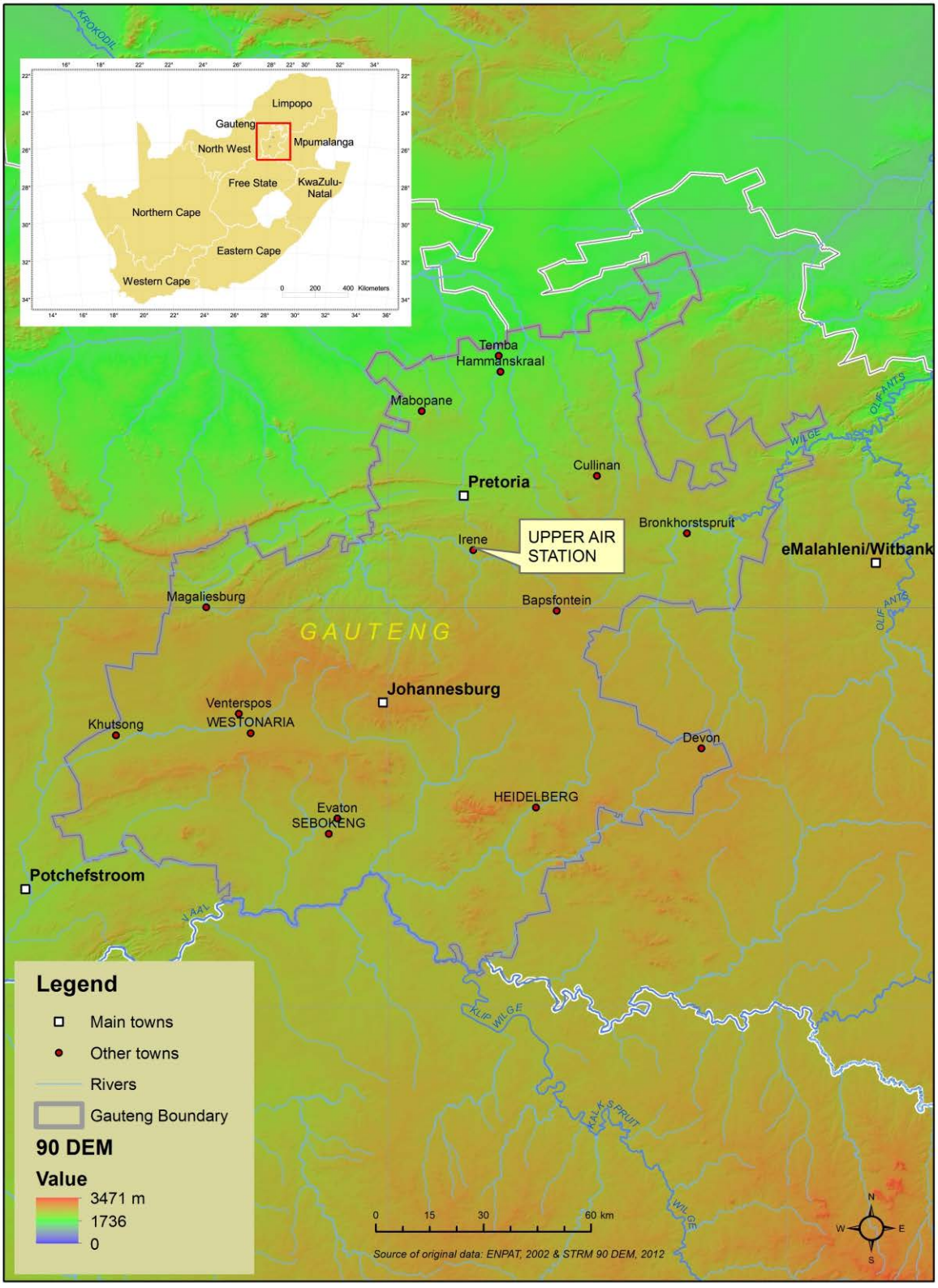


Figure 1-1: Location map of the Gauteng Province in South Africa.

In this research the first approach is to gain better understanding of heavy rainfall and the associated atmospheric conditions by creating climatologies. A heavy rainfall climatology is first constructed and the results from this study are then utilized to create a baseline climatology for the thermodynamic profile of the atmosphere when heavy rainfall occurs. A synoptic climatology is also created where those synoptic-scale weather systems associated with heavy rainfall are identified and described.

For any precipitation to occur, moist air should rise, inducing condensation which leads to cloud formation and precipitation. Once air starts to rise information about the thermodynamic profile of the atmosphere is necessary to determine to what height the ascent will continue. One approach in investigating the thermodynamic profile of the atmosphere is to consider sounding-derived parameters. In this context a sounding refers to a radiosonde connected to a balloon which observed various atmospheric parameters as it rises through the troposphere. In recent years many studies have been undertaken where sounding-derived parameters are associated with severe weather phenomena. Examples are: Dimitrova et al. (2009), Doswell and Schultz (2006), Dupilka and Reuter (2006), Covadonga et al. (2009) and Manzato (2003). Most, if not all, of these studies have been undertaken in the USA or Europe. In each instance a set of ingredients were identified which may be associated with severe weather for a specific geographical location. However the South African elevated plateau (± 1500 m a. m. s. l.) has a significant influence on the thermodynamic profile, moisture content of the atmosphere and values (magnitude) of these severe weather parameters. The threshold values and ingredients identified over other parts of the globe are therefore, unfortunately, not directly applicable to Gauteng. The research described in this thesis is the first investigation of this kind undertaken in South Africa and provides information about parameters associated with heavy rainfall over Gauteng.

1.2 Aim and Objectives

The aim of this study is to describe the atmospheric conditions associated with heavy rainfall over the Gauteng Province and to identify the parameters and understand the processes associated with heavy rainfall. To achieve this aim the following specific objectives are applicable.

1. Classify and describe the spatial and temporal distribution of heavy rainfall over Gauteng by constructing a climatology of heavy daily rainfall over Gauteng using 30-yrs of daily rainfall data. The identified heavy rainfall events are utilized to construct a baseline climatology of sounding-derived parameters.

2. Investigate and identify the synoptic-scale weather systems associated with heavy rainfall over Gauteng and to discuss the seasonal changes in the dominant weather system as the summer season progresses.
3. Investigate the sounding-derived parameters associated with heavy rainfall over Gauteng. Two important features are explored by the use of the radiosonde data. The first is to identify those variables which can best explain the occurrence of heavy rainfall and then to describe how the values of these variables change as the summer season atmospheric circulation progresses from extra-tropical to tropical.
4. Utilize self-organizing maps (SOMs) to construct a climatology of the vertical profiles of temperature, dew point temperature, and wind strength and direction as well as sounding-derived parameters associated with heavy rainfall over Gauteng.
5. Propose a methodology on how to use a combination of circulation criteria and sounding-derived parameters in a SOM as a tool to predict rainfall over Gauteng.

1.3 Outline of the document

Each of the objectives is dealt with separately in the following chapters and the relevant background and scientific content are provided in each of the chapters separately as is the data and methods used. The relevant references for each section are also provided at the end of each chapter.

Chapter 2 deals with the rainfall climatology over Gauteng and most of this chapter is based on a document already published in a peer reviewed journal. The article deals with rainfall up to 2009 but in Chapter 2 some of the rainfall statistics are updated to include data up to the 2011/2012 rainfall season. The results of this Chapter provide the necessary information about heavy rainfall to focus the research of the remainder of the objectives. The synoptic circulation and sounding parameters are analysed on these days in consequent chapters.

Chapter 3 discusses rain bearing synoptic circulation patterns over South Africa. Synoptic maps and atmospheric flow fields are discussed per month for days with heavy rainfall and days which remain dry. Emphasis is placed on how heavy rainfall producing weather systems change throughout the summer rainfall season.

In Chapter 4 a 35-year monthly climatology of basic atmospheric variables (temperature, moisture and winds) throughout the troposphere is provided by using the Irene Weather Office sounding data, followed by a similar climatology of sounding-derived parameters. This chapter then discusses the deviation in value of all these variables during heavy rainfall days as well as days

without any rainfall. Finally a detailed heavy rainfall climatology of the vertical profile of the atmosphere as well as sounding-derived parameters is provided by using SOMs.

In Chapter 5 a method is discussed which combines sounding-derived parameters with circulation criteria in a SOM which is then used to predict rainfall frequency distributions over Gauteng.

Chapter 6 provides a summary and conclusions and places the results from this thesis into context.

1.4 References

ALEXANDER WJR and VAN HEERDEN J (1991). Determination of the risk of widespread interruption of communications by floods. Department of Transport Research Project RDAC 90/16.

BLAMEY RC and REASON CJC (2012). Mesoscale Convective Complexes over Southern Africa. *Journal of Climate*, **25**, 753–766.

COVADONGA, GIAIOTTI, DARIO, STEL, FULVIO, CASTRO, AMAYA, FRAILE and ROBERTO (2010). Maximum hailstone size: Relationship with meteorological variables. *Atmospheric Research,* **96**, 256-265.

DE CONING E, FORBES GS and POOLMAN EP (1998). Heavy precipitation and flooding on 12-14 February 1996 over the summer rainfall regions of South Africa: Synoptic and Isentropic analyses. *National Weather Digest,* **22,** 25-36.

DIMITROVA T, MITZEVA R and SAVTCHENKO A (2009). Environmental conditions responsible for the type of precipitation in summer convective storms over Bulgaria. *Atmospheric Research*, **93**, 30–38.

DOSWELL C A III and SCHULTZ DM (2006). On the use of indices and parameters in forecasting severe storms. *Electronic Journal of Severe Storms Meteorology,* **1**, 1–14.

DUPILKA ML and REUTER GW (2006). Forecasting Tornadic Thunderstorm Potential in Alberta Using Environmental Sounding Data. Part I: Wind Shear and Buoyancy. *Weather and Forecasting,* **21**, 325-335.

DWA (Department of Water Affairs) (2012). Hydrological Services - Surface Water (Data, Dams, Floods and Flows) URL: http://www.dwaf.gov.za/Hydrology/Provincial%20Rain/Provincial%20Rainfall.htm (Accessed 21 August 2012)

DYSON LL and VAN HEERDEN J (2001). The heavy rainfall and floods over the northeastern interior of South Africa during February 2000. *South African Journal of Science,* **97**, 80-86p.

DYSON LL and VAN HEERDEN J (2002). A model for the identification of tropical weather systems. Water SA, **28**(3), 249-258.

GIJBEN M (2012). The lightning climatology of South Africa. *South African Journal of Science*, **108** (3/4), Art. #740, 10 pages.

GUARDIAN (2011). South Africa Flood Death Toll Rises as Government Declares 33 Disaster Zones, 24 January, Johannesburg

HART NCG, REASON CJC and FAUCHEREAU N (2010). Tropical–Extratropical Interactions over Southern Africa: Three Cases of Heavy Summer Season Rainfall. *Monthly Review,* **138**, 2608–2623.

HOLTON JR (1992). *An Introduction to Dynamic Meteorology*. Academic Press. 511 pp.

KRUGER AC (2004). *Climate of South Africa. Climate Regions. WS45*. South African Weather Service, Pretoria, South Africa.

LANDMAN S (2012). A multi-model ensemble system for short-range weather prediction in South Africa. Unpublished MSc Thesis, University of Pretoria.

LANDMAN S, ENGELBRECHT FA, ENGELBRECHT CJ, DYSON LL and LANDMAN WA (2012). A short-range weather prediction system for South Africa based on a multi-model approach, *Water SA*, **38**(5), 765-773.

MANZATO A (2003). A climatology of instability indices derived from Friuli Venezia Giulia soundings, using three different methods. *Atmospheric Research,* **67– 68,** 417– 454.

NEWS24 (2012). Western Cape officials battle snow, floods. URL: http://www.news24.com/SouthAfrica/News/Western-Cape-officials-battle-snow-floods-20120716. (Accessed 29 July 2012)

NGUBO N, MALALA R, MABUSELA X, DE CONING E, ROSENFELD D and KERKMANN J (2008). Hailstorm in Potchefstroom area (South Africa, 27 October 2004). URL: http://www.eumetsat.int/Home/index.htm (Accessed on 14 November 2008).

ORLANSKI I (1975) A rational subdivision of scales for atmospheric processes. *Bulletin of the American Meteorological Society*, **56**, 527–530.

PIELKE RA, S. (2002). *Mesoscale Meteorological Modeling*. San Diego: Academic Press 675 pp.

SAWB (SOUTH AFRICAN WEATHER BUREAU) (1991) History of notable weather events in South Africa 1500-1990. *Caelum.* South African Weather Service, Pretoria, South Africa.

SINGLETON AT and REASON CJC (2007). Variability in the characteristics of cut-off low pressure systems over subtropical southern Africa. *International Journal of Climatology,* **27**(3), 295-310.

TALJAARD JJ (1985). Cut-off lows in the South African region. South Africa. South African Weather Bureau. Technical Note No 14. South African Weather Service, Pretoria, South Africa.

TALJAARD JJ (1996). Atmospheric circulation systems, Synoptic climatology and weather Phenomena of South Africa. Part 6 Rainfall.in South Africa. South African Weather Bureau Technical Paper (32). South African Weather Service, Pretoria, South Africa.

VIVIERS G and CHAPMAN V (2008). Severe thunderstorm at Johannesburg International Airport (03 December 2004). URL: http://www.eumetsat.int/Home/index.htm (Accessed on 14 November 2008).

2 Rainfall climatology over the Gauteng Province in South Africa

2.1 Preface

The major part of this Chapter is an article published in *Water SA* and deals with a heavy rainfall climatology over the Gauteng Province from 1977-2009. Two heavy rainfall classes are developed; the area-averaged heavy rainfall and heavy rainfall at a single rainfall station (see Chapter 2.2.5). The results obtained are essential to the rest of the thesis as days with heavy rainfall are identified over Gauteng. These days are consequently used to analyse the synoptic circulation patterns (Chapter 3) and the sounding-derived parameters (Chapter 4) associated with heavy rainfall. The article also provides a better understanding of the nature of heavy rainfall over the province and the results pertaining to the seasonal change in heavy rainfall during summer contribute to the understanding rainfall variability over a summer rainfall area of South Africa. The article's details are as follows:

> Dyson LL (2009). Heavy daily-rainfall characteristics over the Gauteng Province. *Water SA,* **35**, 627-638.

Area-averaged heavy rainfall and heavy rainfall at single rainfall stations are subdivided into different classes depending on the amount of daily rainfall. 'Significant rainfall' is defined in this article as the 90th percentile for the area-averaged and single station heavy rainfall classes. For the area-averaged heavy rainfall class, 'significant rainfall' occurs when the area-averaged rainfall over Gauteng exceed 10 mm and for single station rainfall when rainfall at any station over Gauteng exceeds 50 mm. Daily 'heavy rainfall' and 'very heavy rainfall' is the 95th and 99th percentile respectively. Days identified as part of the research in this article are then used to investigate atmospheric circulation and thermodynamics associated with heavy rainfall. Sounding parameters are derived from upper air assents done at the Irene weather office at 1200 UT. In Chapter 4 detailed information is provided about the availability of these soundings after Quality Control procedures have been employed. As the soundings are not available every day for analysis, the number of 'heavy rainfall' and 'very heavy rainfall' days as defined in this article was insufficient to provide statistically significant results (See Chapter 4.4.5). It was therefore decided to use days on which 'significant rainfall' occurs for further analysis. In the following Chapters the 'significant rainfall' class in the article is referred to as "area-averaged heavy rainfall" (AHR) and "single station heavy rainfall" (SHR).

In this article rainfall over Gauteng is investigated from 1977 until the 2008/09 summer rainfall season. In the overview of this Chapter some of rainfall statistics are updated to include rainfall over Gauteng up to the 2011/12 summer rainfall season. The statistics provided in Table 2-2 and 2-3 are

not influenced significantly by the inclusion of data for two additional summer rainfall seasons and is left unchanged. Table 2-1 and Figure 2-1 are updated to Table 2-4 and Figure 2-5 to include the latest information and Figure 2-4 is reproduced for the SHR class for these two seasons (Figure 2-6). Furthermore during December and January months in the 2010/11 rainfall season the rainfall was well above normal and merits discussion in the overview.

2.2 Heavy daily-rainfall characteristics over the Gauteng Province[1]

Liesl L Dyson

Published as: Dyson LL (2009). Heavy daily-rainfall characteristics over the Gauteng Province. *Water SA,* **35**, 627-638.

Abstract

Daily rainfall over the Gauteng Province is analysed for the summer months of October to March using 32-yr (1977 to 2009) daily rainfall data from about 70 South African Weather Service stations. The monthly and seasonal variation of heavy rainfall occurrences is also analysed. Three 24-h heavy rainfall classes are defined considering the area-average rainfall. A significant rainfall event is defined when the average rainfall exceeds 10 mm, a heavy rainfall event when the average rainfall exceeds 15 mm and a very heavy rainfall event when the average rainfall exceeds 25 mm. January months have the highest monthly average rainfall as well as the highest number of heavy and very heavy rainfall days. The month with the 2nd highest number of heavy and very heavy rainfall days is February followed by March and October. December has the 2nd highest monthly average rainfall and the most days with rain. However, it is also the month with the lowest number of heavy and very heavy rainfall days. The highest 24-h rainfall recorded at a single station during the 32-yr period was 300 mm in December 2006. However, rainfall exceeding 115 mm at a single rainfall station in the Gauteng Province is very rare and does not occur every year. January months receive these events more than any other month but this only transpires in approximately a third of years. The central and northwestern parts of the Province experience the most events where the rainfall at a single station surpasses 75 and 115 mm. With regard to seasonal rainfall, the 1995/96 summer rainfall season had the highest seasonal rainfall during this 32-yr period followed by the 1999/2000 season. The 1995/96 season had above normal rainfall in both early and late summer but the 1999/2000 season was dry in early summer and very wet in late summer. Significantly high seasonal rainfall is associated with above-normal rainfall in late summer.

Key words: widespread heavy rainfall, Gauteng, daily rainfall, weather forecasting, summer rainfall, climatology of heavy rainfall

[1] Text appears as published but references, figures and tables were adapted to the style of this thesis

2.2.1 Introduction

Rainfall resulting in flooding occurs from time to time over the Gauteng Province. The heavy rainfall events may take place over most of the province and last for days, resulting in widespread flooding and disruption of infrastructure and even loss of life. Examples of such events took place during January and February 1996 (De Coning et al., 1998) and February 2000 (Dyson and Van Heerden, 2001). However, heavy rainfall may also occur in isolated areas over the Gauteng Province from so called mesoscale weather systems, resulting in flash flooding. In these instances the heavy rainfall may be of a short duration (but intense) and is often associated with strong winds and hail (Viviers and Chapman, 2008; Ngubo et al., 2008).

Gauteng Province (hereafter Gauteng) is situated on the interior plateau of South Africa and receives most of its rainfall in summer. The annual average rainfall in Gauteng varies between just over 700 mm on the Witwatersrand (approximately 1 700 m a. m. s. l.) and just over 600 mm north of the Magaliesberg (approximately 1 100 m a. m. s. l.). Most of Gauteng falls into the climate region: Moist Highveld Grassland and is relatively cool with average annual maximum temperatures of about 22˚C in the south but increasing to 25˚C in the north. There are about 100 days with rain in Johannesburg and 85 in Pretoria (Kruger, 2004). The extreme northern parts of Gauteng fall into the Central Bushveld climate region. According to Kruger (2004) the maximum rainfall over Gauteng occurs during the December and January months.

Gauteng is responsible for over a third of South Africa's Gross Domestic Product (GDP) and a tenth of Africa's GDP. Geographically, Gauteng is the smallest province in South Africa, covering approximately 16 500 km^2, but nearly 20 % (9.6 m.) of South Africa's population reside in Gauteng. It is estimated that Gauteng will be home to 14.6 m. people by 2015. There were 405 informal settlements in Gauteng in 2006 (Gauteng Department of Housing, 2006); the overcrowding in these settlements has reached extreme proportions with as many as 24 people sharing a living space of approximately 40 m^2 (Beavon, 2004). The vacant land on the river banks in the informal settlements has also become populated by shacks and these communities are especially vulnerable to flash flooding. This study focuses on Gauteng due to its importance to the economy of South Africa, and the high population density, but also due to the availability of observed meteorological data in the province. In addition it is also one of the regions for which weather forecasts are issued on a daily basis.

In an attempt to understand more about the characteristics of heavy rainfall over Gauteng, observed daily rainfall data were analysed for the summer months (October-March) for a period of 32 years. In this paper early summer refers to October to December and late summer to January to

March. The rainfall at individual stations is investigated, but the emphasis is on the areal average rainfall over Gauteng. One of the forecasting challenges for Gauteng is that the type of weather system responsible for precipitation, and indeed heavy rainfall, differs considerably from early to late summer. During early summer the atmosphere has a distinct extra-tropical nature when weather systems such as cut-off lows are frequent (Singleton and Reason, 2007). However, in late summer (January and February) tropical circulation systems are much more prevalent over South Africa (Dyson and Van Heerden, 2002). In this paper the emphasis is not on the weather systems responsible for the heavy rainfall but rather concentrates on rainfall statistics over Gauteng. The results from this paper form the basis of ongoing research investigating the atmospheric variables and synoptic circulation patterns associated with heavy rainfall over Gauteng.

An example of heavy rainfall 'climatology' in the scientific literature is by Maddox et al. (1979), who described aspects of flash flooding over the USA. More recently, Brooks and Stensrud (2000) created an hourly rainfall climatology over the USA, followed by Schumacher and Johnson (2006), who described characteristics of extreme rain events over the eastern two-thirds of the United States. They found that extreme rain events (where the 24-h precipitation total at 1 or more stations exceeds the 50-yr recurrence amount for that location) are most common in July and in the northern USA these events transpire almost exclusively in the warm season. They also concluded that most of these events (66 %) are associated with mesoscale convective systems while synoptic and tropical systems play a larger role in the south and east. Chen et al. (2007) used a similar statistical approach to investigate heavy rainfall in Taiwan. They found that heavy rainfall occurs with a pronounced afternoon maximum over Taiwan and that the orographic effects are important in determining the spatial distribution of heavy rainfall.

A better understanding and knowledge of the climatology of heavy rainfall will facilitate the forecasting of these extreme events. The main aim of this paper is to make weather forecasters aware of how likely heavy rainfall events are over Gauteng during a particular summer month. Understanding the spatial and temporal distribution of heavy rainfall events is a key aspect in furthering this aim. As flood-producing heavy rainfall events are infrequent, knowledge of the climatology of these events could therefore also aid inexperienced weather forecasters, by providing guidance as to how likely heavy rainfall might be during a particular time of the year.

The data used in this analysis are discussed first and some of the problems encountered in the dataset are highlighted. Information is consequently supplied about the seasonal, monthly and daily rainfall characteristics in Gauteng. Three different heavy rainfall classes are defined for average daily

Gauteng rainfall and the monthly characteristics of these events are examined. Lastly, heavy rainfall characteristics at individual stations are discussed for each of the summer months.

2.2.2 Data and methods

Daily rainfall data were obtained from the South African Weather Service (SAWS) for all summer months from October 1977 to March 2009. SAWS rainfall stations report 24-h accumulated rainfall in the morning (0800 South African Standard Time). All the rainfall stations over Gauteng were investigated for suitability to use in this analysis. Only stations where data were available for at least 75 % of the period were considered. There were 58 stations with a record length spanning between 90-100 % of the period and another 10 stations with record lengths between 75-90 %. However, data from selected rainfall stations with records spanning shorter parts of the period were also included. This was done mainly to capture data in cases where rainfall stations were replaced by new stations with only slightly different locations within the period under consideration. An example is the rainfall station at OR Tambo International Airport which closed on the 31 May 1989 while another station opened on 1 June 1989 at almost the same location. Data from both these stations were then used in the analysis. Data for 5 locations were combined in this way resulting in a total number of 73 stations available for analysis over Gauteng. Not all the rainfall stations were operational every day (i.e., there are discontinuities within the time series of some stations) and consequently the number of rainfall stations which were available for analysis varied between 55 and 73 on any given day within the 32-yr period.

Figure 2-1 depicts the location of the SAWS rainfall stations over Gauteng. The rainfall stations are generally spatially well-distributed throughout the province, with the exception being the northeastern extremes where no rainfall stations were available. There is a higher concentration of rainfall stations in the major metropolitan areas of Gauteng (Pretoria in the north and Johannesburg, about 50 km from Pretoria to the south).

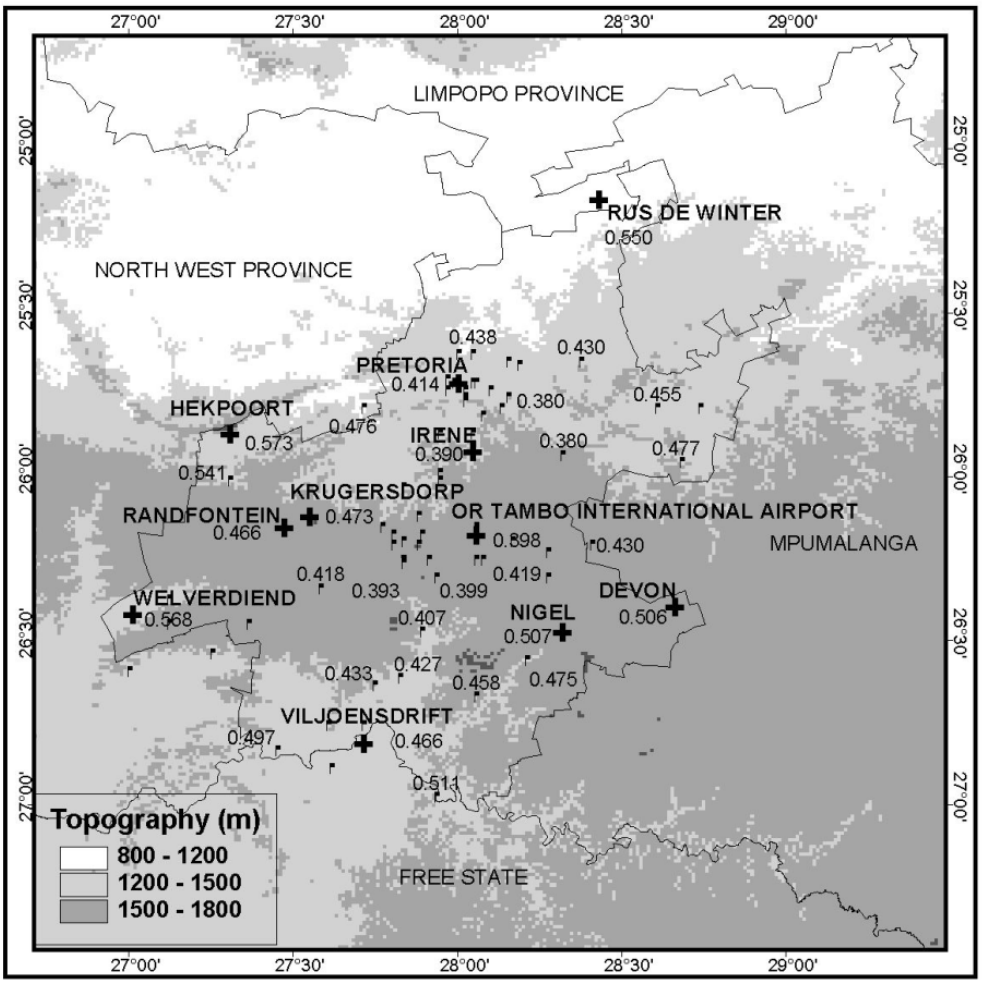


Figure 2-1: The location of the rainfall stations over the Gauteng Province utilised in the study. The shading indicates height (in m) a. m. s. l. The numerical values are the weights assigned to each of the rainfall stations in the calculation of the area-average rainfall.

2.2.3 Quality control of rainfall data

The quality of the rainfall data used, especially in such a large dataset, is of significant concern, and a considerable amount of time was spent performing quality control on the data. Some obvious errors were easy to identify and were removed from the data set. However there were some questionable data values where it was close to impossible to determine the reliability of the observations. The raw rainfall data from SAWS include possible error information, with the daily rainfall values labelled as '**Normal'**, or '**Error**' or '**Accumulated**' (if accumulated over more than 1 d). If the rainfall value was labelled anything other than **Normal** it was not used in the analysis for that particular day. As this research focuses on heavy rainfall it was important to have confidence in the high 24-h rainfall values. Brooks and Stensrud (2000) explain how difficult it is to distinguish between 'rare interesting' rainfall events and 'bad data', as these often look similar. Therefore all rainfall events where 24-h rainfall at a specific station exceeded 115 mm were investigated for possible errors. As will be explained later, 115 mm was identified as a 'single very heavy rainfall event' and

represents the 99th percentile of daily maximum rainfall over Gauteng. It does happen from time to time that rainfall which was accumulated is not identified as such in the raw data set. This was relatively easy to identify in the data set when there were missing data for 1 or more days followed by a day reporting very high rainfall. This high rainfall value was then rejected only after comparison with rainfall from surrounding stations, in the process discussed below.

A second set of errors removed from the data was all the cases where a station reported very heavy rainfall for several days in a row while there was no indication from surrounding stations that this did indeed occur. When rainfall at any station over Gauteng exceeded 115 mm on any particular day, the rainfall values at other stations over Gauteng were also analysed. If there were other stations reporting significant rainfall on that day or if there was a high percentage of rainfall stations over Gauteng reporting rainfall the value was accepted as correct. The last error check was to compare the events remaining, after the elimination of events considered by the previous checks, against other meteorological data and sources such as the SAWS newsletters and website and archived Meteosat 2nd Generation data. This was done in order to identify if there was a physical cause for a heavy rainfall event to occur. The real difficulty was in attempting to detect errors in the much larger number of rainfall events where the daily rainfall at a single station was more than 50 mm but did not exceed 115 mm (later defined as a 'single significant rainfall event'). There were too many of these events to hand-check and it would be difficult to determine the accuracy as there were no other data to compare with. It is therefore possible that the daily rainfall dataset created as part of this research does contain some errors which may have led to some inaccuracies in the results. However, the impact of this would be limited as the research focused on the heavy rainfall events.

2.2.4 Calculation of average daily rainfall

Using the rainfall data from the selected stations, an average daily rainfall value for Gauteng was calculated. Additionally, the percentage of rainfall stations recording more than 0 mm of rainfall was calculated and the highest rainfall measured at any of the stations was also noted. The average Gauteng rainfall was computed using a weighted average method proposed by Tennant and cited in Marx et al. (2003). This method takes the geographical position of each station relative to the other stations into consideration. The following weighting function was applied to the daily rainfall values of all the stations:

$$Wgt = \frac{\sum_{m=1}^{N-1} r_m}{(N-1)r_{\max}}$$

where:

Wgt = weight assigned to station

$\sum_{m=1}^{N-1} r_m$ = sum of the distance between the specific station and all other stations

$r_{\max}$ = maximum distance between the specific station and any other station

N = total number of stations

When rainfall stations are distributed evenly over an area, this method renders results which are very close to the mathematical average (the total rainfall at all the stations divided by the number of rainfall stations). The use of the weighting method becomes important when the rainfall stations are not distributed evenly over an area, as is the case for Gauteng. A rainfall station which is geographically distant (close) to other stations will have a larger (smaller) weight factor and will therefore contribute more (less) in the computation of the average rainfall. The stations with the largest weights were over the western extremes of Gauteng (Fig. 2-1). Hekpoort in the northwest had a weight of 0.573 and Welverdiend in the southwest 0.568. Other stations with weights larger than 0.5 were Rus De Winter in the extreme north (0.55) and Devon (0.506) and Nigel (0.507) in the south-east. These stations contributed more to the calculation of the average rainfall than the stations over central Gauteng such as Irene (0.390), Pretoria (0.414) and OR Tambo International Airport (0.398). The results from the 2 averaging methods are very similar with the weighted average method, generally giving slightly higher daily average values. Of the 5 673 days analysed, the weighted average method produced higher (lower) values than the mathematical average on 753 (560) days. The largest difference occurred on 27 January 1978 when the weighted average was 67 mm and the mathematical average 59 mm. This was a particularly wet day as 23 stations measured more than 50 mm of rain and 11 more than 115 mm. Stations over western Gauteng in particular measured high rainfall values, e.g. Randfontein (100 mm), Krugersdorp (116 mm) and Hekpoort (66 mm). From Fig. 2-1 one can see that these stations also have higher weights in the calculation of the average and therefore the weighted average was higher than the mathematical average for this day. On 18 December 2006 the weighted average rainfall was 18 mm but the mathematical average was 23 mm.

On this day there were only 4 stations with rainfall of more than 50 mm over southern Gauteng, but with 300 mm at Viljoensdrift. Due to the isolated nature of the heavy rainfall, the weights assigned to the stations resulted in the weighted average being lower than the mathematical average; the influence of the extreme rainfall at a single station is therefore de-emphasised.

The average Gauteng daily rainfall was henceforth used to calculate the average Gauteng monthly rainfall and the data standardised in order to identify wet and dry months. Moreover, the rainfall at the individual rainfall stations was investigated in order to identify those locations in Gauteng where heavy rainfall occurs most frequently. This was done by dividing Gauteng into eight 0.5˚ by 0.5˚ grid boxes and the number and percentage of stations that exceeded certain thresholds were calculated for each grid box for each 24-h period considered.

2.2.5 Defining heavy rainfall

An extreme precipitation event is usually defined by using a daily amount exceeding a certain threshold (Zhang et al., 2001). However, different threshold values apply for different parts of the world. One approach is to define heavy rainfall by considering when the areal average rainfall exceeds a particular threshold. For example, Houze et al. (1990) define a 'major rain event' as one in which more than 25 mm rain falls over an area greater than 12 500 km^2 in 24 h. In a South African study, Poolman (contributing to Dyson et al., 2002) defines a heavy rainfall event when more than 25 mm occurs in 24 h in an area of at least 20 000 km^2.

The Gauteng Province is approximately 16 500 km^2 in size. When the average daily rainfall is at least 25 mm over Gauteng it would fall into the major rain event definition provided by Houze et al. (1990). However, analysis of the daily average rainfall data over Gauteng for the 32-yr period reveals that 25 mm is exceeded only 1 % of the time. Over the 32-yr period the daily average rainfall exceeded 25 mm on only 65 occasions. In order to capture 10 % of the heaviest rainfall events, a 'significant rainfall event' is defined here by using the 90th percentile, which is 9 mm. A further classification is made, with a 'heavy rainfall event' defined as the 95th percentile, in this case 13 mm, and a 'very heavy rainfall event', similar to Houze's major rain event, when average daily rainfall exceeds 25 mm. However as these thresholds may be applied in an operational environment they were adjusted slightly to fall in line with thresholds commonly used in the forecasting offices. Therefore a 'significant rainfall event' is classified as rainfall exceeding 10 mm, a 'heavy rainfall event' when the rainfall exceeds 15 mm and a 'very heavy rainfall event' when the rainfall exceeds 25 mm.

Extreme precipitation events are often defined by referring to the rainfall from individual stations. Bradley and Smith (1994) define extreme rainstorms as a 'major rain event' when the daily

rainfall accumulation is at least 125 mm at 1 or more rainfall stations. However, Chen et al. (2007) define a heavy rainfall event in Taiwan when more than 50 mm occurs in 24 h at 1 or more weather stations and an extremely heavy rain event when 130 mm occurs in 1 d (Chen and Yu, 1988). In a recent study Fragoso and Tildes Gomes (2008) identified an extreme rainfall event over southern Portugal when 40 mm occurred in 24 h.

Zhang et al. (2001) define heavy rainfall separately for different stations in Canada by identifying a threshold value that is exceeded by an average of 3 events per year. They also discuss the characteristics of heavy rainfall by examining the 90th percentile of daily rainfall and the 20-yr return values. Extreme rain events in the U.S. were identified when rainfall at 1 or more rain gauges reported a 24-h rainfall total greater than the 50-yr recurrence amount. This spatially-varying threshold is most relevant to identify truly extreme events for this location (Schumacher and Johnson, 2006).

The maximum daily rainfall which occurred at any station over Gauteng was identified for all summer months. Following the definition for the areal average rainfall the 90th (59 mm), 95th (75 mm) and 99th (113 mm) percentile of these values was used to identify a heavy rainfall event at an individual station. However, the forecasters at SAWS issue advisories and warnings for heavy rainfall when more than 50 mm of rain is expected at any location (Rae, 2008). Therefore a 'single significant rainfall' event is defined when the rainfall at any rainfall station exceeds 50 mm. A 'single heavy rainfall' event is when the rainfall exceeds 75 mm at a single station and a 'single very heavy rainfall' event when the rainfall at a single station exceeds 115 mm. This value is close to the 125 mm used by Bradley and Smith (1994) and the 130 mm used by Chen et al. (2007).

Two additional heavy rainfall classes are defined which combined the areal average rainfall and rainfall at individual rainfall stations. A 'major rain event' is defined when the average daily rainfall over Gauteng exceeds 10 mm with at least 50 mm at a single station and an 'extreme rainfall event' is defined when the average Gauteng rainfall exceeds 15 mm with more than 75 mm at a single station.

2.2.6 Seasonal rainfall characteristics over Gauteng

The average Gauteng monthly rainfall was calculated for each summer month from 1978 to 2008. The monthly values were used to calculate an average rainfall value for early summer (October to December) and late summer (January to March) as depicted in Table 2-1. The average summer rainfall (October to March) over Gauteng for this 32-yr period was 587 mm. The highest summer rainfall occurred during the 1995/96 season with 968 mm followed by the summer of 1999/2000

with 793 mm. The driest summer was 1978/79 when the average Gauteng rainfall was 341 mm. The 2nd driest summer was in 2006/07 when the average rainfall was 364 mm.

The totals depicted in Table 2-1 were standardised with respect to the long-term average and standard deviation for early and late summer rainfall and the results are depicted in Fig. 2-2. The 32-yr average early summer rainfall was 278 mm and the late summer rainfall 309 mm. The early summer with the highest average rainfall was in 1995 when 452 mm occurred over Gauteng. There were only 2 other years when the early summer rainfall exceeded 400 mm – 1986 (412 mm) and 1993 (406 mm). The early summer rainfall was less than 200 mm on 5 occasions in the past 32 years. Three of these very dry early summers occurred in the past decade (2002, 2003 and 2005). Also note from Fig. 2-2 the 5 consecutive years of below-normal early summer rainfall (2002 to 2006). This had not transpired before during this 32-yr period, although the early 80s had 3 consecutive dry years.

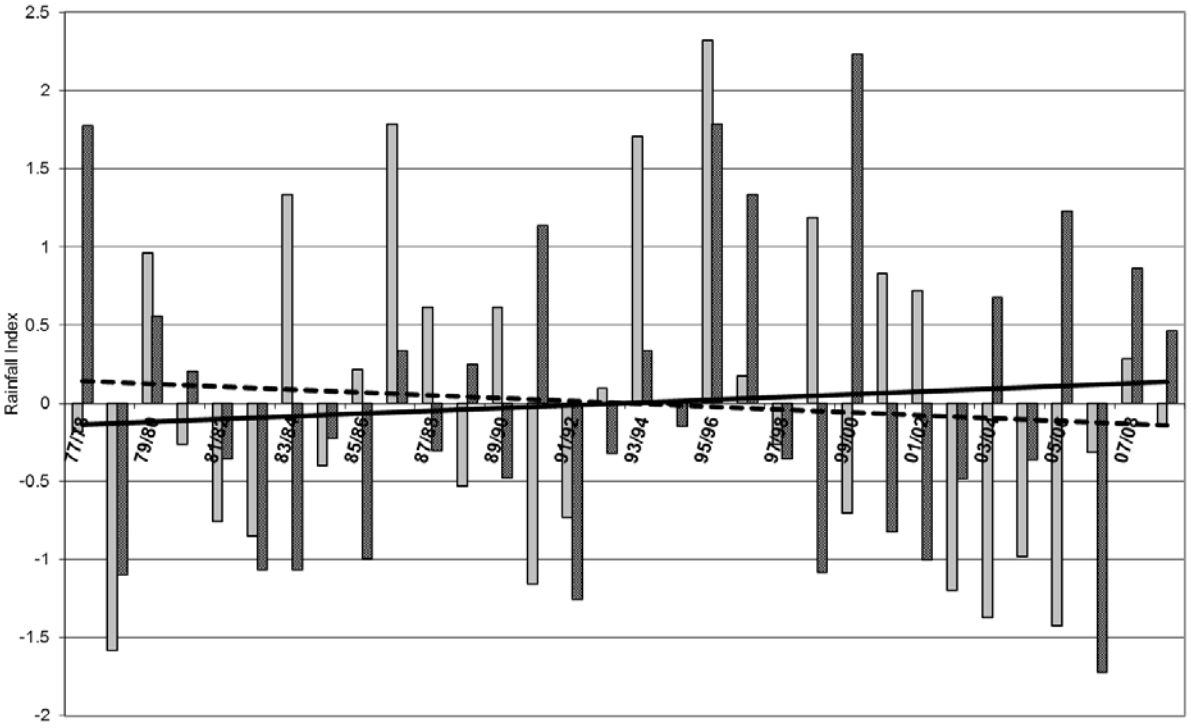


Figure 2-2: Standardized rainfall index for early summer (light grey) and late summer (dark grey) for 1977 to 2009. The early summer trend line is shown in a dotted line while the solid line is the late summer trend line.

Table 2-1: The early summer (October to December) and late summer (January to March) average rainfall totals (mm) for 1977 to 2009. The grey cells show rainfall more than 400 mm while the black cells show rainfall less than 200 mm.

	Rainfall Totals (mm)		
	Early Summer (Oct-Dec)	Late Summer (Jan-Mar)	Total
77/78		515	
78/79	159	182	341
79/80	350	374	724
80/81	258	333	591
81/82	221	268	489
82/83	214	186	400
83/84	378	186	564
84/85	248	283	531
85/86	294	194	488
86/87	412	348	760
87/88	324	274	598
88/89	238	338	576
89/90	324	254	578
90/91	191	441	632
91/92	223	164	387
92/93	285	272	557
93/94	406	348	754
94/95	277	292	569
95/96	452	516	968
96/97	291	464	755
97/98	258	268	526
98/99	367	184	551
99/00	225	568	793
00/01	340	214	554
01/02	332	193	525
02/03	188	253	441
03/04	175	388	563
04/05	204	267	471
05/06	171	452	623
06/07	254	110	364
07/08	299	409	708
08/09	267	363	630
Average	278	309	581

It is no surprise to find that the wettest late summer over Gauteng was in 2000, with 568 mm of rain, as tropical cyclone Eline invaded southern Africa in February 2000 (Dyson and Van Heerden, 2001) and was responsible for widespread heavy rainfall over the entire sub-continent, including Gauteng. Other particularly wet late summers were 1978, 1991, 1996, 1997, 2006 and 2008, with more than 400 mm of rain. The driest late summer in this 32-yr period was in 2007 when a meagre 110 mm occurred. This value is lower than in 1992 (164 mm) and 1983 (186 mm). The latter 2 years

have been identified as the years when the strongest droughts occurred over the summer rainfall area of South Africa (Rouault and Richard, 2003) since the 1921 drought. Although Fig. 2-2 depicts only Gauteng rainfall, it may be noted that a contributing factor to the severity of the drought during these 2 years was that the dry late summers were preceded by dry early summers. Indeed, from considering Fig. 2-2 it seems that notably low average early and late summer rainfall occur rather seldom in 1 rainfall season. The correlation coefficient between early and late summer rainfall is only 0.027 indicating that there is no significant correlation between early and late summer rainfall over Gauteng. The appreciably dry early and late summers of 19978/79, 1982/83 and 1991/92 are therefore quite noteworthy. An equally rare event is a very wet early summer followed by a very wet late summer, which only happened in 1979/80 and 1995/96.

The very high rainfall totals over Gauteng in early 2000 (568 mm) were preceded by a dry early summer rainfall season when only 225 mm occurred. On 4 occasions a noteworthy wet (0.5 on the ordinate on Fig. 2-2) early summer was followed by a very dry late summer (1983/84, 1998/99, 2000/01 and 2001/02). It also happened at least 4 times that very dry early summers were followed by very wet late summers.

The dashed line on Fig. 2-2 is the trend line for early summer rainfall and the solid line the trend line for late summer rainfall. These trend lines indicate a decrease in the early summer rainfall while an increase is observed in the late summer rainfall over Gauteng. Trend analysis was done using the nonparametric Mann-Kendall test (Hcigizoglu et al., 2005). The downward trend in early summer rainfall has a confidence level of only 62 % while the confidence level of the upward trend in late summer is 82 %. Therefore, the trends observed on Fig. 2-2 are considered to be statistically not significant. However, further statistical trend analyses of the rainfall over Gauteng is recommended as recent work by Engelbrecht et al. (2009) found that the model-projected climate change shows an increase in summer rainfall over northeastern South Africa.

2.2.7 Monthly and daily rainfall characteristics over Gauteng

2.2.7.1 January months

January months received the highest average monthly and daily rainfall over Gauteng. They also have the 2nd highest number of days with rain. The standard deviation of the monthly average rainfall was the highest in January and February months indicating the high variability in the monthly average rainfall during these 2 months. This is reflected in the values in Table 2-2 indicating that the average January rainfall is 126 mm but the minimum monthly average rainfall was only 56 mm in 2001. In contrast, in 1978 324 mm was recorded; this is the highest monthly rainfall for any month

during this 32-yr period. In the same month the monthly averaged daily rainfall was 10.5 mm, more than double the average for all January months (4.1 mm). On average rainfall occurs quite often during January months as more than 0 mm of rain was recorded on 23 d. In 2005 there were 28 d with some rain and in 2001 and 2007 only 14 d with rain.

2.2.7.2 December months

December months have the 2nd highest monthly average rainfall (109 mm) and they have the most days with some rain (24 d). The December months of 1988 and 1991, had 29 d on which some rainfall occurred. The standard deviation of the average monthly December rainfall is 30 mm. This value is about half of the standard deviation in January and February months.

2.2.7.3 February and November months

The average monthly rainfall for November months (96 mm) and February months (97 mm) is very similar as is the number of days with more than 0 mm of rain at approximately 20 d. However, the standard deviation in November was 45 mm while it is 59 mm in February. The minimum average monthly rainfall in a November month was 17 mm (2002) with a maximum of 183 mm in 1989. The minimum average monthly rainfall value in February months (25 mm) is similar to November months but the maximum monthly rainfall was significantly higher at 277 mm.

2.2.7.4 March and October months

The average monthly rainfall during March (86 mm) was higher than in October (72 mm) even though the number of days with some rain was very similar (15 to 17 d). The standard deviation of the average monthly rainfall in March was 54 mm and 40 mm in October. The maximum monthly average rainfall in a March month was approximately 100 mm more in March (296 mm) than in October (190 mm).

It is also interesting to note from Table 2-2 that the month with the most rain days did not necessarily coincide with the month with the highest rainfall. For example, the highest rainfall in a December month was in 1995 but it had only 23 d with rain, very close to the monthly average. Some rainfall occurs on average on 118 d in the summer season, 64 % of all days in summer. The 2006/07 rainfall season was particularly dry as it had only 98 d with rain. The 1993/94 season had 142 d with some rain; this is close to 80 % of the summer days.

Table 2-2: Monthly and daily average rainfall statistics over the Gauteng Province. The year of occurrence is shown in brackets and Sv indicates that it occurred in more than 2 years. The seasonal average values are shown in the last row and * indicates the minimum value per season

A – Average Mx - Maximum Mn Minimum

	Average monthly rainfall (mm)				Average daily rainfall (mm)		Mx daily rainfall	No. of days with rainfall > 0 mm			Significant Rainfall				Heavy Rainfall				Very Heavy Rainfall			
											No. days with rainfall > 10 mm.		Mx No. of consecutive days	Percentage of years	No. days with rainfall > 15 mm		Mx No. of consecutive days	Percentage of years	No. days with rainfall > 25 mm		Mx No. of consecutive days	Percentage of years
	A	Mn	Mx	StD	A	Mx	Mx	A	Mn	Mx	A	Mx		%	A	Mx		%	A	Mx		%
Oct	72	15 ('05)	190 ('93)	40	2.3	6.1 ('93)	70 (10/'86)	15	6 ('05)	25 ('07)	2.1	5 ('93,'00)	4 ('07)	81	1.1	4 ('93)	2 (Sv)	63	0.34	2 (Sv)	2 ('85,'00)	19
Nov	96	17 ('02)	183 ('89)	45	3.2	6.1 ('89)	35 (25/'79 01/'85)	20	9 ('85)	28 ('87)	2.6	7 ('79)	4 ('00)	84	1.1	5 ('79, '89)	3 ('79,'95)	50	0.31	2 ('98)	2 ('98)	28
Dec	109	50 ('03)	179 ('95)	30	3.5	5.8 ('95)	46 (16/'95)	24	16 ('84)	29 ('88,'91)	2.9	5 (Sv)	2 (Sv)	97	1.0	4 ('86)	2 (Sv)	63	0.22	2 ('95)	2 ('95)	16
Jan	126	56 ('01)	324 ('78)	59	4.1	10.5 ('78)	67 (27/'78)	23	14 ('01,'07)	28 ('05)	3.8	11 ('06)	6 ('78)	94	1.6	8 ('78)	5 ('78)	63	0.50	4 ('78)	1 (Sv)	34
Feb	97	25 ('07)	277 ('96)	59	3.4	9.6 ('96)	43 (08/'00 10/'00)	19	13 (Sv)	26 ('06)	2.8	10 ('96)	7 ('96)	91	1.7	6 ('89,'96)	5 ('96)	72	0.41	4 ('00)	4 ('00)	22
March	86	22 ('07)	296 ('97)	54	2.8	9.6 ('97)	54 (05/'97)	17	11 (Sv)	26 (Sv)	2.6	10 ('97)	5 ('95)	91	1.3	5 ('97)	3 ('91,'97)	66	0.38	3 ('91)	3 ('91)	25
Season	587	341 ('78/'79)	968 ('95/'96)		3.2	5.3 ('95/'96)		118	98 ('78/'79)	142 ('93/'94)	16.8	31 ('95/'96) 8* (Sv)		100	7.8	18 ('95/'96) 2* ('91/'92)		100	2.16	10 ('95/'96)		72

A South African weather forecaster should be aware of the increase in the average monthly rainfall during the progression of the summer from October to January with a decrease in February and March. On average October will receive some rainfall on about 50 % of the days while in December and January more than 70 % of the days receive some rainfall on average. By March the number of rain days has decreased to about 50 %. Weather forecasters should also take cognisance of the fact that even though the average monthly rainfall in January is higher than in December months, December has slightly more days with rain with a lower variability in the monthly average rainfall. The average monthly rainfall in late summer has high variability with maximum average monthly rainfall values all above 250 mm.

2.2.8 Synoptic circulation in wet and dry seasons

The average 850 and 500-hPa geopotential heights for the wettest and driest early and late summer seasons are displayed in Fig. 2-3. The average early summer rainfall of 1995 was 452 mm and there were 5 days when the area-average rainfall was more than 25 mm, while there were 4 days when rainfall at a single station exceeded 115 mm. At the 850 hPa level a deep low pressure system (1 490 gpm) was located over northern Namibia, extending a trough to the south coast of South Africa (Fig. 2-3A). The Indian Ocean High (IOH) was responsible for the inflow of warm moist air from over the Mozambique Channel into Gauteng. At the 500 hPa level (Fig. 2-3B) the average early summer geopotential height field shows a weak trough west of South Africa and a high over northern Namibia/Botswana. The early summer of 1978 had only 159 mm, there were no days which had area-average rainfall of more than 25 mm and only 2 days which had area-average rainfall of more than 15 mm. There were 5 d where rainfall at a single station exceeded 50 mm (compared to the 20 d in 1995), and there was not a single day that had more than 115 mm. At the 850 hPa level the trough was established over the western interior of South Africa while the low pressure over Namibia was considerable weaker (1 510 gpm) and located further west than in 1995 (Fig. 2-3C). The IOH extended over the eastern interior of the subcontinent. At the 500 hPa level the trough which was present in 1995 was not present and the high pressure over northern Namibia was stronger than in 1995 (Fig. 2-3D).

In the late summer of 2000, 568 mm of rain occurred. There were 5 days when the area-average rainfall was more than 25 mm and there were 7 days when the rainfall at a single station exceeded 115 mm. There are many similarities between the 850 hPa geopotential height fields in the early summer of the 1995/96 season and the late summer of the 1999/2000 season. In both instances the low over northern Namibia was dominant and the trough extended southwards over central and western South Africa. The IOH was located southeast of South Africa in 2000 allowing for

a more direct inflow of moisture from the Mozambique Channel over South Africa (Fig. 2-3E). At the 500 hPa level a trough was present west of South Africa extending northwards and dividing the high pressure over Namibia into 2 cells.

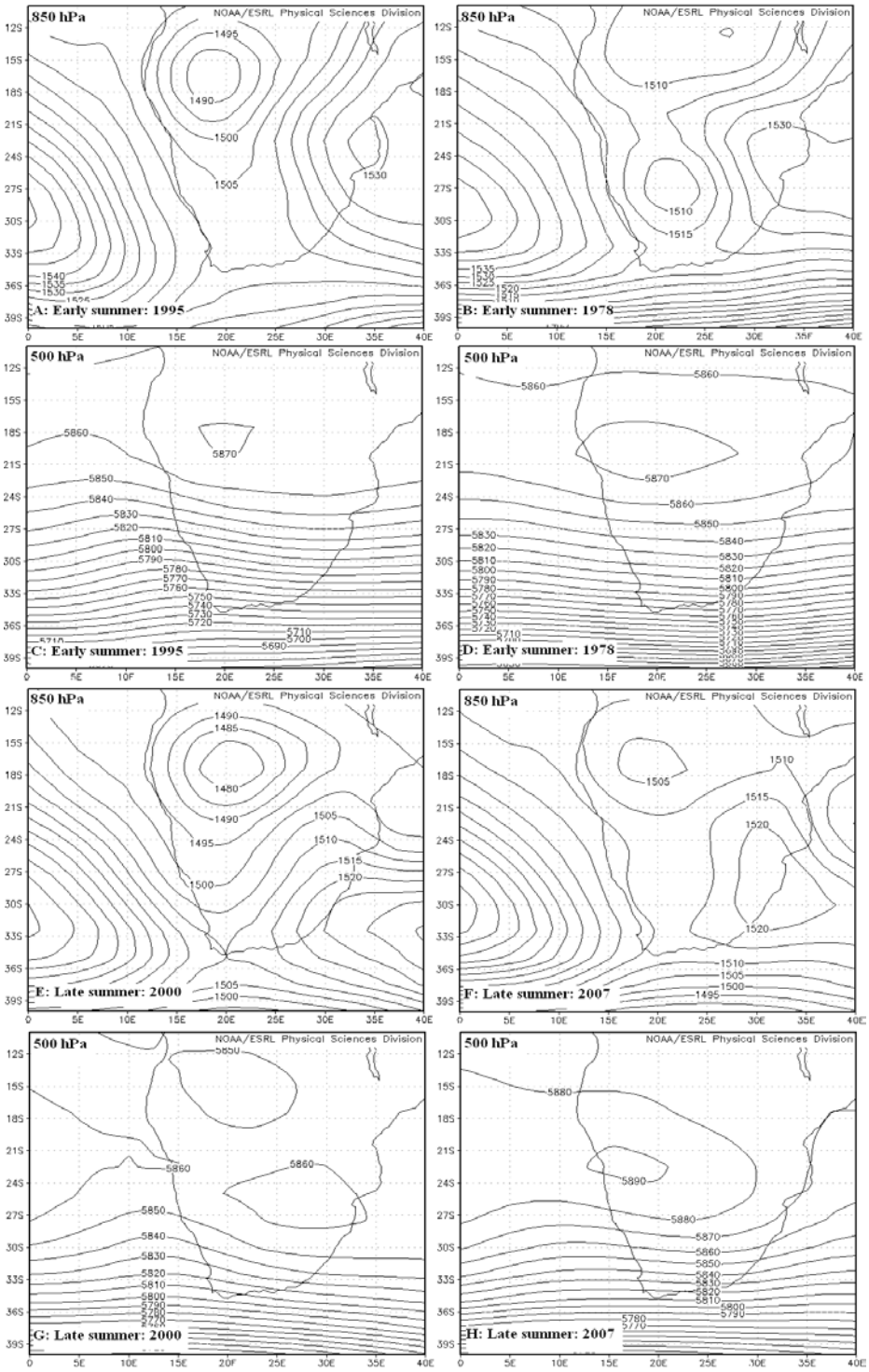


Figure 2-3: The average geopotential heights (in gpm) at 850 and 500 hPa for the early summers of the 1995/96 and 1978, seasons (A-D) as well as for the late summers of the 1999/2000 and the 2006/2007 seasons (E-H).

The eastern cell was located far to the southeast over southern Botswana and northern South Africa. The area average total late summer rainfall over Gauteng In 2007 was only 110 mm. There were no days when the area-average rainfall exceeded 25 mm and 50 mm at a single station was exceeded only on 4 d. In 2000 it happened on 46 days. At the 850 hPa level the high pressure was located over the eastern interior and although the low over Namibia was present (1 505 gpm) it was considerable weaker than in 2000 (1 480 gpm). A strong 500-hPa high (5 890 gpm) was present over Namibia (Fig. 2-3H).

2.2.9 Daily area-averaged heavy rainfall characteristics over Gauteng

The maximum area-averaged rainfall which occurred on any day over Gauteng during this period was 70 mm on 28 October 1986 (Table 2-2). The 2nd highest average daily rainfall occurred on 27 January 1978 (67 mm). The average daily rainfall exceeded 50 mm on only 5 d during this 32-yr period. Two of these events occurred in January months and 2 in March months. On 28 October 1986 the weather system over South Africa was what Taljaard (1996) defined as a southward extending V-shaped trough. There was a well established low pressure system over Botswana at 850 hPa advecting warm humid surface level air into Gauteng. At 500 hPa a weak trough was present over the southwestern parts of the country. All the other weather systems responsible for rainfall in excess of 50 mm over Gauteng were westerly troughs, cold cored in the upper troposphere, the only exception being 27 January 1978 when there was a continental low pressure system over Botswana (Dyson and van Heerden, 2002).

2.2.9.1 Daily area-averaged rainfall in January months

Table 2-2 presents the average number of days per month on which the areal average rainfall exceeds the thresholds of the 3 different heavy rainfall classes (‘significant rain event’, ‘heavy rain event’ and ‘very heavy rain event’). January months have the most ‘significant’ (3.8 d), ‘heavy’ (1.6 d) and ‘very heavy’ (0.5 d) rainfall days. Rainfall exceeding 10 mm occurred quite regularly during January as 94 % of the years had at least 1 such an event. Heavy rainfall occurred in 63 % of the Januaries (the same percentage of October and December months) and very heavy rainfall occurred in 34 % of the Januaries. The average number of significant rainfall days per month may give the weather forecaster an idea of how likely such an event is in any particular month. It may also be instructive to provide information on how often these events occur on consecutive days in any month. January was the month with the highest number of multiple heavy rainfall days for all 3 heavy rainfall classes. There were 11 significant rainfall days in 2006, 8 heavy rainfall days and 4 very heavy rainfall days in 1978. These values were not surpassed in any of the other summer months.

Rainfall exceeding 25 mm was not observed in a January on 2 days in a row but in 1978 there were 5 consecutive days with 'heavy rainfall' and 6 with 'significant rainfall'.

2.2.9.2 Daily area-averaged rainfall in February and March months

The heavy rainfall characteristics in February and March are similar to January. Significant rainfall occurred on average on 2.8 d in February and 2.6 d in March. These events were quite frequent as 91 % of the years had days with significant rainfall and multiple events had been regular in both months. February had the highest number of average heavy rainfall days (1.7 d) and these events occurred in 72 % of the years. In February 1996 there were 5 consecutive days when the average rainfall exceeded 15 mm. Very heavy rainfall occurred in approximately a quarter of February and March months.

2.2.9.3 Daily area-averaged rainfall in December months

December had the 2nd highest number of significant rainfall days (2.9 d) and these events occurred in 97 % of the years. However, December had the lowest average number of heavy rainfall (1 d) and very heavy rainfall (0.22 d) days. Very heavy rainfall is rare in December months as it occurred in only 16 % of the years.

2.2.9.4 Daily area-averaged rainfall in October and March months

The characteristics of heavy rainfall during October and November months are similar. Both months had on average just over 2 d per month with significant rainfall, 1.1 d with heavy rainfall and approximately 0.3 d with very heavy rainfall. Significant rainfall occurred in more than 80 % of the years and both months had 4 consecutive days with these events. There were less multiple days with significant, heavy and very heavy rainfall per month than in late summer. Very heavy rainfall occurred in only 19 % of October months and 28 % of November months.

There are on average 16.8 significant rainfall days per season. This is more than 100 d less than the number of days when the average daily rain exceeds 0 mm (Table 2-2). There are on average 7.8 heavy rainfall days and 2.16 very heavy rainfall days. In the very wet 1995/96 season there were 31 significant rain days 18 heavy rainfall days and 10 very heavy rainfall days.

2.2.10 Heavy rainfall at individual stations

The highest 24-h rainfall which was recorded at any rainfall station during this 32-yr period was 300 mm which occurred on 18 December 2006 at Viljoensdrift, located in southern Gauteng (Fig. 2-1). On 18 December 2006 the atmospheric circulation over South Africa was dominated by a deep surface trough, extending to a low off the southeast coast. There was a strong inflow of surface moisture into the central interior from Botswana and Zambia. At 500 hPa there was a high pressure

system located over northern Namibia causing southwesterly winds over Gauteng. The 2nd highest 24-h rainfall at a single station, 280 mm, occurred on 27 January 1978 at Rietondale in Pretoria. This was the same day which received the highest area-averaged rainfall and as mentioned earlier the weather system on this day was a continental tropical low pressure. There were a further 6 d when the maximum daily rainfall at any station exceeded 200 mm. Two of these events happened in October months (29 October 1995 and 25 October 2001) and 1 in March (19 March 2003). On these days a cut-off low pressure system was present over South Africa. The other 3 events were all caused by westerly troughs. The monthly maximum rainfall was less than 200 mm only in February months as the highest rainfall recorded during a February was 142 mm in 1996.

The average and maximum number of days when rainfall at an individual station exceeds the 3 different heavy rainfall classes is depicted in Table 2-3. The percentage of years in which these events occurred is also indicated on Table 2-3 as well as the maximum number of consecutive days of occurrence. The last row indicates the seasonal averages and extremes.

2.2.10.1 Single station significant rainfall

Rainfall exceeded 50 mm at a single station on nearly 28 d in a summer season. January months had the most days with single station significant rainfall events (7 d), followed by December and February with close to 5 d. In January 2006 there were 17 d when daily rainfall exceeded 50 mm for at least 1 station, and in January 2008 there were 8 consecutive days with single station significant rainfall over Gauteng. This was surpassed only in February of 1996 when 50 mm was exceeded at a single station on 9 consecutive days. The season with the highest number of single station significant rainfall days was the 1999/2000 summer season which had 56 d, 46 of these days occurring in late summer. The lowest seasonal total of significant single station rainfall was in 1985/86 with only 7 of these days.

Figure 2-2 shows that it was the late summer of the 1985/86 season that was particularly dry but during this time there were at least 3 d when rainfall exceeded 50 mm at a single station. Even the extremely dry 1982/83 and 1991/92 had 11 and 15 of these events respectively.

Significant rainfall at single stations occurs frequently, even during years which are considered to be dry. All January months had some days with these events and it occurred in 78 % of October months.

Table 2-3: Monthly and daily average rainfall statistics for rainfall at individual rainfall stations. The year of occurrence is shown in brackets and Sv indicates that it occurred in more than 2 years. The seasonal average values are shown in the last row and *indicates the minimum value per season.

		Single Station Significant Rainfall				Single Station Heavy Rainfall				Single Station Very Heavy Rainfall				Major Rain Event			Extreme Rain Event		
	Mx daily single station rainfall	No of days with single station rainfall > 50 mm		Percentage of years	Mx No. of consecutive days	No of days with single station rainfall > 75 mm		Percentage of years	Mx No. of consecutive days	No of days with single station rainfall > 115 mm		Percentage of years	Mx No. of consecutive days	Ave rain >10 mm, Station rain > 50 mm		Percentage of years	Ave rain >15 mm, Station rain > 75 mm		Percentage of years
	Mx	A	Mx	%		A	Mx	%		A	Mx	%		A	Mx	%	A	Mx	%
Oct	220 (24/'95)	2.3	10 ('93)	78	5 ('00)	0.9	6 ('00)	47	3 ('93)	0.25	3 ('00)	19	1 (Sv)	1.3	5 ('93,'00)	63	0.4	2 (Sv)	28
Nov	200 (26/'01)	3.5	10 ('01)	88	5 ('95,'97)	1.1	4 ('95)	66	2 ('85,'95)	0.09	2 ('95)	6	1 (Sv)	1.5	5 ('79,'89)	63	0.4	3 ('89,'95)	25
Dec	300 (18/'06)	5.1	10 ('01,'06)	97	5 ('06)	1.3	4 (Sv)	53	3 ('93,'95)	0.25	2 ('86)	22	1 (Sv)	1.8	5 ('86)	84	0.4	3 ('86)	25
Jan	280 (27/'78)	7.0	17 ('06)	100	8 ('08)	2.6	7 ('00,'06)	88	5 ('96)	0.59	3 '(96)	38	2 ('96)	3.1	11 ('06)	84	1.0	5 ('78)	56
Feb	142 (26/'96)	5.1	17 ('00)	97	9 ('96)	1.9	10('00)	69	6 ('96)	0.28	5 '(00)	13	3 ('00)	2.3	9 ('96)	78	1.1	5 ('96,'00)	50
March	260 (19/'03)	4.6	15 ('00)	94	7 ('00)	1.7	10 (97)	63	5 ('91)	0.38	4 ('97)	22	2 ('95)	2.1	8 ('97)	81	0.8	5 ('97)	44
Season		27.6	56('99/'00) 7*('85/'86)	100		9.5	24 ('99/'00) 0* ('83/'84)	97		1.84	9 ('95/'96)	66		12.1	23 ('95/'96)	100	4.1	13 ('95/'96)	88

2.2.10.2 Single station heavy rainfall

Rainfall exceeding 75 mm at a single station (single station heavy rainfall) occurred infrequently with the average monthly occurrences being less than 2 d for all summer months, except for January (2.6 d). January months had at least 1 such event in 88 % of the years, while in December these events occurred in about than half of the years (53 %). February 2000 and March 1997 had a maximum of 10 of these days in a single month. During February 1996, there were 6 consecutive days when the rainfall exceeded 75 mm. Single station heavy rainfall events occur on nearly 10 d a season. There were 24 of these days during the 1999/2000 summer and not a single day during the 1983/84 season. Fig. 2-2 shows that the average early summer rainfall of the 1983/84 season was above normal as 377 mm was recorded over Gauteng from October to December. There were 69 d with rain during these 3 months; however, on none of these days did the rainfall at a single station exceed 75 mm.

2.2.10.3 Single station very heavy rainfall

Single station rainfall exceeding 115 mm is very rare over Gauteng. During the 32-yr period, rainfall at any station exceeded 115 mm on only 59 d. On average these events occur on less than 1 d in each of the summer months. January months have the highest average value (0.59 d) but only 38 % of the years recorded at least 1 of these events. There were only 19 d in all the January months where rainfall at a single station exceeded 115 mm. Rainfall at a single station exceeded 115 mm on 3 d during October 2000. However, these events were rare in October as it occurred in only 19 % of the years. The events were most infrequent in November months as only 6 % of the years received some of these events. There were only 2 Novembers, 1995 and 2001, which had days with very heavy rainfall at a single station. Bearing in mind the areal average rainfall as depicted in Table 2-2, February had the 2nd highest number of very heavy rainfall days. However, Table 2-3 shows that very heavy rainfall at a single rainfall station during February months is very rare, as only 13 % of the years had days when the rainfall at a single station exceeded 115 mm. Considering that there were only 59 d during this 32-yr period which had rainfall of more than 115 mm at a single station, the unique rainfall character of the 1995/96 rainfall season is again emphasised (Table 2-3) as there were 10 'single station very heavy rainfall' days during this 1 season alone – 5 d in early summer and 5 in late summer.

2.2.11 Location of heavy rainfall at individual stations

The number of stations which recorded more than 75 mm and 115 mm was calculated for the 8 grid boxes over Gauteng for the entire period. These results are depicted in Fig. 2-4 and are expressed as the total number of events per 100 stations. Single station significant and heavy rainfall

events occur most frequently over the central and northwestern parts of Gauteng. The main watershed in the Witwatersrand which divides the province into 2 major catchments, the Crocodile catchment to the north and the Vaal catchment to the south, lies between 26˚S and 26˚15'S. The 3 grid boxes located south of this watershed receive only half the number of single station heavy and very heavy rainfall events when compared with those grid boxes further to the north. Consider Fig. 2-3 and the average 850 hPa geopotential height fields in the very wet early summer of the 1995/96 season (Fig. 2-3A) and late summer of the 1999/2000 season (Fig. 2-3E). These maps illustrate that during seasons which receive a high number of heavy and very heavy rainfall days there is a deep surface low pressure system over northern Namibia and an IOH located east or southeast of South Africa. This high pressure causes an inflow of warm moist tropical air from the Mozambique Channel which curves around the high pressure system and enters Gauteng from the north. The moisture-laden air is then forced to rise against the Witwatersrand resulting in a higher number of heavy rainfall events at the stations north of the watershed than further to the south.

2.2.12 Major and extreme rain events

The last 2 columns of Table 2-3 depict the results for the major (rainfall at a single station exceeded 50 mm and the average rainfall exceeded 10 mm) and extreme (rainfall at a single station exceeding 75 mm and the average rainfall exceeding 15 mm) rain events. On average January months have 3.1 major rain events per year followed by February which has 2.3 events. October months have the lowest number of events at 1.3 events. There is an average of 12.1 of these events per summer season. The maximum number of days with major rain events occurred in 1995/96 which had 23 such events. The maximum number of these events occurred in January 2006 with 11 d, followed by 9 d in February 1996. Major rain events occur relatively frequently as 84 % of all December and January months had at least 1 of these events compared to only 63 % of October and November months. Extreme rain events occur much less frequently. Fifty-six percent of January months received extreme rain events, 50 % of February months and 44 % of March months. January and February on average both have approximately 1 d with major rain events and all the early summer months 0.4 d. The late summer months all had a maximum of 5 of these events.

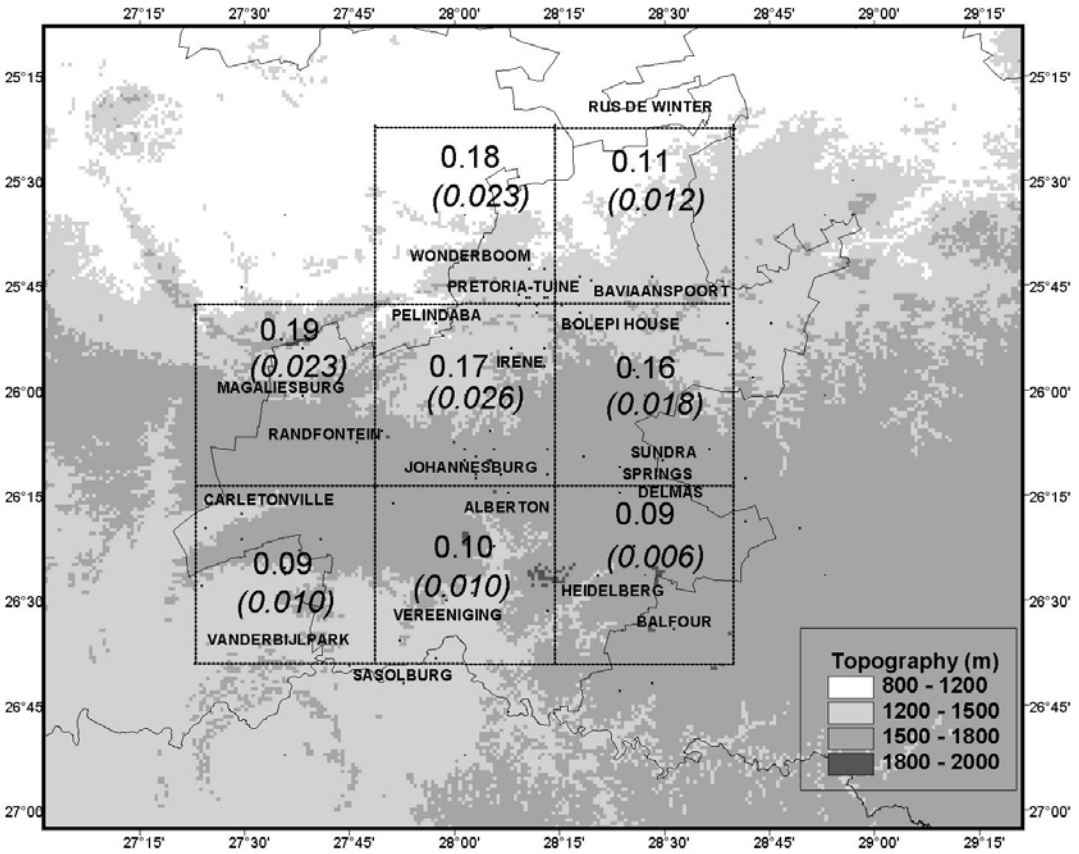


Figure 2-4: The location of eight 0.5˚ by 0.5˚ grid boxes over Gauteng. The values show the total number of rainfall stations which received more than 75 mm and 115 mm (in brackets) in each grid box. The values are standardized to 100 stations per grid box. The shading indicates height (in m) a. m. s. l.

2.2.13 Discussion

January months have the highest average monthly rainfall; January had, on average, 23 d with rain and January also had the most significant, heavy and very heavy rainfall days. Approximately 20 % of all significant, heavy and very heavy rainfall events occurred in January months. Every single January of this 32-yr period had at least 1 d with rainfall of more than 50 mm and there were in total 19 d in January where the rainfall at an individual station exceeded 115 mm. This is a third of the total number of these events.

December had the 2nd highest average monthly rainfall and on average receives 24 d with some rain. December had approximately 3 d per year when the average Gauteng rainfall exceeded 10 mm. However, December had the lowest number of heavy and very heavy rainfall days. The highest 24-h rainfall at any station (300 mm) did occur in a December month but, considering the

rainfall at individual rainfall stations, 115 mm was exceeded on only 7 d during December months and this occurred in about 20 % of all the years.

The average monthly rainfall in February was lower than in December and it had, on average, only 19 d with rain. The average monthly rainfall during February months had a greater variance than in December months as depicted by the standard deviation in Table 2-2. There were fewer days with some rain than in December but more heavy and very heavy rainfall days. However extreme rainfall at an individual station was rare in February. There were only 6 d when the rainfall exceeded 115 mm. This seems to suggest that heavy and very heavy rainfall events during February months are associated with widespread rainfall over the entire province rather than copious rainfall at a single station.

October and March months may be considered the transition months as summer starts in October and ends in March. On average March receives more rainfall than October and also had slightly more days with rain. Heavy rainfall events occur on average approximately on 1 d per month in both months and very heavy rainfall on about 0.35 d per month. The highest average daily rainfall was recorded in an October month at 69 mm. There were on average more days with rainfall in excess of 50 and 75 mm at a single station in March than in October.

Heavy and very heavy rainfall events occur more frequently in late summer, when 60 % of these events were recorded. The same distribution is also present in the rainfall at individual stations. Sixty percent of days with more than 50 mm at a station, 65 % of the days with more than 75 mm, and 63 % of those days with more than 115 mm occur in late summer. Rainfall stations located over the central and northwestern part of Gauteng receive rainfall in excess of 75 and 115 mm more frequently than those in the south and southeast. On 7 of the 8 days when rainfall at a station surpassed 200 mm the atmospheric circulation was dominated by cold upper tropospheric temperatures and troughs or lows in the westerly circulation. On the remaining day a continental tropical low was situated over Botswana.

Whenever the rainfall over Gauteng exceeds 0 mm, the percentage of rainfall stations reporting some rainfall is 40 % on average. For significant, heavy and very heavy rainfall events the percentage of rainfall stations reporting more than 0 mm exceeds 80 %. These events happen when most of Gauteng receives some rainfall and may be classed as widespread heavy rainfall events. However, when the rainfall at an individual rainfall stations exceeds 50 mm (the criteria used by the forecasters to issue warnings of heavy rainfall) the average percentage of rainfall stations reporting some rainfall is approximately 50 %. Heavy rainfall at single stations is therefore an isolated event which does not necessarily occur on days when the entire province receives some rainfall.

2.2.14 Acknowledgements

The author would like to express her sincere appreciation to Colleen de Villiers from the South African Weather Service for the rainfall data. Francois Engelbrecht is acknowledged for his many useful suggestions to enhance this paper and Christien Engelbrecht for the interest shown and many fruitful discussions. The South African Water Research Commission Project No. K5/1333 deals with the early warning of heavy rainfall and its support is acknowledged. The author also acknowledges the 2 anonymous reviewers, whose comments helped to clarify and improve the manuscript.

2.2.15 References

BEAVON KSO (2004). *Johannesburg: The Making and Shaping of the City*. UNISA Press, Pretoria. 373 pp.

BRADLEY AA and SMITH JA (1994). The hydrometeorological environment of extreme rainstorms in the southern plains of the United States. *Journal of Applied Meteorology.* **33,** 1418-1431.

BROOKS HE and STENSRUD DJ (2000). Climatology of heavy rain events in the United States from hourly precipitation observations. *Monthly Weather Review,* **128** (4), 1194-1201.

CHEN C, CHEN Y, LIU C, LIN P and CHEN W (2007). Statistics of heavy rainfall occurrences in Taiwan. *Weather Forecasting* **22** (5), 981-1002.

CHEN GT and YU C (1988). Study of level jet and extremely heavy rainfall over northern Taiwan in the Mei-Yu season. *Monthly Weather Review,* **116,** 884-891.

DE CONING E, FORBES GS and POOLMAN EP (1998). Heavy precipitation and flooding on 12-14 February 1996 over the summer rainfall regions of South Africa: Synoptic and isentropic analyses. *National Weather Digest,* **22** 25-36.

DYSON LL and VAN HEERDEN J (2001). The heavy rainfall and floods over the northeastern interior of South Africa during February 2000. *South African Journal of Science,* **97**, 80-86.

DYSON LL and VAN HEERDEN J (2002). A model for the identification of tropical weather systems. *Water SA,* **28**(3), 249-258.

DYSON LL, VAN HEERDEN J and MARX HG (2002). Short-term forecasting techniques for heavy rainfall. WRC Report No. 1011/1/02. Water Research Commission, Pretoria, South Africa.

ENGELBRECHT FA, McGREGOR JL and ENGELBRECHT CJ (2009). Dynamics of the Conformal-Cubic Atmospheric Model projected climate-change signal over South Africa. *International Journal of Climatology,* **29** (7), 1013-1033.

FRAGOSO M and TILDES GOMES T (2008). Classification of daily abundant rainfall patterns and associated large-scale atmospheric circulation types in Southern Portugal. *International Journal of Climatology,* **28** (4), 537-544.

GAUTENG DEPARTMENT OF HOUSING (2006). Gauteng Housing Annual Report 2006. Gauteng Department of Housing, Marshalltown.

HCIGIZOGLU HK, BAYAZIT M, and ÖNÖZ B (2005). Trends in the maximum, mean, and low flows of Turkish rivers. *Journal Hydrometeorology,* **6,** 280-290.

HOUZE RA, SMULL BF and DODGE P (1990). Mesoscale organization of springtime rainstorms in Oklahoma. *Monthly Weather Review,* **118,** 613-654.

KRUGER AC (2004). *Climate of South Africa. Climate Regions. WS45*. South African Weather Service, Pretoria, South Africa

MADDOX RA, CHAPPELL CF and HOXIT LR (1979). Synoptic and meso-α scale aspects of flash flood events. *Bulletin of the American Meteorological Society,* **60,** 115-123.

MARX HG, DYSON LL and VAN HEERDEN J (2003). Verification of rainfall forecasts for the Vaal Dam catchment for the summer rainfall seasons of 1994 to 1998. *Water SA,* **29,** 195-200.

NGUBO N, MALALA R, MABUSELA X, DE CONING E, ROSENFELD D and KERKMANN J (2008). Hailstorm in Potchefstroom area (South Africa, 27 October 2004). URL: http://www.eumetsat.int/Home/index.htm (Accessed on 14 November 2008).

RAE KJ (2008). Personal communication. Chief Forecaster, National Forecast Centre, South African Weather Service, Bolepi House, Pretoria.

ROUAULT M and RICHARD Y (2003). Intensity and spatial extension of droughts in South Africa at different time scales. *Water SA,* **29**(4), 489-500.

SAWS (SOUTH AFRICAN WEATHER SERVICE) (1998) *Climate of South Africa. Climate statistics 1961-1990. WB42.* South African Weather Service, Pretoria, South Africa.

SCHUMACHER RS and JOHNSON RH (2006). Characteristics of U.S. extreme rain events during 1999-2003. *Weather Forecasting,* **21**(1), 69-85.

SINGLETON AT and REASON CJC (2007). Variability in the characteristics of cut-off low pressure systems over subtropical southern Africa. *International Journal of Climatology,* **27**(3), 295-310.

TALJAARD JJ (1996). Atmospheric Circulation Systems, Synoptic Climatology and Weather Phenomena of South Africa. Part 6: Rainfall in South Africa. South African Weather Service Technical Paper 32. South African Weather Service, Pretoria, South Africa.

VIVIERS G and CHAPMAN V (2008). Severe thunderstorm at Johannesburg International Airport (03 December 2004). URL: http://www.eumetsat.int/Home/index.htm (Accessed on 14 November 2008).

ZHANG X, HOGG WD and MEKIS E (2001). Spatial and temporal characteristics of heavy precipitation events over Canada. *Journal of Climate,* **14,** 1923-1936.

2.3 Overview

The monthly average Gauteng rainfall for all summer months from October 1977 to March 2012 is given in Table 2-4. It is to large extent a repetition of Table 2-1 but here the monthly average rainfall is provided and it includes data for the 2010/11 and 2011/12 rainfall seasons.

The 2010//11 summer rainfall season was characterised by several flood events over Southern Africa. During December and January a number of flash flood events occurred over Gauteng while the Orange River was in flood in January 2011 (SAWDIS, 2012). The normalised rainfall indexes over Gauteng in Fig. 2-5 show that the rainfall in early summer (October-December) and late summer (January-March) was above normal during the 2010/11 season. Some insight into the monthly distribution of the rainfall during 2010/11 rainfall season is provided in Table 2-4.

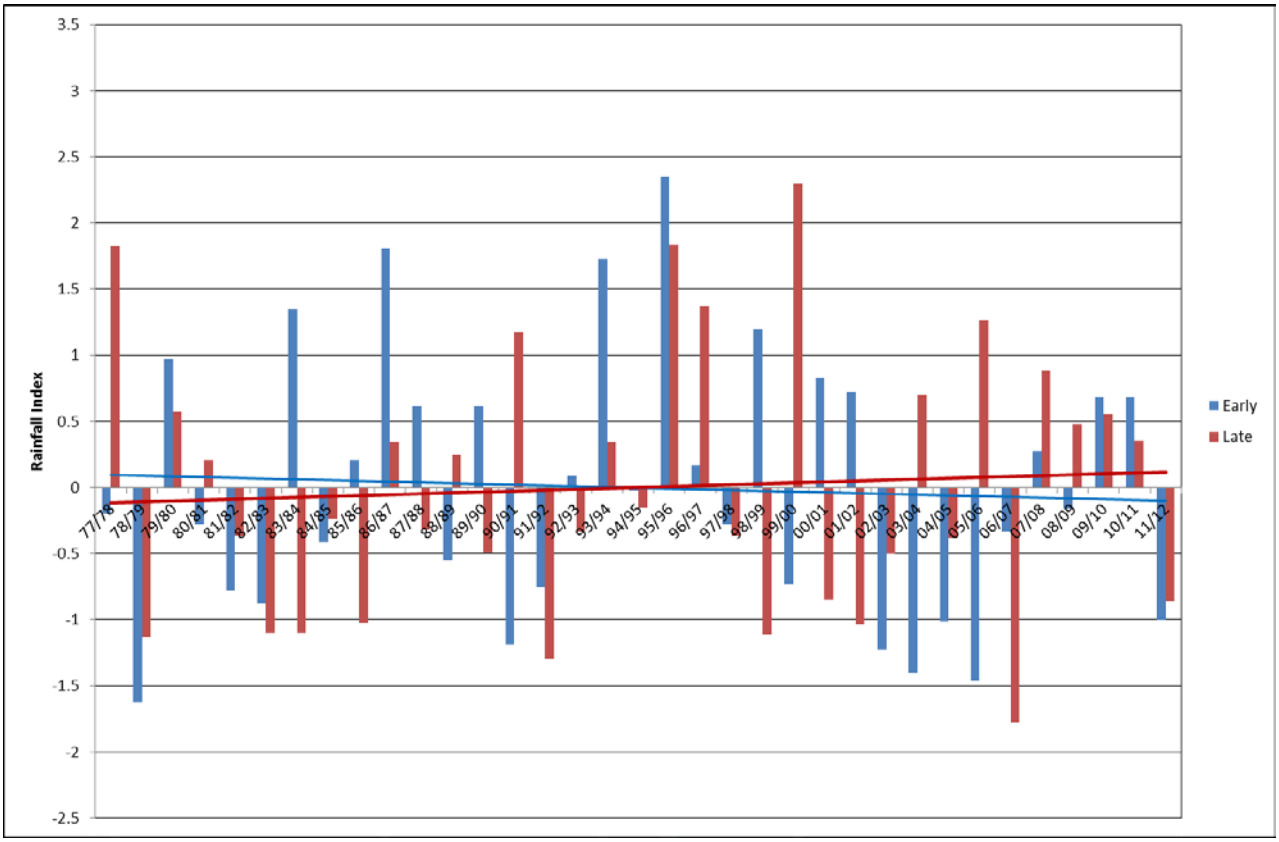


Figure 2-5: Standardized rainfall index for early summer (blue) and late summer (brown) for 1977 to 2012. The early summer trend line is shown in a blue line while the solid brown line is the late summer trend line.

Table 2-4: Monthly rainfall totals (mm) for Gauteng for the summer months of 1977 to 2012. They grey cells indicate the highest monthly average rainfall for any particular month and the black cells the lowest average monthly rainfall. The long term average monthly rainfall (Ave) and standard deviation (Std) are shown in the last two rows.

Season	Oct	Nov	Dec	Jan	Feb	Mar	Total
77/78	78	92	94	324	118	73	779
78/79	67	40	52	75	46	61	341
79/80	96	167	87	190	131	53	724
80/81	19	157	82	135	126	72	591
81/82	60	64	97	157	48	63	489
82/83	68	46	100	91	49	46	400
83/84	85	142	151	61	32	93	564
84/85	93	91	64	137	71	75	531
85/86	104	62	128	83	47	64	488
86/87	132	110	170	120	86	142	760
87/88	59	153	112	88	57	129	598
88/89	86	56	96	98	203	37	576
89/90	36	183	105	58	86	110	578
90/91	46	35	110	176	105	160	632
91/92	54	57	112	70	68	26	387
92/93	33	133	119	90	93	89	557
93/94	190	96	120	136	140	72	754
94/95	63	75	139	87	69	136	569
95/96	104	169	179	183	277	56	968
96/97	89	79	123	113	55	296	755
97/98	37	116	105	126	91	51	526
98/99	64	164	139	85	45	54	551
99/00	25	51	149	166	241	161	793
00/01	135	106	99	56	106	52	554
01/02	111	124	97	84	74	35	525
02/03	64	17	107	110	83	60	441
03/04	53	72	50	110	157	121	563
04/05	23	64	117	159	54	54	471
05/06	15	97	59	218	165	69	623
06/07	40	94	120	63	25	22	364
07/08	131	59	109	212	59	138	708
08/09	50	114	103	175	106	82	630
09/10	77	115	137	205	88	79	701
10/11	23	86	220	191	51	107	678
11/12	59	73	73	83	74	56	418
Ave	**71**	**96**	**112**	**129**	**95**	**86**	**588**
Std	**39**	**43**	**35**	**59**	**57**	**52**	**139**

The rainfall during October and November 2010 over Gauteng was well below the monthly mean with only 23 mm occurring in October 2010. The long-term mean rainfall in December over Gauteng is 112 mm, December 2010 received close to 200 % of the normal rainfall when 220 mm

occurred. This was the highest rainfall for a December month for the entire period under investigation (October 1977 to March 2012), the previous maximum occurred in 1996 (179 mm). However, despite the copious rainfall in December 2010 the early summer rainfall in the 2010/11 rainfall season was exactly the same as the 2009/10 early summer rainfall. (329 mm occurred in these 3 months in both years). The shaded cells in Table 2-4 shows the highest rainfall recorded in that particular month for the period 1977 to 2012 and even though the rainfall for January 2011 was also well above normal (191 mm) the highest rainfall for a January month occurred in 1978 when 324 mm occurred. The February 2011 rainfall was only 51 mm, well below the long-term mean. The extreme rainfall events and associated flooding over Gauteng was mainly confined to the months of December and January. December months usually have many days with rain over Gauteng (24 on average in Table 2-1), in December 2010 there were 27 d with rain and 23 d in January 2011. There were 6 d in December 2010 when the average daily rainfall over Gauteng exceeded 10 mm and 7 d in January 2011. Both months also had 2 days with extremely heavy rainfall (average daily rainfall exceeding 25 mm). During the 2010/11 rainfall season there was approximately 40 rainfall stations available daily for analysis over Gauteng. There was a total number of 78 rainfall stations during this season where the rainfall exceeded 50 mm, 45 of these events occurred in December, 5 in January and 18 in March months. On the 15th and 16th of December alone there were 31 stations where the rainfall exceeded 50 mm over Gauteng. If the rainfall of these two days did not occur, the December 2010 rainfall would have been close the long-term average December value. The influence of a single rainfall event lasting two days was responsible for the well above normal rainfall in December 2010.

In contrast the 2011/12 summer rainfall over Gauteng was below normal during early and late summer months (Fig. 2-5). The monthly average rainfall over Gauteng was less than the long-term average monthly value for all summer months (Table 2-4). There were only 14 d with rain in December 2011, half of the average number of rain days. During the entire season there were only 31 stations where the rainfall exceeded 50 mm. The same number of stations had more than 50 mm on 15 and 16 December 2010 alone.

Fig. 2-5 shows the early and late summer trend lines. The trends identified in Fig. 2-2 persist in Fig. 2-5, with a general decrease in early summer rainfall and an increase in late summer rainfall. However, the the nonparametric Mann-Kendall test shows that these trends are statistically not significant.

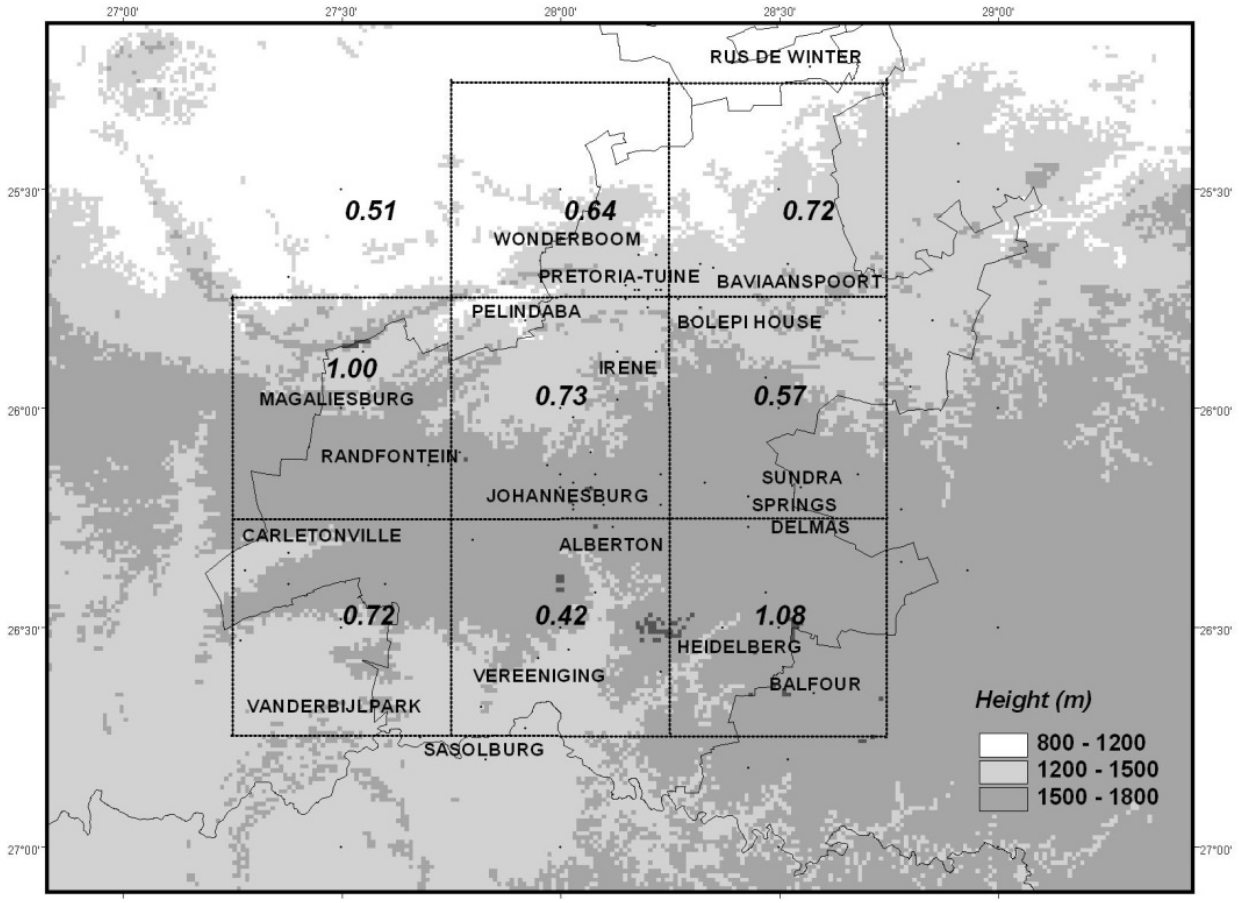


Figure 2-6: The total number of rainfall stations which received more than 50 mm in each grid box for October 2010 to February 2012. The values are standardized to 100 stations per grid box. The shading indicates height (in m) a. m. s. l.

Fig 2-6 shows the percentage of rainfall stations exceeding 50 mm in 0.5° grid blocks over Gauteng for the summer months from October 2010 to February 2012. In the discussion of Fig. 2-4 it was emphasised that those grid boxes over the northwestern parts of Gauteng receive the highest number of heavy rainfall events. In the 2010/11 and 2011/12 rainfall seasons and for single heavy rainfall (SHR) class the grid box over southeastern Gauteng had the highest number of heavy rainfall events. The grid box over western Gauteng received the second highest number of SHR events, followed by central Gauteng. During these two seasons the northern parts of the province still had a relatively high number of SHR events but the southeastern and northeastern grid boxes had more events than the long-term mean values per grid box.

References

SAWDIS (South African Weather and Disaster Information Service) (2012). URL: http://saweatherobserver.blogspot.com/ . (Accessed on 23 Augustus 2012)

3 Synoptic circulation patterns associated with heavy rainfall over Gauteng

3.1 Preface

This Chapter describes the synoptic circulation associated with heavy rainfall over Gauteng. A monthly synoptic climatology is provided first and is then compared to the monthly composite fields complied from days with heavy rainfall and days which remain dry. In Chapter 2 the Gauteng daily rainfall were calculated for a 35-year period and heavy rainfall and dry days were identified and are used here to construct the average monthly maps. Lastly horizontal wind divergence and relative vorticity advection is calculated over Gauteng and vertical profiles of these variables are discussed on heavy rainfall and dry days. In Chapter 4 the thermodynamical profile of the atmosphere is investigated during heavy rainfall at a single location (Irene) but understanding the properties of the larger environment in which heavy rainfall develops is explored in Chapter 3.

3.2 Background

Taljaard (1996) provides a thorough description of the character of rainfall in South Africa. Roughly 85 % of SA receives more than half of its rainfall in summer (October-March). The 50 mm isohyet advances westward from KwaZulu-Natal in August to the Northern Cape Province in March. Gauteng falls within the summer rainfall area and on average receive more than 50 mm of rain in all summer months. Taljaard (1996) stressed the importance of abundant moisture in the production of rainfall and explained how in summer months at surface levels moisture is advected into the interior of South Africa by a high pressure system centred close to Maputo. The Indian Ocean is the primary source of moisture in summer but the southward penetration of air from the ITCZ also plays a significant role in the production of rainfall over Gauteng. For rainfall and indeed heavy rainfall to occur over the interior of South Africa humid tropical maritime air should invade the subcontinent from the east. However, this penetration is inhibited by the semi-permanent anti-cyclone from 700-200 hPa located over the subcontinent. The humid air is injected into the interior when ridging anti-cyclones reach the east coast or as the continental high moves, weakens or strengthens.

The most important weather systems producing rainfall in summer in South Africa were identified by Taljaard (1996) and Tyson and Preston-White (2000) as

- The westerly trough (or TTT)
- The cold front trough
- The southward extending tropical low or V-shaped trough
- The Cut-off low (COL)

- Tropical cyclones (TC)
- The ridging high
- The blocking high or long wave ridge

Not all of these weather systems cause heavy rainfall over South-Africa and only those weather systems that have been shown to be associated with heavy rainfall in summer will be expanded on here. The TTT is a wave in the west-wind regime which connects tropical convective systems to mid-latitude weather systems and is clearly visible as elongated cloud bands on satellite imagery which extends southeastward (Hart et al., 2010). The TTT is part of an important process which causes humid tropical air to flow from the ITCZ into South Africa. Moist tropical air has a large potential for instability as indicated by Convective Available Potential Energy (CAPE) CAPE is dealt with in detail in Chapter 4.3.5. Williams and Reno (1993) calculated CAPE values to be in the order of 2000 - 5000 Jkg^{-1} over much of the tropics. The TTT therefore enhances the potential for convective development over South Africa by causing the inflow of air with greater potential for convective instability. The TTT may cause heavy rainfall over the interior of South Africa as was illustrated by Hart et al. (2010). They stressed the importance of local convection in the production of the heavy rainfall. Furthermore De Coning et al. (1998) identified that a TTT was partially responsible for the very heavy rainfall over the central interior of South Africa during February 1996.

Taljaards' (1996) southward extending tropical low or V-shaped trough is also referred to as a continental tropical low pressure system (Dyson and van Heerden, 2002). This low pressure system is characterised by a low pressure system which stands upright form the surface to 400 hPa but is replaced by a high pressure system in the upper troposphere. The low pressure system is warm cored from 500 hPa upwards but with relatively cool temperatures in the surface to 700 hPa layer. The surface dew point temperatures associated with these systems are in the order of 18-20 °C with precipitable water values of 20 mm. Some noteworthy and very heavy rainfall events were caused by these tropical weather systems for example the Free State floods of February 1988 (Triegaardt et al., 1991) as well as the heavy rainfall and floods over northeastern South-Africa in February 2000 (Dyson and Van Heerden, 2001). Part of the reason for the flooding which may be associated with these weather systems is that they move very slowly- causing heavy rainfall over the same area on consecutive days. In February 1988 heavy rainfall occurred over the central Free State for 3 consecutive days (19-22 February). The convective cells which develop in association with tropical low pressure systems are semi-stationary and associated with little or no vertical wind shear.

Cut-off lows (COLs) are truly significant weather systems of the extra-tropics and are responsible for wide spread rainfall over South Africa and 20 % of COLs produce heavy rainfall which may produce floods (Taljaard, 1985). Most COLs have life-spans of longer than 2 days but may influence the rainfall over South Africa for up to 5 days (Singleton and Reason, 2007a). All together COLs influence the weather over South Africa about 40 days per annum. COLs occur most frequently in March/April and September/October months when approximately 2 systems occur per month. They are least frequent in January and July when on average only 1 system occur per month. These systems are recognisable on synoptic maps by the following:

- Low pressure in the middle troposphere is cut-off from the general westerly circulation and lies equator ward of westerly flow
- Closed cyclonic circulation throughout the troposphere
- Cold cored in the middle troposphere.

These weather systems are clearly baroclinic in nature and heavy rainfall from COLs occur when they are in a development stage. The atmosphere is baroclinic when the atmospheric density is a function of both temperature and pressure and a thermal wind therefore exists (Holton, 1992). A thermal wind indicates vertical wind shear and a COL will be classed as developing when the low in the middle and upper troposphere lies westward of the surface low. Under these conditions vertical wind shear will be present in the atmosphere which should be visible on upper air sounding data over the summer rainfall area (such as at Irene). Singleton and Reason (2007b) described the heavy rainfall associated with a COL over South Africa during March 2003. They concentrate their study on the south coast and adjacent interior but Gauteng also received heavy rainfall during this period (on the 23rd of March 2003 the average rainfall over Gauteng was 11 mm). The COL was lying westward with height on the 23rd when the highest rainfall total occurred over Gauteng for the days under investigation in this study. On the 24th there was a decrease in the vertical wind shear over Gauteng and the average rainfall over Gauteng was only 1 mm.

There are on average 11 tropical depressions over the South West Indian Ocean per year (Jury and Pathack, 1991) but these systems do not make landfall over the southern African subcontinent every year (Malherbe et al., 2012). The normal displacement of tropical cyclones (TC) is to the southwest, south and eventually the southeast when they reach 25° S (Taljaard, 1996). When a system makes landfall it may either move further westward into the interior of southern Africa or it may deflected to the south or north. The influence of tropical cyclones on the rainfall over the eastern interior can be quite contrasting depending on the location of the TC. If centred at 500 km or more from the east coast subsidence results over the eastern parts of the subcontinent with a deep

southerly to southeasterly flow. However, when these systems make landfall torrential rain may fall especially from the coastal belt up to the escarpment. TC Domonia made landfall close to Maputo on 29 January 1984 resulting in 24-h rainfall totals of more than 200 mm over the eastern escarpment (Poolman and Terblanche, 1984). But there was no rainfall over the Gauteng during this event. On the 17th of January 2012 TC Dando made landfall south of Maputo and moved in a northwesterly direction over the subcontinent while weakening (Meteo France, 2012). Daily rainfall in excess of 200 mm occurred over the escarpment of Mpumalanga on the 17th of 18th of January 2012. Gauteng also received no rainfall during this event. Conversely TC Eline made landfall north of Beira, on the Mozambique coast, on 22 February 2000 and tracked slowly westward over Zimbabwe and Botswana during the next few days. More than 100 mm of rain occurred at some stations over Gauteng on the 22nd and about 50 mm on the 23rd and 24th. It is clear that the position of the TC plays a governing role on the amount of rainfall over Gauteng.

The ridging high and blocking high pressure systems contribute to the development of rainfall over the summer rainfall area of South Africa by aiding in the advection of moisture from the Indian Ocean to interior plateau of South Africa. They are not directly responsible for heavy rainfall although if the blocking high occurs in conjunction with a COL, heavy rainfall may result as was the case in September 1988 (Triegaardt el al., 1991).

On the synoptic-scale upward vertical motion, required for the formation of precipitation, is associated with horizontal wind convergence and the advection of relative vorticity (Holton, 1992). The continuity equation (with pressure as vertical coordinate) (equation 3-1) provides the relationship between the change of vertical velocity with pressure (left hand side of equation) and horizontal wind divergence (right hand side of equation).

$$\frac{\partial \omega}{\partial p} = -\overline{\nabla}_H \bullet \overline{V} \qquad \textbf{3-1}$$

where

ω=vertical velocity in Pa s^{-1}

p= pressure

$$\overline{\nabla}_H = \frac{\partial}{\partial x}\bar{i} + \frac{\partial}{\partial y}\bar{j}$$

$$\overline{V} = u\bar{i} + v\bar{j} \qquad \text{(horizontal wind vector)}$$

When wind convergence occurs ($\overline{\nabla}_H \bullet \overline{V} < 0$) it implies that the vertical velocity increases with pressure or decreases with height. If horizontal wind convergence takes place through a deep vertical layer in the atmosphere the upward vertical velocity will increase in that same layer enhancing the probability of rainfall if adequate moisture is available.

Classical meteorological theory as (for example) described as the quasi-geostrophic theory by Holton (1992) provides a relationship between vertical motion and the advection of relative vorticity (equation 3-2). Assuming that weather systems occur on the synoptic-scale and in the extra tropics, cyclonic vorticity advection in the upper troposphere in the southern hemisphere will result in upward motion of air.

Relative vorticity advection can be approximated by (Holton ,1992):

$$\overline{V} \bullet \overline{\nabla}\xi \qquad \textbf{3-2}$$

where

$$\xi = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$$

Irrespective of the type of synoptic circulation over Gauteng, if upper tropospheric wind divergence and cyclonic vorticity advection exists then upward motion in the mid troposphere results and provided that adequate near surface water vapour is available precipitation will occur. These factors in favourable conjunction with some thermodynamic properties, detailed in this thesis, are associated with heavy rainfall over Gauteng Province.

3.3 Data and Method

The daily rainfall over Gauteng were analysed for all summer months from 1977 to 2012 and days with heavy rainfall were identified as described in Chapter 2. Two heavy rainfall classes are considered namely single station heavy rainfall (SHR) and area-average heavy rainfall (AHR) which is the 90th percentile of the daily area-average and single station rainfall events. National Centre for Environmental Prediction (NCEP) data used to investigate the synoptic circulation were found to be unreliable to accurately recreate atmospheric fields prior to 1979 due to the absence of observational data in the Southern Oceans (Tennant, 2004). The analysis preformed here will

therefore only include data from 1979-2012. Table 3.1 lists the number of the AHR and SHR events per month from 1979-2012 as well as the number of days when no rain (No Rain) occurred over Gauteng. January months had the most days with AHR and SHR, followed by December and February months. There were only 63 days in December months when no rainfall occurred over Gauteng, this is just 6 % of all December days. The number of No Rain days was less than AHR and SHR in December, January and February months.

Table 3-1: Number of area-average heavy rainfall (AHR), single station heavy rainfall (SHR) and No Rain days per month.

	Number of AHR events	Number of SHR events	Number No Rain events
October	62	83	250
November	86	126	120
December	102	179	63
January	128	233	79
February	91	168	135
March	86	159	222

NCEP reanalysis data (Kalney et al., 1996) were used to investigate the monthly synoptic circulation associated with AHR, SHR and No Rain days. NCEP data have a horizontal resolution of 2.5 degrees and have 17 levels in the vertical. Variables utilized were geopotential height, the zonal and meridionial winds and mixing ratio. The average height of the interior of South Africa is approximately 1500 m above mean sea level and it is therefore common practice in the forecasting offices in South Africa to use the geopotential heights of the 850 hPa level to represents the surface synoptic circulation; this is also done here. Only 8 of the 17 vertical levels available in the NCEP data set are utilized namely the 850, 700, 600, 500, 400, 300, 250, and 200 hPa levels. Horizontal and vertical resolution of the NCEP data are ideal to research weather systems on the synoptic-scale but small or mesoscale features will not be adequately identified by this data set.

NCEP data were first used to calculate the long-term monthly mean circulation over Southern Africa. The average monthly circulation for AHR, SHR and No Rain days were consequently calculated and these results are discussed. Due to the similarity in the maps for AHR and SHR, only the maps for AHR will be discussed in section 3.4.

In the NCEP data Gauteng is situated approximately in the middle of four grid points (Fig. 3-1). Horizontal wind convergence and relative vorticity advection was calculated and displayed as vertical profiles at these four grid points.

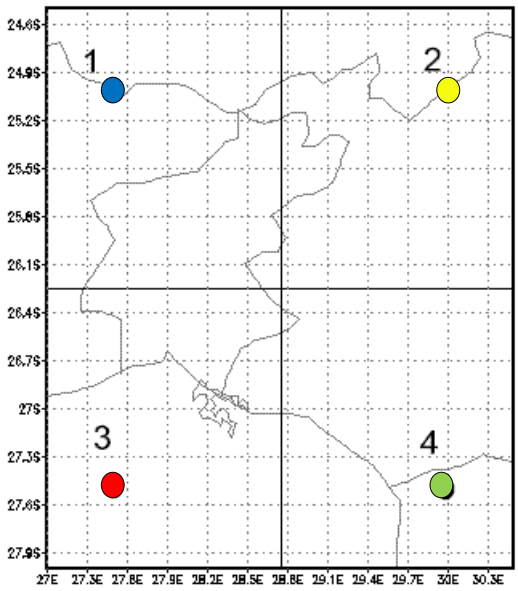


Figure 3-1: Location of the four grid points in the NCEP data set close to Gauteng.

3.4 Synoptic climatology of heavy rainfall over Gauteng

Three types of maps are displayed and discussed in the following 6 figures (Fig. 3-2 to 3-7). In the left hand columns of the figures are the monthly long-term means. The middle column shows the average monthly values for AHR days and in the right hand column of these figures are the average maps for No Rain days.

Variables displayed here are the geopotential heights at 850,700 and 500 hPa (top row of Figs. 3.2-3.7). Shades of blue are the 850 hPa geopotential heights, the 700 hPa geopotential heights are dotted lines and 500 hPa geopotential heights are solid lines. The second row shows the 300 hPa geopotential heights (solid lines) as well as the 500 hPa air temperatures (in °C; shaded contours). Green shades indicate temperatures higher than -7.5 °C. In the third row the arrows represent the 850 hPa moisture flux and the shaded colours the 700 hPa wind divergence (warm colours represents wind convergence).

3.4.1 October months

At 850 hPa on the mean October circulation map a trough of 1520 gpm extends to southern Namibia with a westerly trough of the same depth over the southwest coast (Fig.3-2). A large high pressure system (1540 gpm) is present east of southern Africa over the Indian Ocean (IOH) causing an on-shore flow of moisture into southern Zimbabwe which recurves into the central interior of South Africa. On average, wind divergence occurs over Gauteng at 850 hPa during October months.

At 700 hPa a broad trough is present west of the continent with a high of 3190 gpm over Northern Botswana. The 700 hPa ridge line is located over Gauteng. At 500 hPa the circulation is similar to 700 hPa but with the trough line even further west of the country and a high of 5870 gpm over Botswana. This general pattern repeats itself at 300 hPa but at this level the ridge line was just west of Gauteng with the high north of Namibia and Botswana. At 500 there is a temperature trough over Namibia. Over Gauteng the temperatures are between -10 and -11 °C.

Average maps on October AHR days over Gauteng show that the 850 hPa trough is deeper over Botswana and Namibia than the average. The 1520 gpm contour extends southward to the central and western interior of South Africa. A westerly trough visible on the south coast on the long-term average maps is further south and not visible in this window. The high east of the country has the same strength than on the mean maps (1540 gpm) but now formed a closed cell of anticyclonic circulation further south resulting in an increased flux of moisture into the Limpopo Province and southern Zimbabwe. The deeper 850 hPa trough over western SA results in strong geopotential height gradients over the central interior and stronger winds. This in turn results in a stronger northerly flow into the central interior and Gauteng bringing with it moist air from the Mozambique Channel. At 700-300 hPa the trough west of the country is more noticeable than the long-term average maps and was situated closer to the west coast of southern Africa. Gauteng is situated very close the upper air ridge line at 700 and 500 hPa but the ridge is very slightly east of Gauteng resulting in 700 hPa wind convergence and the advection of cyclonic vorticity (Fig. 3-9 in Chapter 3-5). Even though at 300 hPa the trough west of the country is not very pronounced the ridge present over the central interior on the monthly mean map is now located east of Gauteng over the Indian Ocean. The position of the 500-300 hPa trough (west of Gauteng) means that the circulation at these levels favoured upward motion as cyclonic vorticity is advected into Gauteng (Fig. 3-9 in Chapter 3-5) and horizontal wind divergence occurs at these levels (Fig. 3-8 in Chapter 3-5). Synoptic-scale uplift happens in areas with abundant surface moisture. The combination of the convergence of moist air at 850-700 hPa and uplift above that forces the development of cloud and rainfall. At 700 and 500 hPa the high pressure system clearly visible on the monthly mean October maps over Botswana and Nambia weakens and split into two cells over Zimbabwe and Namibia. This circulation is indicative of Taljaards' westerly trough. The temperature over Gauteng at 500 hPa on AHR days were the same as the average October value, but the temperature trough over the western interior was deeper than the average value. There was also an increased temperature gradient over the area west of Gauteng. Lower temperatures advected into Gauteng in the middle troposphere causes an increase in the temperature lapse rates and therefore an increase in the conditional instability in the atmosphere.

CLIM AHR No Rain

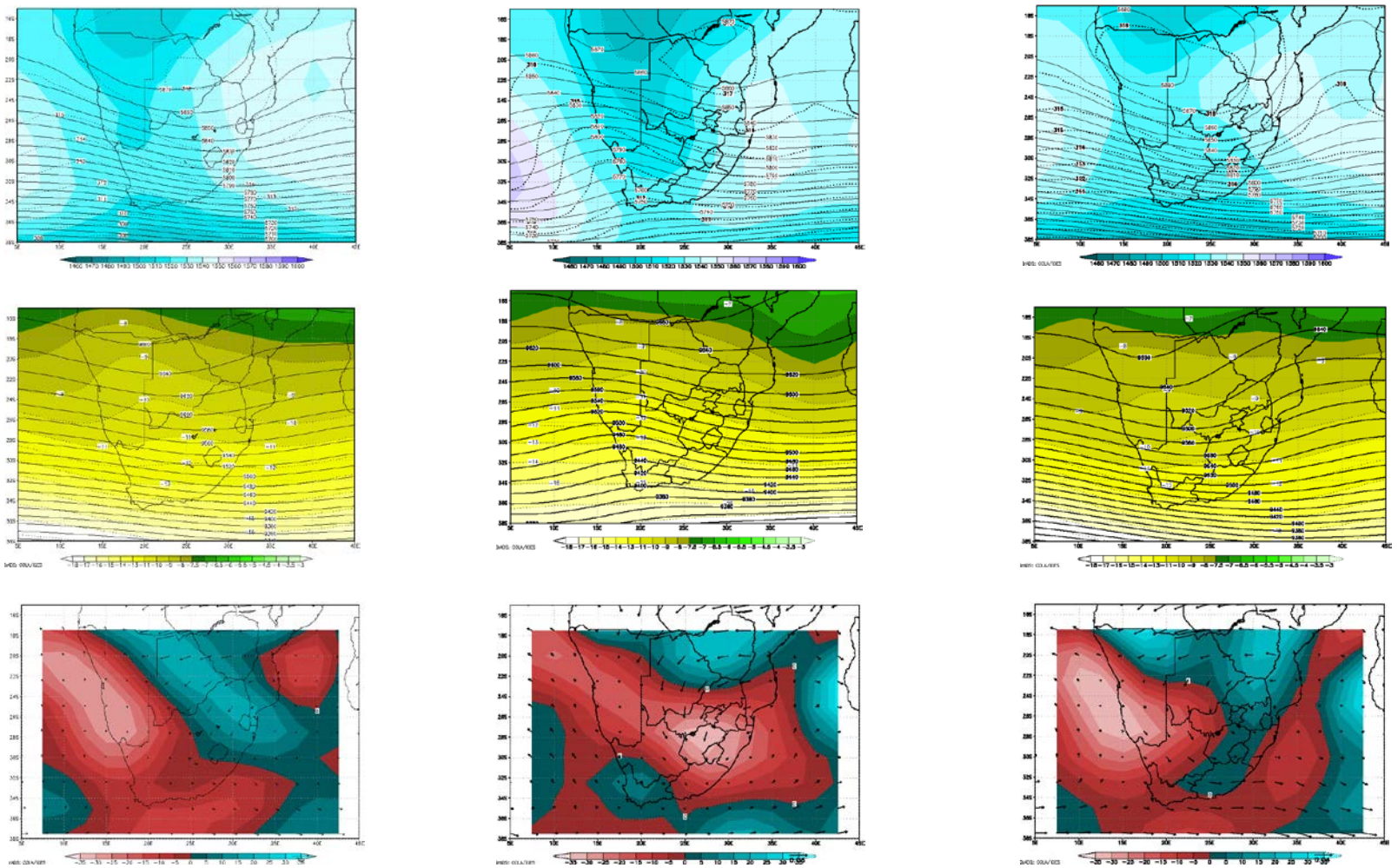


Figure 3-2: Monthly mean geopotential heights for October months at 850, 700 and 500 hPa (top) and the 500 hPa temperatures and 300 hPa geopotential heights (centre) and 700 hPa wind divergence and 850 hPa moisture flux (bottom). The centre column depicts the same maps but for days with heavy rainfall and the right hand column for days with no rain.

CLIM AHR No Rain

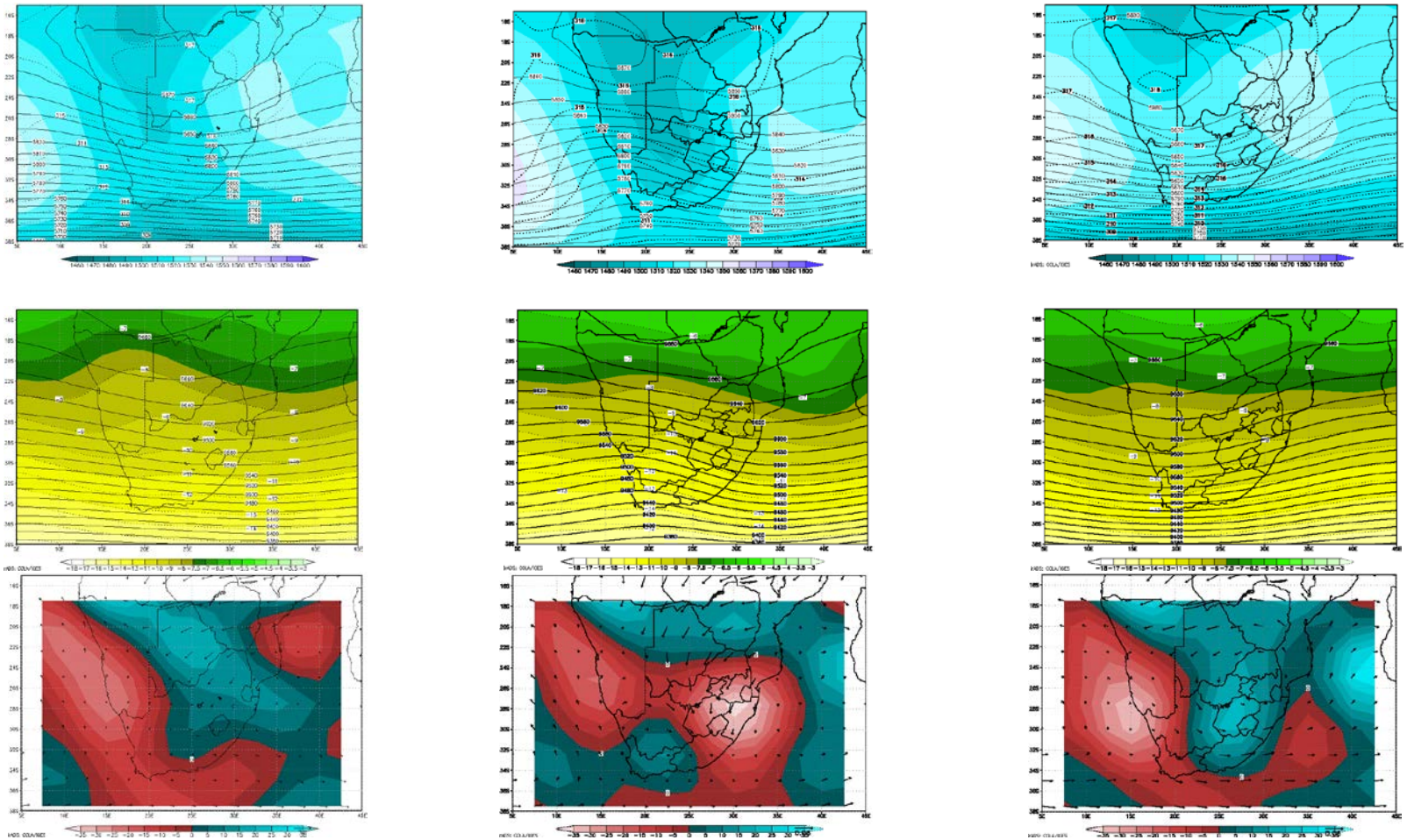


Figure 3-3: : Monthly mean geopotential heights for November months at 850, 700 and 500 hPa (top) and the 500 hPa temperatures and 300 hPa geopotential heights (centre) and 700 hPa wind divergence and 850 hPa moisture flux (bottom). The centre column depicts the same maps but for days with heavy rainfall and the right hand column for days with no rain.

CLIM AHR No Rain

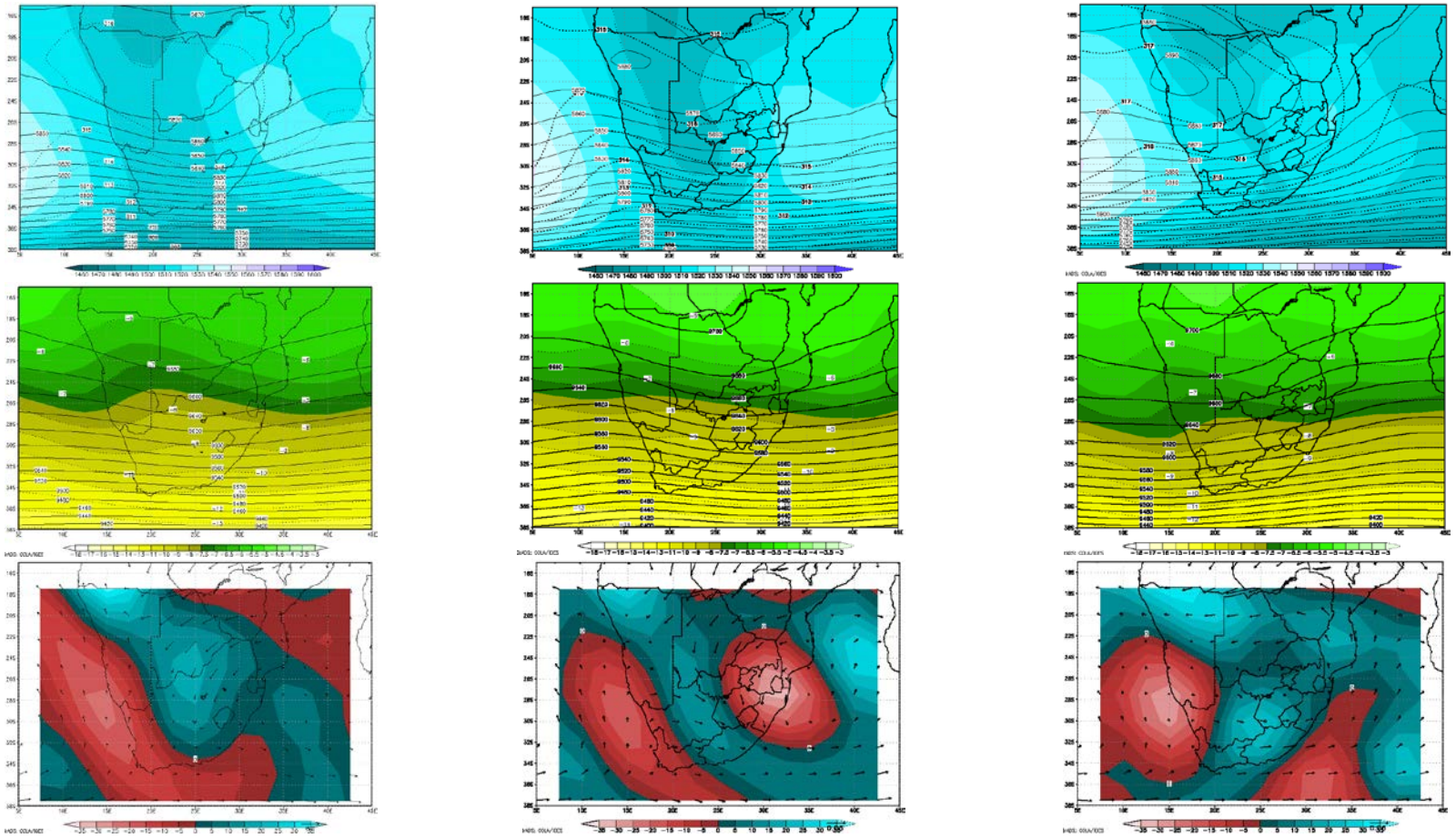


Figure 3-4: : Monthly mean geopotential heights for December months at 850, 700 and 500 hPa (top) and the 500 hPa temperatures and 300 hPa geopotential heights (centre) and 700 hPa wind divergence and 850 hPa moisture flux (bottom). The centre column depicts the same maps but for days with heavy rainfall and the right hand column for days with no rain.

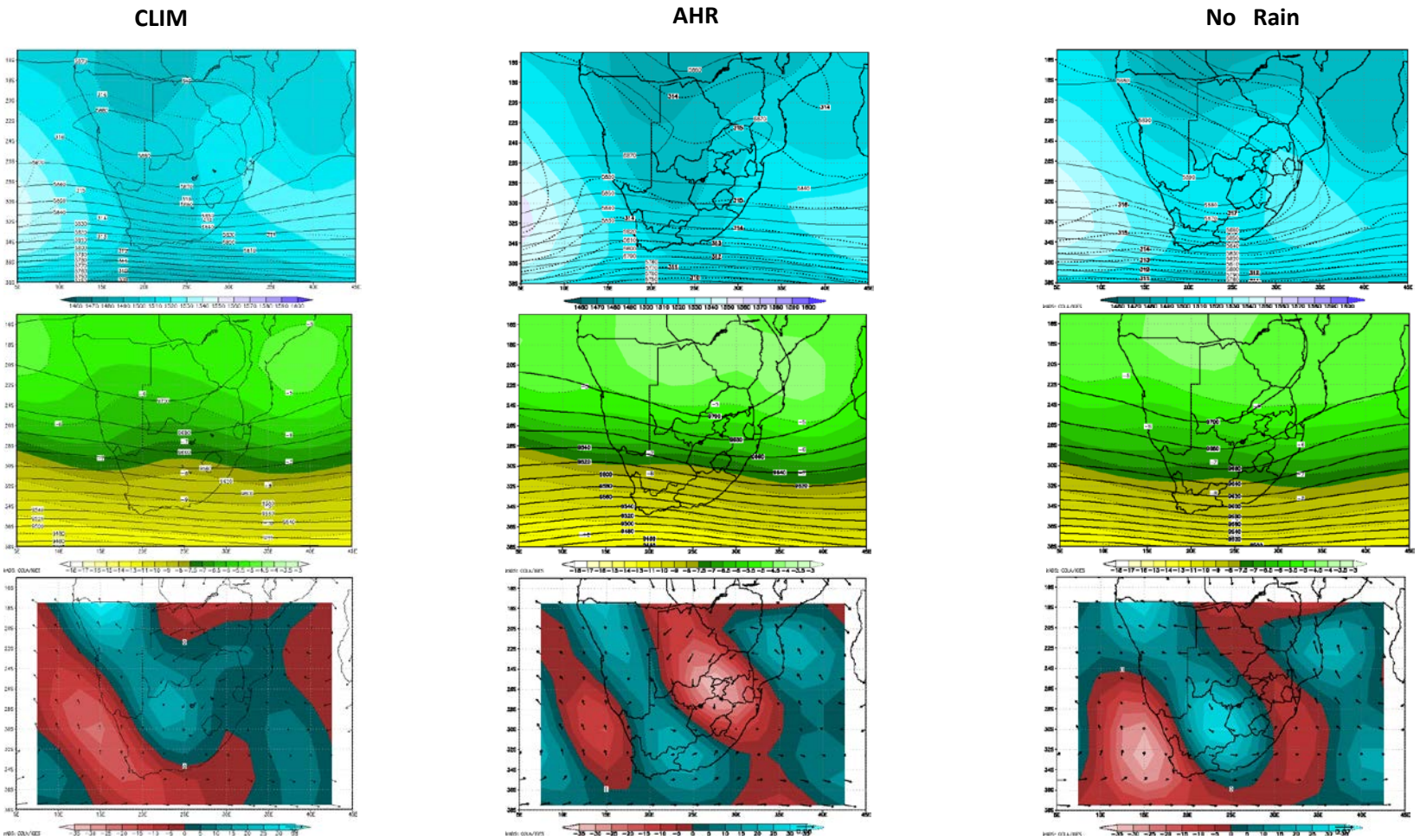


Figure 3-5: : Monthly mean geopotential heights for January months at 850, 700 and 500 hPa (top) and the 500 hPa temperatures and 300 hPa geopotential heights (centre) and 700 hPa wind divergence and 850 hPa moisture flux (bottom). The centre column depicts the same maps but for days with heavy rainfall and the right hand column for days with no rain.

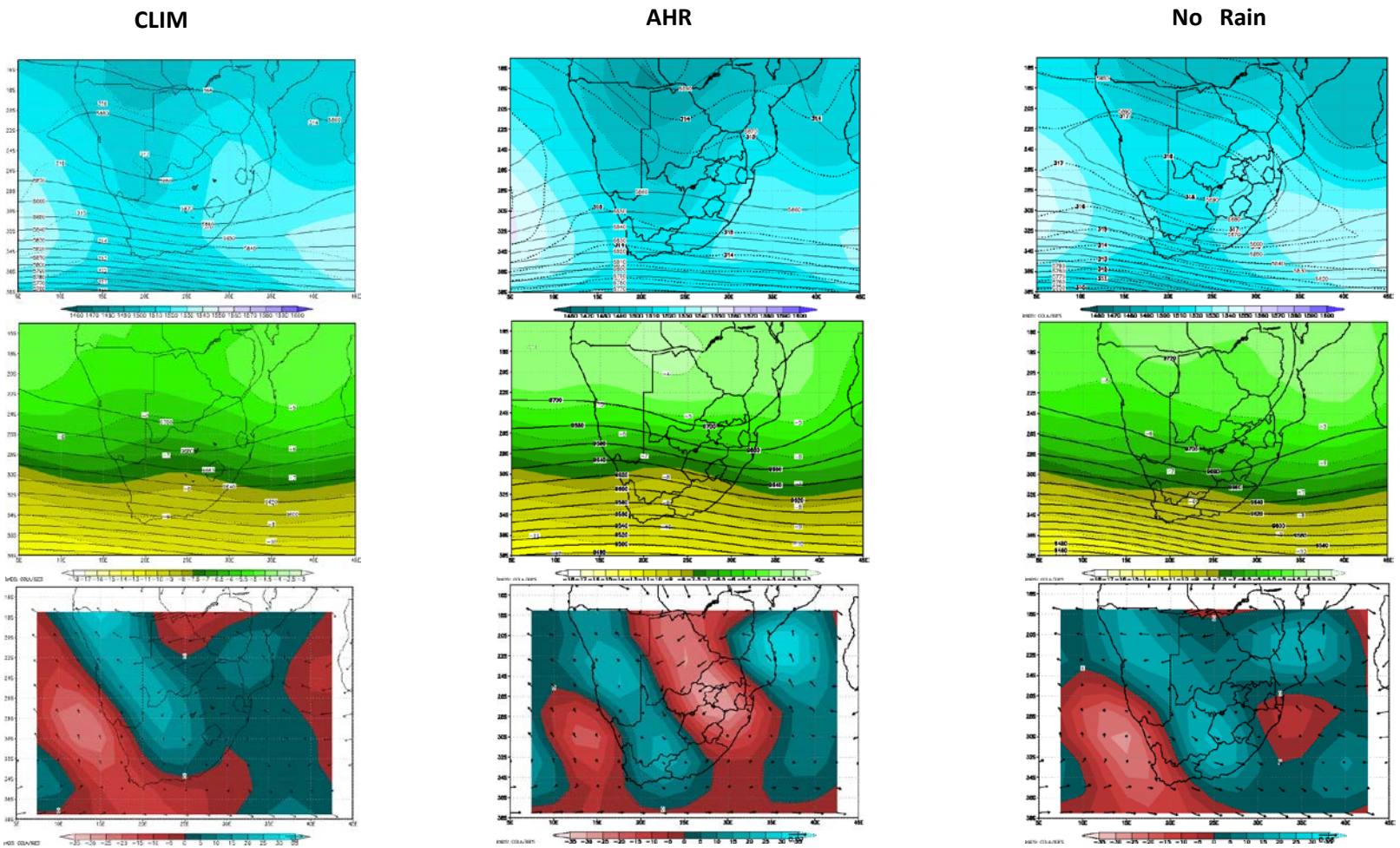


Figure 3-6: : Monthly mean geopotential heights for February months at 850, 700 and 500 hPa (top) and the 500 hPa temperatures and 300 hPa geopotential heights (centre) and 700 hPa wind divergence and 850 hPa moisture flux (bottom). The centre column depicts the same maps but for days with heavy rainfall and the right hand column for days with no rain.

CLIM AHR No Rain

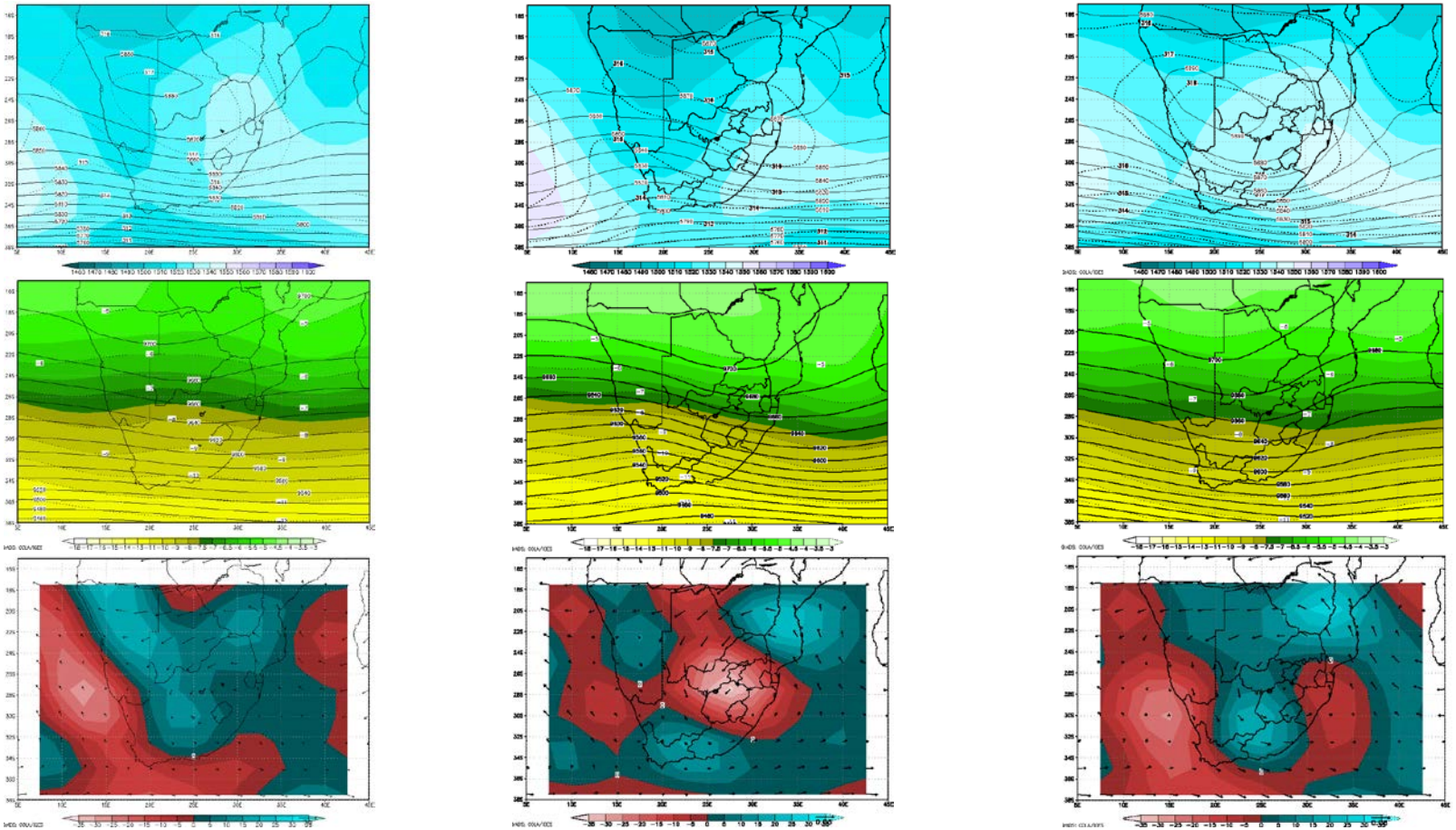


Figure 3-7: : Monthly mean geopotential heights for March months at 850, 700 and 500 hPa (top) and the 500 hPa temperatures and 300 hPa geopotential heights (centre) and 700 hPa wind divergence and 850 hPa moisture flux (bottom). The centre column depicts the same maps but for days with heavy rainfall and the right hand column for days with no rain.

On No Rain days the 850 hPa trough is weaker than the long-term mean trough over Namibia. A very broad trough is present over the western interior and is approximately 30 gpm weaker than during heavy rainfall days. The 1540 gpm high is still present east of the country but is situated further north than normal, as is the westerly trough over the south coast. This circulation indicates that the mid-latitude weather systems are further north than usual. The weaker trough over the western interior and the IOH situated further north cause weaker winds and thus less moisture to invade southern Africa. The 850 hPa ridge also extends into western Botswana resulting in a maximum influx of moisture further westwards over southern Africa. From 700-300 hPa the trough west of the country is weaker and lies further westward over the Atlantic Ocean. The high pressure system over Botswana at 700 and 500 hPa strengthens and extends a ridge to the area west of Gauteng resulting in the advection of anticyclonic vorticity (Fig. 3-9). At 700 hPa wind divergence now takes place over Gauteng and at 300 hPa the ridge is clearly visible west of Gauteng resulting in the advection anticyclonic vorticity and subsiding air. The temperature gradients at 500 hPa is quite slack and over Gauteng the temperature was ± 1 °C warmer than on heavy rainfall days.

3.4.2 November months

Mean circulation at 850 hPa in November months is similar to October months with the exception that the trough over Namibia is 10 gpm deeper and situated slightly further east over the Northern Cape/Namibia border (Fig. 3-3). A westerly trough is still present but slightly further to the southeast on the south coast and is also some 10 gpm deeper than in October months. Synoptic circulation at 700-300 hPa is very similar than in October months but with a slight southward displacement of the high pressure system over Botswana. At 700 hPa this high pressure weakens by about 10 gpm but at 300 hPa the high strengthens over Angola and Zambia.

The 850 hPa circulation for AHR days in November months is similar to AHR days in October months but the surface trough is deeper and located further eastwards over the central interior. The high pressure in the Indian Ocean is also located further east and the circulation around its periphery results in a strong northeasterly influx of moisture into Gauteng. From 700-300 hPa the circulation of heavy rainfall November months is the same as for October. The 500 hPa temperature trough lies further east than in October months and temperatures over Gauteng is 2 °C warmer than in AHR days in October months.

Synoptic circulation on No Rain days in November and October months was similar with the Botswana high stronger and located further west from 700-300 hPa resulting in a ridge west of Gauteng.

3.4.3 December months

At 850 hPa the trough over the western interior extends even further south and merges with the westerly trough on the south coast resulting in a broad trough of 1510 gpm over the western parts of the subcontinent (Fig. 3-4). The high on the eastern sea board is situated in the approximate same position than during the two preceding months but is ±10 gpm weaker. However, with the deepening of the trough over the western interior there is an increase in moisture flux into Gauteng from the northeast. From 700-500 hPa over Botswana the southward displacement of the high pressure systems continued with an additional 10 gpm weakening of the high at 700 hPa. The 500 hPa temperature trough moved from Namibia in October to Botswana as temperatures increased by another 1 °C over most of this window from November.

For heavy rainfall days the 850 hPa trough deepens to 1500 gpm and extends to the western Free State while the high east of the country has a strength of 1520 gpm resulting in a strong northeasterly influx of moisture into Gauteng. At 700 and 500 hPa the continental high pressure is further north than the long-term mean December high and at 500 hPa the high is 10 gpm stronger than the long-term December mean. A steeper trough west of the country together with the position of the Botswana high results in the ridge not being as well established west of Gauteng than on the long-term mean maps. The position of the high at 700 hPa results in easterly winds over the Limpopo Province and Botswana and weak geopotential gradients over Gauteng. There are two distinct areas of wind convergence over Southern Africa at 700 hPa, the one maximum is located over Swaziland (extending to Gauteng) and the other on the west coast in associated with the westerly trough west of South Africa. The 500 hPa temperature trough which was present over the western interior in October and November on AHR days is now very weak, west of Gauteng with 500 hPa temperatures over Gauteng in the order of -7 °C.

On No Rain days the trough also extends to the western Free State but is ± 10 gpm higher in depth than for heavy rain days. The Atlantic Ocean High (AOH) and IOH are well established over the interior with very little flux of moisture southwards into South Africa. At 700 hPa there is also a trough west of the country but the ridge extends over Gauteng resulting in 700 hPa wind divergence (Fig. 3-4). There is no 500 hPa trough west of the country with a stronger than normal high over Botswana and an absence of a trough at 300 hPa with a weak ridge west of Gauteng. Temperatures over Gauteng for AHR days and for No Rain days are both higher than -7 °C and higher than the long-term mean. There is a much more pronounced temperature trough over the western interior and Namibia than on AHR days in December months.

3.4.4 January months

The 850 hPa circulation during January months is very similar to December months with a broad trough over the western interior (Fig. 3-5). The IOH in the east moves a few degrees southwards and retreats from the continent to be situated over the Indian Ocean causing a sustained influx of moisture from the northeast into Gauteng. At 700 hPa the high pressure system over Botswana remains in the same position as during the month before and reaches its minimum strength during the summer months at 3160 gpm. At 500 and 300 hPa this high strengthens and extends a ridge over Gauteng at 700 and 500 hPa and west of Gauteng at 300 hPa. Temperatures at 500 hPa increases further with the temperature trough situated over the central interior.

On AHR days on the 850 hPa level the 1500 gpm trough extends southward to the western Free State and the Angola Low deepens to 1490 gpm. The IOH retreats further eastward to the Indian Ocean, but the influx of moisture into Gauteng is maintained. The 700 hPa trough is still present west of South Africa but now situated further south. There is a clear distinction at 700 hPa between the westerly circulation in the south and the easterly circulation further north separated by a weak ridge of ±3155 gpm over the central interior and with very slack geopotential gradients over Gauteng. This flow nevertheless results in wind convergence over the eastern interior. The Botswana High at 500 hPa depicted in the mean field is weaker and split in two, resulting in a weak trough being present at the same location as the 850 hPa trough and the 300 hPa ridge. This is indicative of a tropical low pressure system (Dyson and Van Heerden, 2002). A weak trough was also present west of South Africa at 300 hPa and the 9700 gpm high over the sub-continent was located further south than in wet December months. The presence of the 500 hPa temperature trough, dominant in early summer for heavy rainfall days, has disappeared with further increase in 500 hPa temperatures over Gauteng.

Synoptic patterns for No Rain January days were similar to dry December days. At 850 hPa the trough over the Mozambique Channel at 850 hPa is deeper than normal, resulting in a strong southerly flux of moisture into Mozambique and also resulting in southeasterly winds in Gauteng. This may be due to the presence of tropical cyclones over the Mozambique Channel which if located over the southern Mozambique Channel result in subsidence over Gauteng (Taljaard, 1996).

3.4.5 February months

In February at 850 hPa the surface trough over the western interior weakens by about 10 gpm while the high in the east re-established itself over the southeastern interior resulting in an easterly

flow into Gauteng (Fig. 3-6). From 700-500 hPa the Botswana High pressure strengthened further and is well established over Gauteng.

During AHR February days the Angola Low reaches its minimum value during summer months at 1480 gpm. However, the trough over the central interior weakens somewhat and the IOH is in its most southerly position. Moisture flux in February months is still from the east and northeast into Gauteng but not as favourable as in wet January months. The 700 hPa trough west of the country is even further south than in January months with a weak high present over Gauteng at this level. At 500 and 300 hPa the circulation is very similar to wet January months. At 500 hPa the warmest summer temperatures occur over Gauteng and over northern Botswana, with -5 °C and -4 °C respectively. The temperature trough is present over the western interior but more pronounced in the south than over Namibia.

The circulation during No Rain February and January months is very similar and characterised by a dominant high centred over southeastern Botswana.

3.4.6 March months

At 850 hPa the circulation in March is similar to October months as the trough over the western interior weakens and moves westward with the high pressure re-establishing itself over the eastern interior (Fig. 3-7). Moisture flux into Gauteng remains easterly but weakens. There is also evidence of a westerly trough over the southwest coast. However, at 700 and 500 hPa the high remained dominant over the central interior (more like January and February months than October months). At 300 hPa the high starts to weaken and moves slightly northward. The temperature trough at 500 hPa disappears and temperatures over Gauteng start to decrease.

During wet March months the surface trough is deeper than the mean for March months but not as deep as during wet January and February months. Southeast of the country the high strengthens but is still located further south than its mean March position. The 700 hPa trough on the west coast is located further north than in January, but with a high of 3160 gpm still located over Gauteng. At 500 hPa the continental high merged over Botswana but with a more pronounced trough west of the country.

Dry March months was characterised by a strong high at all pressure levels, even at 850 hPa the surface trough was situated over the Atlantic Ocean.

3.5 Horizontal wind convergence and advection of relative vorticity

Vertical profiles of horizontal wind convergence and relative vorticity advection at the four grid points indicated in Fig. 3-1 are displayed in Figs. 3-8 and 3-9 respectively. Grid point 1 and 3 blue and red respectively, lie to the west of Gauteng and capture westerly circulation systems moving into Gauteng. Yellow (grid point 2) and green lines (grid point 4) could capture the approach of tropical weather systems in late summer as well as the ridging of the IOH. October months are representative of the early summer months and January of the late summer and the vertical profiles for only these two months are displayed here.

In early summer AHR is associated with wind convergence from 850 to 500 hPa and with wind divergence above that (Fig. 3-8 top left). Convergence in the lower layers of the atmosphere is associated with the low pressure troughs over the western interior in October (Fig. 3-2). The largest values of wind convergence occur at the surface with the level of zero convergence close to 500 hPa. Grid point 1 (red) has the largest values of surface convergence and grid point 2 (yellow) the smallest values. Grid point 1 is situated closest to the surface trough while grid point 2 is nearer the centre of IOH. In Fig. 3-2 No Rain days in October (Fig. 3-8 top right) also experience surface convergence at all four grid points but the level of zero convergence is much closer to 700 hPa. Fig. 3-2 does indicate a surface trough at 850 hPa but the continental high pressure dominates the circulation at higher levels. The shallow layer of convergence will result in upward motion to be limited to the lower levels of the atmosphere.

For late summer AHR the general pattern of the vertical profile of wind divergence is similar to early summer but with the values of convergence at the surface and divergence in the upper troposphere smaller than in early summer (Fig. 3-8 bottom left). This is so despite the deepening of the trough at 850 hPa (Fig. 3-5). The level of zero convergence occurs higher than in October months at approximately 450 hPa. In No Rain January months the two grid points to the south of Gauteng, 3 (red) and 4 (green), have convergence at the surface but with very small values of convergence at grid point 4 (green). The two grid points to the north of Gauteng experience divergence throughout most of the troposphere but with small values of convergence above 400 hPa.

During both early and late summer cyclonic vorticity advection occurs through nearly the entire troposphere at all four grid points (Fig. 3-9 right) but increasing in value at 300 hPa. The value of vorticity advection at 300 hPa in October is about 3 times as large as during January months. In Fig. 3-2, strong geopotential gradients are seen over Gauteng for AHR days but these gradients slacken considerably during January months (Fig. 3-5) resulting in weaker winds and therefore smaller values of vortcity advection. Grid points to the south of Gauteng (red and green) have larger values of

cyclonic vorticity advection than the two grid points to the north. The winds south of Gauteng are stronger than to the north, because of the location of the high pressure system over Botswana.

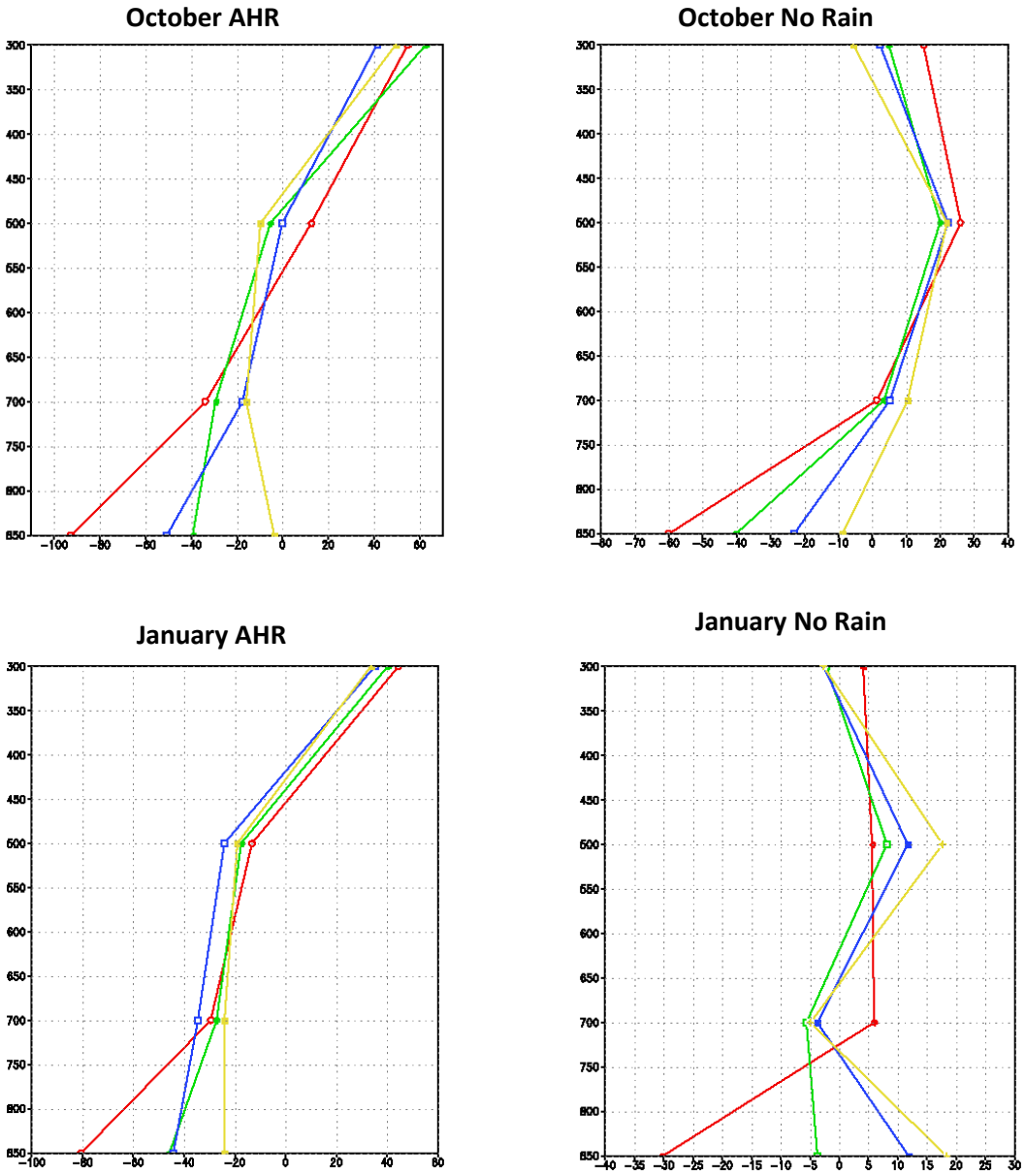


Figure 3-8: Vertical profiles of horizontal wind divergence at the four grids points (Fig. 3-1) surrounding Gauteng on AHR days (left) and No Rain days (right) for October and January months. Divergence values in 10^{-6} ms^{-1}. Negative values indicate wind convergence and positive values wind divergence. The blue vertical profile is for grid point 1 in Fig 3-1, the yellow for grid point 2, the red for grid point 3 and the green for grid point 4.

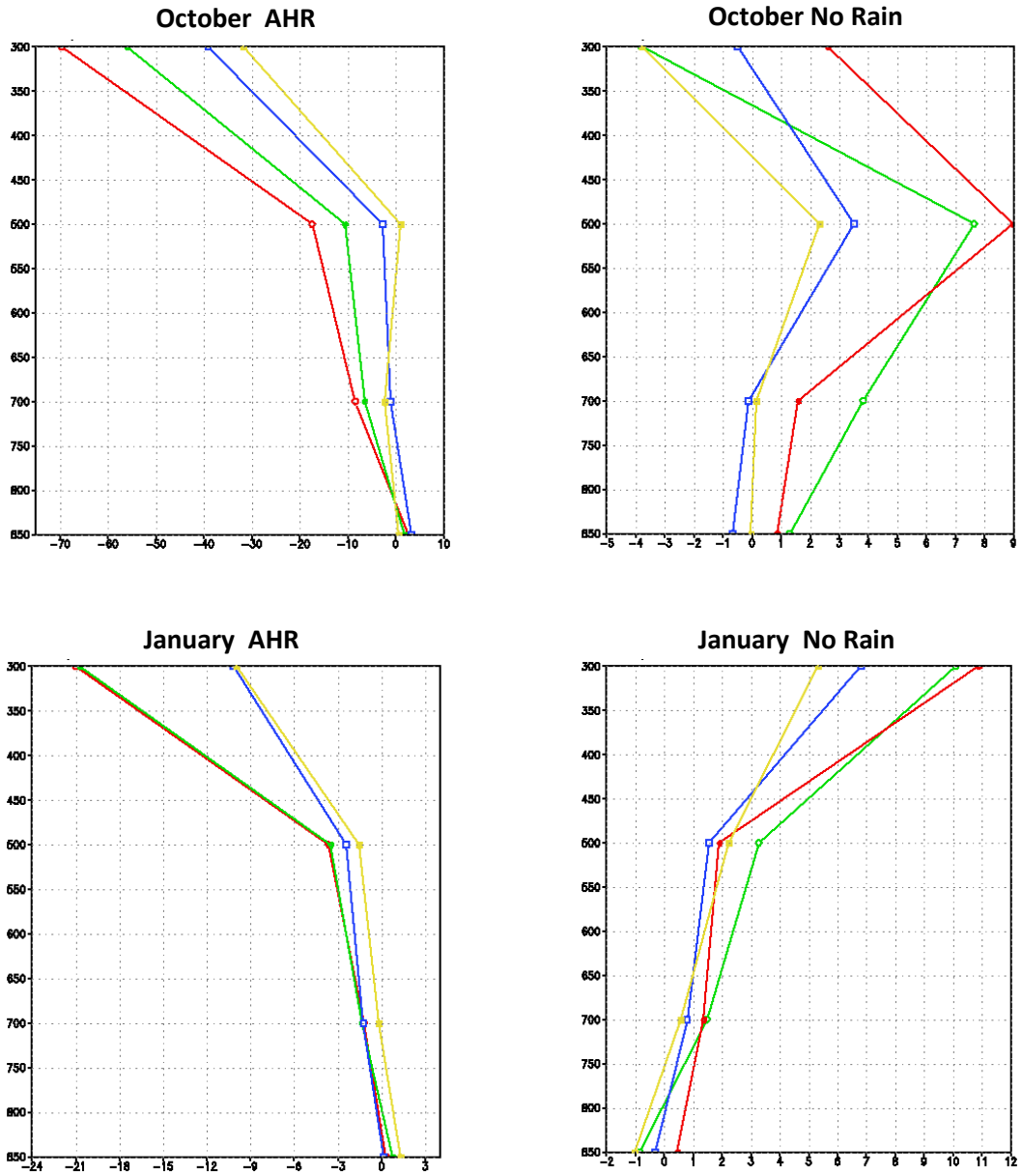


Figure 3-9: : Vertical profiles of relative vorticity advection at the four grids points (Fig. 3-1) surrounding Gauteng on AHR days (left) and No Rain days (right) for October and January months. Relative vorticity advection values in 10^{-10} s^{-2}. Negative values indicate cyclonic vorticity advection and positive values anticyclonic vorticity advection. The blue vertical profile is for grid point 1 in Fig 3-1, the yellow line for grid point 2, the red line for grid point 3 and the green line for grid point 4.

During No Rain months in early summer (Fig. 3-9 top right) cyclonic vortcity advection also occurs at 300 hPa but these values are an order of a magnitude smaller than the values during AHR October months. There is a layer of anticyclonic vorticity advection between 700 and 400 hPa in October. In No Rain January months (Fig. 3-9 bottom right) anticyclonic vorticity advection occurs throughout the troposphere with the largest values at 300 hPa. The high pressure system is well

established west of Gauteng at pressure levels less than 700 hPa causing advection of anticyclonic vorticity (Fig. 3-5).

3.6 Summary of main synoptic-scale features during wet and dry months

The synoptic circulation associated with heavy rainfall over Gauteng begins to change from an atmosphere with significant baroclinic flow features to an atmosphere approaching one with tropical flow features by December months. In October and November the westerly trough in the middle and upper troposphere is dominant west of the country on AHR days and extends to the north coast of Namibia. There are also tighter than normal temperature gradients at 500 hPa over and west of Gauteng with a temperature trough west of Gauteng. It is during these months that westerly troughs and COLs occur with higher frequency over SA (Taljaard, 1996). According Holton (1992) the gradient of temperature on a pressure level is indicative of a baroclinic atmosphere, as this gradient can only exists if atmospheric density is a function of both pressure and temperature. Under these circumstances a thermal wind will exist which simply means that the geostrophic wind will change with pressure. During AHR December days the westerly trough is still present west of the subcontinent but the continental high pressure systems moves southward to over and west of Gauteng in the middle and upper troposphere with a decrease in the temperature gradient at 500 hPa. Furthermore, the temperature trough over the western interior during the previous two months is no longer present.

The temperature trough is an important indicator for increased conditional instability (see Chapter 4.3) as lower temperatures in the middle troposphere cause an increase in temperature lapse rates and larger values of CAPE. During January and February months this cool trough at 500 hPa is absent and temperatures at 500 hPa over Gauteng increase to -6 °C and with a much reduced temperature gradient. This indicates that the atmosphere is becoming increasingly tropical while the conditional instability over Gauteng diminishes. The 700 and 500 hPa flow becomes very weak and even though the atmosphere must strictly still be classified as barcolinic, the decrease in the thermal winds also indicates the decrease in the baroclinic properties of the atmosphere or an increase in the tropical nature of the atmosphere. General circulation for AHR days in January and February months over the central interior of SA clearly represents a tropical low pressure system, as the low stand approximately upright from 850-500 hPa and with a 300 hPa high pressure system above that (see also Dyson and Van Heerden, 2002).

The surface troughs are deeper than normal in all months when AHR situations occur. This trough moves eastward from Namibia in October to the Free State in January and February months. The position of the IOH is further south than normal in all AHR months causing an influx of moisture from the Mozambique Channel into Gauteng. During AHR days in October-December and March the maximum 700 hPa wind convergence takes place south of Gauteng. In January and February months this convergence maximises northwest of Gauteng. All wet months have 850 hPa winds from the north (early summer) to northeast (late summer) and dry late summer months have southeasterly 850 hPa winds.

Horizontal wind convergence occurs through a deep layer of the atmosphere (at least up to 500 hPa) during early and late summer months when AHR occurs. In late summer this convergence takes place up to nearly 400 hPa. The values of convergence are larger in early summer than in late summer as wind speeds diminish. During both periods 300 hPa wind divergence occurs on AHR days. Cyclonic vorticity advection occurs in the upper troposphere on days with AHR but in early summer these values are twice as large as the values in late summer.

No Rain October, November and December months have weaker than normal troughs at all levels and a stronger high pressure system over Namibia at 700-300 hPa. This high extends a ridge to the west of Gauteng and this causes the advection of anticyclonic relative vorticity and downward motion. The IOH is located further north than during wet months and is also slightly stronger. Moisture influx associated with this high is into the central and western interior rather than into Gauteng. For dry January-March months the deeper than normal trough in the Mozambique Channel results in southeasterly winds over the eastern seaboard of South Africa and little moisture advection into Gauteng.

3.7 Overview

Daily Gauteng rainfall, calculated as part of the research in Chapter 2 was used to differentiate between monthly mean, heavy rainfall and No Rain circulation patterns. Chapter 3 provides information on the large scale environment in which heavy rainfall develops over Gauteng. Circulation criteria were identified that helped to classify heavy rainfall during the progression of the summer season. The change in the atmosphere from extra-tropical to tropical during summer is most evident in the decrease in mid-tropospheric temperatures, temperature gradients and with very slack geopotential gradients over Gauteng in late summer. It was shown how deep levels of horizontal wind convergence exist when heavy rainfall occurs in association with cyclonic vorticity advection at 300 hPa and northerly 700 hPa moisture flux. These favourable criteria are further

explored in Chapter 4 when investigating sounding-derived parameters. In Chapter 5 some of these favourable circulation criteria are exploited to predict rainfall frequencies over Gauteng.

3.8 References

DE CONING E, FORBES G. and POOLMAN EP (1998). Heavy precipitation and flooding on 12-14 February 1996 over the summer rainfall regions of South Africa: Synoptic and Isentropic analyses. *National Weather Digest,* **22**, 25-36.

DYSON LL and VAN HEERDEN J (2001). The heavy rainfall and floods over the northeastern interior of South Africa during February 2000. *South African Journal of Science,* **97,** 80-86

DYSON LL and VAN HEERDEN J (2002). A model for the identification of tropical weather systems. *Water SA,* **28**(3), 249-258.

HART NCG, REASON CJC and FAUCHEREAU N (2010). Tropical–Extratropical Interactions over Southern Africa: Three Cases of Heavy Summer Season Rainfall. *Monthly Weather Review,* **138**, 2608–2623.

HOLTON JR (1992). *An Introducion to Dynamic Meteorology*. Academic Press. 511 pp.

JURY MR and PATHACK B (1991). A study of climate and weather variability over the tropical southwest Indian Ocean. *Meteorology and Atmospheric Physics,* **47**, 37–48.

KALNAY et al. (1996). The NCEP/NCAR 40-year reanalysis project, *Bulletin of the American Meteorological Society*, **77**, 437-470.

MALHERBE J, ENGELBRECHT FA, LANDMAN WA and ENGELBRECHT CJ (2012). Tropical systems from the southwest Indian Ocean making landfall over the Limpopo River Basin, southern Africa: a historical perspective. *International Journal of Climatology*, **32**, 1018–1032.

METEO FRANCE (2012). URL: http://www.meteo.fr/temps/domtom/La_Reunion/meteoreunion2/#. (Accessed on 22 August 2012).

POOLMAN EP and TERBLANCHE D (1984). Tropiese siklone Demonia en Imboa. *South African Weather Bureau News Letter*, **420**, 37-45. South African Weather Service, Private Bag X097, Pretoria, 0001

SINGLETON AT and REASON CJC (2007a). Variability in the characteristics of cut-off low pressure systems over subtropical southern Africa. *International Journal of Climatology,* **27**(3), 295-310.

SINGLETON AT, and REASON CJC (2007b). A Numerical Model Study of an Intense Cutoff Low Pressure system over South Africa. *Monthly Weather Review,* **135**, 1128–1150.

TALJAARD JJ (1985.) Cut-off lows in the South African region. *South African Weather Service Tech. Pap.* **14**, 153 pp. South African Weather Service, Private Bag X097, Pretoria, 0001.

TALJAARD JJ (1996). Atmospheric circulation systems, Synoptic climatology and weather Phenomena of South Africa. Part 6 Rainfall.in South Africa. South African Weather Bureau Technical Paper (32). South African Weather Service, Pretoria, South Africa.

TENNANT WJ (2004). Considerations when using pre-1979 NCEP/NCAR reanalysis in the southern hemisphere. *Geophysical Research Letters,* **31**, L11112

TRIEGAARDT DO, VAN HEERDEN J and STEYN PCL (1991). Anomalous precipitation and floods during February 1988. South African Weather Service Technical Paper (23). South African Weather Service, Private Bag X097, Pretoria 0001.

TYSON P and PRESTON-WHITE RA (2000). *The weather and climate of South Africa.* Oxford University Press. 408 pp.

WILLIAMS E and RENO N (1993). An analysis of the conditional instability of the tropical atmosphere. *Monthly Weather Review,* **121**, 21–36.

4 Sounding-derived parameters associated with heavy rainfall over Gauteng

4.1 Preface

In this Chapter, Irene sounding data are used to develop a climatology of sounding-derived parameters over Gauteng. First a general climatology of atmospheric variables over Gauteng is provided. The last official SAWS publication which deals with upper air data was published in 1990 where a 20-yr climatology of temperatures, dew point temperatures and winds are provided (SAWB, 1990). Section 4.3 provides an updated monthly climatology of these variables over Gauteng for a 35-yr period and also introduces a climatology of parameters derived from these basic atmospheric variables. Next detailed information is provided about these variables when average heavy rainfall (AHR) and single station heavy rainfall (SHR) occur over Gauteng and these results are also compared to days which remain dry. The final contribution of this Chapter is to provide a sounding climatology where the vertical profiles of temperature, dew point temperatures, winds and a few of the sounding parameters are provided on heavy rainfall days by utilizing SOMs.

4.2 Background

Rasmussen and Blanchard (1998) explained how weather forecasters often rely on subjective experience when identifying favourable parameters for thunderstorm forecasting. Certain parameters are in general use in an operational environment and are often utilized to research heavy rainfall or severe storms. These parameters and their threshold values are deployed in case studies and forecasting but without thorough climatological verification. A baseline climatology provides objective information on which parameters are favourable and what values constitutes climatologically large or extreme values. The climatology can also provide guidance to weather forecasters on the parameters associated with heavy rainfall over Gauteng. In recent years many studies have been undertaken where sounding-derived parameters are associated with severe weather phenomena. Examples are: Manzato (2003), Craven and Brooks (2004), Doswell and Schultz (2006), Dupilka and Reuter (2006a and 2006b), Dimitrova et al. (2009) and Covadonga et al. (2009). Most, if not all, of these studies have been undertaken in the USA or Europe. In each instance a set of ingredients were identified which may be associated with severe weather for a specific geographical location.

This research followed the approach by Harnack et al. (1998) where the results obtained represents the effect of large scale weather systems on the production of heavy rainfall. It was determined whether the sounding data can be used to discriminate between conditions occurring on a day with heavy rain and those of a climatological mean.

Height above sea level and the complexity of the terrain has a significant influence on circulation and thermodynamical parameters. For example Bunkers et al. (2010) state that using the 700 hPa temperature as a proxy for the capping inversion is useful over the USA but not in elevated terrain. Bosart et al. (2006) propose that significant tornados over mountainous terrain require a deeper mesocyclone than on lower elevations. They also state that convective available potential energy (CAPE) values are larger over the plains than in the elevated terrain. Daniel (2006) proposes specific techniques to forecast the mixed-layer height over mountain terrains. The elevated plateau (± 1500 m a. m. s. l.) over the interior of South Africa likewise has a significant influence on the thermodynamic profile and moisture content of the atmosphere. Threshold values and ingredients identified over other parts of the globe are therefore not necessarily directly applicable to Gauteng. No such investigation has yet been undertaken in South Africa and this research provides ingredients associated with heavy rainfall over Gauteng. This study is conducted over a 35-yr period and serves as an instrument to investigate the intra-seasonal variability of the thermodynamical profile of the atmosphere during heavy rainfall.

4.3 Sounding-derived parameters

The sounding-derived parameters that were calculated are listed in Table 4-2 and the notations and units of the variables used are given in Table 4-1. These parameters were chosen as they are either capture the change in characteristics of the atmosphere from early to late summer or they have been shown to be in some way or another an ingredient of heavy rainfall and severe thunderstorms. Although the results from this research have shown that some of these parameters were not associated with heavy rainfall over Gauteng, they are included here as they are ingredients to heavy rainfall forecasting elsewhere. In Table 4.2 the following notations are used

Table 4-1: Notations and units of variables used in Table 4-2.

Symbol	Variable
T	Temperature (°C)
Td	Dew point temperature (°C)
Z	Geopotential height (gpm)
q	Mixing ratio (gkg^{-1}) but in $kgkg^{-1}$ in precipitable water (PW) calculation
p	Pressure (Pa in all calculation)
v	Meridionial wind speed (ms^{-1})
u	Zonal wind speed (ms^{-1})

Table 4-2: Equation units and abbreviations of sounding parameters

Variable	Units	Abbreviation	Equation
Average 500-300 hPa temperatures	°C	T_{53}	Average temperature in all pressure levels between 500 and 300 hPa
Temperature lapse rate	°C km^{-1}	$TL_{p2,p1}$	$1000 * \frac{T_{p2} - T_{p1}}{Z_{p2} - Z_{p1}}$ where p2>p1
Vertical temperature difference	°C	$VTD_{p2,p1}$	$T_{p2} - T_{p1}$ where p2>p1
Geopotential thickness	gpm	TH	$TH = Z_{p1-}Z_{p2}$ where p2>p1
Freezing level height	m	FL	Height of the 0 °C isotherm
Average mixing ratio	gkg^{-1}	M	Average mixing ratio in 100 hPa above surface level
Precipitable water	mm	PW	$PW = \frac{1}{g} \sum_{i=850\,hPa}^{i=300hPa} \frac{q(p_i)+q(p_{i+1})}{2} \quad (p_i - p_{i+1})$
Average tropospheric equivalent potential temperature	K	$\Theta_{e\ ave}$	Average Θ_e at all pressure levels between the surface and 300 hPa
Equivalent Potential Temperature lapse rate	K	$\Delta\Theta_e$	$\Theta_{e\ p2} - \Theta_{e\ p1}$ where p2>p1
Average wind speed in the 600-400 hPa pressure levels	ms^{-1}	W_{64}	Average wind speed at all pressure levels between 600 and 400 hPa
Average meridionial wind speed in the 800-600 hPa pressure levels	ms^{-1}	WV_{86}	Average meridionial wind speed at all pressure levels between 800 and 600 hPa
Bulk wind shear	ms^{-1}km^{-1}	$WS_{p1,p2}$	$\frac{\sqrt{(u_{p2} - u_{p1)}{}^2 + (v_{p2} - v_{p1)}{}^2}}{(Z_{p2} - Z_{p1)}}$ where p2<p1
Magnitude of wind shear[2]	ms^{-1}	$MWS_{p1,p2}$	$\sqrt{(u_{p2} - u_{p1)}{}^2 + (v_{p2} - v_{p1)}{}^2}$ where p2<p1
K -Index	°C	KI	T_{850}-T_{500}+Td_{850}-(T_{700}-Td_{700})
Elevated *K*-Index	°C	EKI	T_{700}-T_{500}+Td_{700}-(T_{600}-Td_{600})
Total Totals Index	°C	TTI	T_{850}-T_{500}+Td_{850}-T_{500}
Elevated Total Totals Index	°C	ETTI	T_{700}-T_{500}+Td_{700}-$T_{500}{}^2$
Showalter index	°C	SI	Parcel temperature lifted from 850 hPa (calculated using mean mixing ratio in lowest 100 hPa) – environmental temperature at 500 hPa
Mean Layer Convective Available Potential Energy	J kg^{-1}	CAPE	CAPE calculated with the average dew point temperature in the lowest 50 hPa above the surface
Mean Layer warm cloud depth	hPa	WCD	Pressure of freezing level – pressure of lifted condensation level (calculated using mean mixing ratio in lowest 100 hPa)

[2] The average wind in the 100 hPa above the ground is used as the surface wind

4.3.1 Temperature parameters

As summer progresses from early to later in the season, temperatures in the atmosphere increase. The rate of increase in these temperatures and their vertical distribution will impact on the stability of the atmosphere. The average temperature in the 500-300 hPa layer (T_{53}) is a parameter which represents the upper tropospheric temperatures and is used to indicate the seasonal change in temperatures at these levels (Table 4-2). A warm core of T_{53} has also been shown to be an ingredient for identifying continental tropical low pressure systems over South Africa; a weather system that has been responsible for wide spread and heavy rainfall over South Africa (Dyson and Van Heerden, 2002). On the other end of the spectrum, cut-off lows (COLs) are cold cored at these pressure levels while they may also cause heavy rainfall to occur. T_{53} is used to aid in the identification of the characteristics of the atmosphere in which heavy rainfall occurs.

Temperature lapse rates (TLs) are used to define the conditional instability of the atmosphere. When the TL is larger than the dry-adiabatic lapse rate the atmosphere is absolutely unstable while a TL less than the moist-adiabatic lapse rate is said to be absolutely stable (Schultz et al., 2000). Larger TLs are associated with increased conditional instability in the atmosphere and a higher likelihood of convective development (Doswell et al., 1985). Craven and Brooks (2004) arbitrarily define a steep temperature lapse rate as one where the lapse rate is greater than 7 °C km^{-1}. It is common practise by the authors mentioned above to use the TL equation as in Table 4-2 with units °C km^{-1}. In this dissertation a second temperature lapse rate is used and referred to as vertical temperature difference (VTL). VTL is simply the difference in temperature between two pressure levels and was used in instability indices such as the Total Totals Index (TTI) or the *K*-Index (KI).

4.3.2 Moisture variables

Several moisture parameters have been used to investigate heavy rainfall; these include dew point temperatures and mixing ratios at different pressure levels, precipitable water or relative humidity through different layers in the atmosphere (Doswell et al., 1996; Harnack et al., 1998; Dupilka and Reuter, 2006b). In this research two moisture variables are used. They are the average mixing ratio (M) (and dew point temperature) in the lowest 100 hPa above the surface and the precipitable water (PW) calculated from the surface to 300 hPa.

Instability indices such as convective available potential energy (CAPE), the Showalter Index (SI) or convective cloud base height all rely on a surface dew point temperature (or mixing ratio). (The definitions of the instability indices are dealt with in Section 4.3.5). The surface mixing ratio has a significant influence on the value of the instability indices and choosing the most appropriate value

is essential when investigating convection. In this research the so called "mean layer" mixing ratio is used (Craven and Brooks, 2004). This is the average mixing ratio in the lowest 100 hPa (M) and is more representative of a well-mixed boundary layer as Craven et al. (2002) cautions against the use of surface mixing ratio in calculating thermodynamic indices. An example of the impact of a surface mixing ratio on the calculation of convective indices is shown in Fig. 4-1 where there is a sudden decrease in dew point temperature just above the ground. This often happens during afternoon sounding when the temperature lapse rate near the surface is dry-adiabatic or even super-adiabatic. On this day the CAPE value was 2021 Jkg^{-1} when the surface mixing ratio was used but only 1313 Jkg^{-1} when the average mixed layer mixing ratio was used.

Where M provides information about the availability of moisture at the surface, the PW is suitable to quantify the amount water vapour available in the troposphere. PW is the mass of water vapour in a column with unit area. Harnack et al. (1998) found a mean PW of 25 mm to be associated with heavy summer rainfall over Utah while Dupilka and Reuter (2006b) found similar values related with severe tornados over Alberta Canada.

4.3.3 Equivalent potential temperature parameters

The equivalent potential temperature of a saturated parcel is given by (Holton, 1992)

$$\theta_e = \theta exp^{\frac{L_c q_s}{c_p T}} \qquad \textbf{4-1}$$

where

L_c Latent heat of condensation

q_s saturation mixing ratio

c_p specific heat of dry air at constant pressure

θ potential temperature

Equation 4-1 may also be applied to an unsaturated parcel when T is the temperature that the parcel would have if brought to saturation by an adiabatic expansion.

Θ_e is often used to determine the convective or potential instability of a layer in the atmosphere (Schultz et al., 2000). The atmosphere is said to be convectively unstable if $\frac{d\theta_e}{dz}<0$. The change of Θ_e with height was investigated by Cohen et al. (2007) in a study related to mesoscale

convective system (MCS) environments. They found that 0-7 km $\Delta\Theta_e$ as well as the ($\Theta_{e\,min} - \Theta_{e\,max}$) do well in distinguishing between weak and strong MSC environments. Furthermore, Θ_e can be used to identify airmass changes (Houston and Wilhelmson, 2012). Here $\Theta_{e\,ave}$ is used to describe the change in the properties of the airmass over Gauteng from early to late summer, while $\Delta\Theta_e$ will be utilized to investigate the convective instability in the atmosphere.

4.3.4 Wind parameters

Doswell et al. (1996) state that flash floods are associated with slow moving storms, as such situations increase the precipitation rate of convective cells. In order to investigate whether heavy rainfall over Gauteng develops from slow moving storms, the average wind speed in the 600-400 hPa layer (W_{64}) was calculated.

It was shown in Chapter 3.2 how moisture associated with rainfall over the summer rainfall area results from the surface high pressure system located close to Maputo or the invasion of tropical air from Botswana and Angola. Both these generalised circulation patterns cause winds with a northerly component over Gauteng. For this reason an association was sought between the average meridionial wind in the lowest 200 hPa (WV_{86}) of the atmosphere and days on which heavy rainfalls occur.

Bulk wind shear (WS) is a parameter which is often used when discussing atmospheric conditions favourable for the development of MCS (e.g. Doswell et al., 1996; Craven and Brooks, 2004; Dupilka and Reuter, 2006a and Cohen et al., 2007). In the absence of wind shear buoyancy is the dominant factor in the development of deep moist convection. When wind shear is present convective development is influenced through the change in vertical transport of horizontal momentum (e.g. Wu and Yanai, 1994). Wind shear organizes convection in lines (e.g. squall lines) or if it the relationship between buoyancy and wind shear is optimal supercells may develop. It is generally accepted that supercells over the USA are associated with wind shear values of > 20 ms^{-1} in the lowest 4-6 km AGL. Wind shear values larger than 25 ms^{-1} in the lowest 2-3 km AGL was found to be associated with bow echos in the USA (Craven and Brooks, 2004). Dupilka and Reuter (2006a) found that the magnitude of wind shear in the 900-500 hPa as well as the 900-800 hPa layer provides the best guidance in distinguishing between environments favouring the development of significant vs. weak tornados in Alberta, Canada. In their study the mean value of the wind shear for significant tornados in the 900-500 hPa level was approximately 5 $ms^{-1}km^{-1}$. In this thesis WS is used to ascertain if atmospheric conditions associated with severe storms are also related to heavy rainfall. WS is also used to identify the tropical nature of the atmosphere since WS should not exist in a barotropic atmosphere (see Chapter 3.2).

4.3.5 Convective indices

Several indices have been developed to exploit the character of a conditionally unstable atmosphere such as the Total Totals Index (TTI), the Showalter Index (SI) and the *K*-Index (KI) (Henry, 1999). The TTI was first proposed by Miller (1967) and the *K*-Index was developed by George (1960). These two indices consider the ambient temperature and dew point while the SI (Showalter, 1952) also takes the lift of the parcel into consideration.

The KI is used to identify convective and heavy-rain-producing environments and a threshold value of 35 °C is used in the USA to identify widespread convection. The TTI is used to identify severe storms and in the USA a TTI of > 50 °C is used to identify possible severe thunderstorms (Henry, 2000). For both these indices the temperature difference between 850 and 500 hPa (highlighted in grey in Table 4-2) is one of the terms used in the calculation. Increased values in the temperature difference will increase the KI and TTI. Other terms in the KI index deals with surface and mid-level moisture, while in the TTI the 500 hPa temperatures are also taken into consideration. The Elevated *K*-Index (EKI) and Elevated Total of Totals Index (ETTI) is an attempt to take the height of the interior plateau of South Africa into consideration and were first proposed by Todd[3] (2010). The interior of South Africa lies close to 1500 m a. m. s. l. and the surface pressures are close to 850 hPa. In EKI and ETTI all the pressure levels (except 500 hPa) were adapted downward by 100-150 hPa. Here the temperature lapse rate between 700-500 hPa helps to determine the conditional instability of the atmosphere.

SI was calculated by using the mean layer mixing ratio. Lower atmospheric temperatures at 500 hPa will result in larger negative SI values which indicate increased Conditional Instability. De Rubertis (2005) stated that values < -3 °C indicate a moderate unstable atmosphere while values of <-6 °C are indicative of a very unstable atmosphere.

Convective available potential energy (CAPE) is the energy a parcel will have when lifted from the level of free convection (LFC) to the level of neutral buoyancy (LNB) and is calculated by a vertical integration of the parcel-to environment temperature between these levels (De Rubertis, 2005).

[3] Todd M P. Personal communication. Senior Forecaster, National Forecast Centre, South African Weather Service, Bolepi House Pretoria. markpatricktodd@gmail.com

$$\mathrm{CAPE} = g\int_{\mathrm{LFC}}^{\mathrm{LNB}}\left(\frac{T_p(z)-T_e(z)}{T_e(z)}\right)dz \qquad \textbf{4-2}$$

where

Tp(*z*)- temperature of the parcel

Te(*z*)- temperature of the environment

g acceleration due to gravity.

In this thesis the CAPE algorithm provided by Brimelow (1999) is used and the upper boundary is either the LNB or 300 hPa if the LNB has not been reached at pressure level 300 hPa. One of the advantages of using CAPE as a convective index is that it provides a measure of stability integrated over the depth of the sounding, as opposed to the SI or KI (for example) that use data from only a few mandatory levels. The atmosphere is moderately unstable if CAPE values are from 1000-2500 Jkg^{-1} and very unstable if CAPE> 2500 Jkg^{-1}.

Doswell and Rasmussen (1994) propose using the most unstable CAPE (MUCAPE) where the virtual temperature of the most unstable parcel in the lowest 300 hPa is used. The most unstable parcel is lifted from the level with the highest Θ_e value. The use of the MUCAPE is most effective to assess the potential of elevated convection and is especially useful overnight when the surface layer of air cools and a temperature inversion develops, but above the inversion the atmosphere remains unstable (Bunkers, et al., 2002). They found that in the afternoon (close to the time when maximum surface temperature occurs) MUCAPE values are equal to CAPE values when the surface temperature and dew point temperature values are used. In section 4.4.5 it is explained that in this thesis only 1200 UT soundings are used to calculate the convective variables. In South Africa 1200 UT are close to the time of maximum surface temperatures and MUCAPE values would therefore be equal to the CAPE values when the surface temperature and dew point temperature are used. As discussed in section 4.3.2 the use of a mean layer mixing ratio is preferred.

CAPE is essentially a measure of the conditional instability of the atmosphere and is often used together with wind shear to identify severe storm environments e.g. Craven and Brook (2004) and Dupilka and Reuter (2006a). The combination of wind shear and CAPE has also been used to identify short and long lived severe thunderstorms in a tropical atmosphere e.g. Williams and Renno (1993), Chaudhari et al. (2010) and Tyagi et al. (2011).

The warm cloud depth (WCD) is the difference between the pressure of the lifted condensation level (LCL) and the freezing level. This is an important parameter because it defines the layer of the cloud in which collision-coalescence is the dominant process for precipitation droplet growth (Market et al., 2003). Schumacher and Johnson (2008) discuss how the warm cloud depth is one of the ingredients for high precipitation rates. In this thesis the correlation between the WCD and the average wind speed in the 600-400 hPa layer is investigated as primary ingredients for high rain rates.

4.4 Data and Method

4.4.1 Irene sounding data as a proximity sounding for Gauteng

Brooks et al. (1994) provide a detailed discussion on the difficulties of defining a proximity sounding. The challenge is to sample the atmosphere in which the event formed but the main concerns are that the atmosphere is not spatially homogeneous and is changing with time. Darkow (1969) cited in Brooks et al. (1994) suggested that unrepresentative soundings should be eliminated but Brooks et al. (1994) explained some of the pitfalls in this approach. Placing strict constrains on the spatial and temporal variability of the atmosphere may result in eliminating a large number of soundings limiting the statistical significance of the results obtained from the remaining soundings. Darkov considered a proximity sounding when the sounding was done within 80 km of an event (tornados in his investigation) but Rasmussen and Blanchard (1998) preferred 400 km. Craven and Brooks (2004) considered that 185 km distance between the event and sounding location to be a proximity sounding while Groenemeijer and Van Delden (2007) chose a maximum distance of 100 km in order to retain a reasonable number of soundings for different types of severe weather events. The Irene sounding could therefore be considered to be a proximity sounding for Gauteng as the boundaries of the province are within 100 km from Irene.

4.4.2 Sounding data

SAWS Irene upper air data were acquired from http://www.weather.uwyo.edu/ for the 35 austral summers from October 1977 to March 2012. There is no evidence for consistent quality control (QC) of these data over the entire period. Another possible source of data is the Integrated Global Radiosonde Archive (Durre et al., 2006) and the advantage of using data from this source is that the data have been QC. However, it was found that very peculiar errors which occur in South African assents (i.e.: very dry dew point temperature spike just above the surface Fig. 4-2) is not dealt with adequately in their QC process which left the data unusable for the purpose of this research. These erroneous dew point temperatures will influence especially those sounding parameters which

deal with convection such as the Lifted Index or CAPE. Furthermore the IGRA dataset is incomplete as it does not contain data from all the soundings that were available.

Quality control procedures were employed here to create a 35-yr radiosonde dataset for the SAWS Irene Weather Office. The original data set contained 11 227 radiosonde observations valid for 0000 UT (5503) and 1200 UT (5724). After QC process discussed in Chapter 4.4.3 there were 8364 soundings available for analyses, 3800 at 0000 UT and 4564 at 1200 UT. The raw data include pressure, temperature, geopotential height, dew point depression, wind direction and wind speed data on standard and significant pressure levels.

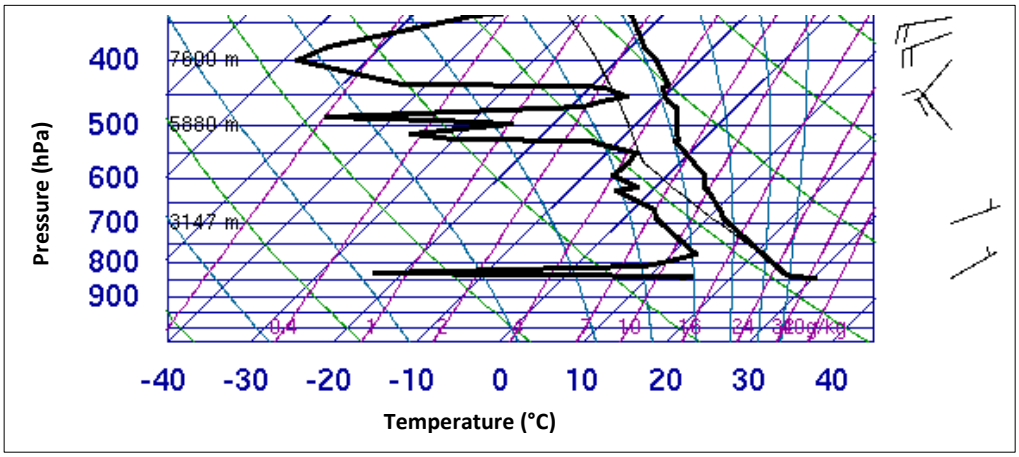


Figure 4-2: Irene skew-t-diagram on 23 December 2008 at 0900 UT. The arrow indicates a dew point temperature spike. Source http://www.weather.uwyo.edu/.

4.4.3 Quality control of upper air data

This research focuses on extreme events and it is therefore important to have confidence in the quality of the data used to calculate the parameters. As this research utilizes a 35-yr climatology of the sounding parameters it is possible, following the logic of Craven and Brooks (2004) that erroneous or contaminated data will be masked by the statistical analysis. However, very heavy rainfall events are rare over Gauteng and bad data may not be masked efficiently due to the small number of these events. A further challenge in interpreting any data, as Brooks and Stensrud (2000) explained is that rare interesting events and bad data often look similar. Similar QC processes as described by the Global Radiosonde Archive (Durre, 2006) were implemented on the Irene sounding data. These checks were only preformed at pressures higher than 300 hPa.

Table 4.3 lists the QC procedures employed and also provides information on the action taken for each of the checks. The *data completeness* test removes entire soundings if any of the

circumstances occur. A total number of 304 soundings were eliminated by this check, 187 of which was a result of the test requiring that the pressure difference between 2 sequential pressure levels be less than 200 hPa. This problem occurred mostly during the early part of the study period as 72 % of these events occurred prior to 1990. There were only 9 such events after 2000. The second highest number of soundings removed by the *data completeness test* was due to data not being available up to at least 300 hPa and because dew point temperatures were missing at pressures higher than 300 hPa. Soundings with missing winds were left in the data set but individual elements were replaced by a missing data flag.

The second set of errors removed from the data is referred to as *fundamental sanity checks* by Durre (2006). The *fundamental sanity* checks employed here are similar but are adapted for local climatological conditions. This check ensures that data values falls within known extremes and only the erroneous data were removed and not the entire sounding. There were 6 600 data values removed with these checks and of these 6184 were as a result of spikes in dew point temperatures.

Table 4-3: Quality control procedures in the Irene sounding data (after Durre, 2006).

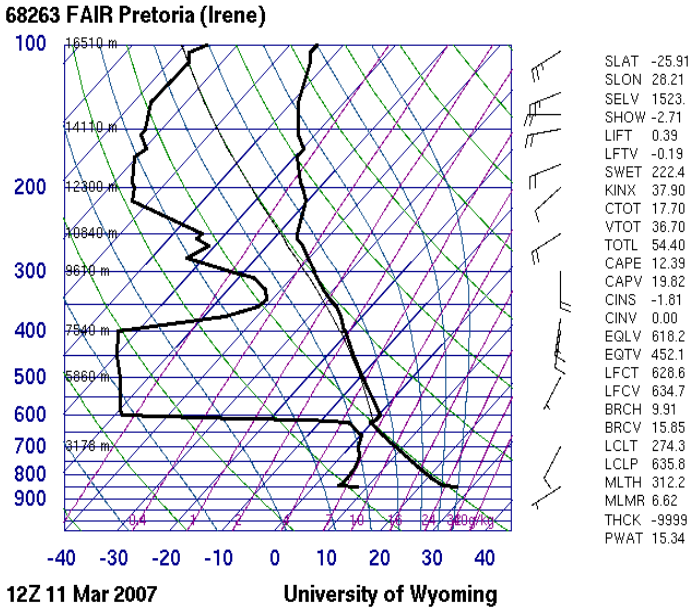
Type of check and *items deleted*	**QC Test**
Data completeness check - *remove soundings*	• < 10 pressure levels between surface and 300 hPa • T and Td not available from surface to 300 hPa • > 200 hPa pressure difference between 2 sequential pressure levels • Pressure does not decrease with height • Surface pressure < 830 hPa • T at surface < -10 °C • T lapse rate in the mid levels > dry adiabatic lapse rate
Fundamental sanity check – *remove individual elements*	• -120 °C > T > 45 °C • T < -65 °C when P > 300 hPa • Wind speed > 200 kt • 0 > wind direction > 360 ° • rh > 100 % • 70 °C < T-Td <0 °C • T > 0 °C, when P < 500 hPa • T > 10 °C when P < 550 hPa • T and Td spike checks

The third set of errors was identified by the *climatological checks*. Here long-term average values of temperature, dewpoint temperature and wind strength were calculated at 50 hPa pressure intervals. The data in each sounding were compared to the climatological average values and was rejected if it differed by more than six standard deviation from the average. However, this was only done after visual inspection of these soundings. In some cases these values were not rejected; for example the surface dew point temperature of -13 ˚C on the 17th of November 2002. Two hundred and twenty data values were eliminated in the *climatological test* but 77 data values were preserved despite failing the *climatological test*.

After all of these checks were completed the geopotential heights were recalculated at each of the available pressure levels for every sounding by using the hypsometric equation.

It is recommended here that radiosonde data not be used without first undergoing a quality control process. Special care should be taken in using these data to calculate convective parameters and when performing trend analysis. Due to the large number of errors in especially surface and lower level dew point temperatures, convective variables (CAPE, Lifted Index) should be calculated after the QC procedure and should not be taken directly from a raw data set such as, from for example, University of Western Wyoming web site. Further investigation is needed into two types of errors which started to occur in the data in the past 10 yrs. These are the anomalously low dew points in the middle troposphere (Fig. 4.3 top) and atmospheric lapse rates larger than the dry adiabatic lapse rate which sometimes occurs above the cloud tops (Fig. 4.3 bottom).

As part of this research, techniques were developed to QC sounding data at the Irene Weather Office. Most of these techniques were adapted from Durre (2006) although a few additional checks were implemented to address problems not picked up by the checks by Durre (2006) or which was peculiar to the Irene data. Identifying errors in such a large dataset is a process which requires significant technical application to deal with the large number of individual data elements. On any particular day there could be data available on 30 pressure levels from the surface to 300 hPa. On each pressure level there are 6 different variables, each of which has to be tested individually. Despite these objective techniques, the practical experience in analysing sounding data was very useful in identifying errors. Visual inspection of the soundings helped to identify erroneous data at the commencement of this process, it served as a verification that the QC procedure identified the errors correctly and lastly it was used to confirm whether data falling outside of the climatological mean should be retained in the data set.



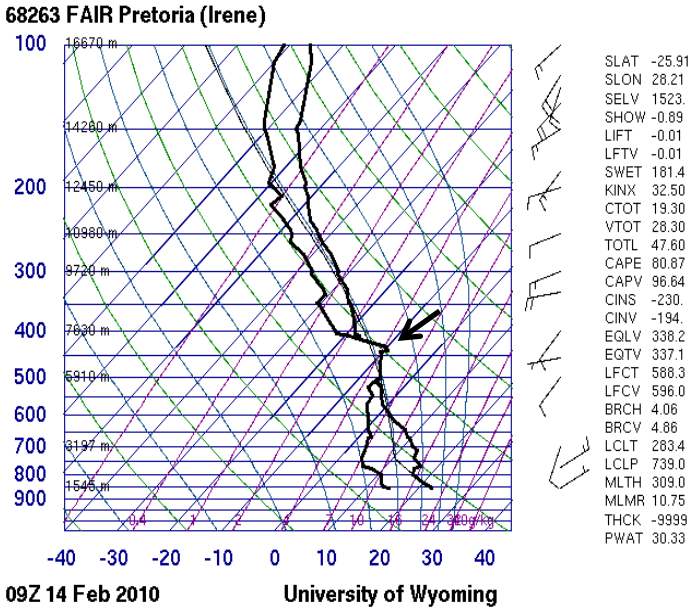


Figure 4-3: Examples of very dry dew point temperatures in the middle troposphere (top) and an absolutely unstable layer (arrow bottom). Source http://www.weather.uwyo.edu/.

4.4.4 Elimination of soundings done in cloud

One of the aims of this thesis is to determine the atmospheric conditions associated with heavy rainfall. When the radiosounding is performed in the presence of clouds it contaminates the sounding as the cloud is sampled and not the atmosphere. For this reason it was attempted to remove all soundings done in cloud by eliminating soundings where the dew point depressions of 5 °C or less occur for a depth of at least 200 hPa. There were 1887 soundings removed by employing this algorithm. When heavy rainfall occurs over Gauteng, these conditions will prevail at some time during the heavy rainfall event. An example of a saturated sounding, during a heavy rainfall event, removed from the data set is the 10th of February 2000 (Fig. 4-4) when the average rainfall over Gauteng was 43 mm and the maximum rainfall at a station 126 mm.

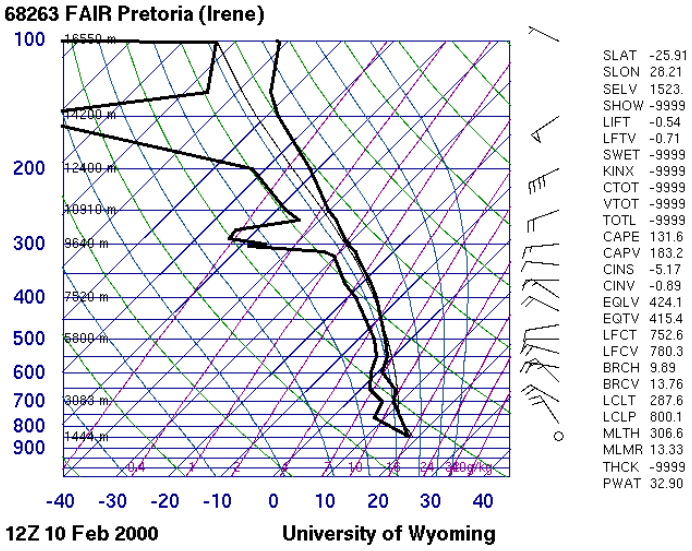


Figure 4-4: Skew-t diagram at Irene weather office on the 10th of February 2000. This sounding is an example of one with a deep saturated layer that was removed from the data set. Source http://www.weather.uwyo.edu/.

4.4.5 Methodology

Daily rainfall (24 h) is reported at 0600 UT in South Africa and is assigned to the day before. The 0000 UT sounding on any particular day therefore falls after the period of observation of the 24 h rainfall. Furthermore, rainfall over the summer rainfall area in South Africa occurs most frequently just after 1800 UT (Rouault et al., 2012). The 1200 UT sounding would therefore generally be done prior to the onset of rainfall while the 0000 UT sounding is more likely to be contaminated by cloud. As part of the QC procedure described in section 4.4.2 soundings done in cloud were eliminated from the data set. There were 1197 0000 UT soundings eliminated in this way and only

690 1200 UT soundings. Moreover, Dyson et al. (2012) indicate that AHR over Gauteng is associated with an above normal number of lightning strikes and heavy rainfall is thus predominantly convective in nature. Considering the convective nature of rainfall over Gauteng, where surface temperatures play an important role and in an attempt to sample the atmosphere prior to the occurrence of heavy rainfall, it was decided to use only 1200 UT soundings in the investigation into heavy rainfall. The convective temperature was the surface temperature in the sounding data set.

It was stated Chapter 2.2 that the number of soundings available for the 'very heavy rainfall' and 'heavy rainfall' classes defined in Chapter 2 were too few to use into an investigation of this kind. There were only 109 soundings available for analysis in the 'heavy rainfall' class for all months together and as few as 10 in March months (Table 4-4). In the 'very heavy rainfall' class from Chapter 2 there were only a total number of 25 soundings available for analysis (not shown in Table 4-4). The low number of soundings available in these two classes underlies the necessity for the 'significant rainfall' class to be used when investigating the sounding parameters associated with heavy rainfall. In the AHR class (area average rainfall > 10 mm) there were 290 soundings available. November, December and January months all had more than 50 AHR soundings and October and March less than 40. There was more than double the number of soundings available in the SHR class (Single station rainfall > 50 mm). December and January months had more than 100 SHR soundings and only 54 in October months. In section 4.7 soundings with heavy rainfall are compared to soundings where no rainfall (No Rain) occurred and the total number of these soundings was 774. It is interesting to note that in December and January months there were less No Rain soundings than SHR and AHR soundings. In Table 2.2 and 2.3 (Chapter 2) it was stated that December and January months were the months with the most days with some rain and this explains the small number of No Rain soundings in these two months. Table 4-4 also lists the total number of soundings used to create the baseline climatology (CLIM) where all available soundings were used.

Table 4-4: The number of soundings available in the different rainfall classes for the 35-year period between October 1977 and March 2012.

Month	No of CLIM soundings	No of *Heavy Rainfall* soundings	No of AHR soundings	No of SHR soundings	No of *No Rain* soundings
October	882	17	36	54	235
November	754	18	50	79	118
December	864	24	66	130	52
January	762	19	58	129	69
February	728	21	42	91	122
March	787	10	38	82	178
Total	4777	109	290	565	774

The Mann-Whitney test is used in section 4.7 to test whether there is a significant difference between the mean values of the sounding-derived parameters for AHR, SHR and No Rain soundings. The Mann-Whitney test is a nonparametric rank randomization two sample test which test the null hypothesis that two samples have the same median (e.g. Steyn et al., 1994). This test does not make the assumption that the values have a normal bell-shaped distribution like other tests, such as the Student's *t*-test. The Mann-Whitney test also only requires only 8 samples or more for reliable results.

4.4.6 Self-organizing maps

The artificial neural network system of SOMs was first described in detail by Kohonen (2001). SOMs provide a mechanism for visualizing a collection of atmospheric states after dominant modes within a dataset are identified (Tennant and Hewitson, 2002). SOM search for a number of nodes within a data space so that the distribution of the nodes characterises the observed distribution. The SOM technique identifies representative nodes spanning the data space so that individual data elements may be associated with a node. Hewitson and Crane (2002) provided detailed information on how SOMs could be used to create a synoptic climatology. The process begins by subjectively choosing the number of nodes after which the reference vectors of the nodes in the SOM array are initialised using random numbers. The SOM calculates the similarity between each data record and each of the node reference vectors. The reference vector of the node which best represents the data record is consequently modified by a user-defined factor, or learning rate. The SOM is trained twice, during the first training process random input vectors are used which produces a first set of nodes. During the second training of the SOM the node vectors identified during the first training process are used as the input vectors. In this manner the SOM provides a simplification into a small number of archetypes (Hewitson and Crane, 2002). When applying this technique to circulation patterns, or as in this case sounding data, the node maps represent physical spatial patterns. The SOM places similar circulation types neighbouring each other and very different types far apart in the SOM space. The nodes represent a non-linear distribution of overlapping, non-discreet, circulation types. Each node in a SOM represents a group of similar sounding profiles that were present in the original dataset. Any day's actual data can be uniquely associated with a closest-matching SOM state, allowing, for example, the study of frequency of occurrence (Reusch et al., 2005).

Hewitson and Crane (2002) were one of the first to apply SOMs in a meteorological application by creating a synoptic climatology of sea level pressure data over the northeastern United States. Tennant and Hewitson (2002) applied SOMs to describe rainfall characteristics over South Africa. Liu and Weisberg (2011) list the wide spectrum of application of SOMs to meteorological data since the

early 2000s. SOMs are often used to create synoptic climatologies of sea level pressure or geopotential heights at different pressure levels. SOMs are also applied to explain extreme climate events and use meteorological variables such as air temperature or wind. Van Schalkwyk and Dyson (2012) used SOMs to identify synoptic circulation types associated with different types of fog at Cape Town International Airport. Jensen et al. (2012) used ozonesondes in a SOM to discuss the vertical profile of ozone at Ascension Island and KwaZulu-Natal in South Africa. In this thesis, SOMs will be used on sounding data to create a heavy rainfall sounding climatology over Gauteng.

The data arrays which were used to train the SOM on days with AHR and SHR separately consisted of temperatures, dew point temperatures, the zonal and meridionial wind components as well as 12 sounding-derived parameters. Temperatures, dew point temperatures and wind components were interpolated to 12 pressure levels at 50 hPa intervals from 850 to 300 hPa. Twelve sounding-derived parameters were selected and included in the daily array of data. These parameters were chosen because research undertaken for this thesis, see section 4.7, indicated that they were capable of distinguishing AHR and SHR from CLIM days or because they are significant to certain types of weather systems and severe weather. The 12 parameters are:

- Precipitable water (PW)
- Mean layer dew point temperature Td_{100}
- Average temperatures in the 500-300 hPa layer (T_{53})
- Convective Available Potential Energy (CAPE)
- Vertical temperature difference from 700-500 hPa (TD_{75})
- Warm cloud depth (WCD)
- Elevated *K*-Index(EKI)
- Meridionial component of the wind in the 800-600 hPa layer (WV_{86})
- Average wind speed in the 600-400 hPa layer (W_{64})
- Magnitude of wind shear from the surface to 400 hPa (MWS_{s4})
- Change of equivalent potential temperatures ($\Delta\Theta_e$) from the surface to 500 hPa
- Average equivalent potential temperature in the troposphere ($\Theta_{e\ ave}$)

The SOM method does not require that the input variables be independent of each other and therefore makes this the ideal method to do an analysis of this type. Many of the variables and sounding parameters used to train the SOM are highly correlated and would have been problematic to use in principle component analysis. The advantage of choosing the SOM method is that any variables or parameters may be used.

All the variables used to train the SOM were standardised with respect to the average and standard deviation prior to running the SOM. Days with missing temperatures or dew point temperatures were not used as part of the training data set but days when there were missing wind data were left in the data set but were replaced by a missing data flag (x in the SOM_PAK software).

The SOM was created by using the SOM software which is freely available from (http://www.cis.hut.fi/research/som-research) and is referred to as SOM_PAK (Kohen et al., 1996). After some experimentation it was found that 20 nodes (5x4) captures the variability of the thermodynamic profile of the atmosphere sufficiently. Using more nodes would provide finer detail and less of a generalization in the thermodynamic profile but due to the relatively small number of AHR days using 20 nodes maximizes the number of days mapped to each node and allows for the calculation of meaningful statistics. The SOM_PAK software (Kohen et al., 1996) allows for user defined settings. In creating the thermodynamical climatology the following setting were utilized:

- The topology was set to rectangle
- The neighbourhood function type was set to bubble
- The training length (rlen) was set to 1 000 000 during the first and second training phases
- The learning rate parameter (alpha) was set to 0.04 during both training phases
- The radius was set to 3 during the initial training phase and 2 during the second phase.

4.5 A monthly climatology of basic variables

The temperature, dew point temperature and zonal and meridionial winds were interpolated to 850 (surface), 800, 700, 600, 500, 400 and 300- hPa pressure levels for every sounding. Monthly average values were then calculated on each of these pressure levels and the results are displayed in skew-t-gram plots (Fig. 4-5) as well as vertical time section plots (Figs 4-6 - 4-8).

Fig. 4-5 depicts the steady increase in temperature throughout the troposphere from October (black) to December (green). Fig. 4-6 shows that the seasonal temperature increase is most pronounced in the upper troposphere with an increase from -10 °C at 500 hPa in October to -7.5 °C in December and -6 °C in February. The largest temperature increase (5 °C) takes place at 300 hPa where the temperatures increase from -37 °C in October to just over -32 °C in February months. At pressures higher than 500 hPa there is on average a 2 °C increase in temperature from October to

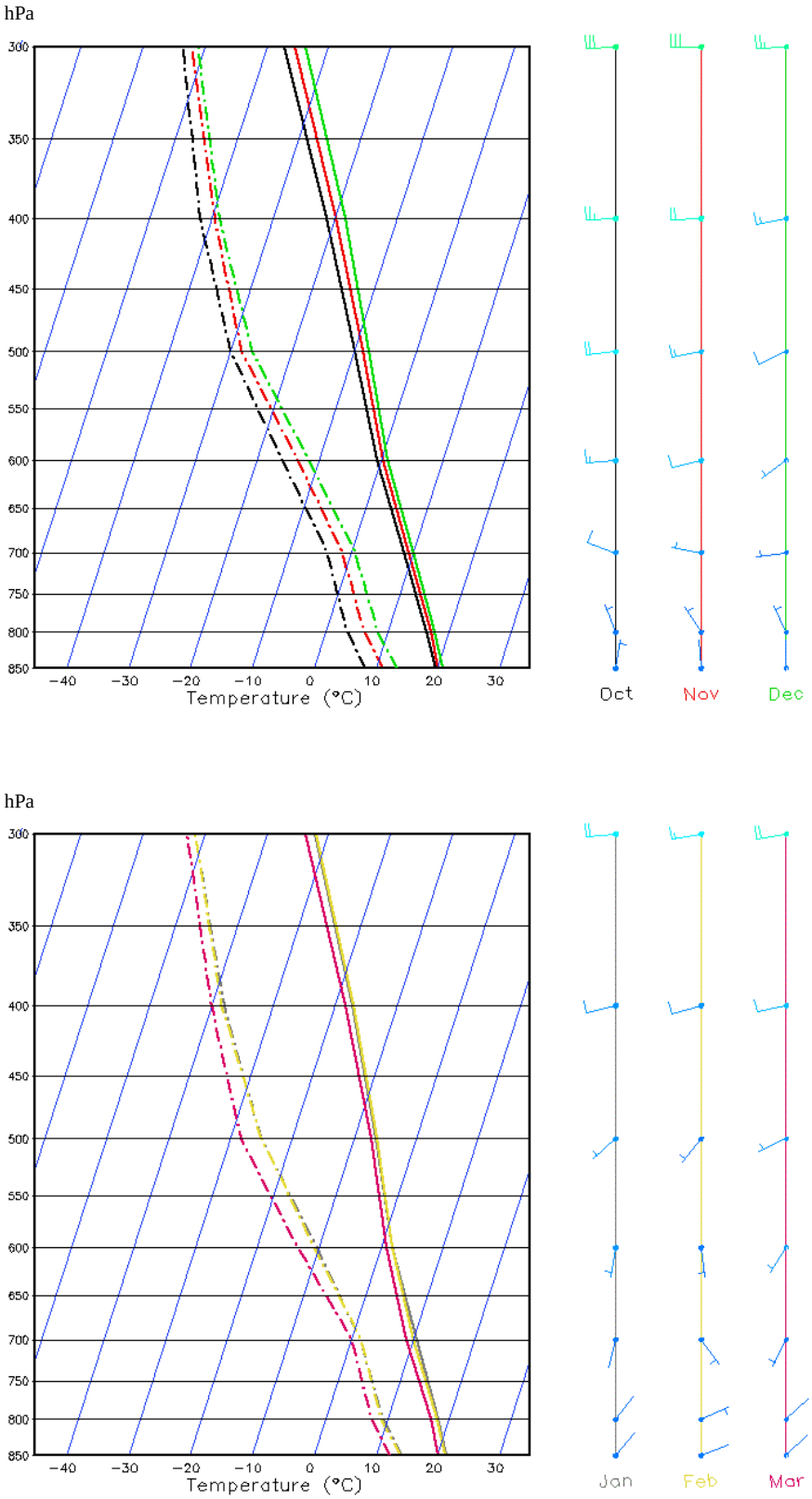


Figure 4-5: Vertical profile of monthly average 1200 UT temperature (solid lines), dew point temperature (dashed lines), and wind speed (knots) and direction at Irene for October to December (top) and January to March (bottom).

January, decreasing slightly in February and March. Temperature increase at the surface is very slight as there is only a 1 °C difference in the average surface temperature from October to January. This is also shown in Fig. 4-5 where there is very little change in surface temperature from October to March. The temperature and dew point temperature profiles for January and February months appear very similar (grey and yellow in Fig. 4-5 (right)). Detailed investigation of Fig. 4-6 reveals that from 600 hPa to the surface, the highest seasonal temperatures occur in January months but at pressures lower than 600 hPa the highest temperatures are recorded in February months. In the lower levels of the atmosphere the highest temperature occurs approximately a month before the highest temperatures occur in the middle and upper troposphere.

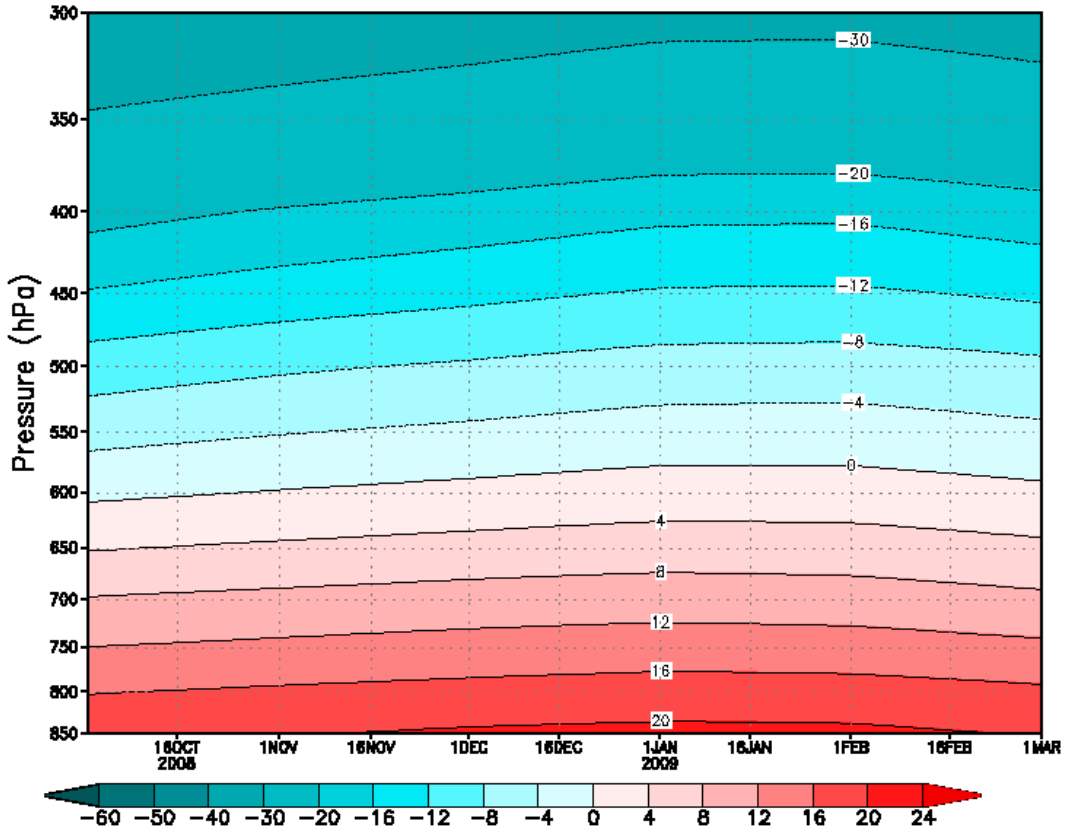


Figure 4-6: Vertical cross section of monthly average 1200 UT temperatures (°C) for October to March at Irene.

The dew point temperatures also increase throughout the troposphere from October to January and decrease again in March. On average the surface dew point temperatures are just higher than 8 °C in October months increasing to more than 14 °C in January and February months. Fig. 4-7 depicts the steady increase in RH values during the summer season so that in the last three months of the season the RH values exceed 55 % in the lowest 3 km of the atmosphere. In the upper troposphere the RH values increase by almost 5 % from October to February decreasing again in March.

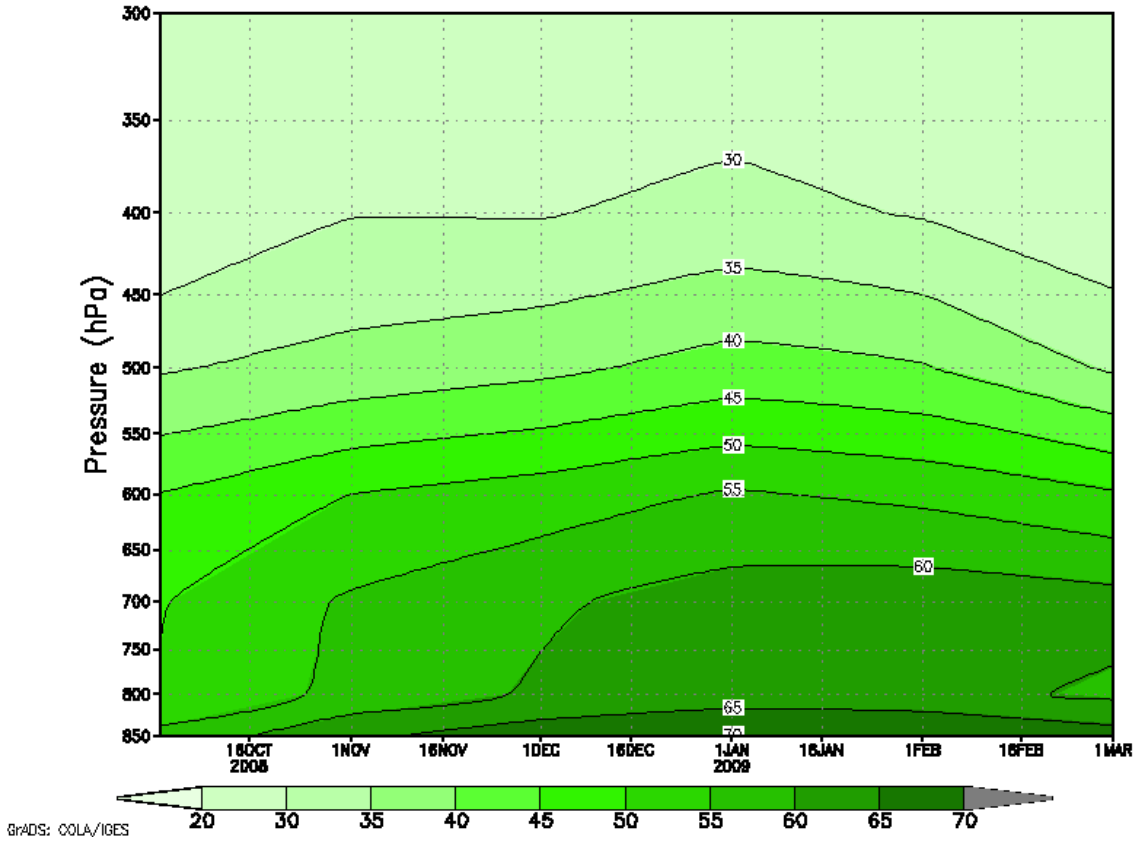


Figure 4-7: Vertical cross section of monthly average 1200 UT relative humidity (%) for October to March at Irene.

The winds at pressures lower than 700 hPa are mostly westerly in the early summer season and increasing in strength with altitude (Fig. 4-8). From 500 to 400 hPa the winds back slightly from October to December, resulting in southwesterly winds at these pressure levels in January and February (Fig. 4-5 and Fig. 4-8). At 700 hPa the winds back during the season to southerly and even southeasterly in February. At the surface the winds are very light and change in direction from northwesterly to northeasterly during the season. The wind strength at all pressure levels decreases from October to February. At 400 hPa the wind strength decreases from about 25 knots in October to 8 knots in February. In January and February months the average wind speeds throughout the troposphere is very light and only at pressures lower than 400 hPa does the wind speed exceed 10 knots. In February months the winds are less than 5 knots from the surface to 500 hPa. There is only a slight increase in wind strengths in March months and this is more pronounced in the upper troposphere.

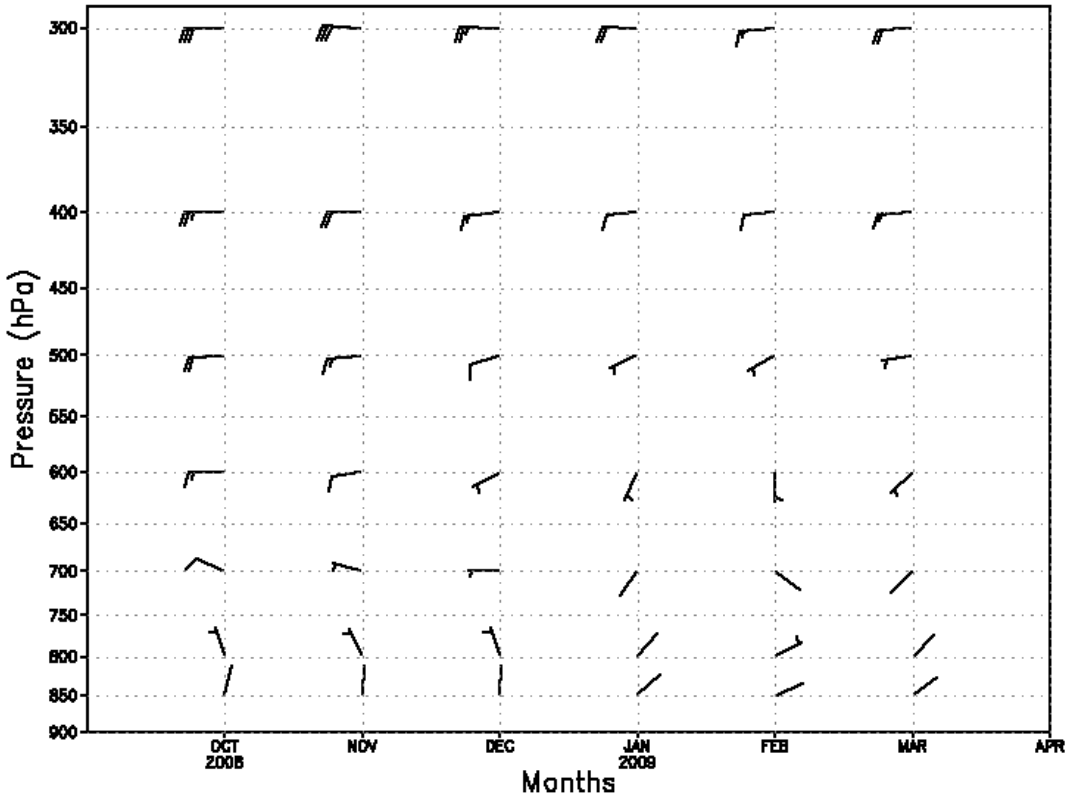


Figure 4-8: Vertical cross section of monthly average 1200 UT wind speed (knots) and direction for October to March at Irene.

4.6 A monthly climatology of sounding-derived parameters.

4.6.1 Temperatures and temperature lapse rates

Considering the upper and mid-tropospheric temperatures, the average 500-300 hPa temperatures increase from -22 °C in October to -17.5 °C in January and February months. The same variation in temperatures is visible at 500 hPa and 700 hPa where the temperatures changes from 8.5 °C in October months to 10.5 °C in January (Fig. 4-6). Temperatures at the surface and 700 hPa increase by about 2 °C from minimum to maximum during the season while at 500 hPa the increase in average temperature is about 4 °C and 5 °C at 300 hPa. The temperature increase in the middle and upper troposphere is larger than at lower pressure levels and this causes a decrease in temperature lapse rates in the late summer. The decrease in temperature lapse rate indicates a decrease in the conditional instability in the atmosphere (Doswell et al., 1985). Conditional instability is defined as when the atmospheric lapse rate is steeper than a moist adiabat but less steep than a dry adiabat. The typical moist adiabatic lapse rate is 5.5 °C km^{-1}; the standard atmospheric lapse rate is 6.5 °C km^{-1} while the dry adiabatic lapse rate is 9.8 C km^{-1}. The average 600-300 hPa temperature lapse rate (TL) (brown in Fig. 4-9) is highest in October (7.1 °C km^{-1}) and the value approaches that of the dry adiabatic temperature lapse rate. The lowest TL in the 600-300 hPa level occurs in February months (6.5 °C km^{-1}) when the lapse rate, on average, is closer to that of the standard atmosphere and confirms the decrease in the conditional instability in the atmosphere (Fig. 4-9). February is also the month with the highest temperatures at pressure levels lower than 600 hPa (Fig. 4-6). The surface to 600 hPa TL is larger than the 600-300 hPa TL for all months. Both lapse rates have a similar seasonal variation as the lowest TL_{63} value also occurs in February months when the value is just lower than 7 °C km^{-1}.

4.6.2 Geopotential thickness and freezing levels

The geopotential thickness between 700-300 hPa is 6420 gpm in October months and as the temperatures in the troposphere increase towards January the thickness also increases to more than 6510 gpm in January and February months. Likewise the height of the freezing level (0°C isotherm) increases from 4360 m to 4820 m a. m. s. l. in February (Fig. 4-10). The height of the freezing levels in December and March are very similar at ± 4650 m

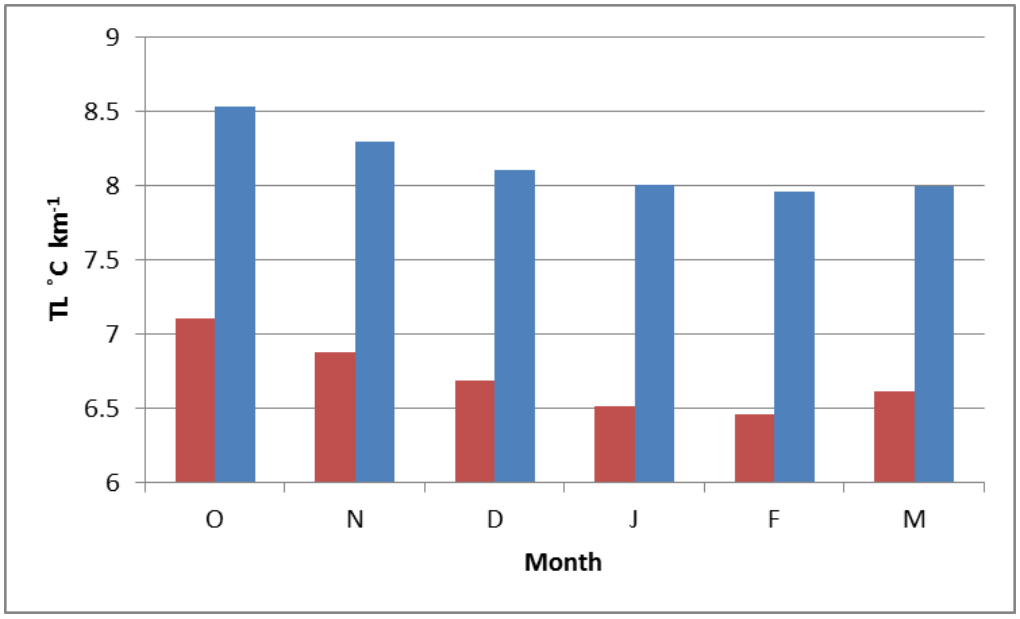


Figure 4-9: The average monthly 1200 UT temperature lapse rates (TL) (°C km^{-1}) between 600-300hPa (blue) and the surface to 600 hPa (brown) for October to March at Irene.

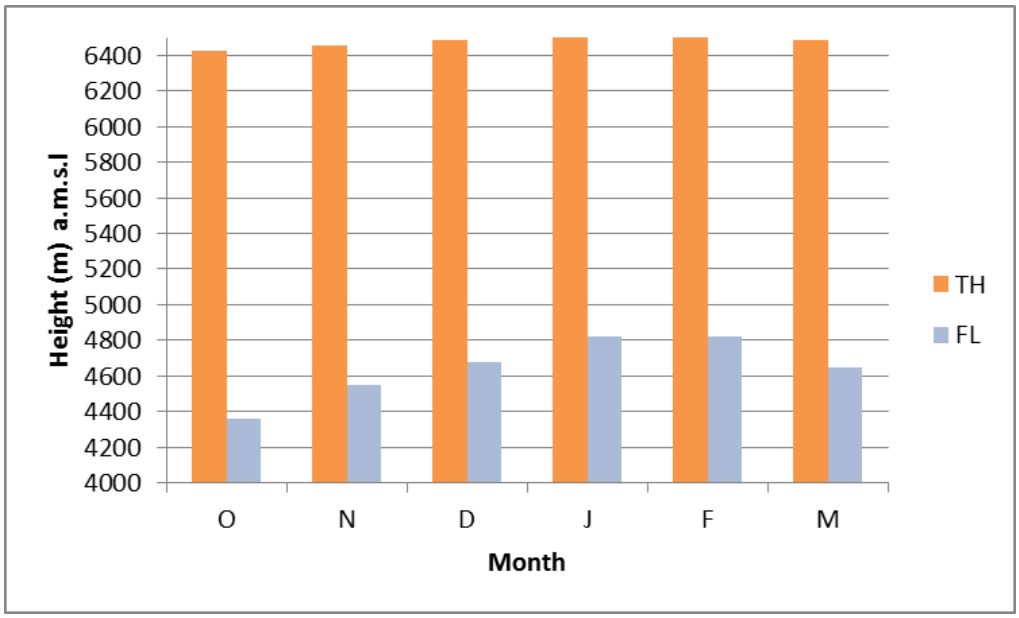


Figure 4-10: The average monthly 1200 UT 700-300 hPa geopotential thickness (TH) (orange) and freezing level height (FZ) (blue) from October to March at Irene.

4.6.3 Moisture variables

In Fig. 4-11 the monthly average values of the mean layer mixing ratio (M) and the precipitable water (PW) show how the moisture content of the atmosphere increases towards late summer. The mixing ratio is 7 gkg^{-1} in October and above 10 gkg^{-1} in January and February months. Precipitable water values increase from 15.5 mm in October to about 22 mm in January and February months decreasing again in March. The maximum precipitable water value which occurred at Irene during this 34-yr period was 37.7 mm on the 12th of January 2009 and the lowest 1.85 mm on the 2nd of October 2002 (excluding soundings done in cloud). No rainfall occurred on the 2nd of October 2002 and on the 12th of January 2009 the average rainfall over Gauteng was 13 mm and the maximum rainfall recorded at any particular station was 62 mm. The PW value exceeded 30 mm on only 154 occasions; this is less than 2 % of available soundings and none of these days were in an October month. Fig. 4-11 shows how the average daily maximum rainfall over Gauteng increases during the season as the precipitable water increases.

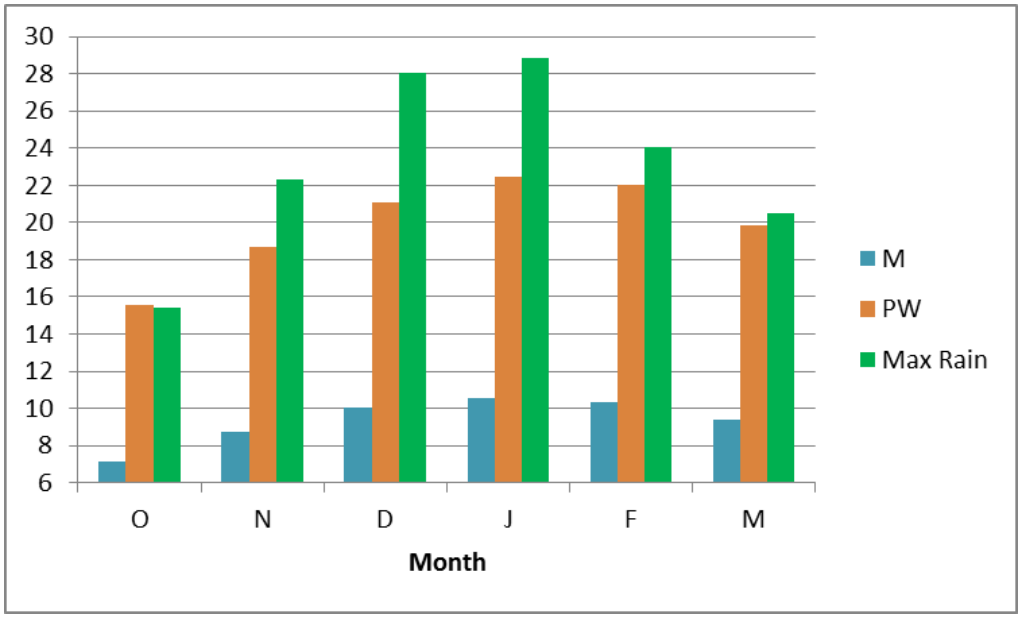


Figure 4-11: The average monthly 1200 UT mixing ratio (M) (gkg^{-1}) (blue), precipitable water (PW) (mm) (orange) and average maximum rainfall (Max) at a single station (mm) (green) for October to March at Irene.

4.6.4 Equivalent potential temperatures

Fig. 4-12 depicts the increase of 7 K in the average tropospheric Θ_e values from October to January. Considering that Θ_e is a function of temperature and moisture (eq. 4-1) this increase during the season corresponds to the increase in moisture and temperature during the same period. Fig. 4-13 depicts the $\Delta\Theta_e$ values between the surface and 500 hPa. In October months the $\Delta\Theta_e$ values are -14 K increasing to -20 K from December to February. The conditional instability of the atmosphere over Gauteng decreases as the summer season progresses (and the temperature lapse rates decrease). Conversely the atmosphere becomes increasingly convectively unstable towards late summer as $\Delta\Theta_e$ values increase. The ingredients required for deep moist convection is lift, instability and moisture (Schultz et al., 2000) and they argue that two of the three ingredients are accounted for in a single combined parameter either CAPE or potential (convective) instability.

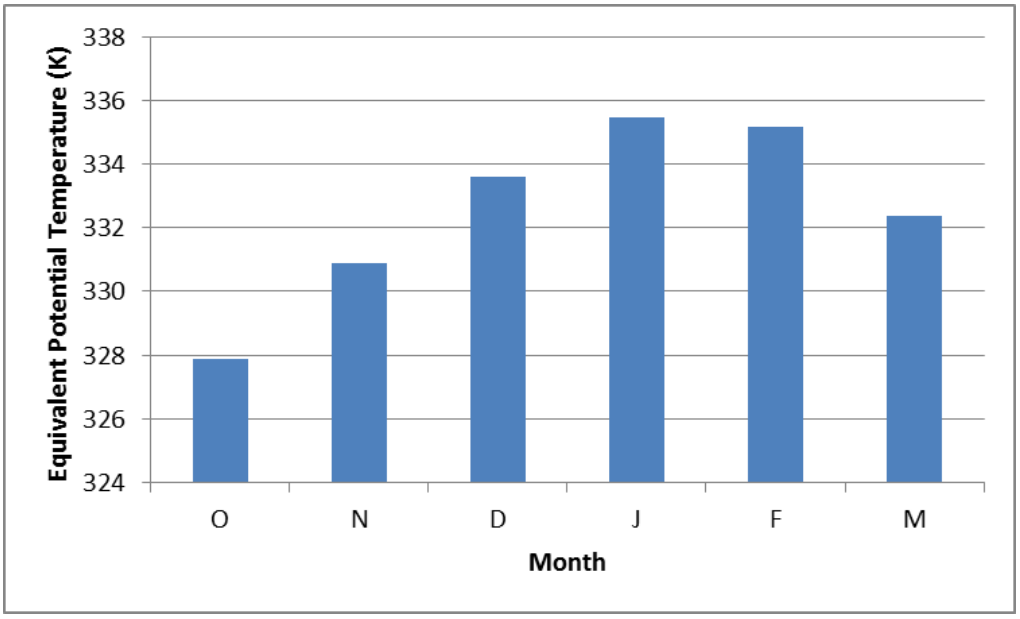


Figure 4-12: Average monthly 1200 UT equivalent potential temperatures in the troposphere (Θe ave) for October to March at Irene.

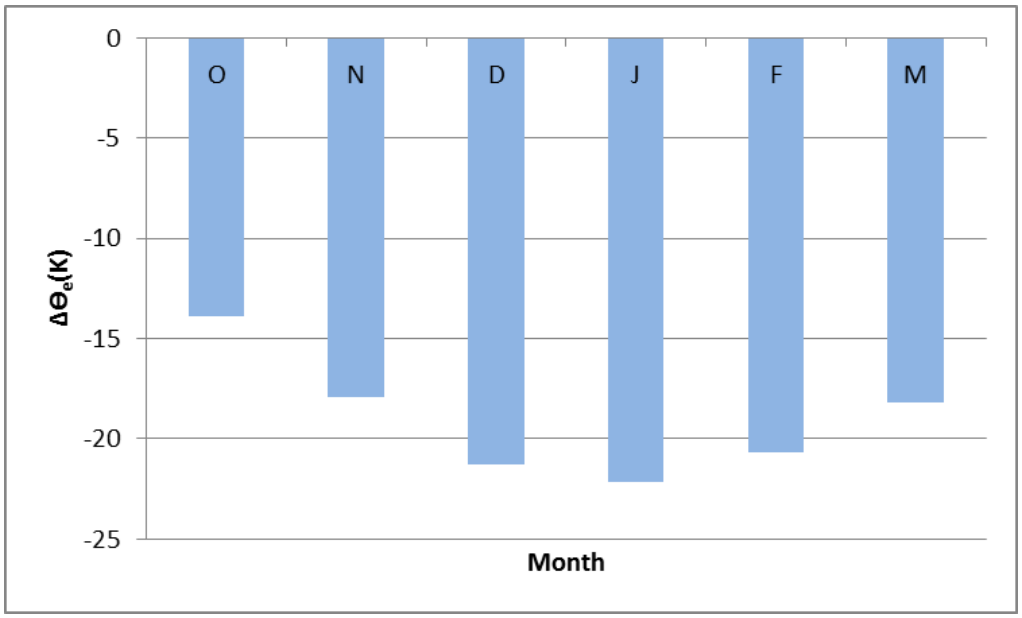


Figure 4-13: Average monthly 1200 UT equivalent potential temperature lapse rate ΔΘe (K) for the surface to 500 hPa for October to March at Irene.

4.6.5 Winds

The speed of movement of convective systems is determined by wind speeds and slower moving systems have higher precipitation rates (Doswell et al., 1996). Fig. 4-8 depicts the decrease in wind speeds during the summer season throughout the entire troposphere. Mesoscale convective systems (MCS) which develop under light wind conditions would move very slowly and could lead to high precipitation rates and heavy rainfall. Fig. 4-14 shows that the average wind speeds in the 600-400 hPa layer decreases from 11 ms^{-1} in October months to nearly half that value (6 ms^{-1}) in February months. MCS in late summer would move slower and potentially result in higher rainfall.

Fig. 4-14 shows that during early summer (October-December) the average meridional wind direction from 800 to 600 hPa is northerly (negative values) and southerly in late summer. The northerly winds in October are double the strength of the light southerly winds in December, January and February months. It will be shown later how low-level wind directions during heavy rainfall events are northerly.

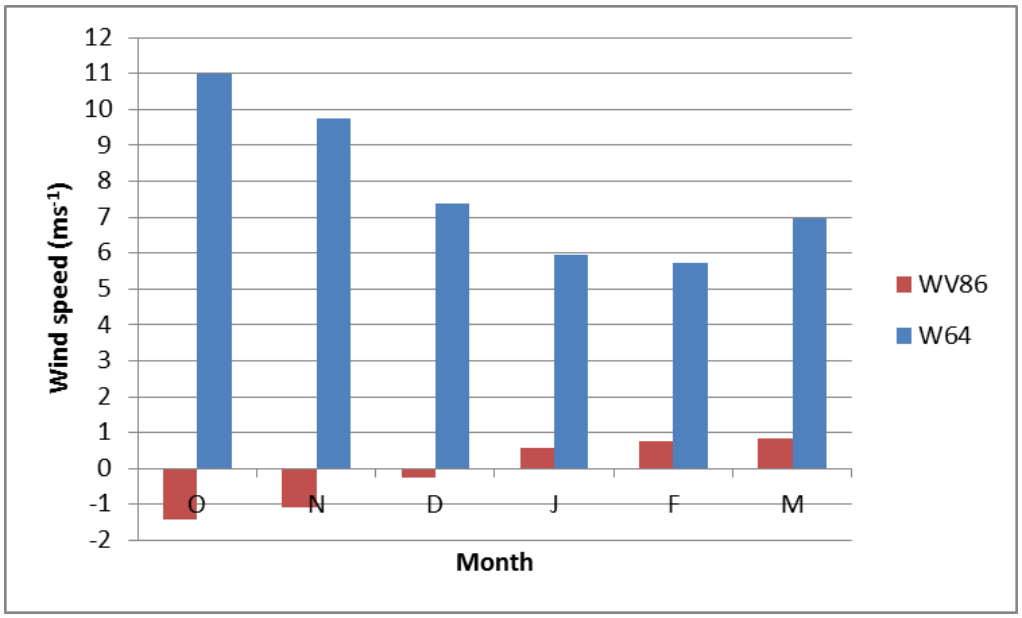


Figure 4-14: The average monthly 1200 UT wind speed in the 800-600 hPa level (W64) (blue) and the average meridionial wind in the 800-600 hPa layer (WV86) for October to March at Irene. All winds are in ms^{-1}.

At Irene the average monthly magnitude of the wind shear in the surface to 700 hPa layer (WS_{s7}) decreases from about 5.8 ms^{-1} in October months to a minimum of 4.6 ms^{-1} in February months (Fig. 4-15). Considering the magnitude of wind shear values in the level from the surface to 400 hPa (WS_{s4}) (400 hPa is approximately 6 km above ground level (AGL) over Gauteng) the values are again largest in October months (13 ms^{-1}) and reach a minimum of about 7.5 ms^{-1} in February. The average WS_{s6} are smaller than the WS_{s4} values and the minimum values are also observed in February months.

4.6.6 *K*-Index, Elevated K-Index, Total Totals Index and Elevated Total Totals Index

The monthly average value of the KI is 29 °C in October but above 30 °C in all other summer months (Fig. 4-16) with the maximum value in January months (37 °C). The average values of the KI in December, January and February are higher than the threshold that has previously been found to identify heavy convective rainfall (De Rubertis, 2006). The values for the EKI are much smaller than the KI and may even have negative values when the lower and mid-levels are very dry. Average values are negative in October months increasing to 5 °C in January months. As elaborated on in section 4.6.1 - the temperature lapse rate decreases in late summer but the moisture at the surface increases causing an associated increase in the KI and EKI.

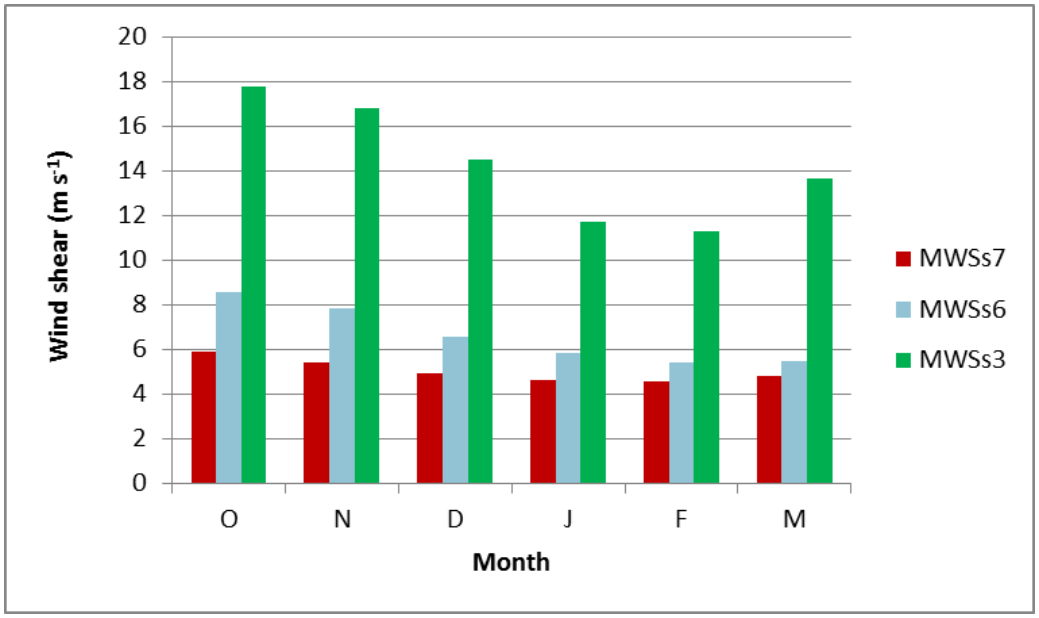


Figure 4-15: The average monthly 1200 UT magnitude of surface to 700 hPa wind shear (MWS_{s7}) (m s^{-1}) (brown) the bulk wind shear from the surface and 600 hPa (WS_{s7}) ($ms^{-1}km^{-1}$) (blue) and the surface and 300 hPa (WS_{s3}) (green) for October to March at Irene.

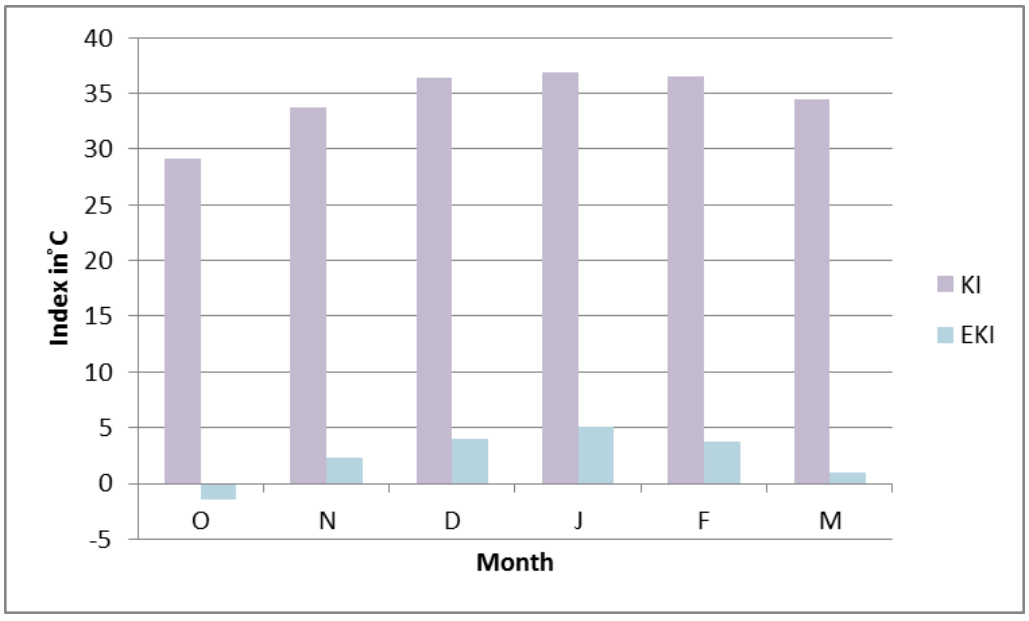


Figure 4-16: The average monthly 1200 UT K-Index (KI) (purple) and Elevated K-Index (EKI) (blue) for October to March at Irene.

The TTI relies on the same principles as the KI but this time not taking the mid-level (700 hPa) moisture into consideration. Larger temperature lapse rates, lower 500 hPa temperatures and increased surface moisture will give higher TTI values. TTI is designed in such a manner that the 500 hPa temperatures have an increased influence on the final value of the index. The lowest values occur in March months (50 °C) and the highest in December months (53 °C) (Fig. 4-17). Values vary only slightly during the season. The influence of the decreasing temperature lapse rates and increasing 500 hPa temperatures during the season will lower the TTI values, however the increase in the 850 hPa dew point temperatures ensures that the value does increase towards late summer, even if it is just slightly. Elevated Total Totals Index (ETTI) has lower values but a similar seasonal variation to the TTI. The average monthly ETTI values do not vary much as the values hover around 24 °C from October to January decreasing slightly in February and March.

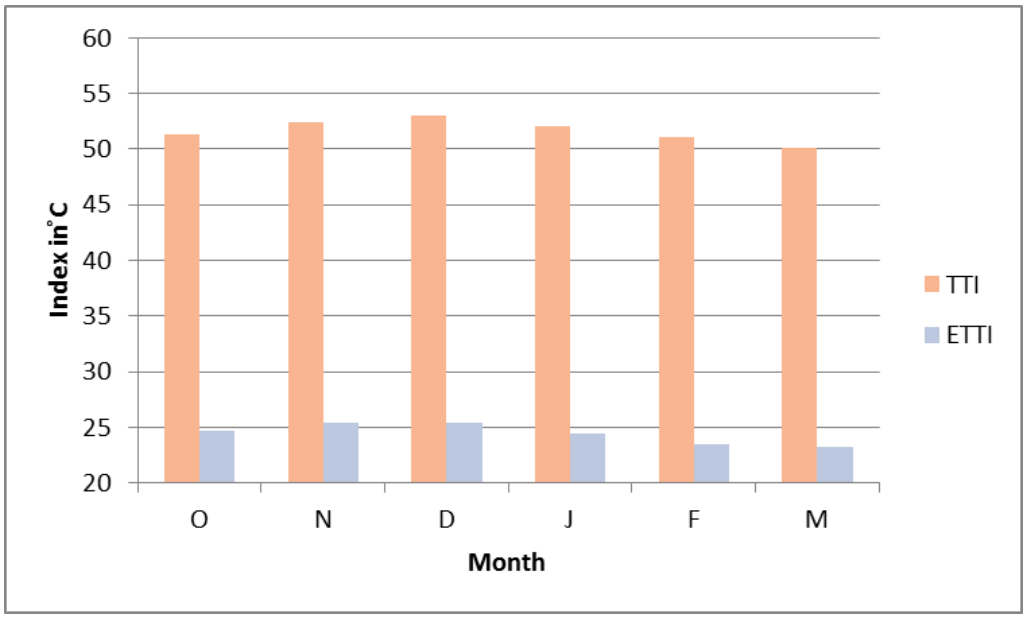


Figure 4-17: The average monthly 1200 UT Total Totals Index (TTI) (orange) and Elevated Total Totals Index (ETTI) (blue) for October to March at Irene.

4.6.7 Convective variables

The seasonal variation of Convective Available Potential Energy (CAPE) at 1200 UT is shown in Fig. 4-18. The average CAPE value is just less than 500 Jkg^{-1} in October but more than 800 Jkg^{-1} in December and January months when the highest average CAPE value occurs. CAPE is a function of the lifted parcel moisture content and temperature lapse rate. DeMott and Randall (2004)

investigate whether CAPE trends are driven by temperature lapse rate changes or by an increased low-level water vapour contents and conclude that trends in atmospheric temperature lapse rate play a secondary role to low-level water vapour content. Furthermore they found that CAPE and PW is generally highly correlated. Considering the TL (Fig. 4-9) and PW (Fig. 4-11) values during the summer season at Irene, it is evident that both variables play an important role in the seasonal distribution of CAPE. The TL is highest in October months, but the CAPE and SI (Figs. 4-18 and 19) values are the lowest in this month due to the absence of moisture. The temperatures in the middle and upper troposphere increase from October to December decreasing the TL as the season progress (Fig.4-6 and Fig. 4-9). However, in December months the TL is still relatively large contributing the large CAPE values. The dew point temperature at the surface increases by 5 °C from October to December and a combination of the relatively high TL and high surface moisture content result in the highest monthly average CAPE values in December months. Even though the atmosphere contains more moisture in February than in November months (Fig. 4-11) the CAPE values are larger in November months when the TL are larger. It is, therefore, important to know whether temperature lapse rates or low-level moisture content are more dominant when identifying climatological significant CAPE values for predicting heavy rainfall.

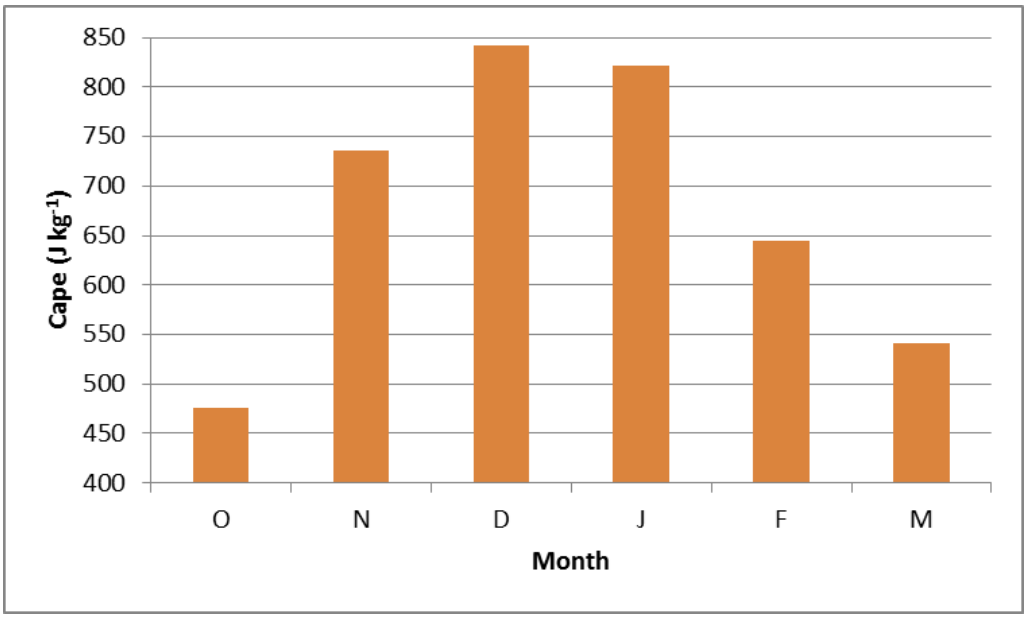


Figure 4-18: The average monthly 1200 UT Convective Available Potential Energy (CAPE) (J kg^{-1}) for October to March at Irene.

Williams and Renno (1993) and Murthy and Sivaramakrishnan (2006) expand on the relationship between CAPE and equivalent potential temperatures in the boundary layer. Average monthly Θ_e values in the boundary layer increase from about 330 K in October to 340 K in January (not shown). There is also good association between the increase in convective instability during the season (Fig. 4-13) and the increase in CAPE.

The Showalter Index (SI) is negative throughout the summer, but with larger negative values later in the season, the minimum monthly value occurring in December months (-2) (Fig. 4-19). In Fig. 4-6 the increase in the temperatures at pressure levels less than 600 hPa is clearly seen from October to December and is also reflected in a decrease in the temperature lapse rates (Fig. 4-9). Also evident in Fig. 4-5 is the increase in lower tropospheric moisture which in turn causes an increase in the convective cloud base pressure (or lifted condensation level (LCL)) as well as the pressure of the level of free convection (PFC) (Fig. 4-20). The PFC increases in pressure from 620 hPa to 675 hPa in February months. Fig. 4-21 shows the increase in warm cloud depth from 50 hPa in October to 130 hPa in January and February months.

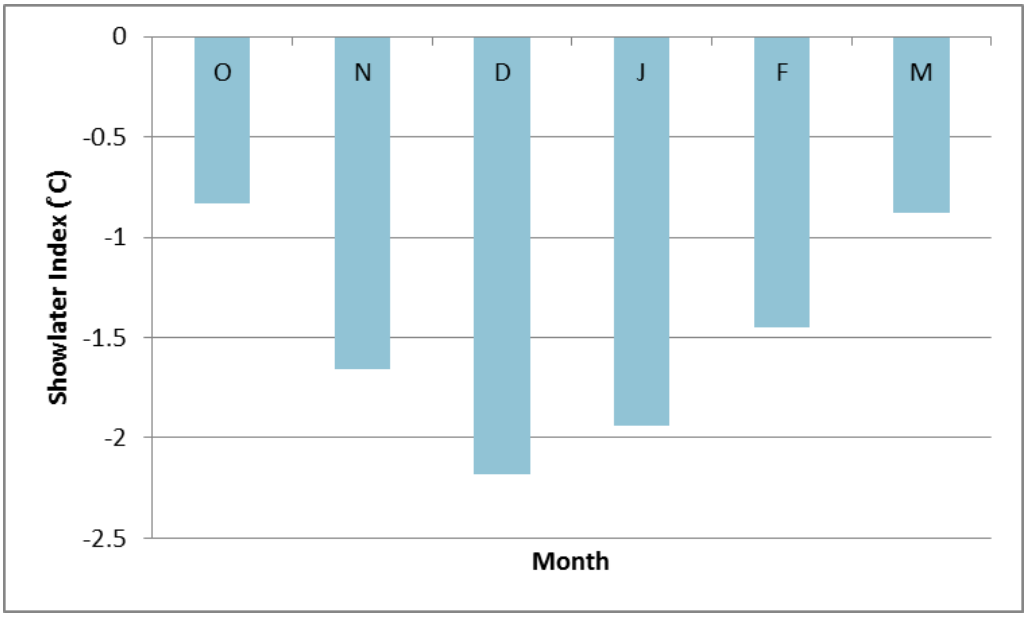


Figure 4-19: The average monthly 1200 UT Showlater Index (SI) for October to March at Irene.

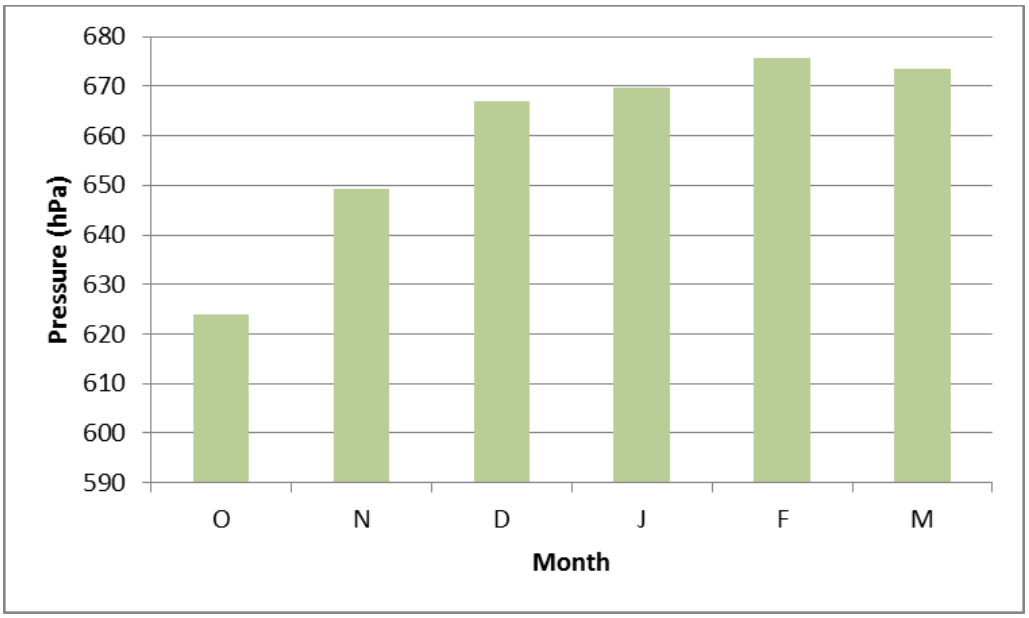


Figure 4-20: The average monthly 1200 UT Pressure of Free Convection (PFC) for October to March at Irene.

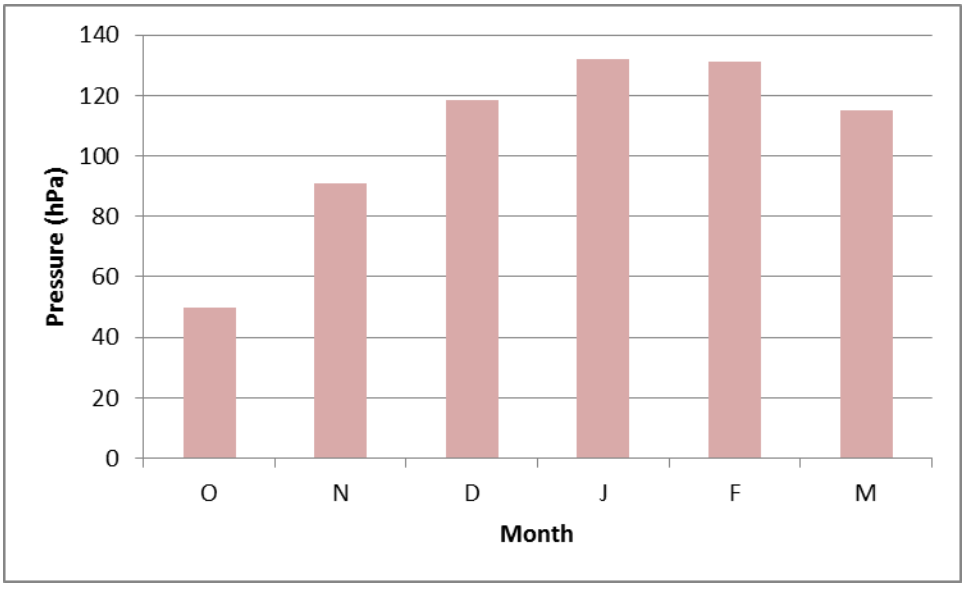


Figure 4-21: Average monthly 1200 UT Warm Cloud Depth (hPa) for October to March at Irene.

4.7 Sounding-derived parameters associated with heavy rainfall and dry days

Due to the large number of graphs generated in this analysis, the average heavy rainfall (AHR) class will be discussed here and the single station heavy rainfall (SHR) will only be included in the discussion if there are significant differences with the AHR. The parameters for AHR soundings are compared to soundings without rain (No Rain) as well as to the long-term average of climatological values (CLIM)

4.7.1 Moisture

For all months the dew point temperatures are higher for AHR soundings than for No Rain soundings. This is true for the entire troposphere and for AHR soundings the average dew point in the middle troposphere is 6 (and more) °C higher than the average value (Fig. 4-22). The same is true for SHR but the mid-tropospheric dew point temperature anomaly is about 4 °C. Fig. 4-22 also shows that from the surface to 700 hPa the dew point temperatures are about 2 °C higher than the CLIM values when AHR and SHR occur (not shown). Investigation of Fig.4-23 reveals that in October (November) months AHR soundings have surface dew point temperatures 5 °C (3 °C) higher than normal. In December, January and February the dew point temperatures for AHR are only about 1 °C higher than the CLIM values. The Mann-Whitney test shows that all the differences in surface dew point temperature between No Rain and AHR soundings are significant with 1 % of freedom. For SHR soundings the surface dew point temperatures have the same positive anomalies than for AHR.

Fig. 4-24 shows the box and whisker diagrams for No Rain and SHR soundings for PW. The Mann-Whitney test found statistically differences at the 1 % significance level between SHR and No Rain for all months. The PW values are significantly higher for SHR soundings than for dry soundings for all months and it is also at least 3 mm higher than the CLIM value for any particular month. For No Rain soundings the mean monthly values are about 3 mm less than the CLIM values. This graph is a clear indication of how the seasonal variation in PW should be taken into consideration when setting threshold values to predict heavy rainfall. The CLIM value in January months (22 mm) is higher than the SHR value in October months (20 mm).

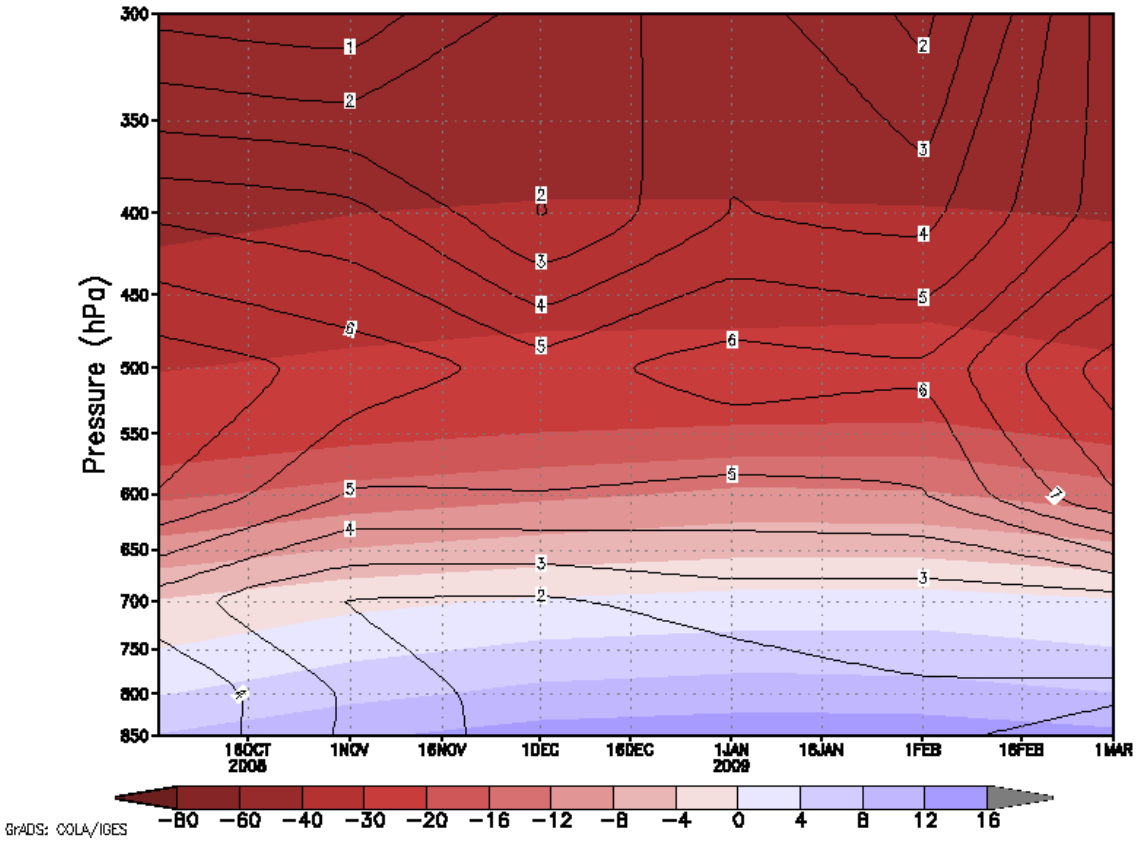


Figure 4-22: Vertical profile of the CLIM dew point temperatures at 1200 UT at Irene (shaded) as well as the anomalies for area-average heavy rainfall AHR (contours).

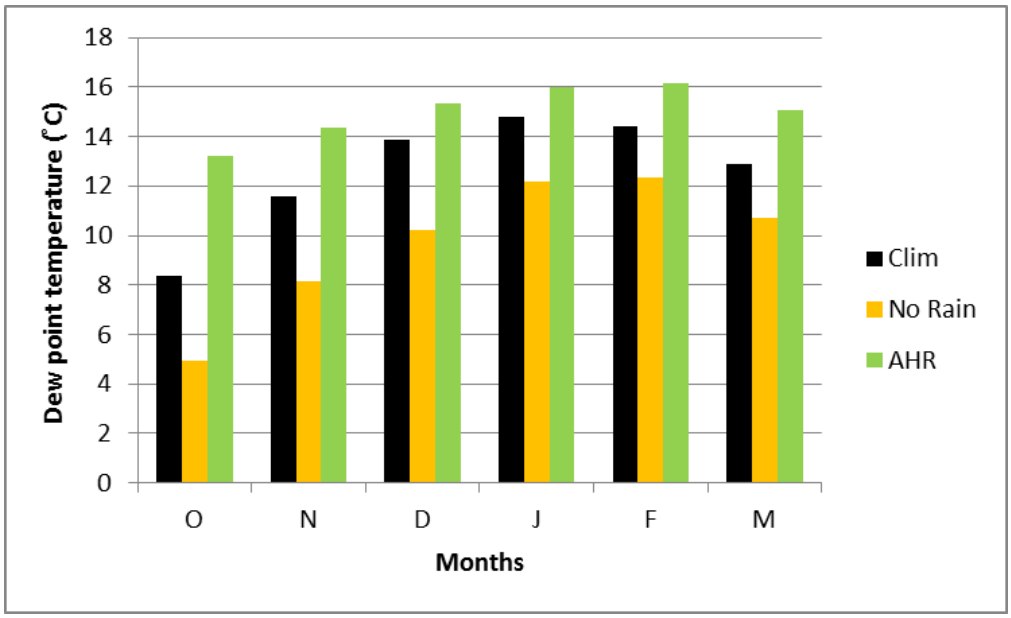


Figure 4-23: Average monthly surface dew point temperatures for 1200 UT CLIM soundings (black), No Rain (yellow) and area-average heavy rainfall AHR (green).

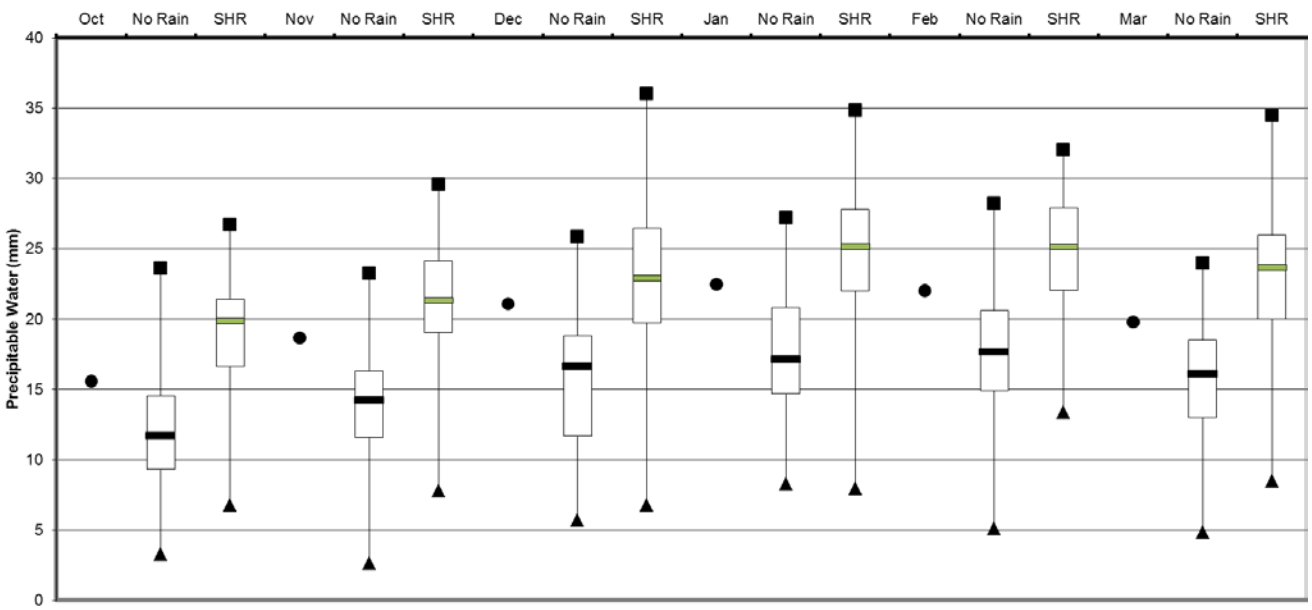


Figure 4-24: Monthly box and whisker diagram of PW at Irene at 1200 UT. The black bars are for soundings with No Rain and the green for single station heavy rainfall (SHR). The circles are the CLIM values at 1200 UT. The bars indicate the mean value, the rectangles the maximum and the triangle minimum values. The top of the clear rectangles are the third quartile and the bottom the first quartile. The star inside the open rectangle indicates that the Mann-Whitney test did not find the mean between the dry and heavy rainfall sounding to be significant with at most a 5 % significance level.

Table 4-5 shows the monthly average PW values for CLIM, AHR and SHR soundings. During the early season PW values of about 20 mm could help to identify days with heavy rainfall, however in late summer a value of 25 mm is more appropriate.

Table 4-5: Average monthly 1200 UT PW values for AHR and AHR soundings as well as long-term average values (CLIM) at 1200 UT.

Month	PW (mm) for AHR (SHR)	PW (mm) for CLIM
October	21 (19)	16
November	22 (21)	19
December	24 (23)	21
January	26 (25)	23
February	26 (25)	22
March	25 (23)	20

4.7.2 Temperatures

Fig. 4-25 depicts the seasonal progression of AHR soundings. Much like the CLIM soundings (Fig. 4-5) the temperatures throughout the troposphere for AHR soundings increase during the season.

For AHR soundings the temperatures are slightly lower than normal (Fig. 4-26) at pressures higher than 400 hPa in early summer but higher than normal for late summer throughout the troposphere, except for the surface. However, these temperature deviations are very small except in the surface layers. The temperature deviation is a maximum in March months (0.8 °C) in the higher levels of the troposphere. At the surface AHR events are associated with below normal surface temperatures for all months but these deviations are larger in early summer. In October and November months the surface temperatures for AHR are 2 °C lower than normal.

For No Rain soundings the temperatures are lower at 700 hPa with as much as 0.8 °C in November and January months. In the mid-levels (600-400 hPa) the temperatures are slightly higher for all months. At pressures lower than 400 hPa the temperatures are slightly below normal in October, November and March months but above normal in January February and March.

The 500 hPa temperature values for SHR soundings are slightly lower than for No Rain soundings but very similar to the CLIM values (Fig. 4-27). The Mann-Whitney test found that only in October, January and February months are the differences between No Rain and SHR soundings significant with a level of least 5 %. The box and whisker diagram for the average 500-300 hPa temperatures shows that the mean values between No Rain soundings and soundings with heavy rainfall are very similar but slightly higher for both heavy rainfall classes (Fig. 4-28). However, these differences are hardly significant as the Mann-Whitney test found that only in March months are differences significant to a level of 5 %. SHR is generally associated with lower temperatures in the middle troposphere but warmer temperatures in the 500-300 hPa layer.

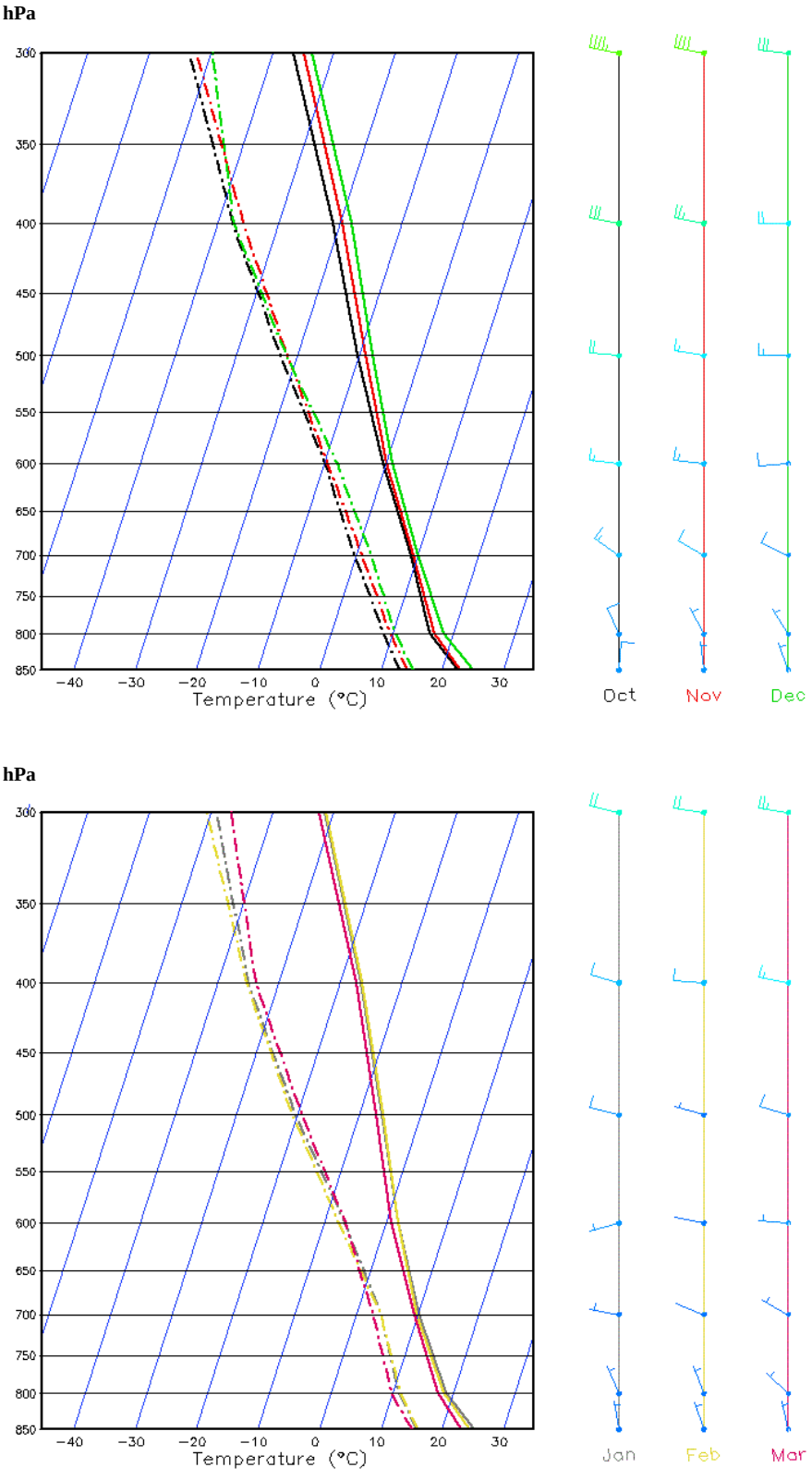
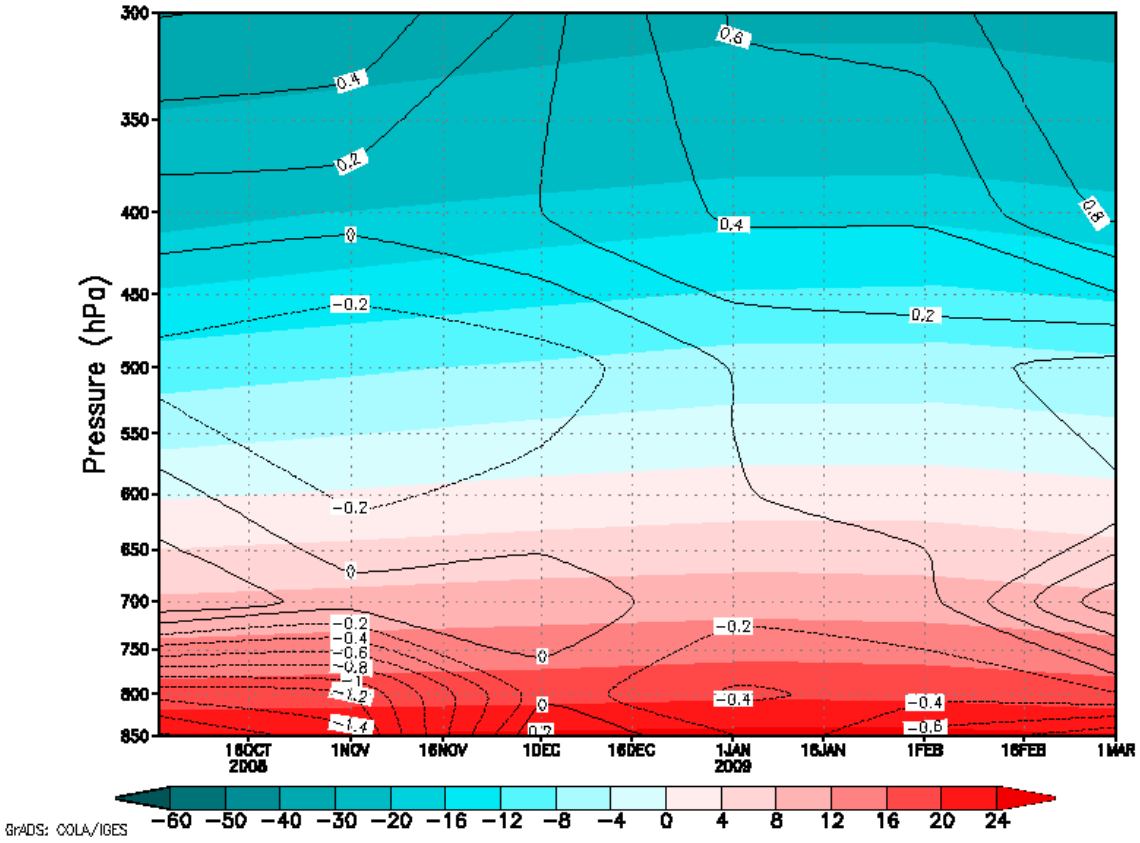


Figure 4-25: The 1200 UT vertical profile of monthly average temperature (solid lines), dew point temperature (dashed lines), and wind speed (kts) and direction at Irene for October to December (top) and January to March (bottom) for area-average heavy rainfall (AHR) days .



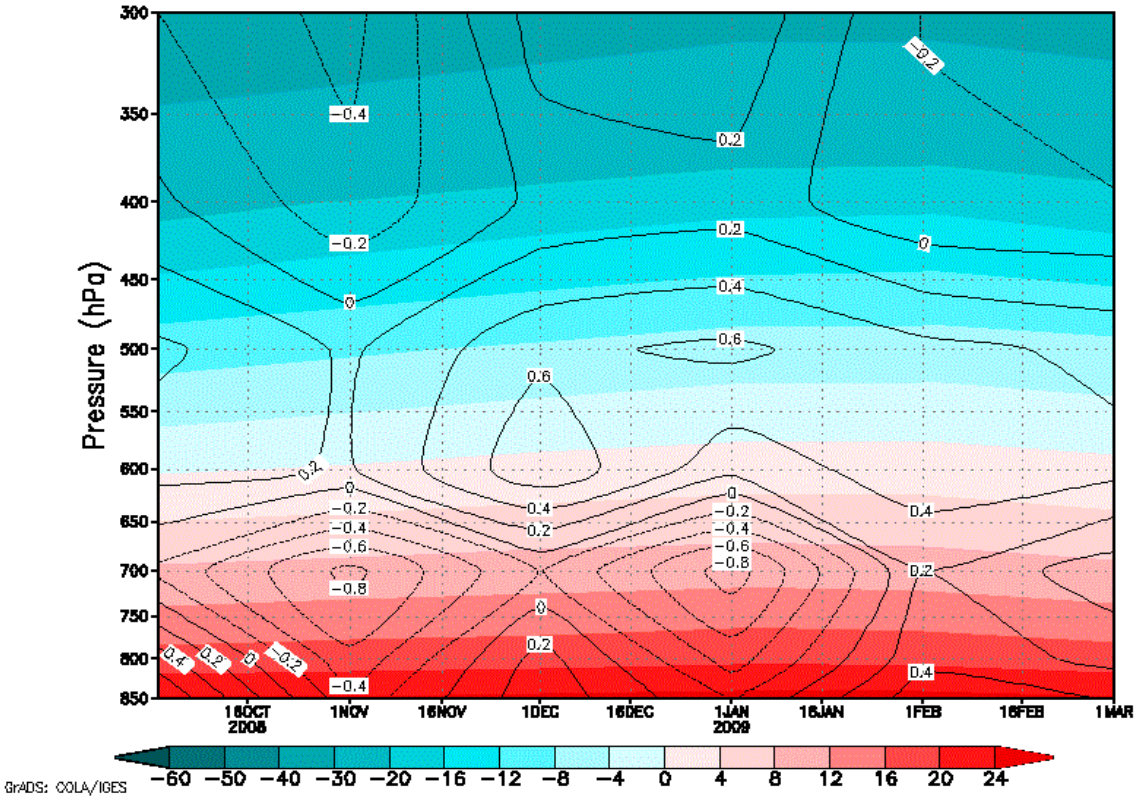


Figure 4-26: The 1200 UT vertical profile of the monthly CLIM temperatures at Irene (shaded) as well as the anomalies for area-average heavy rainfall (AHR) soundings (contours left) and no rain soundings (right).

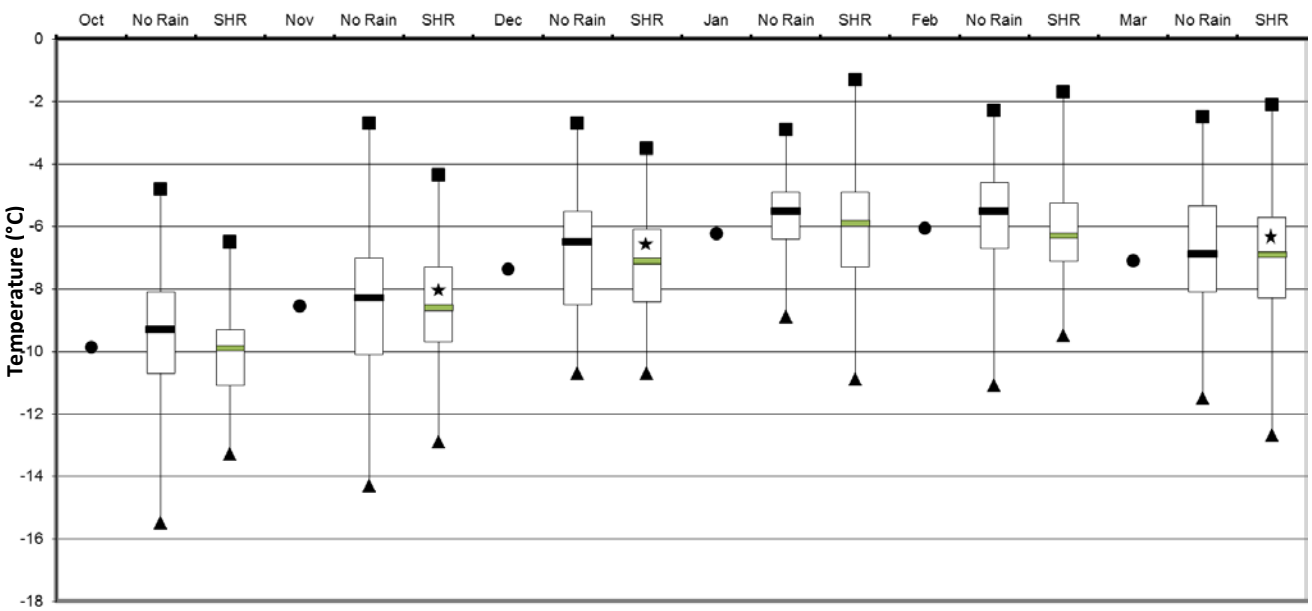


Figure 4-27: Monthly box and whisker diagram of 500 hPa temperatures (°C) at Irene at 1200 UT. The black bars are for soundings with No Rain and the green for single station heavy rainfall (SHR). The circles are the CLIM values at 1200 UT. The bars indicate the mean value, the rectangles the maximum and the triangle minimum values. The top of the clear rectangles are the third quartile and the bottom the first quartile. The star inside the open rectangle indicates that the Mann-Whitney test did not find the mean between the dry and heavy rainfall sounding to be significant with at most a 5 % significance level.

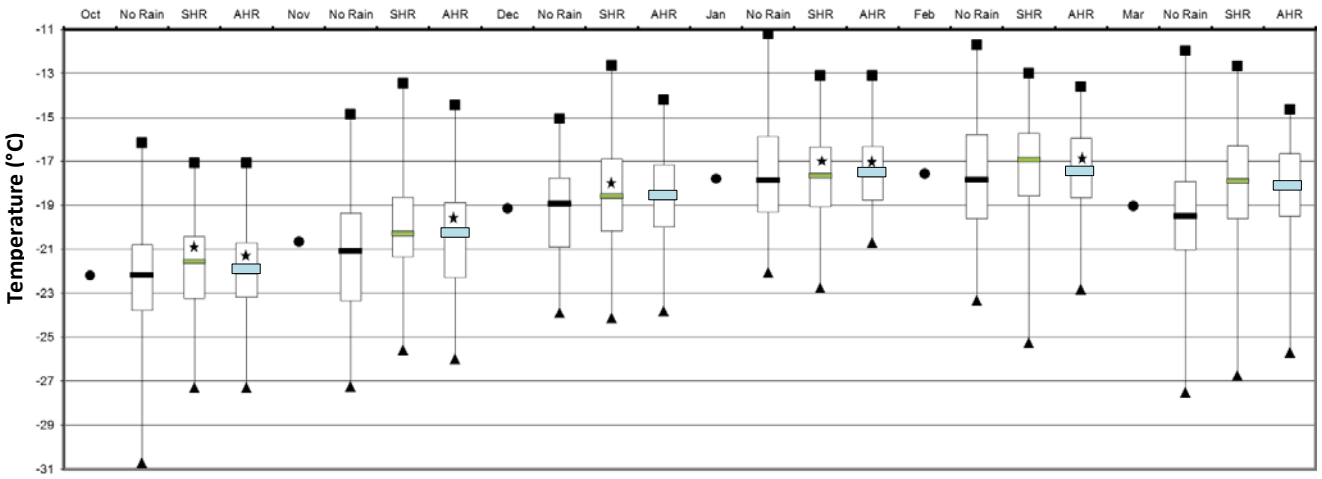


Figure 4-28: Monthly box and whisker diagram of average 500-300 hPa temperatures (°C) at Irene at 1200 UT. The black bars are for soundings with No Rain, the green bars for single station heavy rainfall (SHR) and the blue bars for area-average heavy rainfall (AHR). The circles are the CLIM values at 1200 UT. The bars indicate the mean value, the rectangles the maximum and the triangle minimum values. The top of the clear rectangles are the third quartile and the bottom the first quartile. The star inside the open rectangle indicates that the Mann-Whitney test did not find the mean between the dry and heavy rainfall sounding to be significant with at most a 5 % significance level.

4.7.3 Temperature lapse rates and related variables

In section 4.6.1 it was shown that temperature lapse rate (TL) in the atmosphere decreases as the summer progresses and the conditional instability of the atmosphere decreases. The question remains how the atmosphere behaves on days when heavy rainfall occurs. Are the TL rates on these days larger than normal so that the heavy rainfall develops in an atmosphere with increased conditional instability on that particular day? TLs and temperature differences (TDs) are investigated between several pressure levels and heavy rainfall soundings compared to soundings with no rainfall. These include the TD between the surface and 700 hPa, between the surface and 600 hPa, the surface and 500 hPa, between 700-400 hPa and between 600-300 hPa. As an example, Fig. 4-29 depicts the average monthly values of TD between the surface and 500 hPa for No Rain soundings as well as soundings with AHR and SHR. Only during December and January months are the TD from the surface to 500 hPa for heavy rainfall soundings higher than for dry soundings. However, these differences are not found to be significant to the 5 % significant level by the Mann-Whitney test. In all the other months the TL were lower for heavy rainfall soundings than for dry soundings but Mann-Whitney found that these differences were not significant with the exception of the AHR class in October and November months. The TD investigated between all the other pressure levels showed the same tendency, generally having lower lapse rates for heavy rainfall soundings than for No Rain soundings and with little significance in these differences. The one exception being the TD between the 700-500 hPa pressure levels. The TD between these two pressure levels is consistently higher for both AHR and SHR classes (Fig. 4-30). These differences are significant to at least the 5 % significant level, the exceptions being January and February months. When investigating the upper air conditions associated with heavy rainfall in Utah, Harnack et al. (1998) also found the TD between these two levels to be significant in distinguishing between heavy precipitation and the climatological mean.

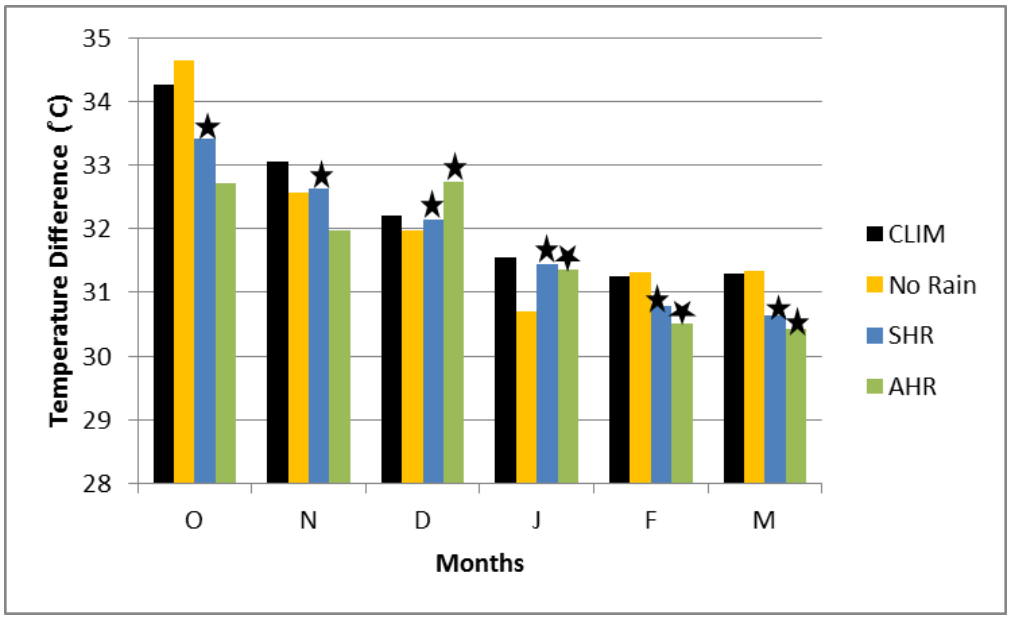


Figure 4-29: The monthly average surface to 500 hPa temperature difference (TD) at Irene 1200 UT. CLIM values are black, the average for No Rain soundings are shown in yellow, SHR and AHR soundings respectively, in blue and green. The star indicates that the Mann-Whitney test did not find the mean between the No Rain and heavy rainfall sounding to be significant with at least a 5 % significance level.

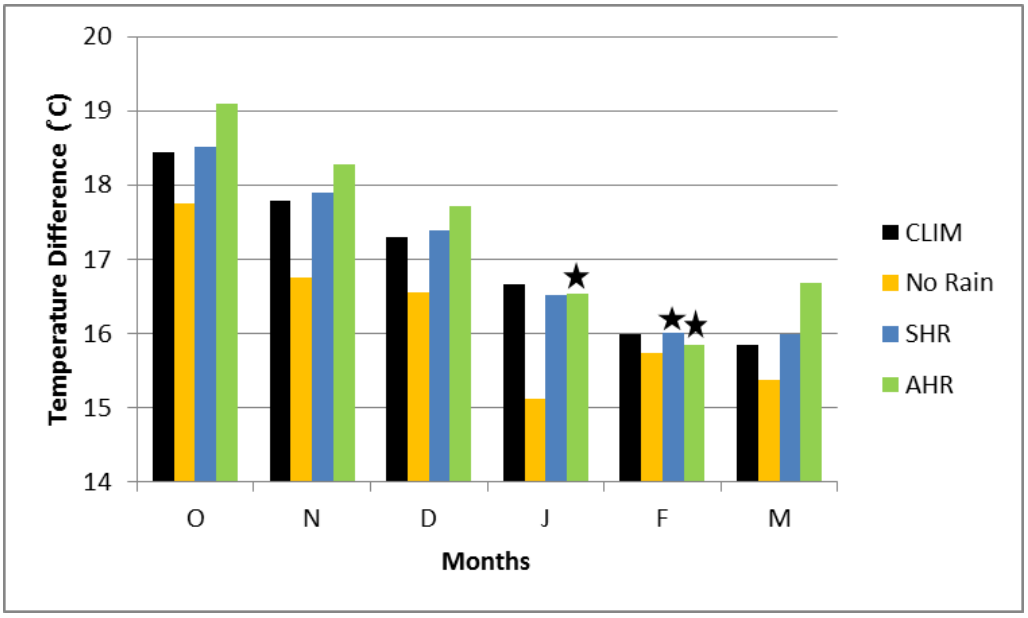


Figure 4-30: The monthly average 700 to 500 hPa temperature difference (TD) at Irene at 1200 UT. CLIM values are black, the average for No Rain soundings are shown in yellow, SHR and AHR soundings respectively, in blue and green. The star indicates that the Mann-Whitney test did not find the mean between the No Rain and heavy rainfall sounding to be significant with at least a 5 % significance level.

It is useful to investigate how these characteristics of the temperature lapse rate impact on the KI and TTI, because both use TD in its calculations.

Long-term average TTI values are above 50 °C for all months reaching a maximum of 53 °C in December (Table 4-6). The TTI for AHR and SHR soundings are only slightly higher than the CLIM value. In January months there is only a 1 °C difference between the CLIM and SHR soundings.

Fig. 4-27 shows that for SHR soundings the 500 hPa temperatures are slightly lower than for No Rain soundings but with values very close to the CLIM values. Moreover, there is no significant difference in the surface to 500 hPa TD between No Rain soundings and AHR and SHR soundings. Furthermore, TDs for AHR soundings are often less than for soundings which remain dry. The higher TTI values for heavy rainfall soundings are therefore largely influenced by the surface dew point temperatures (Fig. 4-23) and especially in late summer. Months with the largest differences in surface dew point temperature between AHR and dry soundings are also the months with the largest differences in TTI.

The ETTI partially relies on the TD between 700-500 hPa and, as was expanded on in Fig. 4-30, these TDs are larger for heavy rainfall soundings than for dry soundings. These differences are very small in value and will have only a slight influence on the value of the ETTI. In this case the higher values for ETTI are largely a function of the increase in 700 hPa dew point temperatures.

Table 4-6: The monthly CLIM values at 1200 UT values for TTI, ETT, KI an EKI, as well as monthly values for AHR (*) soundings and SHR (**) soundings.

Month	TTI	ETTI	KI	EKI
October	51 55* 54**	25 29* 27**	29 36* 34**	-1 9* 5**
November	52 55* 54**	25 28* 27**	34 38* 37**	2 10* 8**
December	53 55* 54**	25 28* 26**	36 40* 38**	4 11* 8**
January	52 55* 53**	24 27* 26**	37 41* 40**	5 12* 9**
February	51 52* 52**	24 26* 26**	37 41* 40**	4 10* 10**
March	50 52* 51**	23 27* 25**	35 38* 37**	1 12* 8**

The KI also uses the surface to 500 hPa TD and suffers from the same limitation as the TTI. In this instance the higher surface dew point as well as the low 700 hPa dew point depression causes the higher than normal KI for AHR soundings in late summer. The EKI has the largest differences between heavy rainfall and dry soundings. Even though the 700-500 hPa TD preforms better to distinguish heavy rainfall and dry soundings, the largest impact on the increased values for heavy

rainfall soundings is due to the higher 700 hPa dew point temperatures as well the low dew point depression at 600 hPa. The monthly EKI values for AHR are slightly higher than for SHR (Table 4-6). Box and whisker diagrams for the EKI reveals that the mean values for No Rain soundings are less than zero in all months and the third quartile for soundings without rain is equal to, or lower than, the first quartile for heavy rainfall soundings (Fig. 4-31).

Considering these 4 indices it therefore is recommended that the EKI is used to identify soundings associated with heavy rainfall at Irene and that values of 10 °C or more are considered an important indicator for wide spread heavy rainfall. However, as with the KI this index also does not indicate the severity of convection. It also considers only three levels in the atmosphere and does not account for dry or stable layers which occurs below or in-between these levels.

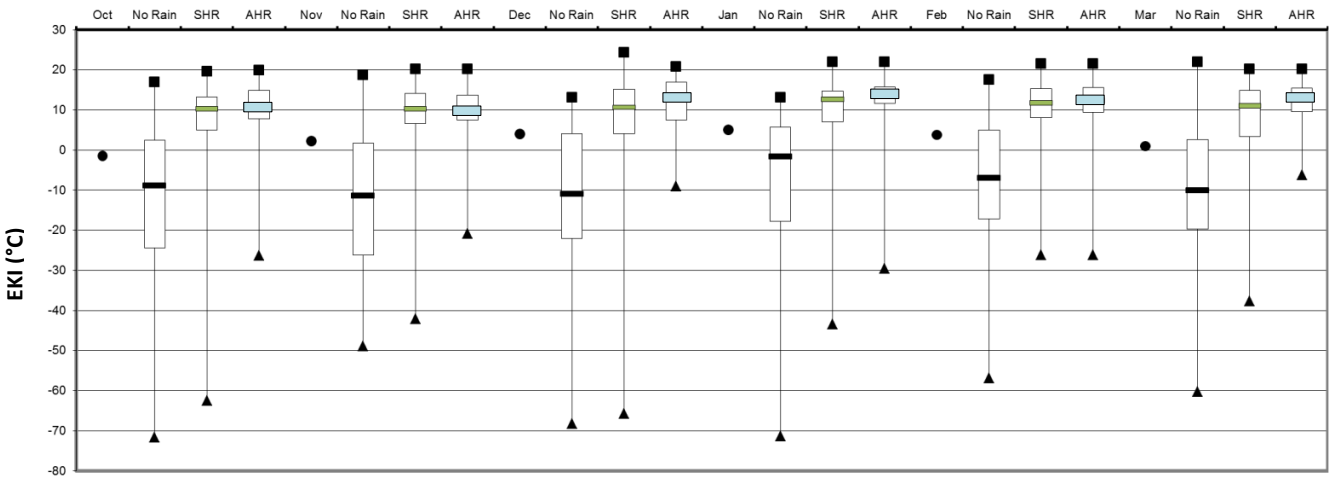


Figure 4-31: Box and whisker diagram of monthly average EKI (°C) at Irene at 1200 UT. The black bars are for soundings with No Rain, the green bars for single station heavy rainfall (SHR) and the blue bars for area-average heavy rainfall (AHR). The circles are the CLIM values at 1200 UT. The bars indicate the mean value, the rectangles the maximum and the triangle minimum values. The top of the clear rectangles are the third quartile and the bottom the first quartile.

4.7.4 Wind speed and direction

The wind direction for AHR soundings generally veers in respect to the wind direction for No Rain soundings (Fig. 4-32). Differences between the wind direction for No Rain soundings and AHR soundings increase as the season progresses. At 500 hPa the AHR wind direction is westerly throughout the season but the average wind strength decreases from 20 knots in October to 5 knots in February. At pressures lower than 500 hPa the wind directions for heavy rainfall soundings are northwesterly for the entire season. At pressures higher than 500 hPa, in early summer, the wind directions for AHR soundings are also northwesterly with an increasing northerly component close to the surface. During late summer at pressures higher than 600 hPa, the AHR winds become very light, but are almost northerly at the surface. In late summer between 600-700 hPa the wind direction for No Rain soundings is southeasterly. The southeasterly winds are caused by the stronger than normal high pressure system southeast of the country extending a ridge into the southeastern interior while a deep trough is situated over the Mozambique Channel (Figs 3-5 and 3-6).

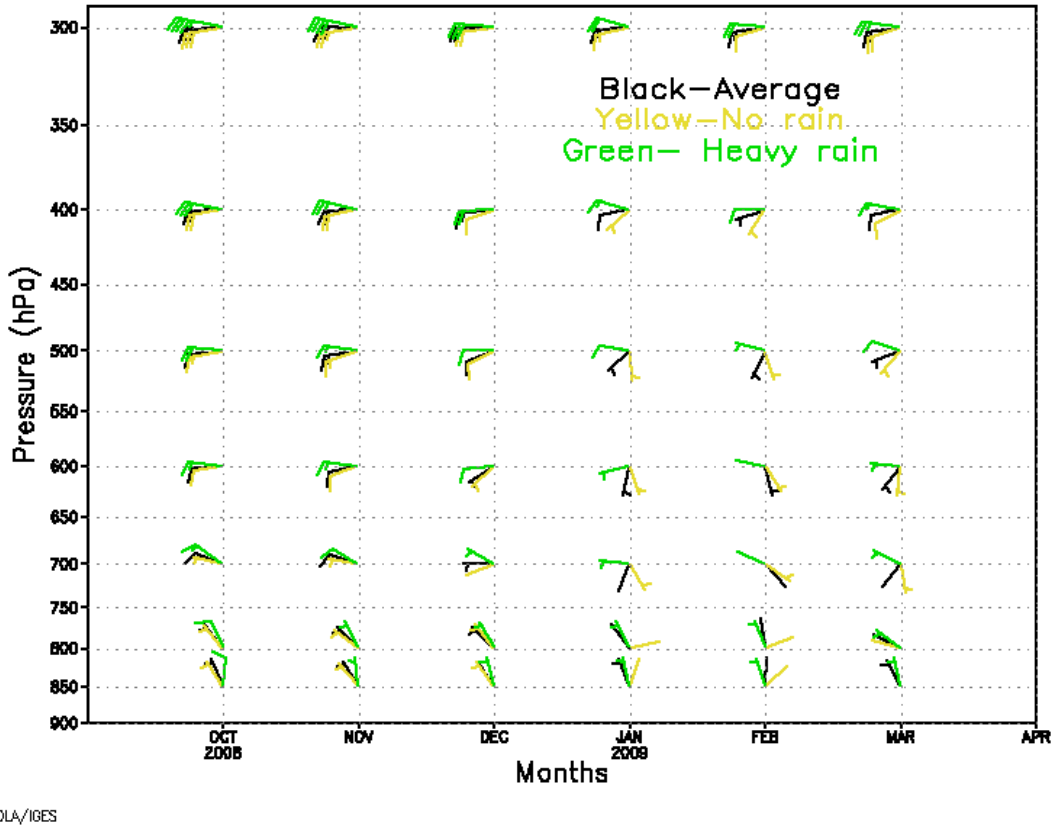


Figure 4-32: Vertical profile of the monthly average wind direction and speed (knots) for October to March at Irene at 1200 UT Black barbs represent the CLIM winds, yellow No Rain soundings and green area-average heavy rainfall (AHR) soundings.

At pressures lower than 500 hPa the wind strengths are stronger for heavy rainfall soundings than the long-term average value (Fig. 4-33). In October months at 300-400 hPa the wind strength for AHR soundings are 10 knots higher. In late summer the winds remain stronger than the average wind strength at these levels, but the higher winds speeds are now only about 5 knots above normal. In all the months except January and February months the winds are stronger than the norm through the entire troposphere for AHR soundings but by only 1 or 2 knots at the surface. In January and February, the wind speeds decrease by 1-2 knots during heavy rainfall events from the surface to 500 hPa. For SHR rainfall events, the winds strengths are very similar to the AHR but the high above normal anomalies in the upper troposphere in October and November is absent.

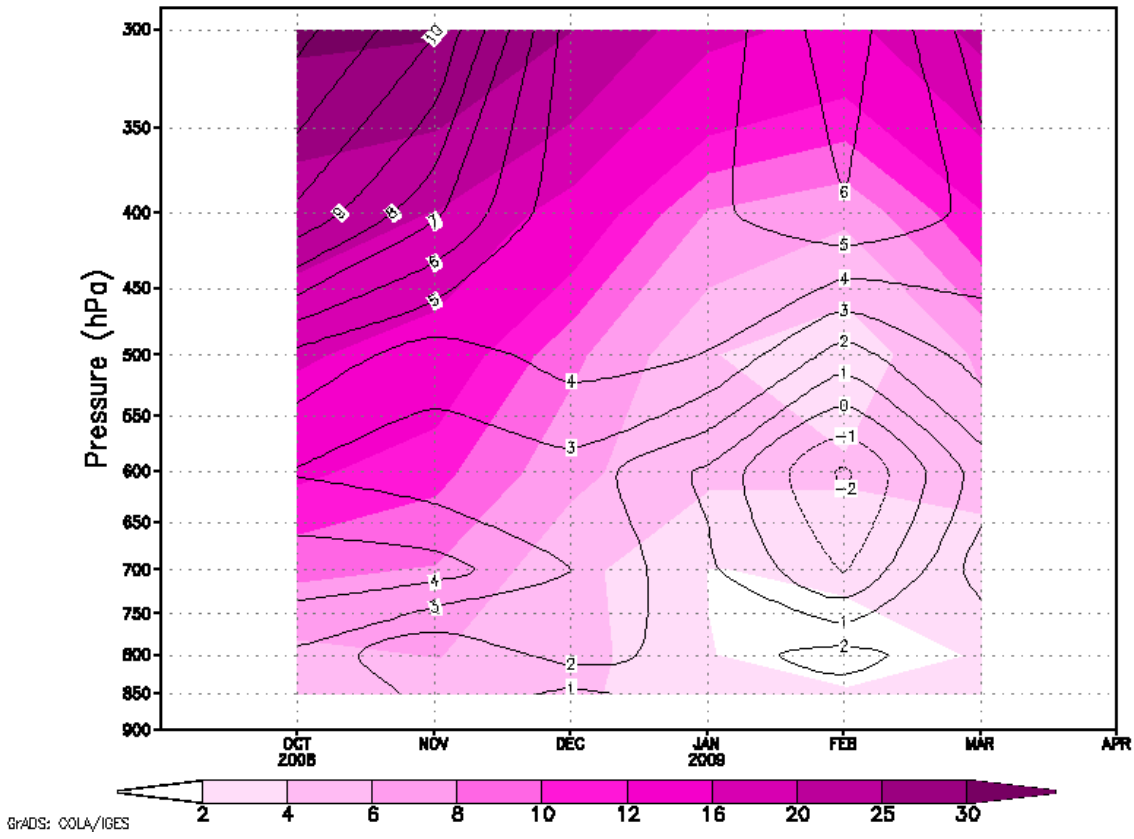


Figure 4-33: Vertical profile of the monthly average wind speed (knots) for October to March at Irene at 1200 UT. The contours are the anomalies for area-average heavy rainfall (AHR).

The average 600-400 hPa winds (W_{64}) were calculated to attempt to isolate months when slow moving convective systems stimulate the production of heavy rainfall. Fig. 4-34 shows that the only months when the mean 600-400 hPa winds are stronger for AHR than for No Rain soundings are October months. For all other months the mean wind strengths for heavy rainfall soundings are equal to or less than the dry soundings. The Mann-Whitney test found that these differences are not significant with the exception of the AHR soundings in February months. There is thus no clear indication that heavy rainfall over Gauteng forms when the wind strength is less than the CLIM value for that particular month. Because of the general decrease in wind speed in late summer, heavy rainfall which develops then is associated with much weaker winds than heavy rainfall which develop in early summer. This is emphasised in Fig. 4-35 which are scatter plots where the single station rainfall are plotted with the average 600-400 hPa wind speed for early (top) and late (bottom) summer. Comparing these images it is clear that the wind speeds in early summer are stronger than in late summer, therefore SHR are also associated with stronger winds in early than in late summer. In early summer the strongest wind speed associated with heavy rainfall was 30 ms^{-1} and in late summer it was 20 ms^{-1}. In early summer SHR is associated with a wider range of wind speeds than late summer. In early summer 61 % of SHR wind speeds are less than 10 ms^{-1} and 85 % in late summer.

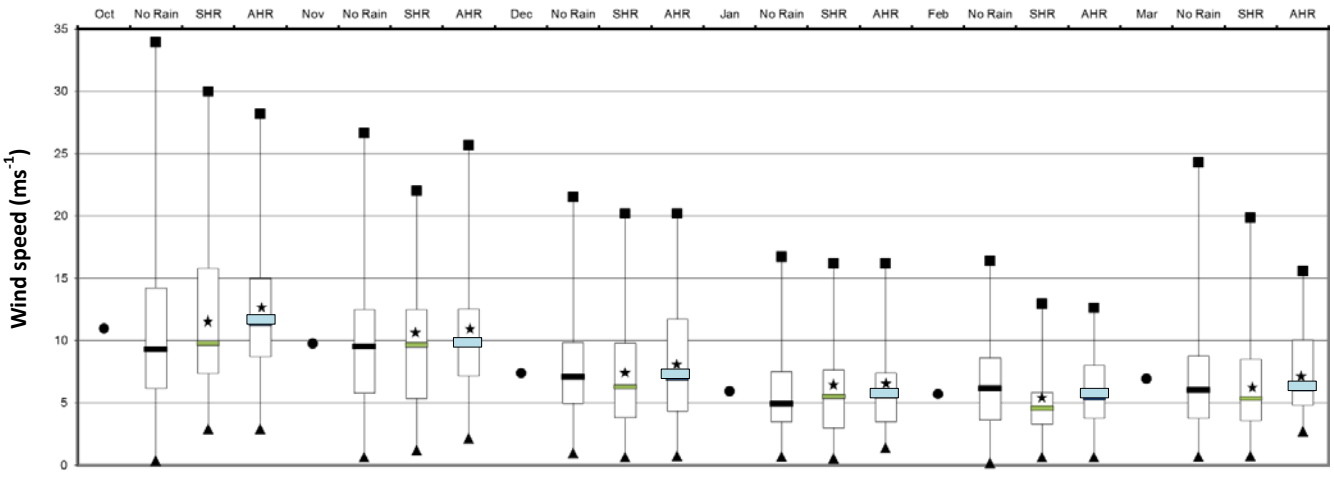


Figure 4-34: Box and whisker diagram of monthly average 600-400 hPa wind speed (W64; ms^{-1}) at Irene at 1200 UT. The black bars are for soundings with No Rain, the green bars for single station heavy rainfall (SHR) and the blue bars for area-average heavy rainfall (AHR). The circles are the CLIM values at 1200 UT. The bars indicate the mean value, the rectangles the maximum and the triangle minimum values. The top of the clear rectangles are the third quartile and the bottom the first quartile. The star inside the open rectangle indicates that the Mann-Whitney test did not find the mean between the dry and heavy rainfall sounding to be significant with at most a 5 % significance level.

The average meridionial wind component in the 800-600 hPa (WV_{86}) pressure level shows that heavy rainfall soundings for both classes and for all the months are associated with northerly winds (Fig. 4-36). These differences are all significant with 1° of freedom according to Mann-Whitney test with the exception of the SHR in October months where the degree of freedom is 10%. In late summer the average wind direction in this level is southerly but it switches to light northerly winds when heavy rainfall occurs. The importance of these northerly winds is that they bring surface moisture into the province. The scatter plot of the average meridionial wind in the 800-600 hPa layer and average daily rainfall (Fig. 4-37) shows that most SHR events are associated with northerly winds (negative values).

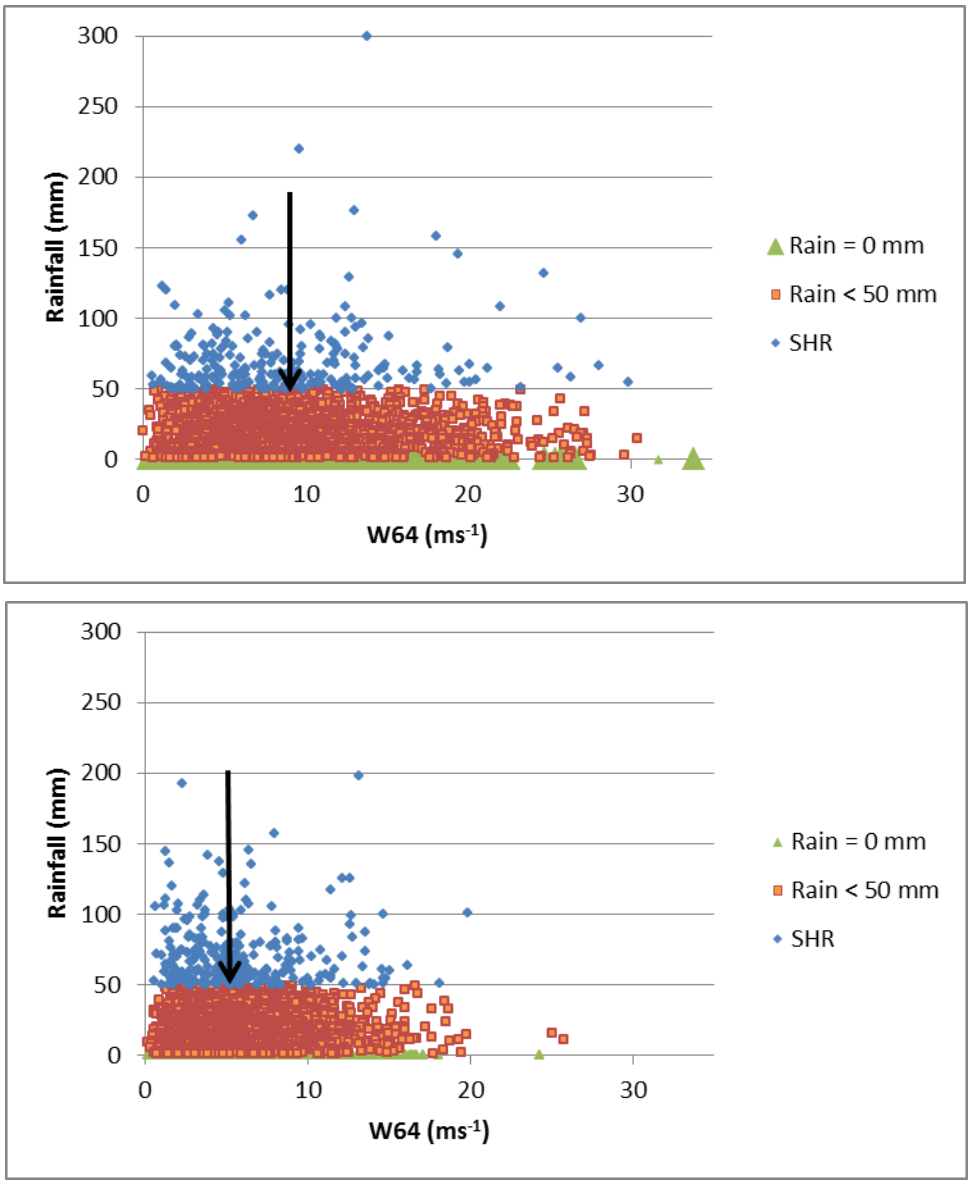


Figure 4-35: Scatter plots of single station daily rainfall and average 600-400 hPa wind speed (W64) for early summer (top) and late summer (bottom) at Irene at 1200 UT. The arrows indicate the average W64 value for SHR.

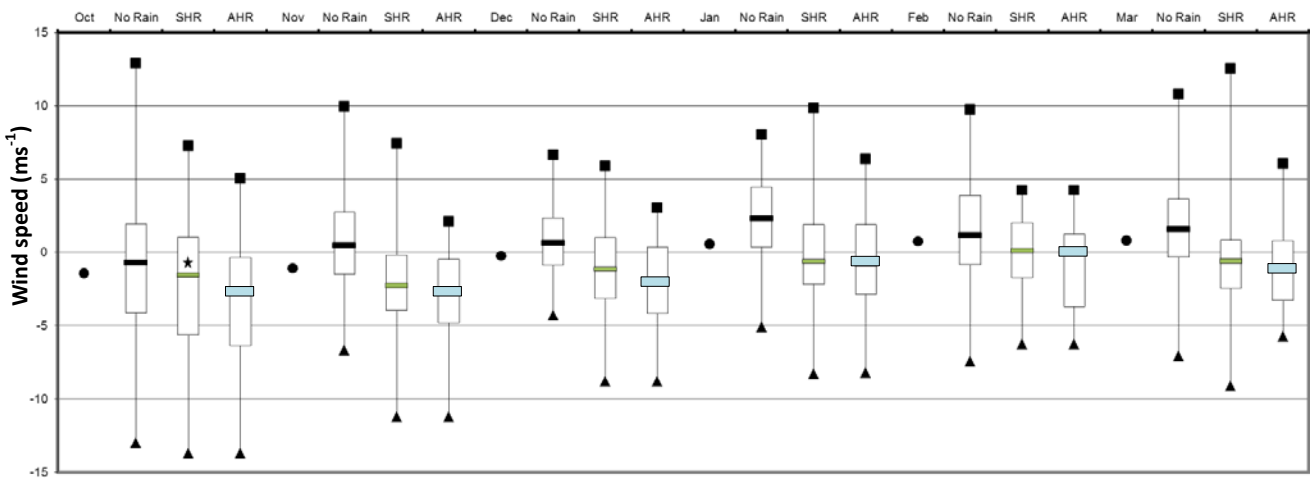


Figure 4-36: Box and whisker diagram of monthly average meridionial wind component in the 800-600 hPa (WV_{86}) pressure level at Irene at 1200 UT. The black bars are for soundings with No Rain, the green bars for single station heavy rainfall (SHR) and the blue bars for area-average heavy rainfall (AHR). The circles are the CLIM values at 1200 UT. The bars indicate the mean value, the rectangles the maximum and the triangle minimum values. The top of the clear rectangles are the third quartile and the bottom the first quartile. The star inside the open rectangle indicates that the Mann-Whitney test did not find the mean between the dry and heavy rainfall sounding to be significant with at most a 5 % significance level.

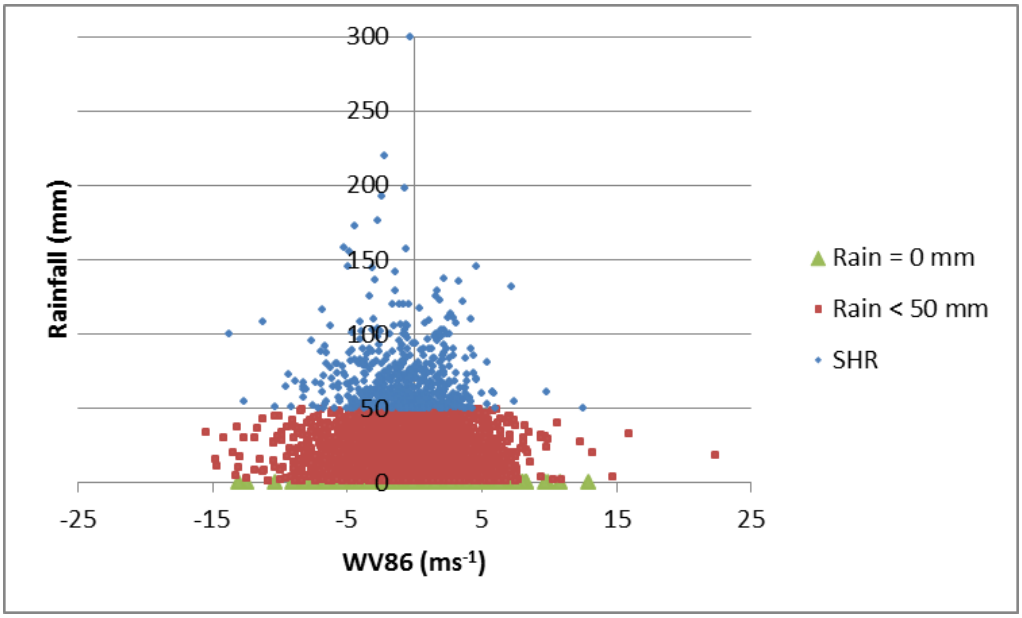


Figure 4-37: Scatter plot of single station daily rainfall and average meridionial wind component in the 800-600 hPa (WV_{86}) pressure level at Irene for the summer season at Irene at 1200 UT

4.7.5 Wind shear

Average bulk wind shear (WS_{s7}) values in the surface to 700 hPa layer show that on average there is very little association between wind shear values and heavy rainfall over Gauteng (Fig. 4-38). WS values associated with AHR are generally higher than the average monthly value as well as for dry soundings but these differences are only significant (according to the Mann-Whitney test) in October and November. The shear values between other pressure levels generally show the same tendency having significant higher values only in the first two months of summer. When heavy rainfall occurs in early summer, wind shear values are between 4-5 $ms^{-1}km^{-1}$. These values are similar to what Dupilka and Reuter (2006a) found to distinguish between weak and significant tornados in Canada. Heavy rainfall environments in early summer over Gauteng are therefore associated with high levels of conditional instability and strong wind shear. In Fig. 4-39 the temperature difference (TD) in the 700-500 hPa layer is plotted with the magnitude of the wind shear (MWS) from the surface to 400 hPa for some early and late summer months for SHR soundings. From this graph it can be seen how in October and November months SHR tend to occur with larger MWS and TD values while in January and February months the TD and MWS values are smaller.

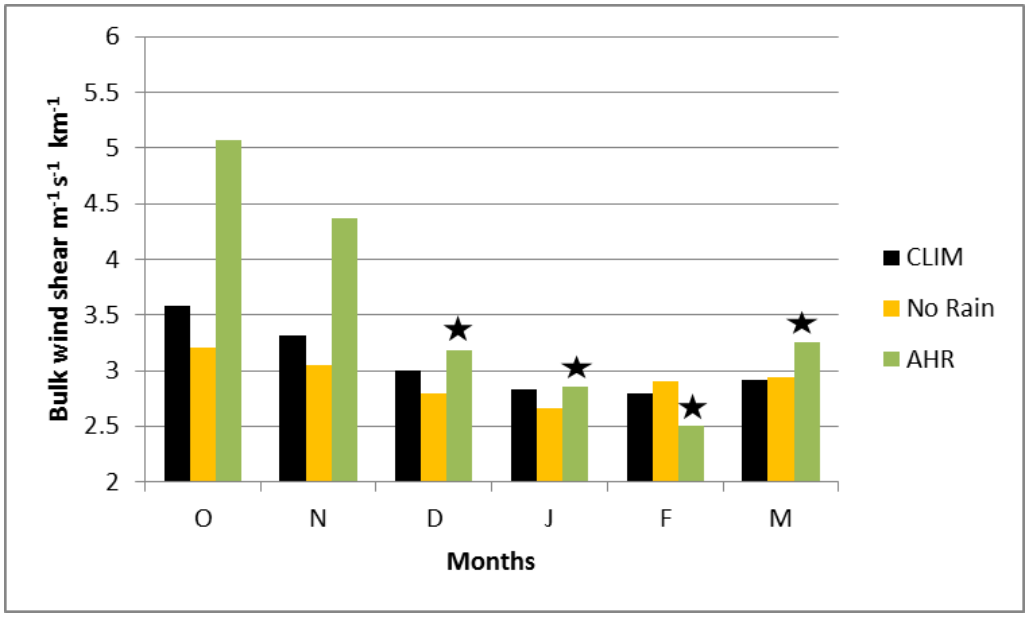


Figure 4-38: The monthly average surface to 700 hPa bulk wind shear (WS_{s7}) (m^{-1} s^{-1} km^{-1}) at Irene at 1200 UT. CLIM values are black, the average for No Rain soundings are shown in yellow, SHR and AHR soundings respectively, in blue and green. The star indicates that the Mann-Whitney test did not find the mean between the No Rain and heavy rainfall sounding to be significant with at least a 5 % significance level.

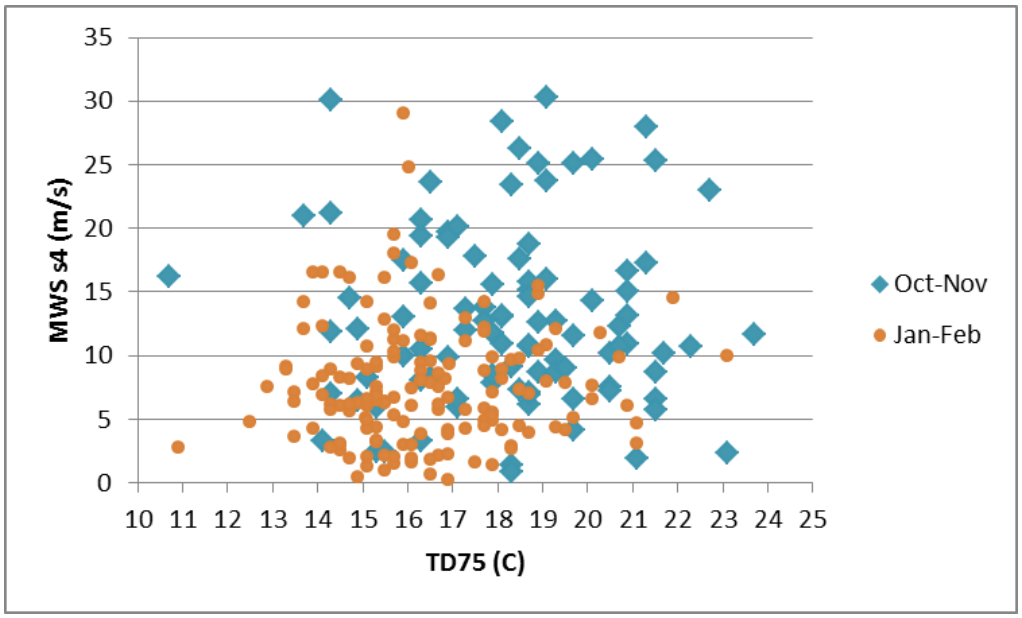


Figure 4-39: Scatter plot of the 700-500 hPa temperature difference (TD_{75}) vs the magnitude of wind shear (MWS_{s4}) for SHR soundings for October and November months (blue) and January and February months (orange) at Irene at 1200 UT

4.7.6 Average tropospheric Θ_e and $\Delta\Theta_e$

The monthly $\Theta_{e\ ave}$ values for SHR and AHR are consistently higher than the CLIM and No Rain sounding values (Fig. 4-40). The Mann-Whitney test found that these differences are significant with 1° of freedom for all the months. The $\Theta_{e\ ave}$ values for SHR soundings are generally 1-1.5 K higher than the CLIM values and AHR sounding have 2-3 K higher values.

$\Delta\Theta_e$ values were calculated between several pressure levels and heavy rainfall soundings were compared to soundings without rain. These include the surface-700 hPa $\Delta\Theta_e$, surface-600 hPa $\Delta\Theta_e$, surface to 500 hPa $\Delta\Theta_e$, surface to 400 hPa $\Delta\Theta_e$, surface to 300 hPa $\Delta\Theta_e$, 700-400 hPa, 700-500 hPa as well as $\Delta\Theta_e$ between the maximum and minimum Θ_e values. $\Delta\Theta_e$ values distinguish well between heavy rainfall and No Rain soundings. The only $\Delta\Theta_e$ values which performed poorly in distinguishing between heavy rainfall and dry soundings were the $\Delta\Theta_e$ between the maximum and minimum values. Fig. 4-41 shows that for $\Delta\Theta_e$ between the surface and 500 hPa both AHR and SHR soundings the values are consistently larger negative than soundings without rain. Except for October months the mean value for AHR and SHR soundings are less than -20 K indicative of the potential instability of the atmosphere.

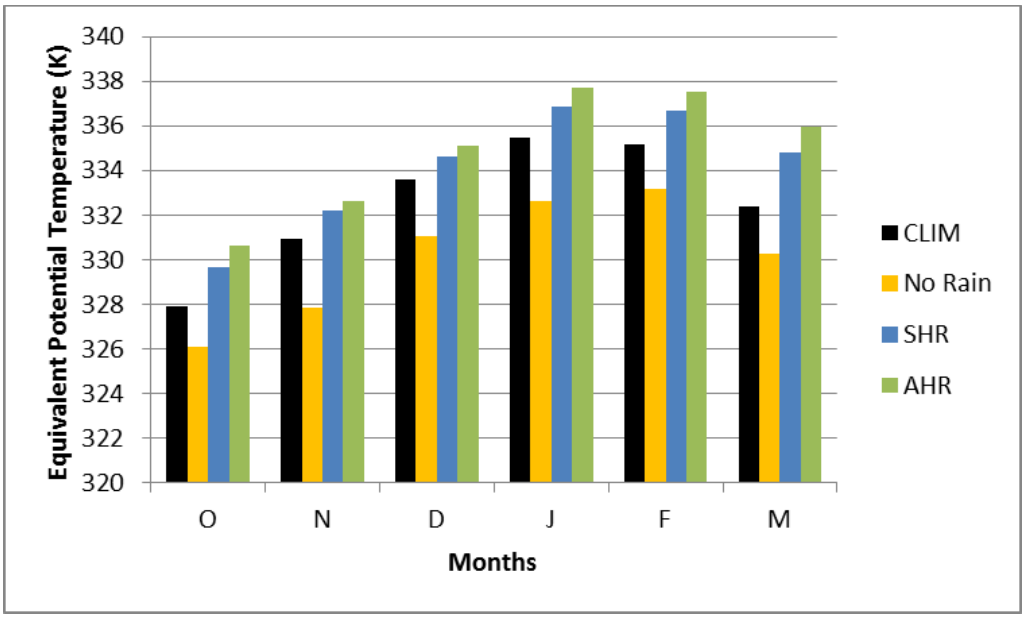


Figure 4-40: The monthly average tropospheric equivalent potential temperature ($\Theta_{e\ ave}$) (K) at Irene at 1200 UT. CLIM values are black, the average for No Rain soundings are shown in yellow, SHR and AHR soundings respectively, in blue and green.

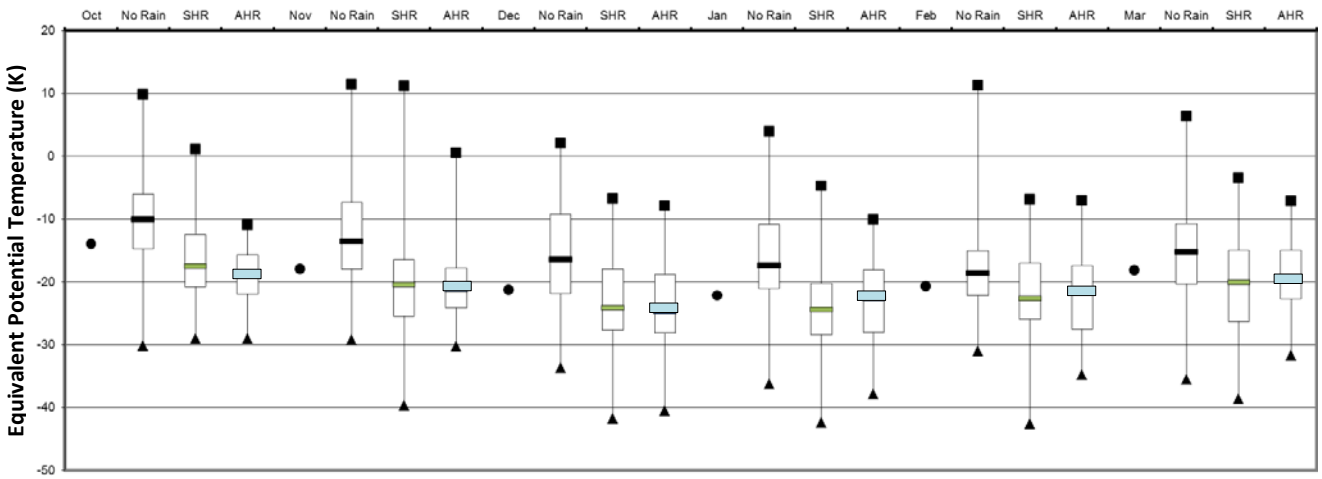


Figure 4-41: Box and whisker diagram of monthly average surface to 500 hPa equivalent potential temperature lapse rate ($\Delta\Theta_e$; K) at Irene at 1200 UT. The black bars are for soundings with No Rain, the green bars for single station heavy rainfall (SHR) and the blue bars for area-average heavy rainfall (AHR). The circles are the CLIM values at 1200 UT. The bars indicate the mean value, the rectangles the maximum and the triangle minimum values. The top of the clear rectangles are the third quartile and the bottom the first quartile.

4.7.7 Convective variables

For both heavy rainfall classes the CAPE values are significantly higher than the long-term average and No Rain soundings (Fig. 4-42). The Mann-Whitney test found all the differences to be significant with 1 % confidence level for both classes and all the months. The values for AHR are generally higher than for SHR with values higher than 1100 Jkg $^{-1}$ in November and December months.

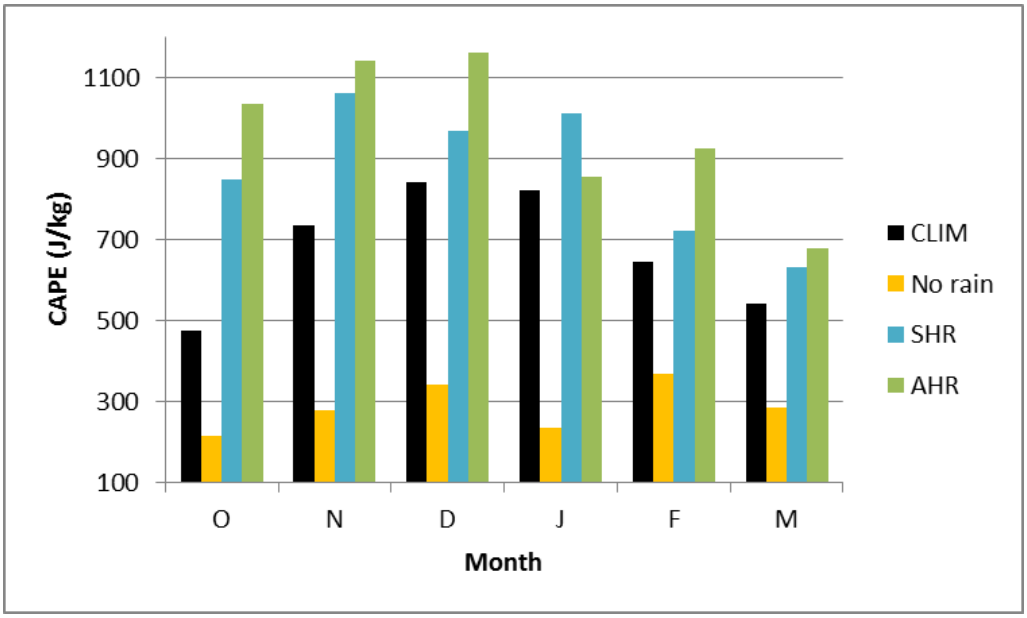


Figure 4-42: The monthly average CAPE values at Irene at 1200 UT. CLIM values are black, the average for No Rain soundings are shown in yellow, SHR and AHR soundings respectively, in blue and green.

The largest negative values in the Showalter Index (SI) occur in early summer when AHR soundings have average values of less than -3. As was the case with CAPE the LI values for AHR are larger than for SHR (Fig. 4-43)

In Fig. 4-44 it can be seen how in when CAPE values > 1000 Jkg^{-1} occur the TD_{57} values are generally larger in early summer than in late summer. Conversely in Fig. 4-45, the larger negative values of $\Delta\Theta e_{s5}$ occurs in late summer when Cape values are larger than 1000 Jkg^{-1}. This is indicative of the atmosphere which changes from one where conditional instability dominates to one where convective instability plays an increasingly important role.

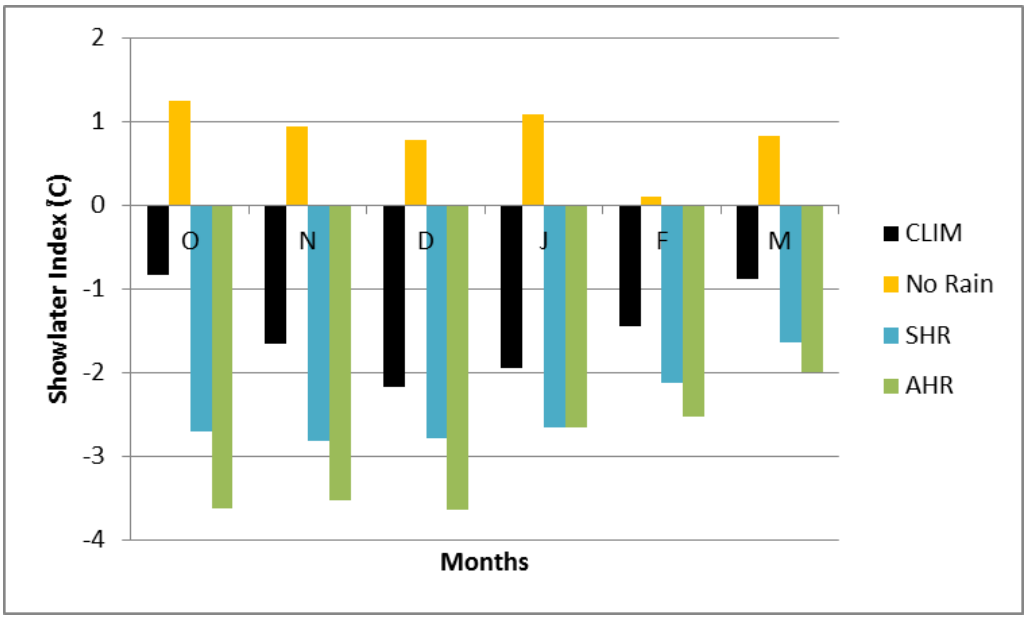


Figure 4-43: The monthly average Showalter Index (°C) values at Irene at 1200 UT. CLIM values are black, the average for No Rain soundings are shown in yellow, SHR and AHR soundings respectively, in blue and green.

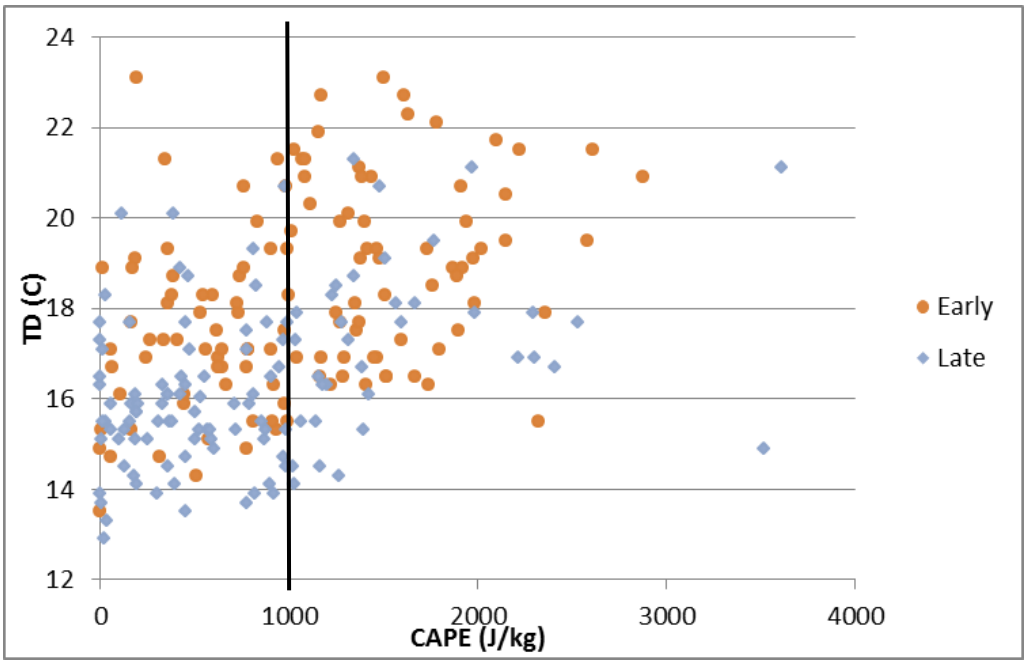


Figure 4-44: Scatter plot of CAPE (J kg^{-1}) vs 700-500 hPa temperature difference TD_{75} for area-average heavy rainfall (AHR) soundings for Early (orange) and Late summer (blue) at Irene 1200 UT.

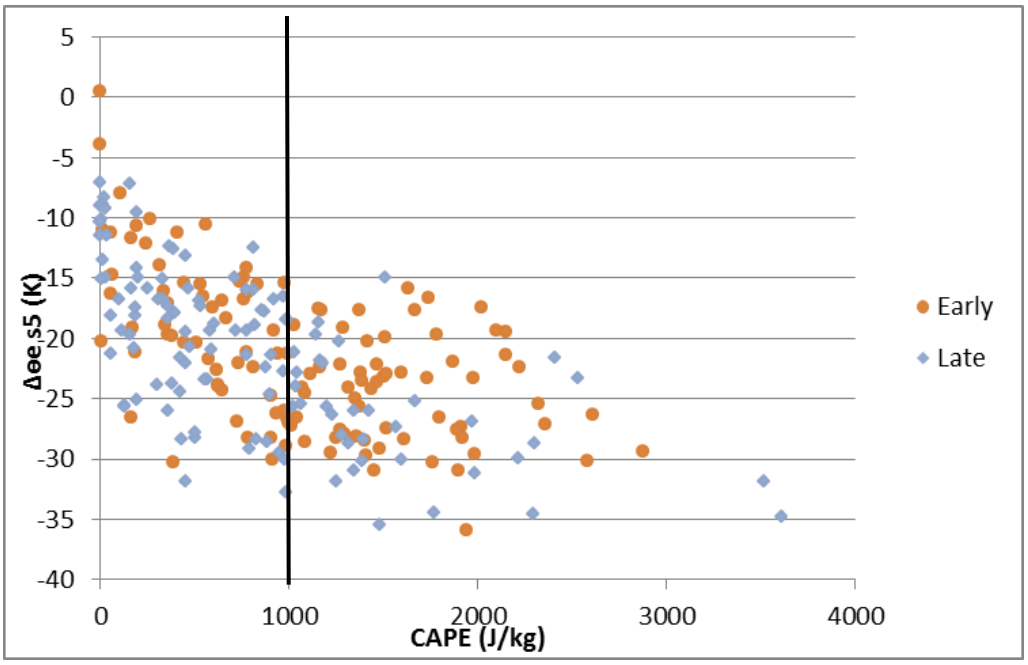


Figure 4-45: Scatter plot of 1200 UT CAPE (J kg^{-1}) vs surface to 500 hPa equivalent potential temperature difference ΔΘe for area-average heavy rainfall (AHR) soundings for Early (orange) and Late summer (blue) at Irene.

During early summer the WCD for AHR and SHR rainfall soundings are significantly larger (according to the Mann-Whitney test) than the CLIM and No Rain values (Fig. 4-46). However, during late summer the WCD is very similar to the CLIM values. The WCD values for heavy rainfall reach a maximum in February months when WCD is nearly 160 hPa. Comparing the WCD to MWS_{s4} in early and late summer in Fig. 4-47 it can be noted that in late summer, when the MWS values are smaller than in early summer the WCD tends to be higher. In early summer when the MWS is > 15 ms^{-1} the WCD is 150 hPa and less. It is thus in late summer when the slow moving precipitation efficient weather systems cause heavy rainfall over Gauteng.

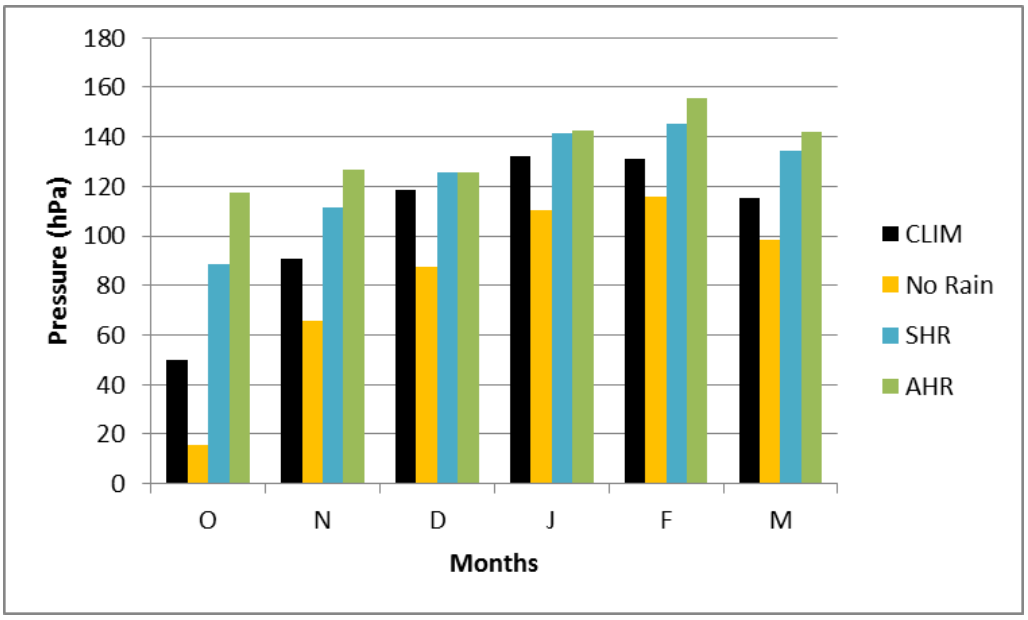


Figure 4-46: The monthly average 1200 UT Warm Cloud Depth (WCD) (hPa) values at Irene. CLIM values are black, the average for No Rain soundings are shown in yellow, SHR and AHR soundings respectively, in blue and green.

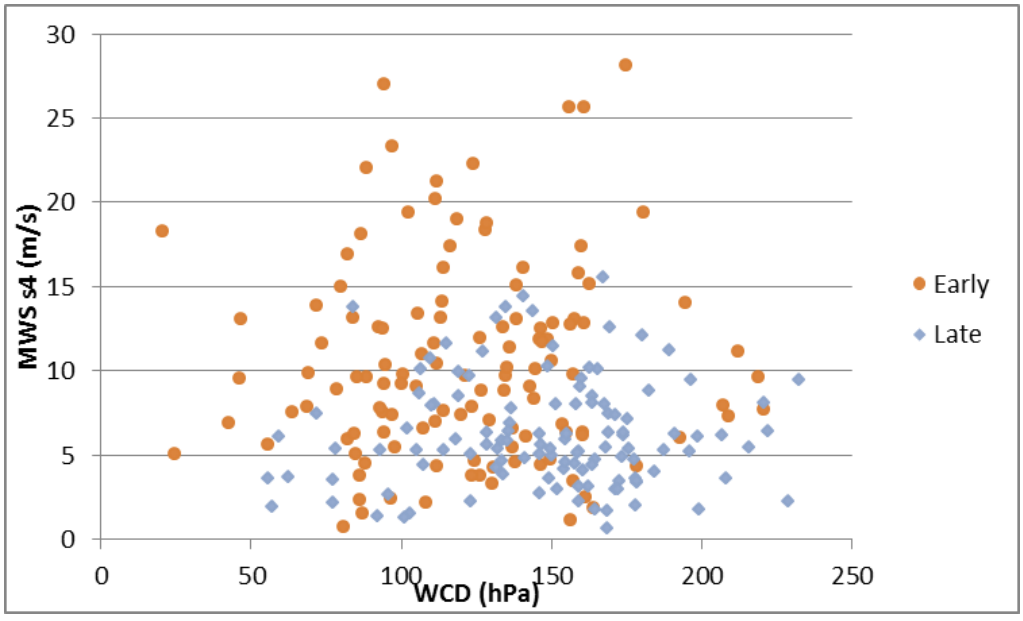


Figure 4-47: Scatter plot of 1200 UT Warm Cloud Depth (WCD) (hPa) vs magnitude of wind shear in the surface to 400 hPa layer MWS_{s4} (ms^{-1}) for area-average heavy rainfall (AHR) soundings for Early (orange) and Late summer (blue) at Irene.

4.8 A heavy rainfall sounding climatology using Self-organizing maps

4.8.1 Average heavy rainfall sounding climatology

Fig. 4-48 is a two dimensional representation of the inter-node distance across the data space of the SOM. The nodes are concentrated in the bottom right hand side of the SOM and are more widely spaced in the top left hand side. Fig. 4-48 is a graphical representation for AHR of temperature, dew point temperature and winds on skew-t gram plots of the 5 by 4-node SOM while Table 4-7 are the associated sounding-derived parameters for each of the 20 nodes.

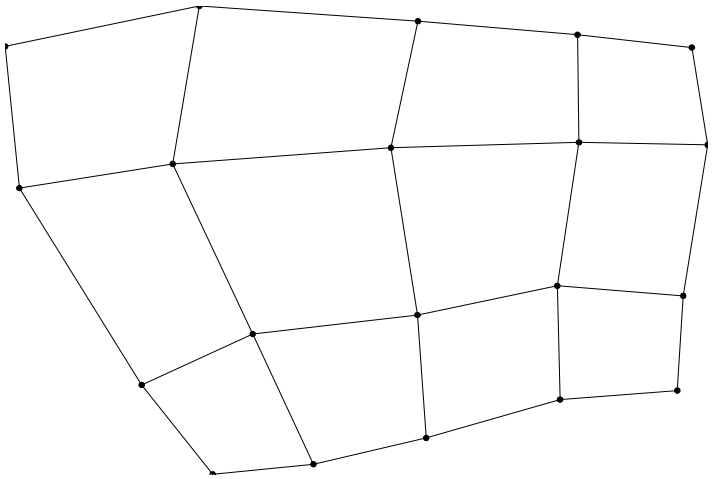


Figure 4-48: Two-dimensional portrayal of the distances between SOM nodes for average heavy rainfall (AHR) soundings from 1977-2012.

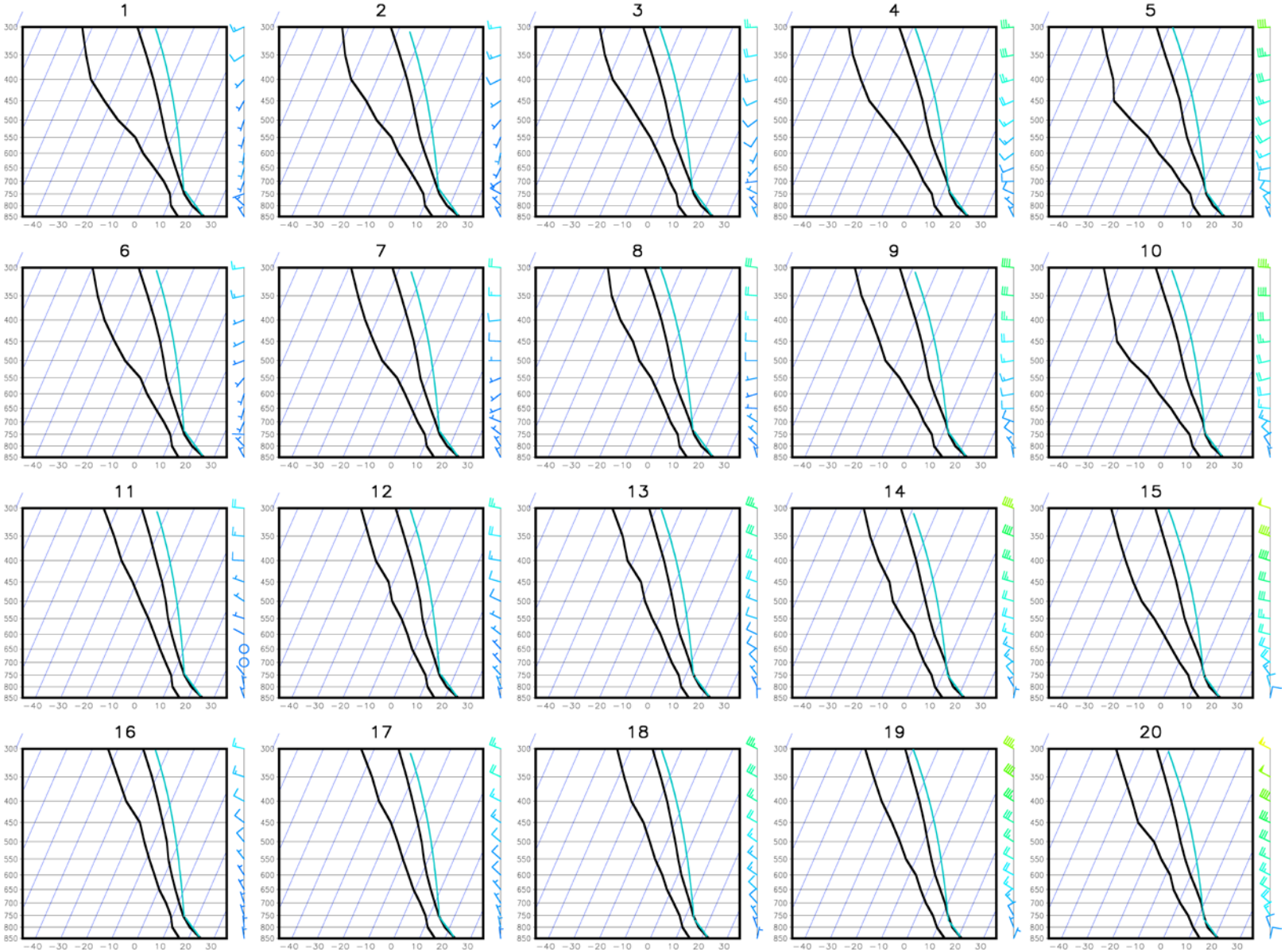


Figure 4-49: SOM of skew-t gram plots of the AHR soundings from 1977-2012. The solid blue lines show the track of a parcel if lifted from the surface to the lifted condensation level by the dry adiabatic lapse rate and then follows the saturated adiabatic lapse rate to 300 hPa.

Table 4-7: The SOM of the sounding derived parameters of AHR soundings from 1977-2012.

Node No	PW	Td100	T53	CAPE	TD75	WCD	EKI	WV86	W64	MWSs4	Δθe s5	θe ave
1	25	13	-18	1313	17	136	11	1	4	7	-26	337
2	23	12	-19	1306	18	122	10	1	5	7	-25	335
3	22	10	-20	1234	19	113	12	0	6	9	-22	333
4	21	10	-20	1024	19	116	9	0	9	12	-22	332
5	20	11	-20	1103	18	125	6	-1	11	15	-24	332
6	26	13	-17	1264	17	140	12	1	4	6	-25	338
7	25	12	-18	1332	17	134	13	-1	5	8	-24	336
8	23	11	-20	1158	18	116	12	-1	6	9	-21	334
9	21	10	-21	980	18	115	9	-2	10	13	-21	332
10	20	11	-21	1030	19	124	6	-2	12	16	-23	331
11	29	13	-17	924	16	153	14	-1	5	7	-21	340
12	27	12	-17	952	16	141	13	-2	6	8	-21	338
13	24	11	-19	853	17	135	12	-3	8	12	-19	335
14	22	10	-20	811	18	122	11	-4	12	16	-19	332
15	21	10	-21	939	19	119	8	-5	16	19	-20	331
16	29	13	-16	656	15	157	14	-2	5	7	-19	340
17	28	12	-17	626	16	156	13	-3	7	9	-18	339
18	26	12	-18	590	16	152	12	-4	10	12	-17	337
19	23	11	-19	643	17	143	11	-5	14	17	-17	334
20	22	11	-21	894	19	130	11	-6	17	20	-19	332

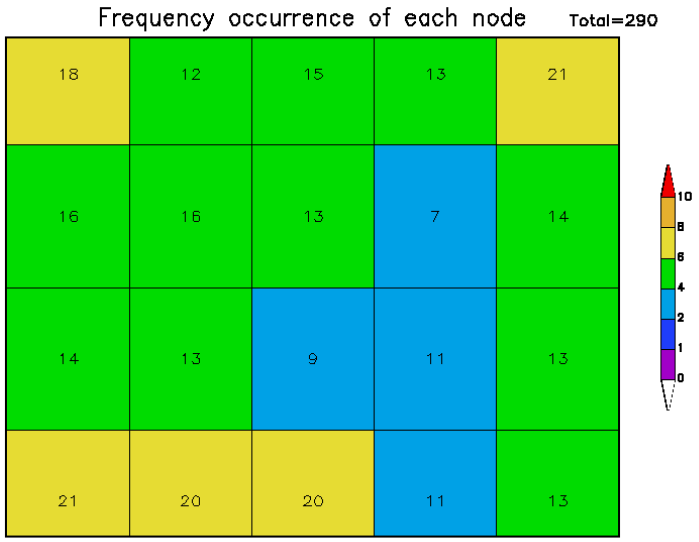


Figure 4-50: The numbers in each block are the total number of AHR soundings which was mapped to each node of the SOM in Fig-4-49, and the shades denote the frequency (%) of occurrence of each node.

The percentage of observations mapped to each node is shown in Fig.4-50. There were 290 AHR soundings mapped in this SOM. Percentage of occurrence of each node varies between 2 and 8 % with the highest frequencies in the top and bottom rows of the SOM. Node 5 and 16, on opposite sides of the SOM had the highest number of days mapped to it, 21. The nodes indicated by blue had the lowest frequencies as there were only between 7-11 days mapped to each of these nodes.

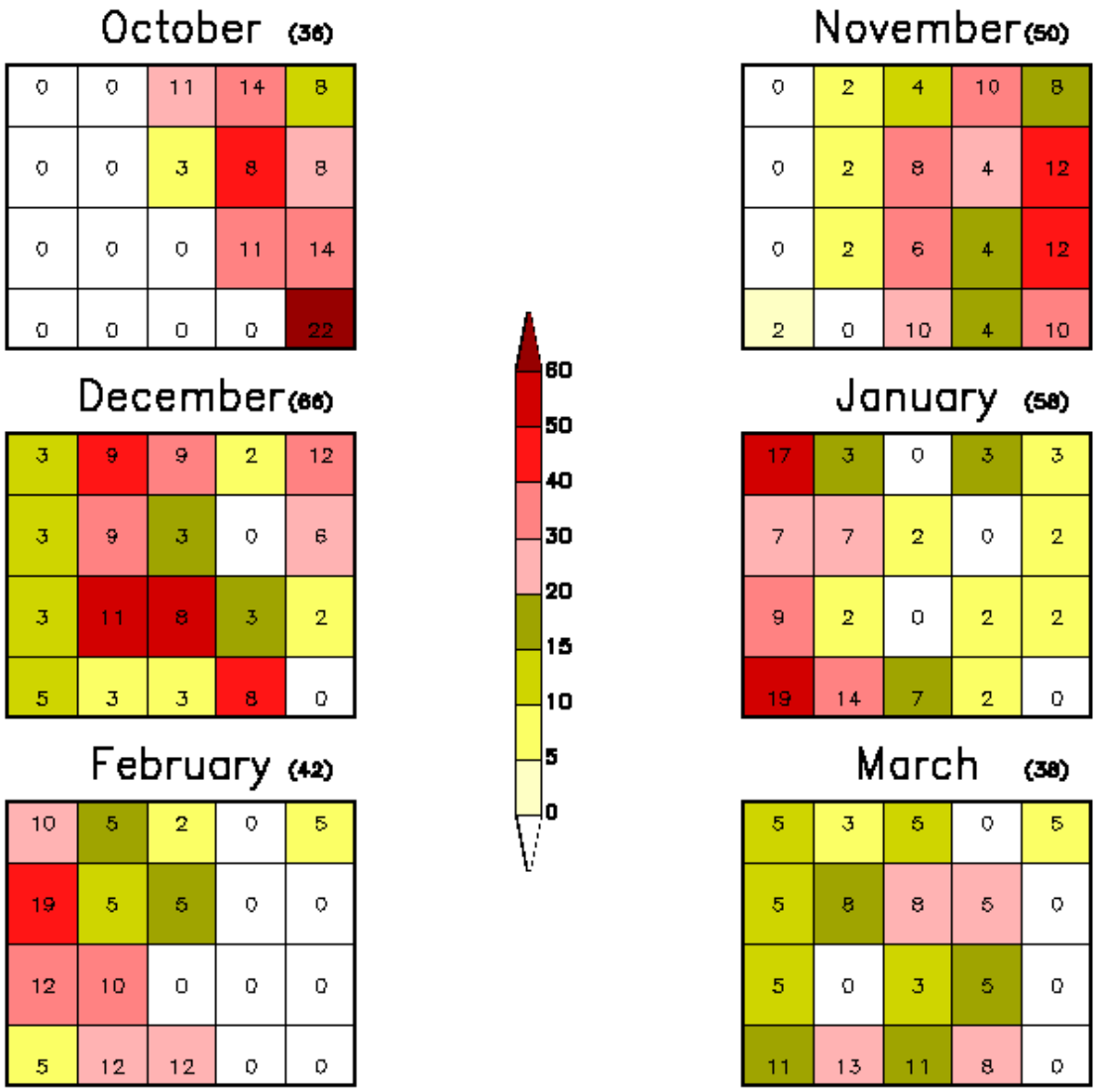


Figure 4-51: The number in each block is the frequency (%) of occurrence of each node relative to the total number of occurrences of AHR in that particular month. The shades are the frequency (%) of occurrence of each node expressed relative to the total number of occurrences of that node during the entire period.

In October the nodes occur exclusively on the right hand side of the SOM (Fig.4-51). Node 20 is the most frequent node when AHR occurs in October months (22 %) and this node occurs almost exclusively in October months. More than 60 % of the time that this node occurs it happens in an October month. The only other month in which this node occurs is November months (30-40 %). The nodes on the right hand side of SOM are associated with stronger winds throughout the troposphere than the nodes on the left hand side of the SOM (Fig. 4-49). The temperatures at the surface are also lower in the right hand nodes as is the upper and middle tropospheric temperatures. In node 20, 10 knot northeasterly winds occur at the surface and the wind backs with height to become northwesterly aloft. At 300 hPa the wind strength associated with node 20 is 55 knots, the strongest wind of all the nodes. Node 20 also have the strongest meridionial wind in the 800-600 hPa layer at -6 ms^{-1}, the strongest winds in the 600-400 hPa layer (17 ms^{-1}) and the largest wind shear values from the surface to 400 hPa at 20 ms^{-1} (Table 4-7). Of all the nodes which occur in October months, node 20 has the highest PW (22 mm). The nodes above node 20 in the SOM have similar characteristics to node 20 but with slightly weaker winds and the atmosphere become progressively drier towards the top row. In the October nodes in the top two rows, the winds still backs with height but with westerly and even southwesterly winds in the middle and upper troposphere. In Table 4-7 it can be noted that the PW in the October nodes are between 20-22 mm with mean layer dew point temperatures around 10 °C. These are the nodes with the coldest temperatures in the 500-300 hPa layer (± -20 °C). The drier October nodes of the top row has the largest CAPE values > 1000 Jkg^{-1}, while the October nodes in the bottom two rows have CAPE values of < 1000 Jkg^{-1}. The October nodes are the nodes with the largest TD_{75} of all the nodes (18-19 °C). $\Theta_{e\ ave}$ values are all below 334 K, the lowest values of all the nodes. The October nodes identified by the SOM are all indicative of a conditionally unstable atmosphere with large temperature lapse rates and strong wind shear

In November months the highest frequency of nodes still occur on the right hand side of the SOM but with an increasing number of nodes in the second and third column. These nodes are moister and have weaker winds than other nodes in the same row, but to the right hand side of the SOM. The PW values of these nodes increase to as high as 27 mm in node 12, with mean layer dew point temperatures closer to 12 °C. The temperatures in the 500-300 hPa layer are higher than the left hand side nodes.

In December and January months nearly all the nodes occur but there is a greater frequency of occurrence on the left hand side of the SOM. In December months nodes 12 and 13 occurs most frequently and in January nodes 1 and 16. In February months only nodes on the left hand side of the SOM occur, with the exception of node 5-which occurs in every month. The nodes in the first column

of the SOM all have surface temperatures of about 25 °C accompanied by relatively high mid-tropospheric temperatures (T_{53} varies between -18 to -16 °C). These nodes are the moistest nodes in the SOM and the highest PW value of 29 mm occurs in node 16 with mean layer dew point temperatures of 13 °C. The TD_{75} is between 15-17 °C, the lowest values in the SOM but with the largest values of $\Delta\Theta_e$, -26 K in node 1. $\Theta_{e\ ave}$ values are close to 340 K in these nodes. The winds in these nodes are relatively weak throughout the troposphere and only in node 11 do the upper tropospheric winds reach 20 knots. In node 16 the surface winds are light northerly with slight backing with height so that the winds in the upper troposphere are northwesterly. In the nodes above node 16, the winds are very weak (< 5 knots) up to 450 hPa. In nodes 1 and 6, the winds turn southwesterly in the upper troposphere. These two nodes are the only nodes where WV_{86} is southerly. W_{64} in these 4 nodes is only 4 or 5 ms^{-1} and MWS_{s4} 6 and 7 ms^{-1}. CAPE values of above 1200 Jkg^{-1} occur in nodes 1 and 6 but in node 16 the CAPE value is only 650 Jkg^{-1}.

Node 1 is the node with the highest CAPE value (1313 Jkg^{-1}). This coincides with relatively warm T_{53} temperatures (-18 °C) and low TD_{75} (17 °C). However, $\Delta\Theta_e$ has the largest negative value of -16 K. This is indicative of an atmosphere which is convectively unstable with decreased conditional instability.

Nodes 5 and 8 are the only 2 nodes which occur in every month. These nodes are very similar except that node 8 has more moisture and less wind shear than node 5. Node 8 has the highest frequency of occurrence in November and March months. This node is quite moist with PW values of 23 mm and mean layer dew point temperatures of 11 °C. The W_{64} is only 6 ms^{-1} and the MWS_{s4} 9 ms^{-1}. Node 5 which has the highest frequency of occurrence in December months, followed by October and November months, has W_{64} values nearly double of node 8 and with MWS_{s4} of 15 ms^{-1}.

4.8.2 Single station heavy rainfall climatology

The two dimensional representation of the inter-node distance across the data space of the SOM is shown in Fig 4-52. Figure 4-53 is the graphical representation of the 20 node SOM as skew-t gram plots and Table 4-8 depicts the sounding-derived parameters of the SOM. The nodes on the right hand side of the SOM are more closely spaced than on the left hand side. There were a total number of 565 days mapped in the SOM; this is about double the number of days mapped in the AHR SOM. There are many similarities between the two SOMs but the atmospheric conditions when 50 mm or more occur at a station over Gauteng ares not necessarily the same as when the average Gauteng rainfall exceeds 10 mm. In this discussion the differences between the two SOMs will be highlighted.

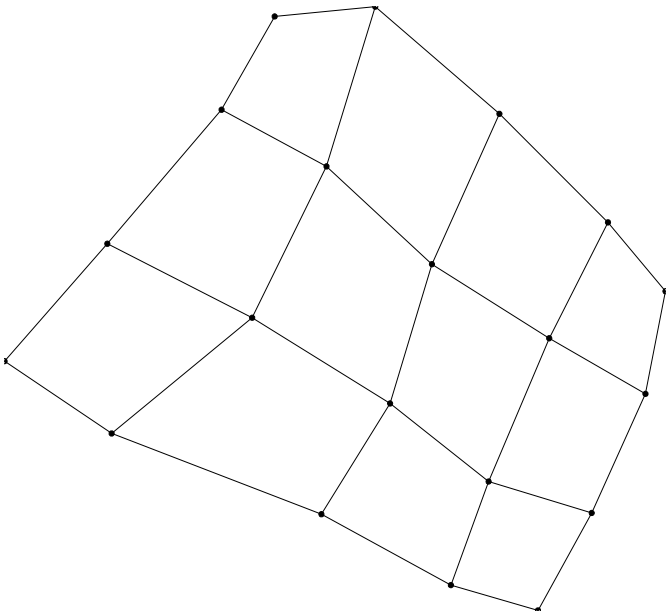


Figure 4-52: Two-dimensional portrayal of the distances between SOM nodes for single station heavy rainfall (SHR) soundings from 1977-2012.

The frequency of occurrence of each node is given in Fig. 4-54. There are three nodes: 1, 2 and 15 which occur between 6-8 % of the time (> 30 days). The rest of the nodes all occur between 4-6 % of the time and only nodes 9, 12 and 18 occurred on less than 23 days.

The nodes on the right hide side of SOM occur most frequently in October months (Fig.4-55). In this SOM the nodes with the stronger winds are in the top row with decreasing wind strength towards the bottom row of the SOM (Fig. 4-53). Nodes 4, 5 and 9 have similar characteristics of the AHR October nodes but the October nodes in the bottom two rows have different features. Nodes 13 and 14 are still quite moist with PW values of above 20 mm and mean layer dew point temperatures of above 10 °C, nodes 18-20 have PW of less than 20 mm and dew points of < 10 °C (Table 4-8). These 3 nodes have the coldest T_{53} of all the nodes with TD_{75} of around 17 °C. There are southerly winds in the 800-600 hPa layer and CAPE values of less than 800 Jkg^{-1}. The atmospheric conditions depicted in these 3 nodes are generally unfavourable for precipitation and yet more than 50 mm occurred at least at one station over Gauteng on the days mapped to these nodes. The mean value of the average daily rainfall over Gauteng for all the days mapped to this node was only 5 mm. The percentage of stations reporting rainfall > 0 mm on the days in these nodes was 34 %.

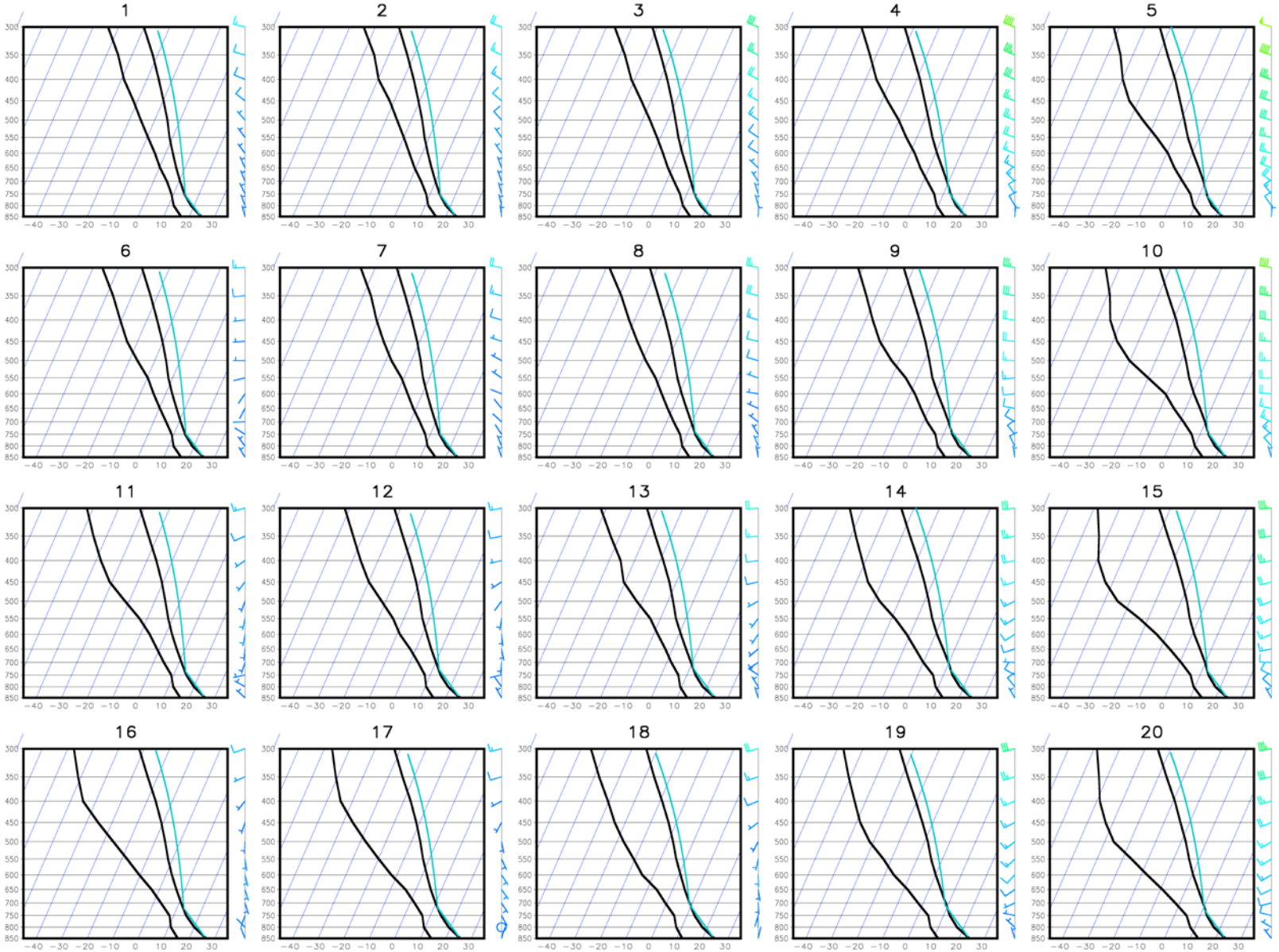


Figure 4-53: SOM of skew-t gram plots of the SHR soundings from 1977-2012. The solid blue lines plots the track of a parcel if lifted from the surface to the lifted condensation level by the dry adiabatic lapse rate and then follows the saturated adiabatic lapse rate to 300 hPa.

Table 4-8: The SOM of the sounding derived parameters of SHR soundings from 1977-2012.

Node No	PW	Td100	T53	CAPE	TD75	WCD	EKI	WV86	W64	MWSs4	Δθe s5	$\Theta_{e\ ave}$
1	29	13	-16	742	15	162	14	-2	5	7	-21	340
2	27	12	-17	636	16	155	13	-2	6	7	-19	338
3	25	11	-18	665	16	145	12	-3	8	10	-18	336
4	23	10	-19	712	18	126	10	-5	13	15	-18	333
5	21	10	-20	768	18	122	9	-5	15	17	-21	332
6	28	13	-17	1076	16	153	14	0	4	6	-23	339
7	26	12	-18	841	16	147	12	-1	4	7	-21	337
8	25	11	-19	933	17	132	12	-2	6	8	-20	335
9	23	11	-20	1066	18	119	11	-2	9	12	-22	334
10	21	11	-20	1038	18	122	8	-3	13	15	-25	333
11	26	13	-17	1323	17	142	12	1	4	6	-26	338
12	24	12	-18	977	17	129	9	1	4	6	-22	336
13	23	10	-19	1011	18	114	10	0	5	8	-21	334
14	21	10	-20	1046	18	107	6	0	8	10	-22	332
15	20	10	-20	1018	18	117	3	0	10	13	-25	332
16	23	12	-17	1147	17	127	7	1	4	6	-25	336
17	21	11	-18	758	16	116	4	1	4	6	-22	334
18	19	9	-20	605	17	97	1	1	5	7	-18	331
19	18	8	-21	693	17	90	-1	1	8	11	-18	330
20	18	9	-20	764	17	107	-3	1	10	12	-22	330

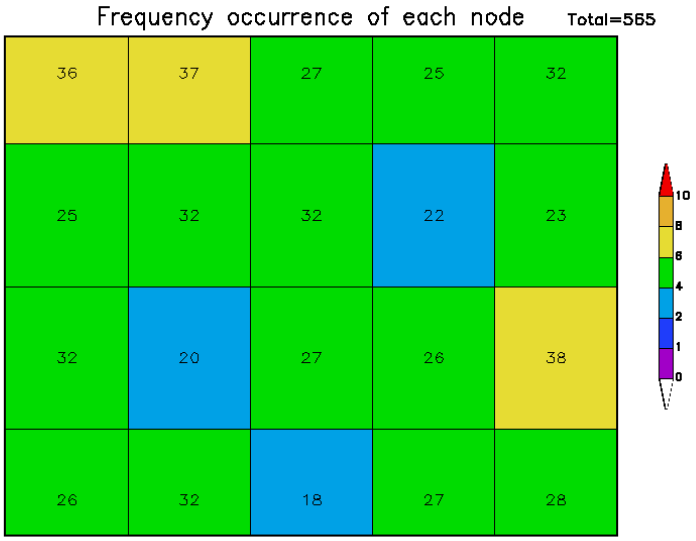


Figure 4-54: The numbers in each block are the total number of SHR soundings which was mapped to each node of the SOM in Fig-4-53, and the shades denote the frequency (%) of occurrence of each node.

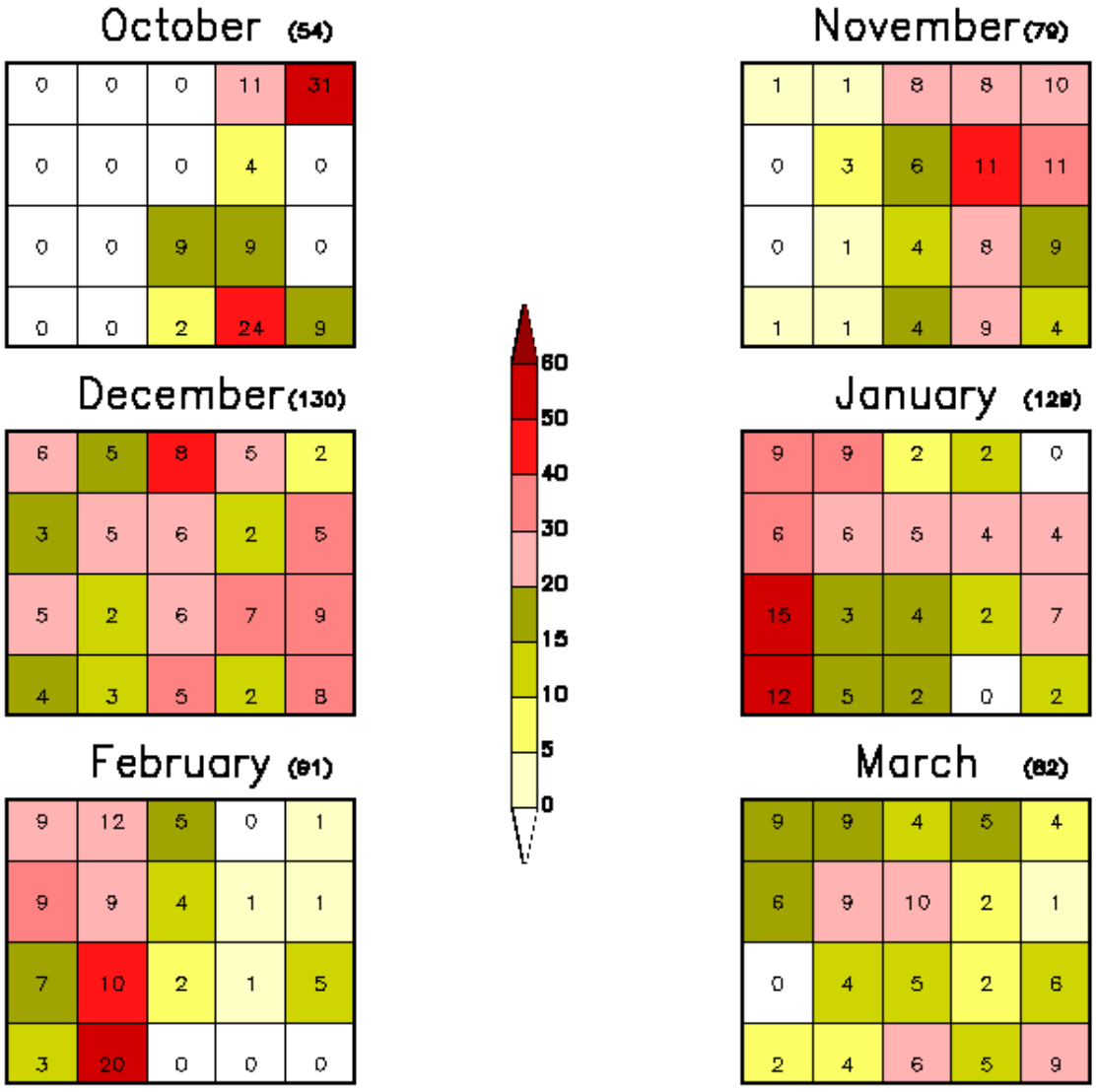


Figure 4-55: The number in each block is the frequency (%) of occurrence of each node relative to the total number of occurrences of SHR in that particular month. The shades are the frequency (%) of occurrence of each node expressed relative to the total number of occurrences of that node during the entire period.

In contrast the mean of the average Gauteng daily rainfall for nodes 4, 5 and 9 was 12 mm with 60 % of the stations reporting rainfall on average on the days mapped to these nodes. A possible explanation the SHR which occurred on days mapped to node 18-20 could be that the rainfall on these days would have resulted from isolated heavy events which occurred over only a portion of the province. A further point to ponder is how representative the 1200 UT sounding is of

the heavy rainfall which could have occurred any time during a 24-hr period. These three nodes occurred during all the months except for February.

During the remaining 5 months of the summer rainfall season, nearly all the nodes occur in every month. In November months the highest frequency of occurrence is in the right hand side of the SOM while in December months the distribution is fairly even over the entire SOM. In January and February months the highest frequency of occurrence is in the left hand side of the SOM.

Nodes 11, 12, 16 and 17 are the four nodes with the highest frequency of occurrence in January and February months. Nodes 11, 12 and 16 are very similar to the most frequent nodes which occur in January and February for AHR. In nodes 16 and 17 there are less moisture in the atmosphere (PW values 21-23 mm) but in node 16 the CAPE value is 1147 Jkg^{-1} with a large $\Delta\Theta_e$ value of -25 K. Node 17, similar to nodes 18-20, is unfavourable. Node 16 and 17 are the only 2 nodes where lower tropospheric winds are southeasterly veering with height to become southwesterly aloft.

The SHR SOM provides the same information as the AHR SOM but with the additional information that SHR can occur under less favourable conditions than AHR.

4.8.3 Summary of the SOM heavy rainfall climatology

The 20 node SOM was capable of capturing the main climatological features of the two heavy rainfall classes as was also discussed in Chapter 4.7. The SOM was capable of distinguishing the atmospheric conditions in early summer from those during late summer. The conditional instability dominant in early summer is replaced by a convective unstable atmosphere in late summer. This is reflected in the skew-t gram plots but also in the twelve sounding-derived parameters. The advantage of using the SOM climatology is that it provides more detailed information than a monthly climatology can provide and also relates the vertical profile of temperature, moisture, wind, and sounding-derived parameters into one node. The SOM of the sounding-derived parameters in Tables 4-7 and 4-8 could be used by operational weather forecasters to aid in the prediction of heavy rainfall. Critical values of the parameters could be adjusted during the season to guide them in identifying favourable atmospheric conditions for heavy rainfall.

4.9 Overview

The objectives addressed in this Chapter were to construct a climatology of soundings and sounding-derived parameters associated with heavy rainfall over Gauteng during summer. Emphasis was placed on distinguishing significantly large values of the parameters during heavy rainfall from

the climatological mean. The seasonal change of the thermodynamical profile over Gauteng was illustrated and adapting critical values to identify heavy rainfall during the season was accentuated. SOMs were used to construct a sounding climatology during heavy rainfall over Gauteng and this technique was capable of identifying the seasonality of heavy rainfall over the province. Moreover, it was adept in combining several different parameters into a two dimensional environment where the relationship between variables was captured. In Chapter 5, twelve sounding-derived parameters were and are combined with circulation criteria discussed in Chapter 3 and a method is proposed to forecast rainfall frequencies by using SOMs.

4.10 References

BOSART LF, SEIMON A, LAPENTA KD and DICKINSON MJ (2006). Supercell tornadogenesis over complex terrain: the Great Barrington, Massachusetts, tornado on 29 May 1995. *Weather and Forecasting*, **21**(6), 897-922

BRIMELOW JC (1999). Numerical modelling of hailstone growth in Alberta storms. M.S. thesis, Dept. of Earth and Atmospheric Sciences, University of Alberta, 153 pp.

BROOKS HE, DOSWELL CA III and COOPER J (1994). On the environments of tornadic and nontornadic mesocyclones. *Weather and Forecasting*, **9**, 606–618.

BROOKS HE and STENSRUD DJ (2000). Climatology of heavy rain events in the United States from hourly precipitation observations. *Monthly Weather Review,* **128** (4), 1194-1201.

BUNKERS MJ, KLIMOWSKI BA, ZEITLER JW, THOMPSON RL and WEISMAN ML (2002). The importance of parcel choice and the measure of vertical wind shear in evaluating the convective environment. Preprints, *21st Conference on Severe Local Storms*, San Antonio, TX, American Meteorological Society, 379–382.

BUNKERS MJ, WETENKAMP JR JR, SCHILD JJ and Fischer A. (2010). Observations of the relationship between 700-mb temperatures and severe weather reports across the contiguous United States. *Weather and Forecasting*, **25**(2), 799-814.

CHAUDHARI H, SAWAISARJE G, RANALKAR M and SEN P (2010). Thunderstorms over a tropical Indian station, Minicoy: Role of vertical wind shear. *Journal of Earth System Science,* **119,** 603-615.

COHEN AE, CONIGLIO MC, CORFIDI SF and CORFIDI SJ (2007). Discrimination of mesoscale convective system environments using sounding observations. *Weather and Forecasting*, **22**, 1045–1062.

COVADONGA, GIAIOTTI, DARIO, STEL, FULVIO, CASTRO, AMAYA, FRAILE and ROBERTO (2010). Maximum hailstone size: Relationship with meteorological variables. *Atmospheric Research,* **96**(2), 256-265

CRAVEN JP AND BROOKS HE (2004). Baseline climatology of sounding-derived parameters associated with deep moist convection, *National Weather Digest,* **28**, 13-24.

CRAVEN JP, JEWELL RE, and BROOKS HE (2002). Comparison between observed convective cloud-base heights and lifting condensation level for two different lifted parcels. *Weather and Forecasting*, **17**, 885–890.

DANIEL EZ (2006). Forecasting mixed-layer height over complex terrain. Preprints of the 12th Conference on Mountain Meteorology. https://ams.confex.com/ams/pdfpapers/114269.pdf (Accessed on 22 January 2012)

DE RUBERTIS D (2006). Recent trends in four common stability indices derived from U.S. radiosonde observations. *Journal of Climate*, **19**, 309–323.

DEMOTT A and RANDALL D (2004). Observed variations of tropical convective available potential energy. *Journal of Geophysical Research,* **109,** D02102*, doi:10.1029/2003JD003784.*

DIMITROVA T, MITZEVA R and SAVTCHENKO A (2009). Environmental conditions responsible for the type of precipitation in summer convective storms over Bulgaria. *Atmospheric Research,* **93,** 30–38

DOSWELL CA III, BROOKS HE and MADDOX RA (1996). Flash flood forecasting: An ingredients-based methodology. *Weather and Forecasting*, **11**, 560-580.

DOSWELL CA III, CARACENA F and MAGNANO M (1985). Temporal evolution of 700-500-mb lapse rate as a forecasting tool–A case study. Preprints, *14th Conf. on Severe LocalStorms*, Indianapolis, IN, American Meteorological Society, 398-401.

DOSWELL CA III and RASMUSSEN EN (1994). The effect of neglecting the virtual temperature correction on CAPE calculations. *Weather and Forecasting*, **9**, 625-629

DOSWELL CA III and SCHULTZ DM (2006). On the use of indices and parameters in forecasting severe storms. *Electronic Journal Severe Storms Meteorology,* **1**, 1–14.

DUPILKA ML and REUTER GW (2006a). Forecasting Tornadic Thunderstorm Potential in Alberta Using Environmental Sounding Data. Part I: Wind Shear and Buoyancy. *Weather and Forecasting,* **21,** 325–335 .

DUPILKA ML and REUTER GW (2006b). Forecasting Tornadic Thunderstorm Potential in Alberta Using Environmental Sounding Data. Part II: Helicity, Precipitable Water, and Storm Convergence. *Weather and Forecasting,* **21,** 336–346.

DURRE I, VOSE RS and WUERTZ DB (2006). Overview of the Integrated Global Radiosonde Archive. *Journal of Climate*, **19**(1), 53-68.

DYSON LL, ENGELBRECHT Cj, TURNER K AND LANDMAN S (2012). A short term heavy rainfall forecasting system for South Africa with first implementation over the Gauteng Province. Water Research Commission. In Press

DYSON LL and VAN HEERDEN J (2002). A model for the identification of tropical weather systems. *Water SA,* **28**(3), 249-258.

GEORGE JJ (1960). *Weather Forecasting for Aeronautics.* Academic Press, 673 pp.

GROENEMEIJER PH and VAN DELDEN A (2007). Sounding-derived parameters associated with large hail and tornadoes in the Netherlands. *Atmospheric Research*, **83**, 473–487

HARNACK RP, JENSEN DT CERMAK III JR (1998). Investigation of upper-air conditions occurring with heavy summer rain in Utah. *International Journal of Climatology*, **8**, 701-732.

HENRY NL (2000). A static stability index for low-topped convection, *Weather and Forecasting*, **15**, 246-254.

HEWITSON BC and CRANE RG (2002). Self-organizing maps: applications to synoptic climatology. *Climate Research,* **22**, 13–26.

HOLTON J.R. (1992). *An Introducion to Dynamic Meteorology*. Academic Press. 511 pp.

HOUSTON AL and WILHELMSON RB (2012). The Impact of Airmass Boundaries on the Propagation of Deep Convection: A Modeling-Based Study in a High-CAPE, Low-Shear Environment. *Monthly Weather Review*, **140**, 167–183.

JENSEN AA, THOMPSON AM, and SCHMIDLIN FJ (2012). Classification of Ascension Island and Natal ozonesondes using self-organizing maps, *Journal Geophysical Research*, **117**, D04302, doi:10.1029/2011JD016573

KOHONEN T (2001) *Self-organizing maps.* Third Edition, Springer, 511 pp

KOHONEN T, HYNNINEN J, KANGAS J and LAAKSONEN J (1996). SOM_PAK: The Self-Organizing Map Program Package. Technical Report A31, Helsinki University of Technology, Laboratory of Computer and Information Science, FIN-02150 Espoo, Finland.

LIU Y and WEISBERG RH (2011). A review of Self-Organizing Map applications in meteorology and oceanography. In *Self-Organizing Maps - Applications and Novel Algorithm Design*, Edited by: Mwasiagi JI 253–272. Rijeka, Croatia: InTech. ISBN 978-953-307-546-4

MANZATO A (2003). A climatology of instability indices derived from Friuli Venezia Giulia soundings, using three different methods. *Atmospheric Research*, **67–68**, 417–454.

MARKET PSA, SCOFIELD R, KULIGOWSKI R and GRUBER A (2003). Precipitation Efficiency of Warm-Season Midwestern Mesoscale Convective Systems. *Weather and Forecasting*, **18**, 1273–1285.

MILLER RC (1967). Notes on analysis and severe storm forecasting procedures of the Military Weather Warning Center. AWS Technical Report 200 (revised), 170 pp. (Available from Headquarters, Air Force Weather Agency, Scott AFB, IL 62225.)

MURTHY BS and SIVARAMAKRISHNAN S (2006). Moist convective instability over the Arabian Sea during the Asian summer monsoon, 2002. *Meteorological Applications*, **13**, 63–72.

RASMUSSEN EN and BLANCHARD DO (1998). A Baseline Climatology of Sounding-Derived Supercell and Tornado Forecast Parameters. *Weather and Forecasting*, **13**, 1148–1164.

REUSCH DB, ALLEY RB and HEWITSON BC (2005). Relative performance of self-organizing maps and principal component analysis in pattern extraction from synthetic climatological data. *Polar Geography*, **29**, 188–212.

ROUAULT M, ROY SS and BALLING RC (2012). The diurnal cycle of rainfall in South Africa in the austral summer. *International Journal of Climatology*, doi: 10.1002/joc.3451

SAWB, SOUTH AFRICAN WEATHER BUREAU (1990). Climate of South Africa. Part 14: Upper air statistics from 1968-1987. South African Weather Service, Pretoria, South Africa.

SCHULTZ DM, SCHUMACHER PN and DOSWELL CA III (2000). The intricacies of instabilites. *Monthly Weather Review,* **128**, 4143-4148

SHOWALTER AK (1953). A stability index for thunderstorm forecasting. *Bulletin of the American Meteorological Society,* **34,** 250–252.

STEYN AGW, SMIT CF, DU TOIT SHC and STRASHEIM C (1994). *Modern statistics in practice.* J.L Van Schaik, 761 pp.

TENNANT W and HEWITSON BC (2002). Intra-seasonal rainfall characteristics and their importance to the seasonal prediction problem. *International Journal of Climatology,* **22**, 1033–1048.

TYAGI B, NARESH KRISHNA V and SATYANARAYANA A (2011). Study of thermodynamic indices in forecasting pre-monsoon thunderstorms over Kolkata during STORM pilot phase 2006–2008. *Natural Hazards* **56,** 681-698.

VAN SCHALKWYK L and DYSON LL (2012). Climatological characteristics of fog at Cape Town International Airport. Accepted by *Weather and Forecasting.*

WILLIAMS E and RENNO N (1993). An analysis of the conditional instability of the tropical atmosphere. *Monthly Weather Review,* **121**, 21–36.

WU X and YANA M (1994). Effects of vertical wind shear on the cumulus transport of momentum: Observations and parameterization. *Journal Atmospheric Science,* **51,** 1640–1660.

5 Self-organizing maps as a predictive tool.

In this chapter a methodology is proposed where the twelve sounding-derived parameters identified in Chapter 4 and circulation criteria used in Chapter 3 are combined in a SOM. The output from the SOM is then used to predict the daily frequency distribution of rainfall over Gauteng.

5.1 Preface

Since Kohonen (1989) first proposed SOMs it has been in widespread use in several disciplines to analyse, cluster and visualise data. For instance in the biological sciences (Tamayo, 1999), in the financial sector (Deboeck and Kohonen ,1998), in human sciences (Frenkel et al., 2012) and in sport sciences (Grunz et al., 2012). SOMs have been used to forecasts exchange rates (Khashei and Bijari, 2012), time series (Simon et al.; 2005, Hsu et al., 2010) and traffic flows (Van der Voort et al., 1996) to name but a few.

SOMs have also been widely used in meteorological applications since the start of the 21st century but with limited forecasting application. There have been significant contributions to synoptic climatologies and synoptic patterns of atmospheric circulation (Liu and Weisberg, 2011). Climate variability and extreme events have been analysed with SOMs (Liu and Weisberg, 2011) and trend analyses performed (Bermejo and Ancell, 2009). Methods were developed to downscale both seasonal and short-range weather forecasts (Gutierrez et al., 2005) and to utilize SOMs to forecasts monthly rainfall (Rivera et al., 2012). Tadross et al. (2005) used SOMs to illustrate the differences in synoptic states associated with early and late onset of the maize growing season over southern Africa. Some of the meteorological forecasting applications of SOMs include predicting electrical power load using temperature distributions (Nagi et al., 2008 and Senjyu et al., 2000). Power load forecasts work on the assumption that the SOM correctly identifies to which node a temperature distribution belong and then uses the previously identified power usage associated with the specific node to forecast the power load usage. Esmaeili (2011) proposed a method to predict wind speeds by using the SOM to find the relationship between the wind speed of the previous day and the day being forecast for. Ismail et al. (2010) used rainfall data to predict the river flow in the Bernam River in Malaysia. Nishiyama et al. (2007) used the wind components at 850 hPa and precipitable water values in a SOM to predict rainfall over Japan. They found that the SOM could successfully extract heavy rainfall events by correctly identifying the synoptic circulation associated with heavy rainfall.

This Chapter builds on the work done by Nishiyama et al. (2007) but uses sounding-derived parameters and circulation criteria in a SOM to predict rainfall. Several types of synoptic circulation systems may cause heavy rainfall to occur over Gauteng. These weather systems were expanded on

in Chapter 3 and it was shown how weather systems associated with conditional instability and wind shear is responsible for heavy rainfall in early summer but in late summer the atmosphere becomes convectively unstable and wind shear values become very small. It was also shown that in the NCEP data set horizontal wind convergence generally occurs from the surface to 500 hPa at four grid points surrounding Gauteng whenever heavy rainfall occurs, irrespective of the month. The magnitude of the convergence differs from early to late summer but the general pattern remains one of a deep (± 5 km) layer of wind convergence replaced by wind divergence at 300 hPa. It was also shown that on average cyclonic vorticity advection occurs at these four grid points at 300 hPa when heavy rainfall occurs.

In Chapter 4 it was shown how 12 sounding-derived parameters identify the general atmospheric conditions when heavy rainfall occurs over Gauteng. It was also shown how the values of these variables vary during the season. Furthermore SOMs were found to be very useful in creating a sounding climatology and this method was capable of capturing the seasonal change in atmospheric conditions associated with heavy rainfall.

The capability of the SOM to convert complex non-linear features into simple two-dimensional relationships is exploited in this Chapter. Different variables are combined in the SOM and the results are used to predict rainfall frequencies.

5.2 Data and methodology

The twelve sounding-derived parameters identified in Chapter 4 were calculated at Irene for every day that 1200 UT sounding data were available in the summer months from January 1979 to March 2005. NCEP data valid for 1200 UT were used to calculate horizontal wind divergence at the four grid points surrounding Gauteng (Fig. 3-1) at 850, 700, 500 and 300 hPa as well as the relative vorticity advection at 300 hPa for the same period. These 12 sounding-derived parameters and 20 circulation criteria values were used to construct the arrays used in the SOM (Fig. 5-1). The sounding-derived parameters contribute to just over one third of the elements in an array and the circulation values to more than 60 %. The circulation criteria values contribute about double the number of elements because they capture a larger geographical area and are also available at several levels in the atmosphere. By choosing not to allow the 1 to 2 ratio to be exceeded, it was attempted to not let the circulation criteria dominate the data arrays used in the SOM. All the data elements were standardised with respect to the average and standard deviation prior to running the SOM. This 3858 day period will be referred to as the training period (Fig. 5.1). A 5 x 4- or 20 node SOM was again constructed.

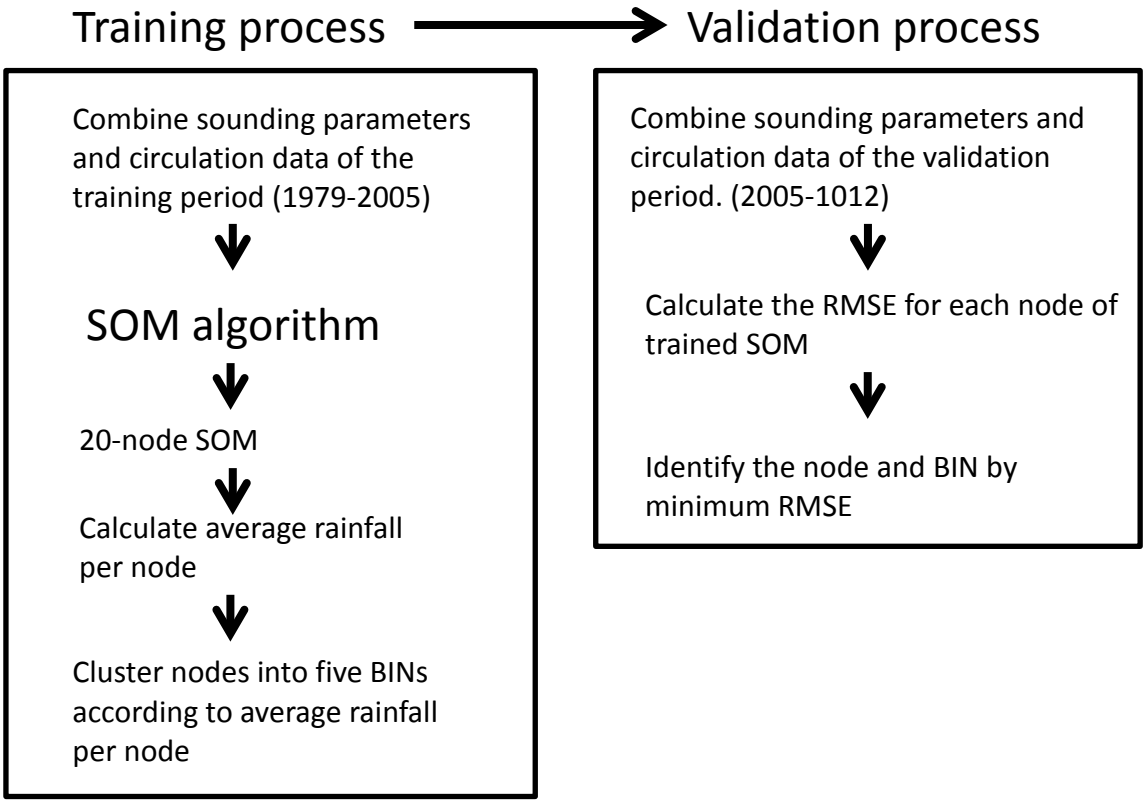


Figure 5-1: The training process involving the classification of the sounding derived parameters and circulation criteria and the validation process involved in predicting rainfall frequency distributions.

As illustrated in Fig. 5.1 after the SOM was run the mean Gauteng rainfall was calculated in each of the 20 nodes. Two Gauteng rainfall values are considered similar to the definitions provided in Chapter 2.1, the area-average daily rainfall as well the rainfall at a single station. These nodes were then clustered together into BINs where BIN1 represents the four nodes with the highest mean rainfall, BIN2 the four nodes with the second highest mean rainfall, up to BIN5 which contains the four nodes with the lowest mean rainfall.

The next step in this prediction experiment was to construct similar data arrays used in the training period for a validation period (Fig. 5.1). The validation period was for all summer days (803 days) when 1200 UT sounding data were available from October 2005 to March 2012. The root mean square error (RMSE) was then calculated between the arrays created for every day in the validation period and each of the 20 nodes created by the SOM in the training period. Prediction is then made that the day being tested will be in the same node as the one for which the RMSE (equation 5-1) is the smallest (Willks, 2006). The forecast assumes that the rainfall for the day being tested will have the same rainfall characteristics as that of the BIN identified during the validation period.

$$RMSE = \sqrt{\frac{\sum_{m=1}^{M}(y_m - o_m)^2}{M}} \qquad 5\text{-}1$$

where

M=32 (total number of elements in an array)

y_m= elements of the SOM node (20 nodes)

o_m= elements of a day in the validation period

The distribution of the rainfall of the training and validation period is compared in three ways. First frequency distributions were created for each of the BINs and the validation and training period frequencies are plotted together for graphical inspection. Furthermore, the two-sample Kolmogorov-Smirnov test (K-S test) is performed on each of the BINs in the training and validation period in order to test whether the samples were drawn from the same distribution (Willks, 2006). The K-S test is a nonparametric test which compares two samples. It is sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples. The confidence level for the test was chosen at 99 %. Lastly the mean and standard deviation were calculated for each of the BINs in the training and validation period.

Overfitting occurs when a statistical model is too complex by having too many predictors relative to the number of observations (Tetko, 1995). In the methodology proposed in this chapter, 32 variables were used to predict rainfall frequency at a single location. Willks (2006) warned against the danger of simply increasing the number of predictors until the apparent best relationship is achieved. A model which has been overfit will generally have poor predictive performance. Willks (2006) provides a set of guidelines when choosing predictors. They are:

- Use only physically meaningful predictors
- Develop the statistical relationship on a set of sample data and test this relationship on a totally independent data set. A very large difference in the performance between the initial data and the independent data shows that the statistical relationship has been overfit.

- The initial (training) data set needs to be large enough so that the statistical relationship is stable. The stability can be determined by testing on independent data.

In this methodology the 32 variables identified to train the SOM were all shown to be associated with heavy rainfall over Gauteng in one way or another and the first requirement as set out by Willks(2006) was therefore adhered to. The second and third requirement to avoid overfitting were also dealt with as the SOM was trained on 3858 days and consequently tested on an independent data set which contains 803 days. The performance of the training and validation period was very similar (see Section 5.3).

When using principle component or regression analysis problems may arise when the predictors are dependent and mutually correlated as some of the predictors may contain redundant information. Analysing data with SOMs does not require that the input variables are independent and the nodes represents the distribution of the observed data (Kohonen, 2001). Furthermore, there is no limit to the number of variables that the SOM can be trained with, see for instance Skupin and Hagelman (2005) and Fabrikant and Skupin (2005).

5.3 The training period

5.3.1 Self-organizing map results

Table 5.1 is the SOM of sounding parameters of the training period and Figs. 5-2 and 5-3 are the SOMs of the circulation criteria. In Fig. 5-2 the vertical profile of horizontal wind divergence is provided at the 4 grid points surrounding Gauteng and in Fig. 5-3 the 300 hPa vortcity advection is given. In nodes 1 to 10 the atmosphere has relatively high moisture content with PW values above 20 mm and mean layer dew point temperatures > 10 °C. Nodes 16-20 are very dry with mean layer dew point temperatures as low at 4 °C.

Table 5-1: The SOM of the sounding derived parameters for all Irene soundings for the training period from January 1979 to March 2005.

Node No	PW	Td100	T53	CAPE	TD75	WCD	EKI	WV86	W64	MWSs4	Δθe s5	θe ave
1	24	12	-18	840	16	146	8	-1	6	8	-21	335
2	24	12	-18	1016	17	143	9	0	6	8	-23	335
3	23	12	-18	1132	17	137	8	0	6	8	-24	335
4	23	11	-19	1158	17	129	9	1	7	8	-24	335
5	23	11	-19	1164	18	124	8	1	8	9	-24	335
6	23	11	-18	613	16	142	6	-1	6	8	-19	334
7	22	11	-18	793	16	137	6	0	6	8	-21	335
8	22	11	-19	934	17	127	7	1	6	8	-22	335
9	22	11	-19	992	18	117	8	0	8	9	-23	334
10	22	11	-19	1065	18	113	8	-1	9	10	-23	334
11	19	9	-19	390	16	123	-2	0	6	9	-16	331
12	19	9	-19	476	16	113	-1	1	6	9	-18	332
13	18	8	-20	574	17	94	0	1	8	10	-18	331
14	18	8	-20	645	18	83	3	-1	10	12	-19	331
15	19	8	-20	736	19	78	4	-2	11	13	-19	332
16	15	7	-21	230	16	102	-11	0	7	10	-13	329
17	15	6	-21	235	16	83	-12	1	8	11	-13	328
18	14	5	-21	266	17	58	-9	1	9	12	-13	328
19	14	4	-21	311	18	39	-5	-1	12	14	-14	328
20	15	5	-21	433	19	40	-1	-2	13	14	-15	329

Table 5-2: A) The area-average daily Gauteng rainfall per node and B) the maximum single station daily rainfall per node. The colours indicate to which BIN each node was assigned. Dark blue BIN1, light blue BIN2, light brown BIN3, darker brown BIN4 and orange BIN5.

A

4.2	4.4	3.2	3.2	4.4
2.4	2.8	3.1	4.1	4.5
1.4	1.1	2.0	2.7	3.5
0.4	0.6	0.6	0.4	1.1

B

32.1	36.2	29.2	28.2	32.8
24.6	27.1	28.6	33.0	33.1
17.1	15.1	19.6	22.3	24.8
8.6	8.7	9.5	8.1	16.4

Comparing Table 5-1 with Table 5-2 shows that node 1-10 had the highest daily area-average as well as daily single station rainfall. These nodes all have abundance of moisture relatively large CAPE values and even though the temperature difference between 700 and 500 hPa (TD_{75}) is quite small the change of the equivalent potential temperature with height ($\Delta\Theta_e$) has favourable (large negative) values. These nodes nearly all have northerly winds in the 800-600 hPa layer with relatively small average wind speeds in the 600-400 hPa layer (W_{64}) and with large warm cloud depth (WCD)

values. These are all favourable conditions for the formation of rainfall. The nodes in the first 2 rows are also associated with a deep layer of horizontal wind convergence (Fig. 5-2) and becoming increasingly favourable towards the right hand side of the SOM as the values of wind convergence increase. These 10 nodes are all also, at least partially, under the influence of cyclonic vorticity advection at 300 hPa (Fig. 5-3). It is interesting to note that the nodes with the highest area-average daily and maximum single station rainfall (dark blue in Table 5-2) all have cyclonic and anticyclonic 300 hPa vorticity advection occurring together over Gauteng. Note that nodes 6 and 7 have lower daily area-average rainfall values than the other nodes from 1-10. Although these nodes are moist (PW > 22 mm) the TD_{75} values are relatively low while the $\Delta\Theta e$ values are relatively large. It was shown in Chapter 4 that one of these two variables should maximise for heavy rainfall to occur. Furthermore Fig. 5-2 indicates that Gauteng experiences wind convergence only in the southwest when these two nodes occur.

The average rainfall for nodes 11-20 is less than the rainfall in the first two rows of the SOM. Table 5-1 shows how these nodes become drier with small CAPE values, negative EKI values and relatively unfavourable values of $\Delta\Theta_e$. The nodes in the bottom left hand side of the SOM (11, 12, 16 and 17) experience predominantly wind divergence, making it unfavourable for the development of rainfall. This despite the cyclonic vorticity advection at 300 hPa in these nodes. The nodes in the bottom right hand side of the SOM (14, 15, 19 and 20) have deep layers of convergence but have relatively strong anticyclonic vorticity advection.

Node number 15 is interesting as it has a relatively high daily area-average rainfall value even though Table 5-1 shows that this node is quite dry, with relatively low T_{53} and large TD_{75} values. The CAPE value is above 700 Jkg^{-1} indicative of atmospheric instability and this is associated with a large magnitude of wind shear (MWS_{s4}) value. A further favourable factor for the development of rainfall is that this node is associated with horizontal wind convergence throughout most of the troposphere (Fig. 5-2)

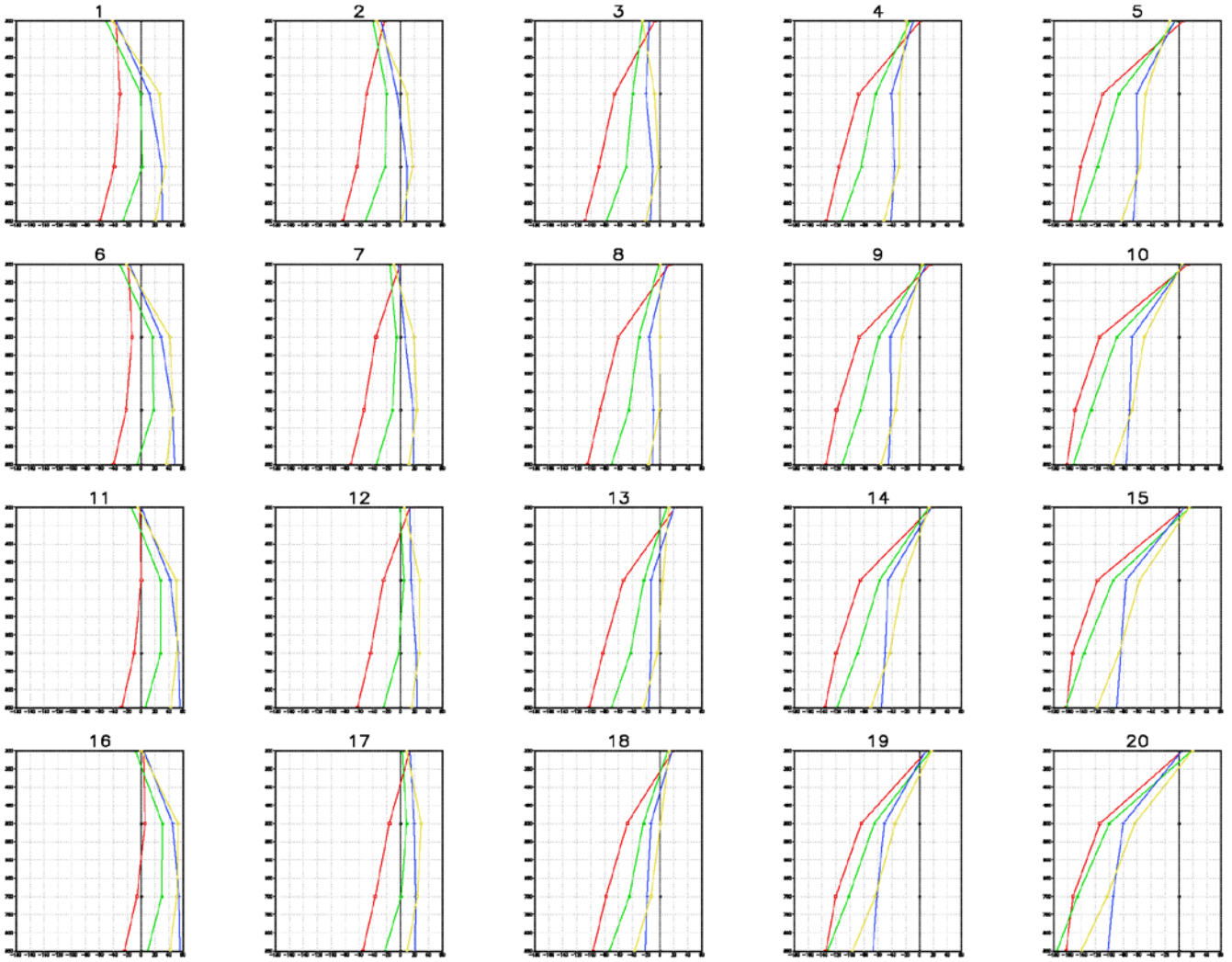


Figure 5-2: SOM of the vertical profiles of horizontal wind divergence at the four grid points surrounding Gauteng for the training period. Divergence values in 10^6 ms^{-1}. Negative values indicate wind convergence and positive values wind divergence. The blue line is for grid point 1 in Fig 3-1, the yellow line for grid point 2, the red line for grid point 3 and the green line for grid point 4.

Figure 5-3: SOM of the 300 hPa vortcity advection at the four grid points surrounding Gauteng. Cyclonic vorticity advection is indicated in red and anticyclonic vorticity advection is indicated in yellow The vorticity advection values are expressed in 10^{-10} s^{-2}.

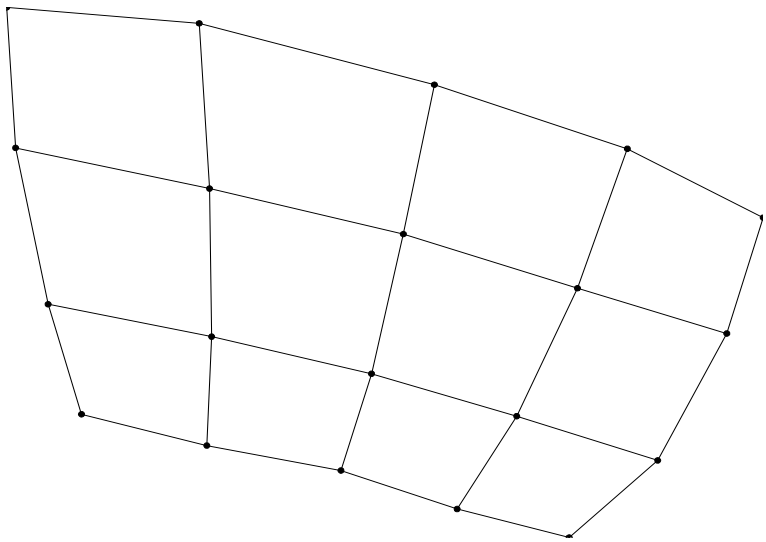


Figure 5-4: Two-dimensional portrayal of the distances between SOM nodes for the training period.

5.3.2 Frequency distribution of rainfall during the training period

The rainfall in Table 5-2 was binned together into 5 classes with BIN1 containing the four nodes with the highest rainfall (dark blue), BIN 2 the four nodes with the second highest rainfall (light blue) up to BIN5, which contains the four nodes with the lowest amount of rainfall (orange). The frequency distribution of these 5 BINS for the area-average daily rainfall and the maximum single station rainfall is shown in Fig. 5-5 and Fig. 5-6 respectively.

Whenever a day was mapped to the nodes in BIN1 and BIN2, the area-average daily rainfall exceeded 0 mm more than 80 % of the time while in the long-term average frequency distribution (black line in Fig. 5-5) 0 mm is exceeded 60 % of the time (Fig. 5-5). For all other rainfall thresholds BIN1 and BIN2 also have the highest frequency of occurrence. On average 10 mm is exceeded around 4 % of the time but when the nodes in BIN1 occurs it is exceeded 10 % of the time and 6 % for BIN2. From Fig. 5-5 it is clear that the days mapped to the nodes in BIN1 and BIN2 have a higher probability to be associated with heavy rainfall. Whenever the days which were mapped to BIN4 and BIN5 occur, the daily area-average rainfall was 0 mm, 54 % and 78 % of the time respectively. Even though daily area-average rainfall of more than 10 mm still occurs on days which were mapped to these nodes, it happens less than 1 % of the time. These two BINs are associated with dry conditions. The frequency distribution of BIN3 lies very close to the long-term average frequency distribution.

The frequency distributions of the five BINs for single station daily rainfall have a similar pattern than those for the area-average daily rainfall. On the days mapped to BIN 1 and 2, rainfall > 0 mm occurred at a station over Gauteng more than 90 % of the time (the average value is also quite high at 85 %). More than 10 mm was exceeded more than 80 % of the time, compared to the average value of 65 %. Fifty mm is exceeded on average 10 % of the time but nearly 20 % of the time when the days mapped to the nodes in BIN1 occur. Whenever the days are mapped to BIN3 and 4, more than 0 mm was still exceeded 81 % and 58 % of the time, respectively. More than 10 mm precipitated on 54 % of the days when BIN4 occurred, but only on 30 % of the days when BIN5 occurred. More than 50 mm was recorded on only 5 % (BIN4) and 2 % (BIN5) of the days when mapped to these nodes.

The output of the SOM process comprises metrological variables which can be related to rainfall. Heavy rainfall has a higher frequency of occurrence in some of the nodes while drier conditions prevail when other nodes occur.

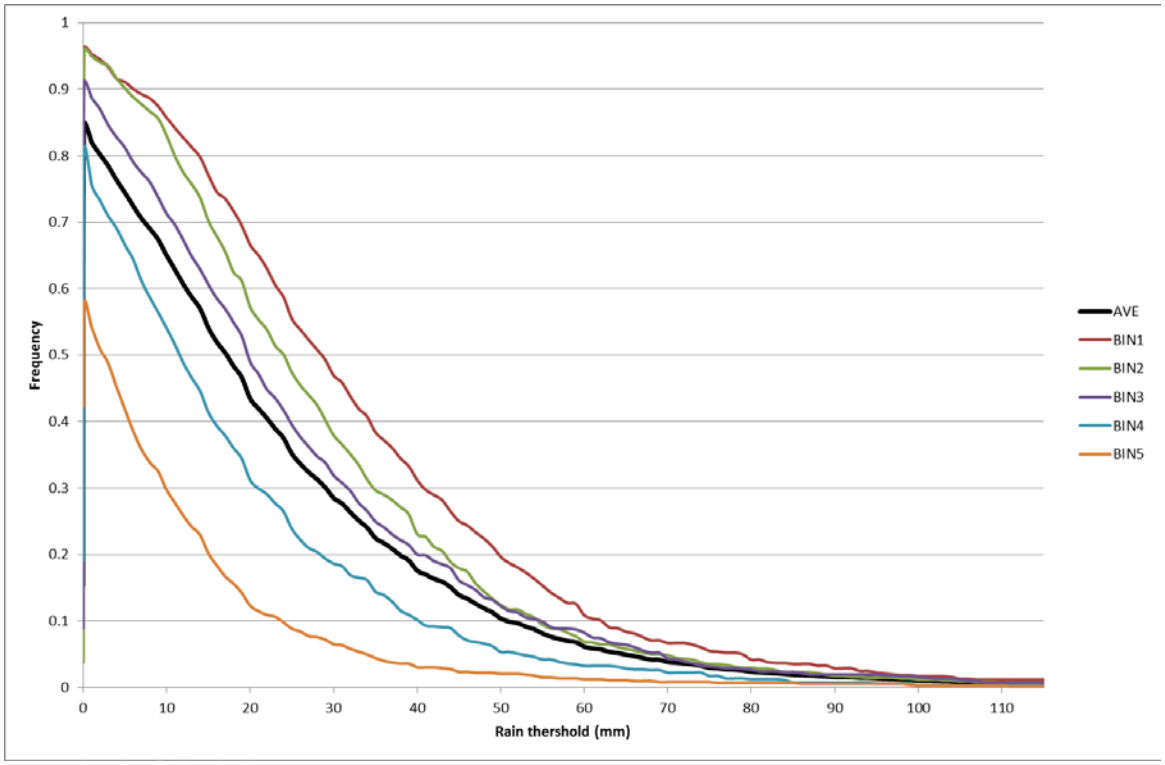


Figure 5-5: Frequency distribution of daily average maximum rainfall for the five BINs from January 1979 to March 2005.The black line indicates the climatological frequency distribution where all days are considered.

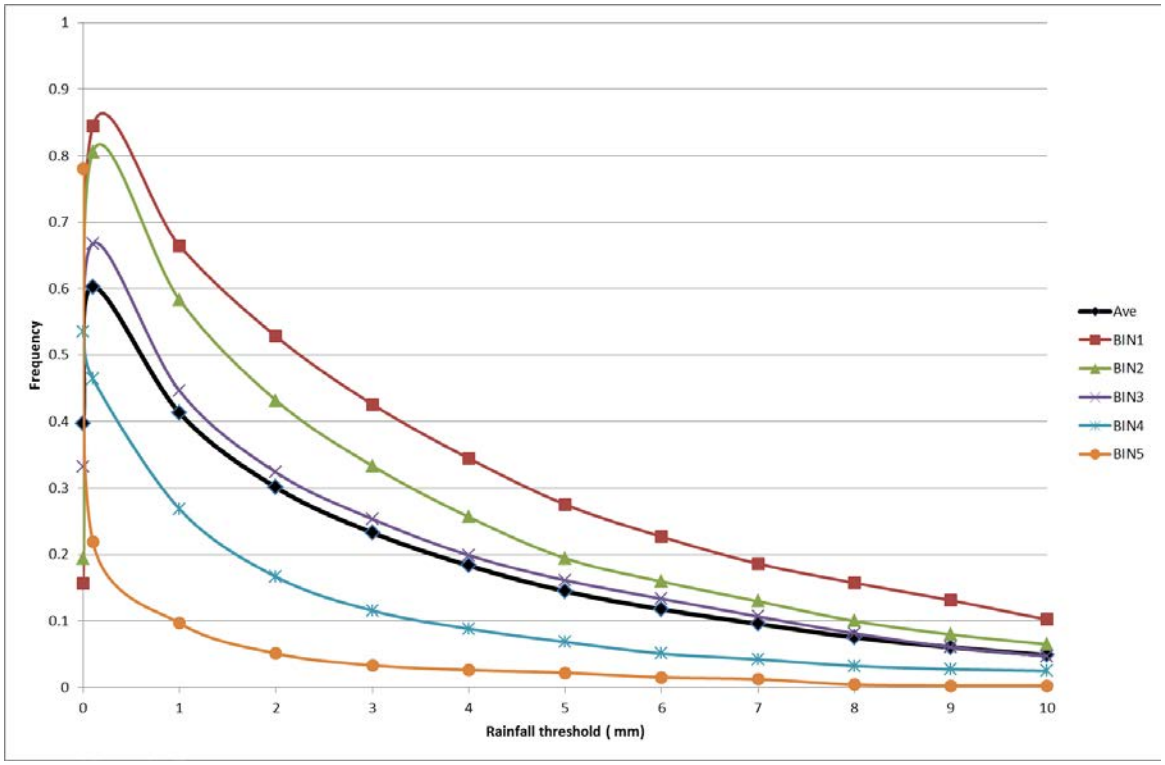


Figure 5-6: Frequency distribution of daily area-average rainfall for the five BINs from January 1979 to March 2005.The black line indicates the climatological frequency distribution where all days are considered.

5.4 Predicting the frequency distribution of rainfall in a validation period

Each day in the validation period was assigned to one of the nodes identified during the training period with the procedure described in Chapter 5-2. The days mapped to the nodes in the validation period were binned together in the same way as in the training period and rainfall frequencies calculated (Fig. 5-1). Figs. 5-7 and 5-8 show the frequency distribution of the BINs for daily average rainfall and maximum single station rainfall, respectively.

The frequency distribution of all days in the validation period and all days in the training period was compared first to ensure that the rainfall distributions during the two periods were similar. The thick black and dotted black lines in Figs. 5-7 and 5-8 indicate that these distributions were very similar. For average daily rainfall (Fig. 5-7) the frequency in the validation period is slightly lower for the higher rainfall thresholds than the frequencies in the training period. In the maximum single station daily rainfall class the frequency of the validation period is slightly higher than for the training period for rainfall values > 30 mm (Fig. 5-7). The mean value of the average daily rainfall during the training period when all the days were used (Ave) was 2.5 mm for the training period and 2.4 mm in the validation period (Table 5-3). The standard deviations of the average daily rainfall for the training and validation periods were also very similar. The mean and standard deviation of the single station daily rainfall when all days were used (Ave) was also very close. The K-S test indicated that the training and validation data samples were from the same distribution for the area-average daily and maximum single station daily rainfall classes.

Table 5-3: The mean rainfall and the standard deviation for the area-average daily (left) and single station daily rainfall (right) for the 5 BINs as well as the average (Ave) where all the days are used. The white cells are for the training period and the shaded cells for the test period.

	Average daily rainfall				Maximum daily rainfall			
	Mean		StdDev		Mean		StdDev	
Ave	2.5	2.4	4.1	3.8	22.9	24.3	23.2	26.0
BIN1	4.4	4.3	5.1	5.2	33.8	33.5	26.4	26.8
BIN2	3.5	3.6	4.3	4.1	28.8	31.9	21.7	30.4
BIN3	2.8	2.2	4.3	2.7	25.6	27.5	23.9	25.3
BIN4	1.4	1.1	2.9	2.5	17.1	16.2	18.9	20.9
BIN5	0.5	0.5	1.4	1.6	8.8	10.2	14.3	15.1

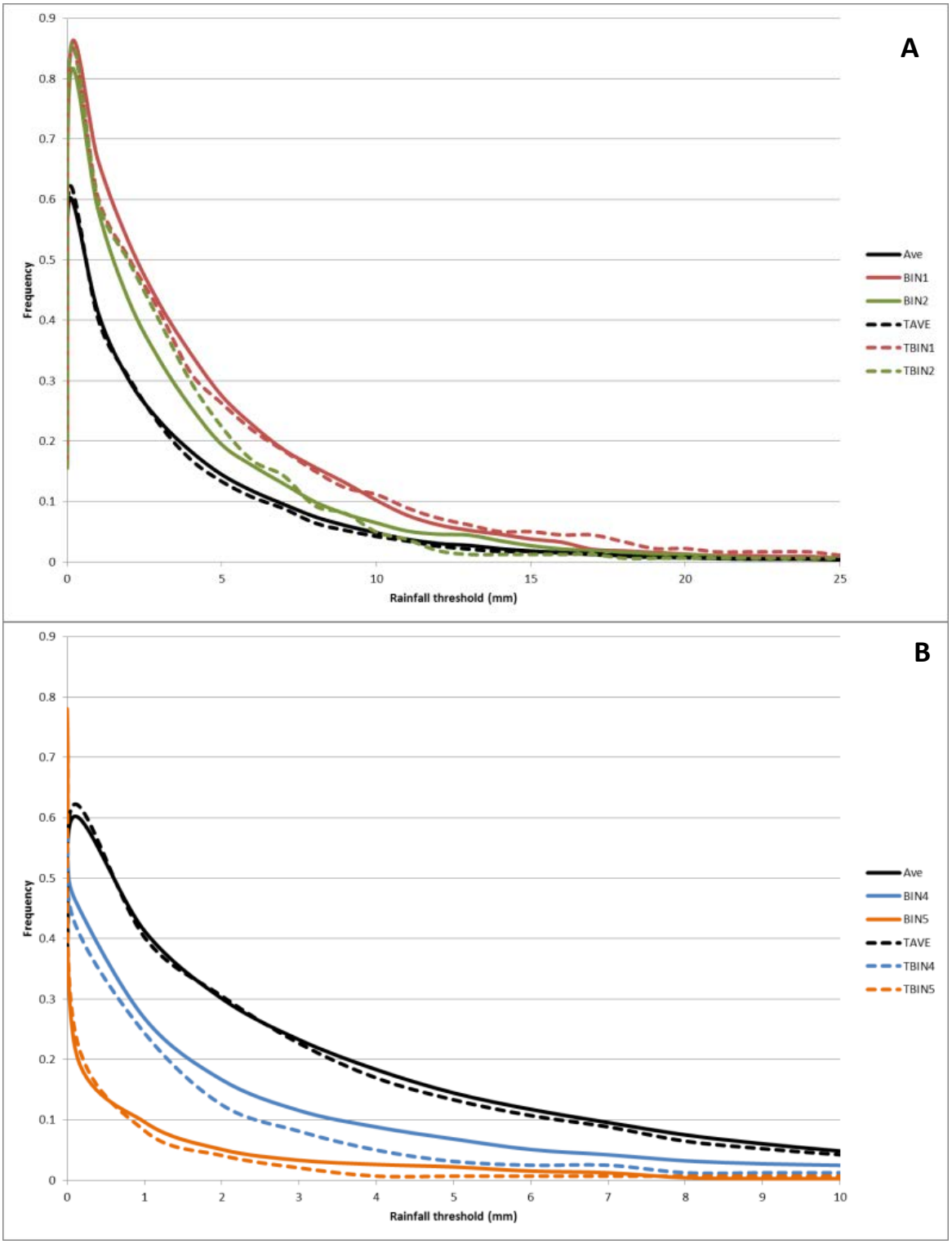


Figure 5-7: A) Frequency distribution for area-average daily rainfall of BIN1 and BIN 2 of the training period (solid lines) and the testing period (dashed line). B) Frequency distribution of the BIN5 and BIN 6 of the training period (solid lines) and the testing period (dashed lines). The solid black line indicates the climatological frequency distribution of all days in the training period and the dashed black line the frequency distribution of all days in the testing period.

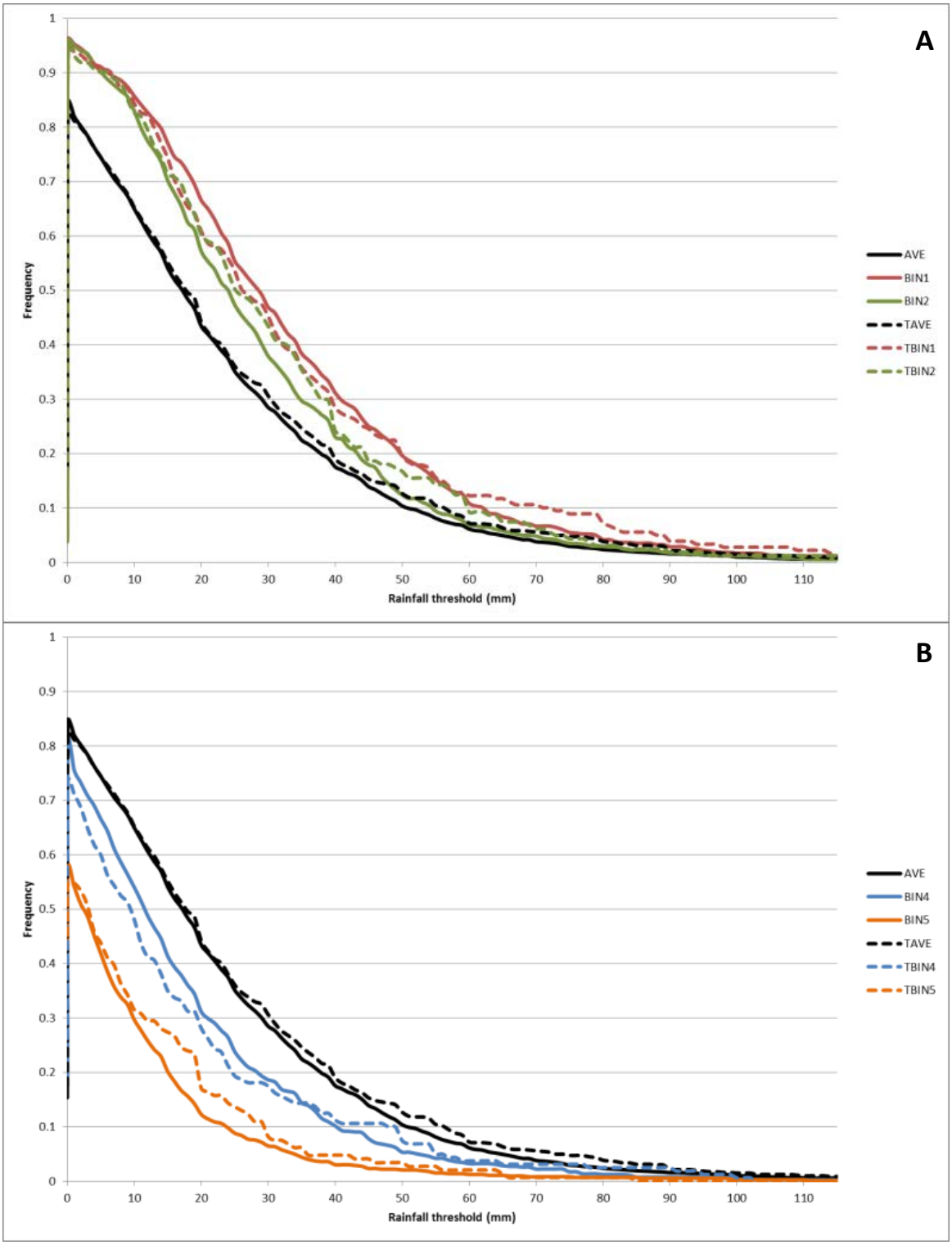


Figure 5-8: A) Frequency distribution for daily maximum rainfall of BIN1 and BIN 2 of the training period (solid lines) and the testing period (dashed line). B) Frequency distribution of the BIN5 and BIN 6 of the training period (solid lines) and the testing period (dashed lines). The solid black line indicates the climatological frequency distribution of all days in the training period and the dashed black line the frequency distribution of all days in the testing period.

Figs. 5-7 and 5-8 depict that there is a remarkable relationship between the frequency distributions of the validation and training periods for all the BINs and for both rainfall classes. For the rainfall threshold < 5 mm in the area-average daily rainfall class for BIN2 (green line in Fig. 5-7) the validation period has slightly higher frequencies than the training period. For BIN4 (blue in Fig. 5-7) the frequencies are slightly lower for the validation period than for the training period for all rainfall thresholds. Table 5-3 shows that for the area-average rainfall class the means and standard deviations for all the BINs for the training and validation period was very similar. BIN3 has the largest differences in the mean and standard deviation between the training and validation periods. However, the K-S test found that for any particular BIN in the area-average daily rainfall class the data samples of the training and validation period were from the same distribution.

The K-S test also found that for all the BINs in the maximum single station daily rainfall class the training and validation data samples had the same distribution. Fig. 5-8 shows that the frequencies of BIN1 (brown) for rainfall thresholds > 60 mm the validation period had a higher frequency than the training period. The mean maximum single station daily rainfall in BIN1 in the validation period was very close to the training period value. BIN2 (green Fig. 5-8) had higher frequencies for the validation period than the training period for rainfall thresholds < 60 mm. The mean value of the maximum single station daily rainfall in BIN2 for the validation period was 3 mm higher than for the training period and with the largest difference in the standard deviation for all the BINs. This indicates that for BIN2 the validation period data had larger variability than the training data. BIN4 (blue in Fig. 5-8) generally had lower frequencies in the lower rainfall thresholds in the training period than the validation period while BIN5 (orange) had higher frequencies in the training period than the validation period. The mean rainfall for BIN5 was slightly higher for the validation than training period.

5.5 Summary

A combination of sounding-derived and circulation criteria values were used to train a 20-node SOM for a 27-year period. The SOM was capable of distinguishing between nodes with above normal rainfall and nodes with below normal rainfall. The sounding-derived parameters provide information about the thermodynamical profile of the atmosphere over Gauteng while the circulation criteria capture the larger scale dynamical forcing. The SOM shows that above normal rainfall are associated with nodes where a combination of the sounding-derived parameters and circulation criteria are favourable. However, the SOM also allows for one of the two sets of input data into the SOM to be unfavourable when above normal rainfall occurs. This allows a better understanding of detailed processes which lead to heavy rainfall over Gauteng.

The SOM was used to predict rainfall frequencies over Gauteng by calculating to which node to assign a specific day in a 7 year validation period. This method showed that the binned rainfall frequencies in the validation period had very similar distributions that the training period. An operational forecaster may benefit from this method by using the output of the prediction system to aid in the decision making process to predict heavy rainfall over Gauteng. NCEP data has a horizontal resolution of 2.5 °, but using NWP data with a higher resolution could further improve this method.

5.6 Overview

This Chapter summarises and combines the other objectives of this thesis into a method to predict rainfall frequencies. Favourable circulation criteria (Chapter 3) and sounding-derived parameters (Chapter 4) associated with heavy rainfall are combined and used to train a SOM (Chapter 4). The daily rainfall values from Chapter 2 (Objective 1) are then used to first assign rainfall values to different nodes and are then used to validate the rainfall frequencies after the prediction was made.

5.7 References

BERMEJO M and ANCELL R (2009). Observed changes in extreme temperatures over Spain during 1957–2002 using Weather Types. *Revista de Climatologia*, **9**, 45-61.

DEBOECK G and KOHONEN T (1998). *Visual explorations in finance: with self-organizing maps* (Vol. 2). London: Springer-Verlag.

ESMAEILI MA (2011). Self-Organizing Map (SOM) in Wind Speed Forecasting: A New Approach in Computational Intelligence (CI) Forecasting Methods. Proceedings of the 7th Annual GRASP Symposium, Wichita State University, 69.

FABRIKANT S and SKUPIN A. (2005). Cognitively Plausible Information Visualization. (Eds.) Dykes J., MacEachren A., Kraak M.J. (Eds.). *Exploring Geovisualization*. Amsterdam: Elsevier. 667-690.

FRENKEL A, BENDIT E and KAPLAN S (2012). The linkage between the lifestyle of knowledge-workers and their intra-metropolitan residential choice: A clustering approach based on self-organizing maps. *Computers, Environment and Urban Systems*.

GRUNZ A, MEMMERT D, and PERL J (2012). Tactical pattern recognition in soccer games by means of special self-organizing maps. *Human Movement Science*, **31**(2), 334-343.

GUTIERREZ JM, CANO R, COFINO AS and SORDO C (2005). Analysis and downscaling multimodel seasonal forecasts in Peru using self-organizing maps. *Tellus A*, **57**, 435-447.

HSU YT, HUNG HF, YEH J and LIU M C (2010). Forecast of financial time series based on grey self-organizing maps. *International Journal of Innovative Computing, Information and Control,* **6**, 475-486.

ISMAIL S, SAMSUDIN R and SHABRi A (2010). River Flow Forecasting: a Hybrid Model of Self Organizing Maps and Least Square Support Vector Machine. *Hydrology and Earth System Sciences Discussions*, **7,** 8179-8212.

KHASHEI M and BIJARI H (2012). Exchange rate forecasting better with hybrid artificial neural networks models. *Journal of Mathematical and Computational Science*, ***1***(1).

KOHONEN T (1989). *Self-organization and associative memory*. Springer-Verlag, 312 pp.

LIU Y and WEISBERG RH (2011). A review of Self-Organizing Map applications in meteorology and oceanography. In *Self-Organizing Maps - Applications and Novel Algorithm Design*, Edited by: Mwasiagi JI 253–272. Rijeka, Croatia: InTech.

NAGI J, YAP KS, TIONG S K and AHMED SK (2008). Electrical Power Load Forecasting using Hybrid Self-Organizing Maps and Support Vector Machines. *Training*, **99**(1), 31.

NISHIYAMA K, ENDO S, JINNO K, BERTACCHI UVO C, OLSSON J, and BERNDTSSON R (2007). Identification of typical synoptic patterns causing heavy rainfall in the rainy season in Japan by a Self-Organizing Map. *Atmospheric Research*, **83**(2), 185-200.

RIVERA D, LILLO M, UVO CB, BILLIB M and ARUMÍ JL (2012). Forecasting monthly precipitation in Central Chile: a self-organizing map approach using filtered sea surface temperature. *Theoretical and Applied Climatology*, **107**(1), 1-13

SENJYU T, TAMAKI Y and UEZATO K (2000). Next day load curve forecasting using self organizing map. Proceedings International Conference on Power System Technology, **2**, 1113-1118.

SIMON G, LENDASSE A, COTTRELL M, FORT JC and VERLEYSEN M (2005). Time series forecasting: Obtaining long term trends with self-organizing maps. *Pattern Recognition Letters*, **26**(12), 1795-1808.

SKUPIN A and HAGELMAN R (2005). Visualizing Demographic Trajectories with Self-Organizing Maps. *GeoIn*formatica, **9**(2), 159-179

TADROSS MA, HEWITSON BC and USMAN MT (2005). The Interannual variability of the onset of the maize growing season over South Africa and Zimbabwe. *Journal of Climate,* **18**, 3356-3372.

TAMAYO P, SLONIM D, MESIROV J, ZHU Q, KITAREEWAN S, DMITROVSKY E and GOLUB TR (1999). Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proceedings of the National Academy of Sciences*, **96**(6), 2907-2912.

VAN DER VOORT M, DOUGHERTY M and WATSON S (1996). Combining Kohonen maps with ARIMA time series models to forecast traffic flow. *Transportation Research Part C: Emerging Technologies*, **4**(5), 307-318.

WILKS D (2006). *Statistical methods in the atmospheric sciences*. 2nd Edition, Elsevier Academic Press, California, United States of America, 661 pp.

6 Summary conclusions and recommendations

This chapter provides a general summary of the different topics dealt with in each of the chapters and highlights the major results. The results from this thesis are assessed and placed into the broader scientific context.

6.1 Study area

The Gauteng Province in South Africa was the focus of this research and was chosen as a study area as it offers an opportunity to investigate meteorological phenomena in a relatively similar area. Although Gauteng is an area defined by political boundaries there are several meteorological factors which makes it an ideal study area. Gauteng is situated on the interior plateau of South Africa with an average height of 1500 m a. m. s. l. It is a climatological relative uniform area although slightly cooler and with higher annual average rainfall in the elevated areas of the Witwatersrand in the south (Kruger, 2004). The entire province receives predominantly summer rainfall and from weather systems conducive to convective development (Gijben, 2012). There are approximately 75 rainfall stations scattered over Gauteng and many of these stations have rainfall records spanning more than 30 years. Upper air soundings done at Irene weather office situated in the central part of the province allows to investigate the thermodynamic profile of the atmosphere as the sounding can be considered to be representative of the atmosphere over the entire province (Brooks et al., 1994). There are data available from this sounding for more than 30 years. Rainfall over Gauteng provides water to two major catchments, the Crocodile catchment to the north and the Vaal catchment to the south. Both catchments have experienced major flooding events during the past decade (SAWDIS, 2012). Gauteng is the smallest of South Africa's nine provinces but with the highest population density and is responsible for more than one third of South Africa's GDP. Gauteng is therefore a very important contributor to the economy of South Africa and the disruption of infrastructure due to heavy rainfall and floods could result in large financial losses. There are also several vulnerable communities who live in shacks on vacated land close to rivers. Early warning of heavy rainfall could mitigate the impact that floods and flash floods have on these communities.

Considering that the aim of this study was to describe the atmospheric conditions associated with heavy rainfall over Gauteng five objectives were identified and are summarised here.

6.2 Characteristics of heavy rainfall over Gauteng: Objective 1

The first objective described the characteristics of heavy rainfall over Gauteng and formed the basis which allowed the other objectives to be dealt with. The results obtained provide information

about days, months and seasons with heavy rainfall and this was used to analyse the synoptic circulation patterns (Objective 2) and sounding-derived parameters (Objective 3).

A comprehensive daily rainfall climatology was constructed for the Gauteng Province. The climatology results were summarised in user-friendly reference tables. Three 24-hour heavy rainfall classes were defined considering the daily area-average rainfall. A 'significant rainfall event' (90^{th} percentile) was defined when the daily area-average rainfall exceeded 10 mm, a 'heavy rainfall event' (95^{th} percentile) when the daily area-average rainfall exceeded 15 mm and a 'very heavy rainfall event' (99^{th} percentile) when the daily area-average rainfall exceeded 25 mm. The 25 mm threshold of daily area-average rainfall over Gauteng is similar to the 'major rain event' defined by Houze et al. (1990). Over the interior of South Africa and in those regions which receive summer rainfall, the months with the highest rainfall totals are December and January months (in the east) and February months further west (Taljaard, 1996). The results obtained in this thesis indicate that January months has the highest monthly average rainfall as well as the highest number of heavy and very heavy rainfall days over Gauteng. The month with the second highest number of heavy and very heavy rainfall days is February followed by March and October. December has the second highest monthly average rainfall and the most days with rain. However, it is also the month with the lowest number of heavy and very heavy rainfall days.

Over the southern plains of the USA, Bradley and Smith (1994), defined a 'major rain event' at a single station when the 24-hr rainfall exceeded 125 mm and in Taiwan, Chen and Yu (1988) defined an 'extremely heavy rainfall event' when the rainfall exceeded 130 mm. Others like Zhang et al. (2001) defined heavy rainfall by examining the 90^{th} percentile or the 20-yr return values. This approach was also used in this thesis and a 'single significant event' (90^{th} percentile) is defined when rainfall at a single station exceeds 50 mm, a 'single heavy event' (95^{th} percentile) when the daily rainfall exceeds 75 mm and a 'single very heavy rainfall event' when the rainfall exceeds 115 mm (99^{th} percentile). This value for very heavy rainfall is slightly lower but in the same order of magnitude as the values found by Bradley and Smith (1994) and Chen and Yu (1988). Rainfall exceeding 115 mm at a single rainfall station in the Gauteng Province is very rare and does not occur every year. January months receive these events more than any other month but this happens on only approximately a third of the years. The central and northwestern parts of the province experience the most events where the rainfall at a single station surpasses 50, 75 and 115 mm. This is significant as the annual average rainfall in the southern and southeastern parts of the province is higher than over these areas. The highest 24-hour rainfall recorded at a single station during the 32-year period was 300 mm in December 2006.

Considering the seasonal rainfall, the 1995/96 summer rainfall season had the highest seasonal rainfall during the period under investigation followed by the 1999/2000 season. The 1995/96 season had above normal rainfall in both early and late summer but the 1999/2000 season was dry in early summer and very wet in late summer. Truly significant high seasonal rainfall is associated with above-normal rainfall in late summer.

December 2010 had the highest monthly December rainfall (220 mm) over Gauteng for the entire period starting in 1977. However this anomalously high monthly rainfall was a result of the rainfall on only three days during this month (14^{th}- 16^{th}) when 100 mm occurred over Gauteng.

6.3 Synoptic circulation associated with heavy rainfall over Gauteng: Objective 2

In Objective 2, the days with heavy rainfall which were identified in Objective 1 are used to investigate the monthly synoptic circulation patterns when heavy rainfall occurs. The synoptic circulation patterns influencing the weather over South Africa have been well documented by amongst others Taljaard (1995) and Tyson and Preston-White (2000). Heavy rainfall producing weather systems have also been well documented and mostly by conducting case studies of heavy rainfall events for example De Coning et al., 1998; Dyson and van Heerden, 2001; Hart et al., 2010 and Viviers and Chapman, 2008. In this thesis emphasis is placed on general monthly circulation patterns when heavy rainfall occurs over Gauteng rather than individual events. Furthermore the inter-seasonal variability of the synoptic circulation is explored.

The synoptic circulation over Gauteng changes from an atmosphere with significant baroclinic flow features to an atmosphere with tropical flow features by January months. The synoptic circulation patterns associated with heavy rainfall also changes significantly from early to late summer. In all months heavy rainfall is associated with a deeper than normal surface trough over the interior of Southern Africa. However, the average position of the surface trough when heavy rainfall occurs differs from month to month. It lies over the western interior in October months but over the Free State in January and February months. In early summer heavy rainfall over Gauteng is associated with a westerly trough west of South Africa at 700 and 500 hPa. In late summer this trough is located further south with a weak 700 hPa ridge over Gauteng. There is also an easterly inflow of moisture north of Gauteng in late summer when heavy rainfall occurs. At 500 hPa the geopotential height and temperature gradients become very slack in late summer with a weak tropical trough over the central interior.

Holton (1992) explained how cyclonic vorticity advection in the upper troposphere causes upward motion on the synoptic-scale. Furthermore if horizontal wind convergence occurs through a

deep layer in the atmosphere upward motion will also result (continuity equation). In turn upward motion results in cloud formation and rainfall if adequate surface moisture is available. The results in this thesis show that whenever heavy rainfall transpires over Gauteng, horizontal wind convergence occurs from the surface up to at least 500 hPa over the province. In early summer the magnitude of the convergence is larger than in late summer when the convergence occurs as high as up to 400 hPa. Heavy rainfall is also associated with cyclonic vorticity advection throughout the entire troposphere but larger in magnitude at 300 hPa.

6.4 Sounding-derived parameters associated with heavy rainfall over Gauteng: Objective 3

Sounding data from the Irene weather office were used to first construct a baseline climatology of basic and derived parameters for a 35-year period over Gauteng. These parameters were then investigated on days with heavy rainfall and days which remain dry. Twelve parameters were identified which helps to discriminate between days with heavy rainfall and days which remain dry.

More than 30-years of Irene weather office upper air data were obtained and put through a vigorous Quality Control procedure similar to the process described by Durre et al. (2006). These data were then first used to create a monthly climatology of the vertical profile of basic variables as well as sounding-derived parameters over Gauteng. A striking feature of the seasonal variation of these parameters is how the atmosphere changes from one with a very distinct extra tropical nature in October to one which is clearly tropical in February months. This is most evidently demonstrated by the increase in temperatures throughout the troposphere and the decrease in wind strength and vertical wind shear during the season. In February months the average wind speed is less than 5 knots from the surface up to 500 hPa. It is also interesting to note how the temperatures increase more at lower pressures than at higher pressures closer to the ground. This in turn influences the conditional instability of the atmosphere and parameters which employ temperature lapse rates as an ingredient becomes less useful especially in late summer. In late summer using parameters which identify the convective instability in the atmosphere are more useful such as the change of equivalent potential temperature with height ($\Delta\Theta_e$) between the surface and 500 hPa. Convective variables such as CAPE and the Showalter Index reach maximum favourable values in December months.

The parameters which could be used to identify heavy rainfall over Gauteng were consequently identified. There is a clear seasonal variation in the values of most sounding

parameters associated with heavy rainfall and single threshold values should not be used to identify heavy rainfall. As an example, the average precipitable water value associated with heavy rainfall in October months is 21 mm but 26 mm in January and February months.

Twelve sounding-derived parameters were identified as being able to distinguish heavy rainfall days from days without rainfall. They are

Precipitable water (PW)

Mean layer dew point temperature Td_{100}

Convective Available Potential Energy (CAPE)

Average temperatures in the 500-300 hPa layer (T_{53})

Vertical temperature difference from 700-500 hPa (TD_{75})

Warm cloud depth (WCD)

Elevated K-Index (EKI)

Meridionial component of the wind in the 800-600 hPa layer (WV_{86})

Average wind speed in the 600-400 hPa layer (W_{64})

Magnitude of wind shear from the surface to 400 hPa (MWS_{54})

$\Delta\Theta_e$ from the surface to 500 hPa

Average Θ_e in the troposphere

The parameters associated with moisture provide the best guidance in distinguishing days with heavy rainfall over Gauteng. PW values are consistently significantly higher (according to the Mann-Whitney test) on heavy rainfall days than on No Rain days and CLIM days for every month. Dupilka and Reuter (2006) stated that in Alberta, Canada in summer (May-September) the average PW water value is about 15 mm and values above 20 mm would have a higher association with severe thunderstorms. Over Gauteng the average PW value is 15 mm in October months but closer to 23 mm in January months. Heavy rainfall are associated with values ranging from 21 mm (October) to 26 mm in January. The January heavy rainfall PW value over Gauteng is similar to what Harnack et al. (1998) found to be associated with heavy rainfall in Utah, in the USA in summer. The mean layer dew point temperature (Td_{100}) on heavy rainfall days is also higher than on No Rain days for every month but in late summer the heavy rainfall values are not significantly higher than CLIM values. When heavy rainfall occurs the average Td_{100} value is consistently higher than 12 °C irrespective of the month but reaches values of more than 16 °C in January and February months. In Utah (USA) the mean surface dew point temperature associated with heavy rainfall was 10.8 °C (Harnack et al.,

1998) while the average value was only 6.2 °C. In Gauteng the lowest long-term average (CLIM) Td_{100} value occurs in October months at 8 °C.

The average 500-300 hPa temperatures (T_{53}) are generally higher on heavy rainfall days than on No Rain and CLIM days but these differences are mostly not significant. It was found that the only pressure levels between which the temperature difference (TD) had larger values on heavy rainfall days than on dry days were in the 700-500 hPa layer. Heavy rainfall days had significantly higher TD_{75} values than No Rain and CLIM days in early summer but not in January and February months. Harnack et al. (1998) also considered the temperature differences between the 700-500 hPa levels and found values of 18.5 °C to be associated with heavy rainfall in Utah. Heavy rainfall over Gauteng have TD_{75} values larger than 18 °C only in October and November months with values less than 16 °C in February.

Average monthly CAPE values are significantly larger on heavy rainfall days than No rain and CLIM days for all months. It was shown that large CAPE values in early summer are associated with large TD_{75} values but in late summer large CAPE values are associated with large negative values of $\Delta\Theta_e$. This indicates how atmospheric instability changes from conditional unstable in early summer to convective unstable in late summer. However, no direct relationship between CAPE values and heavy rainfall exists, as heavy rainfall may also develop in low CAPE environments (Craven and Brooks, 2004; Clark, 2009). The $\Delta\Theta_e$ values are consistently larger negative on heavy rainfall than on No Rain and CLIM days for all summer months. On heavy rainfall days the $\Delta\Theta_e$ are larger than -20 K in all months except October. These values are comparable to $\Delta\Theta_e$ values found by Cohen et al. (2007) in the area east of the Rocky Mountains when mesoscale convective systems develop. The average tropospheric equivalent potential temperature ($\Theta_{e\ ave}$) values are larger when heavy rainfall occurs than on No Rain days but in late summer there is no significant difference between these values and CLIM values. Of all the variables which employ lapse rate as an ingredient, the Elevated K-Index performed the best in distinguishing between heavy rainfall and No Rain and CLIM days and values of 10 and more are favourable for heavy rainfall.

The warm cloud depth (WCD) is consistently greater on days with heavy rainfall than on No rain days. In early summer there is a significant difference in the WCD between CLIM and heavy rainfall days but these differences diminish in late summer. It was shown how in late summer larger WCD values occur in conjunction with weaker winds in the 600-400 hPa layer. This is indicative of slow moving weather systems where coalescence plays a dominant role in the rainfall efficiency (Doswell et al., 1996). Heavy rainfall over Gauteng is often associated with winds with a northerly component in the 800-600 hPa layer. This is indicative of moisture being advected into southern

Africa from the Mozambique Channel or when the ITCZ moves southward to about 20 °S in late summer (Taljaard, 1995). In early summer heavy rainfall is associated with large wind shear values but in late summer the wind shear values are very small and there is no significant difference between wind shear values on heavy rainfall days, No Rain and CLIM days.

6.5 A heavy rainfall sounding climatology using Self-organizing maps: Objective 4

Self-organizing maps (SOMs) have been used successfully by several authors to create synoptic climatologies (see Hewitson and Crane, 2002; Tennant and Hewiston, 2002 and Alexander et al., 2010). In objective 4, SOMs were utilized to create a sounding climatology over Gauteng. The SOM was trained with data which contained information about temperature, dew point temperature and the meridionial and zonal components of the wind on 12 pressure levels as well as the 12 sounding-derived parameters. A 20-node SOM was trained for all area-averaged heavy rainfall (AHR) and single-station heavy rainfall (SHR) summer days for the period of 1977 to 2012. The SOM provides detailed information about the vertical profile of the basic variables as well as the sounding-derived parameters on heavy rainfall days. The SOM was capable of distinguishing the favourable atmospheric conditions for heavy rainfall in early summer from those in late summer. It has the added advantage of providing information about how the different variables combine when heavy rainfall occurs. A SOM of the sounding-derived parameters provides guidance on critical values for heavy rainfall during the progression of the summer rainfall season.

6.6 Self-organizing maps as a predictive tool: Objective 5

Although SOMs have been used widely in meteorology since about the year 2000, limited work has been done to use SOMs to forecast meteorological phenomena. Objective 5 exploits the capability of SOMs to convert complex non-linear features into simple two-dimensional relationships. These relationships are then used to predict rainfall frequencies. The results from Objective 2 (circulation criteria associated with heavy rainfall) and Objective 3 (sounding-derived parameters) are combined to train the SOM and then to use the basic relationships to predict rainfall. In this Objective the same approached was used than that of Nishiyama et al. (2007) but the data used to train the SOM differs.

Twelve sounding-derived parameters were combined with 20 circulation criteria values and this was used to train a SOM for all available summer months for a 27-year period from 1979 to 2005. The circulation criteria are horizontal wind divergence values ar four grid points surrounding Gauteng at 850, 700, 500 and 300 hPa as well as vorticity advection values at 300 hPa. The 20-node SOM was capable of distinguishing a combination of sounding-derived parameters and circulation

criteria associated with heavy rainfall. The rainfall of 20-nodes was binned together into 5 rainfall classes with BIN1 containing the four nodes with the highest average rainfall and BIN5 the nodes with the lowest average rainfall. Frequency distributions of each of these 5 BINs were consequently constructed and it was shown how BIN1 and BIN2 are associated with above normal rainfall and BIN4 and BIN5 with below normal rainfall. The SOM also provided a range of possible favourable conditions associated with heavy rainfall and has the added advantage of still identifying favourable heavy rainfall conditions even if either the sounding-derived parameters or circulation criteria are unfavourable.

The same data arrays were created for an 8-year validation period and it was determined to which node a specific day should be mapped. The rainfall of the nodes was clustered together in the same way as before and the frequency distributions of the different BINs were compared to those in the training period. For all BINs there was a very close relationship between the frequency distribution of the testing and training period. This indicated that applying a SOM in such a way has potential of being useful as a predictive tool.

6.7 Assessing the scientific contribution of this study

This study investigated the character of heavy rainfall over Gauteng and uses this information to construct a sounding climatology associated with heavy rainfall. It was mentioned several times in this thesis that Gauteng lies approximately 1500 m a. m. s. l. This height above sea level has a significant influence on especially the amount of water vapour which is available in the atmosphere over Gauteng. The diminished values of water vapour will therefore necessarily impact on the amount of rainfall which will fall. The first scientific contribution of this research is to provide information about the characteristics of heavy rainfall over an elevated area such as Gauteng. Very heavy rainfall events over Gauteng were compared to heavy rainfall elsewhere in the world and it was generally found that extreme rainfall values over Gauteng is lower. Detailed monthly information is also provided about area-average daily rainfall and single station rainfall over Gauteng. The average early summer rainfall is less than the average late summer rainfall and it is shown how the atmosphere changes from extra-tropical to tropical during this period. Most heavy rainfall events occur when the atmosphere has distinct tropical characteristics.

The elevated plateau over the interior of South Africa also has a significant influence on the thermodynamic profile of the atmosphere. The threshold values and ingredients identified over other parts of the globe are therefore not directly applicable to Gauteng. No such investigation has yet been undertaken in South Africa and this research provides ingredients associated with heavy rainfall over Gauteng. Twelve heavy rainfall sounding-derived parameters were identified and critical

values specific to Gauteng are proposed. A localised parameter which searches for winds from the north in the 800-600 hPa layer when heavy rainfall occurs is introduced. It was also shown how some commonly used parameters such as the *K*-Index is not very useful to identify atmospheric conditions associated with heavy rainfall over Gauteng, especially in late summer. This study is conducted over more than 30-years and therefore also serves as an instrument to investigate the intra-seasonal variability of the thermodynamical profile of the atmosphere during heavy rainfall. It was shown how conditions associated with heavy rainfall change from conditionally unstable in early summer to convectively unstable in late summer.

SOMs have been used increasingly in meteorological applications during the past decade and have been used successfully to create synoptic climatologies. This thesis contributes scientifically by applying SOMs to a combination of the vertical profile of basic atmospheric variables and sounding-derived parameters. This is one of the first SOM applications of its kind and provides significant meteorological information. The SOMs provide detailed climatological information of the relationship between the different sounding parameters during different months in the summer rainfall season. The detailed differences and relationship between parameters were not readily available from the monthly climatology conducted initially.

The final scientific contribution is to propose a method to use SOMs to predict above and below normal rainfall over Gauteng. This a very robust method and the frequency distributions of the training period are compared to frequency distributions of a test period. These frequency distributions of the training and test period have remarkable similar distributions and create the expectation that SOMs may be used to predict daily rainfall.

6.8 Recommendations for future research

It is recommended that two specific types of errors which occur in the sounding data be further investigated. They are the anomalously low dew points in the middle troposphere and atmospheric lapse rates larger than the dry adiabatic lapse rate which sometimes occurs above the cloud tops. These, probably, are instrument related and not true representation of atmospheric conditions.

Research to create similar climatologies for severe storms and also for anomalously high lightning strikes over Gauteng should be supported.

The methodology to use SOMs to predict rainfall frequencies warrants further research. A good starting point could be to use the rainfall frequencies of the different BINs, determined in this research, to create probabilistic daily rainfall forecast for the Gauteng Province.

A prediction technique, based on the SOMs, (but not necessarily with the same variables), should be researched and applied to seasonal rainfall forecasts where above normal, normal and below normal probabilistic forecast is the norm.

6.9 References

ALEXANDER L V, UOTILA P, NICHOLLS N and LYNCH A (2010). A new daily pressure dataset for Australia and its application to the assessment of changes in synoptic patterns during the last century. *Journal of Climate,* **23**, 1111-1126.

BRADLEY AA and SMITH JA (1994). The hydrometeorological environment of extreme rainstorms in the southern plains of the United States. *Journal of Applied Meteorology.* **33,** 1418-1431.

BROOKS HE, DOSWELL CA III and COOPER J (1994). On the environments of tornadic and nontornadic mesocyclones. *Weather and Forecasting*, **9**, 606–618.

CHEN GT and YU C (1988). Study of level jet and extremely heavy rainfall over northern Taiwan in the Mei-Yu season. *Monthly Weather Review,* **116,** 884-891.

CLARK MR (2006). The southern England tornadoes of 30 December 2006: Case study of a tornadic storm in a low CAPE, high shear environment. *Atmospheric Research*, **93**, 50-65

COHEN AE, CONIGLIO MC, CORFIDI SF and CORFIDI SJ (2007). Discrimination of mesoscale convective system environments using sounding observations. *Weather and Forecasting*, **22**, 1045–1062.

CRAVEN JP AND BROOKS HE (2004). Baseline climatology of sounding-derived parameters associated with deep moist convection. *National Weather Digest,* **28**, 13-24.

DE CONING E, FORBES GS and POOLMAN EP (1998). Heavy precipitation and flooding on 12-14 February 1996 over the summer rainfall regions of South Africa: Synoptic and isentropic analyses. *National Weather Digest,* **22,** 25-36.

DOSWELL CA III, BROOKS HE and MADDOX RA (1996). Flash flood forecasting: An ingredients-based methodology. *Weather and Forecasting*, **11**, 560-580.

DUPILKA ML and REUTER GW (2006). Forecasting Tornadic Thunderstorm Potential in Alberta Using Environmental Sounding Data. Part II: Helicity, Precipitable Water, and Storm Convergence. *Weather and Forecasting,* **21,** 336–346.

DURRE I, VOSE RS and WUERTZ DB (2006). Overview of the Integrated Global Radiosonde Archive. *Journal of Climate*, **19**(1), 53-68.

DYSON LL and VAN HEERDEN J (2001). The heavy rainfall and floods over the northeastern interior of South Africa during February 2000. *South African Journal of Science,* **97**, 80-86.

GIJBEN M (2012). The lightning climatology of South Africa. *South African Journal of Science*; **108**, (3/4). Art. #740, 10 pages

HARNACK RP, JENSEN DT CERMAK III JR (1998). Investigation of upper-air conditions occurring with heavy summer rain in Utah. *International Journal of Climatology*, **8**, 701-732.

HART NCG, REASON CJC and FAUCHEREAU N (2010). Tropical–Extratropical Interactions over Southern Africa: Three Cases of Heavy Summer Season Rainfall. *Monthly Weather Review,* **138**, 2608–2623.

HEWITSON BC and CRANE RG (2002). Self-organizing maps: applications to synoptic climatology. *Climate Research,* **22**, 13–26.

HOLTON JR (1992). *An Introducion to Dynamic Meteorology*. Academic Press. 511 pp.

HOUZE RA, SMULL BF and DODGE P (1990). Mesoscale organization of springtime rainstorms in Oklahoma. *Monthly Weather Review,* **118,** 613-654.

KRUGER AC (2004). *Climate of South Africa. Climate Regions. WS45*. South African Weather Service, Pretoria, South Africa.

NISHIYAMA K, ENDO S, JINNO K, BERTACCHI UVO C, OLSSON J, and BERNDTSSON R (2007). Identification of typical synoptic patterns causing heavy rainfall in the rainy season in Japan by a Self-Organizing Map. *Atmospheric research*, **83**(2), 185-200.

SAWDIS (South African Weather and Disaster Information Service) (2012). URL: http://saweatherobserver.blogspot.com/. (Accessed on 23 Augustus 2012)

TALJAARD JJ (1995). Atmospheric circulation systems, Synoptic climatology and weather Phenomena of South Africa. Part 3 Synoptic Climatology in South Africa in January and July. South African Weather Bureau Technical Paper (29). South African Weather Service, Pretoria, South Africa.

TALJAARD JJ (1996). Atmospheric circulation systems, Synoptic climatology and weather Phenomena of South Africa. Part 6 Rainfall.in South Africa. South African Weather Bureau Technical Paper (32). South African Weather Service, Pretoria, South Africa.

TENNANT W and HEWITSON BC (2002). Intra-seasonal rainfall characteristics and their importance to the seasonal prediction problem. *International Journal of Climatology,* **22**, 1033–1048.

TYSON P and PRESTON-WHITE RA (2000). *The weather and climate of South Africa.* Oxford University Press. 408 pp.

VIVIERS G and CHAPMAN V (2008). Severe thunderstorm at Johannesburg International Airport (03 December 2004). URL: http://www.eumetsat.int/Home/index.htm (Accessed on 14 November 2008).

ZHANG X, HOGG WD and MEKIS E (2001). Spatial and temporal characteristics of heavy precipitation events over Canada. *Journal of Climate,* **14,** 1923-1936.