A heavy rainfall sounding climatology over Gauteng, South Africa, using self-organising maps

by

Liesl L Dyson

Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria, South Africa

Corresponding author's details:

Dr Liesl Dyson
Department of Geography, Geoinformatics and Meteorology
Geography Building
University of Pretoria
Lynnwood Road
Hatfield
South Africa
0028
Email: liesl.dyson@up.ac.za
Telephone: +27 12 420 2469

Abstract

The daily weather at a particular place is largely influenced by the synoptic circulation and thermodynamic profile of the atmosphere. Heavy rainfall occurs from a particular subset of synoptic and thermodynamic states. Baseline climatologies provide objective information on heavy rainfall-producing circulation patterns and thermodynamic variables. This is how climatologically large or extreme values associated with heavy rainfall are identified. The aim of this research is to provide a heavy rainfall sounding climatology in austral summer over Gauteng, South Africa, using self-organising maps (SOMs). The results show that the SOM captures the intra-seasonal variability of heavy rainfall soundings by clearly distinguishing between the atmospheric conditions on early summer (October to December) and late summer (January to March) heavy rainfall days. Conditions associated with heavy early summer rainfall are large vertical wind shear and conditional instability, while the atmosphere is drier and cooler than when heavy rainfall occurs in late summer. Late summer heavy rainfall conditions are higher convective instability and small vertical wind shear values. The SOM climatology shows that some heavy rainfall days occur in both early and late summer when large-scale synoptic weather systems cause strong near-surface moisture flux and large values of wind shear. On these days, both the conditional and convective instability of the atmosphere are low and heavy rainfall results from the strong synoptic forcing. In contrast, heavy rainfall also occurs on days when synoptic circulation is not very favourable and the air is relatively dry, but the atmosphere is unstable with warm surface conditions and heavy

rainfall develops from local favourable conditions. The SOM climatology provides guidelines to critical values of sounding-derived parameters for all these scenarios.

Key words: sounding climatology, self-organising maps, heavy rainfall, South Africa

1. Introduction

There are several examples of instances where sounding-derived parameters are associated with severe weather (Craven and Brooks 2004; Doswell and Schultz 2006; Dimitrova et al. 2009) and where those sounding parameters with climatological extreme values are isolated when severe weather occurs (Dupilka and Reuter 2006a, 2006b; Groenemeijer and Van Delden 2007; Covadonga et al. 2010). Harnack et al. (1998) and Dyson et al. (2014) identified sounding parameters that distinguish between days with heavy rainfall and the climatological mean. Over the province of Gauteng in South Africa (the focus area of this research), Dyson et al. (2014) (henceforth DYSON14) stressed how the character of the atmosphere changes during austral summer from one with extra-tropical characteristics in early summer (October to December) to one with distinctly tropical characteristics in late summer (January to March). They found that heavy rainfall over Gauteng in October and November is associated with sounding parameters that exploit the conditionally unstable character (Doswell et al. 1985) of the atmosphere, such as large values of wind shear and convective available potential energy (CAPE) (see Table 1). In late summer, the atmosphere displays distinctly tropical characteristics and those parameters that deal with convective instability (Schultz et al. 2000), such as the equivalent potential temperature lapse rate, distinguish heavy rainfall from the climatological mean. In early summer, heavy rainfall over Gauteng is associated with atmospheric conditions that are conducive to the development of severe storms (Johns and Doswell 1992). In late summer, slow-moving precipitation-efficient storms (Doswell et al. 1996) cause heavy rainfall. DYSON14 found that those parameters that provide information about the moisture content of the atmosphere, such as the mean layer dew point temperature (Td_{100}) and precipitable water (PW) distinguish between the climatological mean and heavy rainfall days during all months. The only other variables (of those investigated) that were capable of making this distinction during all six summer months were the average meridionial wind in the 800–600 hPa layer ($V_{800,600}$), the mean layer equivalent potential temperature ($\Theta_{e\ 100}$), the Showalter Index (SI), the K-Index and a K-index adapted for the interior plateau of South Africa, namely the Elevated K-Index (EKI). DYSON14 stressed how typical values of parameters associated with heavy rainfall vary during the season. This variation needs to be considered when identifying climatological large values.

A question that remains unanswered by DYSON14 is what the relationship between the sounding-derived parameters is when heavy rainfall occurs. For instance, when heavy rainfall occurs in a large CAPE environment, what is the value of the vertical temperature lapse rate and low-level dew point temperature and does this relationship remain the same throughout the season?

The aim of this paper is to use self-organising maps (SOMs) to identify the relationship between the soundings and the parameters derived from them and to investigate the intra-seasonal variability of the thermodynamic profile of the atmosphere during heavy rainfall.

2. Heavy rainfall-producing weather systems over Gauteng

The focus of this paper is to provide a thermodynamic profile of the atmosphere over Gauteng when heavy rainfall occurs. However, the temperature and moisture profile of the atmosphere, as well as the associated sounding-derived parameters, are principally caused by the larger-scale synoptic flow. There is a wealth of information detailing synoptic circulation patterns over southern Africa (Taljaard 1985; Taljaard 1996; Tyson and Preston-White 2000; Singleton and Reason 2007; Hart et al. 2010) and describing those synoptic systems that cause heavy rainfall (De Coning et al. 1998; Dyson and Van Heerden 2002; Malherbe et al. 2012).

Taljaard (1996) stressed the importance of abundant moisture in the production of rainfall and explained how a high pressure system centred close to Maputo advects moisture at surface levels into the interior of South Africa in the summer months. The Indian Ocean is the primary source of moisture in summer, but the southward penetration of air from the Inter-tropical Convergence Zone (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, which is 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 that produce rainfall in summer in South Africa were identified by Taljaard (1996) and Tyson and Preston-White (2000). These weather systems are the westerly trough or tropical temperate trough (TTT), the southward-extending tropical low or V-shaped trough, cut-off lows (COLs), tropical cyclones (TCs) and the ridging high.

The TTT is a wave in the west wind regime that connects tropical convective systems to mid-latitude weather systems and is clearly visible on satellite imagery as elongated cloud bands that extend south-eastward (Hart et al. 2010). The TTT is part of an important process that causes humid tropical air to flow

from the ITCZ into South Africa. Moist tropical air has a greater potential for convective instability. The TTT may cause heavy rainfall over the interior of South Africa, as Hart et al. (2010) illustrate. They stress the importance of local convection in the production of the heavy rainfall from these systems. Furthermore, De Coning et al. (1998) found that a TTT was partially responsible for the very heavy rainfall over the central interior of South Africa during February 1996. Hart et al. (2010) explained how the rainfall produced by TTTs is preceded by a north-easterly moisture transport over Botswana, which is set up by the Angola low. The Angola low is defined as a heat low situated over northern Namibia/southern Angola in summer, which reaches maximum strength in January or February (Reason et al. 2006).

The southward-extending tropical low or V-shaped trough of Taljaard (1996) is also referred to as a continental tropical low pressure system (Dyson and Van Heerden 2002). This system is characterised by a low pressure system that stands upright from the surface to 400 hPa and 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 low temperatures in the surface to the 700 hPa layer. The surface dew point temperatures associated with these systems are in the order of 18–20 °C, with PW values of more than 20 mm. Some noteworthy and very heavy rainfall events over South Africa 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 north-eastern South Africa in February 2000 (Dyson and Van Heerden 2001). Part of the reason for the flooding 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 three consecutive days (19 to 22 February). The convective cells that develop in association with tropical low pressure systems are semi-stationary and are associated with little or no vertical wind shear.

COLs are truly significant extra-tropical weather systems and are responsible for widespread rainfall over South Africa and 20% of COLs produce heavy rainfall, which may cause floods (Taljaard 1985). Most COLs have life spans of longer than two days, but they may influence the rainfall over South Africa for up to five days (Singleton and Reason 2007). On synoptic maps, these systems are characterised by a cold cored low pressure in the middle troposphere that is cut off from the general westerly circulation, as well as closed cyclonic circulation throughout the troposphere. These weather systems are clearly baroclinic and heavy rainfall from COLs occur when they are in the developmental stage. The atmosphere is baroclinic when the atmospheric density is a function of both temperature and pressure and, therefore, a thermal wind exists (Holton 1992). A COL is 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).

On average, there are 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 influence of TCs 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 south-easterly flow. However, when these systems make landfall, torrential rain may fall, especially from the coastal belt up to the escarpment. On 17 January 2012, TC Dando made landfall south of Maputo and moved in a north-westerly direction over the subcontinent, while weakening (Meteo France 2012). Daily rainfall in excess of 200 mm occurred over the escarpment of Mpumalanga on 17 and 18 January 2012. Gauteng received only light rainfall during this event. Conversely, TC Eline made landfall north of Beira on the Mozambican coast on 22 February 2000 and tracked slowly westward over Zimbabwe and Botswana during the next few days. More than 100 mm of rain was recorded at some stations over Gauteng on 22 February 2000 and about 50 mm on 23 and 24 February 2000 (Dyson and Van Heerden 2001).

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 the interior plateau of South Africa. These systems are not directly responsible for heavy rainfall, but 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).

3. Self-organising maps

Since Kohonen (1989) first proposed SOMs, they have been widely used in several disciplines to analyse, cluster and visualise data. For instance, SOMs have been used in the biological sciences (Tamayo et al. 1999), in the financial sector (Deboeck and Kohonen 1998), in the human sciences (Frenkel et al. 2012) and in the sport sciences (Grunz et al. 2012). SOMs have been used to forecast 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).

SOMs have also been widely used in meteorological applications since the start of the 21st century. There have been significant contributions to synoptic climatologies and the synoptic patterns of atmospheric circulation (Hewitson and Crane 2002; Tennant and Hewitson 2002; Liu and Weisberg 2011; Engelbrecht

et al. 2014). Climate variability and extreme events have been analysed with SOMs (Liu and Weisberg 2011) and trend analyses performed (Bermejo and Ancell 2009; Lennard and Hegerl 2014).

Methods to downscale both seasonal and short-range weather forecasts were developed (Gutierrez et al. 2005). Tadross et al. (2005) used SOMs to explain the differences in synoptic circulation associated with the early and late onset of the maize-growing season over southern Africa, and Van Shalkwyk and Dyson (2013) associated different synoptic circulation features with the occurrence of fog at Cape Town International Airport. So far, there has been limited application of SOMs to sounding data, with the exception of Jensen et al. (2012), who used ozonesondes to discuss the vertical profile of ozone at Ascension Island in the South Atlantic Ocean, as well as KwaZulu-Natal in South Africa. Among other things, Seefeldt and Cassano (2008) investigated the vertical profiles of winds and potential temperatures over the greater Ross Ice Shelf region using SOMs. In the research presented in this paper, SOMs will be used on sounding data to create a heavy rainfall sounding climatology over Gauteng.

Lennard and Hegerl (2014) explain that SOMs are ideal for synoptic classification, as the data are treated as a continuum and do not rely on correlation, cluster or eigenfunction analysis. Therefore, it does not assume that weather states can be divided into categories with distinct boundaries. Hewitson and Crane (2002) point out that one of the advantages of SOMs in the development of meteorological climatologies is that they span the entire multidimensional data space and attempt to search for a number of nodes within a data space so that the distribution of the nodes characterises the observed distribution. In this way, local surface variables can be quantitatively related to the atmosphere by retaining the non-linear characteristics of the data. When applying this technique to circulation patterns or, as in this case, sounding data, the nodes represent physical spatial patterns unlike principal component analysis (PCA), which produces patterns of variance, rather than direct physical states of the atmosphere (Reusch et al. 2005). Hewitson and Crane (2002) also point out that the SOM technique makes no assumption about the distribution of the data. SOMs therefore provide a mechanism for envisioning an assortment of atmospheric states after dominant modes in a dataset are identified (Tennant and Hewitson 2002). The SOM technique identifies representative nodes that span the data space so that individual data elements may be associated with a specific node.

Kohonen (2001) provides a detailed description of the artificial neural network of SOMs. In this paper, the SOM was created using software that is freely available from http://www.cis.hut.fi/research/som-research, referred to as SOM_PAK (Kohen et al. 1996). The first step is to choose a number of nodes where random numbers are used to initialise the reference vectors. The number of nodes is chosen subjectively and depends on the degree of generalisation required (Lennard and Hegerl 2014). The second stage is the map training process where the SOM software calculates the similarity (weight) between each

data record and each of the node reference vectors. The reference vector of the node that best represents the data record (having a minimum Euclidian distance between the data vector and node vector) is consequently modified by a user-defined factor, or learning rate. The learning rate determines how fast the weights move towards the data points. The reference vector that best represents the data vector (winning node) is updated continuously during the training process and the weights of the nodes surrounding the winning node are adjusted so that each vector converges to the input pattern (Lennard and Hegerl 2014). The radius of influence, which also needs to be set during the map training, determines how many nodes surrounding the winning node are updated during each step in the training process.

The SOM is trained twice. During the first training process, random input vectors are used, which produce a first set of nodes. During the second phase of training, the winning nodes identified during the first training process are used as input vectors. This second training process allows final convergence of the map, thereby allowing a simplification into a small number of archetypes (Hewitson and Crane 2002). The SOM places similar nodes neighbouring each other and very different nodes far apart in the SOM space. The nodes represent a non-linear distribution of overlapping, non-discreet, thermodynamic profiles. Each node in a SOM represents a group of similar sounding profiles that were present in the original dataset. The study of the frequency of occurrence is made possible as any day's actual data can be distinctively associated with a closest-matching node in the SOM (Reusch et al. 2005).

The user-defined settings used in the SOM_PAK software are rectangular topology, bubble neighbourhood function type, 1 000 000 training length during the first and second training phases, 0.1 learning rate parameter during the first training phase and 0.01 during the second phase, as well as a radius of 1 during the initial and second training phases.

4. Data and methodology

4.1 Study area

This research focuses on daily heavy rainfall events during austral summer over Gauteng, South Africa (Fig. 1). Gauteng falls in the summer rainfall area of South Africa and receives 600–700 mm of rain per annum. Gauteng is cool for its latitude (average annual maximum temperatures between 22 °C and 25 °C) because of its height above mean sea level (AMSL) of between 1 350 m and 1800 m (Kruger 2004). There are approximately 50 rainfall stations across Gauteng and daily upper air soundings are performed at the Irene weather office of the South African Weather Service (SAWS), which is located in the centre of the province (Fig. 1).

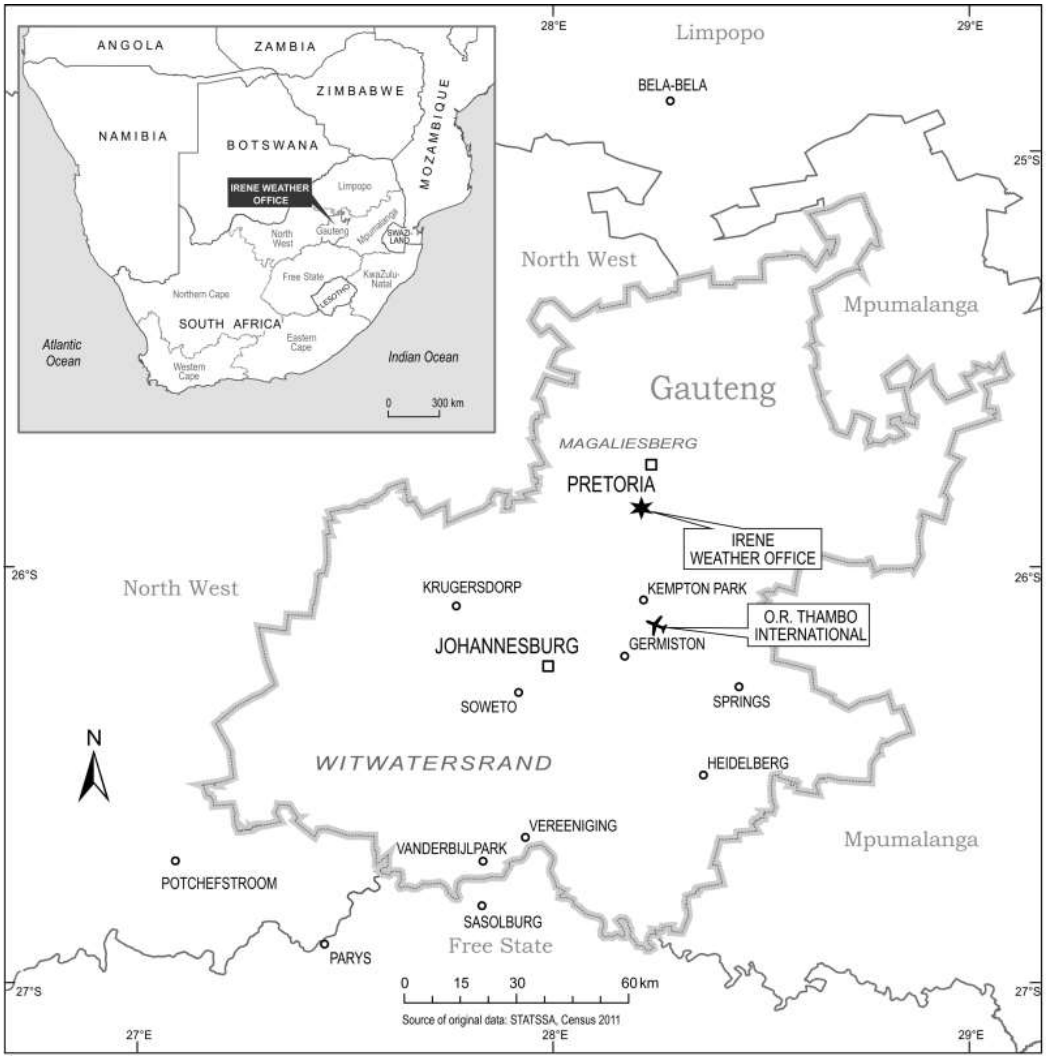

Figure 1: Geographical location of southern Africa and Gauteng.

4.2 Definition of heavy rainfall

Daily rainfall data for approximately 50 SAWS rainfall stations across Gauteng were analysed for all summer days from October 1977 to March 2012 to understand the characteristics of daily heavy rainfall over the province (Dyson 2009). Daily heavy rainfall is defined as when at least one rainfall station over Gauteng receives more than 50 mm, the average rainfall of the 50 rainfall stations is more than 3 mm (average Gauteng rainfall) and at least 30% of the rainfall stations receive more than 0 mm. Forecasters at SAWS issue advisories and warnings of heavy rainfall when more than 50 mm is expected in 24 h. The 50 mm threshold is also close to the 90th percentile of 24 h summer rainfall over Gauteng (Dyson 2009). The average Gauteng rainfall, as well as the 30% distribution criteria, was used to ensure that the rainfall event

was not restricted to a single station, that the rainfall occurred over at least a third of the province and that significant rainfall was also recorded at other rainfall stations.

4.3 Sounding data and parameters

DYSON14 expand on the difficulties in identifying proximity soundings, as the main challenge is to sample the atmosphere in which the event formed. The Irene sounding is considered to be a proximity sounding for Gauteng, as the boundaries of Gauteng are all within 100 km of Irene (Fig. 1) and it adheres to the spatial criteria stipulated by Craven and Brooks (2004) and Groenemeijer and Van Delden (2007). The rainfall over Gauteng is predominantly convective (Gijben, 2012) and occurs most frequently just after 16:00 UTC (Rouault et al. 2013). In South Africa local time is UTC + 2 h. Only daily rainfall data are available for the entire period under investigation and it is therefore assumed that the 12:00 UTC sounding is most representative of the atmosphere prior to the onset of rainfall and only 12:00 UTC sounding data were therefore analysed.

SAWS upper air data for the Irene weather office in Gauteng were obtained from http://www.weather.uwyo.edu for 35 austral summers from October 1977 to March 2012. The data comprised temperatures, dew point temperatures, wind direction and wind speeds on standard and significant pressure levels. The quality control procedures described by DYSON14 were implemented, and those soundings that were done on days with heavy rainfall were identified after soundings that were contaminated by cloud were excluded from the dataset. There were 389 heavy rainfall soundings identified for further analysis, 177 soundings in early summer (October to December) and 212 in late summer (January to March). Temperatures, dew point temperatures and wind components were interpolated to 12 pressure levels at 50 hPa intervals from 850–300 hPa. Twelve sounding-derived parameters were calculated and 389 arrays were constructed. Each array consisted of a 60-element vector: temperature, dew point temperature, meridionial and zonal winds on 12 pressure levels, as well as 12 sounding parameters. The final arrays, which were used to run the SOM, were prepared by standardising each data element of the array using the mean and standard deviation.

After some experimentation, it was found that 12 nodes (4 x 3) capture the variability of the thermodynamic profile of the atmosphere sufficiently. Using more nodes would provide finer detail and less generalisation in the thermodynamic profile, but it was found that using 12 nodes maximises the number of days mapped to each node and allows for the calculation of meaningful statistics.

Table 1: Equations, units and abbreviations of the sounding parameters. The following symbols are used: (T) Temperature (°C), (Z) Geopotential height (gpm), (q) Mixing ratio ($kgkg^{-1}$), (p) Pressure (Pa in all calculations), (v) Meridionial wind speed (ms^{-1}), (u) Zonal wind speed (ms^{-1}). Adapted from Dyson et al. (2014).

The 12 sounding parameters that were included in the data arrays and used to train the SOM were selected

Parameter	**Units**	**Abbreviation**	**Equation**
Temperature difference in the vertical	°C	$TD_{p2,p1}$	where p2>p1
Mean layer dew point temperature	°C	Td_{100}	Average dew point temperature in the 100 hPa above ground level
Precipitable water	mm	PW	$PW = \frac{1}{g} \sum_{i=850hPa}^{i=300hPa} \frac{q(p_i)+q(p_{i+1})}{2} \quad (p_i - p_{i+1})$
Wind speed at 300 hPa	ms^{-1}	W_{300}	
Bulk wind shear between 850 and 400 hPa	ms^{-1}	$BS_{850,400}$	Magnitude of vector difference
Mean meridionial wind speed in the 800-600 hPa layer	ms^{-1}	$V_{800,600}$	Average meridionial wind speed of all pressure levels between 800 and 600
Mean layer equivalent potential temperature	K	$\Theta_{e\ 100}$	Average Θ_e in the 100 hPa above ground level
Equivalent potential temperature lapse rate	K	$\Delta\Theta_{e\ 850,400}$	
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
Showalter index	°C	SI	Parcel temperature lifted from 850 hPa (calculated using mean mixing ratio in lowest 100 hPa) – environmental temperature at 500 hPa
Elevated K-Index	°C	EKI	T_{700}-T_{500}+Td_{700}-(T_{600}-Td_{600})
Warm cloud depth	hPa	WCD	Pressure of freezing level – pressure of local condensation level (LCL)

because DYSON14 showed that they were capable of distinguishing heavy rainfall days from the climatological mean or because they are significant to certain types of weather systems and severe weather (Table 1). Two parameters that deal with moisture, Td_{100} and PW, were included and were capable of distinguishing between the climatological mean and heavy rainfall for all months. DYSON14 indicated that heavy rainfall in early summer is associated with atmospheric conditions conducive to the development of severe storms and some of the parameters that help identify these conditions are the

temperature difference between 850 and 700 hPa ($TD_{850,700}$), bulk wind shear between 850 and 400 hPa ($BS_{850,400}$) and the wind strength at 300 hPa (W_{300}). These three parameters were capable of distinguishing between heavy rainfall and the climatological mean in October and November only. All the other parameters listed in Table 1 distinguish between the climatological mean and heavy rainfall in all the months, with the exception of CAPE, which was not capable of doing this in January and March.

5. Results

5.1 Frequency of occurrence and seasonal distribution of SOM archetypes

Figure 2 is a Sammon map, which is a two-dimensional representation of the inter-nodal distance across the data space of the SOM and indicates how much the vectors associated with each node differ from the vectors of surrounding nodes (Lennard and Hegerl 2014; Malherbe et al. 2013). The nodes on the right of the SOM are slightly more concentrated than on the left, but it is nevertheless a well-ordered Sammon

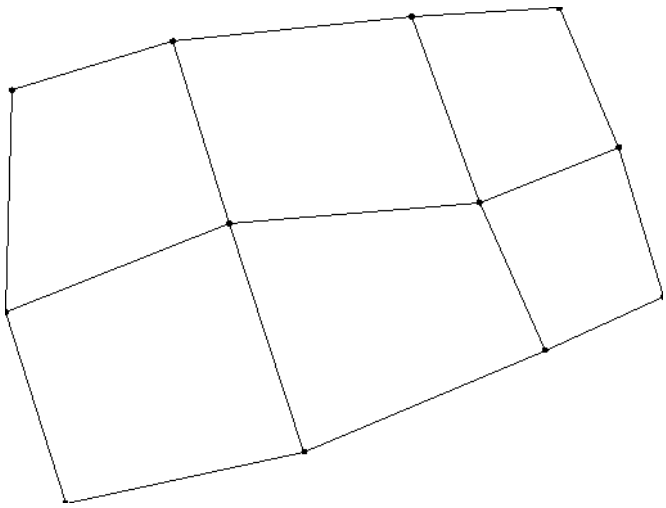

Figure 2: The Sammon map representing the distances between SOM nodes on a two-dimensional plain for heavy rainfall soundings from 1977-2012 at Irene.

map that allows for vigorous interpretation of the nodal relationships. Figure 3 is a graphic representation of temperature, dew point temperature and winds on skew-t gram plots of the 4 x 3 node SOM. Table 2 shows the associated sounding-derived parameters for each of the 12 nodes.

One of the advantages of the SOM is that each day can be related to one of the 12 nodes, allowing the calculation of the frequency which each node was mapped to over the 35-year period. Node 9 was mapped

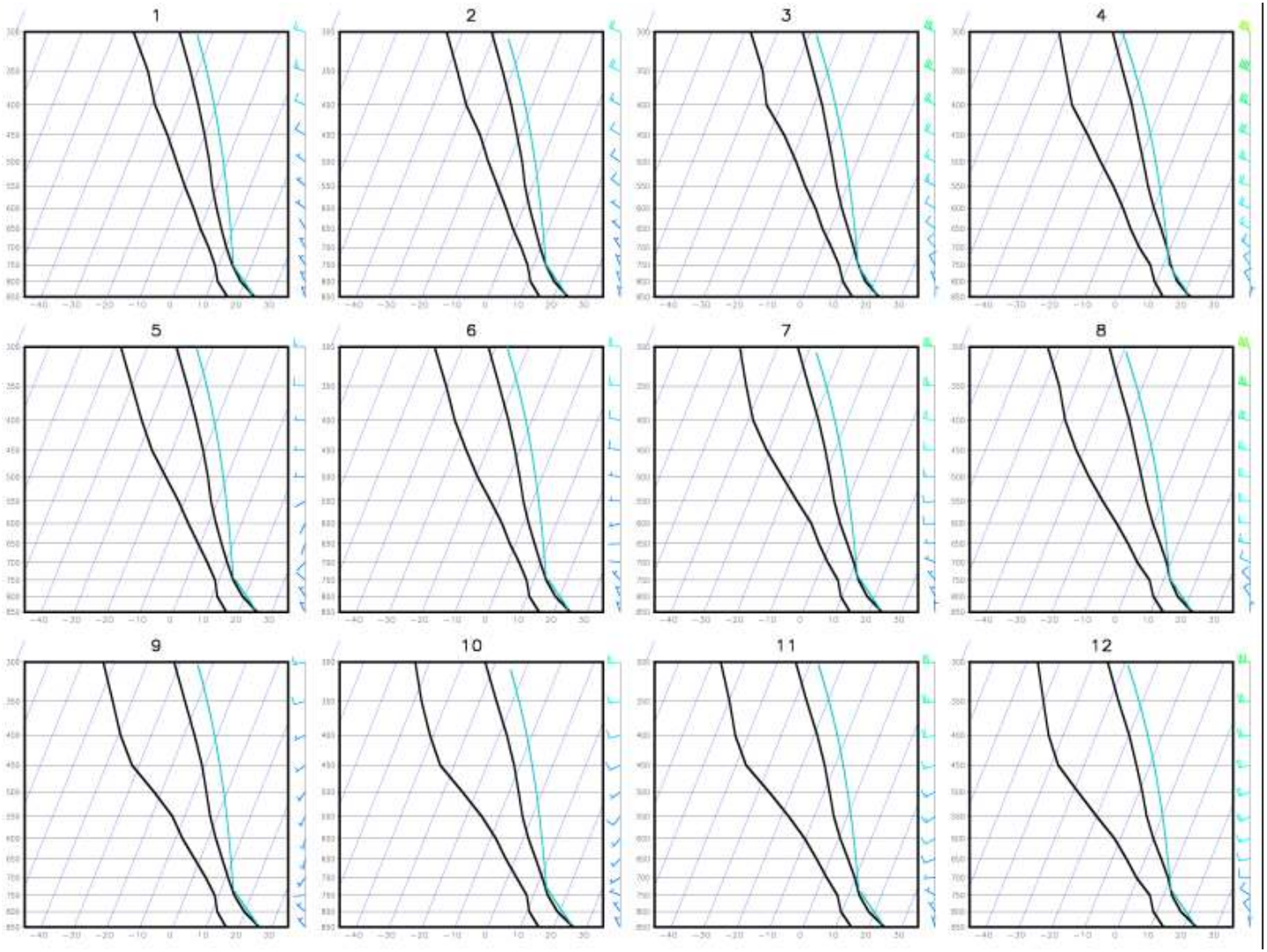

Figure 3: Archetypal skew-t gram plots of austral summer heavy rainfall soundings from 1977-2012 over Gauteng. The solid blue lines show the track of a parcel if lifted from the surface to the lifted condensation level by the dry adiabatic lapse rate and then follows the saturated adiabatic lapse rate to 300 hPa.

to most frequently (> 11%), followed by Node 1 (10%) (Fig. 4a). Nodes on the left side of the SOM were mapped to more regularly (yellow and green shades) than the nodes on the right (blue shades).

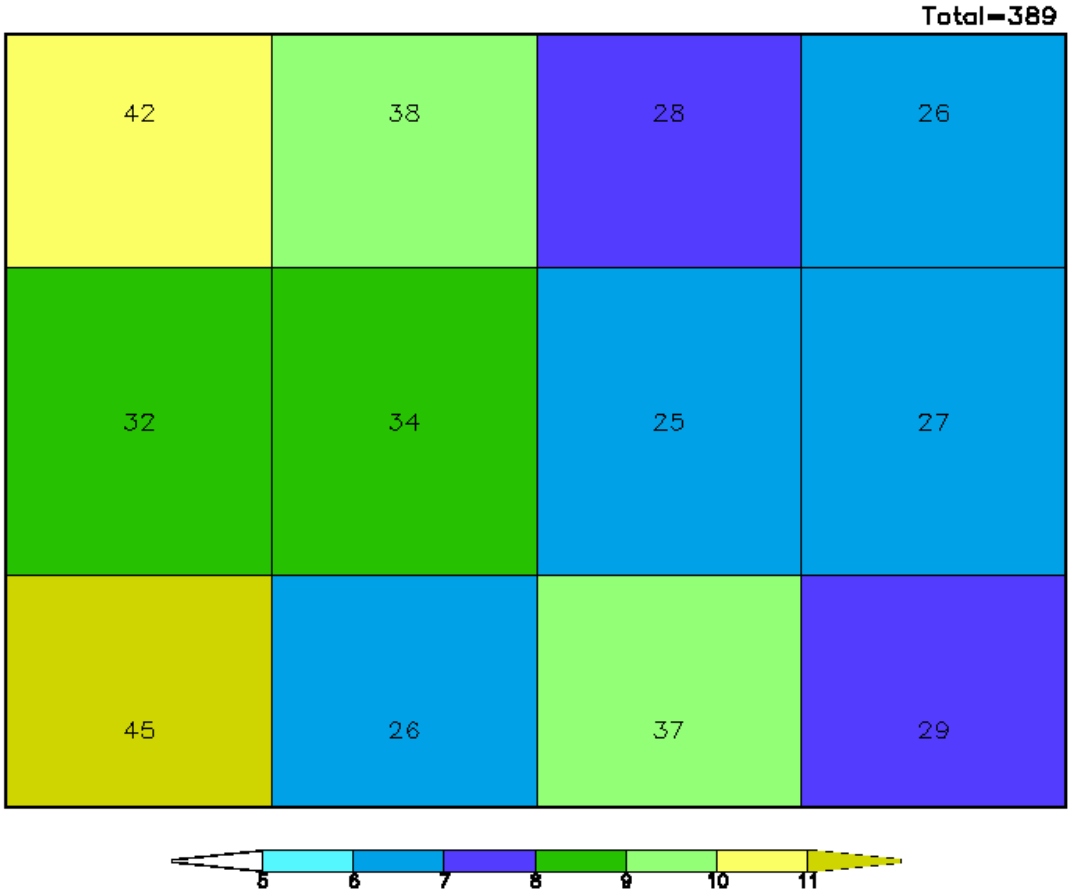

Figure 4a: Frequency of occurrence (%) of each node (shaded) and the total number of days mapped to each node of the SOM (numbers in each block). Numbering matches Fig. 3 where the top left square corresponds to Node 1 and the bottom right square to Node 12.

The nodes on the right side of the SOM are mapped to most frequently in early summer (October to December), while the nodes on the left are mapped to most frequently on late summer (January to March) heavy rainfall days (Fig. 4b). There were 177 heavy rainfall days in early summer and 67% of these days were mapped to nodes in the last two columns (early summer nodes). The nodes in the last column were mapped to on 82 days and 81% of these days were early summer days. There were 212 late-summer heavy rainfall days and 75% of these days were mapped to the nodes in the first two columns (late summer nodes). The first two columns were mapped to on 217 days and 72% of these days were late summer days.

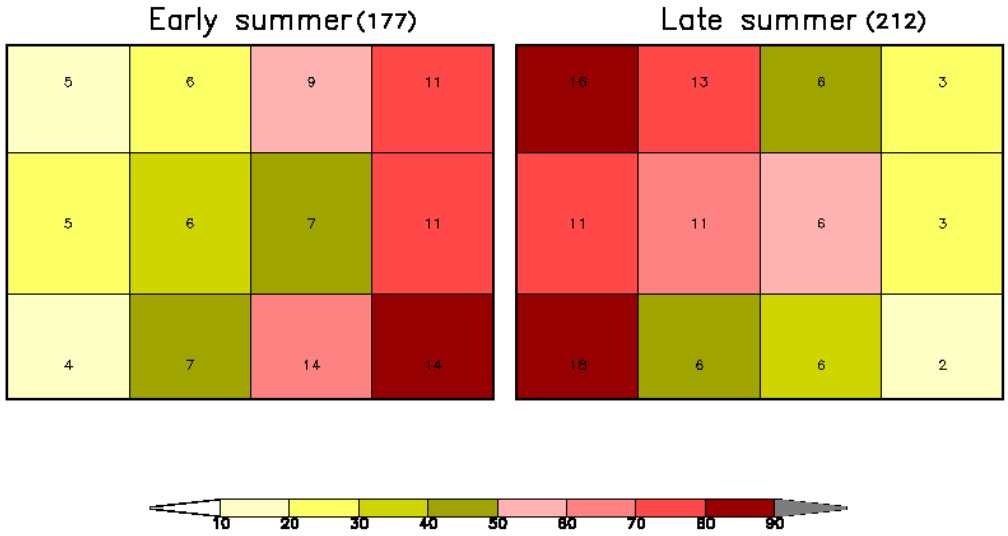

Figure 4b: The frequency (%) of occurrence of each node in early summer (left) and late summer (right) expressed relative to the total number of occurrences of that node during the entire period (shaded). The number in each block is the frequency (%) of occurrence of each node relative to the total number of occurrences of heavy rainfall during that particular season.

During October and November, there is a tendency towards mapping heavy rainfall days to the right of the SOM, and during January and February to the left (Fig. 4c). In October the nodes in the last column are mapped to most frequently with almost no heavy rainfall days mapped to the nodes on the left hand side of the SOM. The distribution in November is similar but now more nodes in the first 2 columns are mapped to. There is a near-even distribution of mapped-to nodes (Fig. 4c) in December and March. In December and March months the nodes in the two centre columns are mapped to most frequently, although the bottom row has a higher frequency of being mapped to in December and the top row in March. The first two columns (late summer columns) are mapped to on 56% (57%) of heavy rainfall December (March) days and the last two columns (early summer columns) are mapped to on 44% (43%) of the December (March) days. There is a near-even distribution between heavy rainfall days being mapped to the typical early summer nodes and the days being mapped to late summer nodes for these two months. During December, the atmosphere changes from extra-tropical to tropical and returns to an atmosphere with typical extra-tropical characteristics in March. In January and February months the late summer nodes in are mapped to more frequently. Node 9 is mapped to on more than 40% of heavy rainfall days in January.

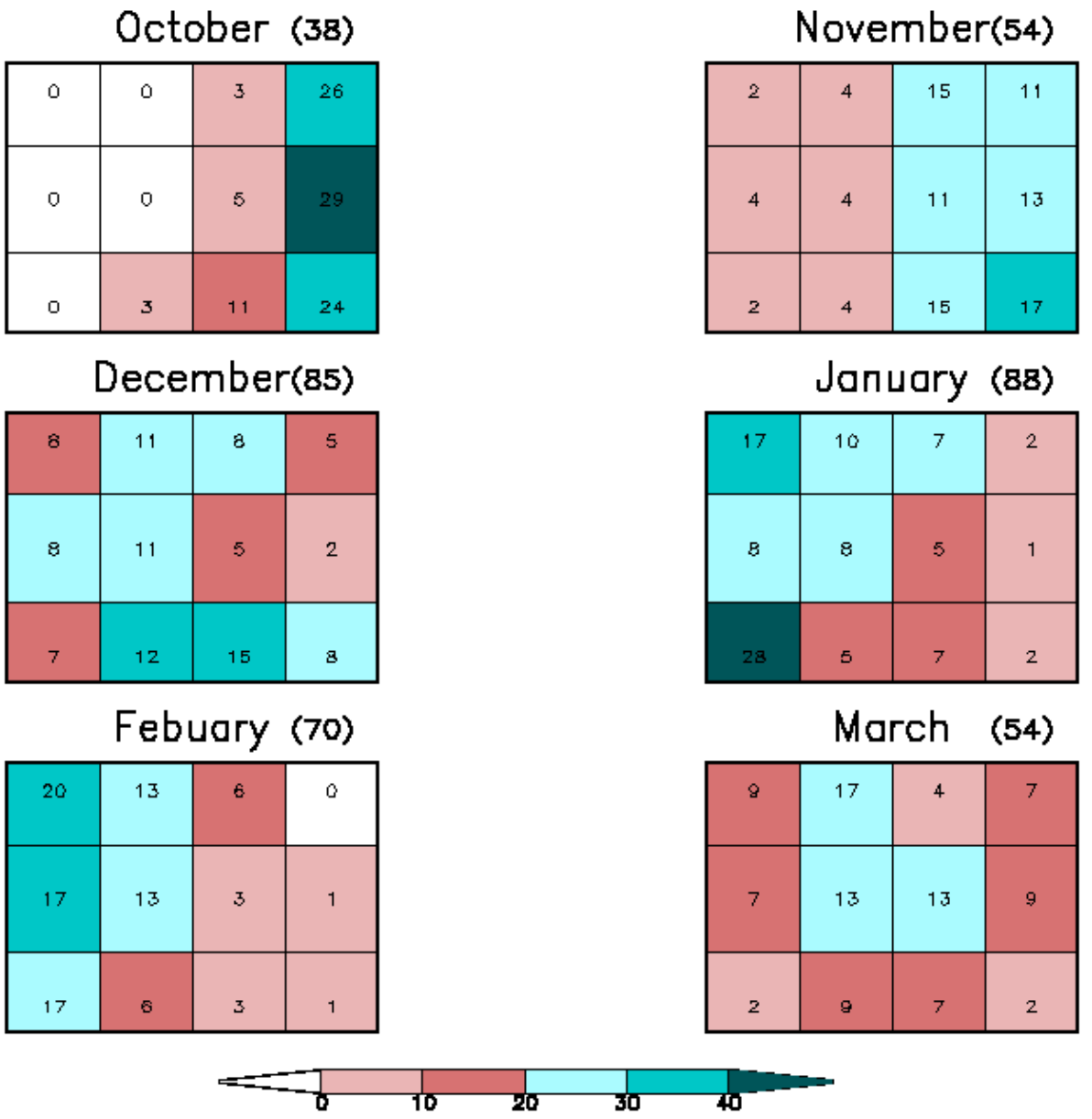

Figure 4c: The frequency (%) of occurrence of each node expressed relative to the total number of occurrences of that node during the entire period (shaded). The number in each block is the frequency (%) of occurrence of each node relative to the total number of occurrences of heavy rainfall during that particular month.

5.2 Sounding and derived parameter archetypes

The nodes in the first two columns are warm, with high water vapour content and small wind shear (Table 2 and Fig. 3), which are typical characteristics of a tropical atmosphere. The nodes in the last two columns have significantly larger wind shear, are cooler and have less moisture than the nodes on the left.

Table 2: Archetypal values of sounding-derived parameters for austral summer heavy rainfall soundings over Gauteng from 1977-2012. Ave denotes the mean daily Gauteng rainfall, Max the mean of the daily maximum rainfall and Dist the daily mean percentage of rainfall stations that reported more than 0 mm on all the days mapped to a node. The last two rows are the mean values on all heavy rainfall days in early summer (Early) and late summer (Late).

Node	$TD_{850,700}$	PW	Td_{100}	W_{300}	$BS_{850,400}$	$V_{800,600}$	$\Theta_{e\,100}$	$\Delta\Theta_{e\,850,400}$	CAPE	SI	EKI	WCD	Gauteng Rain		
													Ave (mm)	Max (mm)	Dist (%)
1	14.7	28	13	8	7	-1.4	348	18	753	-2.1	13	156	12	71	68
2	14.5	27	12	10	8	-2.2	346	16	715	-2.1	13	148	14	78	70
3	13.7	24	11	16	12	-3.7	342	15	653	-2.2	11	142	15	75	74
4	13.2	22	10	21	15	-4.4	339	14	628	-2.2	9	129	17	77	73
5	15.2	27	13	7	6	-0.1	348	20	1063	-2.8	12	143	8	72	52
6	15.2	25	12	10	8	-0.8	346	18	970	-2.9	12	136	10	71	58
7	14.7	23	11	14	11	-1.7	342	17	946	-3.0	10	122	10	66	67
8	13.9	21	10	20	15	-2.4	339	16	816	-2.9	7	118	12	72	67
9	15.5	25	13	7	6	0.8	349	22	1333	-3.6	11	137	10	75	61
10	15.5	24	12	9	8	0.3	347	21	1314	-3.6	10	124	10	77	67
11	15.1	21	11	14	11	-0.3	343	20	1185	-3.7	8	118	9	74	64
12	14.6	20	10	17	14	-0.8	340	18	1012	-3.3	6	109	11	68	66
Early	14.6	23	11	14	11	-2.2	342	17	1031	-3.2	9	120	12	73	67
Late	15.0	26	13	10	9	-0.4	348	21	1100	-3.0	11	143	11	75	65

The temperatures throughout the troposphere decrease from left to right in each row of the SOM (Fig3) resulting in the nodes with the highest temperatures being on the left, and the nodes with the lowest temperatures being on the right of the SOM. There is also a decrease in the dew point temperatures throughout the troposphere (Fig. 3) as well as Td_{100} and PW values (Table 2) from left to right in each row. Node 1 has the highest PW value (28 mm) of all the nodes, while in Node 12, far removed from Node 1 in the SOM space, the PW is only 20 mm. The wind strength throughout the troposphere increases from left to right in each row of the SOM and W_{300} values are the highest in the last column. In Node 4, W_{300} is 21 ms^{-1}, but a third of this value at 7 ms^{-1} in Node 5 and Node 9. In the nodes on the bottom left of the SOM, the winds are very light, less than 2.5 ms^{-1} in the lower levels of the atmosphere. These same

nodes also have very low wind shear values with Node 5 and Node 9 having $BS_{850,400}$ values of only 6 ms^{-1}. The three nodes in Column 4 have wind shear values of 14 ms^{-1} and more.

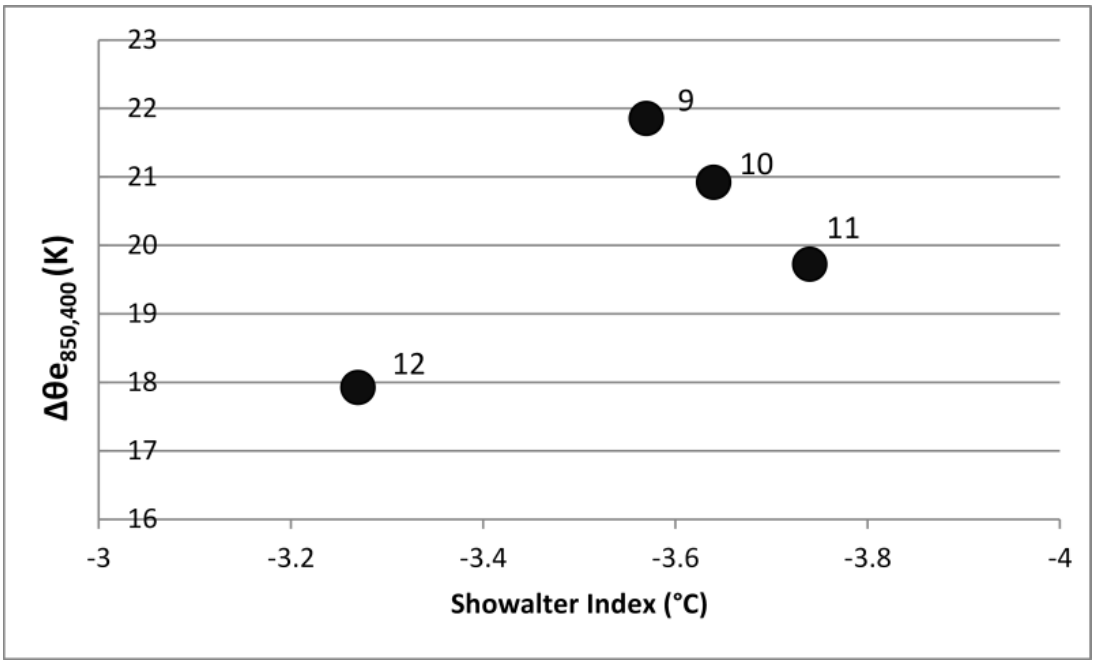

Figure 5a: Scatter plot of the Showalter Index (SI) and 850 hPa to 400 hPa equivalent potential temperature lapse rate ($\Delta\theta_{e\,850,400}$) for the nodes in the third row of the SOM. The number on the graph is the number of the node.

Conditional instability is essentially a lapse rate definition where the temperature lapse rate has to be steeper than the moist adiabatic lapse rate (Doswell et al. 1985), while the atmosphere is convectively unstable if $\Delta\theta_e > 0$ through a significantly deep layer of the atmosphere (Shultz et al. 2000). The value of $\Delta\theta_{e\,850,400}$ is used here as an indication of the convective instability of the atmosphere, while the SI is an indicator of the conditional instability. Consider Fig. 5a which shows the relationship between the SI and $\Delta\theta_{e\,850,400}$ in the third row of the SOM (Fig. 3). As the atmosphere becomes increasingly conditionally unstable (larger negative SI values) from left to right in Row 3, the convective instability decreases as $\Delta\theta_{e\,850,400}$ values become correspondingly smaller. Node 12 is out of sequence, as both the SI and $\Delta\theta_{e850,400}$ are relatively low in this node. Fig. 5b examines the relationship between $\Delta\theta_{e\,850,400}$ and $TD_{850,700}$ (black), as well as $TD_{850,500}$ (grey) for the nodes in the first row of the SOM. The increase in $\Delta\theta_{e\,850,400}$ is directly proportional to the shallow-layer vertical temperature difference ($TD_{850,700}$), but indirectly proportional to a deeper-layer vertical temperature difference ($TD_{850,500}$). DYSON14 indicated that the seasonal increase in mean temperatures in the lower troposphere is only about 2 °C from October to January, while the temperatures in the corresponding period increase by 4 °C at 500 hPa. The late summer nodes (Node 1 and Node 2) have higher temperatures throughout the troposphere than the early summer nodes (Node 3

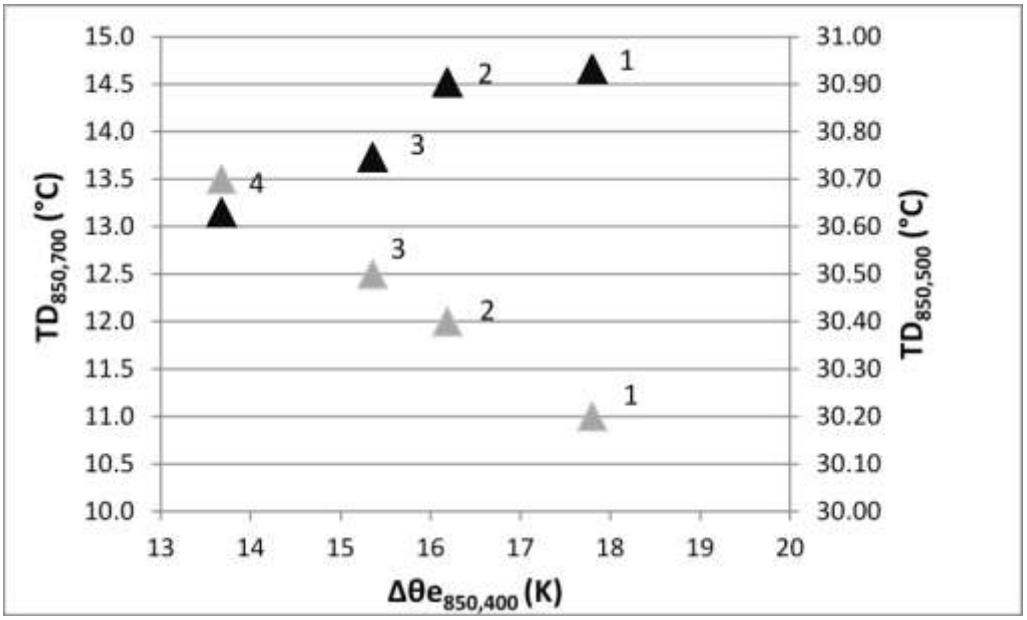

Figure 5b: Scatter plot of the 850 hPa to 400 hPa equivalent potential temperature lapse rate ($\Delta\theta_{e\,850,400}$) and the temperature difference between 850 and 700 hPa ($TD_{850,700}$ black triangles and left axis), as well as 850 and 500 ($TD_{850,500}$ grey triangle and right axis) for the first row of the SOM. The number on the graph is the number of the node.

and Node 4). The difference in temperature between early and late summer nodes is larger in the middle troposphere than at levels closer to the ground. Therefore a temperature difference over a deep layer of the

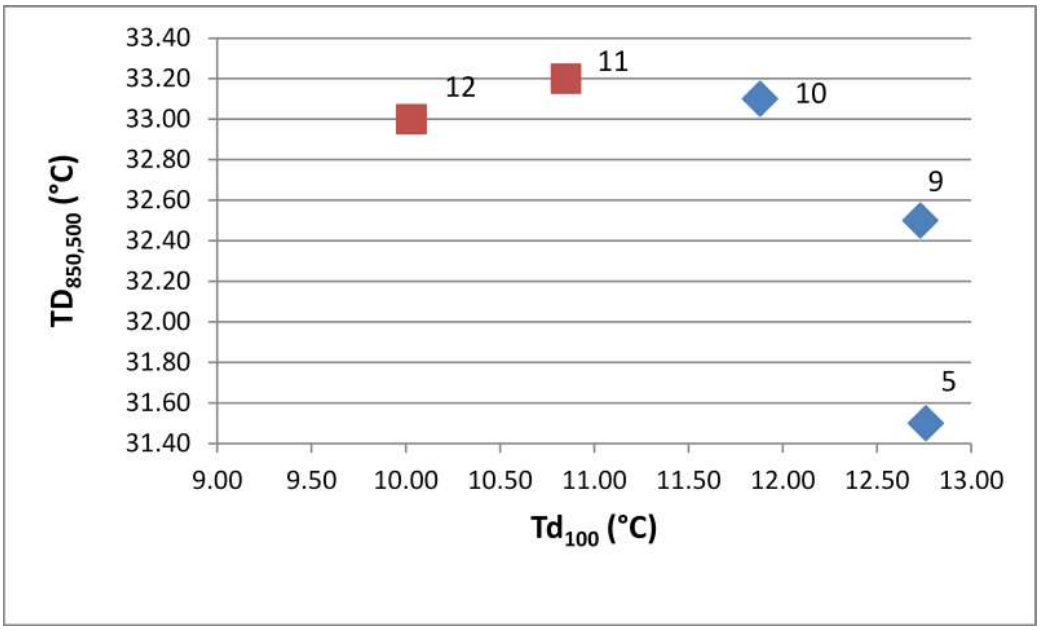

Figure 5c: Scatter plot of the mean layer dew point temperature (Td_{100}) and the temperature difference between 850 and 500 ($TD_{850,500}$) for early-summer (brown) and late-summer (blue) nodes when the CAPE value of the node > 1 000 Jkg^{-1}. The number on the graph is the number of the node.

atmosphere (850–500 hPa) is a better indicator of the conditional instability of the atmosphere than a temperature difference over a shallow layer (850–700 hPa). The late summer nodes are convectively unstable (large values of $\Delta\theta_{e\,850,400}$) and the early summer nodes conditionally unstable (large values of . $TD_{850,500}$).

CAPE and $\Delta\theta_e$ represent similar physical atmospheric processes (Cohen et al. 2007) and CAPE values decrease similarly to $\Delta\theta_{e\,850,400}$ from left to right in each row of the SOM (Table 2). Fig. 5c is a scatter plot of the Td_{100} and $TD_{850,500}$ values for the early summer (brown) and late summer (blue) nodes when the CAPE value of the node exceeds 1 000 Jkg^{-1}. In early summer, when large CAPE occurs, the Td_{100} values are relatively small (< 11 °C), but $TD_{850,500}$ values are large with values of 33–34 °C. In late summer, with CAPE > 1 000 Jkg^{-1}, $TD_{850,500}$ values are generally less than in early summer, but the dew point temperatures are 1–2 °C higher. The scatter plot illustrates how CAPE is a function of both the moisture content of the lifted parcel and the atmospheric lapse rate. In early summer, $TD_{850,500}$ values are dominant for large CAPE and, in late summer, high parcel dew point temperatures play a leading role. However, CAPE and $TD_{850,700}$ have a direct relationship (Table 2), and higher $TD_{850,700}$ values are associated with higher CAPE values, irrespective of the season. Large temperature differences through a deep layer of the atmosphere (850–500 hPa) are associated with large CAPE in early summer, while large shallow-layer temperature differences (850–700 hPa) are associated with larger CAPE throughout the summer season. Over South Africa in summer, the convective cloud base height is close to the 700 hPa pressure level (Dyson and Van Heerden 2002) and large temperature lapse rates between the ground and the cloud base are conducive to larger CAPE.

Buoyancy and wind shear are considered essential for deep moist convection. Convection is organised in lines by wind shear (for example, squall lines), and super cells may develop when the relationship between buoyancy and wind shear is ideal (Wu and Yanai 1994). If wind shear is absent, 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. Figure 5d shows the relationship between CAPE and $BS_{850,400}$ for the nodes in the third row of the SOM. CAPE decreases from left to right, but $BS_{850,400}$ increases from Node 9 to Node 12. In late summer, buoyancy is dominant, while wind shear is more important in early summer. The atmospheric conditions associated with heavy rainfall over Gauteng vary from low-shear, high-CAPE environments (Node 9) to high-shear, low-CAPE environments (Node 4).

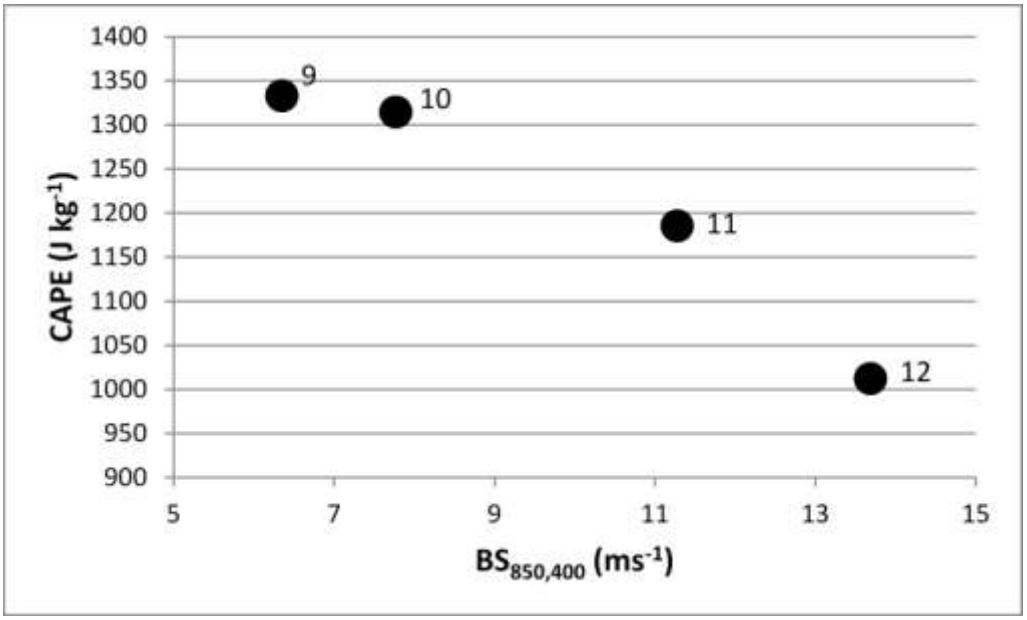

Figure 5d: Scatter plot of the wind shear from 850 to 400 ($BS_{850,400}$) and CAPE for the nodes in the third row of the SOM. The number on the graph is the number of the node.

DYSON14 indicated how precipitation-efficient storms with large warm cloud depth (WCD) and low wind shear values occur on heavy rainfall days in late summer. The SOM was able to capture the same relationship, because low wind shear values and high WCD values are present in the late summer nodes (Table 2). $BS_{850,400}$ values clearly distinguish between early and late summer nodes with large $BS_{850,400}$ values in early summer and smaller values in late summer nodes. The WCD values in early summer are small, while these values are large in late summer.

5.3 Early summer heavy rainfall

The nodes of Column 1 and Column 4 are analysed to illustrate the variability that is present in heavy rainfall conditions during late (Column 1) and early (Column 4) summer heavy rainfall.

The nodes in the last column (nodes 4, 8 and 12) are almost exclusively mapped to on early summer days. More than 50% of heavy rainfall days in October and November are mapped to these three nodes. These nodes have the strongest winds (W_{300} in Table 2 and Fig.3) and highest $BS_{850,\ 400}$ values of all the nodes, although there is a slight decrease in value of these variables from Node 4 to Node 12. The $V_{800,600}$ values are all negative, as north-westerly winds generally occur in the lower troposphere, but the absolute value of $V_{800,600}$ decreases from Node 4 to Node 12 as the winds become westerly to south-westerly in Node 12. The surface dew point temperatures in these three nodes are very similar, but Node 4 has a higher PW

value than Node 12, where it becomes quite dry in the mid-levels. Node 4 has the lowest CAPE value of all the nodes (628 Jkg^{-1}), while the CAPE in Node 12 is more than 1 000 Jkg^{-1}. All three these nodes are conditionally unstable, as reflected by the negative SI, but with increasing conditional instability from Node 4 to Node 12. The elevated EKI is below 10 in all three nodes, and θe_{100} is close to 340 K. In Node 12, the WCD is low due to the relatively dry atmosphere.

Figure 6 is a composite map that reflects the average circulation for all the days mapped to a node. It should be noted that the interior plateau of South Africa rises to 1 500 m over extensive areas, and the 850 hPa geopotential height field therefore represents the surface circulation. Node 4 is characterised by a deep surface trough (1 490 gpm) over the central interior of southern Africa and the Indian Ocean High (IOH) (1 520 gpm) is located over Madagascar. The high pressure system is located in the Mozambique Channel at 700 hPa, causing a strong flux of moisture from Angola and Zambia into Gauteng. A sharp westerly trough lies to the west of Gauteng at 500 hPa. This trough is in the most northern position of all the nodes. Of all the days mapped to Node 4, 57% were days on which COL pressure systems were present west of Gauteng. The remainder of the days all had sharp westerly troughs at 500 hPa. The average Gauteng rainfall on all the days mapped to Node 4 was 17 mm, with 73% of the rainfall stations over Gauteng reporting rainfall of more than 0 mm (Table 2). These values are the highest of all the nodes and result from the well-developed extra-tropical weather systems approaching Gauteng. Even though the atmosphere is not particularly unstable, the high wind shear values and strong moisture flux accompanied by favourable synoptic forcing causes widespread and heavy rainfall on the days mapped to this node.

In Node 12, the surface trough is still present west of Gauteng, but it is around 30 gpm weaker than in Node 4 (Fig.6). The IOH is stronger than in Node 4 and is located over the eastern extremes of South Africa and Mozambique, while the centre of the high is over southern Zimbabwe at 700 hPa. The moisture flux into Gauteng is weaker than Node 4 and the anti-cyclonic circulation crosses the Atlantic Ocean. Taljaard (1996) stated that the dry and stable air over the ocean modifies the air mass so that the air that eventually reaches Gauteng is less favourable for convective development. In this node, the 500 hPa trough is weak and lies east of Gauteng, while a high pressure extends a ridge into central South Africa. However, a trough in the 500 hPa temperature field (temperature trough) is present west of Gauteng, which causes the advection of colder temperatures into the province aloft, enhancing the conditional instability of the atmosphere. This can also be seen by the higher CAPE and SI values in this node. In Node 12, the synoptic forcing is not as favourable as in Node 4, but local conditions are favourable for the development of thundershowers, as the surface temperatures are higher and the atmosphere is

conditionally unstable. The average Gauteng rainfall on days mapped to Node 12 is 11 mm, with 66% of rainfall stations receiving some rain.

The advantages of using the SOM archetypes in a climatology of soundings and derived parameters are that the variation of conditions conducive to heavy rainfall are captured by the nodes and that the relationship between the different variables are also available. The heavy rainfall climatology, as provided by DYSON14, provide average monthly values when heavy rainfall occurs and therefore provides insight into climatological extreme values when heavy rainfall occurs. The SOM climatology complements this by also providing a range of possible conditions for heavy rainfall. Compare the sounding-derived parameters of Node 4 and Node 12 with the average early summer values of the sounding-derived parameters when heavy rainfall occurs (Early in Table 2). In Node 4, θe_{100}, $\Delta\theta_{e\,850,400}$, CAPE and SI values are less favourable for convective development than the average early summer heavy rainfall values, but with adequate moisture and strong wind shear, conditions remain favourable for heavy rainfall to develop. In Node 12, there is less moisture than the average early summer heavy rainfall values indicate. However, in this case, the large CAPE and favourable SI and $\Delta\theta e_{850,400}$ values explain the occurrence of heavy rainfall.

5.4 Late summer heavy rainfall

The nodes in the first column are mapped to on 54% of heavy rainfall days in January and February. These three nodes all have weak winds (W_{300} and $BS_{850,\,400}$ in Table 2). In Node 5, the mid-level winds are particularly weak with $V_{800,600}$ only -0.06 ms^{-1}. Node 1 has north-westerly winds through the troposphere, while Node 9 has light southerly to south-westerly winds at pressure levels lower than 750 hPa. CAPE, SI and $\Delta\theta_{e\ 850,400}$ are more favourable in Node 9 than in Node 1, but with the EKI higher in Node 1, as the atmosphere is moister in this node. The PW in Node 1 is 28 mm, the highest of all the nodes, while the value in Node 9 is 25 mm.

The composite maps reveal a surface trough of 1 500 gpm extending from southern Angola and Zambia to the west of Gauteng in Node 1 (Fig. 6). There is also a notable 850 hPa low over the Mozambique Channel, while the IOH of 1 520 gpm is located close to the coast of KwaZulu-Natal. The centre of the 700 hPa high is located over north-eastern South Africa and causes an influx of moisture from the Mozambique Channel into southern Africa. There is also strong northerly moisture flux directly to the east of the 700 hPa low over northern Namibia/southern Angola. This low has been associated with heavy rainfall over the summer rainfall area of South Africa, as tropical air flows from the ITCZ into the

southern subcontinent (Reason et al. 2006). The 500 hPa high pressure centre is in the same location at the 700 hPa high, but with very weak geopotential gradients over South Africa. Although not clearly visible on this image, a 500 hPa trough is located over southern Botswana and with the warm core of 500 hPa temperatures over Botswana, the synoptic-scale weather system represented by these conditions is the continental tropical low, which has been associated with heavy rainfall over Gauteng (Dyson 2002). The average Gauteng rainfall of the days mapped to Node 1 is 12 mm with 68% of stations receiving some rainfall (Table 2).

The synoptic circulation in Node 9 is less favourable, as the Angola low is weaker and the surface trough does not extend as far south into South Africa as it does in Node 1. The 700 hPa high is situated over the province of North West. Even though there is still an onshore flow from the Mozambique Channel, this high, together with the position of the Angola Low, forces the air over the Atlantic Ocean, prior to invading central South Africa. There is evidence of a temperature trough west of Gauteng at 500 hPa, which enhances the conditional instability of the atmosphere (also see Table 2). Strong synoptic forcing does not cause the heavy rainfall over Gauteng in Node 9, as is the case in Node 1. It is caused by the instability of the atmosphere, together with very low wind shear. The thunderstorms that develop under these conditions will be very slow-moving, and will therefore be precipitation-efficient (Doswell et al. 1996). The average rainfall in Node 9 was 10 mm and 61% of rainfall stations reported some rainfall (Table 2).

The sounding parameters of Node 1 and Node 9 are compared to the average late summer heavy rainfall sounding-derived parameters (Late in Table 2). The instability parameters of Node 1 are not as favourable as the average late summer heavy rainfall values, but the moisture parameters are higher than normal. In Node 9, the moisture parameters are similar to the long-term mean values, but the instability parameters are more favourable with larger-than-normal CAPE and $\Delta\theta_{e\,850,400}$ values.

6. Summary and discussion

The SOM methodology has been used successfully to produce synoptic climatologies (Hewitson and Crane 2002; Liu and Weisberg 2011), to investigate significant weather phenomena and to explore their characteristics (Tennant and Hewitson 2002; Van Schalkwyk and Dyson 2013; Lennard and Hegerl 2014). In this paper, it is shown how SOMs may also be used to create a heavy rainfall sounding climatology using observed upper air data and parameters derived from these data. This paper is one of the very first that describes the use of sounding data in SOMs to create a climatology. The days mapped to a particular node of the SOM have similar thermodynamical profiles. Further exploration of the synoptic circulation

patterns of the days mapped to a node reveals that these similarities are also present in the synoptic circulation patterns.

Irene sounding data were obtained for 35 austral summers, and 289 1200 UTC soundings were isolated on days when heavy rainfall occurred. These soundings, as well as sounding-derived parameters, were used to create a 12-node SOM to produce a sounding-derived parameter climatology over Gauteng. Some of the limitations of this study are the assumption that the 1200 UTC Irene sounding is representative of the rainfall for a 24-h period and that the upper air ascent at Irene is representative of the atmospheric conditions over the entire province.

The seasonal frequency maps show that the SOM is capable of distinguishing heavy rainfall days that occur in early summer (October to December) from those that occur in late summer (January to March). The mapped-to nodes on heavy rainfall days in December and March are evenly distributed in the SOM. December sees the onset of predominantly tropically induced heavy rainfall, while conditions return to typical extra-tropical circulation in March. The nodes that represent the early summer heavy rainfall are cooler and drier than the nodes that represent the late summer heavy rainfall. In the late summer nodes, the winds are weak with small values of wind shear, while strong winds occur throughout the troposphere in early summer, and wind shear values are as much as three times stronger than the late summer nodes. The late summer nodes are convectively unstable with larger values of $\Delta\theta_{e\,850,400}$, while the SI values are higher in early summer, indicating higher conditional instability. In early summer, nodes with large CAPE values are associated with larger values of $TD_{850,500}$ and smaller Td_{100} values, while the dew point temperatures are higher and the temperature lapse rates lower in late summer. The temperature lapse rate beneath the cloud base ($TD_{850,700}$) correlates well with CAPE throughout the entire summer. Temperature differences over a deep layer ($TD_{850,500}$) is a better representation of conditional instability than temperature differences in the boundary layer. The SOM climatology complements the traditional climatology of sounding-derived parameters (DYSON14) by being able to describe the intra-seasonal variability of heavy rainfall soundings. It also provides information about the relationship between different parameters as the season changes.

DYSON14 provided a monthly climatology of sounding-derived parameters associated with heavy rainfall and the SOM climatology provides additional information on how atmospheric conditions in early (or late) summer vary when heavy rainfall occurs. In both seasons, there are some days with heavy rainfall when significant weather systems cause strong moisture flux into Gauteng with large wind shear, but with relatively unfavourable instability. In these instances, the dynamic forcing of the larger-scale weather

systems is responsible for the heavy rainfall. There are nodes on the other side of the spectrum where the synoptic circulation patterns are not particularly favourable, with less moisture and weaker wind shear. These nodes are slightly colder aloft and warmer at the surface, which causes enhanced atmospheric instability and aids convective development. The archetypical values of the sounding-derived parameters provide typical values associated with the different conditions under which heavy rainfall occurs.

An investigation into the atmospheric flow fields associated with each of the nodes reveals that the days mapped to each of the nodes have similar synoptic circulation patterns. In early summer, the highest average Gauteng rainfall is caused by strong baroclinic weather systems, such as sharp westerly troughs or COLs west of Gauteng. In late summer, the highest rainfall occurs when weather systems associated with the Angola low, in association with the continental tropical low, are present over southern Africa. The synoptic circulation of days mapped to Node 12 in early summer is where the westerly wave system lies to the east of Gauteng in the upper air with a high over Namibia, which extends a ridge over Gauteng. In late summer, days with an upper air high pressure located over central South Africa are mapped to Node 9. For heavy rainfall to occur over Gauteng, the average atmospheric flow fields reveal that a surface trough that causes north-westerly winds at the surface should be present west of Gauteng. Furthermore, a 500 hPa temperature trough should be present west of Gauteng, as the cold air advection aloft enhances the conditional instability of the atmosphere.

It is recommended that the capability of the SOM to "convert complex non-linear features into simple two-dimensional relationships" (Nishiyama et al. 2007) is investigated to use the soundings and derived parameters in an SOM as an objective rainfall-forecasting tool.

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.

Covadonga P, Dario G, Stel F, Castro A and Fraile R (2010) Maximum hailstone size: Relationship with meteorological variables. Atmos Res 96:256–265. doi: http://dx.doi.org/10.1175/WAF915.1.

Craven JP and Brooks HE (2004) Baseline climatology of sounding-derived parameters associated with deep moist convection. Natl Wea Digest 28:13–24.

Deboeck G and Kohonen T (1998) Visual explorations in finance: with self-organizing maps. Vol. 2. Springer-Verlag, London.

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. Natl Wea 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. Atmos Res 93:30–38. doi: 0.1016/j.atmosres.2008.10.010

Doswell CA III and Schultz DM (2006) On the use of indices and parameters in forecasting severe storms. Electron J Severe Storms Meteorol. 1:1–14.

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 Conference on Severe Local Storms, Indianapolis, IN. Am Meteorol Soc. 398–401.

Doswell CA III, Brooks HE and Maddox RA (1996) Flash flood forecasting: An ingredients-based methodology. Wea Forecasting 11:560–580. doi: http://dx.doi.org/10.1175/1520-0434(1996)011%3C0560:FFFAIB%3E2.0.CO;2.

Dupilka ML and Reuter GW (2006a) Forecasting tornadic thunderstorm potential in Alberta using environmental sounding data. Part I: Wind shear and buoyancy. Wea Forecasting 21:325–335. doi: http://dx.doi.org/10.1175/WAF921.1.

Dupilka ML and Reuter GW (2006b) Forecasting tornadic thunderstorm potential in Alberta using environmental sounding data. Part II: Helicity, precipitable water, and storm convergence. Wea Forecasting 21:336–346. doi: http://dx.doi.org/10.1175/WAF922.1.

Dyson LL (2009) Heavy daily rainfall characteristics over the Gauteng Province. Water SA 35:627–638. doi: http://dx.doi.org/10.4314/wsa.v35i5.49188.

Dyson LL and Van Heerden J (2001) The heavy rainfall and floods over the north-eastern interior of South Africa during February 2000. S Afr J Sci 97:80–86.

Dyson LL and Van Heerden J (2002) A model for the identification of tropical weather systems. Water SA 28:249–258. doi: http://dx.doi.org/10.4314/wsa.v28i3.4892.

Dyson LL, Van Heerden J and Sumner PD (2014) A baseline climatology of sounding-derived parameters associated with heavy rainfall over Gauteng, South Africa. Int J Climatol 35:114-127. doi: 10.1002/joc.3967.

Engelbrecht CJ, Landman WA, Engelbrecht FA and Malherbe J (2014) A synoptic decomposition of rainfall over the Cape south coast of South Africa. Clim Dynamics. DOI 10.1007/s00382-014-2230-5

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. doi: 10.1016/j.compenvurbsys.2012.09.001.

Gijben M (2012) The lightning climatology of South Africa. S Afr J Sc. doi: 10.4102/sajs.v108i3/4.740.

Groenemeijer PH and Van Delden A (2007) Sounding-derived parameters associated with large hail and tornadoes in the Netherlands. Atmos Res 83:473–487. doi: http://dx.doi.org/10.1016/j.atmosres.2005.08.006.

Grunz A, Memmert D and Perl J (2012) Tactical pattern recognition in soccer games by means of special self-organizing maps. Hum Mov Sci 31:334-343. doi: 10.1016/j.humov.2011.02.008.

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.

Harnack RP, Jensen DT and Cermak JR III (1998) Investigation of upper air conditions occurring with heavy summer rain in Utah. Int J Climatol 8:701–732. http://dx.doi.org/10.1175/1520-0434(1997)012%3C0282:IOUACO%3E2.0.CO;2.

Hart NCG, Reason CJC and Fauchereau N (2010) Tropical–extratropical interactions over Southern Africa: Three cases of heavy summer season rainfall. Mon Weather Rev 138:2608–2623. doi: http://dx.doi.org/10.1175/2010MWR3070.1.

Holton JR (1992) An Introducion to dynamic meteorology. Academic Press, London.

Hewitson BC and Crane RG (2002) Self-organizing maps: applications to synoptic climatology. Clim Res 22:13–26. doi: 10.3354/cr022013.

Hsu YT, Hung HF, Yeh J and Liu MC (2010) Forecast of financial time series based on grey self-organizing maps. Int J Innovative Comput, Inf Control 6:475–486.

Jensen AA, Thompson AM, and Schmidlin FJ (2012) Classification of Ascension Island and Natal ozonesondes using self-organizing maps. J Geophys Res. 17: D04302. doi:10.1029/2011JD016573.

Johns R and Doswell CA III (1992) Severe local storms forecasting. Wea Forecasting 7:588-561. doi: http://dx.doi.org/10.1175/1520-0434(1992)007<0588:SLSF>2.0.CO;2.

Jury MR and Pathack B (1991) A study of climate and weather variability over the tropical southwestern Indian Ocean. Meteorol Atmos Phys 47:37–48.

Khashei M and Bijari H (2012) Exchange rate forecasting better with hybrid artificial neural networks models. J Math Comput Sci 1:103–125.

Kohonen T (1989) Self-organization and associative memory. Springer-Verlag, London

Kohonen T (2001) Self-organizing maps, 3rd edn. Springer, London.

Kruger AC (2004) Climate of South Africa. Climate Regions. South African Weather Service WS45:19.

Lennard C and Hegerl G (2014) Relating changes in synoptic circulation to the surface rainfall response using self-organising maps. Clim Dynamics. doi: 10.1007/s00382-014-2169-6.

Liu Y and Weisberg RH (2011) A review of Self-Organizing Map applications in meteorology and oceanography. In: Mwasiagi JI, Self-Organizing Maps - Applications and novel algorithm design. Rijeka, pp 253–272.

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. Int J Climatol 32:1018–1032. doi: 10.1002/joc.2320.

Malherbe J, Landman WA and Engelbrecht FA (2013) The bi-decadal rainfall cycle, Southern Annular Mode and tropical cyclones over the Limpopo River Basin, southern Africa. Clim Dynamics. 42:3121-3138. doi 10.1007/s00382-013-2027-y.

Meteo France (2012) URL: http://www.meteo.fr/temps/domtom/La_Reunion/meteoreunion2/#. Accessed 22 August 2012.

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. Atmos Research. 83:185–200. doi: 10.1016/j.atmosres.2005.10.015.

Reason CJC, Landman W and Tennant W (2006) Seasonal to decadal prediction of southern African climate and its links with variability of the Atlantic Ocean. Bull Am Meteorol Soc 87:941–955. doi: http://dx.doi.org/10.1175/BAMS-87-7-941.

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. doi: 10.1080/789610199.

Rouault M, Roy SS and Balling RC (2013) The diurnal cycle of rainfall in South Africa in the austral summer. Int J Climatol 33:770–777. doi: 10.1002/joc.3451.

Seefeldt MW and Cassano JJ (2008) An analysis of low-level jets in the greater Ross Ice Shelf region based on numerical simulations. Mon Weather Rev 136:4188–4205. doi: http://dx.doi.org/10.1175/2008MWR2455.1.

Schultz DM, Schumacher PN and Doswell CA III (2000) The intricacies of instabilities. Mon Weather Rev 128:4143–4148. doi: http://dx.doi.org/10.1175/1520-0493(2000)129<4143:TIOI> 2.0.CO;2.

Simon G, Lendasse A, Cottrell M, Fort JC and Verleysen M (2005) Time series forecasting: Obtaining longterm trends with self-organizing maps. Pattern Recognit Lett 26:1795–1808. doi: 10.1016/j.patrec.2005.03.002.

Singleton AT and Reason CJC (2007) A numerical model study of an intense cut-off low pressure system over South Africa. Mon Weather Rev 135:1128–1150. doi: http://dx.doi.org/10.1175/MWR3311.1.

Tadross MA, Hewitson BC and Usman MT (2005). The Interannual variability of the onset of the maize growing season over South Africa and Zimbabwe. J of Clim 18:3356-3372. doi: http://dx.doi.org/10.1175/JCLI3423.1.

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 96:2907–2912.

Taljaard JJ (1985) Cut-off lows in the South African region. South Africa. S Afr Weather Bur Technical Paper No 14:153.

Taljaard JJ (1996) Atmospheric circulation systems, synoptic climatology and weather phenomena of South Africa. Part 6: Rainfall in South Africa. S Afr Weather Bur Technical Paper No 32:89.

Tennant W and Hewitson BC (2002) Intra-seasonal rainfall characteristics and their importance to the seasonal prediction problem. Int J Climatol 22:1033–1048. doi: 10.1002/joc.778.

Triegaardt DO, Van Heerden J and Steyn PCL (1991) Anomalous precipitation and floods during February 1988. S Afr Weather Bur Technical Paper No 23:25.

Tyson P and Preston-White RA (2000) The weather and climate of South Africa. Oxford University Press, Johannesburg.

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 4:307–318. doi: 10.1016/S0968-090X(97)82903-8.

Van Schalkwyk L and Dyson LL (2013) Climatological characteristics of fog at Cape Town International Airport. Wea Forecasting. 28:631–646. doi: 10.1175/WAF-D-12-00028.1.

Wu X and Yanai M (1994) Effects of vertical wind shear on the cumulus transport of momentum: Observations and parameterization. J Atmos Science 51:1640–1660. doi: http://dx.doi.org/10.1175/1520-0469(1994)051<1640:EOVWSO>2.0.CO;2.