INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. (2010)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/joc.2266

RMetS
Royal Meteorological Society

Recurrent daily rainfall patterns over South Africa and associated dynamics during the core of the austral summer

Julien Crétat,[a]* Yves Richard,[a] Benjamin Pohl,[a] Mathieu Rouault,[b] Chris Reason[b] and Nicolas Fauchereau[c]

[a] *Centre de Recherches de Climatologie, CNRS UMR 5210/Université de Bourgogne, Dijon, France*
[b] *Department of Oceanography, University of Cape Town, Cape Town, South Africa*
[c] *CSIR/NRE, Ocean Systems and Climate, CHPC, Rosebank, South Africa*

ABSTRACT: This paper investigates the influence of some modes of climate variability on the spatio-temporal rainfall variability over South Africa during the core of the rainy season, December to February (DJF). All analyses are based directly on the rainfall field instead of atmospheric processes and dynamics. An original agglomerative hierarchical clustering approach is used to classify daily rainfall patterns recorded at 5352 stations from DJF 1971 to DJF 1999. Five clusters are retained for analysis. Amongst them, one cluster looks most like the rainfall and circulation mean picture. Another one, representing 37% of the days, describes strong negative rainfall anomalies over South Africa resulting from a regional barotropic trough-ridge-trough wave structure and moisture divergence. These dry anomalies are more frequent in El Niño years (with a 0.75 correlation between ENSO and the seasonal occurrences of that cluster) and are associated with a weakening and/or a regional shift of the Walker-type circulation and the subtropical jet stream. Three clusters, which describe, respectively, a decrease, an increase, and a conservation of the seasonal mean eastward rainfall gradient, are all associated with strong cyclonic moisture flux anomalies centred over Angola or northern Botswana and the anomalous convergence located over the largest positive rainfall anomalies, suggesting the importance of thermal low pressures. One of them is clearly associated with tropical temperate troughs. The rainfall patterns of these three clusters tend to propagate eastward, and their seasonal frequencies tend to increase in La Niña years.

KEY WORDS atmospheric dynamics; daily rainfall patterns; ENSO; hierarchical clustering; South Africa

Received 10 November 2009; Revised 19 October 2010; Accepted 25 October 2010

1. Introduction

Southern Africa (SA), south of 15 °S, is a semi-arid region mainly experiencing its larger rainfall amounts in austral summer between November and February. Because of the predominance of rain-fed agriculture over SA, large departures from the average seasonal cycle (either floods or droughts) may have detrimental effects on the economies and societies of the region. The southern African rainfall field is known to show strong spatio-temporal variability at different scales, materialising the influence of distinct rain-bearing processes that themselves depend on various modes of atmospheric variability.

At the global scale, El Niño Southern Oscillation (ENSO) (Lindesay, 1988; Jury *et al*., 1994; Reason *et al*., 2000; Reason and Rouault, 2002; Rouault and Richard, 2003, 2005) is known to impact South African rainfall variability, especially since the late 1970s (Richard *et al*., 2000, 2001). As ENSO signal is phase locked to the annual cycle over most parts of SA (e.g. Camberlin *et al*., 2001), the peak of ENSO-associated atmospheric, oceanic, and rainfall anomalies typically occurs in austral summer (Nicholson and Selato, 2000; Reason *et al*., 2000), i.e. the main rainy season over the region. In tropical SA, wet (dry) conditions tend to prevail in La Niña (El Niño) years during austral summer. Nevertheless, correlations values remain generally weak (typically 0.4–0.5) and some El Niño years, like the strong 1997–1998 event, are not accompanied by particularly dry conditions.

At a more regional scale, interannual rainfall fluctuations were also shown to be modulated by the Sea Surface Temperature (SST) variations in the south Atlantic and south Indian Oceans (Mason, 1995; Reason and Mulenga, 1999; Reason, 2001, 2002; Rouault *et al*., 2003). South Atlantic SST partially influence moisture fluxes between the south Atlantic Ocean and SA (Vigaud *et al*., 2007, 2009). The Indian Ocean is known to be the major source of moisture for SA (Jury *et al*., 1996; Reason, 2001, 2002; Rouault *et al*., 2003; Washington and Preston, 2006). For instance, positive SST anomalies in the southwest Indian Ocean are known to be associated with wetter conditions over eastern and central South

* Correspondence to: Julien Crétat, Centre de Recherches de Climatologie, UMR 5210 CNRS/Université de Bourgogne, 6 Bd Gabriel, 21000 Dijon, France. E-mail: julien.cretat@u-bourgogne.fr

Africa (Reason, 2001, 2002; Washington and Preston, 2006).

The high-frequency variability of South African rainfall is mainly related to the so-called tropical-temperate troughs (TTT), corresponding to synoptic-scale cloud bands that link a tropical disturbance over the subcontinent (typically, a lower-layer easterly wave) with an upper-tropospheric frontal system embedded in the mid-latitude westerly circulation (Washington and Todd, 1999; Todd and Washington, 1999; Todd *et al*., 2004; Hart *et al*., 2010). Although about 30–60% of summer rainfall over subtropical SA results from such TTT systems, the factors influencing the rainfall spatial and temporal variability are very complex and still largely matter of debate. The remaining 40–70% of the summer rainfall amounts are provided by rain-bearing mechanisms mainly linked with tropical convection (Harrison, 1984a, 1984b), such as regional thermal low pressures (Tyson and Preston-Whyte, 2000; Reason *et al*., 2006) or the Madden-Julian Oscillation (MJO) (Pohl *et al*., 2007), but these mechanisms remain poorly misunderstood.

Recently, Fauchereau *et al*. (2009, hereafter F09) and Pohl *et al*. (2009, hereafter P09) investigated the interactions between TTT systems, the MJO and ENSO, by applying clustering analyses on daily Outgoing Longwave Radiation (OLR), a proxy for deep atmospheric convection. They showed that (1) TTT events are the main rain-bearing system in austral summer and tend to propagate eastwards with time, and (2) MJO activity in the tropics does not influence their probability of occurrence. They also confirmed the relative weakness of linear correlations between usual ENSO descriptors and South African rainfall, suggesting the implication of non-linear processes.

The respective influence of all these mechanisms on the rainfall spatio-temporal variability remains to be assessed and quantified. This paper contributes to investigate this issue, by using a method based on the analysis of the daily spatial distributions of South African summer rainfall instead of atmospheric processes and dynamics. Our goal here is to verify whether a clustering technique applied to a relatively dense network of daily rain-gauge records allows us to detect the recurrent gradients linked with the main modes of South African climate variability, and to quantify them. To that end, an objective hierarchical clustering technique is applied to daily rainfall amounts recorded in a network covering the southern part of SA for the 1971–1999 period.

The paper is organised as follows. Section 2 presents the data and methodology used in this study. Sections 3 and 4 describe, respectively, the detected anomalous rainfall patterns, their persistence and evolution, and the associated atmospheric dynamics at regional and near-global scales. Section 5 investigates the teleconnections between the interannual frequency of the rainfall patterns and SST patterns. Section 6 contains the summary and discussion.

2. Data and method

2.1. Daily rain-gauge records

Daily rainfall amounts are provided at high resolution by the rain-gauge records compiled in the Water Research Commission database by Lynch (2003). We have extracted the 5352 stations receiving at least 40% of the annual rainfall during the DJF season from the total of 7665 stations already used in Pohl *et al*. (2007). Thus, we only focus on the summer rainfall regions because this empirical threshold excludes the winter rainfall (western Cape region of South Africa) or the all-year long rainfall (south coast of South Africa) regions. Over the 1971–1999 DJF period, there are 29 summer (DJF) seasons representing 2610 days, without missing values. We have also excluded each 29th of February.

2.2. Atmospheric dynamics

We make use of the 12-hourly (00 UTC and 12 UTC) ERA-40 reanalyses (Uppala *et al*., 2005) for the 1971–1999 DJF period to describe the atmospheric circulation associated with the recurrent rainfall patterns. They are provided by the European Center for Medium-Range Weather Forecasts (ECMWF) on a $2.5^\circ \times 2.5^\circ$ regular grid.

Within tropical and subtropical areas of southern and central Africa, rainfall regimes are largely dependent on deep convective processes and moisture convergence at different levels of the troposphere. The zonal (U) and meridional (V) components of the wind and specific humidity (Q) are used to compute daily moisture fluxes, vertically integrated between 1000 and 200 hPa. This proxy effectively captures the moisture transport in the atmosphere. We also compute moisture convergence from the integrated zonal and meridional daily moisture fluxes. The 500 hPa vertical velocity (W) describes the approximate centre of mass of the troposphere and allows for an insight into large-scale vertical motion within the whole troposphere, indicative of deep convective upward motion.

At the near-global scale, we examine the large-scale components of the rotational and divergent 850 and 200 hPa wind, depicted by the stream function (Ψ) and velocity potential (Φ). These fields are respectively defined as $\nabla^2\Psi = \nabla \times \nu$ and $\nabla^2\Phi = \nabla \times \nu$, where ν is the two-dimensional wind vector (U and V). These fields are used to describe the large-scale circulations such as the zonal Walker circulation in the tropics and the Hadley cell.

2.3. ENSO index and sea surface temperature

The Multivariate ENSO Index (MEI: Wolter and Timlin, 1993) is used to measure the relationship between the interannual frequency of each cluster and ENSO. This index is provided by the Climate Diagnostic Center, and is available at the bimonthly timescale. As it is based on both atmospheric and oceanic fields, it is particularly suitable to describe the coupled nature of

the ENSO phenomenon. In order to evaluate to what extent the results are sensitive to the choice of the ENSO descriptors, we correlate the MEI with more traditional indices, such as the Southern Oscillation index and the Niño3.4 index. Interannual DJF correlation coefficients reach −0.93 and 0.98, respectively, indicating moderate uncertainties associated with this parameter. These small differences could be partly explained by the presence of positive trends in SST indices, like Niño3.4, while none is significant for atmospheric indices, such as the SOI.

Monthly SST provided by the United Kingdom Meteorological Office (HadISST: Rayner *et al*., 2003) are used to document the background climate conditions favouring the seasonal occurrences of the successive rainfall patterns. They are available on a $1^\circ \times 1^\circ$ regular grid.

2.4. Clustering technique applied on daily rain-gauge records

The agglomerative hierarchical clustering (AHC) has been successfully employed with rainfall data (e.g. Ramos, 2001; Tennant and Hewitson, 2002; Muñoz-Diaz and Rodrigo, 2004, 2006). Here, we use the AHC to classify the raw daily rainfall patterns over the Republic of South Africa as well as neighbouring Lesotho, Swaziland and southern Namibia. This procedure creates a nested sequence of partitions of the patterns from a dissimilarity matrix, and proceeds by series of fusions of the n objects (the 2610 days here) into groups (Gong and Richman, 1995). The AHC produces a series of partitions of the data, $P_n, P_{n-1}, \ldots, P_1$. In this study, P_n consists of 2610 single object clusters, and P_1 consists of a single group containing the 2610 days. At each stage, the AHC regroups the two clusters that are closest according to a Euclidean distance metrics. Consider a matrix X containing 2610 days × 5352 rainfall stations, with *days* describing the time dimension (observations) and *rainfall stations* describing the space dimension (variables), then the Euclidean distance between entities X_{day1} and X_{day2} is given by:

$$d_{day1,day2} = \sum_{sta=1}^{5352} [(X_{day1} - X_{day2})'(X_{day1} - X_{day2})]^{1/2}. \quad (1)$$

For merging two clusters including more than one day, or a cluster of one day with a cluster including more than one day, we use Ward's method (Ward, 1963) because it minimises the within-group variance at each stage of merging and optimises an objective statistic. At each step, the intra-cluster sum of squares is minimised over all partitions obtainable by merging two clusters from the previous step. If C_K and C_L are two clusters that are merged to form cluster C_M, the combinatorial formula that defines the distance between the new cluster and another cluster C_J is:

$$d_{J,M} = \frac{(n_J + n_K)d_{JK} + (n_J + n_L)d_{JL} - n_J d_{KL}}{n_J + n_M}, \quad (2)$$

where n_J, n_K, n_L and n_M are the number of objects in clusters J, K, L and M, and d_{JK}, d_{JL}, d_{KL} are the distances between *JK, JL* and *KL*, respectively (Ramos, 2001). Thus, Ward's algorithm can be implemented through updating a stored Euclidean distance between cluster centroids.

To sum up, the AHC procedure enables us to obtain partitions of rainfall patterns based on objective distance computations. Anomalous rainfall (dynamic) patterns are computed by subtracting the mean seasonal rainfall (dynamic) field from cluster averages.

3. Results of the daily rainfall rain-gauge clustering

3.1. Description of the clusters

At the end of the AHC procedure, i.e. once the 2610 days are regrouped into a single cluster, the number of clusters is subjectively chosen according to a clustering tree, also known as dendrogram (Figure 1(a)). A clustering tree illustrates the fusions made at each successive stage of the partitioning. To limit intra-cluster heterogeneity, one needs to choose the number of clusters just before the distance between merged clusters increases strongly. Basing on Figure 1(a), we choose here to retain six clusters. In order to validate their spatial and temporal robustness, we applied the AHC method to 40, 50 and 60% of the 5352 rain-gauge stations, randomly chosen and on the entire period. We also applied the AHC to all rain-gauge stations for the separated 1971–1984 and 1985–1999 periods. It appears that the number of days spent in cluster #2 varies noticeably compared to the other clusters, while cluster #1 is almost stationary. The rainfall patterns associated with the six clusters appears, nonetheless, to be relatively stable, which indicates a relatively high robustness.

The 2610 days are distributed unequally within the six clusters (Figure 1(a)). This seems in accordance with the temporal rainfall distribution of semi-arid regions where dry or light rainfall days predominate. For our purposes, this is an advantage of the AHC compared to non-hierarchical clustering techniques such as the *k-means* algorithm which tends to generate equiprobable clusters.

Before analysing the rainfall patterns, we briefly compare the six clusters by examining their internal heterogeneity. To this end, we compute the spatial average of each day spent in each cluster (Figure 1(b)). For cluster #1 (the most robust cluster in terms of sensitivity to the partitioning), 80% of the 2610 days are rather dry with an average of 0.86 mm/day, *versus* 3.16 mm/day for cluster #2 (the most sensitive cluster). Other clusters are wetter. Rainfall averages reach 6 mm/day for clusters #3 and #4, 19 mm/day for cluster #5 and 9 mm/day for cluster #6. Due to spatial averaging, there are few days without rainfall (observed rarely in cluster #1) and extreme rainfall events are smoothed (the maximum value does not exceed 25 mm/day; cluster #5). Figure 1(b) also shows that the spatial averages have non-Gaussian distributions and are

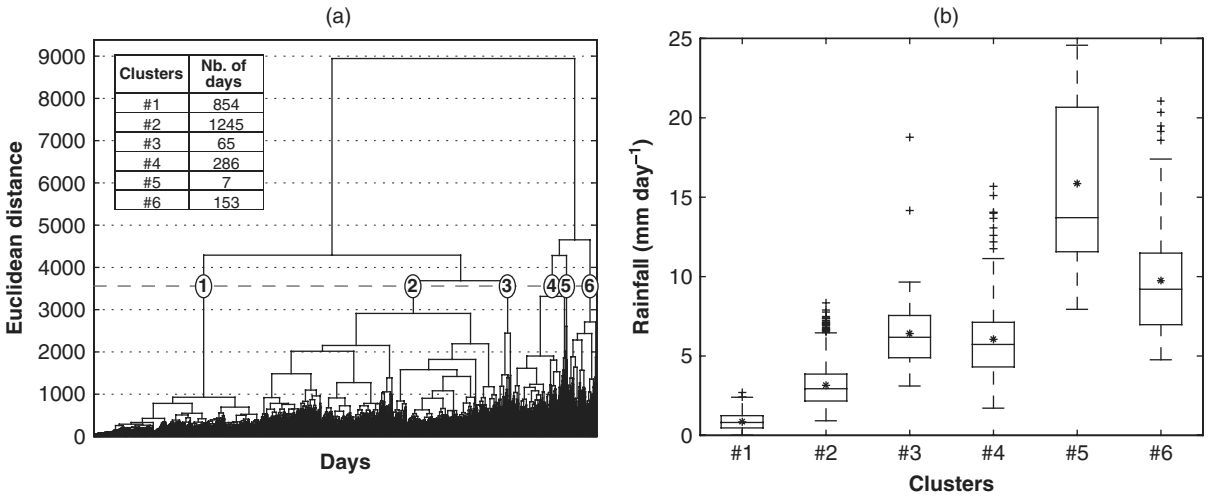

Figure 1. (a) Clustering tree of 2610 days for the 1971–1999 DJF period. Numbers on tree branches correspond to clusters. The table gives the size of each cluster. (b) Box plots of daily spatial mean of the 5352 rainfall stations for the six clusters. The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers are lines extending from each end of the boxes to show the extent of the range of the data within 1.5 by inter-quartile range from the upper and lower quartiles. Stars are mean values. Outliers are displayed by a 'plus' sign.

skewed towards extreme positive values. Mean values are higher than median values, except for cluster #1.

The distributions of the spatial averages attributed to each cluster have a very large variance which may, however, be related to an exactly similar pattern but shifted in amplitude. Therefore, it does not restrict our study the aim of which is to focus on rainfall gradients. Cluster #5, which involves only 7 days and has too-skewed a distribution of the spatial mean of daily rainfall, is not analysed furthermore.

3.2. Recurrent daily rainfall patterns

Figure 2(a) displays the seasonal mean rainfall of the 5352 rain-gauge stations for the 1971–1999 DJF period. Except for the desert areas, the west–east rainfall gradient is typical of most subtropical continents. It results essentially from the presence of semi-permanent anticyclones over the neighbouring oceans and associated SST contrasts. Regarding SA, the south Atlantic or St Helena anticyclone located off the west coast limits incoming moisture fluxes, while the Mascarene High or south Indian Ocean anticyclone off the east coast favours moisture fluxes towards SA. The regional topography modifies such moisture fluxes (Van Heerden and Taljaard, 1998; Tyson and Preston-Whyte, 2000), while a heat low over southern Angola/northern Namibia facilitates the low-level penetration of moisture from the tropical southeast Atlantic, north of 10 °S (Rouault *et al.*, 2003; Reason *et al.*, 2006). However, the western Indian Ocean remains the major moisture source for most of SA (Vigaud, 2007).

Figure 2(b)–(f) presents the significant composite anomalies associated with each cluster. For cluster #2, only 60% of the retained rain-gauge stations have significant composite anomalies (Figure 2(c)) and the anomalies are weak. They do not exceed −3 mm/day in the northeast and +5 mm/day elsewhere. Hence, the rainfall pattern of cluster #2 is close to the DJF climatology, except abnormally dry conditions in the Lowveld (northeastern part of South Africa). By contrast, composite rainfall anomalies associated with clusters #1 (Figure 2(b)) and #6 (Figure 2(f)) are significant over almost all the domain. They are strongly negative for cluster #1, especially in the east, resulting in generalised dryness and reduced climatological rainfall gradient. On the other hand, they are positive for cluster #6, reaching up to +20 mm/day north of Lesotho, but rarely exceed +5 mm/day elsewhere. Composite anomalies associated with clusters #3 and #4 (Figure 2(d)–(e)) are strong and mainly significant when they are positive. Cluster #3 depicts positive rainfall anomalies in central South Africa and negative rainfall anomalies in the northeast. Cluster #4 is associated with positive rainfall anomalies in the east and negative rainfall anomalies in the west. Compared to the DJF climatology rainfall values (Figure 2(a)), cluster #3 (#4) thus describes a weakening (enhancement) of the seasonal mean eastward gradient.

3.3. Persistence and evolution of the clusters

Table I presents transitions between the clusters, i.e. the possible lead/lag relationships between clusters. The high percentages observed on the diagonal give an indication of the first-order persistence of each cluster. Figure 3 displays the persistence of each sequence affiliated to each cluster. More than 60% of clusters #1 and #2 are followed by themselves, denoting their strong persistence. Clusters #1 and #2 both last more than 2 days in 50% of the cases, and more than 4 days and 3 days, respectively, in 25% of the cases. This is different from the remaining clusters, for which rainfall events are less persistent, with a length of 1 day in 50% of the cases and rarely higher than 2 days, suggesting transient patterns. Table I supports this hypothesis and indicates a possible west-to-east propagation of enhanced rainfall related to these clusters (Figure 2): in 30% of the cases, cluster #3 is

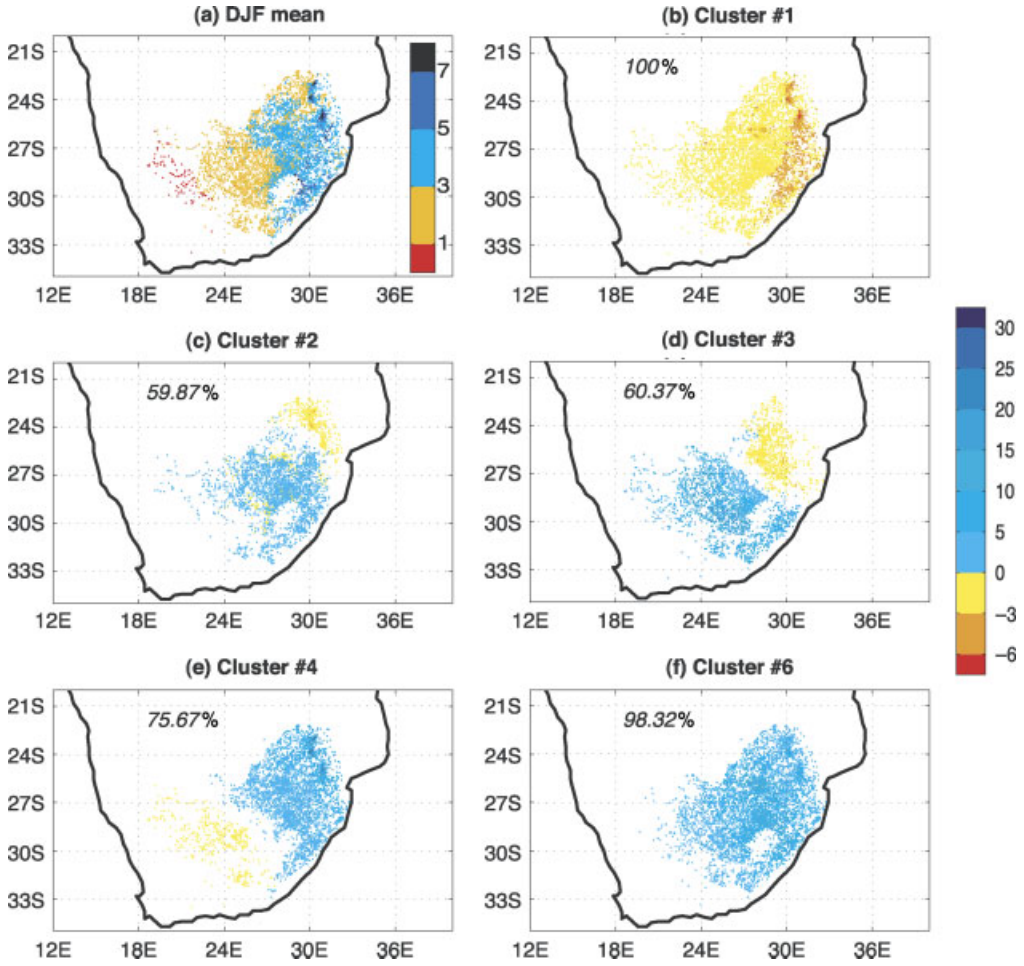

Figure 2. (a) Daily mean of the 5352 rainfall stations for the 1971–1999 DJF period. (b)–(f) Significant composite anomalies associated with clusters #1, #2, #3, #4, and #6 over the same period at the 95% confidence level using a t-test applied to the square root of rainfall time series in order to be obtain Gaussian distributions. The part (%) of stations showing significant rainfall anomalies during the occurrences of the cluster are labelled on the figure. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

Table I. Percentage of days spent in each cluster followed by the same or another cluster. Values in bold correspond to positive and significant associations between observed and expected frequencies of clusters according to a Pearson chi-square test at the 95% confidence level.

	Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #6
Cluster #1	**64.9**	32.5	0.5	1.5	0.3
Cluster #2	20	**63.2**	2.5	9.5	4.7
Cluster #3	3.1	28.5	**28.5**	9.5	**30.1**
Cluster #4	14.9	34.8	2.1	**40.5**	7.4
Cluster #6	3.2	39.4	3.2	**21**	**32.8**

followed by cluster #6, and for 21% of the cases, cluster #6 is followed by cluster #4. This result will be further discussed below.

4. Associated atmospheric circulation anomalies

In this section, we first describe the atmospheric circulation at the regional scale and try to identify associated physical mechanisms. At this scale, moisture fluxes and descriptors of deep convective activity are the most relevant variables for regional rainfall variability. Then, we focus on the near-global scale to assess whether there are any coherent changes in the Walker-type or Hadley circulations. To that end, stream function and velocity potential are used to describe large-scale vorticity and divergence, respectively. In this section, anomalous dynamic patterns are computed by subtracting the mean DJF field from cluster averages.

4.1. Regional scale

Composite anomalies of moisture fluxes associated with cluster #1 extend from the tropics to the mid-latitudes (Figure 4(a)). They are strongest south of 15 °S where a well defined west-to-east trough-ridge-trough wave structure is evident. This pattern is also apparent in the geopotential height anomalies at all vertical levels (not shown). Moisture flux anomalies are anticyclonic over most of the subcontinent, inducing relative moisture divergence in most areas (Figure 4(a)) and middle-tropospheric subsidence (Figure 5(a)). Anomalous moisture convergence occurs near the two troughs

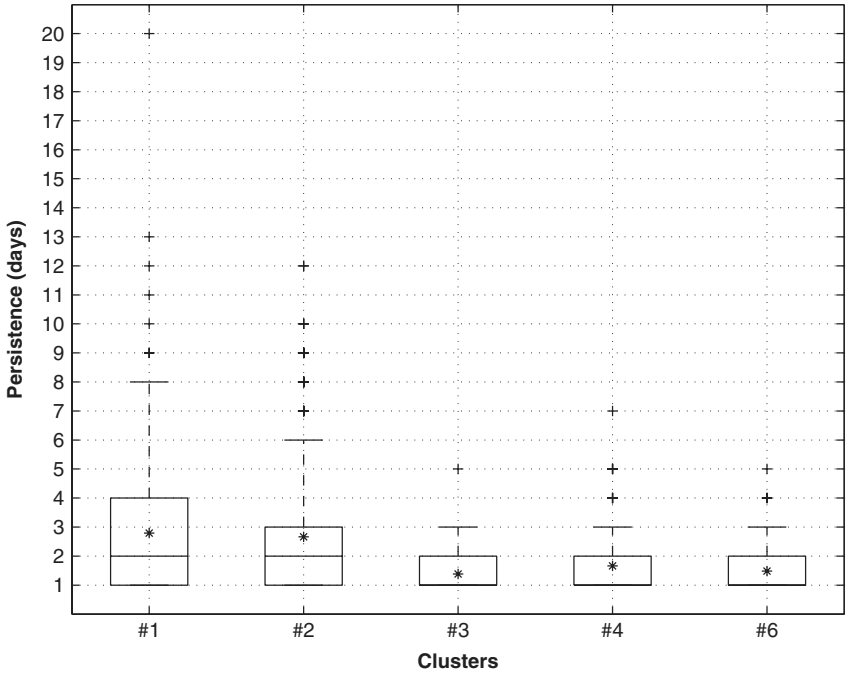

Figure 3. As Figure 1(b), but for the persistence of each sequence associated with each cluster.

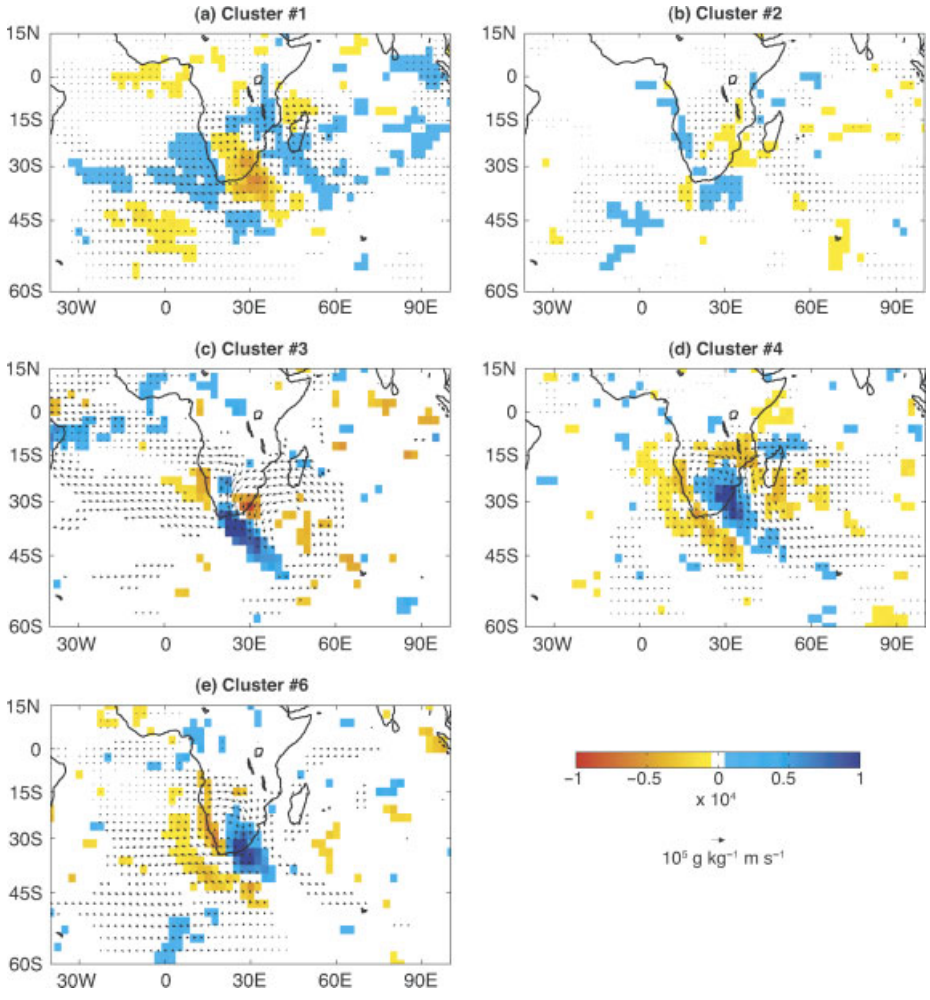

Figure 4. Composite anomalies of moisture convergence and fluxes integrated between 1000 hPa and 200 hPa ($g.kg^{-1}$ $m.s^{-1}$) for each cluster. Cold (hot) colors display moisture convergence (divergence) anomalies. Only anomalies of convergence (moisture fluxes) that are significant at the 95% level according to a Student two tailed *t*-test (Hotelling test) are represented. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

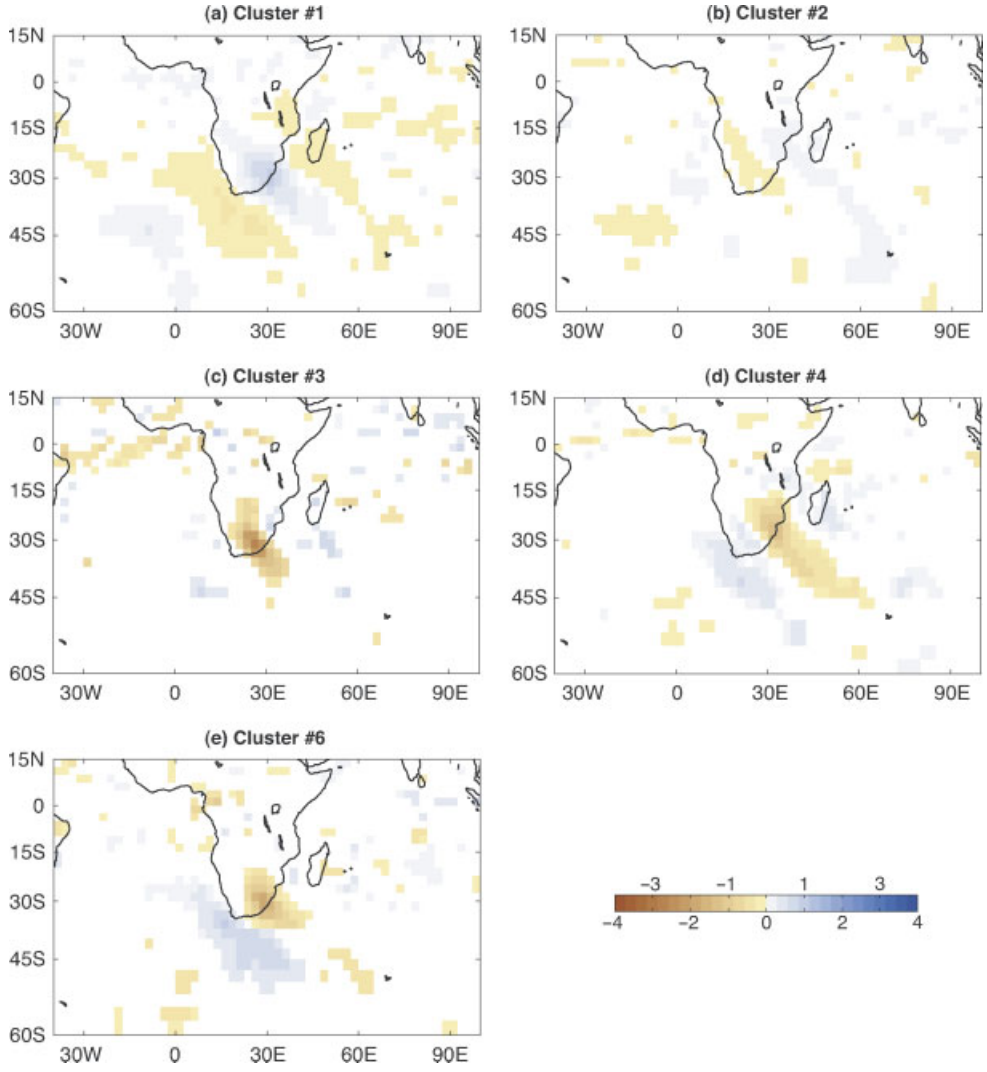

Figure 5. Composite anomalies of mid-troposphere (500 hPa) vertical velocity (m s^{-1}) associated with each cluster. Significance tested and shown as for Figure 4. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

(Figure 4(a)) together with anomalous ascending motion over the neighbouring southeast Atlantic and southwest Indian Oceans (Figure 5(a)). The wave structure denotes a weakening of the mean moisture flux originating from the north (both the Mozambique Channel and the Congo basin). Thus, this configuration favours low water vapor amounts in the air mass over SA, reduces tropical moisture advections, and acts to suppress moist convection. Consequently, the likelihood of mid-latitude disturbances linking up to form a TTT over South Africa is also reduced, as well as more regional features such as thermal low pressures, which is consistent with the generalised negative rainfall anomalies associated with cluster #1 (Figure 2(b)).

Moisture flux anomalies associated with cluster #2 (Figure 4(b)) are of weak amplitude and barely significant. A northwest–southeast zone of anomalous convergence extends from eastern Angola towards central South Africa, in association with anticyclonic moisture flux anomalies. Anomalous uplift at 500 hPa (Figure 5(b)) extends southeastwards across SA with a zone of relative subsidence extending from Zimbabwe to the Prince Edward Islands. These rather weak atmospheric anomalies are consistent with those in the rainfall field (Figure 2(a)–(c)).

Composite anomalies associated with clusters #3, #4, and #6 (Figure 4(c)–(e)) are relatively strong and extend from the tropics to the mid-latitudes in each case. For cluster #3, a large anticyclonic moisture flux anomaly is found over the southwest Indian Ocean and southeastern Africa, suggesting an enhancement of the mean moisture advection originating from the subtropical south Indian Ocean. Further west, cyclonic anomalies take place over the south Atlantic Ocean and southwestern Africa with a stronger Angola heat low (Mulenga, 1998; Reason *et al*., 2006) in the lower troposphere (found in 850 hPa wind and geopotential height anomalies, not shown). This pattern leads to an increase in low-level moisture convergence over central SA (Figure 4(c)), associated with deep convection over South Africa as depicted by vertical velocity at 500 hPa (Figure 5(c)). Clusters #4 (Figure 4(d)) and #6 (Figure 4(e)) are associated with

a wave structure oriented along a northwest–southeast direction. This structure consists of a well defined ridge-trough-ridge structure with cyclonic wind anomalies over the landmass (not shown) and anticyclonic moisture flux anomalies over the neighbouring oceans. Over South Africa, the cyclonic anomalies are linked with moisture convergence (Figure 4(d)–(e)). This wave structure is more or less opposite to that associated with cluster #1 (Figure 4(a)). Anomalies in the OLR field (not shown) indicate enhanced atmospheric convection along the trough. The wave structure is also noted in the mid-tropospheric vertical velocity (Figure 5(d)–(f)). Clusters #4 and, to a lesser extent #6, seem to be associated with a trough that extends roughly from Botswana (24 °S, 25 °E) to the Prince Edward Islands (46 °S, 38 °E) and which is apparent throughout the troposphere. Both these clusters, particularly cluster #4, are hypothesised to reflect TTT activity over SA. The mean duration found for these clusters (Figure 3) matches that of the TTT described in F09 (i.e. between 3 and 6 days).

The main common feature between clusters #4 and #6 is the cyclonic moisture flux anomaly cell, centred over Botswana, which acts to reduce the export of moisture towards Namibia. It also leads to moisture convergence over Mozambique and inputs of moisture into the trough. The main difference consists of a generalised eastward shift of the atmospheric patterns. For instance, anticyclonic anomalies cover almost the whole south Atlantic Ocean for cluster #6 but are confined to the southeastern Atlantic and the southwest Indian Ocean for cluster #4, where they are much stronger. This difference is also noted in the 500 hPa vertical velocity anomalies (Figure 5(e)). Thus, cluster #6 is less obviously a TTT feature than cluster #4, or it could represent its decaying phase in accordance with, e.g., F09.

4.2. Near-global scale

Cluster #2 shows very weak 850 hPa and 200 hPa stream function and velocity potential anomalies (Figures 6(b)–(g) and 7(b)–(g), respectively). Its main characteristics involve large-scale convergence (divergence) over the Indian Ocean in the lower (upper) layers (Figure 7(a)–(f)). Velocity potential anomalies are not significant elsewhere. This suggests little modification of the Hadley and Walker circulations.

Over South Africa and the surrounding oceans, stream function anomalies of cluster #1 present lower-layer patterns that are relatively similar to moisture fluxes anomalies (Figure 4(a)), and that show a well-defined trough-ridge-trough wave structure (Figure 6(a)). The latter is still perceptible in the upper layers (Figure 6(f)), though with weaker values. Except for this near-barotropic structure, lower- and upper-tropospheric anomalies are of reversed sign, indicative of baroclinicity. At 850 hPa, cyclonic (anticyclonic) prominent features describe anticyclonic (cyclonic) anomalies over Australia (the Philippines) together with cyclonic (anticyclonic) anomalies in the southern tropical Pacific (north of the equator and south of 60 °S). At 200 hPa, clear cyclonic anomalies continuously prevail in the Southern Hemisphere between the subtropics and the mid-latitudes, while more regionalised anticyclonic anomalies are found in the northern Tropics. Concomitant velocity potential anomalies show regional scale divergence (convergence) in the lower (upper) troposphere, consistent with the generalised dryness occurring in SA (Figure 2(b)). Large-scale patterns show upper-tropospheric divergence (convergence) over the Indian Ocean basin and the Americas (the Maritime Continent and Africa). The reverse occurs in the lower layers. These wavenumber-2 circulation anomalies demonstrate that the rainfall anomalies associated with cluster #1 are embedded in large-scale climate variability patterns. In particular, these results suggest that this cluster is related to large-scale weakening/reversal of the Walker circulation, suggestive of a possible connection with El Niño events.

Like cluster #1, the core of the anomalous stream function associated with clusters #3, 4 and 6 (Figure 6(c), (d) and (e)) is stronger in the lower troposphere. The spatial extension of the anomalies is however generally wider in the upper layers (Figure 6(h), (i) and (j)), where they are mainly confined to the inter-tropical zone and the southern mid-latitudes. Anomalies of stream function and velocity potential patterns associated with these three clusters are more or less reversed to those of cluster #1 (Figures 6(h), (i) and (j)) and (Figure 7(h), (i) and (j), respectively), especially over SA. Velocity potential anomalies are both weaker and spatially more limited. Those associated with cluster #6 are compatible with a possible increase of the Walker-type circulation. It should nonetheless be noted that the strongest signals associated with these cluster are clearly regional by nature. They consist of lower (upper) layer convergence (divergence), consistent with associated wet anomalies over SA (Figure 2(c) and (f)). This makes the possible involvement of ENSO uncertain for these clusters.

5. SST teleconnections and possible ENSO connections

This section aims to quantify to what extent the seasonal occurrences of each rainfall pattern is modulated by the state of ENSO and/or more regional SST anomalies.

Figure 8(a) and (e) shows linear correlations between seasonal occurrences of each cluster and synchronous seasonal SST. In addition, Table II shows linear correlations between the clusters seasonal frequencies/mean durations and the MEI. For cluster #1, correlation patterns (Figure 8(a)) display strong positive (negative) correlations in the eastern (western) Pacific and positive correlations over the tropical Indian Ocean. Correlation coefficients between this cluster and the MEI reach +0.75 (for seasonal occurrences) and +0.4 (for the spell duration), shown in Table II. El Niño (La Niña) events tend to be associated with an increase (decrease) of the occurrence of cluster #1, and less clearly with an increase

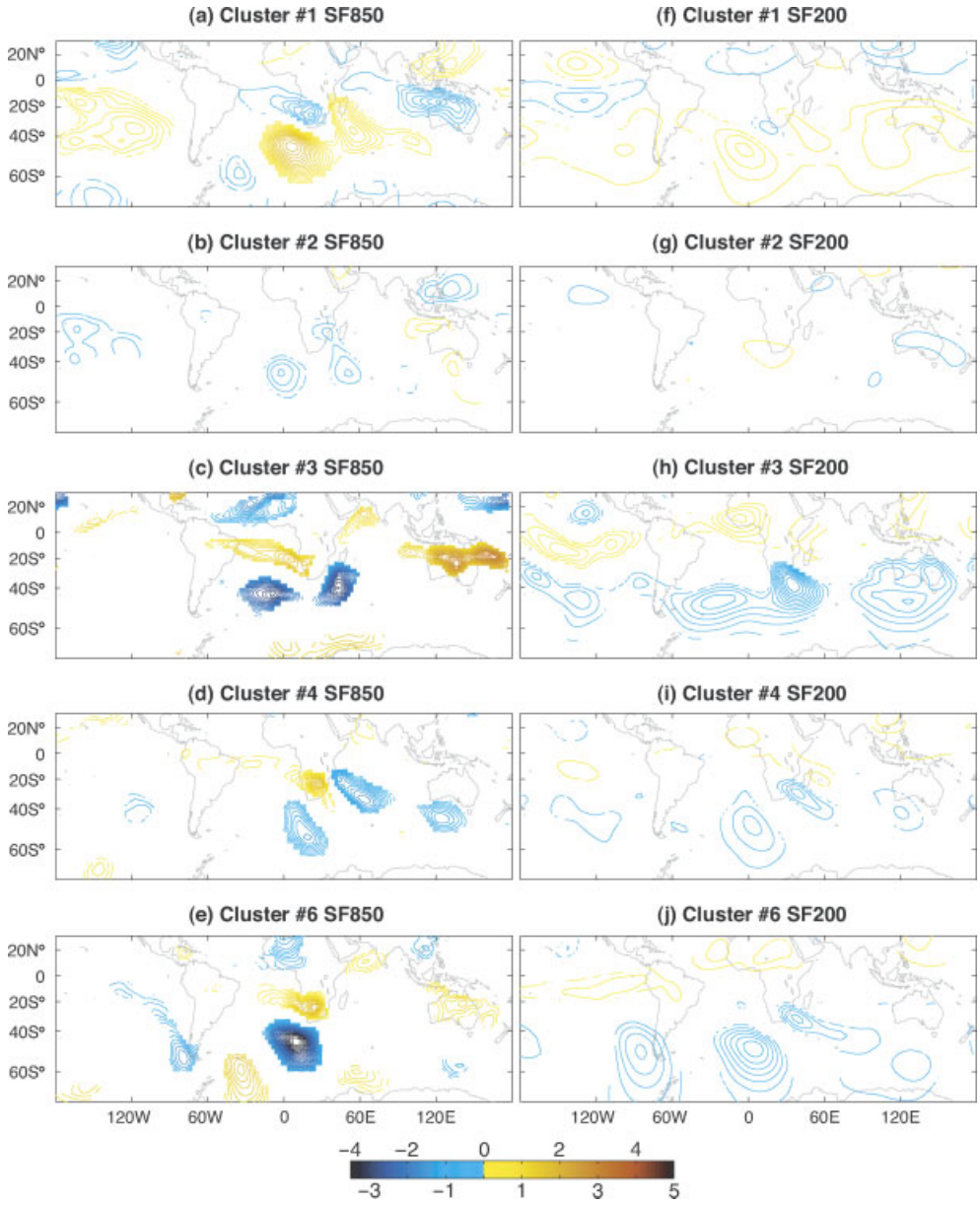

Figure 6. Composite anomalies of stream function ($m^2\ s^{-1}$) associated with each cluster at 850 hPa (left hand column) and 200 hPa (right hand column). Only anomalies significant at the 95% level according to a Student's two-tailed t-test are shown. Contour interval is $1 \times 10^6\ m^2\ s^{-1}$. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

(decrease) in the persistence of generalised dry days over South Africa. Clusters #3–4 (Figure 8(c)–(d)), and to lesser extent #2–6 (Figure 8(b)–(e)), display opposite patterns in the Pacific basin and more contrasted correlation values over the Indian Ocean. These patterns are rather reminiscent of La Niña conditions, as confirmed by correlation values computed with the seasonal MEI (Table II). Persistence of cluster #4 is the only to be significantly correlated with the MEI. These results suggest that ENSO may be associated with the origin of strong modulations of the relative frequency of all five clusters from one year to another. The well known involvement of El Niño events in SA seasonal droughts (Lindesay, 1988; Jury *et al*., 1994; Reason *et al*., 2000; Reason and Rouault, 2002; Rouault and Richard, 2003, 2005), is largely reinforced when recurrent South African generalised dryness are considered at daily timescale (cluster #1). Except for cluster #1, and in spite of the statistical significance of these signals, the correlation values shown in Table II remain nevertheless moderate. This relative weakness could be attributed to the non-linear nature of this relationship (F09) or to distinct regional effects of ENSO 'flavours' (Kao and Yu, 2009), as suggested by Hart *et al*. (2009).

To disentangle ENSO and more regional SST effects partly affected by the latter, we compute partial linear correlation coefficients between the clusters seasonal frequencies and seasonal SST (Figure 8(f)–(j)) after removal of the variance associated with the synchronous seasonal MEI. Qualitatively similar results are obtained using either the Niño3.4 or SOI indices (not shown). Correlation patterns for all clusters are barely modified

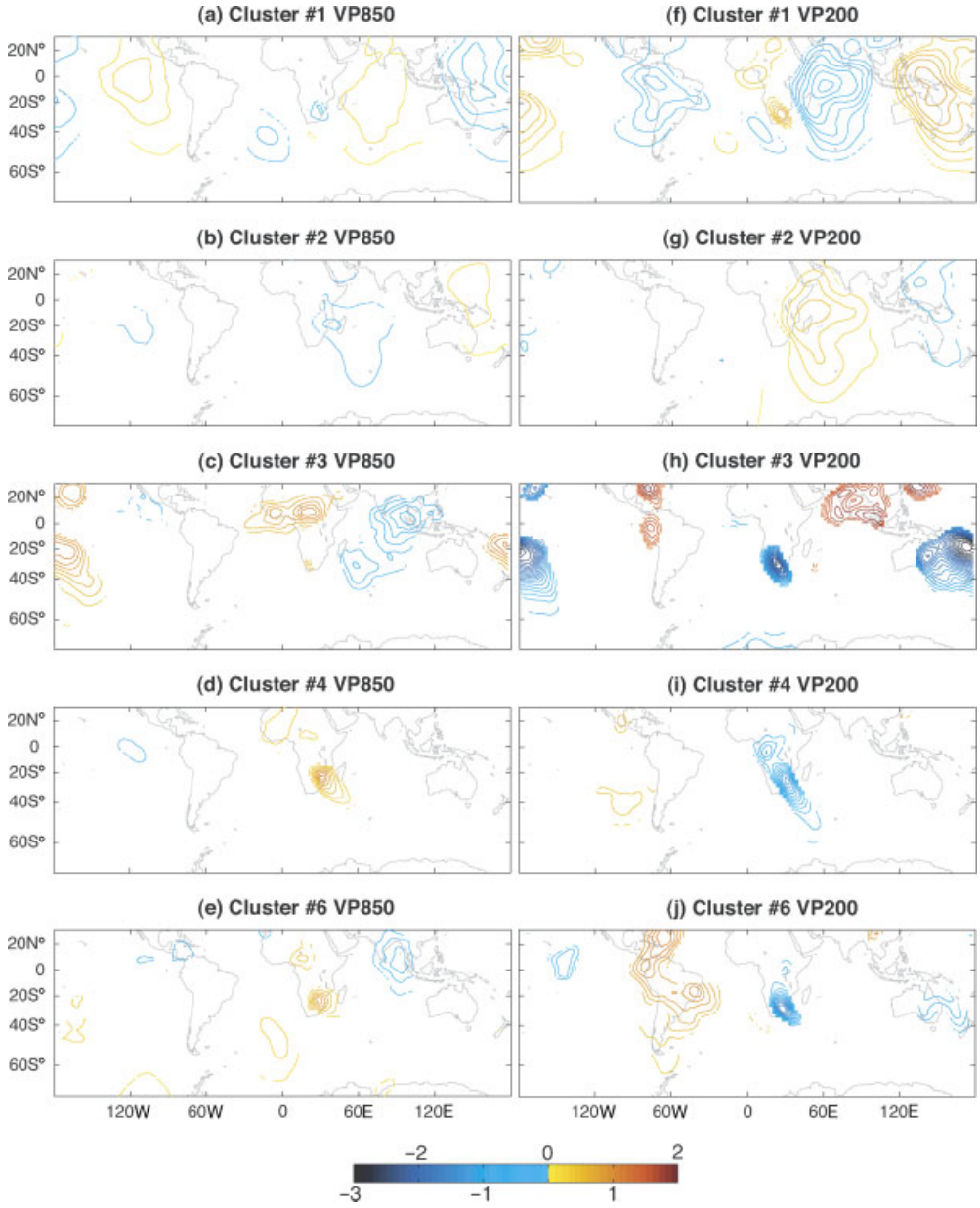

Figure 7. As Figure 6 but for velocity potential ($m^2\ s^{-1}$) at 850 hPa (left hand column) and 200 hPa (right hand column). Contour interval is $1 \times 10^5\ m^2\ s^{-1}$. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

in the mid-latitude (south of 50 °S), suggesting relative statistical independence of ENSO.

Results are more contrasted in the Tropics. Cluster #2 presents the singularity to be favoured by global warm conditions when ENSO effects are removed (Figure 8(g)). Positive correlations in the tropical Atlantic Ocean (Figure 8(b)) are enhanced (Figure 8(g)) and negative correlations elsewhere change sign (Figure 8g). More or less, the symmetric pattern is found for cluster #1. The strong positive correlations found in Figure 8(a) are essentially attributed to the tropical warm anomalies associated with El Niño events. Once this effect is removed, cold conditions in the tropics are found to be favourable for the occurrences of this cluster. This is particularly true for the western sides of the three major ocean basins. This questions the regional versus remote effects of ENSO in SA (i.e. through an atmospheric bridge linking the Pacific basin to Africa, or through a modulation of regional SST over the adjacent ocean basins: Reason *et al.*, 2000; Nicholson *et al.*, 2001). Addressing this problem is, however, beyond the scope of this study. Both clusters show in-phase teleconnections with the three sub-tropical ocean basins, a pattern already depicted by Fauchereau *et al.* (2003), and attributed to a southward shift and a strengthening of the subtropical high-pressure systems over the southern basins. Such mode of variability appears to be independent of ENSO (Figure 8(f) and (g)). Its positive phase (i.e. in-phase warm anomalies in the three basins) favours seasonal occurrences of cluster #2 and, thus, wet conditions in SA (Figure 2), the reverse being found for the negative phase and cluster #1.

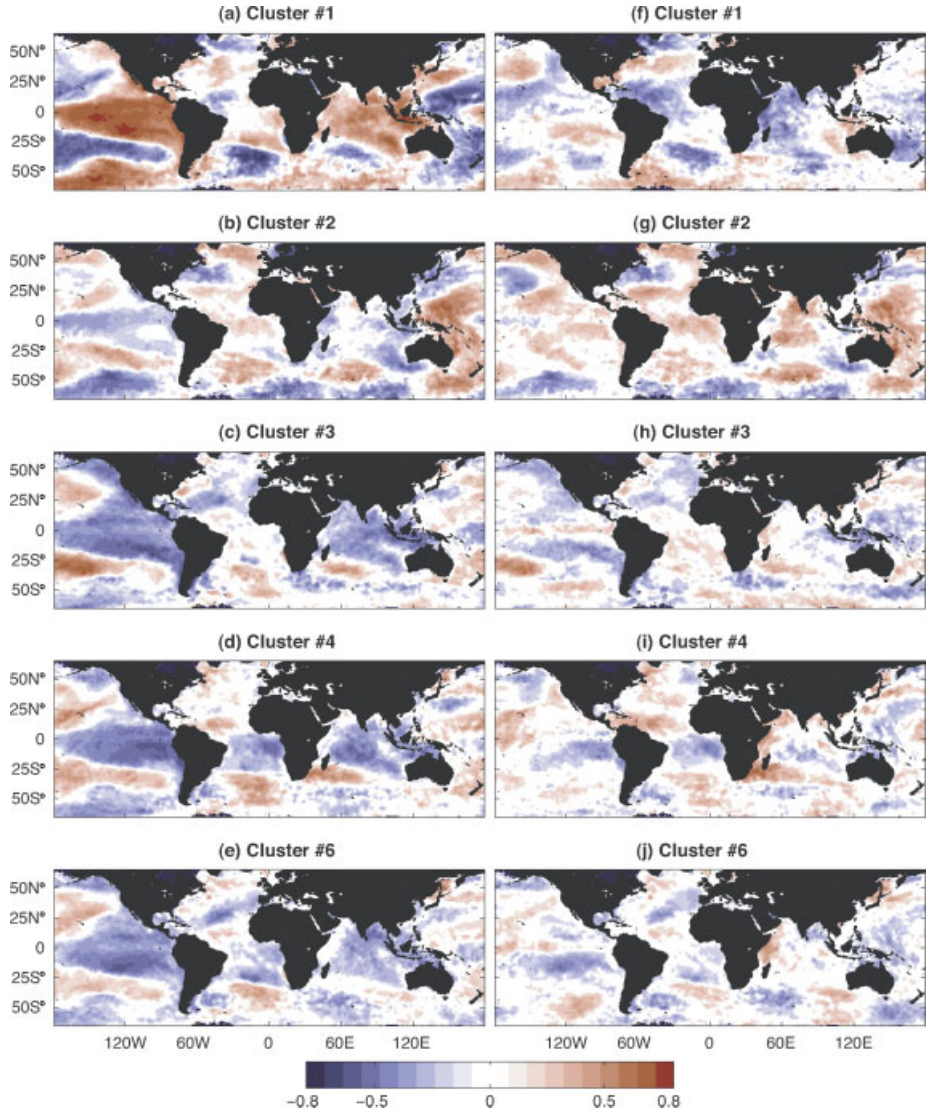

Figure 8. (a)–(e) Linear correlation coefficients between the seasonal frequencies of the 5 clusters and seasonal SST. (f)–(j) Same as (a)–(e) for partial linear correlation coefficients. The seasonal MEI index is used to remove ENSO effects. Only significant correlation coefficients are shown at the 95% level according to a Bravais-Pearson test. This figure is available in colour online at wileyonlinelibrary.com/journal/joc

Table II. Significant linear correlation coefficients between the number spent in each cluster by DJF season (top)/the mean spells duration associated with each cluster (bottom) and the bi-monthly MEI index averaged from December/January to January/February at the 95% level.

	Cluster #1	Cluster #2	Cluster #3	Cluster #4	Cluster #6
MEI	0.75 0.40	−0.36	−0.45	−0.48 −0.39	−0.38

Clusters #3, 4, and 6 were found to be favoured by La Niña conditions in the Pacific (Figure 8(c), (d) and (e)), the latter being typically associated with cold SST anomalies in the tropical Indian Ocean basin (Reason *et al*., 2000; Xie *et al*., 2002). In addition, clusters #4–6 showed significant negative (positive) teleconnections with south Atlantic and Indian SST at the tropical (mid-) latitudes (Figures 8(d)–(i) and 8(e)–(j), respectively), which are reinforced during La Niña events.

Our results confirm thus the positive teleconnection between subtropical south Atlantic/southwestern Indian Oceans and South African rainfall (such as Mason, 1995; Reason, 2001, 2002; Washington and Preston, 2006). However, negative correlation patterns between SA summer rainfall and central Indian Ocean SST (Jury *et al*., 1996), change sign when ENSO effects are removed. It seems therefore of primary importance to both consider full and partial teleconnections after removal of ENSO variability to disentangle the specific influence of regional SST variability and the regional SST signature of ENSO.

6. Summary and Discussion

An agglomerative hierarchical clustering was applied to the daily rainfall recorded at 5352 stations over

South Africa and the neighbouring countries of Lesotho, Swaziland, and southern Namibia. The aims were to investigate the nature of recurrent rainfall patterns during the core of the rainy season (December to February) and to relate them with the atmospheric circulation. The clustering technique led to the identification of six recurrent rainfall patterns among which the weakest (to which only 7 days were affiliated) was not analysed in detail.

One cluster (cluster #2), regrouping about 50% of the days, shows a rainfall pattern close to the seasonal mean, i.e. a gradient from dry conditions in the west to wetter conditions in the east. This cluster is also associated with excessively weak atmospheric anomalies on both regional and global scales, consistent with the fact that it reflects the climatology.

Cluster #1 regroups about 37% of the days and is associated with strong negative rainfall anomalies over South Africa, resulting from a regional trough-ridge-trough wave structure and moisture divergence over the region. At the global scale, this cluster is clearly associated with El Niño events and corresponds to a reduction or a possible shift of the Walker-type circulation and in the subtropical jet stream. The very strong teleconnections between its seasonal frequency and tropical SST are dramatically weakened when the influence of ENSO is removed, which confirms the implication of the latter. Thus, this clustering approach of the rainfall patterns appears as a useful tool to extend the results of Lindesay (1988) and Van Heerden *et al.* (1988) who found a covariance of 20% between ENSO and the South African summer rainfall amounts.

These two clusters (#1 and 2) are the most persistent. This is particularly true for cluster #1 whose persistence is even increased during El Niño events. This is in agreement with Tennant and Hewitson (2002), who showed that wet austral summer seasons in SA are not necessarily associated with a decrease in the persistence of dry days.

Regrouping, respectively, about 2.5, 11 and 6% of the days, the three remaining clusters (#3, 4, and 6) are both wetter and less persistent. They, respectively, correspond to a decrease, an increase, and maintenance of the seasonal mean eastward rainfall gradient. In each case, strong cyclonic moisture flux anomalies centred over Angola or northern Botswana and anomalous convergence located over the largest positive rainfall anomalies explain such wet conditions in SA. Among these clusters, two (clusters #3–6) are more or less all related to tropical/extra-tropical interactions, and one (cluster #4) is clearly associated with TTT events. In addition, and consistently with previous studies (F09, P09), the transitions between these clusters tends to depict an eastward propagation of heavy rainfall structures. Analysis of large scale circulation patterns associated with these clusters suggests a weak but significant effect of La Niña. The implication of ENSO is again confirmed by full and partial teleconnections with tropical SST.

Like all classification algorithms, the AHC partitioning has some limitations, and it cannot be considered as a satisfactory solution to discretise the spatio-temporal variability of the rainfall field into five or six recurrent classes. Nonetheless, it has the advantage of being concise and, in the present case, it allows us to demonstrate that the three wettest rainfall patterns could be linked to some extent either to tropical convection anomalies (clusters #3–6) or to tropical–temperate interactions (cluster #4). Possible causes for regional rainfall variability may also include tropical perturbations such as easterly waves or meso-scale convective systems, or the establishment of semi-permanent thermal lows over the margins of the Kalahari Desert (southern Angola, northern Namibia, and Botswana). Further work is still required to better understand how these different mechanisms interact or interfere at the synoptic timescale.

Acknowledgements

ECMWF ERA-40 data used in this study have been freely obtained from the ECMWF Data Server (http://data-portal.ecmwf.int/data/d/era40_daily/). Monthly HadISST is provided by the Hadley Center Meteorological Office. Calculations were performed using HPC resources from DSI-CCUB (Université de Bourgogne). Mathieu Rouault thanks ACCESS, CNRS, Nansen Tutu Center, NRF, and WRC for funding. We thank the WRC for making the rainfall data available. Similarly, we thank the two anonymous reviewers for their helpful comments.

References

Camberlin P, Janicot S, Poccard I. 2001. Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical Sea-Surface Temperature: Atlantic VS ENSO. *International Journal of Climatology* **21**: 973–1005.

Fauchereau N, Traska S, Richard Y, Roucou P, Camberlin P. 2003. SST co-variability in the Southern Atlantic and Indian Oceans and its connections with the atmospheric circulation in the Southern Hemisphere. *International Journal of Climatology* **23**: 663–677.

Fauchereau N, Pohl B, Reason CJC, Rouault M, Richard Y. 2009. Recurrent daily OLR patterns in the Southern Africa/Southwest Indian Ocean Region, implications for South African rainfall and teleconnections. *Climate Dynamics* **32**: 575–591.

Gong X, Richman MB. 1995. On the application of cluster analysis to growing season precipitation data in North America East of the Rockies. *Journal of Climate* **8**: 897–931.

Harrison MSJ. 1984a. A generalized classification of South African summer rain-bearing synoptic systems. *International Journal of Climatology* **4**: 547–560.

Harrison MSJ. 1984b. Comparison of rainfall time series over South Africa generated from real data and through principal component analysis. *International Journal of Climatology* **4**: 561–564.

Hart NCG, Fauchereau N, Reason CJC. 2009. El Nino's Flavors: implications for South African rainfall. Proceedings of the South African Society for Atmospheric Sciences, 25th *annual conference*, 9-10 sept. 09, Tulbagh: South Africa: 19–20.

Hart NCG, Reason CJC, Fauchereau N. 2010. Tropical-extratropical interactions over southern Africa: three cases of heavy summer season rainfall, *Monthly Weather Review* **138**: 2608–2623.

Jury MR, Mc Queen C, Levey K. 1994. SOI and QBO signals in the African Region. *Theoretical and Applied Climatology* **50**: 103–115.

Jury MR, Pathack B, Rautenbach CJDW, Vanheerden J. 1996. Drought over South Africa and Indian Ocean SST: statistical and GCM results. *The Global Atmosphere and Ocean System* **4**: 47–63.

Kao HY, Yu JY. 2009. Contrasting Eastern-Pacific and Central-Pacific types of ENSO. *Journal of Climate* **22**: 615–632.

Lindesay JA. 1988. South African rainfall, the Southern Oscillation and a southern hemisphere semi-annual cycle. *International Journal of Climatology* **8**: 17–30.

Lynch SD. 2003. Development of a raster database of annual, monthly and daily rainfall for Southern Africa. Report No. 1156/1/04, Water Research Commission: Pretoria, South Africa; 78 pp.

Mason SJ. 1995. Sea-Surface Temperature – South African rainfall associations, 1910-1989. *International Journal of Climatology* **15**: 119–135.

Mulenga HM. 1998. Southern African climatic anomalies, summer rainfall and the Angola low. PhD Thesis, University of Cape Town, Cape Town: South Africa; 232 pp.

Muñoz-Diaz D, Rodrigo FS. 2004. Spatio-temporal patterns of seasonal rainfall in Spain (1912-2000) using cluster and principal component analysis: comparison. *Annales Geophysicae* **22**: 1435–1448.

Muñoz-Diaz D, Rodrigo FS. 2006. Seasonal rainfall variations in Spain (1912-2000) and their links to atmospheric circulation. *Atmospheric Research* **81**: 94–110.

Nicholson SE, Selato JC. 2000. The influence of La Niña on African rainfall. *International Journal of Climatology* **20**: 1761–1776.

Nicholson SE, Leposo D, Grist J. 2001. The relationship between El Niño and drought over Botswana. *Journal of Climate* **14**: 323–335.

Pohl B, Richard Y, Fauchereau N. 2007. Influence of the Madden-Julian oscillation on Southern African summer rainfall. *Journal of Climate* **20**: 4227–4242.

Pohl B, Fauchereau N, Richard Y, Rouault M, Reason CJC. 2009. Interactions between synoptic, intraseasonal and interannual convective variability over Southern Africa. *Climate Dynamics* **33**: 1033–1050.

Ramos MC. 2001. Divisive and hierarchical clustering techniques to analyse variability of rainfall distribution patterns in a Mediterranean Region. *Atmospheric Research* **57**: 123–138.

Rayner NA, Parker DE, Horton EB, Folland CK, Alexander LV, Rowell DP, Kent EC, Kaplan A. 2003. Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late Nineteenth Century. *Journal of Geophysical Research* **108**, DOI:10.1029/2002JD002670.

Reason CJC, Mulenga H. 1999. Relationships between African Rainfall and SST anomalies in the Southwest Indian Ocean. *International Journal of Climatology* **19**: 1651–1673.

Reason CJC, Allan RJ, Lindesay JA, Ansell TJ. 2000. ENSO and climatic signals across the Indian Ocean basin in the global context. Part I: Interannual composite patterns. *International Journal of Climatology* **20**: 1285–1327.

Reason CJC. 2001. Subtropical Indian Ocean SST dipole events and Southern African rainfall. *Geophysical Research Letters* **28**: 2225–2227.

Reason CJC. 2002. Sensitivity of the Southern African circulation to dipole Sea-Surface Temperature patterns in the South Indian Ocean. *International Journal of Climatology* **22**: 377–393.

Reason CJC, Rouault M. 2002. ENSO-like decadal variability and South African rainfall. *Geophysical Research Letters* **29**: 16.1–16.4.

Reason CJC, Landman W, Tennant W. 2006. Seasonal to decadal prediction of Southern African climate and its links with variability of the Atlantic Ocean. *Bulletin of the American Meteorological Society* **87**: 941–955.

Richard Y, Trzaska S, Roucou P, Rouault M. 2000. Modification of the Southern African rainfall variability/ENSO relationship since the late 1960s. *Climate Dynamics* **16**: 883–895.

Richard Y, Fauchereau N, Poccard I, Rouault M, Trzaska S. 2001. 20th century droughts in Southern Africa: spatial and temporal variability, teleconnections with oceanic and atmospheric conditions. *International Journal of Climatology* **21**: 873–885.

Rouault M, Florenchie P, Fauchereau N, Reason CJC. 2003. South East Tropical Atlantic warm events and Southern African rainfall. *Geophysical Research Letters* **30**, DOI:10.1029/2002GL014840.

Rouault M, Richard Y. 2003. Intensity and spatial extension of drought in South Africa at different time scales. *Water S.A.* **29**: 489–500.

Rouault M, Richard Y. 2005. Spatial extent and intensity of droughts in Southern Africa, *Geophysical Research. Letters* **32**, DOI:10.1029/2005GL022436.

Tennant WJ, Hewitson BC. 2002. Intra-seasonal rainfall characteristics and their importance to the seasonal prediction problem. *International Journal of Climatology* **22**: 1033–1048.

Todd MC, Washington R. 1999. Circulation anomalies with tropical-temperate troughs in Southern Africa and the South West Indian Ocean. *Climate Dynamics* **15**: 937–951.

Todd MC, Washington R, Palmer PI. 2004. Water vapour transport associated with tropical-temperate trough systems over Southern Africa and the Southwest Indian Ocean. *International Journal of Climatology* **24**: 555–568.

Tyson PD, Preston-Whyte RA. 2000. *The weather and climate of Southern Africa* (2nd edition), Oxford University Press: Cap Town, South Africa; 396 pp.

Uppala SM, Kallberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Van de Berg L, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Holm E, Hoskins BJ, Isaksen L, Janssen PAEM, Jenne R, McNally AP, Mahfouf JF, Morcrette JJ, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Vitterbo P, Woollen J. 2005. The ERA-40 re-analysis. *Quartely Journal of the Meteorological Society* **131**: 2961–3012.

Van Heerden J, Terblanche DE, Schulze GC. 1988. The Southern oscillation and South African summer rainfall. *International Journal of Climatology* **8**: 577–597.

Van Heerden J, Taljaard JJ. 1998. Meteorology of the Southern Hemisphere. 1998. *American Meteorological Society Monograph*, Karoly DJ, Vincent DG (eds), Boston, Massachussetts; 307–336.

Vigaud N. 2007. Water Vapour Transport from the South Atlantic and Indian Oceans and Summer Rainfall in Southern Africa. PhD Thesis, University of Cape Town, Cape Town, South Africa; 222 pp.

Vigaud N, Richard Y, Rouault M, Fauchereau N. 2007. Water vapour transport from the Tropical Atlantic and summer rainfall in Tropical Southern Africa. *Climate Dynamics* **28**: 113–123.

Vigaud N, Richard Y, Rouault M, Fauchereau N. 2009. Moisture transport between the South Atlantic Ocean and Southern Africa: relationships with summer rainfall and associated dynamics. *Climate Dynamics* **32**: 113–123.

Ward JH. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association* **58**: 236–244.

Washington R, Todd MC. 1999. Tropical-temperature links in Southern African and Southwest Indian Ocean satellite-derived daily rainfall. *International Journal of Climatology* **19**: 1602–1616.

Washington R, Preston A. 2006. Extreme wet years over Southern Africa: role of Indian Ocean Sea Surface Temperatures. *Journal of Geophysical Research* **111**, DOI:10.1029/2005JD00672415.

Wolter K, Timlin MS. 1993. Monitoring ENSO in COADS with a seasonally adjusted principal component index. 17th Climate Diagnostics Workshop Norman, OK; 52–57.

Xie SP, Annamalai H, Schott F, McCreary JP Jr. 2002. Structure and mechanisms of South Indian Ocean climate variability. *Journal of Climate* **15**: 864–878.

