

Validation of a coupled GCM and projection of summer rainfall change over South Africa, using a statistical downscaling method

Yan Zhao^{1,2,*}, Pierre Camberlin¹, Yves Richard¹

¹Centre de recherches de climatologie, CNRS UMR 5080, Université de Bourgogne, Sciences Gabriel, 6 Bd Gabriel, 21000 Dijon, France

²Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing PO Box 9804, 100029 Beijing, China

ABSTRACT: In this study, the southern African climate response to increasing amounts of greenhouse gases is investigated, based on the dataset of a 150-yr climate change experiment following the IPCC Special Report on Emissions Scenarios marker scenario B2 (SRES-B2) performed with the coupled ARPEGE/OPA/GELATO general circulation model (GCM). The method of canonical correlation analysis (CCA) is adopted to validate the ability of the GCM to simulate the present-day climate over the southern African region and project the late-summer rainfall change over South Africa at the end of the 21st century. The model validation shows that the ARPEGE/OPA/GELATO GCM is able to capture the observed link between rainfall over South Africa and adjacent sea-level pressure (SLP), despite the existence of some systematic errors. The structure and variability of SLP are reproduced by the GCM in a realistic way. The major controlling mechanism of rainfall over South Africa can be identified in the GCM. The projection of rainfall indicates a drying trend during the 21st century over most parts of South Africa, in particular the central interior. Compared to present-day climatology, the overall late-summer rainfall will decrease by 8.2% by the end of 21st century as derived from GCM grid-point output, and by 16.1% from the downscaling model.

KEY WORDS: Climate change · Model evaluation · Statistical downscaling · Canonical correlation analysis · South Africa

—Resale or republication not permitted without written consent of the publisher—

1. INTRODUCTION

The African continent is particularly vulnerable to the impacts of climate change because of factors such as widespread poverty, recurrent droughts, inequitable land distribution, and overdependence on rainfed agriculture (IPCC 2001). About 45% of southern Africa (south of 15° S) receives on average less than 500 mm precipitation yr⁻¹ and is therefore highly vulnerable to rainfall variability and change. With socio-economic development and population increase, the need for additional water resources is becoming more urgent. Therefore, projecting rainfall change is of major importance to the southern African community.

Future annual warming across Africa ranges from 0.2°C decade⁻¹ (B1 — low scenario) to more than 0.5°C

decade⁻¹ (A2 — high scenario) (Hulme et al. 2001). This warming is greatest over semi-arid margins like central southern Africa. But the inter-model range is also greatest over the interior of southern Africa. Future changes in mean seasonal rainfall are even less well defined. For the Orange Basin, for example, which is fundamental for the South African economy, Arnell (1999) estimates that the change in precipitation will be between -5 and +5%, and the change in runoff between -10 and +10%. A regional analysis of ensemble simulations shows that for southern Africa (as in 20 to 30% of all worldwide cases) the simulated sensitivity does not have the same sign in all realizations of the ensemble (Giorgi & Francisco 2000). Thus, large uncertainties remain concerning the future evolution of South African rainfall and water resources. How-

*Email: yan.zhao@cea.fr

ever, there is evidence of increased interannual variability in recent decades over the region (Richard et al. 2001), which fuels further concern about that evolution.

General circulation models (GCMs) are the most fundamental tool for studying climate change. Whereas GCMs do a creditable job on the large scale, they have much lower skill at the regional scale. However, it is at these higher resolution scales that climate change information is most needed. A possible approach to bridging the scale gap is downscaling, which uses dynamical or statistical models to relate large-scale information from GCMs to regional parameters (Karl et al. 1990, Giorgi & Mearns 1991, Joubert et al. 1999, Zorita & Von Storch 1999). Considering that the application of regional climate models is still in its infancy and computationally very demanding, statistical downscaling is a practical means to address the immediate needs of the southern African region.

Applying statistical downscaling methods to climate change analysis is becoming quite popular (Von Storch et al. 1993, Cui et al. 1995, Crane & Hewitson 1998, Busuioc et al. 2001). However, downscaling has received relatively little attention in Africa. Hewitson & Crane (1998) and Hewitson & Joubert (1998, available at www.egs.uct.ac.za/fccc/) have applied the statistical downscaling method of artificial neural networks (ANNs) to generate climate scenarios for South Africa. In their study, the climate scenario used is $2 \times \text{CO}_2$ which is an idealized situation. With new scenario experiments available, it is necessary to project the future rainfall change under different conditions and with different GCMs.

In this study, the scenario experiment follows the IPCC Special Report on Emissions Scenarios marker scenario B2 (SRES-B2) (Nakicenovic & Swart 2000). The statistical downscaling method is based on canonical correlation analysis (CCA) which tends to offer clearer physical interpretation than ANNs (Barnett & Preisendorfer 1987, Von Storch 1995).

The purposes of this paper are (1) to use CCA as a diagnostic tool to verify the performance of a climate model on the regional scale and (2) to apply a statistical downscaling model based on CCA to a scenario experiment in order to project future summer rainfall change over South Africa.

Sections 2 and 3 of the paper present the data, GCM experiment and methodology used in this study. Validation of the GCM simulation of the present-day climatology is described in section 4. In section 5, we compare the estimates of rainfall change with respect to rainfall over South Africa, as derived directly from the SRES-B2 simulations and indirectly using the downscaling model. Section 6 completes the paper with a brief summary and concluding remarks.

2. DATA AND GCM EXPERIMENT

2.1. Data

Most of South Africa experiences summer rainfall (October to March). Observation shows that circulation patterns associated with the rains differ between the early and late summer (D'Abreton & Lindesay 1993). Compared to early summer, late summer (January to March [JFM]) variability is internally more consistent (Richard et al. 2002). In addition, this season receives a large proportion (40%) of the total annual rainfall over most parts of southern Africa. Therefore in this study, late summer is chosen as the target season.

The original monthly precipitation data were obtained from the South African Weather Service (SAWS). Covering the whole of South Africa, 93 rainfall districts are defined by SAWS. For each one, a spatial mean using a variable number of rain-gauges was computed. Series start in 1920, and have undergone quality control by SAWS. We retain 55 of these districts, where JFM seasonal rainfall accounts for over 40% of the annual total. They are used as the dependant variables in the statistical downscaling model. For the southern African region and adjacent oceans, we additionally use Climate Prediction Center Merged Analysis of Precipitation data (CMAP, Xie & Arkin 1996) and the Hulme monthly precipitation data from the Climatic Research Unit (CRU, Hulme 1994, updated), available as a 2.5×3.75 degree grid. Monthly mean sea-level pressure (SLP), which is used as the independent variable in the statistical downscaling model, is from NCAR/NCEP reanalysis data for 1948 to 1999 (Kalnay et al. 1996). NCEP SLP has been shown to be close to the GMSLP (Global Mean Sea Level Pressure) observed data from the Hadley Centre (Fauchereau et al. 2003). In addition, we also use grid-point precipitation and SLP data from the GCM.

2.2. GCM experiment

The GCM used here is the coupled atmosphere-ocean-sea-ice model ARPEGE/OPA/GELATO assembled at Météo-France. The atmospheric general circulation model (AGCM) is ARPEGE version 3.0 (Déqué et al. 1994, Déqué et al. 1999). It has a T63 spectrum truncation corresponding to a 128×64 Gaussian grid (roughly 270×270 km over South Africa) and 45 vertical levels distributed between the earth's surface and the upper stratosphere. At the earth's surface, the ISBA (interactions between soil biosphere and atmosphere) land surface scheme is incorporated (Mahfouf et al. 1995, Noilhan & Mahfouf 1996, Douville et al. 2000). The ocean model is the OPA GCM version 8.0 devel-

oped at Laboratoire d'Océanographie Dynamique et de Climatologie (LODYC), and the sea-ice sub-model is the GELATO model from CNRM (Salas-Mélia, 2002). These 2 models are coupled to the AGCM by means of the OASIS coupler developed at CERFACS (Terray et al. 1995). The ARPEGE/OPA/GELATO coupled model was found to skillfully simulate El Niño Southern Oscillation (ENSO) features in the Pacific, as well as large-scale teleconnections (Latif et al. 2001, Ashrit et al. 2003, Camberlin et al. 2004).

In the scenario experiment used here, the greenhouse gases (GHG) and aerosol concentrations are prescribed but updated each year according to observations from 1950 to 2000, and to the SRES marker scenario B2 afterwards. B2 is one of the most widely used scenarios in climate change simulations (IPCC 2001, Hulme et al. 2001), thus enabling straightforward comparisons. Under this scenario, the CO₂ concentration reaches 620 ppm in 2100, which is less than double the current concentration. The global distribution of sulphate aerosols is updated every 10 yr. Both the direct and indirect effects of aerosols are taken into account in the radiative code, according to a parameterisation derived from the work of Boucher & Lohmann (1995). More details about the GCM experimental design can be found in Royer et al. (2002).

We chose 2 main 30-yr time slices for the study: 1970 to 1999 representing present-day climate, and 2070 to 2099 for future climate. The difference between these 2 periods is hereafter referred to as climate change. Global scale patterns and teleconnections are presented in Douville et al. (2002) and Camberlin et al. (2004). Here we focus on the southern African region. Note that there are some limitations to the results, since they are based on 1 experiment only using a single GCM. Intermediate 30-yr timeslices have also been considered; these suggest that the climate trends detected based on the years 2070 to 2099 are somewhat robust, but confidence in the results will require consideration of additional experiments and other GCMs.

3. METHOD

In this study, a statistical downscaling procedure based on CCA is used to validate the performance of the GCM and infer rainfall information over South Africa from SLP fields. CCA has the ability to show relationships between 2 sets of space-time-dependent variables by identifying the optimum linear combination between the 2 sets with maximum correlation being produced. One of the main advantages of the method is that it delivers spatial patterns that normally lend themselves to a clear physical interpretation. In

this sense, it is also a straightforward way to investigate whether the GCM is able to reproduce the observed linear relationship between the 2 fields.

The CCA downscaling technique has recently been used for monthly/seasonal rainfall forecast studies in South Africa (Landman & Tennant 2000, Landman & Goddard 2002). The predictors are usually SLP or geopotential heights (1000 hPa, 850 hPa and 500 hPa) or a combination thereof. For example, Landman & Tennant (2000) use SLP and 500 hPa height to downscale monthly rainfall for December to February; in another study (Landman & Goddard 2002), 850 hPa height was chosen as the predictor. These studies indicate that a statistical downscaling method based on CCA can produce satisfactory rainfall forecasts.

For climate change issues, SLP is considered to be a more suitable atmospheric circulation variable compared to geopotential heights (Zorita & von Storch 1999). Geopotential heights tend to be much more affected by global warming, but these changes may be related to changes in the mean atmospheric density and not necessarily to changes in atmospheric circulation. Therefore, we choose SLP rather than geopotential heights as the independent variable to develop the statistical downscaling model. Although downscaling without the incorporation of key variables such as atmospheric water vapour in the relationship may lead to difficulties, significant regional climate change information may still be derived (Joubert & Hewitson 1997).

The performance of a statistical downscaling model may be sensitive to the predictor domain chosen. For example, when a smaller domain excluding the Atlantic section (10° W to 70° E) was used, we found that the first 3 SLP EOF patterns for NCEP reanalysis data did not agree very well with those for the GCM. Therefore, a wider domain was used. Based on careful assessments, we chose the domain 40° W to 100° E and 55° S to the equator, which covers much of the tropical–subtropical Atlantic and Indian Ocean, to construct the CCA downscaling model.

In this study, projection of the future rainfall change is made in 2 steps. Firstly, we construct a CCA statistical downscaling model based on the present-day observed link between SLP and rainfall:

$$\hat{R}(x,t) = \sum_{i=1}^n \alpha(i) \cdot s(i,t) \cdot \mathfrak{R}(i,x) \quad (1)$$

where $\hat{R}(x,t)$ are the time series of estimated rainfall anomalies for a given district x , regression coefficients $\alpha(i)$ are related to canonical correlation coefficients, $s(i,t)$ are the time-dependent amplitudes of the canonical correlation patterns for SLP, $\mathfrak{R}(i,x)$ are the canonical correlation patterns for rainfall, and n is the number of canonical patterns retained in the statistical downscaling model.

Secondly, keeping α and \mathfrak{R} unchanged, which is the downscaling assumption that these canonical correlation relationships will also remain valid in the future climate, the rainfall change estimated by the statistical downscaling model is just a linear combination of the canonical patterns. This linear combination changes only for the different time steps, as given by the time series $s(i, t)$:

$$Slpa(x, t) = \sum_i^n [Slp(i, x) \cdot s(i, t)] + \varepsilon \quad (2)$$

where $Slpa(x, t)$ are the SLP changes between future and present-day, $Slp(i, x)$ are the canonical SLP patterns determined in the first step and ε are the residuals. By minimizing the variance of the residuals, $s(i, t)$ is found. Applying $s(i, t)$ to Eq. (1), future rainfall change is projected.

More detailed information on the procedure can be found in Barnett & Preisendorfer (1987), von Storch et al. (1993) and Zorita & von Storch (1999); relevant standard calculation routines are available from public software packages.

The above procedure indicates that the projected rainfall changes depend closely on the quality of the

statistical downscaling model. Thus, in the first step, extensive sensitivity experiments were conducted to determine $\alpha(i)$ and $\mathfrak{R}(i, x)$, which are sensitive to many factors, such as the domain chosen to identify the stable large-scale patterns, the number of EOFs retained to extract enough large-scale information and filter unnecessary 'noise', and the number of CCA components used for the construction of the downscaling model.

4. GCM SIMULATION OF PRESENT-DAY CLIMATE

4.1. Mean states

Fig. 1 shows the seasonal mean rainfall over southern Africa and adjacent areas. Compared with CMAP data (1979 to 2000 mean, Fig. 1a), the GCM reproduces the general features well, that is rainfall is centred over tropical areas, Madagascar and part of the sub-tropical region. The 'V' pattern of the Intertropical Convergence Zone (ITCZ)—located at higher latitude around the Mozambique Channel and closer to the equator over the tropical Indian Ocean and the African continent—is simulated in a realistic way. However, rainfall is overestimated over southern Africa, as it is by the Hadley Centre HadAM3H (Hudson & Jones 2002) and many other GCMs (Joubert & Hewitson 1997). Note that the NCEP reanalysis data does not exhibit this feature, except along the south-eastern coast (Poccard et al. 2000). GCM rainfall overestimation may be due to unrealistic representation of orography and sub-grid parameterization.

The global SLP pattern (not shown) is captured by the GCM in a realistic way. For example, the position of the main semi-permanent high pressure cells of the subtropics is generally well reproduced. There are larger errors at high latitudes (roughly north of 65°N and south of 50°S) which may partly be related to the treatment of sea-ice in the coupled model. Ashrit et al. (2003) found an underestimation of the sea-ice feedback in similar experiments with the same GCM. The simulation of SLP over low and middle latitudes is in good agreement with the NCEP climatology, as demonstrated in Fig. 2 for southern Atlantic and Indian Ocean region. Compared to NCEP, the simulated SLP is slightly higher over the interior of South Africa, but the difference is generally less than 1 hPa. Around southern Africa, the 2 major sub-tropical highs, the Mascarene and St-Helena anticyclones, are reproduced well in terms of both location and shape, but the intensities of these 2 highs are 1 to 2 hPa lower in the GCM. Therefore, the SLP gradient between the sub-tropical oceans and the African continent, as well as with the mid-high latitudes, is somewhat weaker than in NCEP.

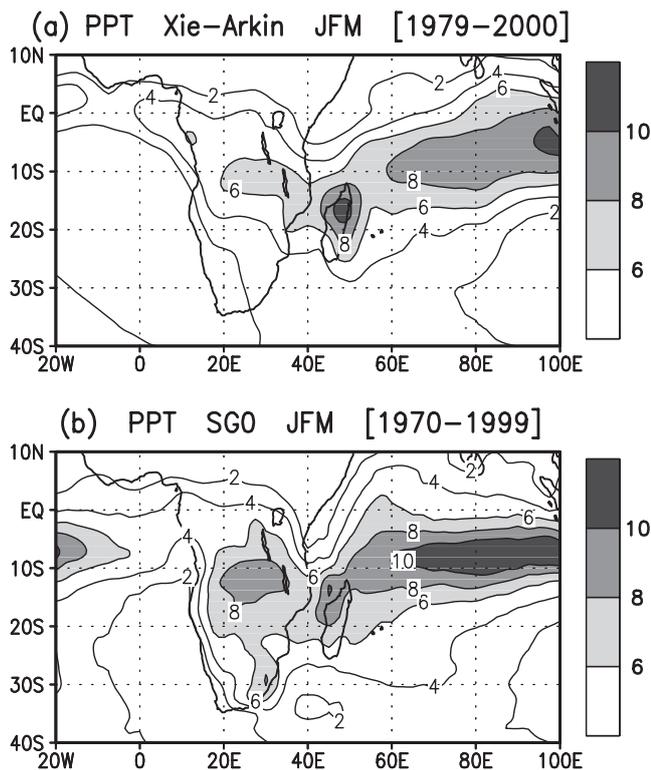


Fig. 1. Seasonal mean precipitation (PPT) January to March (JFM, mm d⁻¹). (a) CMAP (1979 to 2000; Xie-Arkin, Xie & Arkin 1996), (b) ARPEGE/OPA/GELATO simulation (1970 to 1999). Contour interval is 2 mm d⁻¹; values exceeding 6 mm d⁻¹ are shaded

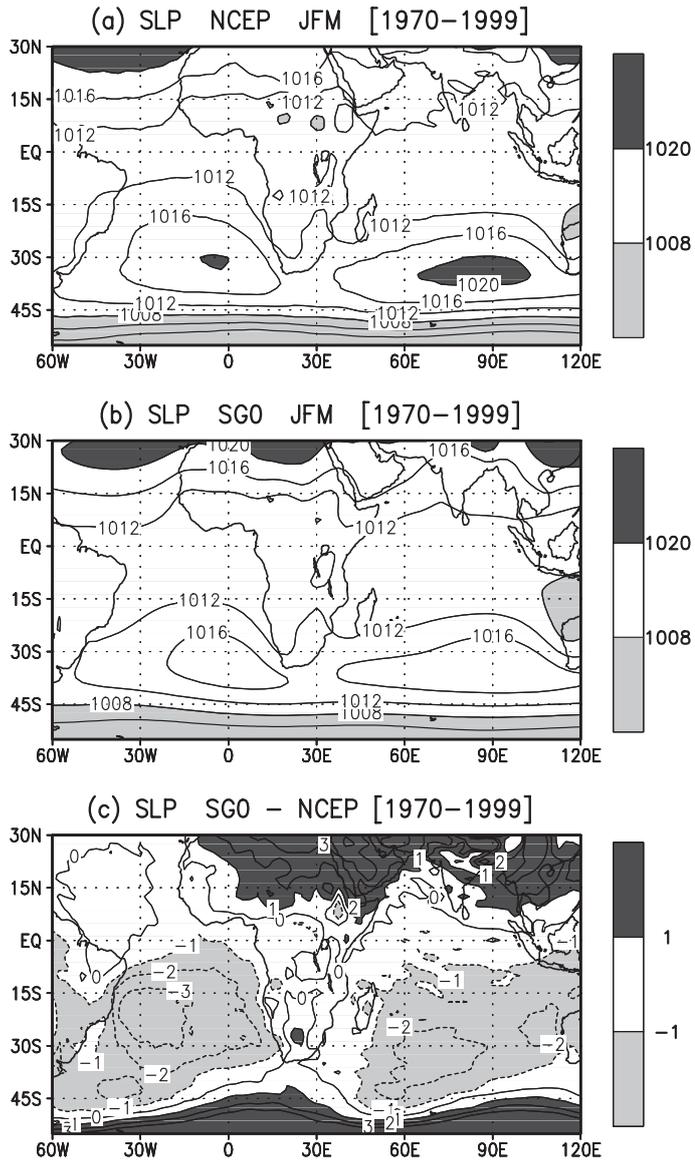


Fig. 2. Seasonal mean SLP (sea-level pressure, hPa) January to March, 1970 to 1999. (a) NCEP reanalysis, (b) ARPEGE/OPA/GELATO simulation and (c) difference between GCM (general circulation model) and NCEP. In (a) and (b) contour interval is 4 hPa, values below 1008 are lightly shaded and values exceeding 1020 are heavily shaded. In (c) contour interval is 1 hPa. NCEP: National Centers for Environmental Prediction reanalysis data. For other abbreviations see Fig. 1

4.2. Teleconnection patterns

The Southern African Rainfall Index (SARI, Richard et al. 2000) is used to investigate relationships between observed rainfall over southern Africa and global SLP. To define the SARI, 28 grid points selected by an EOF were extracted from the CRU dataset (Hulme 1994). These grid points cover an area, including the part of South Africa under study here, with strong covariance.

After standardising each of the 28 series, a spatial mean was computed (Richard et al. 2000). A ‘quasi-SARI’ index was computed as the spatial mean of 20 GCM grid points covering approximately the same area. The correlation map between SARI and NCEP reanalysis (1970 to 1999) shows that rainfall over southern Africa is significantly associated with extended SLP anomalies (Fig. 3a). A dipolar pattern of high SLP over the southwest Indian Ocean and low SLP over South Africa is associated with above normal rainfall over South Africa. This dipole was previously observed using outgoing longwave radiation data (Jury 1992). Over the tropical Pacific area, negative correlations are found in the west, and positive in the east, a pattern characteristic of ENSO variability. Southern African rainfall variability is known to be associated with ENSO (Lindesay 1988).

The southwest Indian Ocean dipolar pattern and the larger-scale teleconnections are correctly reproduced by the GCM at present-day (Fig. 3b), and remain robust in the future (Fig. 3c), although these relationships seem to be stronger than those found from the NCEP reanalysis. This is consistent with simulated teleconnection signals between ENSO and global rainfall being slightly stronger than observed (Camberlin et al. 2004), though the GCM adequately reproduces the patterns and seasonality. In particular, the GCM simulates a distinct drying over southern Africa in warm ENSO years, in agreement with the observations (Camberlin et al. 2004).

Over southern Atlantic high latitudes (40 to 60° S), there is also a significant positive correlation area with SLP, which the GCM fails to display or misplaces towards the southwest Indian Ocean. Considering the relatively sparse observed SLP data over high latitude oceans in the Southern Hemisphere, this discrepancy may be partly data-dependent. However, when we choose the downscaling domain, the southern boundary is limited to 45° S so as to reduce the possible effects of this discrepancy.

The correlation map of SARI with 850 hPa geopotential height is similar to the pattern of SARI with SLP, and this pattern is also well reproduced by the GCM (not shown here).

4.3. EOF patterns

4.3.1. Precipitation

For observed precipitation, the original 55 district rainfall data were first regridded on a similar grid to that of the GCM, in order to enable meaningful comparisons. The first EOF pattern, as computed from the covariance matrix, explains 65.8% of the variance. It

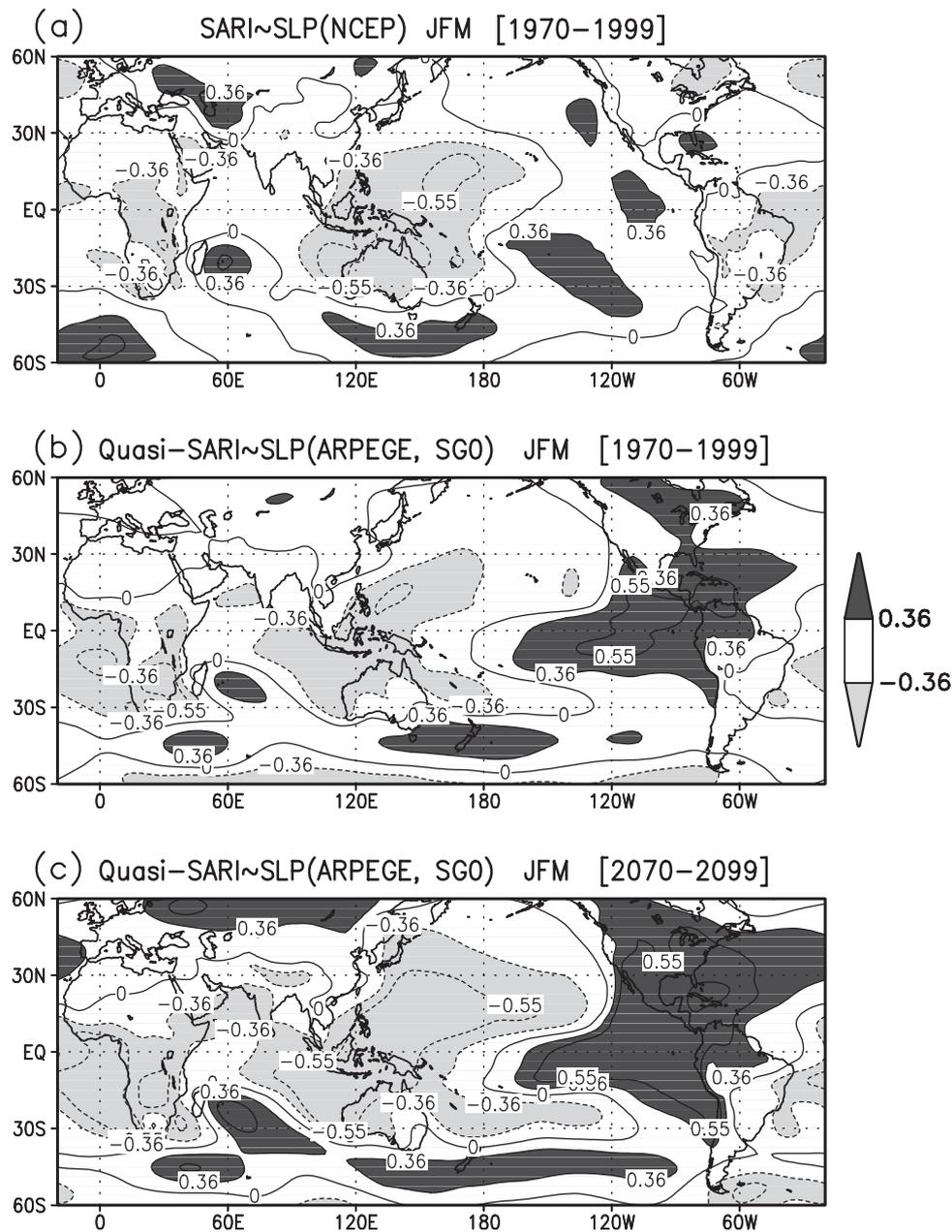


Fig. 3. Correlation between (quasi-) SARI (Southern African Rainfall Index) and SLP. (a) NCEP reanalysis (1970 to 1999), (b) ARPEGE/OPA/GELATO (1970 to 1999) and (c) ARPEGE/OPA/GELATO (2070 to 2099). Light (dark) shading indicates areas of negative (positive) correlation. Contour values are 0, ± 0.36 , and ± 0.55 ; 0.36 and 0.55 are locally significant correlation thresholds at 95 and 99.9% level, respectively. For abbreviations see Figs. 1 & 2

bears the same sign over the whole of South Africa with slightly heavier positive loadings in the central interior (Fig. 4a). There exists over southern Africa an extensive mode of rainfall variability, extending beyond South African boundaries (e.g. Richard et al. 2000).

The second pattern (Fig. 4b), explaining only 17.2% of the total variance, displays opposite loadings between the southwest (positive) and the northeast (neg-

ative). It should be interpreted as a weighing factor associated with the first EOF: a systematic analysis of JFM rainfall anomalies in South Africa since 1920 did not indicate such a recurrent dipole (Rouault & Richard 2003). The first and second EOFs together account for 83% of the total seasonal precipitation variance.

The GCM seemingly fails to reproduce the observed features, since the uniform variations displayed in EOF1 are not found (Fig. 4c,d). However,

the first 2 EOFs depict regional patterns which coincide well with the 2 poles displayed in the observed EOF2 (Fig. 4b). Moreover, when a rotation is performed on the first 2 observed EOFs (Fig. 4e,f), we obtain spatial patterns which are very close to those obtained from the GCM output, either unrotated (Fig. 4c,d) or rotated (Fig. 4g,h). The total variance

explained by the first 2 EOFs does not differ much between the observations (83%) and the GCM (79.7%). With the exception that the large-scale variability common to the whole region is weak in the GCM, one can consider that the GCM fairly well approximates the large-scale patterns in the observations.

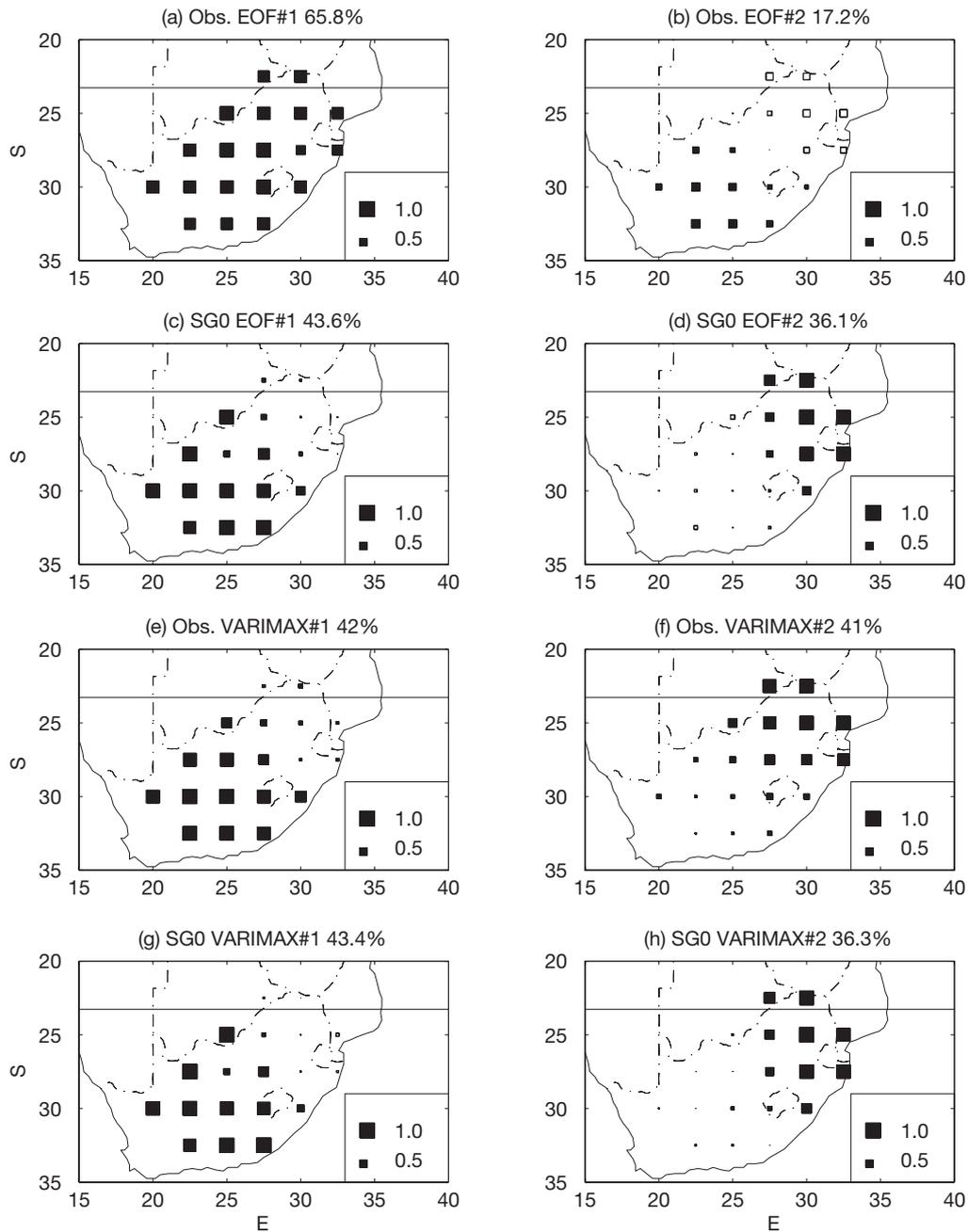


Fig. 4. EOF spatial pattern of precipitation (1970 to 1999). (a) 1st and (b) 2nd PC from upscaled observations (Obs.); (c) 1st and (d) 2nd PC from ARPEGE/OPA/GELATO; (e) 1st and (f) 2nd PC from upscaled observations with Varimax rotation; and (g) 1st and (h) 2nd PC from ARPEGE/OPA/GELATO with Varimax rotation. Shaded: positive loadings; white: negative loadings. SGO: control experiment

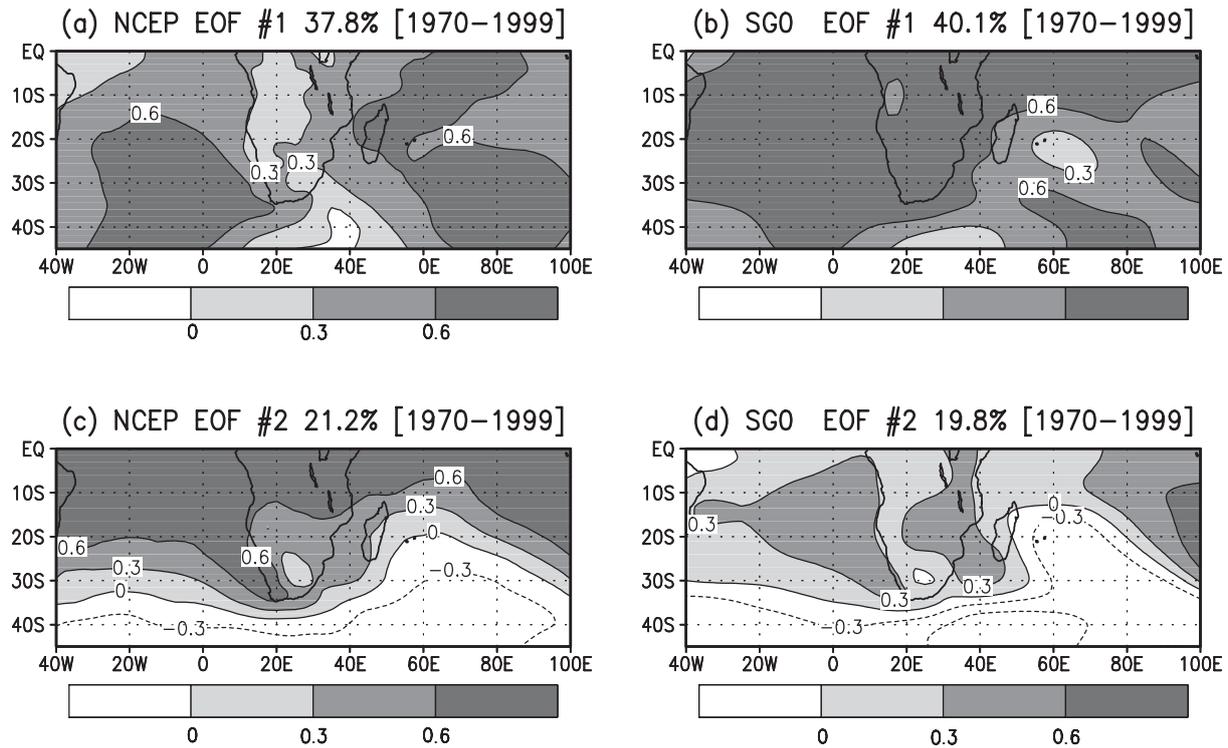


Fig. 5. EOF spatial pattern of SLP (1970 to 1999). (a) 1st and (c) 2nd PC from NCEP; (b) 1st and (d) 2nd PC from ARPEGE/OPA/GELATO. For abbreviations see Figs. 1, 2 & 4

4.3.2. SLP

The first unrotated EOF pattern from NCEP reanalysis data is represented by a positive sign over almost all the domain, with loadings heaviest in the south Atlantic and south Indian Oceans (Fig. 5a). The GCM leading EOF pattern is also dominated by a positive sign everywhere (Fig. 5b). However, the patterns do not perfectly agree, since loadings over land are weaker (though still positive) in the GCM than in NCEP reanalysis. The explained variances from NCEP reanalyses and the GCM are 37.8 and 40.1%, respectively. This mode, depicted in a similar way by both the reanalyses and the GCM, is not an artefact: in the observation (Nicholson & Kim 1997) and in the GCM (not shown), in-phase variations are shown over both basins at times of ENSO events during JFM, and the present mode is partly a reflection of the ENSO impact on the regional SLP field.

The second EOF patterns from NCEP reanalyses and ARPEGE are also similar; both of them display an out-of-phase relationship between the tropics and subtropics (Fig. 5c,d). The explained variances are 21.2 and 19.8%, respectively. These patterns broadly coincide with those depicted in Fauchereau et al. (2003). They exhibit a pressure distribution over the Atlantic Ocean which is similar to that of the Indian Ocean,

including anomalies of opposite signs between the tropical and the subtropical-temperate latitudes.

4.4. CCA patterns

Fig. 6 shows the first CCA features from observed rainfall and NCEP reanalysis SLP. Wetter than normal conditions over most of South Africa (Fig. 6a) are associated with a pattern of low SLP covering South Africa and 2 high SLP nuclei in the southern Indian and southern Atlantic Oceans, bringing moist air masses onto the continent (Fig. 6b). This pattern is similar to that obtained by Miron & Tyson (1984) for the wettest seasons. This association is emphasized by the high correlation (0.77) between the SLP and rainfall CCA time scores (Fig. 6c).

CCA2 (Fig. 7a) should be viewed as a weighting factor associated with the NE/SW modulation of South African rainfall as displayed in the CCA1 pattern. The dominant SLP feature of CCA2 (Fig. 7b) is located over the ocean southeast of the subcontinent, with enhanced (reduced) SLP at times of above (below) normal rainfall over the central interior (northeast). The positive SLP anomalies correspond with an anticyclonic/anticlockwise circulation anomaly. In agreement with CCA2, Reason (1998) shows that in the

CSIRO GCM such a circulation feature, forced by warm sea surface temperature (SST) south of South Africa, leads to regionally contrasted rainfall anomalies over South Africa, with dry conditions over the northeast.

Given the fact that the GCM fails to replicate the in-phase variability over the region (section 4.3.1), it is no surprise that the CCA features at first sight look quite different for the GCM simulation. In the GCM, CCA1 depicts rainfall over the central interior (Fig. 8a), and

CCA2 over the northeast plateau (Fig. 9a). The rainfall and SLP time scores are similar, with correlations of 0.88 (Fig. 8c) and 0.72 (Fig. 9c), for CCA 1 and 2 respectively. A common feature associated with above normal rainfall in both regions (Figs. 8b & 9b), well shown in the observations as well (our Figs. 6b & 7b, July 1992), is a SLP high east of Madagascar. However, south of the continent, the 2 simulated CCA modes display different features, i.e. a high (low) associated with a wet central interior (northeast). These features are

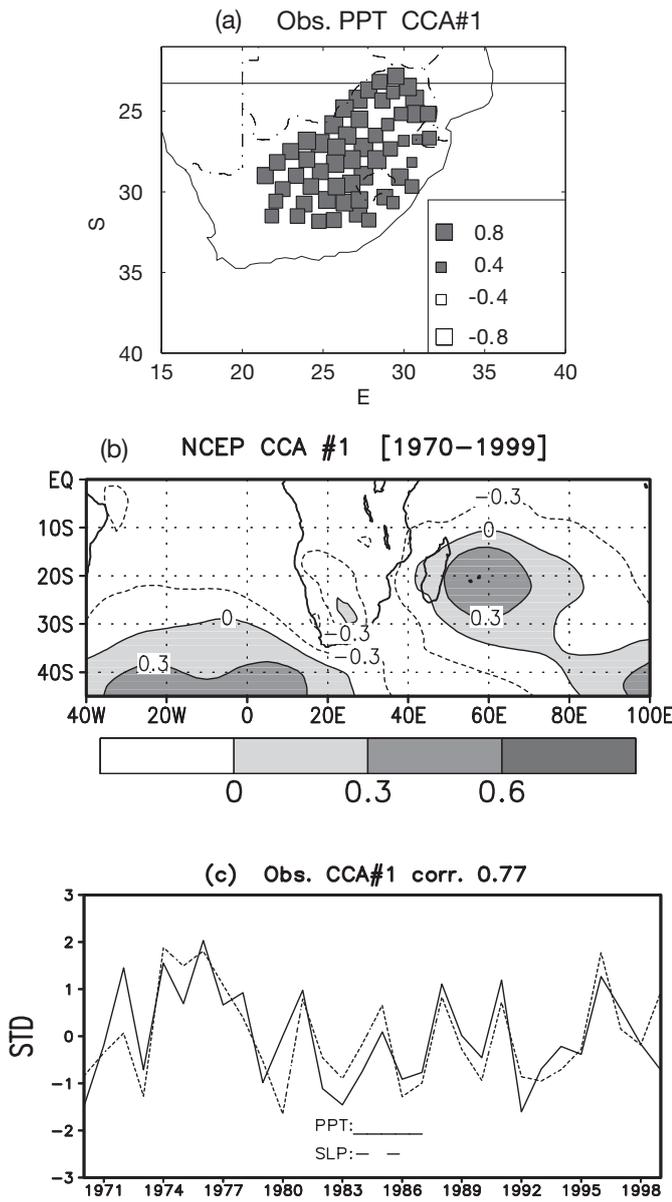


Fig. 6. First canonical correlation analysis (CCA) pairs from observed precipitation and NCEP SLP (1970 to 1999). (a) Precipitation correlation map, (b) SLP correlation map and (c) time scores. Obs.: observations; PPT: precipitation; SLP: sea-level pressure; corr: correlation coefficient; STD: standard deviation

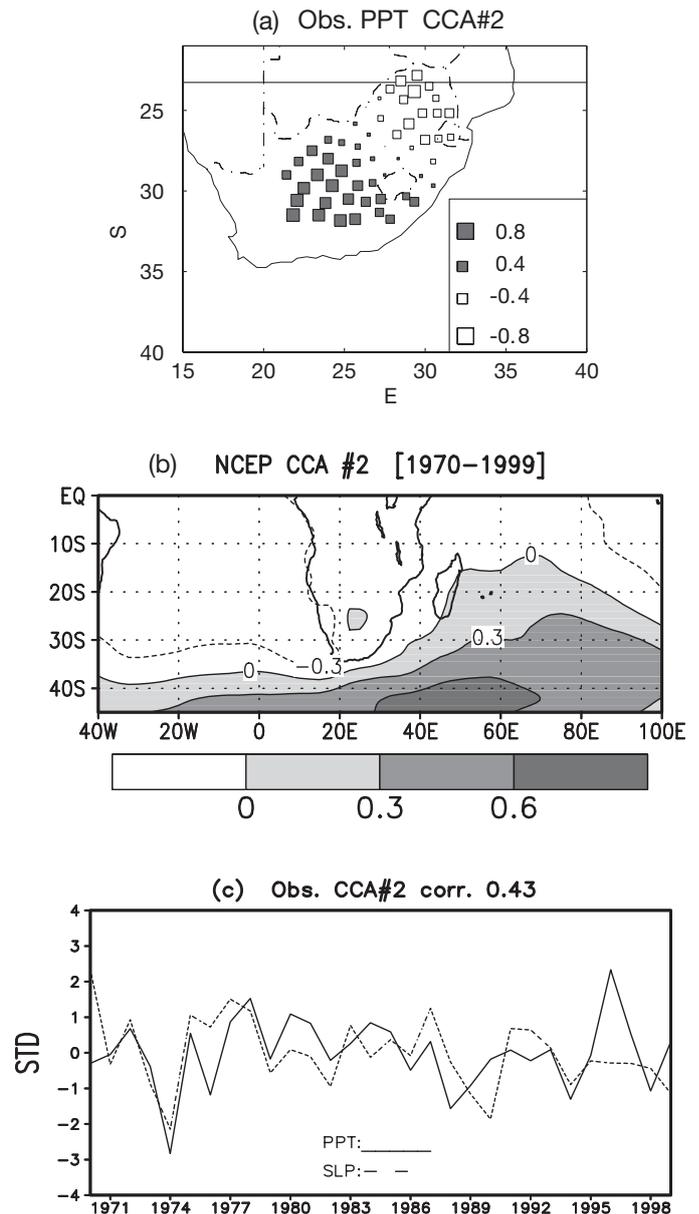


Fig. 7. Second CCA pairs from observed precipitation and NCEP SLP (1970 to 1999). (a) Precipitation correlation map, (b) SLP correlation map and (c) time scores. See Fig. 6 for abbreviations

actually very similar to those seen as a combination of CCA 1 and 2 in the observation. Thus, on the whole, the GCM is considered to have some skill in simulating the mechanisms that control regional precipitation through SLP. In addition, since it is a combination of the 2 CCA modes which are used to define the downscaling model, the fact that there are some differences between the GCM and observed individual modes does not matter much and can be bypassed by the mixed empirical/dynamic approach described in the next section.

5. DOWNSCALING MODEL AND PROJECTION OF SUMMER RAINFALL CHANGE

5.1. Downscaling model

Based on the observed link between rainfall and SLP, the statistical downscaling model is developed with 1970 to 1999 as the training period and 1948 to 1969 as the verification.

Prior to the CCA, the original data are projected onto their EOFs in order to filter out atmospheric noise (Bar-

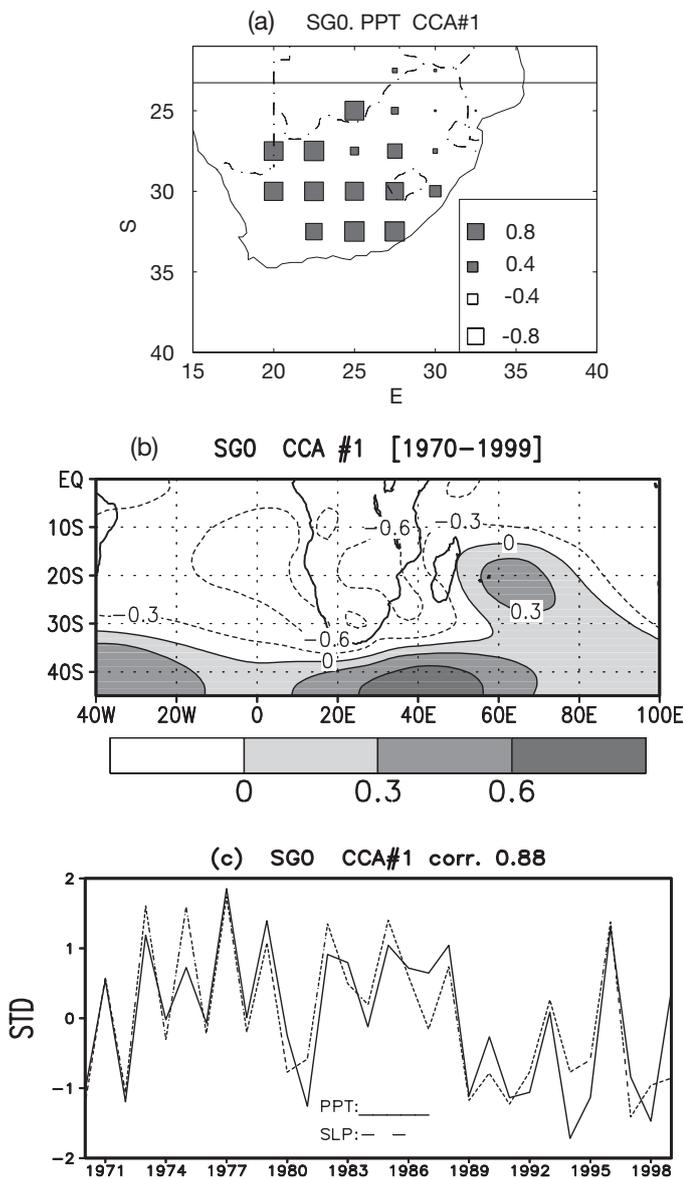


Fig. 8. First CCA pairs from ARPEGE/OPA/GELATO simulation (1970 to 1999). (a) Precipitation correlation map, (b) SLP correlation map and (c) time scores. SGO: control experiment. For other abbreviations see Figs. 1 & 6

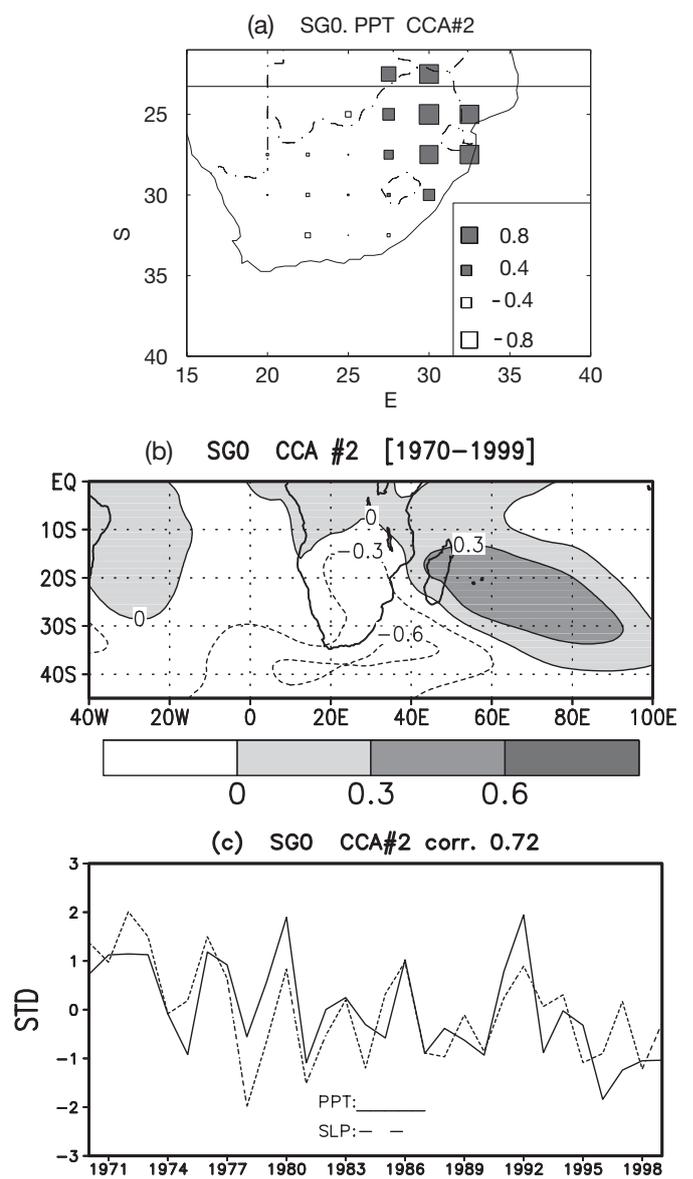


Fig. 9. Second CCA pairs from ARPEGE/OPA/GELATO simulation (1970 to 1999). (a) Precipitation correlation map, (b) SLP correlation map and (c) time scores

nett & Preisendorfer 1987). The performance of the downscaling model is sensitive to the number of EOFs retained for the CCA and the number of CCA components used in the regression model. In the present study, these numbers are determined simultaneously through training and testing skill sensitivity tests, i.e. checking that the skill of the model is high and does not substantially change after the addition of new components. Here, the skill of the downscaling model is measured by 2 indexes: the correlation between observed and reconstructed rainfall, and the variance explained by reconstructed rainfall. The combination of 2 EOF modes of precipitation (79.4%) from SAWS, 8 EOF modes of SLP (91.7%) from NCEP data, and 2 CCA modes produces the optimal skill. In the training period, the average correlation between observed and reconstructed time-series is 0.69, and the variance of reconstructed precipitation explains 37.5% of total local variance. It is interesting to note that if we replace NCEP SLP with SLP from the GCM and keep the '2-8-2' combination in the training period, the correlation between estimated and observed rainfall is 0.59 and the explained variance is 34.4%, not decreasing much compared to the skill based on NCEP data. This comparable skill is attributed to the realistic simulation of SLP in the GCM. However, the correlation is overestimated during the training period from a finite sample (e.g. von Storch 1995). Hence, we also use the independent period 1948 to 1969 (22 yr), based on SAWS rainfall and NCEP SLP data, to test the statistical model.

For the testing period, a Monte Carlo process (Livezey & Chen 1983) is used to estimate the skill level of the CCA downscaling model. This is done by randomly resampling the rainfall estimated from the CCA downscaling model for each observed district and correlating each resampled series with the observed rainfall series. After 1000 repetitions, the absolute correlations for each district obtained from the resampling process are ranked, and the 950th correlation value is the lower limit for the rainfall correlation for the particular district which is significant at the 95% level. For the period 1948 to 1969, the threshold value for a 95% confidence level is about 0.35. The skill (even based on NCEP SLP) is generally low: only 38.2% (21 out of 55) of the districts have statistically significant skill. Most of the skill is concentrated over the central interior where 14 out of 20 districts have significant skills, and the average correlation coefficient between observed and fitted time-series is 0.41. The skill is quite low in the northeast area where only 2 out of 13 districts reach statistical significance. The CCA downscaling model displays a general weak dry bias (as compared to SAWS observed rainfall) through the verifying period. The rainfall is mostly underestimated by the CCA model (about 5%) over the central interior, and slightly overestimated (less than 5%) over the northeast area.

The low skill of the CCA model is attributed to (1) lower reliability of NCEP reanalysis SLP data in the African region before 1968 (Poccard et al. 2000) and (2) decadal variability in the relationship between atmospheric circulation and southern African rainfall (Richard et al. 2000, 2001). The rainfall/SLP relationship as demonstrated by the SARI-NCEP SLP correlations during 1948 to 1967 is rather weak. For instance, the out-of-phase relationship between southern Africa rainfall and SLP over the southwest Indian Ocean cannot be identified in this period. These results partly reflect the lower quality of NCEP data for the region prior to 1968. As to the second possible reason, it is partly an inherent caveat of statistical downscaling methods based on linear regression, such as CCA procedures (Zorita & von Storch 1999).

5.2. Projection of future rainfall change

Assuming that the relationships between precipitation and SLP remain valid in the future, we can apply the CCA downscaling model (calibrated on 1970 to 1999) to project rainfall change over South Africa under the SRES B2 scenario. By solving Eq. (2), the time series related to the simulated SLP change are found and then applied to Eq. (1), which enables the future rainfall change to be projected. Fig. 10a shows the percentage difference between future (2070 to 2099) and present-day (1970 to 1999) rainfall from the direct GCM grid-point output; Fig. 10b shows that derived from the CCA downscaling model. Both of them demonstrate a drought tendency. The CCA downscaling model indicates a drier (−16.1% on average) situation than the GCM direct estimation (−8.2% on average). However, both of them are within the range of estimates presented in Hulme et al. (2001) and in IPCC (2001). The latter, partly based on Giorgi & Francisco (2000), used 5 AOGGCMs, including that of the Hadley Centre, for which 4 realizations were available. Under a 1% yr^{−1} CO₂ increase with sulphate forcing, the simulated change ranges between −25 and +10% for December to February rainfall over a southern African window larger than ours. Our analysis indicates that over the central interior rainfall is likely to decrease more than over other areas, as is indicated by both the GCM direct output and, more clearly, the downscaling model. Over the northeastern region, the downscaling model indicates a rainfall decrease of about 15%, whereas the change from the GCM direct output is very modest at less than 5%.

We have repeated this procedure for other 30-yr periods, taking 2020 to 2049 for example, and find a similar rainfall change pattern (not shown here): an overall dry tendency occurs over the domain of interest, especially the central interior. For 2020 to 2049, the average rain-

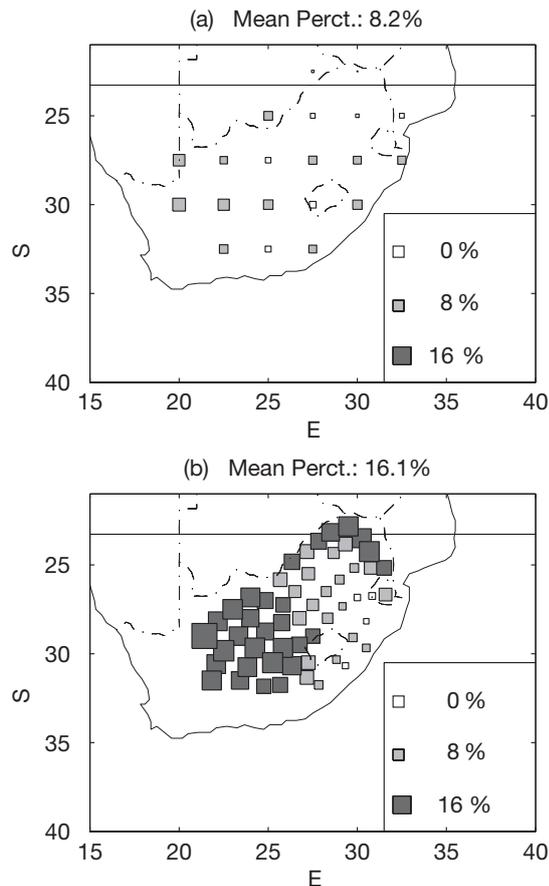


Fig. 10. Mean precipitation percentage change in South Africa between 1970 to 1999 and 2070 to 2099 from ARPEGE/OPA/GELATO (SRES-B2). (a) Direct grid-point values, (b) estimation from the CCA downscaling model. Dark grey squares: below -16% ; light grey squares: -8 to -16% ; white squares: 0 to -8%

fall decrease is 2.6% from the GCM direct estimation and 7.4% from the CCA downscaling model. Over the northeastern region, the GCM direct output indicates a very small increase, which is in contrast to an almost 5% decrease as derived from the CCA downscaling model. This is consistent with the result from the period 2070 to 2099 that more uncertainty of rainfall change exists over the northeastern part of South Africa.

The discrepancy between GCM and downscaled rainfall changes may be attributed to the fact that only atmospheric dynamics (indirectly, through SLP) are considered in the downscaling model. Rainfall amounts also depend on atmospheric water vapour content, this effect being fully addressed by the GCM. Since temperature and evaporation (at least over the ocean) show a clear rising trend during the 21st century in the simulation, the resulting enhancement of specific humidity may account for the fact that the rainfall deficit is less pronounced in the GCM than in the downscaling model. Given that only SLP-induced precipitation changes are considered

in the latter model, how confident are we about the projections? SLP patterns are well simulated by the GCM, which makes this variable a robust candidate for assessing changes in the regional climate. By contrast, evaporation over the region is largely overestimated (this problem being shared by many other GCMs, e.g. Hudson & Jones 2002). This is a result of much too high rainfall amounts, but reciprocally this impacts on the availability of water in the atmosphere. Though the downscaling model only considers purely dynamical effects and still misses some part of the rainfall variance, possibly related to other components of the water cycle, the results are more robust than if humidity was used as a predictor, given the biases in the simulation of this quantity. In addition, the fact that both the GCM gridpoints and the downscaling model indicate a moderate decrease in rainfall gives more confidence in the sign of change (which is uncertain according to previous studies; Giorgi & Francisco 2000, IPCC 2001).

Which modifications of the SLP field may explain the overall rainfall decrease? The GCM simulation displays for the 21st century a marked SLP increase over the subcontinent and a slight decrease south of 40°S and east of Madagascar (not shown). The pattern is strongly reminiscent of CCA1 (with a reversed sign) in the GCM, and is conducive to a strong divergence anomaly over the Kalahari. This leads to weakened moisture advection from the southwest Indian Ocean towards central South Africa.

6. CONCLUDING REMARKS

In this study, we first use a CCA downscaling procedure to validate the ability of a coupled GCM (ARPEGE/OPA/GELATO) to simulate the present-day climatology over the southern African region. Then we apply a CCA statistical downscaling model to project rainfall change over South Africa under the SRES-B2 emissions scenario. The major conclusions are as follows:

- The GCM simulates the general late-summer rainfall pattern over southern Africa, but overestimation is found over the continent, especially in South Africa, with almost a doubling compared to observations.
- The structure and variability of SLP are reproduced by the GCM in a fairly realistic way. Both the spatial patterns and time series of the first 2 EOFs match those from NCEP reanalysis data reasonably well. However, some discrepancies exist, such as subtropical high pressures over the south Atlantic Ocean and south Indian Ocean being weaker than observed.
- The GCM is able to capture the observed link between rainfall over South Africa and adjacent SLP, especially the pattern of anomalously low SLP over the central subcontinent and anomalously high SLP

over the southwest Indian Ocean, which implies a major controlling mechanism of rainfall.

- Both the GCM grid-point output and the CCA downscaling model indicate a drying trend in the 21st century over most parts of South Africa, in particular the central interior. Compared to present-day climatology, in the GCM used in this study, the overall rainfall at the end of 21st century is projected to decrease by 8.2% as deducted from GCM grid-point data, and by 16.1% from the downscaling model. For South Africa, these moderate decreases are consistent with Joubert & Hewitson (1997), who suggested that mean summer rainfall will decrease by between 10 and 20%. While such a change may appear small, it has to be considered in combination with the enhanced evaporation associated with warming, as well as the increased water requirements in the context of population rise and economic growth.

Large-scale circulation can only explain part of rainfall variance (37.5% in the CCA downscaling model). Some regional processes such as local soil-moisture feedback also contribute part of rainfall variability which, however, cannot be included in the downscaling model. In addition, this CCA model mainly depends on only one possible climate state (the present-day state in this case) without accounting for much of the decadal variability. All of these issues create uncertainty in the projected rainfall change from the downscaling model. The CCA downscaling procedure should be applied to several runs (to account for varying initial conditions), emissions scenarios and to other GCMs (to evaluate consensus between GCMs), and hence strengthen the levels of confidence for the regional projection.

Decadal variability of observed rainfall is a relatively well known feature for the 20th century; in particular a 18-yr oscillation is common throughout southern Africa (Tyson 1986). It is not verified for most other climate variables, whose time-series only cover 30 to 50 years. The GCM used here exhibits a 60-yr rainfall cycle, which is not evident in the observation, but no 18-yr periodicity. Decadal variability may also be sensitive to the SRES scenario, though the GHG evolution is fairly steady in scenario B2, which is the one used here. It is found that the regional projection for the mid 21st century agrees, in a more or less linear way, with that for the late 21st century, which brings some confidence to our conclusions because the consistency in the trend of precipitation change is likely to be the effect of the rather linear increase in GHG in the B2 scenario.

Nevertheless, the CCA downscaling model may still need refining. The model skill could be improved by adding other parameters, such as water vapour or winds, but this also probably indicates increasing uncertainty in the projection of future climate change

because these parameters are also subject to large errors in GCM simulations. Additionally, data availability (e.g. water vapour) is likely to limit this option. The empirical downscaling approach should also be developed alongside the dynamical modelling approach. As regional climate modelling skills and techniques develop, the 2 approaches should be evaluated against each other for requirements of both researchers and policy planners.

Acknowledgements. We would like to thank Dr. Eduardo Zorita from Institute of Hydrophysics, GKSS Forschungszentrum, Geesthacht, Germany, who provided some of the downscaling codes. We acknowledge the help given by Dr. E. Roeckner from MPI for Meteorology, Hamburg, Germany for several discussions. We are also grateful to J.F. Royer, F. Chauvin, H. Douville, and S. Tyteca from CNRM (Météo-France Toulouse) who provided the scenario experiment data. This study was supported by the French Education and Research Ministry grant ACI No. 67625 entitled 'Rainfall Variability in South Africa'.

LITERATURE CITED

- Arnell NW (1999) A simple water balance model for the simulation of streamflow over a large geographic domain. *J Hydrol* 217:314–335
- Ashrit RG, Douville H, Rupa Kumar K, Royer JF, Chauvin F (2003) Response of the Indian Monsoon and ENSO-monsoon teleconnection to enhanced greenhouse effect in the CNRM coupled model. *J Meteorol Soc Jpn* 81:759–777
- Barnett TP, Preisendorfer RW (1987) Origins and levels of monthly and seasonal forecast skill for United States air temperature determined by canonical correlation analysis. *Mon Weather Rev* 115:1825–1850
- Boucher O, Lohmann U (1995) The sulfate-CCN-cloud albedo effect: A sensitivity study with two general circulation models. *Tellus (Ser B)* 47:281–300
- Busuioac A, Deliang C, Hellstrom C (2001) Performance of statistical downscaling models in GCM validation and regional climate change estimations: application for Swedish precipitation. *Int J Climatol* 21:557–578
- Camberlin P, Chauvin P, Douville H, Zhao Y (2004) Simulated ENSO-tropical rainfall teleconnections in present-day and under enhanced greenhouse gases conditions. *Clim Dyn* 23:641–657
- Crane RG, Hewitson BC (1998) Doubled CO₂ precipitation changes for the Susquehanna Basin: downscaling from the GENESIS general circulation model. *Int J Climatol* 18: 65–76
- Cui M, Von Storch H, Zorita E (1995) Coastal sea level and the large-scale climate state: a downscaling exercise for Japanese Islands. *Tellus* 47A:132–144
- D'Abreton PC, Lindesay JA (1993) Water vapour transport over southern Africa during wet and dry early and late summer months. *Int J Climatol* 13:151–170
- Déqué M, Dreveton C, Braun A, Cariolle C (1994) The ARPEGE/IFS atmosphere model, a contribution to the French community climate modelling. *Clim Dyn* 10: 249–266
- Déqué M, Braun A, Piedelievre JP, Marquet P and 8 others (1999) Documentation ARPEGE-CLIMAT. Centre National de Recherches Météorologiques, Météo-France, Toulouse

- Douville H, Planton S, Royer JF, Stephenson DB, Tyteca S, Kergoat L, Lafont S, Betts RA (2000) Importance of vegetation feedback in doubled-CO₂ time-slice experiments. *J Geophys Res* 105:14841–14861
- Douville H, Chauvin F, Planton S, Royer JF, Salas-Melia D, Tyteca S (2002) Sensitivity of the hydrological cycle to increasing amounts of greenhouse gases and aerosols. *Clim Dyn* 20:45–68
- Fauchereau N, Trzaska 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. *Int J Climatol* 23: 663–677
- Giorgi F, Francisco R (2000) Uncertainties in regional change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM. *Clim Dyn* 16:169–182
- Giorgi F, Mearns MR (1991) Approaches to the simulation of regional climate change: a review. *Rev Geophys* 29: 191–216
- Hewitson BC, Crane RG (1998) Regional scale daily precipitation from downscaling of data from the GENESIS and UKMO GCMs. In: Proc 14th conf probability and statistics in the atmospheric sciences, Phoenix, AZ, 11–16 Jan 1998. American Meteorological Society, Boston, MA, p J48–J50
- Hewitson BC, Joubert A (1998) Climate downscaling: current South African projections. www.egs.uct.ac.za/fccc/
- Hudson DA, Jones RG (2002) Simulations of present-day and future climate over Southern Africa using HadAM3H. Hadley Centre technical note 38, Bracknell
- Hulme M (1994) Validation of large-scale precipitation fields in general circulation models In: Desbois M, Desalmand F (eds) Global precipitations and climate change. NATO ASI Series. Springer-Verlag, Berlin, p 387–406
- Hulme M, Doherty R, Ngara T, New M, Lister D (2001) African climate change: 1900–2100. *Clim Res* 17:145–168
- IPCC (Intergovernmental Panel on Climate Change) (2001) Climate change: the scientific basis. Contribution of Working Group I to the Third Assessment Report of the IPCC. Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Xiao D (eds) Cambridge University Press, Cambridge
- Joubert AM, Hewitson BC (1997) Simulating present and future climates of southern Africa using general circulation models. *Prog Phys Geog* 21:51–78
- Joubert AM, Katzfey JJ, Mcgregor JL, Nguyen KC (1999) Simulating mid-summer climate over Southern Africa using a nested regional climate model. *J Phys Geog* 21: 51–78
- Jury MR (1992) A climatic dipole governing the interannual variability of convection over the SW Indian Ocean and SE Africa region. *Trends Geophys Res* 1:165–172
- Karl TR, Wang WC, Schlesinger ME, Knight RW, Portman D (1990) A method of relating general circulation model simulated climate to the observed local climate. Part I: Seasonal statistics. *J Clim* 3:1053–1079
- Kalnay E, Kanamitsu M, Kistler R, Collins W and 18 others (1996) The NCEP/NCAR 40-year Reanalysis Project. *Bull Am Meteorol Soc* 77:437–471
- Landman WA, Goddard L (2002) Statistical recalibration of GCM forecasts over Southern Africa using model output statistics. *J Clim* 15:2038–2055
- Landman WA, Tennant WJ (2000) Statistical downscaling of monthly forecasts. *Int J Climatol* 20:1521–1532
- Latif M, Sperber K, Arblaster J, Braconnot P and 30 others (2001) ENSIP: the El Niño simulation intercomparison project. *Clim Dyn* 18:255–276
- Lindesay JA (1988) South African rainfall, the Southern Oscillation and a southern hemisphere semi-annual cycle. *J Climatol* 8:17–30
- Livezey RE, Chen WY (1983) Statistical field significance and its determination by Monte Carlo techniques. *Mon Weather Rev* 111:46–59
- Mahfouf JF, Manzi AO, Noilhan J, Giordani H, Déqué M (1995) The land surface scheme ISBA within the METEO-France climate model ARPEGE. Part I: implementation and preliminary results. *J Clim* 8:2039–2057
- Miron O, Tyson PD (1984) Wet and dry conditions and pressure anomaly fields over South Africa and the adjacent oceans, 1963–1979. *Mon Weather Rev* 112:2127–2132
- Nakicenovic N, Swart R (2000) Special report on emissions scenarios. Cambridge University Press, Cambridge
- Nicholson SE, Kim J (1997) The relationship of the ENSO to African rainfall. *Int J Climatol* 17:117–135
- Noilhan J, Mahfouf JF (1996) The ISBA land surface parameterization scheme. *Global Planet Change* 13:145–159
- Poccard I, Janicot S, Camberlin P (2000) Comparison of rainfall structures between NCEP/NCAR reanalyses and observed data over tropical Africa. *Clim Dyn* 16:897–915
- Reason CJC (1998) Warm and cold events in the southeast Atlantic/southwest Indian Ocean region and potential impacts on circulation and rainfall over southern Africa. *Meteorol Atmos Phys* 69:49–65
- Richard Y, Trzaska S, Roucou P, Rouault M (2000) Modification of southern African rainfall variability/El Niño Southern Oscillation relationship. *Clim Dyn* 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. *Int J Climatol* 21:873–885
- Richard Y, Camberlin P, Fauchereau N, Mulenga H (2002) Cohérence intrasaisonnière de la variabilité pluviométrique interannuelle en Afrique du Sud. *L'Espace Géographique* 31:63–72
- Rouault M, Richard Y (2003) Intensity and spatial extension of drought at different time scales since 1921 in South Africa. *Water SA* 29:489–500
- Royer JF, Cariolle D, Chauvin F, Déqué M and 11 others (2002) Simulation des changements climatiques au cours du XXI^e siècle incluant l'ozone stratosphérique (Simulation of climate change during the 21st century including stratospheric ozone). *C R Geoscience* 334:147–154
- Salas-Mélia D (2002) A global coupled sea ice-ocean model. *Ocean Model* 4:137–172
- Terray L, Thual O, Belamari D, Déqué M, Dandin P, Delecluse P, Levy C (1995) Climatology and interannual variability simulated by the ARPEGE-OPA coupled model. *Clim Dyn* 11:487–505
- Tyson PD (1986) Climatic change and variability in Southern Africa. Oxford University Press, Cape Town
- von Storch H (1995) Spatial patterns: EOFs and CCA. In: von Storch H, Navarra A (eds) Analysis of climate variability: application of statistical techniques. Springer-Verlag, Berlin, p 227–258
- von Storch H, Zorita E, Cubasch U (1993) Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *J Clim* 6: 1161–1171
- Xie P, Arkin PA (1996) Analysis of global monthly precipitation using gauge observations, satellite estimates, and numerical model precipitations. *J Clim* 9:840–858
- Zorita E, Von Storch H (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *J Clim* 12:2474–2489