A Methodology for Developing Regional Climate Change Scenarios from General Circulation Models

BC Hewitson

Report to the Water Research Commission
by the
Department of Environmental and Geographical Science
University of Cape Town

WRC Report No 594/1/97

A METHODOLOGY FOR DEVELOPING REGIONAL CLIMATE CHANGE SCENARIOS FROM GENERAL CIRCULATION MODELS

by

BC HEWITSON

Dept of Environmental and Geographical Science
University of Cape Town
Cape Town

Report to the Water Research Commission on the Project
"Regional climate change scenarios for precipitation and temperature
from General Circulation Models"

WRC Report No : 594/1/97
ISBN No : 1 86845 249 2

A METHODOLOGY FOR DEVELOPING REGIONAL CLIMATE CHANGE SCENARIOS FROM GENERAL CIRCULATION MODELS

B.C. Hewitson
University of Cape Town

Executive Summary

Background and Motivation

South Africa's climate is subject to high inter-annual variability and presents significant vulnerabilities in the face of probable global climate change. Regional impacts from global changes are most likely to be manifest in the hydrological aspects which will impact water resources and water dependent infrastructures most strongly. In the face of the long lead times required for planning, and the conflicting demands placed on policy developers and politicians, it is vital that an accurate understanding be developed of the probable regional consequences of global environmental change for the next five to ten decades.

The most viable tools for such investigations are General Circulation Models (GCMs) which attempt to simulate the global geophysical system encompassing the land, ocean, and atmosphere. However, while GCMs manage to simulate global-scale processes with considerable success, GCMs are also relatively coarse resolution tools and fail to resolve the detail of regional climate variability, especially with regard to hydrological variables. It is particularly at these higher resolution scales which the GCM fails to adequately capture that climate change information is most needed.

This shortcoming may be addressed through the use of *downscaling* which utilises the dependence of local-scale climates on the large scale circulation to derive regional information. In essence, downscaling uses observed data to derive relationships between the larger scale atmospheric circulation and local climate parameters. These relationships may then be used in conjunction with the more viable large scale circulation from GCMs to derive local climate information based on the GCM simulation of atmospheric circulation.

Objectives

The primary aim was to derive regional scenarios for global climate change impacts on regional precipitation and temperature over South Africa as a consequence of anthropogenic increases of atmospheric CO_2. In doing do, three objectives were targeted:

1. To validate GCM simulations of atmospheric controls on regional precipitation and temperature over South Africa.
2. To develop a new non-linear methodology for cross-scale climate relationships.
3. To develop quantified estimates of the regional changes for use as input to hydrological and agricultural models.

Results and Conclusions

This project developed a new downscaling procedure for South Africa based on Artificial Neural Nets (ANNs) to derive relationships between circulation and local precipitation and temperature on a 1° latitude-longitude grid. These derived relationships were evaluated for accuracy, and demonstrated that the relationships do successfully capture both short and long term variability in the regional climates as a function of the larger scale circulation. For both the downscaled precipitation and downscaled temperatures, the nature of the near-daily event is well captured, as well as the seasonal timing and amplitude for both variables.

GCM circulation data were then evaluated for validity over South Africa using a Principal Components Analysis (PCA) technique. The circulation in terms of the spatial distribution of variance was shown to be a reasonable reflection of present day circulation. The ANN relationships between circulation and local climate were subsequently applied to the GCM daily circulation data from a simulation of present day conditions, and the regional climates derived. The regional climates generated by the GCM circulation were shown to demonstrate a high degree of similarity to the observed regional climates and further indicates that the GCM reasonably simulates the large scale circulation forcing over South Africa. The relationships were then applied to data from a GCM simulation of future conditions under doubled atmospheric CO_2, and the results differenced from the present day simulation to generate the CO_2 induced changes in regional climates.

The primary results of consequence for climate change impacts indicate that, as a function of changes in the GCMs circulation forcing, there are reductions in the regional summer precipitation over the summer rainfall region extending from the east coast through to the central and northern regions of the country. These reductions are of the order of a 10-15% decrease in mean monthly precipitation. However, the results are preliminary, based on only one model, and should be used only as an initial possible indication of change.

Primarily this work has demonstrated the viability and value of the methodology developed, and that the empirical downscaling technique is a valuable approach toward developing more robust climate change scenarios at suitable spatial and temporal scales for use by South African climate change impacts researchers. The objectives of the project have thus been well met, in particular with respect to the evaluation of the GCM performance and development of the downscaling methodology. In addition, the primary aim of deriving regional climate change scenarios was achieved. However, these should be seen as one models projection for the future, and the methodology now needs to be applied more widely to evaluate the consensus of results among different models.

The methodology offers a technique by which researchers may derive local climate data as a function of the GCM predicted global change, and offers a significant advancement over the climate change data currently available from other sources. Further results based on the methodology should in the future be of direct interest to South African researchers in a wide spectrum of disciplines. While the initial scenarios developed here are consistent with broad expectations of climate change impacts, they should not be considered too definitive as further work is needed to develop confidence levels for the scenarios.

Recommendations

While follow-on projects have already been initiated, these techniques should continue to be developed and take advantage of new model results that become available. While the techniques applied here are at the forefront in this field, there is still a large scope for improvement to narrow the gap between the form of climate change data output from

GCMs, and the needs of other impacts researchers. It is recommended that the following are specifically addressed in the further work:

- The methodology should be refined with regard to the specifics of including atmospheric water vapour in the relationships, adding direct radiative forcing values to derived temperatures, and improving the spatial and temporal resolution of the scenarios.
- The procedure should be applied to other GCMs in order to evaluate consensus between GCMs, and hence strengthen the levels of confidence for the regional scenarios.
- The results need to be further analysed in terms of the fundamental dynamical controls that give rise to the regional climate change indicated by the downscaling.
- The results of new scenarios should be made available to researchers, particularly in the hydrological and agricultural fields, for further application to understanding the impacts of the regional climate change on society. In addition, the essence of the results should be used to further inform the development of policy and planning at the political level in South Africa.
- As a longer term agenda, the empirical downscaling should be developed alongside the more computationally demanding nested modelling approach (only practical with international collaboration given limited computing resources in South Africa). As nested modelling skills and techniques are developed, the two approaches should be evaluated against each other for suitability to the research requirements of the South African community and for policy developers and planners.
- The methodology developed is computationally suited to the infrastructures in neighbouring countries, and extension of the technique to other southern Africa nations and scientists could assist in developing a broader picture of future regional change, as well as foster stronger collaborative ties with the other countries.

ACKNOWLEDGEMENTS

The research in this report emanated from a project funded by the Water Research Commission and entitled:

"Regional climate change scenarios for precipitation and temperature from General Circulation Models"

The steering Committee responsible for this project consisted of the following persons:

Dr. G.C. Green	Water Research Commission (Chairman)
Prof. P.D. Tyson	University of the Witwatersrand
Ms. H.A. Riphagen	SA Weather Bureau
Prof. R.E. Schulze	University of Natal
Mr S. van Biljon	Department of Water Affairs and Forestry
Dr. B.C. Hewitson	University of Cape Town
Mr. F.P. Marais	Water Research Commission (Committee Secretary)
Mr. H. Maaren	Water Research Commission
Prof. R.F. Fuggle	University of Cape Town

The financing of the project by the Water Research Commission and the contribution of the members of the Steering Committee is acknowledged gratefully.

Table of Contents

A METHODOLOGY FOR DEVELOPING REGIONAL CLIMATE CHANGE SCENARIOS FROM GENERAL CIRCULATION MODELS

1: Introduction

A primary objective of climatology and the analysis of atmospheric dynamics is the application of climate data and atmospheric analysis to issues related to society, and to those aspects of the physical environment that impact on the human system. For example, the motivation behind research ranging from daily forecasting through to drought and extreme event analysis is invariably based on society's need to accommodate and adapt to the impacts of the climate system. This is very evident in the societal need to respond to global environmental change, along with the need to anticipate probable future changes in the climate system. While for some regions of the globe this is of more critical concern than others (for example, poorer nations with more fragile economies and agricultural infrastructures), there is nonetheless a global need for understanding the possible impacts of climate change and variability.

In contrast to the global nature of the atmospheric forcing, the focus when considering the societal response is on regional impacts and the local consequences of large-scale change. It is in this area, however, that analysis techniques are weakest. Although GCMs, the primary tools for investigating global change, provide reasonable simulation accuracy of present climate when viewed from global, hemispheric, and continental scales, data at the models' finest scale of representation (typically 2 - 4 degrees of latitude or longitude) are often found to be highly erroneous, especially in the important area of hydrology. The *skill level* of the model (von Storch et al., 1993) is substantially different from the actual model resolution, with the skill level for many parameters increasing as they are averaged over a larger and larger grid (Grotch and MacCracken, 1991). In contrast, the primary needs of a society attempting to plan for or respond to climate change lie at far finer spatial resolutions (e.g. water catchment basins). For example, while GCM accuracy decreases as spatial scales become finer, in hydrological, as in most other applications, one usually needs *increasing* detail as spatial scales decrease (e.g. information on individual rainfall events in a catchment basin to determine runoff and storage); consequently, the scales of accuracy of GCMs are in direct opposition to that of societal need. Typical needs of researchers dealing with climate change impacts are summarised in Table 1.

Climate Variables	**Spatial Scales**	**Temporal Scales**
Temperature Precipitation Winds Radiation Water Vapour Clouds Snow	100km and better	Event frequency, persistence, intensity, and recurrence at daily through seasonal and inter-annual scales.

Table 1: Scales and variables of primary concern to climate change impact researchers.

A number of approaches can be adopted to meet the needs identified in Table 1. At present, these fit into two categories: process based techniques, involving the explicit solving of the physical dynamics of the system; and empirical techniques that use identified system relationships derived from observational data. Either approach may be adopted for developing regional climate change information, and both have advantages and shortcomings. Common to both processes, however, is the concept of *downscaling*. This concerns the derivation of local-scale information from the larger scale circulation dynamics and is based on the proposition that the dominant character of local climates is driven by the synoptic scale.

In the context of climate change research, process based downscaling techniques are generally found in the form of nested modelling, where high resolution models with limited domains are driven by larger domain and coarser resolution dynamical models. Conversely, empirical downscaling has diverse forms ranging from simplistic weather typing through to a range of more sophisticated quantitative techniques such as employed here.

Nested modelling is probably the most viable technique in the long term, and has the potential to provide detailed information at spatial scales down to 10-20 km and temporal scales of hours or less. However, the technique has significant drawbacks, especially in the context of the South African research community. Most importantly, such models are computationally very demanding and are, therefore, not an easily accessible research avenue within the context of the South African infrastructure. Furthermore, there are still difficulties to be overcome in the interface between the GCM and the nested model (e.g. how to relate the coarse resolution grid cell data of the GCM, often below the GCM skill scale, to boundary conditions for the far finer scale nested model).

In contrast, empirical downscaling offers a useful and computationally inexpensive means to address the immediate needs of the South African regions, and in particular to assist in the development of research on societal vulnerability in the face of global environmental change. Thus this is a practical approach for addressing current needs in the South African climate change community, and also in many of the countries liable to be most sensitive to climate change impacts.

In the empirical approach one seeks to derive quantitative relations in some form of y=f(x), using atmospheric circulation (x) as the independent forcing to derive the local climate (y). The transfer function is derived from observational data using a mathematical or statistical relationship. More traditional forms of synoptic climatology rely on statistically generalised functions that categorise the circulation features and their associated climate characteristics; consequently, they are constrained to some degree by the assumptions of the statistical model used and the large within-group variability that generally occurs. In contrast, the methodology outlined here derives a direct mathematical relationship between the circulation and the local climate using Artificial Neural Networks (ANNs). This removes some of the constraints noted above, and treats the data as the continuum it is. Furthemore, ANNs also have the ability to capture some of the non-linear aspects of the relationships between circulation and local climate.

In global change analysis, these cross-scale relationships are derived from gridded observational data and then applied to a GCM circulation data set to derive local-scale information consistent with the synoptic-scale forcing of the GCM. This highlights an important additional application of downscaling in GCM validation. Through the application of transfer functions (derived from observed data) to the GCM climate, one implicitly tests the internal cross-scale consistency and synoptic scale forcing of the GCM on regional climates. Applying the same observation-derived transfer function to GCM climate change experiments then allows for the development of regional climate change scenarios at sub-GCM grid-scale resolution, with greater reliability than is possible using the GCM results alone.

As the relationship is based on observational data sets, the derived local scale information is thus *consistent* with the synoptic scale forcing -- the scale at which GCMs perform well. However, for application to GCM data and climate change issues, the approach hinges on three basic premises:

1. that the local scale parameter is dominantly a function of synoptic forcing,
2. that the GCM circulation used to drive the derived relationship is valid at the synoptic scale,
3. that the relationship derived using observed data remains valid under a $2xCO_2$ atmosphere.

The first premise is implicitly tested in the derivation of a relationship. The second premise is the subject of much ongoing research and there are a number of procedures to draw on (e.g.: Hewitson and Crane, 1992). For this project, the second premise is tested through the use of Principal Components Analysis (PCA) as outlined later. The third premise, that the relationship holds under a $2xCO_2$ atmosphere is the more difficult to justify, and without waiting for the time when atmospheric CO_2 has doubled, is in fact impossible to prove one way or another. Nonetheless, given that basic dynamics of the atmosphere are not about to be transformed, it would seem a reasonable assumption. The major caveat in this regard is in the case of precipitation where if the atmospheric water vapour content were to be dramatically altered the relationship may significantly change. However, if the primary atmospheric response will be in terms of changes of persistence, intensity, and frequency of synoptic circulation events (as seems likely), then the third premise may be accepted, at least in the interim.

2. Artificial Neural Nets

The derivation of the downscaling relationship becomes all important if this approach is to succeed. In light of this, and that the relationships may well be non-linear in nature, techniques more sophisticated than, for example, simple regression, are needed. Artificial neural nets (ANNs) have particular strengths in this regard.

The concept of ANNs is itself quite basic, and many analogies in the statistical world capture one or other aspect of the ANN. However, the ANN goes further than these techniques in many areas. Consider for the moment an ANN as a black box with inputs and outputs, and which performs some function for mapping the input to an output. Accept that initially the net is in some random state, "untrained," and represents a random function. The first step then is to adapt, or train, the net to learn some mapping/relationship between the input and output. This is accomplished through presenting the net with a sample of known inputs and outputs, which, in conjunction with a learning algorithm, modifies the internal function performed by the net to find a relation between the input and output. Thereafter, as long as one remains within the bounds of the training samples, the net can be applied in a similar manner to further unseen data.

Continuing for the moment with treating the ANN as a black box, a useful analogy is to consider the net as a super-form of multiple regression--in fact, linear multiple regression is a special case of neural nets. In the same way as one performs regression between {X} and {Y}, developing a relation such that $\{y\} = f\{x\}$, so does a net find some function f{x} when trained. The attractiveness of ANNs lies in part in their ability to, in theory, represent any arbitrary non-linear function. Whereas in regression one is tied to a linear relationship, or at best a pre-specified non-linearity, the neural net finds its own function as best it can, given the complexity used in the net, and without the constraint of linearity.

Other attractions of ANNs compared to their more traditional counterparts are the ability to generalise a relationship from only a small subset of data, to remain relatively robust in the presence of noisy inputs or missing input parameters, and to adapt and continue learning in the face of changing environments. In short, the ANN is highly flexible. Finally, the net is not so much a black box--an unfairly acquired reputation--as a 'grey' box, and techniques are available to interpret the function it represents.

Looking inside the grey box of an ANN it is easiest to use a biological analogy. As ANNs were initially developed as models of biological neural systems they retain, in common with the brain, the same basic macro-structure. In the same way as the brain is composed of inter-linked processing elements (neurons), so the ANN has simple processors (nodes) connected by weighted links. Figure 1 shows what a simple node in an ANN might look like.

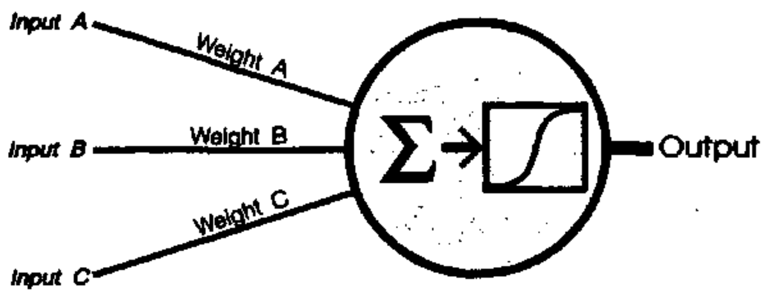

Figure 1: Example of a node within a Neural Net

A node in the ANN sums the weighted inputs from the links feeding into it, and performs some function on the summed value. This function is typically a non-linear bounded function, although a wide variety of options exist, including threshold functions, binary functions, and linear functions. The output is then given to further weighted links leading to other units, and many nodes may be connected together to form a net of processing nodes. Figure 2 shows a simple configuration, termed feed-forward, in which nodes (with possibly different functions) are connected such that the information flows unidirectionally from input to output. The net as a whole will have some input point(s) and some point(s) of output.

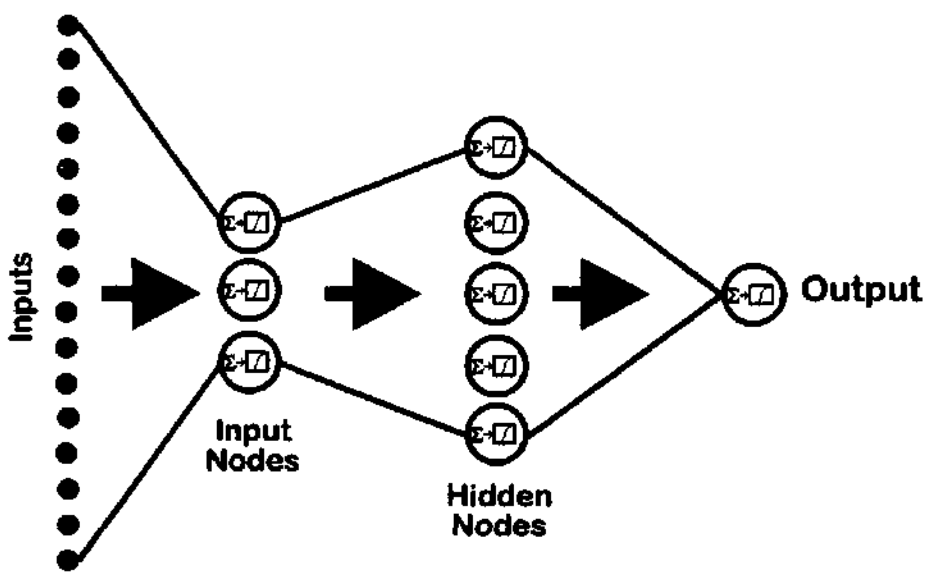

Figure 2: Example of a simple "feed-forward" Neural Net

Typically the nodes are placed in layers, with an input and output layer of nodes, and usually one or more hidden layers (in the sense that they have no connections external to the net). The example shown above in Figure 2 has one hidden layer, and has full connectivity between layers. As can be easily imagined, great complexity is available, with the possibility for links to leapfrog layers, connect to nodes within the same layer, and even feed back to earlier layers.

3. Data and Analysis

In downscaling a number of choices may be made about the spatial and temporal scales to be used. In terms of deriving a relationship of the form $\{y\} = f\{x\}$, the atmosphere forms the independent variables $\{x\}$ whose scale needs to be large enough to capture the synoptic forcing features which are represented by the GCM. However, in the case of the dependent variable $\{y\}$, for example precipitation, one can choose to work on a wide range of scales subject to the constraint that the scale chosen retains a significant relationship to the synoptic forcing. Obviously the finer the spatial or temporal scales the more local effects will be an influence and the less clear will be the synoptic forcing.

In this example a range of spatial and temporal scales were investigated ranging temporally from daily through to monthly, and spatially from 0.25° to 2.5° of latitude/longitude. While initially it was intended to work at two spatial scales, after investigation it was decided to work on the spatial scale of 1° by 1°, and the temporal scale of daily data smoothed with a 5 day filter.

The observed data used for deriving the downscaling relationships was comprised of station derived area average daily precipitation, temperature minimums and maximums, and Sea Level Pressure (SLP) and 500hPa geopotential height (gpm) fields from the Goddard Space Flight Center (GSFC) gridded analyses. The GCM data are from the Genesis GCM v1.02, a derivative of the USA National Center for Atmospheric Research (NCAR) CCM model and incorporates a 2° x 2° land surface model and multi-layer soil and ice models.

The methodology employed is as follows:

- Generation of 1° precipitation and temperature data from station data.
- Validation of the GCM circulation fields using PCA.
- Derivation of downscaling functions for each 1° grid location using ANNs with observed data.
- Validation of the derived relationships by generating regional precipitation from the observed circulation and comparing with the observed precipitation.
- Deriving regional parameters from the GCM control simulation circulation data and comparing with observed circulation derived regional variables -- a GCM validation procedure.
- Deriving regional parameters from the 2xCO_2 simulation circulation. This is then differenced from parameters derived off the control simulation circulation to generate regional climate change scenarios consistent with changes in synoptic forcing under a doubled CO_2 atmosphere.

3.1 Regional observed precipitation and temperature data

The observed precipitation and temperature data are derived from the data base at the Computing Centre for Water Research (CCWR) at the University of Natal, Pietermaritzburg. Daily data for all stations in the data base were extracted and all suspect observations discarded. A number of interpolation routines were evaluated in order to regrid the station data onto a regular grid, including the chloropleth nearest neighbour algorithm which was anticipated to be effective. However, on visual comparison of time series and spatial patterns simple area averaging appeared as effective at significantly lower computational expense. Simple area averaging was thus used to derive gridded data on a 1° by 1° grid with the following constraints:

- A grid cell observation was deemed undefined if less than three stations were available with valid data on the day in question.
- A station had to be have at least one year of data record in order to be included.

On any given day between ~3000 and ~4000 stations across the country were available for precipitation, and approximately half that number for temperature.

As temperature is a spatially continuous field it was considered justified to then interpolate the 1° gridded temperature data to those grid cells with undefined data as a function of the constraints above. The interpolation was based on inverse square distance using spherical distances, modified to consider the angular distribution of data points around the location being interpolated to. Given that precipitation is not spatially continuous, interpolation was not performed. Nonetheless, for both precipitation and temperature this procedure provides data for essentially the entire country, and a significant portion of Namibia. Figure 3 shows an example of the spatial coverage and the 1° resolution for observed precipitation.

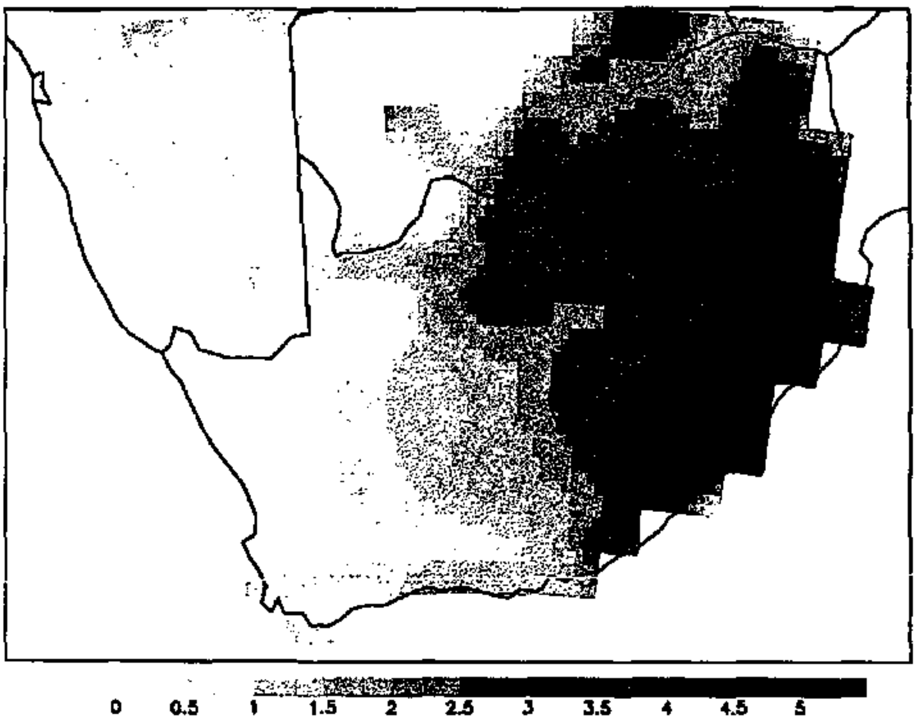

Figure 3: Example of spatial coverage of 1° resolution precipitation (December, January, February, average mm/day).

The relationship between local climate and synoptic circulation is a function of both temporal and spatial resolution. The finer the temporal or spatial resolution, the more local effects dominate each observation in terms of unique information. Through testing different resolutions it was decided that daily data smoothed with a 5 day filter provided the best temporal resolution given a 1° spatial grid and the constraint of only SLP and 500hPa gpm as synoptic circulation variables. While initial investigations suggested that the inclusion of atmospheric water vapour would enhance the accuracy of predicting local climate as a function of the circulation, the water vapour data requested through the GSFC database in the USA was not available in time for the closing stages of this project. This aspect will be addressed in 1997 as a part of a follow-up two year WRC project, and which will also take advantage of the new GCM model available.

3.2 Circulation data and GCM validation.

The GCM used was the Genesis v1.02 GCM (now superseded by version 2.0) developed at NCAR, and run on the Pennsylvania State University Cray supercomputer. This version of the model was run at R15 spectral resolution (~5° latitude by 7° longitude) with a 2° by 2° land surface model, a multi-layered soil and ice model, plume convection scheme, and mixed layer ocean with prescribed horizontal heat transport.

The GCM validation was oriented around the need for the GCM synoptic scale circulation features to be a reasonable representation of reality in order to justify it's use as input in a downscaling function. Of primary importance is the spatial characteristics of the variance in the synoptic field. Validation using PCA components has already been employed in a number of cases whereby the observed and GCM data are reduced to independent sources of variance through the use of rotated Principal Components Analysis (PCA), and the resulting component spatial loading patterns compared. The parameters to be used to describe the circulation, SLP and 500hPa gpm for both observed gridded and GCM data, were extracted on a daily time frame for the entire globe. A further subset of the data was extracted for a window centred over South Africa extending from 44°S to 16°S, and 0°E to 50°E.

The same variables were also extracted from the GSFC gridded assimilation data, and in this case re-interpolated to the GCM grid (using inverse square distance and spherical distances). For each of the observed and GCM data sets, S-mode PCA was performed to dimensionally reduce the data. This resulted in 6 and 7 component score time series for each of the SLP and 500hPa gpm data sets respectively.

For each PCA the validity of the components retained was checked using Priesendorfer's Rule-N test. The GCM and observed component loading patterns are then correlated with each other using a three way correlation: point-wise correlation, correlation of the north-south gradients, and correlation of the west-east gradients. The average correlation is then used as an index of similarity between the observed and GCM data. As the component loading maps represent distribution of coherent variance in the data, this procedure evaluates the spatial behaviour of the circulation between the GCM and observed data. In all cases the circulation was found to be a reasonable match between the observed and GCM data for both sea level pressure and 500hPa gpm. Thus the GCM circulation is judged to be spatially realistic over the southern Africa domain.

Table 1 below gives an example of the correlations between components for the sea level pressure. These are derived through averaging the pointwise correlation of the loadings, and the north-south and east-west gradients of the loading patterns. Overall the GCM is seen to match the distribution of variance in the observed data well. Figure 4 shows an example of the component loading patterns for two matching patterns between the observed and GCM data.

GSFC Component	GCM Component	Average r
3	2	0.87
1	4	0.78
2	1	0.72
4	3	0.72
5	5	0.69

Table 2: Correlations between PCA loading patterns of observed and GCM SLP data.

Note that this validation procedure does not test the temporal validity of the GCM data, however this is intrinsically evaluated in using the downscaling functions, as temporal error will show up in the downscaled variable (temperature or precipitation).

3.3 Derivation and validation of downscaling relationships

The downscaling relationships between circulation and precipitation/temperature were developed with standard feed-forward backpropogation ANNs. For each grid location a separate ANN was trained for each of precipitation, minimum temperature, and maximum temperature. Inputs to the ANN were daily time series of the SLP and 500hPa gpm PCA scores smoothed with a 5 day running average filter, and the target outputs were formed by the 5 day smoothed daily precipitation and temperature. In addition to the temporal smoothing, the input circulation scores are time-lagged over a 5 day period in order to

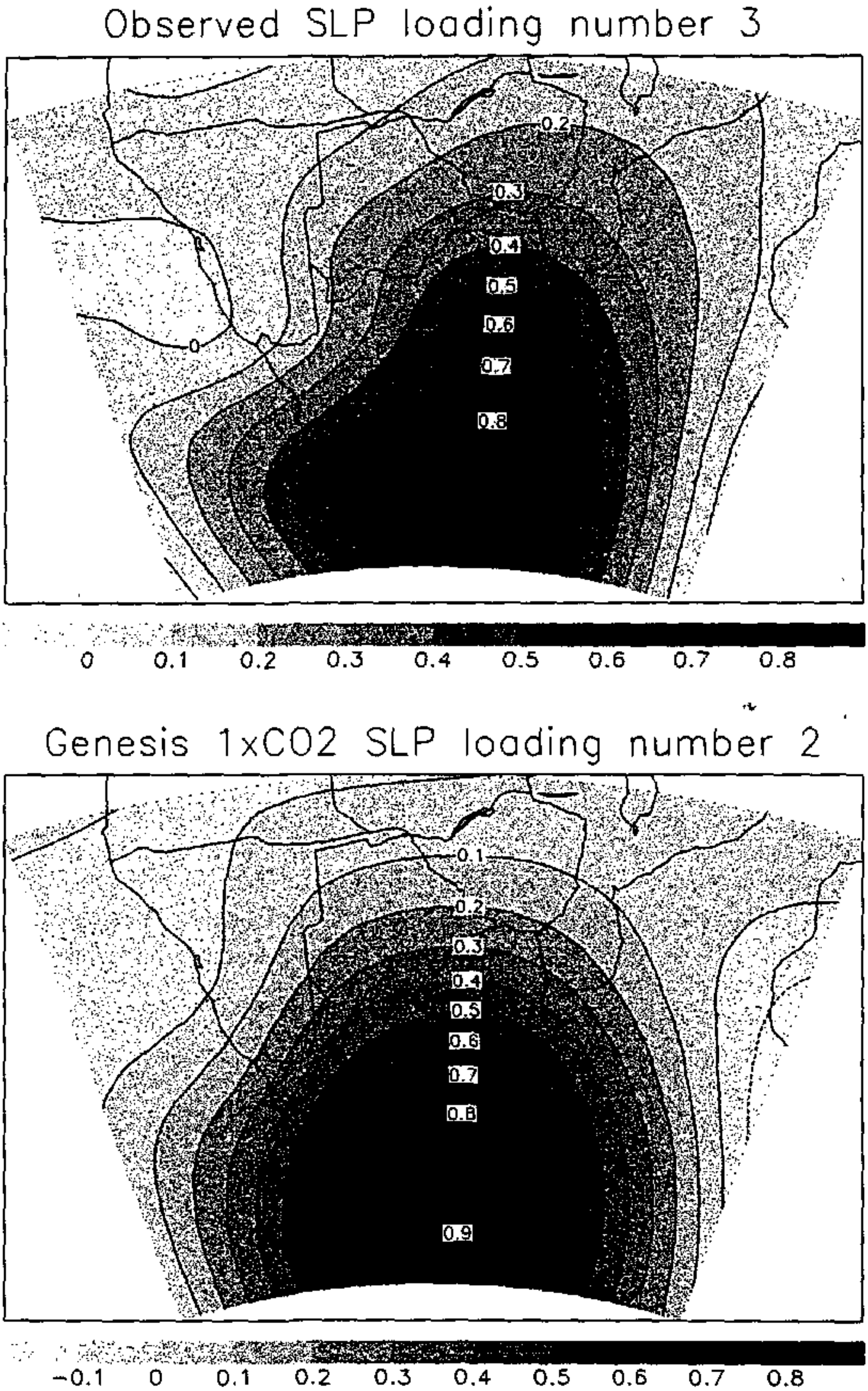

Figure 4: Example of the component loading match between GCM and Observed data loadings.

account for antecedent conditions; most notably the air mass trajectory and air mass history.

In training the ANN 75% of the data observations were used keeping 25% of the days as independent test data with which to evaluate the downscaling function. In training, the ANN function adapts to the relationship between the input circulation data and the target data, iteratively evaluating itself against the test data. The root mean square error (RMSE) of the ANN predicted value with respect to the target value is used as a measure of how well the ANN is predicting the target parameter. When the RMSE of the output with respect to the test data set fails to decrease any further then the ANN is deemed trained with the best possible relationship between the input and target data. Selected grid cells over the spatial domain were then evaluated by comparing the time series of the ANN predicted values with the actual observed data for the same days. Figures 5-7 shows some time series samples of the observed versus predicted values for different locations and different parameters.

As can be seen in Figures 5-7 the ANN captures the phase of the events along with seasonality very well. The amplitude of the events is also captured although high amplitude values are underestimated. This latter aspect is considered to be due partly to the lack of water vapour in the input to the ANN, however, as the relationship is explicitly a *generalisation* of the forcing, one would not expect the peaks to be perfectly matched. Thus the ANN captures the generalised relationship very well while missing the variability of the peak events due to lack of information as to what separates a peak event from the general event.

If one averages the data up to monthly and seasonal scales (e.g.: mid-summer and mid-winter patterns in Figures 8-10) and considers the spatial attributes, it is apparent that the ANN performs well while the under-prediction of peak events is manifest in a small under-prediction of the high precipitation regimes. Notwithstanding this, it is apparent that the ANN has reproduced the seasonality, spatial distribution, and near-absolute values with a high degree of accuracy. Based on these results, and that the GCM simulation of circulation is considered valid, it is accepted that downscaling for these parameters is practical over southern Africa and applicable for use with the GCMs.

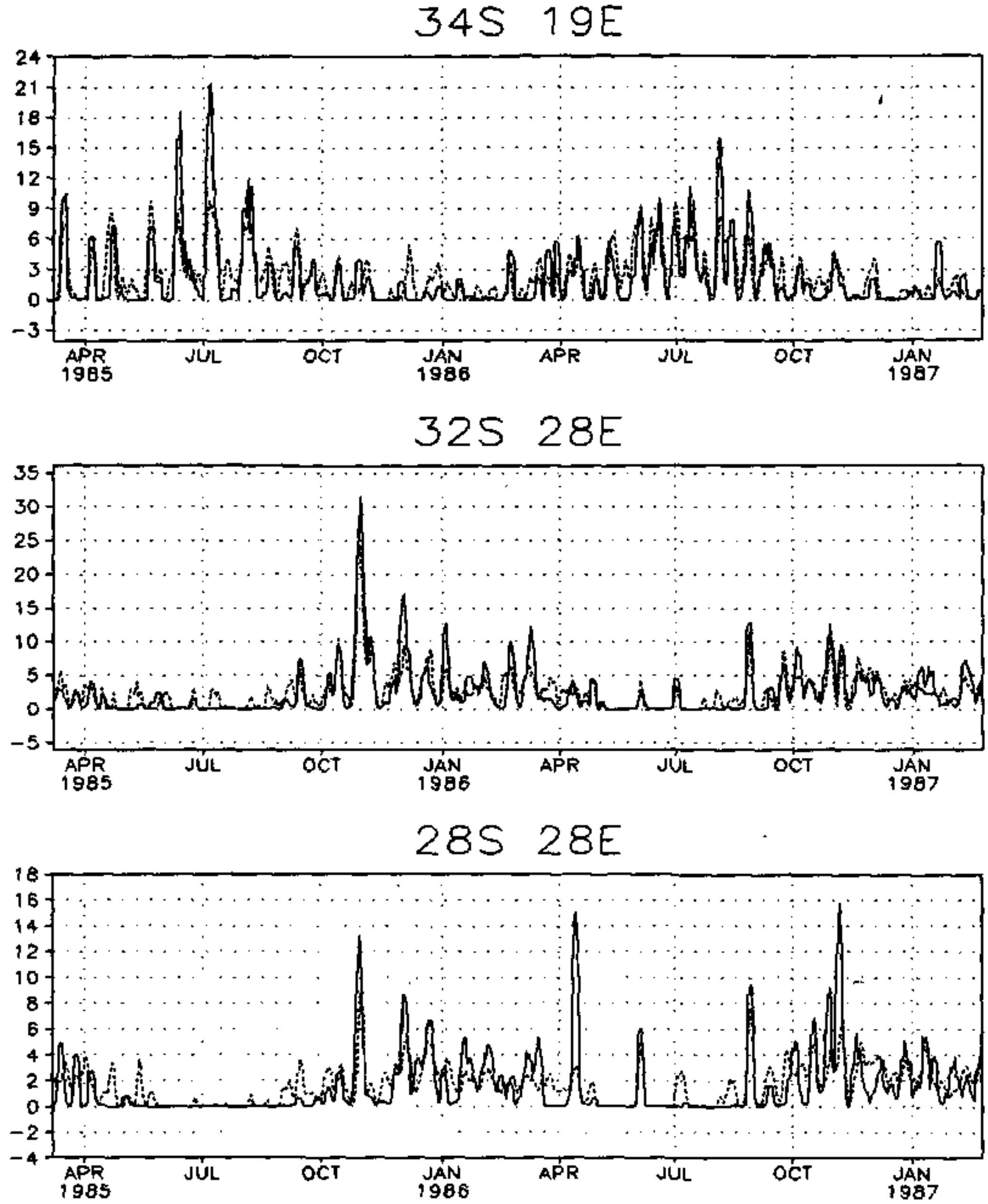

Figure 5: Example of observed (solid) and circulation-predicted (dashed) precipitation (mm/day).

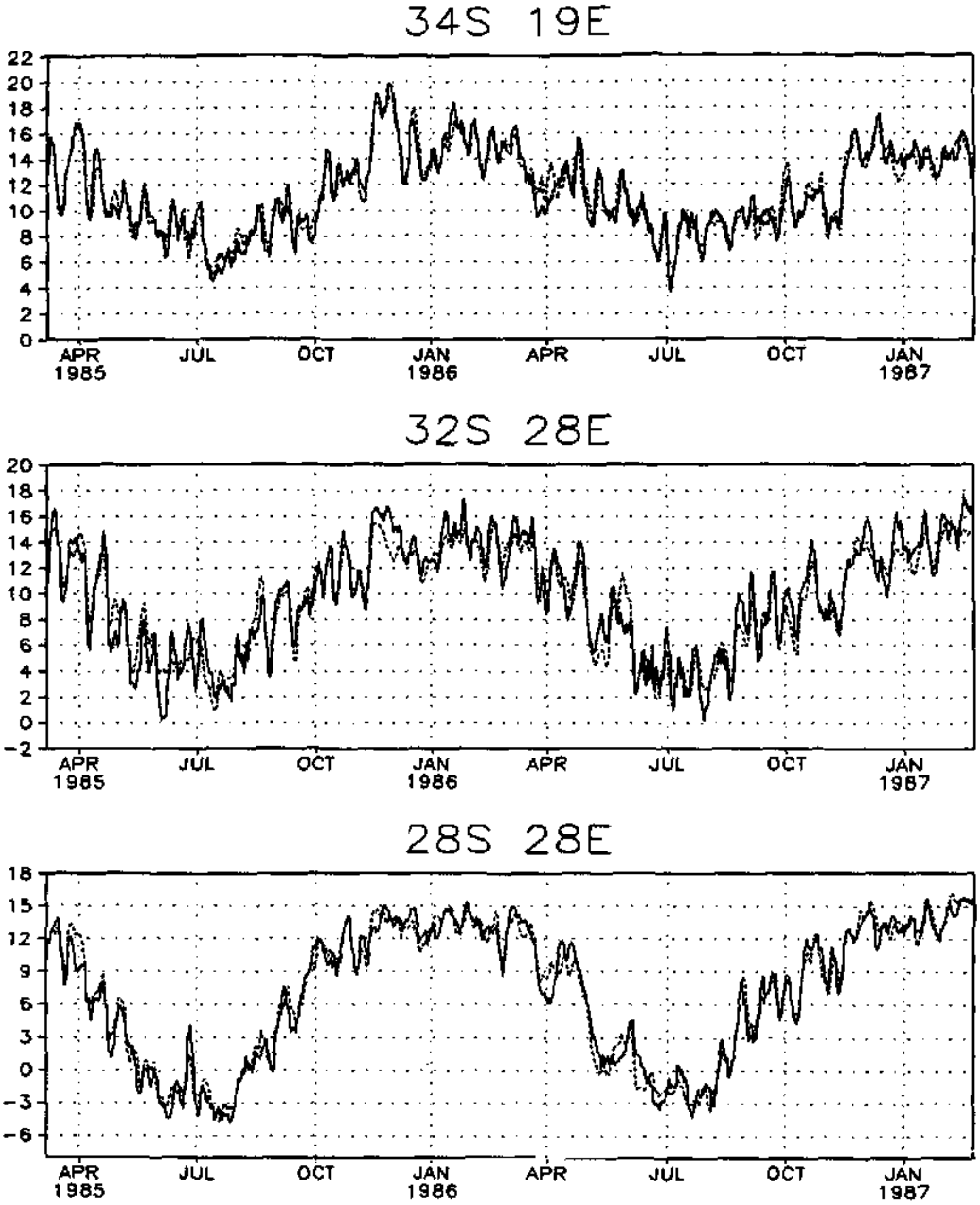

Figure 6: Example of observed (solid) and circulation-predicted (dashed) minimum temperatures (°C).

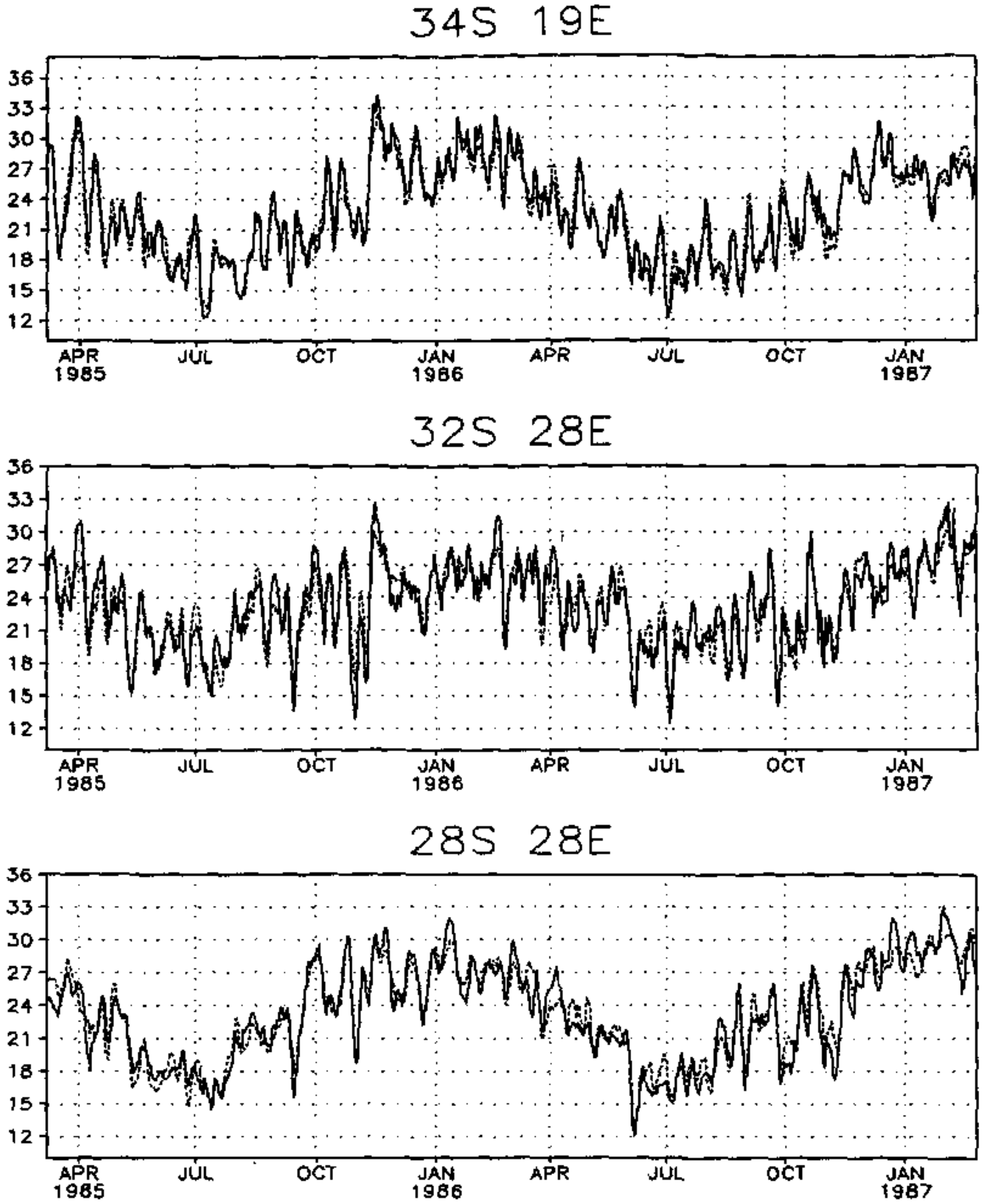

Figure 7: Example of observed (solid) and circulation-predicted (dashed) maximum temperatures (°C).

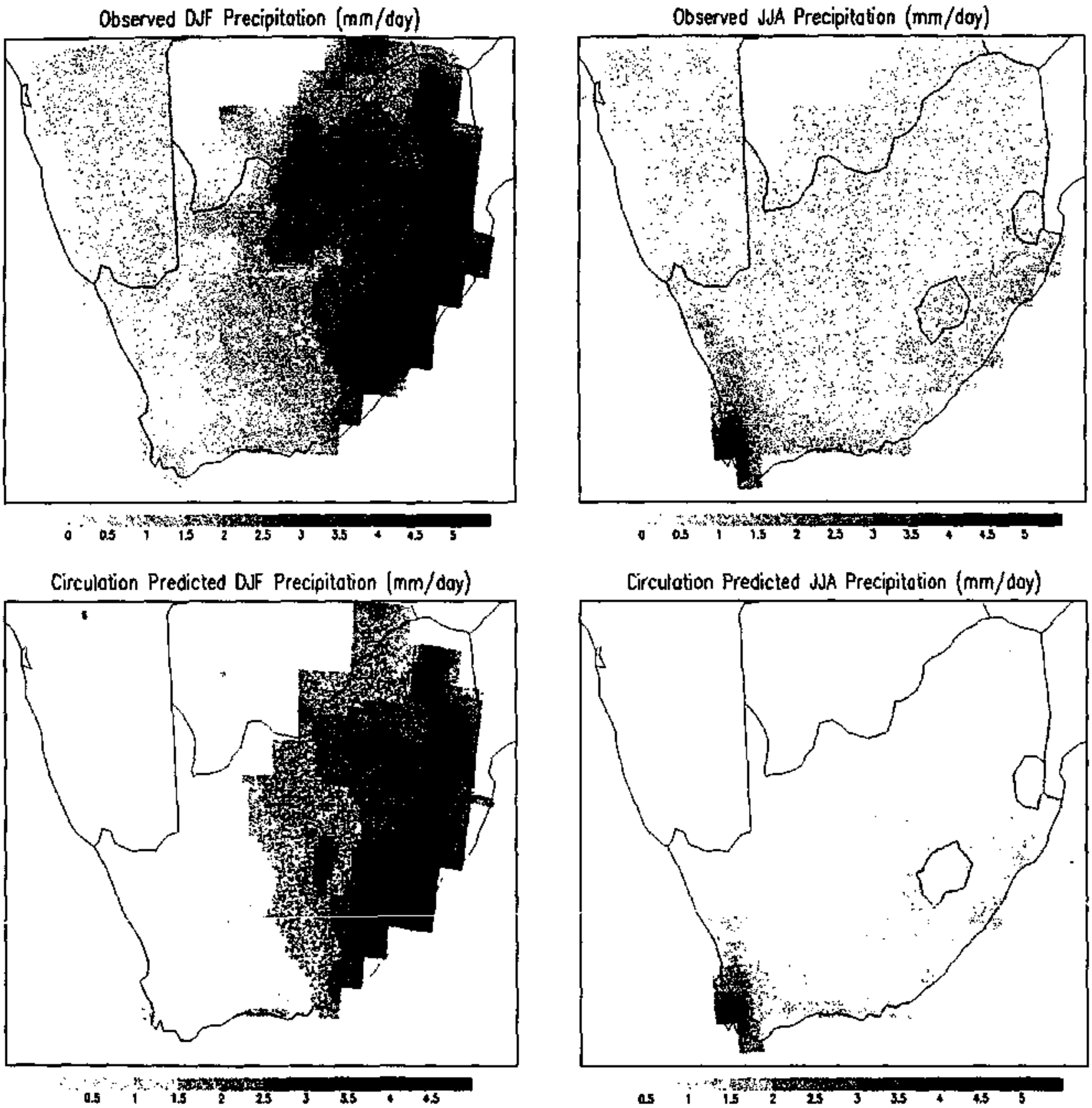

Figure 8: Seasonal averages of observed and circulation-predicted precipitation

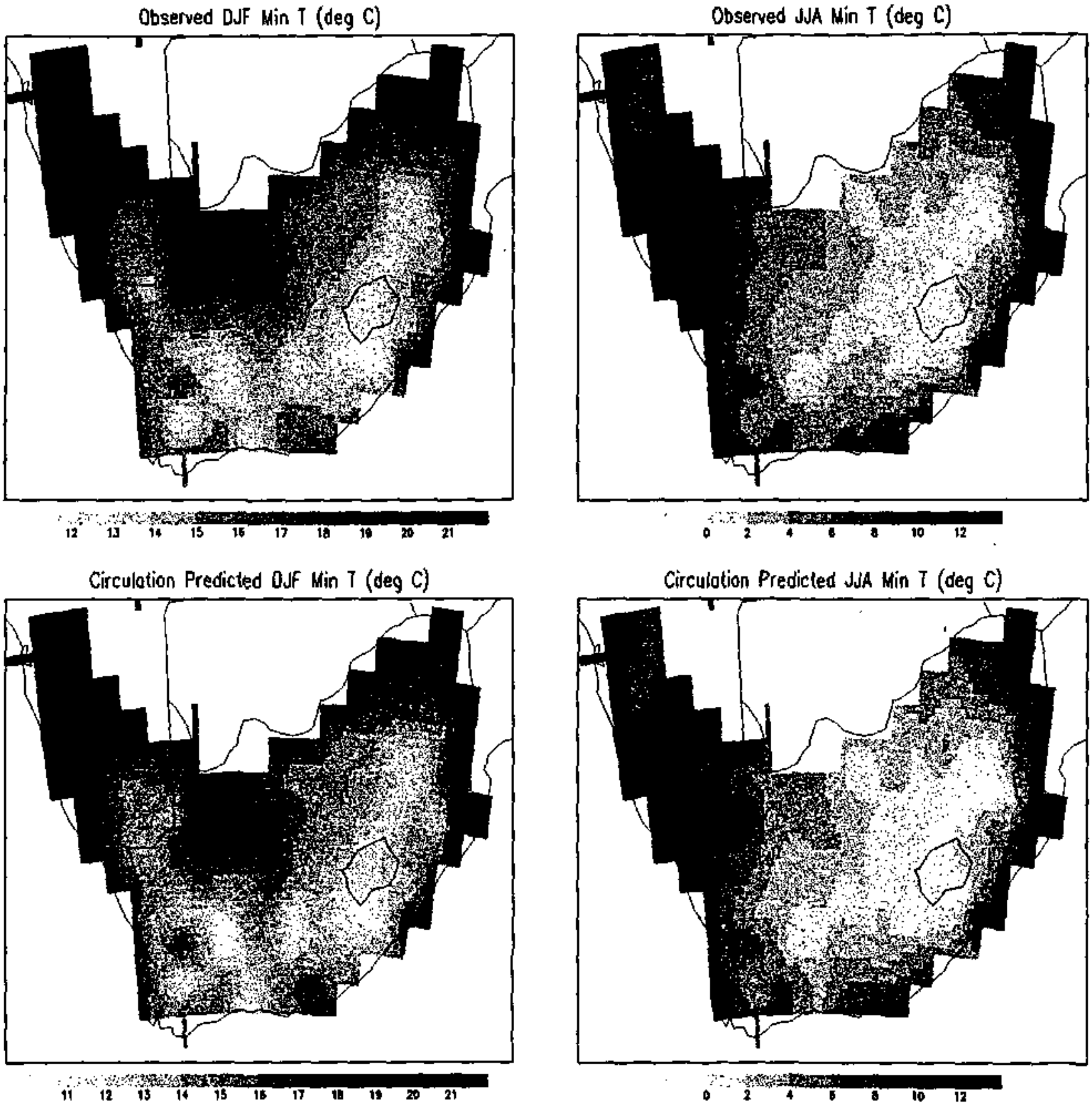

Figure 9: Seasonal averages of observed and circulation-predicted minimum temperature.

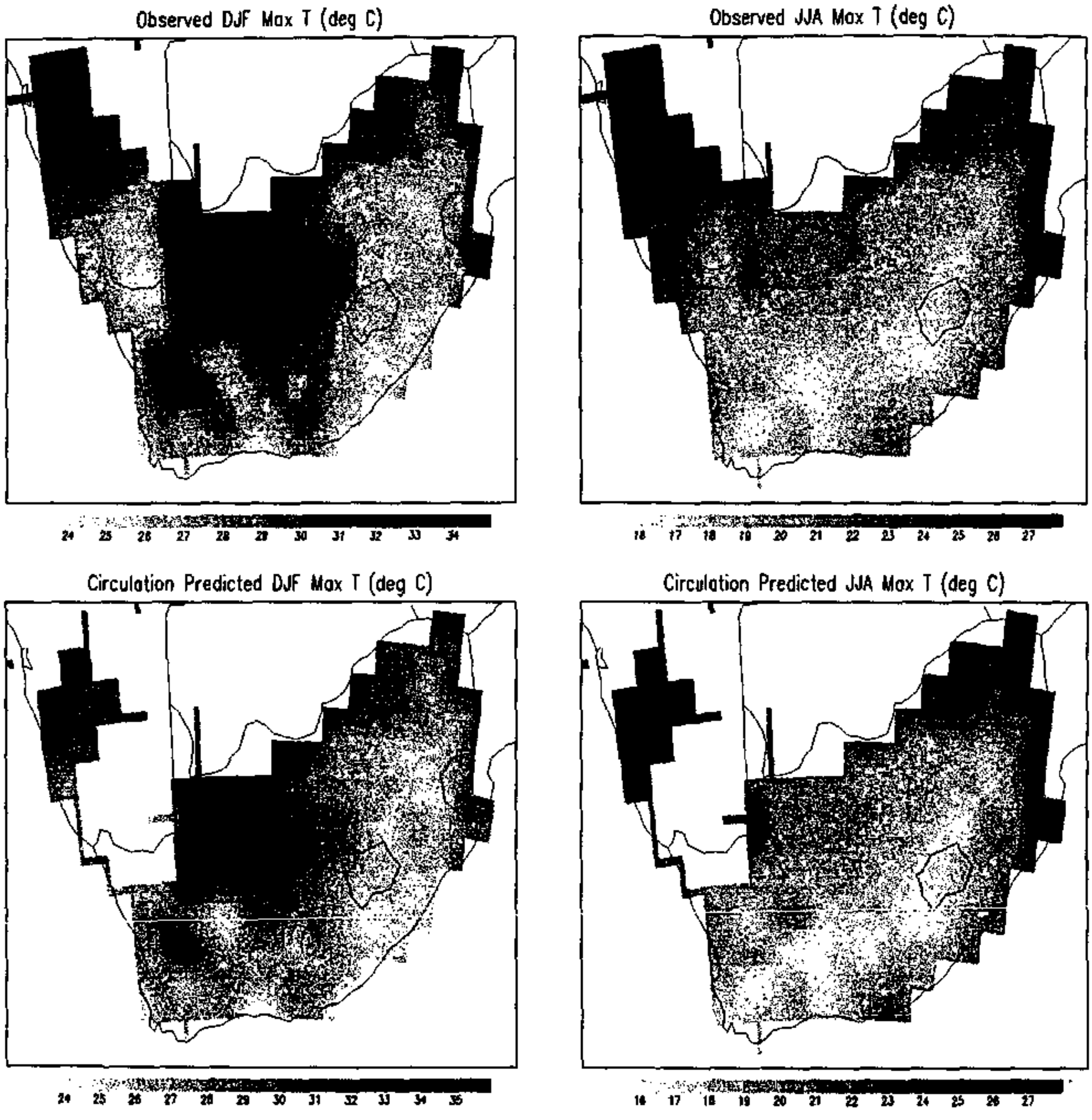

Figure 10: Seasonal averages of observed and circulation-predicted maximum temperature.

3.4 *Application to GCM 1xCO_2 data*

The use of downscaling with the GCM control simulation data derives local climates as a function of the GCM representativeness of the temporal and spatial behaviour of circulation. As such, the application provides an inherent validation of the GCM ability to represent 'reality', and may be used both as a means of indicating the reliability of the results, and as a diagnostic procedure that is of significant use for the modelling community. In this case the comparison 'observed' data set, or 'truth', is the precipitation derived from the observed circulation. This is preferable to using the 'raw' observed precipitation as the circulation derived values represent the regional precipitation as a generalised function of circulation, and thus inherently excludes unique local scale influences and provides a more valid basis of comparison.

The GCM control simulation data is processed in the same manner as the observed data, with the modification that the PCA score time series are generated using the estimated score coefficients from the PCA of the observed data. This ensures that the spatial contribution of variance to each component score is consistent with the data used to train the ANN, and is justified on the basis that the PCA loading patterns of the GCM and observed data match (section 3.2). The score time series of circulation are then applied as inputs to the trained ANN and the regional climate parameters generated consistent with the representation of the GCM synoptic forcing. The results show a startling improvement over the GCM internal values in terms of both the spatial distribution and the seasonality. Whereas before the GCM internal precipitation showed no seasonal cycle over the summer rainfall region, with significant errors in mean rainfall amount, the downscaled values show a very close resemblance to those derived from the observed circulation. Figures 11-13 show the summer and winter values of precipitation and temperature derived from observed and GCM circulation.

3.5 *Application to GCM 2xCO_2 data*

The application to the 2xCO_2 simulation data is the same as with the 1xCO_2 data. However, in this case an additional post-processing step of generating 2xCO_2-1xCO_2 anomaly maps is performed to elucidate the nature of the climate change impacts as a direct result of CO_2 changes. While the GCM demonstrates reasonable ability to simulate present-day climate, there are nonetheless errors in the GCM which affect mean values.

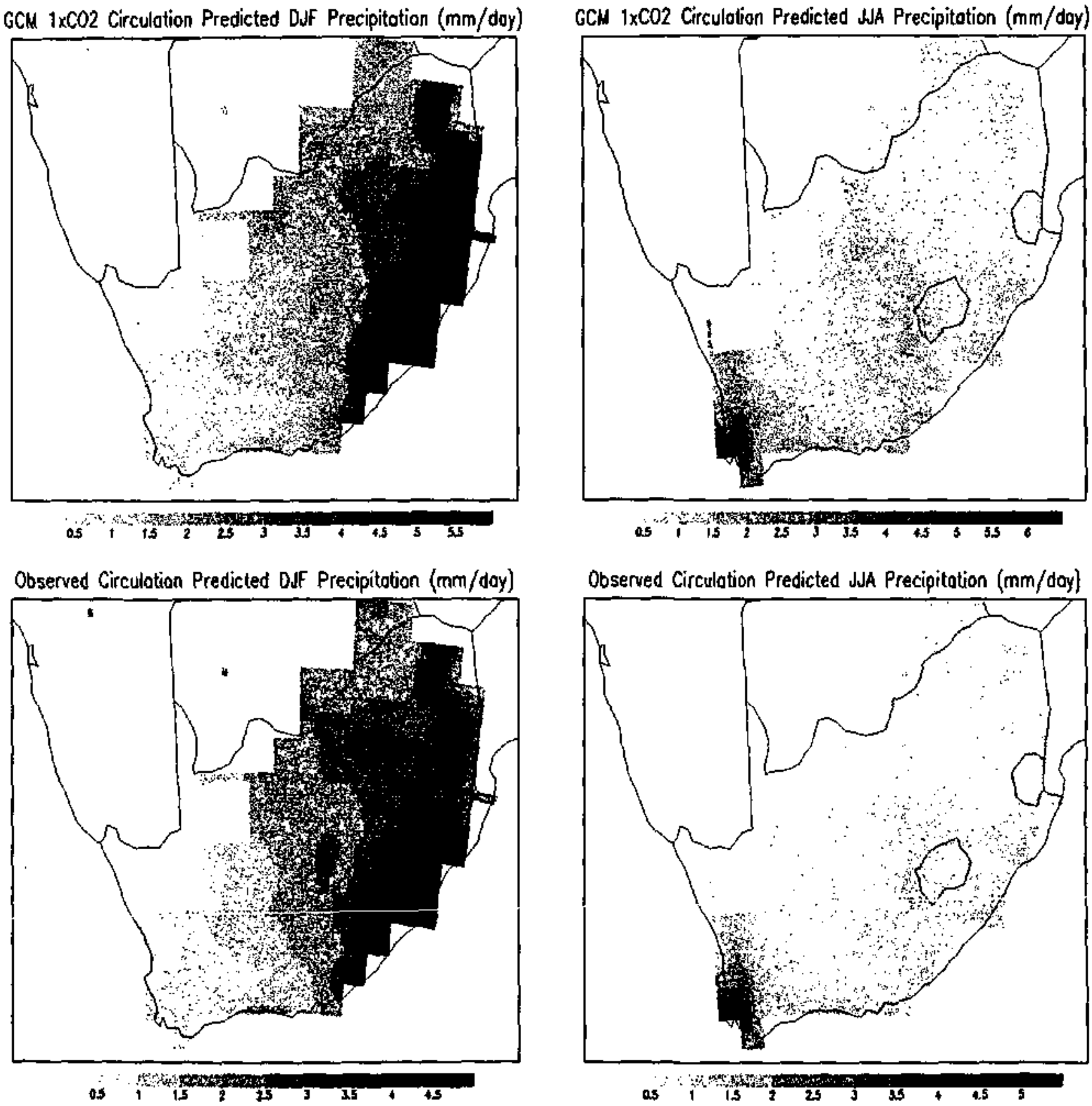

Figure 11: Observed circulation-predicted vs GCM circulation predicted precipitation.

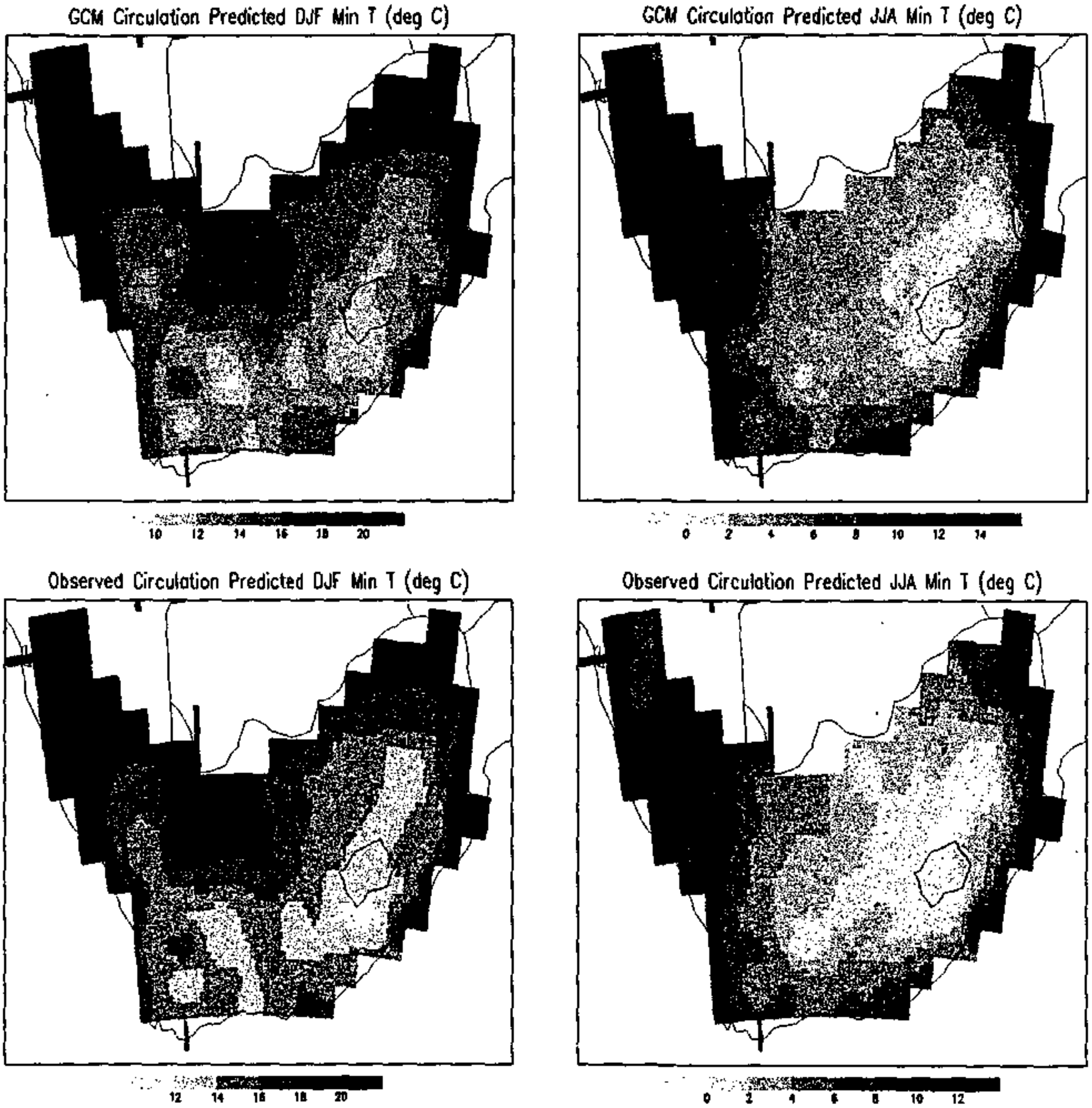

Figure 12: Observed circulation-predicted vs GCM circulation predicted minimum temperature.

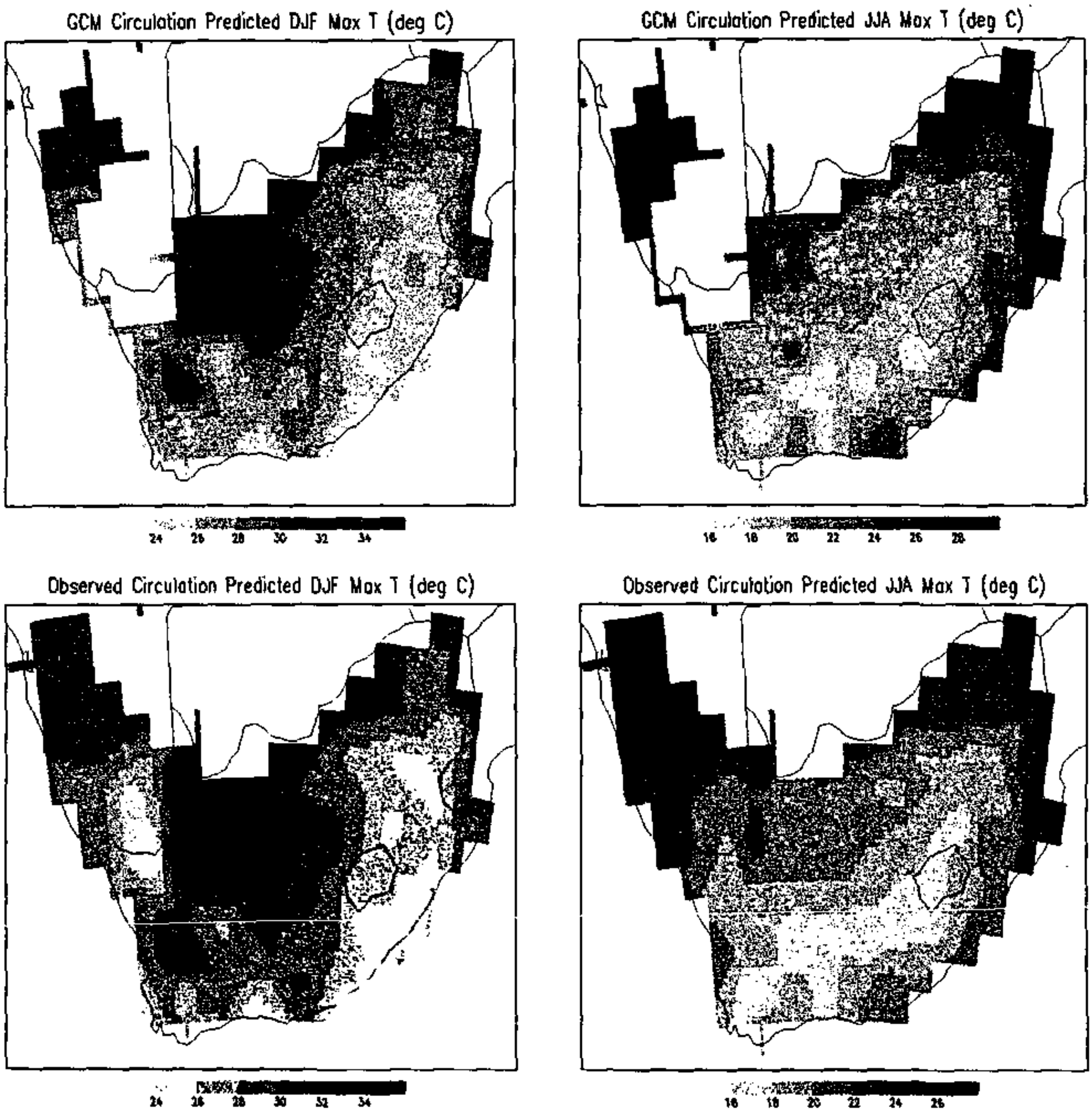

Figure 13: Observed circulation-predicted vs GCM circulation predicted maximum temperature.

Thus by using the anomaly map one represents the CO_2 forced synoptic response of the local climate within the context of the GCM climate simulation. Figures 14-16 show the anomaly maps for the seasonal mean precipitation and temperature changes.

4. Discussion

4.1 Downscaling functions

The downscaling functions show evident ability to represent the regional climate as a function of synoptic scale forcing. The ANN represents the *generalised* response of regional precipitation and temperature to the current and antecedent circulation to the extent that circulation alone is the dominant control. In this regard the aspects that the ANNs miss are the peak events. Thus while the circulation indicates a given state of synoptic forcing, and hence the regional value for the downscaled variable, there is still variability unaccounted for due to other features, for example, atmospheric water vapour.

Nonetheless, the ANN captures the temporal behaviour of the system well. When considering the seasonal mean values and their close resemblance to the observed data it is apparent that the extreme positive and negative values missed by the ANN lead to little bias in the mean, and this at present tends to generate only a nominal under-prediction. As such the downscaling can be accepted as a valid representation of regional climate response to the larger atmospheric system.

The significance of these results are important in the context of future climate change work for South Africa. While this study is preliminary and uses only one GCM simulation set, the validity of the approach is demonstrated and holds promise for further extension to newer simulation data sets from different models. In doing so an evaluation of model consensus between model simulations may be derived which is important for building a basis of credibility for a particular scenario's implications.

4.2 GCM control simulation

Application of the downscaling demonstrates that the Genesis GCM v1.02 simulates synoptic scale forcing with reliable accuracy as shown by the comparison of the downscaled values with respect to those derived from the observed circulation patterns.

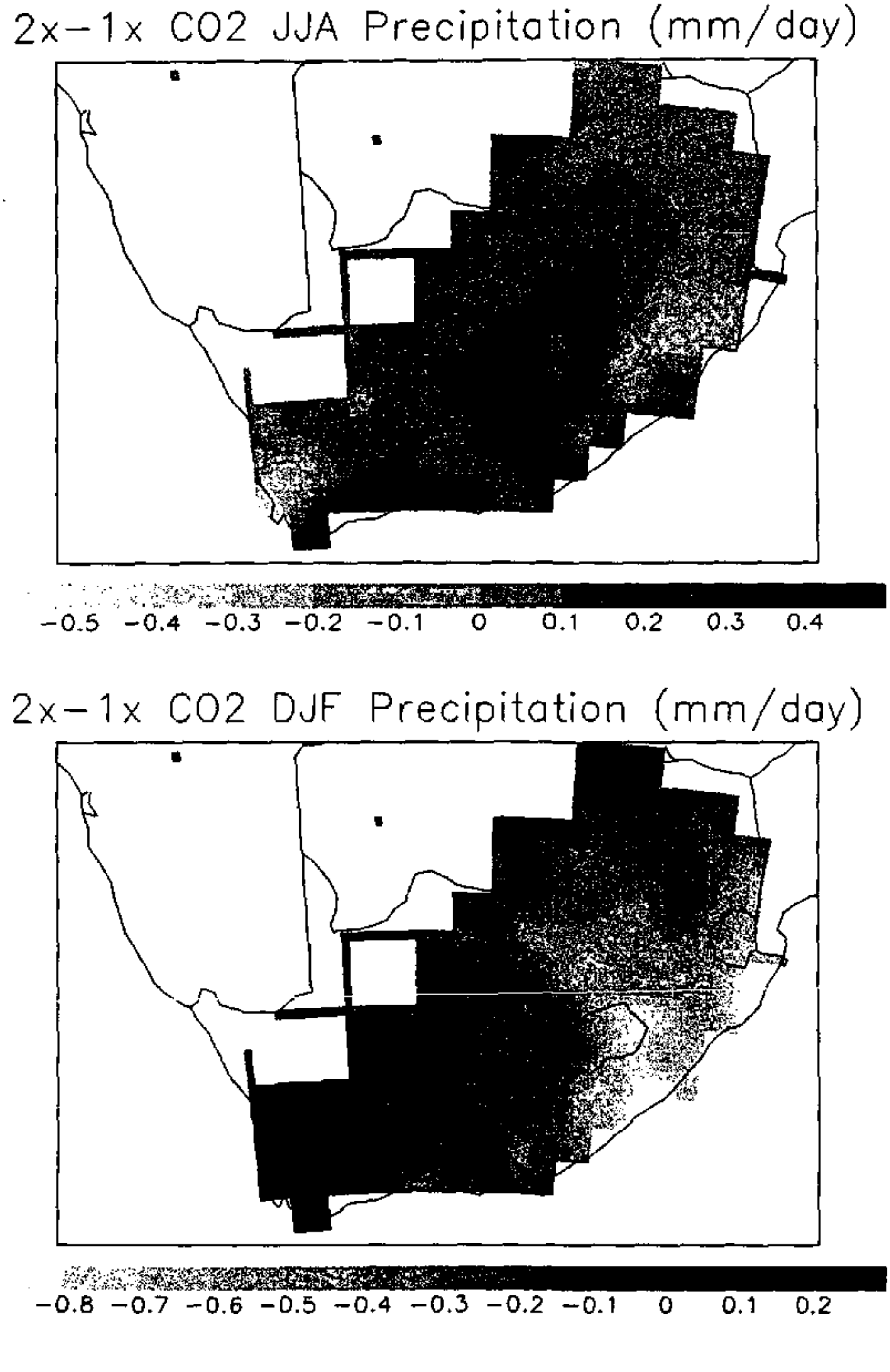

Figure 14: $2xCO_2$ - $1xCO_2$ circulation-predicted anomaly precipitation.

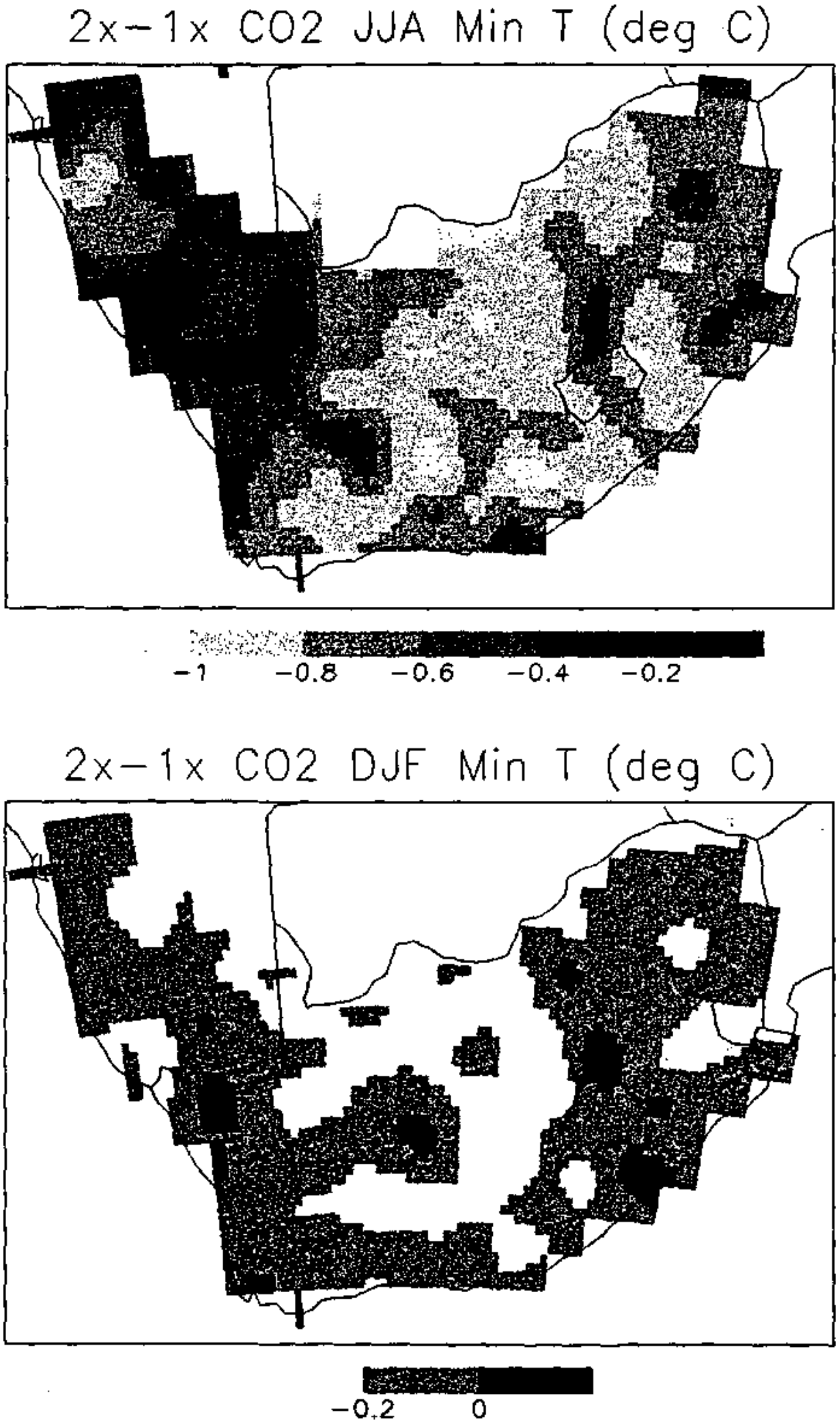

Figure 15: $2xCO_2$ - $1xCO_2$ circulation-predicted anomaly minimum temperature.

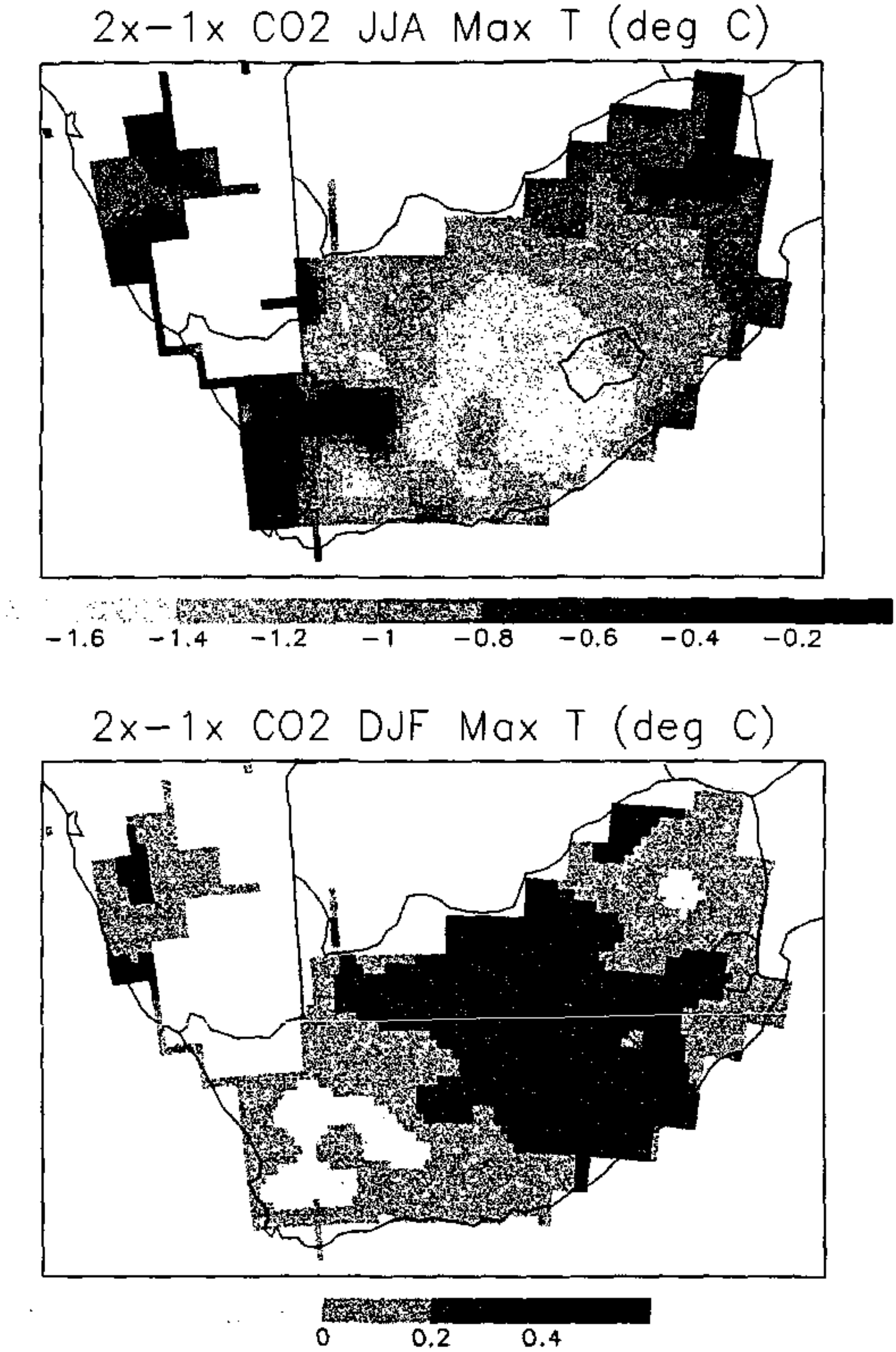

Figure 16: $2xCO_2$ - $1xCO_2$ circulation-predicted anomaly maximum temperature.

The model does show a tendency to over-predict precipitation in some coastal regions, but nonetheless the over-prediction is not excessive and the spatial distribution and seasonal behaviour are both well captured. These results give confidence that the same model can provide reasonable $2xCO_2$ simulations within the constraints of the simplified ocean used in the model.

Initial results from the new Genesis GCM v2.0 also indicate are that the circulation is much improved, and in particular demonstrates better seasonal response. Thus future work with the Genesis 2.0 model are likely to improve on results obtained here.

4.3 GCM $2xCO_2$ regional climate scenarios.

4.3.1: Caveats and concepts

The climate change results presented are preliminary and primarily intended as a basis for the evaluation of the downscaling methodology applied to South Africa. Nonetheless, the premise in developing climate change scenarios is that the GCM represents a response *change* in the circulation with respect to the control simulation. Thus while the control simulation has some minor deviation from observed data (as discussed in the previous section), the fact that the control simulation reasonably represents reality lends credence to using the change in the Genesis $2xCO_2$ simulation as potential changes for the future.

Although such climate change scenarios thus have some credibility and provides a basis for policy development and adaptation, this particular GCM simulation is now somewhat dated in terms of it's internal representation of the ocean which is relatively simplified in the model. In light of the relationship of South Africa's climatic variability to the oceanic variability, it must be expected that the model-derived scenarios are unlikely to capture a significant portion of the inter-annual variation. In particular with regard to the El Nino - Southern Oscillation (ENSO) it is important to recognise that the models do not simulate ENSO. Furthermore, we still do not fully understand ENSO dynamics and possibilities exist that ENSO may move to a preferred stable state in future climates.

A further key caveat is to recognise that the coupled atmosphere-cryosphere-biosphere-hydrosphere system is extremely complex and past records do suggest possibilities of chaotic behaviour. The implications of this are that the system could theoretically move to a new climate state in a step jump. Such possibilities, however, are by nature of the system unpredictable, which leaves one with the only option of using the most sophisticated tools possible to achieve the most likely scenarios -- in this case, using GCM simulations that demonstrate reasonable ability at representing reality.

It should also be noted that any downscaling procedure does not take into account direct radiative forcing changes. This will not significantly affect the precipitation downscaling; however, it could have bearing on the temperature scenarios. As such, the temperature scenarios should be seen as changes due to changes in synoptic forcing, and that a large-scale mean bias field may need to be included. Parallel research on this topic is underway which seeks to determine the large scale bias field to add to the synoptically forced temperature changes.

In the light of these caveats the downscaling performed over South Africa does represent credible scenarios that are consistent with current broad expectations of the nature of climate change. In particular, these expectations are that a greater degree of continentality will be manifest, an expansion of convective regimes, and with respect to temperature, the greatest changes will occur occurring in minimum temperatures. These aspects are all evidenced in the downscaled regional predictions.

4.3.2: Seasonal changes.

The following discussion focuses on the climate change anomaly maps presented in figures 14-16, and reference to these will assist in understanding the implications described in the section below. The discussion is a preliminary description of the changes and represents an evaluation of the procedure rather than a definitive prediction of future climate.

<u>Summer: December, January, February:</u>

For precipitation the summer season shows the most significant climate change impacts of any of the seasons, and presents the most severe potential consequences. For most of the summer convective region, and hence the summer agricultural area, there is a consistent 10-20% decrease in average daily rainfall across the rainfall gradient from the east coast into the interior. The primary significance in this is that a reduction of, say, 15% in rainfall, generally translates into a far greater percentage reduction in surface runoff. In contrast the west coast shows an increase of the order of 20%.

For temperatures, there is essentially no change in the summer minimum temperatures anywhere in the country, while for maximum temperatures there is a broad regional change showing a moderate increase of up to 0.5°C for the central and northern interior and the south-east coast. While 0.5° does not sound significant, it must be born in mind that, firstly, this does not take into account possible direct radiative changes. Secondly, this is the seasonal *mean* change, thus the change is still of consequence and individual events can show marked differences. Overall, the indications are for drier summer rainfall conditions with marginally increased variability.

Winter: June, July, August:

Precipitation in winter shows a nominal 5-10% decrease over the west coast with a further drying in the already dry north-east region. Much of the rest of the country shows a nominal increase. Temperatures show a strong cooling of 0.5 - 1.2°C for the central regions with an overall decrease everywhere. It is tempting to interpret this in terms of greater continentality and a strengthened continental high pressure system; however, further investigation of this is first needed.

5. Conclusions and Recommendations.

Today's society is faced by the dilemma of whether to invest in the significant cost of developing mitigation and adaptation strategies for climate change impacts, or to risk the equally significant consequences of present inactivity should a major human-induced climate change occur in the relatively near future. As the political process must respond to a complex set of often conflicting pressures and objectives, the formulation of global change policy requires thorough understanding of the likely changes that may occur in the global environmental system.

The necessary research base needs to come from the broad field of physical sciences, all of which, to some degree, require climate change scenario information as a fundamental point of departure. For example, water resources are a critical component of the economic structure in many regions of the world; however, developing the infrastructure necessary to adapt to changing water resources often requires long planning lead times. Although, in some cases, it may be possible to identify projects that would be beneficial regardless of whether any climate change actually occurs, for many other projects it may be difficult to justify the expense involved without the availability of reliable climate change information at the required spatial and temporal scales. The only viable tools for such research are GCMs, yet these have intrinsic limitations for which downscaling, either through dynamical process based nested models or as empirical transfer functions, can prove a valuable solution. While the nested modelling avenue promises to be the optimum long-term solution, empirical relationships provide an immediate computationally efficient approach to tackling some of these research questions. The methodology developed demonstrates an effective means with which downscaling can translate from the GCM's skill level down to the spatial scales necessary for assessing societal impacts.

Overall it is demonstrated that empirical downscaling can achieve regional climate scenario results at a significantly higher spatial and temporal resolution than otherwise obtainable from current research methodologies. The procedure captures both intra- and inter-seasonal variability well, and in application to the GCM data, derives regional climates consistent with the GCM synoptic forcing and with far greater detail and accuracy than is possible from using the GCM fields directly.

However, the actual scenarios presented here are preliminary, and the following points are noted:

- The regional scenarios are based on one GCMs simulation of global climate change and is linked to this models ability to simulate the complexity of the global system. Nonetheless, the model does represent a probable dynamic response of the atmosphere to doubling of CO_2, and processes such as ENSO should be considered to be a modulation on top of the basic global change response signal obtained by the GCM.
- A consensus has yet to be shown between these results and those derived by the same process from other models. Nonetheless, independent research in other areas suggests that the Genesis model used here is consistent with other GCMs, and presents a reasonably robust first basis for the scenarios.
- Due to the difficulties with the availability of water vapour data for the project the scenarios do not explicitly take into account possible changes in atmospheric water

vapour content. However, it is considered that the primary circulation response will dominate the climate change impacts prior to the effect of any water vapour changes, and hence the scenarios present a valid outlook for initial understanding of regional changes.

- The scenarios are based on relatively short (in climatological terms) simulation periods which limit the potential impacts of inter-annual variability. Longer term simulations will better reflect the climatological means under a doubled CO_2 atmosphere, and in response to this, further scenarios are currently being developed based on the established methodology.

As a consequence of the results obtained, a number of important research issues are suggested that should be noted for further work.

- Investigation of the physical dynamics underlying the regional changes represented by the GCM. The resultant local climate change, while of value to the impacts researcher, need to be further investigated in terms of the dynamical processes which give rise to the change. This is important for understanding why the change is occurring, and possible secondary effects that may arise as a result.
- New scenarios developed from other GCMs and the new Genesis v2.01 GCM. These are important in order to consider the degree of consensus that exists between different models, and thus provide a basis for confidence limits for a particular scenario.
- Incorporation of more sophisticated ocean representation in the GCM, and incorporation of atmospheric sulphates. Both these aspects stand to alter any scenario developed, and are of particular importance in arriving at a justifiable scenario on which to base long term policy development.
- Inclusion of water vapour in new studies at higher spatial and temporal resolutions. Water vapour changes are fundamental to precipitation potential and are also intrinsically tied to the oceanic system. Initial investigations suggest that water vapour will increase marginally, and thus should be included in future work.

The issues raised above, along with other related research questions, are being followed up in further work as part of the new WRC project starting in 1996, and in collaborative work with the Earth System Science Center at the Pennsylvania State University in the USA. In particular, the question of further refining the downscaling is being addressed, along with disaggregation of the 1° spatial resolution data down to station and catchment basin scales, and the generation of daily unsmoothed data more suitable for hydrological impact studies. This report is also available on the internet at the site: http://www.egs.uct.ac.za

7. References

Grotch, S. L., MacCracken, M. C., 1991, The Use of General Circulation Models to Predict Regional Climatic Change, *J. Climate*, 4:286-303.

Hewitson, B.C., 1995: Needs for downscaling: Applications and examples, *Proceedings of the 6th International Meeting on Statistical Climatology*, University of Galway, Ireland.

Hewitson, B.C., and Crane, R.G. (eds.), 1994: *Neural Nets: Applications in Geography*, Kluwer Academic Publishers, Dordrecht, The Netherlands.

Hewitson, B.C., and Crane, R.G., 1992: Regional climates in the GISS GCM: synoptic scale circulation. *J. Climate*, 5: 1002-1011.

von Storch H., Zorita E. and Cubasch U., 1992, Downscaling of Global Climate Change Estimates to Regional Scales: An Application to Iberian Rainfall in Wintertime, *J. of Climate*, 1993, 6: 1161-1171.