

An assessment of the skill of downscaled GCM outputs in simulating historical patterns of rainfall variability in South Africa

D. A. Hughes, S. Mantel and T. Mohobane

ABSTRACT

Uncertainties associated with General Circulation Models (GCMs) and the downscaling methods used for regional or local scale hydrological modelling can result in substantial differences in estimates of future water resources availability. This paper assesses the skill of nine statistically downscaled GCMs in reproducing historical climate for 15 catchments in five regions of South Africa. The identification of skilled GCMs may reduce the uncertainty in future predictions and the focus is on rainfall skill as the GCMs show very similar patterns of change in temperature. The skill tests were designed to assess whether the GCMs are able to realistically reproduce precipitation distribution statistics and patterns of seasonality, persistence and extremes. Some models are consistently less skilful for the regions assessed, while some are generally more skilful with some regionally specific exceptions. There are differences in the GCMs skill across the different regions and in the skill ranking between coastal areas and inland regions. However, only a limited reduction in uncertainty is achieved when using only the downscaled GCM outputs identified as being skilled in a hydrological model for one of the regions. Further modelling studies are required to determine the general applicability of this observation.

Key words | climate change, downscaling, hydrological modelling, skill

D. A. Hughes (corresponding author)
S. Mantel
T. Mohobane
Institute for Water Research,
Rhodes University,
Grahamstown,
South Africa
E-mail: d.hughes@ru.ac.za

INTRODUCTION

While it is frequently suggested that climate change will have large impacts on hydrology and the water resources of the world (Bates *et al.* 2008), there remains a substantial degree of uncertainty associated with the General Circulation Models (GCMs) themselves, as well as the methods used to downscale (Hewitson & Crane 1996; Segui *et al.* 2010) the outputs for use with hydrological models (Prudhomme *et al.* 2003; Chen *et al.* 2011). Part of the uncertainty lies in the different structures and initial conditions assumed for the different GCMs and various inter-comparison studies have demonstrated that there can be substantial differences in the outputs representing both the present climate (Reichler & Kim 2008) as well as the future (Hughes *et al.* 2011). Pirtle *et al.* (2010) discuss the issue of GCM independence and refer to the use of

model ensembles that average the outputs across a group of models (Hagedorn *et al.* 2005). The assumption behind the use of multi-model ensembles appears to be based on the fact that there are inherent errors in all the models (largely due to the need to simplify highly complex atmospheric physics) and that using multi-model output ensembles provides better coverage of the possible climate outcomes. However, using an ensemble average is essentially equivalent to ignoring the uncertainties and is contrary to the trend in the hydrological sciences to try and account for all uncertainties (Pappenberger & Beven 2006). It is therefore considered necessary to use all members of the ensemble to represent the uncertainties in future climate change projections (Hughes *et al.* 2011). Whether any specific ensemble of models is able to reflect the real

uncertainty and whether the models within the ensemble can be considered independent are further issues that need to be considered (Masson & Knutti 2011).

GCM outputs are inadequate for use with hydrological models applied at the catchment scale as the spatial resolution is too coarse. It is therefore considered necessary to use some form of downscaling (Hewitson & Crane 1996; Bouwer *et al.* 2004; Fowler *et al.* 2007; Segui *et al.* 2010; Frost *et al.* 2011). However, while downscaling can address the spatial scale issues it is likely to introduce different uncertainties into the predictions (Buytaert *et al.* 2010). There are two general categories of downscaling in general use: dynamical and statistical. Dynamical downscaling can be achieved by using a Regional Climate Model (RCM) nested within the GCM simulations (Leung *et al.* 2004), while statistical downscaling is an empirical approach that establishes relationships between the GCM outputs and local scale variables (Hewitson & Crane 1996). In the latter case, the additional uncertainties will be related to the uncertainties in the observations of the local scale variables, as well as the assumptions used in developing appropriate relationships.

Given the inherent uncertainty in the outputs from downscaled GCMs it seems justified to assess their performance in simulating past or present day climates (Wilby 2010) in an attempt to identify which (if any) of a range of models are more skilled than others in a specific region. If this is possible it may be possible to limit the number of GCMs used in the estimation of future possible scenarios which may reduce the range of uncertainty. However, if skill tests are to be undertaken it is essential that they examine aspects of the downscaled GCM outputs that are consistent with the structure and limitations of the GCMs (Huard 2011). It is generally not correct to base the skill assessment on temporal correlations between climate model outputs and observed precipitation and temperature, despite that fact that there are examples of such approaches in the literature (Anagnostopoulos *et al.* 2010). This type of approach is considered invalid because of the way in which GCMs are externally forced, the internal (largely chaotic) dynamics of the climate system and the fact that they are not expected to 'predict climate in a deterministic sense' (Huard 2011). It is therefore necessary to select skill measures that are not only appropriate for the purposes of the hydrological simulations,

but that also are consistent with the expectations of the GCM outputs.

From a hydrological and water resources assessment perspective, it is essential that the GCMs are able to realistically reproduce patterns of precipitation seasonality, persistence and extremes. If they are not able to adequately reproduce these characteristics under present day forcing conditions, it is unlikely that we can have much confidence in their ability to predict changes in these characteristics under different forcing conditions in the future. This is not the same as saying that skilful simulations of the past will lead to skilful predictions of the future. It is, however, suggesting that a lack of skill in simulations of the past is very likely to translate into a lack of skill in future predictions. Past skill therefore becomes a necessary, but insufficient, basis for confidence in future predictions (Knutti 2008).

Various approaches to assessing skill have been reported in the literature (Wilby & Harris 2006; Schmidli *et al.* 2007; Chiew *et al.* 2009) and they have included assessments of the ability of downscaled GCM data to reproduce frequency characteristics of precipitation, mean seasonal patterns, as well as durations above or below certain thresholds. Downscaled GCM data for future hydrological scenarios typically require some form of bias correction if they are to be used in comparison with hydrological simulations of the past using historically observed data (Chen *et al.* 2011). Part of the skill analysis could therefore focus on the assumptions implicit in any bias correction method and whether or not these are equally valid across the range of downscaled data sets being used.

This study forms part of a larger project to investigate the uncertainties in hydrological model outputs based on simulations using historical data (largely parameter and observed forcing data uncertainty) combined with the additional uncertainties associated with both climate change and future water resources development impacts. The original objective of the study was to identify whether the results of skill tests applied to nine downscaled GCMs could be used to either reduce the number of applicable models or to rank their levels of uncertainty. A favourable outcome from a hydrological modelling point of view would be that the uncertainty in the predictions of the future could be reduced by focusing on the more skilful GCMs.

THE GCMs AND DOWNSCALING METHOD

The downscaled data used in this study are those produced by the Climate Systems Analysis Group of the University of Cape Town which are available for the nine GCMs listed in Table 1 for the SRES (Special Report on Emission Scenarios) A2 emission scenario (IPCC 2001). The data are based on statistical downscaling using the methods discussed in Hewitson & Crane (1996, 2006). An earlier version of this data set was used by Lumsden *et al.* (2009) to assess patterns of future change in rainfall statistics across the whole of South Africa. The authors accept that the single emission scenario and projection data used from nine GCMs are not representative of the full uncertainty in future climates. However, these are data that are based on the same methods of downscaling and that have been made readily available to hydrologists in South Africa for translating climate change signals into impacts on water resources availability. The differences in both projections of change (direction and extent) and skill should therefore be treated as a sample of possible variations and uncertainties. It is not suggested that this is a representative sample, it is simply a sample created from consistent downscaling methods that has been made available for hydrological analyses in South Africa. The authors acknowledge that the issues noted by Masson & Knutti (2011) about similarities in climate model genealogy can be important in assessing GCM differences. However, as hydrologists, we are constrained by the data that are recommended and made available by climatologists (Hewitson & Crane 1996) and these issues are not addressed further in this paper.

The specific downscaled data products used in this study are the so-called quinary catchment-based data set. The 1946 quaternary catchments are the divisions used within South Africa for water management purposes (Middleton & Bailey 2008), while there are typically three or more quinary catchments within each quaternary. Quaternary catchments are typically between 100 and 1,000 km² in the wetter parts of the country, but can be much larger in very dry areas. The final downscaled data products are daily time series of rainfall depths, maximum and minimum temperatures at the quinary catchment scale for a baseline period (1961–2000), as well as near (2046–2065) and far future (2081–2100) scenarios. In this study, the quinary scale daily rainfall data have been aggregated to quaternary scale monthly rainfall data using an inverse distance weighting procedure (Wilk *et al.* 2006) based on points representing the location of the quinary catchments. This approach does not affect the statistical properties of the original downscaled rainfall data as there are very few differences in these properties between closely adjacent quinary catchments.

The study has focused on five groups of quaternary catchments: the H10 group in the headwaters of the Breede River in the Western Cape; the K90 group representing the Kromme River basin on the Southern Cape coast; the R20 group in the Buffalo River on the Eastern Cape coast; the C/D group representing the inland (Free State Province) catchments of the Caledon and Modder rivers; and the X31 group in the headwaters of the Sabie River located on the eastern escarpment in Mpumalanga Province

Table 1 | Global climate models used in the study

GCM abbreviation	Source of GCM
CCCMA	Canadian Centre for Climate Modelling and Analysis
CNRM	France Centre National de Recherches Meteorologiques
CSIRO	Australian Commonwealth Scientific and Industrial Research Organisation (CSIRO) Atmospheric Research
GFDL	USA National Oceanic and Atmospheric Administration (NOAA) Geophysical Fluid Dynamics Lab
GISS	USA Goddard Institute for Space Studies
IPSL	France Institut Pierre Simon Laplace
MIUB	Germany Meteorological Institute of the University of Bonn
MPI	Max-Planck Institute for Meteorology
MRI	Japan Meteorological Research Institute

(Figure 1). The sites were chosen partly to cover a variety of climate conditions and rainfall types, but were also selected because of their relevance to parallel studies within the Institute for Water Research of the effects of climate change on water resources.

All of the GCMs for all regions show very similar patterns of change in temperature, but not for rainfall. Figure 2 summarises the variation in projected rainfall changes for the nine GCMs using the per cent difference in the mean annual rainfall between the baseline and near future periods for all of the quaternary catchments in each region (12 for X, 11 for H, 7 for K, 42 for C/D and 11 for R). It is clear that there are substantial variations in both the degree and direction of change across different models and different regions, as well as (in some cases) within regions. Previous studies (Lumsden *et al.* 2009) have suggested that the Western Cape region (H) shows the most consistent projections of reduced rainfall across most GCMs. While the data in Figure 2 indicate that more GCMs suggest lower rainfalls in the H region than other regions, three of the GCMs predict increased rainfall, while one suggests small changes. Using these results as inputs into hydrological models will clearly introduce a large amount of uncertainty in future estimates of water resources availability, particularly for those areas where there is substantial disagreement between GCMs on both direction and degree of rainfall change (X, K and C/D).

Figure 3 illustrates that there are very big differences in the seasonal distributions between the nine baseline rainfall

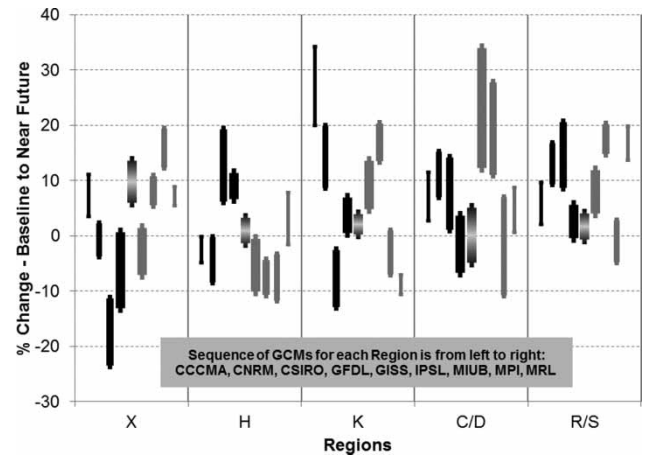


Figure 2 | Summary of per cent change from baseline (1960–2000) to near future (2046–2065) for all nine downscaled GCM data. The bars represent the variability across all of the quaternary catchments within each region.

time series and the WR2005 (Middleton & Bailey 2008) rainfall data that are routinely used in water resources analyses in South Africa. The WR2005 rainfall data were compiled at the scale of the quaternary catchments from all available observed rainfall station data (Middleton & Bailey 2008). While the WR2005 data are less than perfect, largely because of relatively low densities of measurement stations in some areas, they are the best representation of historical rainfall patterns that are available for all parts of the country. While it would have been useful to include uncertainty bounds around the historical WR2005 rainfall data, this has not been possible as the authors do not have access to the original station data nor any detailed

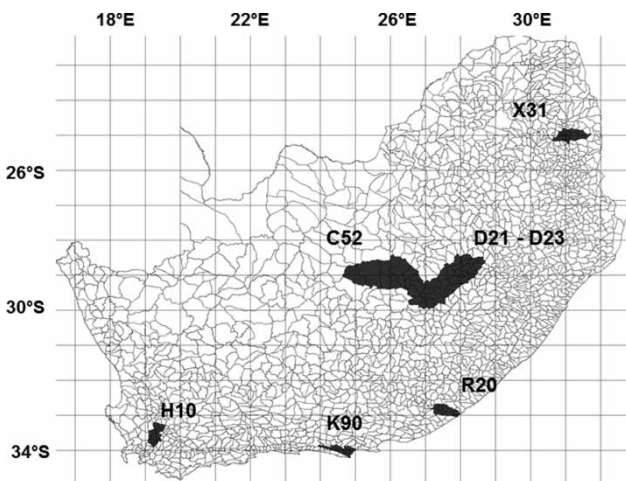


Figure 1 | Map of South Africa showing the six study areas.

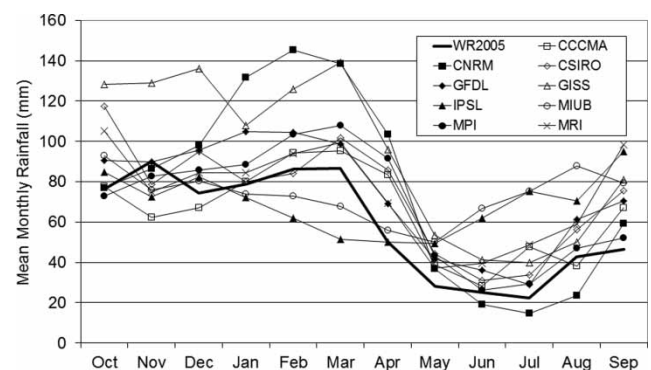


Figure 3 | Mean monthly rainfall depths (mm) for the historical data (WR2005) and the nine downscaled GCMs for the baseline period (1960–2000) for quaternary catchment R20B.

documentation of the methods used. The other regions show similar degrees of bias between the downscaled baseline data and the WR2005 historical data and it is therefore clear that bias correction is necessary before the future projections can be used to assess the potential impacts on water resources. It is also clear that the bias corrections will have to be different for individual months.

RAINFALL BIAS CORRECTION

The ultimate objective of the main study is to use the downscaled rainfall data as inputs to hydrological and water resources yield models and specifically the monthly time-step Pitman model (Hughes *et al.* 2006). Figure 3 illustrates that, compared to the WR2005 data, the baseline rainfall time series exhibit bias in monthly means and, although not shown on the diagram, the same is true for the monthly standard deviations. The method used to remove this bias from the future (near and far) rainfall estimates is to express the future monthly rainfalls as standard deviates of the baseline monthly distributions (using a square root transformation) and to scale the standard deviates with the monthly distribution statistics of the historical rainfall data (Equation (1)). Similar approaches have been applied in other studies (Haerter *et al.* 2011).

$$RFC_{ijk} = \left(RH\mu_j + RH\sigma_j \times \left(RF_{ijk} - RB\mu_{jk} \right) / RB\sigma_{jk} \right)^2 \quad (1)$$

where: RFC_{ijk} = Future rainfall after correction for month i and calendar month j in the time series of GCM k . $RH\mu_j$ = Mean of the square root transformed historical (WR2005) rainfalls for calendar month j . $RH\sigma_j$ = Standard deviation of the square root transformed historical (WR2005) rainfalls for calendar month j . RF_{ijk} = Square root transformed future rainfall for month i and calendar month j in the time series of GCM k . $RB\mu_{jk}$ = Mean of the square root transformed baseline rainfalls for GCM k and calendar month j . $RB\sigma_{jk}$ = Standard deviation of the square root transformed baseline rainfalls for GCM k and calendar month j .

The objective of the correction equation is to remove the bias in the monthly means and variations between the historical and GCM baseline estimates, while preserving the differences between the GCM baseline and future scenarios.

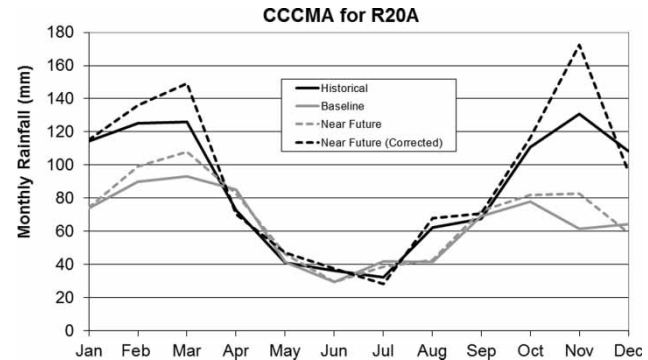


Figure 4 | Example of the rainfall bias correction for downscaled CCCMA data and quaternary catchment R20A.

Several other correction approaches (such as using the cumulative frequency distributions of rainfall) did not preserve the seasonality and structure of the downscaled future rainfalls. The square root transformation was used to account for the generally positive skewness evident in monthly rainfall data. Initial applications of the bias correction method used a natural logarithmic transformation, but this was found to introduce additional bias in some site/GCM combinations (related to the existence of low or even negative skewness values for some months in some GCM data), while the square root transformation was found to be more generally applicable. An example of the results of applying this bias correction is provided in Figure 4 for the downscaled CCCMA rainfall data for quaternary catchment R20A. The patterns of differences between the baseline and historical data (and therefore the degree of bias correction required) are highly variable across the regions and GCMs.

CONVERTING TEMPERATURE DATA TO MODEL EVAPOTRANSPIRATION INPUTS

Evaporation inputs to the Pitman model typically consist of an annual potential evaporation (PE) depth (mm) and a fixed seasonal distribution. While it is possible to use a time series of PE inputs (Hughes *et al.* 2006) there is often not enough historical information to quantify these variations accurately and Sawunyama (2008) noted that their inclusion did not make much difference to the overall model results. In this study, it was decided to use the maximum and minimum temperature data for the baseline and future climate model

scenarios to determine the temperature component of the Hargreaves (Hargreaves & Samani 1985) equation (Equation (2)). The per cent increases in these values, from baseline to future, were then used to scale the seasonal distributions of PE when running the model for future scenarios.

$$HC_k = (TMax_k + TMin_k)/2 \times \sqrt{(TMax_k - TMin_k)} \quad (2)$$

where: HC_k = Temperature component of the Hargreaves equation for GCM k , calculated for baseline and future conditions. $TMax_k$ = Daily maximum temperature for GCM k . $TMin_k$ = Daily minimum temperature for GCM k .

The daily values are used to compute mean monthly values (MHC_{jk} , where j is the month) for all calendar months and the seasonal scaling factor computed from the ratio of the HC_k values for the individual GCM future scenarios to their baseline scenarios.

SELECTED MEASURES OF SKILL

Part of the rainfall transformation process relies, to a certain extent, on the WR2005 and baseline data for the different GCMs both having similar skewness values for the distribution of rainfall depths within each calendar month and that the square root transformation is appropriate (Equation (1)). The first skill measure was therefore based on the absolute value of the relative difference in skewness between the GCM baseline and WR2005 data (Equation (3)):

$$\gamma_skill_{jk} = ABS\left(\left(RB_{\gamma_{jk}} - RH_{\gamma_j}\right)/RH_{\gamma_j}\right) \quad (3)$$

where: γ_skill_{jk} = Skewness skill score for month j , GCM k . $RB_{\gamma_{jk}}$ = Skewness of baseline monthly rainfalls for calendar month j and GCM k . RH_{γ_j} = Skewness of the WR2005 monthly rainfalls for calendar month j .

The seasonality of the rainfall regime is clearly of great importance in hydrological modelling and during the initial phases of the study it had already been observed that some of the downscaled rainfall data did not appear to reproduce historical seasonality patterns very well. A skill measure was therefore adopted that would measure the relative differences between the GCM seasonal rainfall distributions and

the WR2005 data, but with the overall depth bias removed (Equation (4)):

$$Seas_skill_{jk} = ABS\left(\left(RB_{jk} \times (RB_MAP_k/RH_MAP) - RH_j\right)/RH_j\right) \quad (4)$$

where: $Seas_skill_{jk}$ = Season skill score for month j , GCM k . RB_{jk} = Baseline mean monthly rainfall for month j and GCM k . RB_MAP_k = Baseline mean annual rainfall for GCM k . RH_j = WR2005 mean monthly rainfall for month j . RH_MAP = WR2005 mean annual rainfall.

The equivalent of Equation (4) has also been used to calculate a seasonality skill measure for the temperature data. Equation (2) was applied to some historical temperature data (Schulze & Maharaj 2004) as well as the baseline temperature data for the nine downscaled GCMs. The mean monthly values for the historical and GCM data were then used in Equation (4), replacing mean monthly rainfall with the MHC_{jk} values (see above) and the mean annual rainfall values with the temperature equivalents.

The rainfall seasonality skill measure does not adequately account for the variation in rainfalls across different years for the same calendar month and therefore an additional skill measure has been added to account for this and based on the coefficient of variation, or the ratio of standard deviation to the mean (Equation (5)):

$$CV_skill_{jk} = ABS\left(\left(RB_CV_{jk} - RH_CV_j\right)/RH_CV_j\right) \quad (5)$$

where: CV_skill_{jk} = Coefficient of variation skill score for month j , GCM k . RB_CV_{jk} = Coefficient of variation of the baseline monthly rainfalls for calendar month j and GCM k . RH_CV_j = Coefficient of variation of the WR2005 monthly rainfalls for calendar month j .

The final skill measure has been based on calculations of serial auto-correlation within the individual time series using lags of 1, 2, 11, 12 and 13 months. These lags were chosen on the basis of the serial correlation patterns observed in the WR2005 data that demonstrated weak intra-season persistence (for example, 0.25 and 0.10 for lags 1 and 2 for sub-basin R20A), as well as weak persistence across two adjacent seasons (0.19, 0.22, 0.16 for lags 11, 12 and 13, respectively for sub-basin R20A). The baseline rainfall time series for all nine GCMs exhibited similar patterns,

but with quite different correlation values. The serial correlation skill measure was therefore simply the sum (for all five lags) of the absolute differences in serial correlation (Equation (6)):

$$SC_skill_k = \Sigma ABS(SC_{lk} - SC_l) \quad (6)$$

where: SC_skill_k = Sum of the skill values for all five lags. SC_{lk} = Serial correlation coefficient for lag l and baseline rainfall for GCM k . SC_l = Serial correlation coefficient for lag l , WR2005 monthly rainfalls.

During the initial phase of the skill assessment, the GCMs were compared using the annual skill values detailed above ($STAT_skill_k = \Sigma STAT_skill_{jk}$ for all j months, with $STAT$ used to represent the γ_skill , $Seas_skill$ and CV_skill measures). However, many of the annual values are influenced by extremely high (poor skill) values for the dry season months where small absolute differences in rainfall amounts between the WR2005 data and the GCMs result in high values of the skill indices. All of the monthly skill values for the first three measures were therefore weighted by the ratio of the mean monthly WR2005 rainfall to the mean annual WR2005 rainfall (Equation (7)) and then summed for revised annual values:

$$STAT_new_skill_{jk} = STAT_skill_{jk} \times RH_j / RH_MAP \quad (7)$$

The first approach to assessing the variations in skill across the various regions and GCMs was to use a rank

for each annual skill measure and then to sum and multiply the individual skill ranks to obtain two overall skill scores. These were then averaged to obtain a final rank. The process is illustrated in Table 2 for quaternary catchment R20B and these rankings have been used to order the climate models in the presentations of the more detailed results (Figure 5).

A second approach was used to take into account both the total annual skill value, as well as the maximum skill score (applicable to γ_skill , $Seas_skill$ and CV_skill measures only) for the 12 months. For each skill measure, threshold values were calculated such that approximately half of the skill values for all GCM/site combinations fall on either side of the threshold. This analysis was done for the total skill measures (all months for γ_skill , $Seas_skill$ and CV_skill measures and all lag values for the SC_skill) and for the maximum monthly skill measure (γ_skill , $Seas_skill$ and CV_skill). Those GCM/site combinations with values above the thresholds were identified as failing the specific skill test, while those below the thresholds passed the specific test. The use of the terms passing and failing are therefore relative and have no absolute meaning.

RESULTS

The first observation was that there is much less difference in the temperature skill measures across the different climate models for several of the regions where observed historical temperature data were available. It was therefore

Table 2 | Rankings of the GCMs for the different measures of rainfall skill

GCM	Measure of skill				Cumulative measures and ranking				
	Skew	Season	CV	SC	Total multiplied	Rank	Total summed	Rank	Mean rank
CCCMA	1	5	2	7	70	2	15	3	2.5
CNRM	4	7	4	9	1,008	8	24	7	7.5
CSIRO	9	4	5	4	720	6	22	5	5.5
GFDL	6	1	6	1	36	1	14	2	1.5
GISS	7	2	8	8	896	7	25	8	7.5
IPSL	3	9	7	3	567	5	22	5	5
MIUB	8	8	9	5	2,880	9	30	9	9
MPI	5	3	3	2	90	4	13	1	2.5
MRI	2	6	1	6	72	3	15	3	3

Rainfall seasonality

Skill Rank	W Cape Coast				S Cape Coast		E Cape Coast		Inland Caledon			Inland Eastern Escarpment			
	H10A	H10E	H10G	H10L	K90A	K90F	R20B	R20G	C52G	D21B	D23E	X31A	X31E	X31L	X31M
1	GISS	GISS	MPI	MRI	MPI	MPI	GFDL	MPI	CNRM	IPSL	GISS	CNRM	CNRM	GISS	CNRM
2	MRI	CCCMA	MRI	GISS	GISS	GFDL	MPI	MRI	GFDL	MPI	MPI	IPSL	MUB	CNRM	IPSL
3	CCCMA	GFDL	CSIRO	CCCMA	GFDL	IPSL	MRI	GFDL	IPSL	GFDL	CNRM	GISS	IPSL	IPSL	GISS
4	GFDL	IPSL	CCCMA	IPSL	CNRM	CNRM	CCCMA	GISS	GISS	MUB	IPSL	CSIRO	GISS	CSIRO	MUB
5	MPI	MRI	CNRM	MPI	CSIRO	CCCMA	CSIRO	CSIRO	MUB	CCCMA	GFDL	MUB	CSIRO	MUB	MPI
6	IPSL	CSIRO	GFDL	GFDL	MRI	GISS	IPSL	CCCMA	MPI	GISS	MUB	MPI	MPI	MPI	GFDL
7	CSIRO	MPI	IPSL	CNRM	CCCMA	MRI	GISS	IPSL	CSIRO	MRI	CSIRO	CCCMA	GFDL	CCCMA	CSIRO
8	CNRM	CNRM	MUB	MUB	IPSL	MUB	CNRM	MUB	CCCMA	CNRM	MRI	GFDL	MRI	GFDL	CCCMA
9	MUB	MUB	GISS	CSIRO	MUB	CSIRO	MUB	CNRM	MRI	CSIRO	CCCMA	MRI	CCCMA	MRI	MRI

Coefficient of variation

Skill Rank	W Cape Coast				S Cape Coast		E Cape Coast		Inland Caledon			Inland Eastern Escarpment			
	H10A	H10E	H10G	H10L	K90A	K90F	R20B	R20G	C52G	D21B	D23E	X31A	X31E	X31L	X31M
1	GISS	GISS	MPI	MRI	MPI	MPI	GFDL	MPI	CNRM	IPSL	GISS	CNRM	CNRM	GISS	CNRM
2	MRI	CCCMA	MRI	GISS	GISS	GFDL	MPI	MRI	GFDL	MPI	MPI	IPSL	MUB	CNRM	IPSL
3	CCCMA	GFDL	CSIRO	CCCMA	GFDL	IPSL	MRI	GFDL	IPSL	GFDL	CNRM	GISS	IPSL	IPSL	GISS
4	GFDL	IPSL	CCCMA	IPSL	CNRM	CNRM	CCCMA	GISS	GISS	MUB	IPSL	CSIRO	GISS	CSIRO	MUB
5	MPI	MRI	CNRM	MPI	CSIRO	CCCMA	CSIRO	CSIRO	MUB	CCCMA	GFDL	MUB	CSIRO	MUB	MPI
6	IPSL	CSIRO	GFDL	GFDL	MRI	GISS	IPSL	CCCMA	MPI	GISS	MUB	MPI	MPI	MPI	GFDL
7	CSIRO	MPI	IPSL	CNRM	CCCMA	MRI	GISS	IPSL	CSIRO	MRI	CSIRO	CCCMA	GFDL	CCCMA	CSIRO
8	CNRM	CNRM	MUB	MUB	IPSL	MUB	CNRM	MUB	CCCMA	CNRM	MRI	GFDL	MRI	GFDL	CCCMA
9	MUB	MUB	GISS	CSIRO	MUB	CSIRO	MUB	CNRM	MRI	CSIRO	CCCMA	MRI	CCCMA	MRI	MRI

Skewness

Skill Rank	W Cape Coast				S Cape Coast		E Cape Coast		Inland Caledon			Inland Eastern Escarpment			
	H10A	H10E	H10G	H10L	K90A	K90F	R20B	R20G	C52G	D21B	D23E	X31A	X31E	X31L	X31M
1	GISS	GISS	MPI	MRI	MPI	MPI	GFDL	MPI	CNRM	IPSL	GISS	CNRM	CNRM	GISS	CNRM
2	MRI	CCCMA	MRI	GISS	GISS	GFDL	MPI	MRI	GFDL	MPI	MPI	IPSL	MUB	CNRM	IPSL
3	CCCMA	GFDL	CSIRO	CCCMA	GFDL	IPSL	MRI	GFDL	IPSL	GFDL	CNRM	GISS	IPSL	IPSL	GISS
4	GFDL	IPSL	CCCMA	IPSL	CNRM	CNRM	CCCMA	GISS	GISS	MUB	IPSL	CSIRO	GISS	CSIRO	MUB
5	MPI	MRI	CNRM	MPI	CSIRO	CCCMA	CSIRO	CSIRO	MUB	CCCMA	GFDL	MUB	CSIRO	MUB	MPI
6	IPSL	CSIRO	GFDL	GFDL	MRI	GISS	IPSL	CCCMA	MPI	GISS	MUB	MPI	MPI	MPI	GFDL
7	CSIRO	MPI	IPSL	CNRM	CCCMA	MRI	GISS	IPSL	CSIRO	MRI	CSIRO	CCCMA	GFDL	CCCMA	CSIRO
8	CNRM	CNRM	MUB	MUB	IPSL	MUB	CNRM	MUB	CCCMA	CNRM	MRI	GFDL	MRI	GFDL	CCCMA
9	MUB	MUB	GISS	CSIRO	MUB	CSIRO	MUB	CNRM	MRI	CSIRO	CCCMA	MRI	CCCMA	MRI	MRI

Serial auto-correlation

Skill Rank	W Cape Coast				S Cape Coast		E Cape Coast		Inland Caledon			Inland Eastern Escarpment			
	H10A	H10E	H10G	H10L	K90A	K90F	R20B	R20G	C52G	D21B	D23E	X31A	X31E	X31L	X31M
1	GISS	GISS	MPI	MRI	MPI	MPI	GFDL	MPI	CNRM	IPSL	GISS	CNRM	CNRM	GISS	CNRM
2	MRI	CCCMA	MRI	GISS	GISS	GFDL	MPI	MRI	GFDL	MPI	MPI	IPSL	MUB	CNRM	IPSL
3	CCCMA	GFDL	CSIRO	CCCMA	GFDL	IPSL	MRI	GFDL	IPSL	GFDL	CNRM	GISS	IPSL	IPSL	GISS
4	GFDL	IPSL	CCCMA	IPSL	CNRM	CNRM	CCCMA	GISS	GISS	MUB	IPSL	CSIRO	GISS	CSIRO	MUB
5	MPI	MRI	CNRM	MPI	CSIRO	CCCMA	CSIRO	CSIRO	MUB	CCCMA	GFDL	MUB	CSIRO	MUB	MPI
6	IPSL	CSIRO	GFDL	GFDL	MRI	GISS	IPSL	CCCMA	MPI	GISS	MUB	MPI	MPI	MPI	GFDL
7	CSIRO	MPI	IPSL	CNRM	CCCMA	MRI	GISS	IPSL	CSIRO	MRI	CSIRO	CCCMA	GFDL	CCCMA	CSIRO
8	CNRM	CNRM	MUB	MUB	IPSL	MUB	CNRM	MUB	CCCMA	CNRM	MRI	GFDL	MRI	GFDL	CCCMA
9	MUB	MUB	GISS	CSIRO	MUB	CSIRO	MUB	CNRM	MRI	CSIRO	CCCMA	MRI	CCCMA	MRI	MRI

Figure 5 | Overall skill rank and higher (white cells) or lower (grey cells) skill for all climate model/catchment combinations for the four skill criteria.

concluded that the rainfall skills are likely to be far more important in distinguishing between the climate models than the temperature skill differences.

Figure 5 presents the detailed results for the four rainfall skill measures. In all of the diagrams the climate models are listed in order using the ranking approach referred to above (and illustrated in Table 2), while those that passed (white background) and those that failed (grey background) either

the total annual skill score test or the maximum monthly skill score test are also identified. Figure 6 presents the models ranked in the same way as Figure 5 but with the shading designed to illustrate how many of the skill scores were above either of the thresholds. Figure 7 presents histograms of the frequencies of between 0 and 4 skill scores falling into the low skill categories for all climate models across all the catchments (15 in total). All of these diagrams

Number of tests in lower skilled categories



Skill Rank	W Cape Coast				S Cape Coast		E Cape Coast		Inland Caledon			Inland Eastern Escarpment			
	H10A	H10E	H10G	H10L	K90A	K90F	R20B	R20G	C52G	D21B	D23E	X31A	X31E	X31L	X31M
1	GISS	GISS	MPI	MRI	MPI	MPI	GFDL	MPI	CNRM	IPSL	GISS	CNRM	CNRM	GISS	CNRM
2	MRI	CCCMA	MRI	GISS	GISS	GFDL	MPI	MRI	GFDL	MPI	MPI	IPSL	MIUB	CNRM	IPSL
3	CCCMA	GFDL	CSIRO	CCCMA	GFDL	IPSL	MRI	GFDL	IPSL	GFDL	CNRM	GISS	IPSL	IPSL	GISS
4	GFDL	IPSL	CCCMA	IPSL	CNRM	CNRM	CCCMA	GISS	GISS	MIUB	IPSL	CSIRO	GISS	CSIRO	MIUB
5	MPI	MRI	CNRM	MPI	CSIRO	CCCMA	CSIRO	CSIRO	MIUB	CCCMA	GFDL	MIUB	CSIRO	MIUB	MPI
6	IPSL	CSIRO	GFDL	GFDL	MRI	GISS	IPSL	CCCMA	MPI	GISS	MIUB	MPI	MPI	MPI	GFDL
7	CSIRO	MPI	IPSL	CNRM	CCCMA	MRI	GISS	IPSL	CSIRO	MRI	CSIRO	CCCMA	GFDL	CCCMA	CSIRO
8	CNRM	CNRM	MIUB	MIUB	IPSL	MIUB	CNRM	MIUB	CCCMA	CNRM	MRI	GFDL	MRI	GFDL	CCCMA
9	MIUB	MIUB	GISS	CSIRO	MIUB	CSIRO	MIUB	CNRM	MRI	CSIRO	CCCMA	MRI	CCCMA	MRI	MRI

Figure 6 | Number of skill criteria falling into lower skill categories for all climate model/catchment combinations.

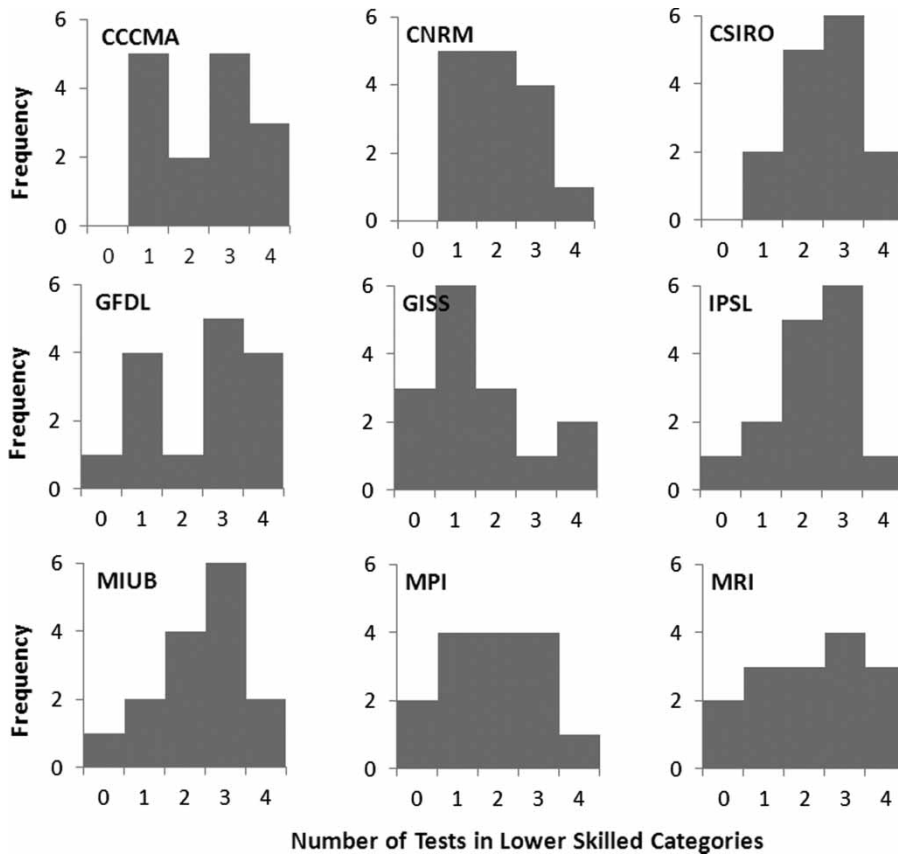


Figure 7 | Frequency of between none and four tests falling into the lower skill categories for the different climate models across all catchments.

should be viewed in the context of the variations in future projections illustrated in Figure 2.

With respect to variations in the highest ranked models (combining all four skill criteria) across the different regions, there are some consistent patterns, but there are also a number of inconsistencies even within individual regions:

- The most consistent result for the coastal areas is that MIUB, CNRM and CSIRO are generally ranked in the lowest half, while MPI, GFDL, MRI and GISS are the models with the most frequent high rankings.
- There are differences in ranking between the W. Cape and the other coastal areas (S. and E. Cape), which

might be related to the differences in weather patterns that result in winter season rainfall in the W. Cape, but spring and autumn rainfall seasons in the S. and E. Cape.

- For the inland catchments, CCCMA, CSIRO and MRI perform the worst, while CNRM, IPSL and GISS perform the best.
- For all regions, CSIRO has one of the biggest variations in predicted rainfall change (Figure 2) and this is also the model that performs the worst in terms of skill ranking. In contrast, GISS which ranks as one of the most skilful models has a relatively lower range of variability in predicted future change.

Perhaps the most immediately apparent result from Figures 5 and 6 is that there is a greater number of inland areas falling into the lower skill groups than for the coastal catchments, further emphasising differences in the skill results between the coastal and inland catchments. The implications, bearing in mind the limited sample size, are that either the climate models or the downscaling methods perform less well for inland areas than for coastal areas of South Africa. This may be related to the differences in rainfall producing weather patterns that occur (Preston-Whyte & Tyson 1988) across the inland and coastal regions of the country, but such speculation would have to be further investigated by climatologists who have a better understanding of the differences in the climate models and downscaling methods than the authors. These observations and conclusions about regional and GCM differences have to be viewed in the context of the limited sample of GCM data used in the study and it is not possible to reach any conclusions about whether they are statistically meaningful.

Figure 7 clearly illustrates that, based on the number of tests in the higher skill groups across all catchments, GISS, CNRM and MPI represent the top three most skilful models, while GFDL, MIUB, CCCMA and CSIRO fall into the least skilful grouping. Figure 2 represents the uncertainties in future projections across the climate models (variations in the vertical position of the bars) and within the regions for the same climate model (length of the bars). Two of the models identified as being generally less skilful (CSIRO and CCCMA) show some of the largest degrees of uncertainty across and within regions, while GISS and CNRM (identified as skilful) show fewer variations. One question

about the potential value and use of the skill measures is whether rejecting the less skilful models would reduce the uncertainty in future predictions of rainfall. For region X, the best models are CNRM, IPSL and GISS which have a narrower band of uncertainty than all nine models, but still represent predictions ranging from >10% increases to over 5% decreases. For the W. Cape region (H), the majority of the models suggest decreases in rainfall and even rejecting just the CSIRO predictions would reduce the uncertainty considerably. Rejecting the four least skilful models for the K region (MIUB, CSIRO, MRI and CCCMA) would similarly reduce the uncertainty, but would remain between approximately +20% and -8%. The impacts on uncertainty of rejecting some climate models is much lower for the E. Cape coastal region (R) because the highest (MRI) and lowest (MPI) predictions are both generated by models identified as skilful. The same conclusion can be reached about the C/D region where IPSL and MPI are considered skilful, resulting in an uncertainty range of between +35% and -11%.

IMPLICATIONS FOR SIMULATING FUTURE WATER RESOURCE AVAILABILITY

One of the major issues associated with the use of down-scaled GCM data within hydrological models is the uncertainty in the future predictions related to the range of results given by, supposedly, equally credible GCMs. The Pitman monthly rainfall-runoff model (Hughes *et al.* 2006) is being applied within an uncertainty framework (Kapangaziwiri *et al.* 2012) to the historical data (1920–2005) as well as to the near future (2046–2065) period for all nine GCMs. The methods used for generating the rainfall (Equation (1)) and PE (based on Equation (2)) inputs to the model have been referred to above and the model has been initially applied to three headwater quaternary catchments of the Buffalo River (R20A to R20C).

The model assumes uncertainty distributions (assumed to be normally distributed in this study) for the main runoff generation and water balance parameters and generates ensemble outputs (typically 1,000–10,000) using independent Monte Carlo sampling from the parameter frequency distributions. The first step is to assign mean and

standard deviations to define the uncertainty distributions of the most important model parameters and assess whether the ensemble results for the historical period (1920–2005) are appropriate given previous simulated flow patterns in the catchment (e.g., WR2005; Middleton & Bailey 2008). Five out of the total 18 model parameters were treated as uncertain and these are the parameters that largely determine the runoff responses at both high and low flows (Kapangaziwiri *et al.* 2012). The information available from stream flow gauges is also used to guide the process of establishing appropriate parameter distributions but it is recognised that the gauged records are impacted by a number of ill-defined upstream developments and abstractions and are therefore inherently uncertain. The same parameter distributions are then used with the rainfall time series and PE seasonal distributions appropriate to the near future scenarios of all nine GCMs.

Figure 8 shows the seasonal distribution of near future climate change effects on mean monthly simulated runoff for all of the nine GCMs listed in order of skill. The data plotted are the percentage deviations of the GCM monthly means from the historical equivalents and the median simulation ensembles are used in all cases. The overall impression is that the wet season runoff is increased and the dry season runoff is decreased. There is, however, a substantial amount of variation in the changes across the nine GCMs and it is very difficult to identify any pattern in the differences between the more skilful GCMs in this region (MPI, GFDL and MRI) and those with less skill.

Figure 9 shows the envelopes of flow duration curves for the simulations: the ensembles of simulations using historical

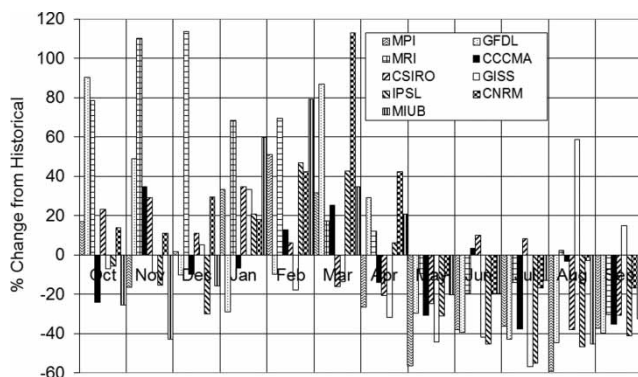


Figure 8 | Seasonal distribution of climate change monthly means for the Buffalo River expressed as a per cent change from historical means (GCMs ordered by skill).

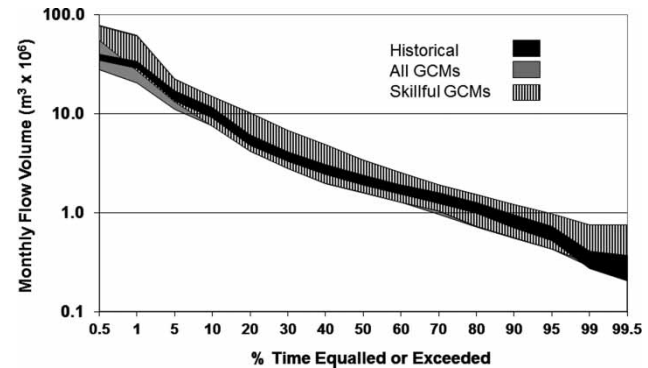


Figure 9 | Flow duration curve envelopes (the upper and lower bounds of 90% of the simulation ensembles) for the near future for all GCMs and the three most skilful, as well as the simulations using historical data.

data; the range of ensemble results for all GCMs; and the ensemble range for the three most skilful models. The results of the near future period simulations suggest that the uncertainty in future water resources availability is substantially increased across the full range of monthly flow volumes. If only the apparently more skilful GCMs are used the uncertainty is marginally reduced only for high flows (equalled or exceeded less than 10% of the time) and the trend is for generally larger high flows. This is a modification of the conclusion reached above (based on only the rainfall data) that suggested there would be little change in uncertainty by selecting skilful models for this region.

CONCLUSIONS

In common with many other studies, this contribution has demonstrated that there is a great deal of variation in the projections of future climates based on downscaled outputs from nine supposedly equally credible climate models. These variations are more evident in the projections of future rainfall patterns than they are in temperature projections. Given the uncertainties associated with using all nine model outputs, this study has investigated the differences in the ability of the models to simulate the characteristics of historical rainfall patterns using four measures of skill. The relatively simple skill measures have been developed to reflect the assumptions used in the bias correction method, to be appropriate for the purposes of hydrological simulations and to acknowledge the constraints of downscaled GCM outputs. They are therefore designed to assess the skill of the different models in

reproducing the main seasonal distribution (Seas_skill), statistical (γ _skill and CV_skill) and serial correlation (SC_skill) characteristics of monthly rainfall time series.

Several methods of assessing the relative differences in GCM skill have been applied for a total of 15 catchments drawn from five regions of South Africa. The general conclusion is that some models are consistently less skilful for the regions assessed (CSIRO and CCCMA), while some are generally more skilful (GISS, CNRM and MPI) with some regionally specific exceptions. There is a substantial variation in skill across the different regions for most models and the results suggest that there are differences in skill, as well as differences in the GCMs skill ranking between coastal areas and inland regions. It was speculated that this result may be a reflection of the well-documented differences in rainfall producing weather patterns between inland and coastal regions of South Africa. However, further investigations with specialist inputs from climatologists familiar with the structure of the GCMs and the downscaling methods would be needed to confirm or reject this possibility.

Figure 2 illustrates the large variation in the near future (2046–2065) projections of rainfall change across the five regions and nine GCMs. Using the skill rankings to limit the GCMs used in each region suggests that the variation and therefore the future uncertainty could be reduced in the X (Eastern Escarpment), H (W. Cape) and K (S. Cape) regions but not in the R (E. Cape) and C/D (Caledon and Modder catchments) regions. To date, the bias corrected near future rainfall data and temperature-based estimates of future evaporation have been used in a hydrological model application (including model parameter uncertainty) within the Buffalo River (R) region. As expected, the results suggest that uncertainties in the availability of water resources based on historical data are substantially increased when considering future water availability. The uncertainties in future high flows are slightly reduced if only the skilful GCMs are included, but the range of uncertainty in the moderate to lower flows is not affected by excluding the less skilful models. These modelling studies will be expanded to the other regions of the country in the near future.

During the review process of this paper, it was pointed out that using uncorrected temperature data to calculate PE, together with bias corrected precipitation data, potentially means that the inputs to the hydrological model may

be physically inconsistent. However, the authors consider that this is unlikely for two reasons. First, there was far greater consistency among the downscaled GCM data sets in the historical temperature estimates than the rainfall. Second, the PE inputs to the model are fixed seasonal distributions and therefore any differences in the predicted time series of temperature are buffered by the use of simple seasonal distributions.

It must be acknowledged that the climate change projections used in this study are not representative of all the possible uncertainties that exist and have been constrained by the data that have been made readily available to the hydrological modelling community by the climate modelling community in South Africa. The study cannot therefore be considered as a comprehensive assessment of climate change uncertainty that includes *inter alia* different time horizons, different emission scenarios or different earth system feedback mechanisms. However, despite these limitations, the study has highlighted the potential impact of climate model uncertainties on future water resources assessment uncertainties, and that was the main purpose of the study. The results suggest that although there are substantial differences in the skill of different downscaled climate model data products, choosing those outputs that can be identified as skilful will not necessarily reduce uncertainty in all regions of the country. Liepert & Previdi (2012) indicate that some of the largest discrepancies in global mean atmospheric water balance are found among the models identified as skilful in this study (such as CNRM and IPSL). Conversely, the least skilful model for South Africa (CSIRO) has a low discrepancy at the global scale. It is therefore important to recognise that the skill assessment results are specific to the geographic scale and the regions used in this study.

ACKNOWLEDGEMENTS

The work undertaken for this study was partially supported through two Water Research Commission of South Africa projects: K5/2018 that is focused on climate change and adaptation strategies for local water boards and K5/2056 that is dealing with several issues of uncertainty in hydrological and water resources assessment modelling. The third author is a post-graduate student supported by

the Carnegie Foundation of New York through the SSAWRN (Sub-Saharan Africa Water Resources Network) component of the RISE (Regional Initiative in Science Education) programme. All of the downscaled GCM data were obtained from the Climate Systems Analysis Group (CSAG) of the University of Cape Town and the authors are particularly grateful to Lisa Coup and Mark Tadross for their assistance in accessing all the data.

REFERENCES

- Anagnostopoulos, G. G., Koutsoyiannis, D., Christpfides, A., Efstratiadis, A. & Mamassis, N. 2010 A comparison of local and aggregated climate model outputs with observed data. *Hydrol. Sci. J.* **56** (7), 1094–1110.
- Bates, B. C., Kundzewicz, Z. W., Wu, S. & Palutikof, J. 2008 Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva.
- Bouwer, L. M., Aerts, J. C. J. H., Van de Coterlet, G. M., Van de Giesen, N., Gieske, A. & Mannaerts, C. 2004 Evaluating downscaling methods for preparing global circulation model (GCM) data for hydrological impact modelling. In: *Climate Change and Contrasting River Basins* (J. C. J. H. Aerts & P. Droogers, eds). CAB International, Wallingford, UK.
- Buytaert, W., Vuille, M., Dewulf, A., Urrutia, R., Karmalker, A. & Célleri, R. 2010 Uncertainties in climate change projections and regional downscaling in the the tropical Andes: implications for water resources management. *Hydrol. Earth Syst. Sci.* **14**, 1247–1258.
- Chen, J., Brissette, F. P. & Leconte, R. 2011 Uncertainty of downscaling methods in quantifying the impact of climate change on hydrology. *J. Hydrol.* **401**, 190–202.
- Chiew, F. H. S., Teng, J., Vaze, J., Post, D. A., Perraud, J. M., Kirono, D. G. C. & Viney, N. R. 2009 Estimating climate change impact on runoff across southeast Australia: method, results, and implications of the modeling method. *Water Resour. Res.* **45**, W10414, doi:10.1029/2008WR007358.
- Fowler, H. J., Blenkinsop, S. & Tebaldi, C. 2007 Linking climate change modeling to impacts studies: recent advances in downscaling techniques for hydrological modeling. *Int. J. Climatol.* **27**, 1547–1578.
- Frost, J. A., Charles, S. P., Timbal, B., Chiew, F. H. S., Mehrotra, R., Nguyen, K. C., Chandler, R. E., McGregor, J. L., Fu, G., Kirono, D. G. C., Fernandez, E. & Kent, D. M. 2011 A comparison of multi-site daily rainfall downscaling techniques under Australian conditions. *J. Hydrol.* **408**, 1–18.
- Haerter, J. O., Hagemann, S., Moseley, C. & Piani, C. 2011 Climate model bias correction and the role of timescales. *Hydrol. Earth Syst. Sci.* **15**, 1065–1079.
- Hagedorn, R., Doblas-Reyes, F. J. & Palmer, T. N. 2005 The rationale behind the success of multi-model ensembles in seasonal forecasting. Part I. Basic concept. *Tellus* **57A**, 219–233.
- Hargreaves, G. H. & Samani, Z. A. 1985 Reference crop evapotranspiration from temperature. *Appl. Eng. Agric.* **1** (2), 96–99.
- Hewitson, B. & Crane, R. 1996 Climate downscaling: techniques and application. *Clim. Res.* **7**, 85–95.
- Hewitson, B. & Crane, R. 2006 Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *Int. J. Climatol.* **26** (10), 1315–1337.
- Huard, D. 2011 Discussion of ‘A comparison of local and aggregated climate model outputs with observed data’. A black eye for the Hydrological Sciences Journal. *Hydrol. Sci. J.* **56** (7), 1330–1333.
- Hughes, D. A., Andersson, L., Wilk, J. & Savenije, H. H. G. 2006 Regional calibration of the Pitman model for the Okavango River. *J. Hydrol.* **331**, 30–42.
- Hughes, D. A., Kingston, D. G. & Todd, M. C. 2011 Uncertainty in water resources availability in the Okavango River Basin as a result of climate change. *Hydrol. Earth Syst. Sci.* **15**, 931–941.
- IPCC 2001 *Emissions Scenarios. A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, New York.
- Kapangaziwiri, E., Hughes, D. A. & Wagener, T. 2012 Incorporating uncertainty in hydrological predictions for gauged and ungauged basins in southern Africa. *Hydrol. Sci. J.* **57** (5), 1000–1019.
- Knutti, R. 2008 Why are climate models reproducing the observed global surface warming so well? *Geophys. Res. Lett.* **35**, L18704, doi:10.1029/2008GL034952.
- Leung, L. R., Quin, Y., Bian, X. M., Washington, W. M., Han, J. & Roads, J. O. 2004 Mid-century ensembles regional climate change scenarios for the western United States. *Clim. Change* **62** (1), 75–113.
- Liepert, B. G. & Previdi, M. 2012 Inter-model variability and biases of the global water cycle in CMIP3 coupled climate models. *Environ. Res. Lett.* **7**, 014006, doi:10.1088/1748-9326/7/1/014006.
- Lumsden, T. G., Schulze, R. E. & Hewitson, B. C. 2009 Evaluation of potential changes in hydrologically relevant statistics of rainfall in Southern Africa under conditions of climate change. *Water SA* **35** (5), 649–656.
- Masson, D. & Knutti, R. 2011 Climate model genealogy. *Geophys. Res. Lett.* **38**, L08703, doi:10.1029/2011GL046864.
- Middleton, B. J. & Bailey, A. K. 2008 Water Resources of South Africa, 2005 Study (WR2005). WRC Report No. TT381/08, Water Research Commission, Pretoria, South Africa.
- Pappenberger, F. & Beven, K. 2006 Ignorance is bliss: or seven reasons not to use uncertainty analysis. *Water Resour. Res.* **42**, W05302.
- Pirtle, Z., Meyer, R. & Hamilton, A. 2010 What does it mean when climate models agree? A case for assessing independence

- among general circulation models. *Environ. Sci. Pol.* **13**, 351–361.
- Preston-Whyte, R. A. & Tyson, P. D. 1988 *The Atmosphere and Weather of Southern Africa*. Oxford University Press, Cape Town, South Africa.
- Prudhomme, C., Jakob, D. & Svensson, C. 2003 Uncertainty and climate change impact on the flood regime of small UK catchments. *J. Hydrol.* **277**, 1–23.
- Reichler, T. & Kim, J. 2008 How well do coupled models simulate today's climate? *Bull. Amer. Met. Soc.* **89** (3), 303–311.
- Sawunyama, T. 2008 Evaluating Uncertainties in Water Resources Estimation in Southern Africa: A Case Study of South Africa. Unpublished PhD Thesis, Rhodes University, Grahamstown, South Africa.
- Schmidli, J., Goodess, C. M., Frei, C., Haylock, M. R., Hurrell, J. W., Ribalaygua, J. & Schmith, T. 2007 Statistical and dynamical downscaling of precipitation: an evaluation and comparison of scenarios for the European Alps. *J. Geophys. Res.* **112**, D04105.
- Schulze, R. E. & Maharaj, M. 2004 Development of a Database of Gridded Daily Temperatures for Southern Africa. WRC Report No. 1156/2/04. Water Research Commission, Pretoria, South Africa.
- Segui, P., Ribes, A., Martin, E., Habets, F. & Boé, J. 2010 Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins. *J. Hydrol.* **2383**, 111–124.
- Wilby, R. L. 2010 Evaluating climate model outputs for hydrological applications. *Hydrol. Sci. J.* **55** (7), 1090–1093.
- Wilby, R. L. & Harris, I. 2006 A framework for assessing uncertainties in climate change impacts: low-flow scenarios for the River Thames, UK. *Water Resour. Res.* **42**, W02419, doi:10.1029/2005WR004065.
- Wilk, J., Kniveton, D., Andersson, L., Layberry, R., Todd, M. C., Hughes, D., Ringrose, S. & Vanderpost, C. 2006 Estimating rainfall and water balance over the Okavango River basin for hydrological applications. *J. Hydrol.* **331**, 18–29.

First received 27 January 2012; accepted in revised form 21 February 2013. Available online 28 March 2013