Evaluating the Consistency between Statistically Downscaled and Global Dynamical Model Climate Change Projections

B. Timbal and P. Hope
Centre for Australian Weather and Climate Research, Bureau of Meteorology, Melbourne, Australia

S. Charles
Commonwealth Scientific and Industrial Research Organisation, Land and Water, Floreat, Australia

(Manuscript received 17 December 2007, in final form 24 March 2008)

ABSTRACT

The consistency between rainfall projections obtained from direct climate model output and statistical downscaling is evaluated. Results are averaged across an area large enough to overcome the difference in spatial scale between these two types of projections and thus make the comparison meaningful. Undertaking the comparison using a suite of state-of-the-art coupled climate models for two forcing scenarios presents a unique opportunity to test whether statistical linkages established between large-scale predictors and local rainfall under current climate remain valid in future climatic conditions. The study focuses on the southwest corner of Western Australia, a region that has experienced recent winter rainfall declines and for which climate models project, with great consistency, further winter rainfall reductions due to global warming. Results show that as a first approximation the magnitude of the modeled rainfall decline in this region is linearly related to the model global warming (a reduction of about 9% per degree), thus linking future rainfall declines to future emission paths. Two statistical downscaling techniques are used to investigate the influence of the choice of technique on projection consistency. In addition, one of the techniques was assessed using different large-scale forcings, to investigate the impact of large-scale predictor selection. Downscaled and direct model projections are consistent across the large number of models and two scenarios considered; that is, there is no tendency for either to be biased; and only a small hint that large rainfall declines are reduced in downscaled projections. Among the two techniques, a nonhomogeneous hidden Markov model provides greater consistency with climate models than an analog approach. Differences were due to the choice of the optimal combination of predictors. Thus statistically downscaled projections require careful choice of large-scale predictors in order to be consistent with physically based rainfall projections. In particular it was noted that a relative humidity moisture predictor, rather than specific humidity, was needed for downscaled projections to be consistent with direct model output projections.

1. Introduction

General circulation models (GCMs) are the cornerstone to generating future projections of climate change. Their global nature, the computational requirements and the challenges in dynamically solving the complexity of the interactions affecting the climate system on small spatial scales mean that direct model outputs (DMOs) are often too coarse to satisfy climate change impact studies (e.g., Wood et al. 2004). Statistical downscaling, establishing a statistical linkage between the large-scale features meaningfully captured by GCMs and the local variables of interest for impact studies (e.g., rainfall), is often needed to bridge this scale gap (Wilby and Wigley 1997). The literature is rich with examples of statistical downscaling models (SDMs) tested and carefully validated for the current climate system [see the extensive review in the Intergovernmental Panel on Climate Change (IPCC) third assessment reports; Giorgi et al. 2001].
However, an ongoing issue is the assumption that the statistical linkage established for the current climate will remain valid in a future (warmer) climate. It is near impossible to directly prove the validity of such an assumption as the climate record required (multicentury, covering different climatic epochs) is not available. However, some attempts have been made by 1) showing that SDMs are capable of reproducing observed shifts in the local surface climate, such as rainfall (Charles et al. 2004; Timbal 2004) or 2) by assessing the validity of the statistical linkage in the surrogate climate of multi-century-long coupled climate model simulations (Frias et al. 2006).

Here we propose to complement previous findings by examining the consistency between physically based projections (using DMOs) and statistically based projections (using two SDM techniques) across a large number of climate models in order to sample how climate models relate large-scale atmospheric dynamics to model rainfall (resulting from the physical parameterization used) and comparing that with the statistical linkage established for the current climate. This consistency is also tested across a range of external forcings covering the likely evolution of the climate during the twenty-first century. There is no a priori expectation that a point-scale statistically downscaled projection will give the same signal as the overlying direct model grid, given that they are fundamentally different (Skelly and Henderson-Sellers 1996). However, averaging across a sufficient number of points to account for within-grid heterogeneities is more directly comparable to the direct model grid value.

Our case study is the southwest corner of Western Australia (SWWA: from 30° to 35°S and from 115° to 120°E), where rainfall in winter decreased in the late 1960s to mid-1970s by about 10%; rainfall has not recovered since, with arguably a further decline since the mid-1990s (IOCI 2002; Ryan and Hope 2005, 2006; Bates et al. 2008). SWWA is a winter rainfall dominated climate, with nearly 75% of annual rainfall falling during the May–October 6-month period. This small region exhibits large heterogeneities with a strong rainfall gradient from wet coastal areas to dry inland (Fig. 1). The timing of the rainfall decline across the region also varies, with the west coast experiencing the strongest decline in the mid-1970s whereas the south coast only recently experiencing a decline (Hope et al. 2006).

2. Data and methods

Gridded daily rainfalls from 1948 onward were obtained from the Australian Bureau of Meteorology’s National Climate Centre (NCC) on 0.25° latitude by 0.25° longitude grid. The data are interpolated from
station data (Jones and Weymouth 1997). Since these data are interpolated, there is some level of spatial homogenization inherent in the gridded values that may minimize the particular local effects at any one station site; in total 31 key data points, evenly distributed across the target region, were selected for this study (Fig. 1). The target region is SWWA, defined as the Indian Ocean Climate Initiative (IOCI) triangle (IOCI 2002). Since the observed gridded rainfall is obtained by interpolation between stations it contains a lot of very small rainfall values. In this study, daily rainfall below 0.3 mm was discarded. The downscaling calibration and projection for two techniques was performed for all 31 grid points and then averaged to obtain a SWWA spatial average covering the region of high winter rainfall, which is comparable to climate model DMOs. We do not expect that DMO absolute values will be correct, as climate models are known to underestimate SWWA rainfall (Hope 2006). The statistical downscaling techniques used here have been shown to correct this underestimation (Timbal 2004; Charles et al. 1999b). The expectation is that the 31 grid points sample the range of local conditions encountered: that is, coastal locations that face either facing the prevailing westerlies or the southern coast, high-elevation locations on ranges, and inland dry areas. Three hypotheses can be made where DMOs and spatially averaged downscaled projections are inconsistent: 1) the spatial averaging does not sample the subgrid heterogeneities well enough and so is not representative of the model grid box, 2) the statistical downscaling is not sufficiently consistent with the large-scale physical forcing, or 3) surface heterogeneity results in unrealistic climate models gridbox averages (e.g., in the vicinity of large mountains, such as the Andes or of the Himalayas, or near land or sea ice). The third explanation is unlikely to be the case for SWWA and the careful selection of the 31 rainfall grid points rules out the first. Thus we are assessing hypothesis 2.

In conjunction with IPCC’s Fourth Assessment of climate change science released in 2007 (Solomon et al. 2007), a new set of global climate model experiments has been produced: the World Climate Research Programme’s (WCRP’s) Coupled Model Intercomparison Project phase 3 (CMIP3) multimodel dataset. Up to 25 GCMs contributed to the CMIP3 dataset; however, because the SDMs rely on daily outputs for the predictors, which were not provided by every modeling group, only a subset of these dataset could be used (Table 1). The models are ranked according to a measure of their sensitivity ($\Delta T$) (last column in Table 1), calculated using

<table>
<thead>
<tr>
<th>Originating group</th>
<th>Country</th>
<th>Acronym</th>
<th>Grid size (km)</th>
<th>$\Delta T$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commonwealth Scientific and Industrial Research Organisation (CSIRO)</td>
<td>Australia</td>
<td>CSIRO-Mark version 3.0 (Mk3.0)</td>
<td>~200</td>
<td>2.11</td>
</tr>
<tr>
<td>National Aeronautics and Space Administration (NASA) Goddard Institute for Space Studies (GISS)</td>
<td>United States</td>
<td>GISS-Model E-R (ER)</td>
<td>~400</td>
<td>2.12</td>
</tr>
<tr>
<td>Canadian Climate Centre</td>
<td>Canada</td>
<td>Canadian Centre for Climate Modelling and Analysis (CCCma) Coupled General Circulation Model, version 3.1 (CGCM3.1) (T47)</td>
<td>~300</td>
<td>2.47</td>
</tr>
<tr>
<td>Meteorological Research Institute</td>
<td>Japan</td>
<td>MRI Coupled General Circulation Model, version 2.3.2 (CGCM2.3.2)</td>
<td>~300</td>
<td>2.52</td>
</tr>
<tr>
<td>Geophysical Fluid Dynamics Laboratory</td>
<td>United States</td>
<td>GFDL Climate Model version 2.1 (CM2.1)</td>
<td>~300</td>
<td>2.53</td>
</tr>
<tr>
<td>Météo-France</td>
<td>France</td>
<td>Centre National de Recherches Météorologiques Coupled Global Climate Model, version 3 (CNRM-CM3)</td>
<td>~200</td>
<td>2.81</td>
</tr>
<tr>
<td>Geophysical Fluid Dynamics Laboratory</td>
<td>United States</td>
<td>GFDL Climate Model version 2.0 (CM2.0)</td>
<td>~300</td>
<td>2.98</td>
</tr>
<tr>
<td>Institut Pierre-Simon Laplace</td>
<td>France</td>
<td>IPSL Coupled Model, version 4 (CM4)</td>
<td>~300</td>
<td>3.19</td>
</tr>
<tr>
<td>Centre for Climate Research</td>
<td>Japan</td>
<td>Model for Interdisciplinary Research on Climate 3.2, medium-resolution version [MIROC3.2(medres)]</td>
<td>~300</td>
<td>3.35</td>
</tr>
<tr>
<td>Max Planck Institute for Meteorology, German Climate Computing Centre [Deutsches Klimarechenzentrum (DKRZ)]</td>
<td>Germany</td>
<td>ECHAM5/Max Planck Institute Ocean Model (MPI-OM)</td>
<td>~200</td>
<td>3.69</td>
</tr>
</tbody>
</table>
the global warming produced by the model using the A1B scenario when approximated by linear regression over the twenty-first century (CSIRO and the Bureau of Meteorology 2007). Only two future emission scenarios for the twenty-first century are used here (up to six future emission scenarios are available in the CMIP3 database): the A2 scenario, a higher-emissions scenario; and the B1 scenario, a lower emission scenario to encompass the likely range of global change in the twenty-first century. Corresponding simulations of the twentieth century were used as a baseline climate.

Daily data were only available for three time slices: 40 yr from 1961 to 2000 for the twentieth century simulation and two 20-yr periods from 2046 to 2065 in the middle of the twenty-first century (designated A2_50 and B1_50) and from 2081 to 2100 at the end of the twenty-first century (designated A2_100 and B1_100) for the A2 and B1 scenarios. All the projections calculated for the future 20-yr time slices are expressed as percentage of the mean values from 1961 to 2000 for both DMOs and downscaled projections. Direct model rainfall changes were calculated using monthly rainfall from the same periods for which daily data were available. Like the observed climatological rainfall distribution, all models simulated a small area of higher annual mean rainfall corresponding to the southwest corner of the Australian landmass compared to dryer inland values. The land grid point with the highest total rainfall in the twentieth century simulation was thus chosen to represent SWWA for each model.

The two statistical techniques used are weather typing schemes based on the view that atmospheric large-scale states relate to a set of local climate variables. They differ from weather generators or transfer function type techniques. The first downsampling technique used is based on a nonhomogeneous hidden Markov model (NHMM). Originally developed to model rainfall occurrence patterns (Hughes et al. 1999) and extended to include amounts (Charles et al. 1999a), the NHMM simulates multisite daily rainfall as a small number of “hidden” (i.e., unobserved) states. Daily transitions between states are conditional on atmospheric predictors. Previous studies for SWWA have shown that the NHMM can reproduce the observed mid-1970s step change in winter rainfall (Charles et al. 2004) and generate projections consistent with a projection from a regional climate model (Charles et al. 1999b). When calibrated to the 31 gridpoint “stations” for May–October using National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) atmospheric reanalyses (Kalnay et al. 1996) for 1958–2003 the optimal predictor set consisted of mean sea level pressure (MSLP), north–south gradient of MSLP, and dewpoint temperature depression at 850 hPa (DTD850). DTD850 is the difference between the air temperature and dewpoint temperature at the 850-hPa level, and thus is a measure of relative humidity in the lower atmosphere that takes into account the ability of a warming atmosphere to hold more moisture. This optimum predictor set was determined by first using local meteorological expert knowledge to produce a list of “candidate” predictors over SWWA, keeping in mind the predictors had also to be available from available climate model output. These were then evaluated by assessing the performance of NHMMs calibrated on various predictor combinations using statistical tests (Bayes information criterion), the subjective assessment of the physical realism of the NHMM states (i.e., SWWA rainfall patterns and associated synoptics), and ability to reproduce at-site and intersite statistics such as wet-day frequencies, wet- and dry-spell length distributions, spatial correlations in occurrences and amounts, intensity distributions, and interannual variability, as outlined in Charles et al. (1999a, 2004). The selected predictors were calculated for the NCEP–NCAR grid points over the SWWA landmass, not exactly the same domain was used for each predictors, the largest domain considered was for MSLP from 116.25° to 123.75°E and from 38.75° to 31.25°S.

The second downsampling approach used is based on the idea of meteorological analogs (Timbal and McAvaney 2001). This is one example of a more general type of SDM based on weather classification methods in which predictands are chosen by matching previous (i.e., analogous situations) to the current weather state. The analog SDM was first developed for daily temperature extremes (Timbal and McAvaney 2001) and then extended to rainfall occurrences (Timbal et al. 2003) and amount (Timbal 2004). As for the NHMM, it was tested for the rainfall in SWWA using NCEP–NCAR reanalyses for 1958–2003 and similar predictors were examined. The optimal combination of predictors was found to be: MSLP, the specific humidity at 850 hPa (Q850), and the zonal component of the wind at 850 hPa (U850). The optimal combination of predictors was chosen by a subjective analysis: first assessing the skill of various individual predictors (limited to variable best reproduced by climate model on a large scale) and then the gain in skill obtained while combining the predictors. A range of statistics was looked at to assess the skill: reproduction of the mean and variance of the observed series, correlation between the observed and reconstructed series on daily and interannual time scale. Only weakly correlated combinations of predictors were tested to avoid overfitting. Several geographical
domains for the predictors were tested, the optimized size was found to be a rectangle centered above SWWA: from 108.5°E to 128.5°E and from 27°S to 39°S, a slightly larger domain than for the NHMM method.

The optimal combination of predictors is notably similar between the two techniques: a synoptic predictor (MSLP) combined with a measure of the zonal air flux (zonal wind or north–south pressure gradient) and a measure of the atmospheric moisture (either specific humidity or dewpoint depression). The third predictor however has the potential to lead to very different signals in a warmer world. Specific humidity has been projected (Held and Soden 2006) to increase with global warming because of an increase in the moisture holding capacity of the warmer atmosphere and strengthening of the hydrological cycle while the concurrent increase in the saturation point will limit the increase in of relative humidity, for example, the dewpoint depression. The impact of the different moisture predictors is tested by adding temperature at 850 hPa ($T_{850}$) to the predictors in the analog model. The combination of $T_{850}$ and the specific humidity at 850 hPa is a closer match to the dewpoint depression at 850 hPa used by the NHMM. Although not selected as the optimal combination of predictors for the analog model, adding $T_{850}$ does not result in a major change in skill. While the day-to-day correlation is unchanged and the interannual correlation slightly reduces from 0.55 to 0.48, there is a better representation of the observed downward trend across the second half of the twentieth century.

3. Rainfall projections and global warming

Climate model rainfall projections for SWWA exhibit very high consistency across the range of climate models compared to other areas of the globe (Solomon et al. 2007). More than 90% of all the CMIP3 climate models indicate a further rainfall decline in winter (CSIRO and Bureau of Meteorology 2007). This very high consistency makes it an interesting case study to evaluate how downscaled projections compare with direct model outputs. First, we examine how much of this consistency can be accounted for by the dependency of the rainfall declines on surface global warming, one of the most robust features of climate model climate change projections (Solomon et al. 2007). This hypothesis is in line with previous analysis of future rainfall projections (Hope 2006) and attribution of the current SWWA rainfall decline to anthropogenic forcing (Timbal et al. 2006).

Rainfall declines for May to October obtained directly from the climate models are plotted as a function of the model global warming for each experiment (Fig. 2), calculated as an anomaly between the future 20-yr periods and 1981 to 2000 (as summarized in Fig. 5 of the summary for policy makers of the IPCC; Solomon et al. 2007). The relationship is strong and explains about half of the variability of the rainfall decline per model simulation ($r^2 = 0.46$). The slope is 9.0% rainfall decline per degree of warming: for the past 50 yr (1956–2005), a decline of 9.9% was observed while the global warming in the same period was of 0.65°C ± 0.15°C.

![Figure 2](image-url)
(Solomon et al. 2007): thus a decline of between 12% and 25% per degree of global warming. There is also an uncertainty attached to the calculated rainfall trend, the figure given here is obtained using monthly gridded data and the full IOCI (2002) “triangle,” as it is the most reliable data. However, the same calculation done on the dataset used here (daily gridded data sampled using 31 points) gives a decline of 5.3% or a rate of 6.5% to 13.5% per degree of global warming. This disparity in trend probably arises from fewer daily stations included in the daily gridded data compared to the monthly network through the 1950s and 1960s. Overall, it appears that the past rainfall decline per degree of warming had likely been larger than the number obtained for future projections. It is not obvious that the larger decline for the current transient climate points toward an underestimation for future model projections. Past studies have shown that only a part of the observed rainfall decline is attributable to anthropogenic forcings (Timbal et al. 2006; Cai and Cowan 2006), while natural variability (Cai et al. 2005) and regional land clearance (Pitman et al. 2004; Timbal and Arblaster 2006) have likely enhanced the signal in the past 50 yr (IOCI 2002).

The rainfall decline—temperature increase relationship is driven by the external forcing strength, ranging from B1_50 (pluses in the top left corner of Fig. 2) to A2_100 (black squares in the bottom right corner of Fig. 2). However in most instances, for a given scenario, there is congruence between the rainfall decline and the measure of model sensitivity used in Table 1 (with a slope between 7% and 10% of rainfall reduction per degree of warming; not shown).

The strength of the relationship between rainfall decline and global warming is of the same magnitude when downscaled projections are considered (expressed as correlation in the fourth row in Table 2). However, most of the linearity comes from the rainfall decline being commensurate with the external forcing, as the relationship between rainfall decline and model sensitivity is weaker for downscaled rainfall projections. The slope of the relationship (fifth row in Table 2) between downscaled projections and global warming varies between 6.6% and 9.6% for May–October average.

4. Consistency between direct and downscaled rainfall projections

As stated, point-scale climate change projections obtained using a statistical downscaling technique and grid average DMOs are fundamentally different and therefore unlikely to be directly comparable. Averag-

| Table 2. Correlation between various rainfall declines (May–October average) from DMOs and using statistical downscaling models: the NHMM; the analog model (Analog); and the analog model with \( T_{850} \) as an additional predictor \((A + T_{850})\) (rows 1–3). Each rainfall decline is also related with the model global warming (row 4) and the slope of the relationship is given (row 5). Correlations are based on 10 global climate models and four experiments, and are all significant at the 99% level. |
|-----------------|-----------------|-----------------|-----------------|
|                 | Pearson correlation (r²) | DMO  | NHMM | Analog  | A + T_{850} |
| DMOs            | 0.75            | 0.63            | 0.65            |
| NHMM            | -0.68           | -0.70           | -0.61           | -0.69       |
| Analog          | -0.68           | -0.70           | -0.61           | -0.69       |
| Slope of rainfall vs global temperature | -9.0% | -9.6% | -6.6% | -9.1% |

ing downscaled results across a large number of points from a gridded high-resolution dataset, however, provides the opportunity to aggregate downscaled projections back up to a scale where they are comparable to DMO projections. Plotting the downscaled rainfall projections averaged across SWWA against the direct projections (Fig. 3a) shows the points are aligned along the diagonal. For any particular point, however, a large difference may exist between DMO and downscaled projections for the given model (up to 20%). Overall the correlation between downscaled and DMO projections varies between 0.57 and 0.75 across techniques (Table 2). The slope (in %/°C) of the best linear fit, intercepting with (0, 0), is 0.75 in the case of the analog approach and 0.93 for the NHMM. For the largest projected changes, the downscaled projections tend to reduce the dramatic rainfall declines projected by the DMOs.

Comparing the two statistical downscaling techniques, the NHMM approach is overall more consistent with the DMOs than the analog approach. NHMM correlations are higher (Table 2) and the slope of the DMO versus downscaled relationship is closer to 1 (Fig. 3). We hypothesize that one reason why the NHMM provides better consistency with DMOs on a larger scale is because the predictors (which differ from those of the analog model) better capture the main large-scale forcings corresponding to the rainfall generation mechanisms producing the DMO rainfall decline.

The consistency between downscaled projections is greatly improved when \( T_{850} \) is added to the optimized combination of predictors of the analog approach (Fig. 3b). The slope of the best fit line between analog based downscaled projections and direct model outputs is not equal to 1, but the differences with the NHMM projections are negligible. The correlation between individual
projections for both techniques remains unchanged (around 0.7) when $T_{850}$ is added. This is not surprising as the correlation between the two analog models is extremely high (above 0.9 in individual months and up to 0.96 across the May to October total rainfall). This suggests that the magnitude of SDM projected rainfall changes has more to do with the combination of large-scale predictors used, rather than the statistical technique itself, providing that the technique has been found suitable when tested on the current climate. This result corroborates the earlier findings of Charles et al. (1999b), where an NHMM using an absolute humidity, rather than relative humidity, predictor could not produce rainfall projections consistent with a dynamic regional climate model. Thus predictor selection should not be based solely on assessing a SDMs performance for current climate. This point, as initially proposed by Charles et al. (1999b), has been recently restated by Vrac et al. (2007) with similar conclusions.

5. Conclusions

Rainfall projections from two statistical downscaling models, averaged across the southwest of Western Australia, were compared with direct model outputs. Eleven CMIP3 climate models were used for two emissions scenarios. The rainfall declines projected by climate models for this region are extremely consistent across models and scenarios, all project further Austral winter rainfall reductions. The projected declines align with the global warming projected by the climate models (about 9% rainfall reduction per degree of global warming).

On a broad scale (i.e., averaging gridded data across several hundred square kilometers), the downscaled projections are highly consistent with the direct model projections. For any particular GCM the downscaled and direct model projections can be quite different but over the large sample of multimodel outputs there is only a small suggestion that downscaled rainfall projections are slightly reduced for the largest rainfall declines projected by GCMs. The NHMM produced better consistency with direct model outputs than the analog approach. A modified analog model with large-scale predictors similar to the NHMM produced a level of consistency equal to the NHMM.

Our results offer additional information on the validity of the statistical linkage used to relate large-scale predictors and local rainfall in a different climate. This is complementary to 1) the validation of these statistical techniques across observed climate change (Timbal 2004; Charles et al. 2004) and 2) the evaluation of the consistency of statistical techniques in the surrogate climates of coupled model simulations (Frías et al. 2006). Our results confirm that statistical linkages based on sound predictor selection, which have a physical basis to drive the local predictands, provide downscaled projections comparable to dynamical model projections. The choice of the large-scale predictors was shown to be critical to make this linkage resilient to climate change and far more important than the choice of statistical downscaling technique.

Acknowledgments. BT is supported by the Australian Greenhouse Office under the Australian Climate Change Science Programme. PH is supported by the government of Western Australia under the Indian Ocean Climate Initiative (IOCI). SC is supported by IOCI and the eWater Cooperative Research Centre, Australia. We acknowledge the modeling groups, the Program for Climate Model Diagnosis and Intercomparison (PCMDI), and the WCRP’s Working Group on Coupled Modelling (WGCM) for their roles in making

![Figure 3](image-url)
available the WCRP CMIP3 multimodel dataset. Support of this dataset is provided by the Office of Science, U.S. Department of Energy. Thanks are due to Aurel Moise and Lawson Hanson from the BoM for transferring the CMIP3 database to Australia. The estimate of model global warming was provided by Julie Arblaster from the NCAR. Useful comments on earlier version of the manuscript were made by Julie Arblaster, Penny Whetton, and three anonymous reviewers.

REFERENCES


