Downscaling of GCM-Simulated Precipitation Using Model Output Statistics

JONATHAN M. EDEN AND MARTIN WIDMANN

School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, United Kingdom

(Manuscript received 18 January 2013, in final form 9 July 2013)

ABSTRACT

Producing reliable estimates of changes in precipitation at local and regional scales remains an important challenge in climate science. Statistical downscaling methods are often utilized to bridge the gap between the coarse resolution of general circulation models (GCMs) and the higher resolutions at which information is required by end users. As the skill of GCM precipitation, particularly in simulating temporal variability, is not fully understood, statistical downscaling typically adopts a perfect prognosis (PP) approach in which high-resolution precipitation projections are based on real-world statistical relationships between large-scale atmospheric predictors and local-scale precipitation. Using a nudged simulation of the ECHAM5 GCM, in which the large-scale weather states are forced toward observations of large-scale circulation and temperature for the period 1958–2001, previous work has shown ECHAM5 skill in simulating temporal variability of precipitation to be high in many parts of the world. Here, the same nudged simulation is used in an alternative downscaling approach, based on model output statistics (MOS), in which statistical corrections are derived for simulated precipitation. Cross-validated MOS corrections based on maximum covariance analysis (MCA) and principal component regression (PCR), in addition to a simple local scaling, are shown to perform strongly throughout much of the extratropics. Correlation between downscaled and observed monthly-mean precipitation is as high as 0.8–0.9 in many parts of Europe, North America, and Australia. For these regions, MOS clearly outperforms PP methods that use temperature and circulation as predictors. The strong performance of MOS makes such an approach to downscaling attractive and potentially applicable to climate change simulations.

1. Introduction

Climate simulations for the twenty-first century with general circulation models (GCMs) show an increase in temperatures and substantial changes in precipitation. However, GCMs are usually not able to reliably provide the local-scale projections that are of greatest relevance to users of climate change information. In the case of precipitation, this is due to model errors and biases on the resolved scales caused by the imperfect parameterization of subgrid processes such as convection, cloud formation, and cloud microphysics, as well as by the coarse spatial resolution and representation of topography.

Downscaling methods are often used to produce output at the smaller spatial scales required by the majority of end users. Over the last 20 years, a large amount of research has focused on the methods and applications of dynamical and statistical downscaling, much of which is summarized in a number of key review papers (e.g., Wilby et al. 2004; Christensen et al. 2007; Maraun et al. 2010). In statistical downscaling, regional climate is assumed to be a function of the large-scale climatic state and physical features of the local environment. The goal is to derive a statistical link between a set of large-scale “predictor” variables (e.g., temperature, humidity, wind, or geopotential height) and a local-scale “predictand.” Dynamical downscaling uses higher-resolution regional climate models (RCMs), which are nested into GCMs over a limited area, in order to represent physical processes at a greater spatial resolution. The dynamical approach is constrained by the availability of RCM simulations and so statistical downscaling is a popular alternative for impact studies because of its relative ease of use and a general performance comparable with output from RCMs.

Typically, in the case of precipitation downscaling for climate studies, a perfect prognosis (PP; also referred to as perfect prog) approach is taken in which a statistical
relationship is established between precipitation observations (the predictand) and simultaneous observed large-scale predictors and then applied to simulated predictors for the future. Although there are general guidelines for the choice of predictors (e.g., Wilby and Wigley 1997; Wilby et al. 2004), there is little consensus about which predictors are most appropriate to ensure that the climate change signal is captured. An alternative to PP statistical downscaling is to derive a statistical correction for simulated precipitation. This approach, known as model output statistics (MOS), requires the statistical link to be derived between precipitation observations and precipitation from some historical simulation; in application, as with PP downscaling, the statistical link is applied to the simulated predictor (in this case, precipitation) for a future scenario. In principle, MOS offers a number of potential benefits over PP. Although only applicable to the numerical model upon which it has been calibrated, MOS is able to explicitly account for model-inherent error and bias. Furthermore, because precipitation parameterizations use large-scale circulation, temperature, and humidity as input, the simulated precipitation can be considered to contain the predictive information of all relevant climate variables.

MOS is commonplace in numerical weather prediction (e.g., Glahn and Lowry 1972; Klein and Glahn 1974; Wilks 2006), but it has yet to be fully exploited in downscaling climate change scenarios. Recently, statistical methods have been developed for linking RCM-simulated precipitation to observations through statistical “bias correction” (e.g., Engen-Skaugen 2007; Graham et al. 2007; Lenderink et al. 2007; Piani et al. 2010b; Themessl et al. 2011). Such statistical corrections may be calibrated on intensity distributions or time series, depending on the type of simulation used in the development stage (Maraun et al. 2010). For instance, when calibration is based on an RCM driven by an atmospheric reanalysis, the temporal evolution of the large-scale weather states, at least in principle, matches that of the real world. Such a simulation allows for the statistical link to be derived between simulated and observed time series for the same period; this can be considered an “eventwise” approach to MOS. Conversely, should calibration be limited to an RCM driven by a standard GCM simulation, which typically does not facilitate data assimilation and thus does not match the specific temporal evolution of observed internal variability, simulation–observation corrections may only be undertaken “distributionwise” (i.e., by mapping long-term means or other aspects of the distribution).

While there is a wealth of recent literature that focuses on MOS correction of RCM output, such an approach is inherently limited by the availability and additional computational expense associated with RCM simulations. An alternative approach is to directly correct the precipitation simulated by the GCMs. Such an approach has been previously applied, for example, in the context of hydrological (Sharma et al. 2007) and crop yield simulations (Ines and Hansen 2006). Statistical correction methods for GCM fields typically operate at the same scale but may, in principle, include both a correction and downscaling step, thus removing the requirement for an additional RCM simulation (e.g., Schmiedl et al. 2006). Piani et al. (2010a) developed a statistical bias correction for GCM-simulated precipitation that was shown to produce more realistic mean and variance compared to the raw model output across the majority of the globe. Given that GCM simulations used for generating climate change scenarios do not assimilate observations when run for a historical period, GCM MOS has generally been limited to distributionwise correction. However, Widmann et al. (2003) demonstrated that it is possible to apply MOS in an eventwise context to correct the precipitation from a reanalysis. The authors showed that the precipitation field from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis, which in this case was considered to be a GCM simulation in which the temporal evolution of day-to-day weather matches reality, is a useful predictor of local-scale precipitation in northwestern United States. However, because MOS corrections are specific to the model upon which they have been calibrated, the statistical corrections of Widmann et al. (2003) cannot simply be applied to other GCMs. With this in mind, Eden et al. (2012) showed that eventwise MOS is possible and has good potential when calibrated on precipitation from a nudged GCM simulation in which the circulation and temperature variables were forced toward corresponding reanalysis fields but the precipitation field is solely determined by model physics.

As part of the development of eventwise MOS for GCMs, Widmann et al. (2003) and Eden et al. (2012) also showed that there is predictive power in the simulated precipitation with respect to temporal variability; the question of how much information about temporal variability the simulated precipitation actually contains is usually not addressed in a distributionwise MOS setup. As discussed in Eden et al. (2012), distributionwise MOS could lead to completely meaningless results if applied in situations where the simulated precipitation does not include information about temporal variability. Eden et al. (2012) also distinguished among three different sources of error in GCM-simulated precipitation: systematic bias in the large-scale circulation or in the response to climate forcings (type 1 errors); random internally
generated variability (type 2 errors); and deficiencies in the convection and precipitation parameterization (type 3 errors). The nudging setup used by Eden et al. (2012) provides a simulation in which the type 1 and type 2 errors are approximately removed, allowing for the type 3 precipitation error to be isolated and quantified. Eden et al. (2012) went on to identify regions where the type 3 error was small and demonstrated that a simple local scaling (bias correction with a downscaling step) has excellent potential as a downscaling method when applied to such regions.

In this paper, we present and evaluate an extended MOS approach for downscaling monthly precipitation from GCM simulations. Using the nudged simulation of ECHAM5, described by Eden et al. (2012), it is possible to perform eventwise calibration between simulated precipitation and gridded observations. Here, we investigate two eventwise methods that attempt to correct location biases in the simulated precipitation field. PP downscaling typically employs so-called nonlocal methods that link local precipitation to spatial patterns (e.g., principal components) of a predictor field for some area of influence. Such methods have previously been applied to derive MOS corrections for a reanalysis (Widmann et al. 2003), but this paper represents the first application of this approach to a standard GCM simulation. Our nonlocal methods, based on maximum covariance analysis (MCA) and principal component regression (PCR), as well as the local scaling method (Eden et al. 2012), which is also considered here, are cross validated and compared against the same nonlocal methods applied in a PP context, using spatial fields from the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40) as predictors.

The remainder of this paper is structured as follows: Section 2 describes the nudged simulation and the observational datasets to be used. The development and cross validation of each MOS method is presented in section 3, with regions of good skill identified. For each region, a comparison with the skill of PP methods is discussed. A discussion is given in section 4, with final conclusions drawn regarding the potential applications for a MOS downscaling correction.

2. Data and methods

Our MOS downscaling corrections are calibrated on monthly observations and precipitation simulated by the nudged ECHAM5 simulation described by Eden et al. (2012). The nudging technique is based on Newtonian relaxation and forces the prognostic fields in ECHAM5 (viz., divergence, vorticity, temperature, and surface pressure) to corresponding daily fields from ERA-40 (Uppala et al. 2005) for the period 1958–2001. Further details about the simulation, including setup and analysis of bias and correlation with observations, can be found in Eden et al. (2012).

Separate models are fitted for each month of the year. In the eventwise setup used here, the temporal evolution of spatial patterns of precipitation is, in principle, synchronous with reality. It is therefore possible to perform a nonlocal MOS correction by extending the spatial domain of the predictor beyond an individual grid point and utilize the predictive information of neighboring areas. Nonlocal MOS can correct precipitation location errors that result from the poor representation of topography in a GCM.

When utilizing predictor information across a spatial domain, it is useful to reduce the dimensionality of the field by either extracting the leading principal components (PCs) or, alternatively, by using predictor patterns that optimally link the predictor field to the predictand. In the case of estimating point-scale precipitation, a standard approach is PCR, in which the leading PCs of the predictor are fed into a multiple linear regression model; this is identical to a one-dimensional canonical correlation analysis (CCA) (Glahn 1968; Bretherton et al. 1992; Cherry 1996; Widmann 2005; Tippett et al. 2008). An alternative to PCR is to use one-dimensional MCA, which is equivalent to using the time expansion coefficient of the regression map of the predictor field on the predictand time series as a predictor in a standard linear regression. It is unclear which method yields the greater skill when applied to independent data (Widmann 2005; Tippett et al. 2008; Eden et al. 2013).

Both MCA and PCR were implemented using ECHAM5-simulated precipitation as the sole predictor (methods referred to as MCA_MOS and PCR_MOS). For fitting, precipitation observations were taken from the Global Precipitation Climatology Centre (GPCC) gridded dataset, which consists of interpolated (land only) rain gauge observations with a resolution of 0.5° × 0.5°. The quality of the dataset is dependent on station density and completeness of observational records, which are both sufficient for the period 1958–2001. A rectangular domain of 20° longitude × 10° latitude centered on each observational grid point was defined as the predictor field. For global applicability and consistency, the position (relative to the predictand) and spatial extent of the predictor domain was constant, although optimizing these parameters may be useful for future studies at specific locations during given seasons. Nonlocal corrections were compared to the results of the simple local scaling method (LS_MOS), which was also fitted using GPCC data and ECHAM5-simulated precipitation. All methods were cross validated using a leave-n-out...
technique in which monthly estimates for a given validation year are determined using an independent fitting period (e.g., Wilks 2006). Here, $n = 7$; the 7-yr period centered on the year to be estimated was omitted from the fitting period so as to reduce the possibility of autocorrelation in neighboring years influencing the statistical model (e.g., Eden et al. 2012, 2013).

To draw conclusions on the performance of MOS relative to PP, the MCA and PCR methods were also used to develop a set of PP downscaling models (MCAPP and PCRPP). In this case, a statistical link was established between GPCC observations and atmospheric predictors from ERA-40. The set of predictors included geopotential height $Z$, temperature $T$, specific humidity $q$, and relative humidity RH at 1000, 850, and 500 hPa. The spatial extent of the predictor domain was considerably larger than that used for the MOS approach (40° longitude × 20° latitude) in order to capture predictive information at the synoptic scale. All PP models were subject to the same cross-validation technique used for the MOS models.

3. Results

a. MOS development and validation

Cross-validated correlation between the downscaled MOS estimates and corresponding observations are shown in Fig. 1. For brevity, we only present results for the months of January and July, so as to best represent a global view of intra-annual variation in downscaling skill. Large-scale correlation patterns are similar for each method, with correlation coefficients greater than 0.7 across much of the Northern Hemisphere throughout the year and consistently greater than 0.8 during the boreal winter months. The regions of smallest long-term bias are also similar, with monthly precipitation totals well captured across the majority of Eurasia and North America during January and July.
The skill of each MOS method varies greatly across different regions and interpretation of model skill must be made on a region-by-region basis. Clearly, all methods lack any discernible skill over large parts of the globe; correlation is weak (around 0.2–0.4) and bias is larger than 10% across the majority of the tropics and areas where observation network quality is poor. Eden et al. (2012) found low correlations between observations and precipitation from the nudged ECHAM5 simulation across the tropics, which the authors attributed to the dominance of convective processes in these regions. While the nudging controls large-scale weather states, the large random component involved in the formation of convective precipitation for a given large-scale atmospheric state means that relatively low correlations may be expected even in the case that the precipitation parameterizations perform well. However, we are able to identify a number of regions where our MOS models perform strongly; each region is given individual focus in section 3b.

b. Regional comparisons of MOS with PP downscaling

It is important to understand not only where MOS downscaling is skillful but also where this approach offers an improvement on PP methods. To date, the majority of downscaling research has focused on Europe and North America (Maruano et al. 2010), and it is possible that MOS can improve precipitation estimations in regions where downscaling has proved difficult. In this section, we systematically compare the skill of MOS and nonlocal PP downscaling models across Europe, North America, and Australia.

Prior to the comparison, we determined which PP predictors perform strongest on a global scale. Table 1 provides a summary of mean correlation statistics for all potential PP predictor variables. It is immediately apparent that the skill of each predictor varies greatly according to both the time of year and the method used. PCR\textsubscript{PP} has slightly more predictive skill overall than MCA\textsubscript{PP} for most variables, particularly for geopotential height and temperature. MCA\textsubscript{PP} appears better suited for models based on atmospheric moisture predictors. In addition to single predictors, four pairs of variables were chosen as combined predictors: 1000-hPa geopotential height and temperature ($Z_{1000}$, $T_{1000}$); 1000-hPa specific humidity and temperature ($q_{1000}$, $T_{1000}$); 1000-hPa geopotential height and specific humidity ($Z_{1000}$, $q_{1000}$); and 1000-hPa geopotential height and 850-hPa specific humidity ($Z_{1000}$, $q_{850}$). Square brackets are used throughout the remainder of the paper to denote where variables have been used as single (e.g., [Z1000]) or combined (e.g., [Z1000,T1000]) predictors. The results in Table 1 provide an overall global performance measure for each (combination of) predictor variable(s).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>January MCA</th>
<th>January PCR</th>
<th>July MCA</th>
<th>July PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z_{1000}$</td>
<td>0.157</td>
<td>0.347</td>
<td>0.151</td>
<td>0.274</td>
</tr>
<tr>
<td>$Z_{850}$</td>
<td>0.157</td>
<td>0.338</td>
<td>0.162</td>
<td>0.280</td>
</tr>
<tr>
<td>$Z_{500}$</td>
<td>0.153</td>
<td>0.276</td>
<td>0.165</td>
<td>0.242</td>
</tr>
<tr>
<td>$T_{1000}$</td>
<td>0.205</td>
<td>0.301</td>
<td>0.199</td>
<td>0.242</td>
</tr>
<tr>
<td>$T_{850}$</td>
<td>0.171</td>
<td>0.273</td>
<td>0.187</td>
<td>0.224</td>
</tr>
<tr>
<td>$T_{500}$</td>
<td>0.124</td>
<td>0.210</td>
<td>0.121</td>
<td>0.164</td>
</tr>
<tr>
<td>$q_{1000}$</td>
<td>0.240</td>
<td>0.314</td>
<td>0.155</td>
<td>0.200</td>
</tr>
<tr>
<td>$q_{850}$</td>
<td>0.227</td>
<td>0.303</td>
<td>0.166</td>
<td>0.192</td>
</tr>
<tr>
<td>$q_{500}$</td>
<td>0.227</td>
<td>0.303</td>
<td>0.166</td>
<td>0.192</td>
</tr>
<tr>
<td>$RH_{1000}$</td>
<td>0.190</td>
<td>0.242</td>
<td>0.147</td>
<td>0.177</td>
</tr>
<tr>
<td>$RH_{850}$</td>
<td>0.210</td>
<td>0.246</td>
<td>0.200</td>
<td>0.238</td>
</tr>
<tr>
<td>$RH_{500}$</td>
<td>0.216</td>
<td>0.255</td>
<td>0.222</td>
<td>0.239</td>
</tr>
<tr>
<td>$Z_{1000}$, $T_{1000}$</td>
<td>0.248</td>
<td>0.359</td>
<td>0.234</td>
<td>0.271</td>
</tr>
<tr>
<td>$q_{1000}$, $T_{1000}$</td>
<td>0.248</td>
<td>0.331</td>
<td>0.219</td>
<td>0.254</td>
</tr>
<tr>
<td>$Z_{1000}$, $q_{1000}$</td>
<td>0.264</td>
<td>0.358</td>
<td>0.217</td>
<td>0.249</td>
</tr>
<tr>
<td>$Z_{1000}$, $q_{850}$</td>
<td>0.260</td>
<td>0.363</td>
<td>0.221</td>
<td>0.261</td>
</tr>
<tr>
<td>$Z_{1000}$, $q_{500}$</td>
<td>0.285</td>
<td>0.392</td>
<td>0.251</td>
<td>0.317</td>
</tr>
</tbody>
</table>

In the following subsections, focus is given to each region individually. We use correlation maps to show how closely the cross-validated mean precipitation estimates from different methods match observations. Again, for brevity, results are not shown for all models fitted on each predictor variable, although we present results that permit a robust assessment of the performance of the MOS models against the strongest performing PP models. The layout of Figs. 2–4 is as follows: In each figure, for both January and July, we present the regional observed GPCC precipitation climatology [panel (i) in each figure] and correlations between observed precipitation and the estimates derived from L5MOS, MCA\textsubscript{MOS}, and PCR\textsubscript{MOS} [panels (ii)–(iv)]. The remaining panels, (v)–(x),
Fig. 2. Correlation statistics for observed and estimated (a) January and (b) July precipitation in Europe (1958–2001). (i) Mean precipitation (mm). (ii)–(iv) Correlation between observed and MOS estimated (LSMOS, MCA\textsubscript{MOS}, and PCR\textsubscript{MOS}) precipitation; correlations are shown where significant at the 5% level. (v)–(x) Differences in correlation of MOS and PP precipitation estimates.
FIG. 3. As in Fig. 2, but for North America (1958–2001).
Fig. 4. As in Fig. 2, but for Australia (1958–2001).
show differences between MOS- and PP-derived correlations. For MCA_{MOS} (PCR_{MOS}), we present differences between those correlations and those derived from MCA_{PP} (PCR_{PP}) using the two strongest performing predictors [panels (vi) and (vii) and panels (ix) and (x), respectively]. As there is no PP equivalent to LSMOS, a comparison is instead made with the MCA_{PP} and PCR_{PP} models fitted with the strongest performing predictors [panels (v) and (viii)].

1) EUROPE

Differences between the correlations of MOS and the most skillful PP downscaling models during January are shown in Fig. 2a. MCA_{MOS} shows consistently higher correlations than the best MCA_{PP} downscaling models, fitted on \([Z_{1000}, T_{1000}]\) and \([Z_{1000}, q_{1000}]\), across much of continental Europe. In eastern Europe, particularly in parts of Ukraine and western Russia, MCA_{MOS} correlations are around 0.4 higher than for MCA_{PP} \([Z_{1000}]\). Geopotential height performs well as a predictor in much of southern Europe, including the Iberian Peninsula and western Balkans. The LSMOS method produces much higher correlations and smaller biases than the best MOS methods, in particular, across much of Europe but is actually outperformed by both PP methods in particularly wet, windward regions of northern Europe and the northern Alps. By contrast, both MCA_{MOS} and PCR_{MOS} exhibit greater skill than the strongest performing nonlocal PP models, particularly in some of Europe’s mountainous regions, including the Alps, Carpathians, and Pyrenees. Schmidt et al. (2007) found a number of PP downscaling models underestimated the magnitude of interannual variations in precipitation across the Alps and so the ability to downscale precipitation in regions of complex topography is a promising trait of the PCR_{MOS} model in particular. It has already been shown that, while the LS_{MOS} model is easily applicable and relatively successful across continental Europe, its skill is lacking in many mountainous areas, where the climate is widely considered to be poorly represented in GCMs in the first instance (Christensen et al. 2007; Randall et al. 2007).

In the summer months, the pattern in European mean precipitation, best described as a north–south divide, resembles the pattern in each MOS correlation map. A possible link between small seasonal precipitation totals and poor skill of downscaled estimates is known to extend to dynamical downscaling; the suite of Prediction of Regional Scenarios and Uncertainties for Defining European Climate Change Risks and Effects (PRUDENCE) RCMs showed greater skill in winter than in summer (Christensen and Christensen 2007; Jacob et al. 2007). Additionally, low correlations across much of central and southern Europe may be attributable to the difficulty of representing precipitation variability associated with convective events. Nevertheless, MOS methods for summer continue to show higher correlation than PP techniques, and the difference is generally larger than during winter. During July, all three MOS models produce similar correlation patterns, with values of at least 0.7 in much of northern and western Europe (Fig. 2b).

Along the western coastlines of the British Isles, Scandinavia, and northern France, MCA_{MOS} correlation is up to 0.6 higher than MCA_{PP} fitted on \([q_{1000}, T_{1000}]\) and \([q_{1000}]\). This suggests that the observed circulation predictors used in the PP models are not able to sufficiently explain the variability of precipitation associated with the prevalence of maritime airmass types, which may in addition have originated in the (sub)tropics at this time of year. Cavazos and Hewitson (2005) noted that precipitation processes during summer may be linked to the lower- and upper-tropospheric circulation more typical of convective, tropical regimes. Similarly, correlation is consistently higher for PCR_{MOS} than PCR_{PP} fitted on \([q_{1000}, T_{1000}]\), although the difference is generally smaller (approximately 0.4) and not so apparent in coastal areas. While mountainous regions are, in general, associated with higher MOS correlations compared to PP, there are some regions such as the southeastern Alps where the strongest PP models outperform all three MOS methods.

It is unclear to what extent orographic effects, which are not captured by MOS, are captured by the PP models in this small area alone.

2) NORTH AMERICA

During the winter months, and to a lesser extent spring and autumn, MOS produces correlations of at least 0.7 across the majority of North America, with the exception of the Rocky Mountains and northern Canada (not shown). Interestingly, LS_{MOS} shows greater skill (i.e., higher correlations and smaller biases) than MCA_{MOS} and PCR_{MOS} to the west of the Rocky Mountains and across the Great Plains. During summer, MOS methods exhibit far less skill, particularly across the dry western United States.

In January, the MOS models all produce high correlations (up to 0.8) across most of the wettest parts of North America (Fig. 3a). Precipitation estimates based on LS_{MOS} show a higher correlation with observations over much of the central United States than either MCA_{MOS} or PCR_{MOS} and the nonlocal pattern-based methods do not appear to add value in mountainous regions. This is the opposite of findings in Europe, where the processes involved in orographic precipitation formation may be better captured in Europe and a correction scheme merely has to account for a location bias.
Along much of the East Coast of the United States, MCAPP exhibits a substantial improvement over MCA_{PP} fitted on [Z_{1000}, q_{850}] and [Z_{1000}, T_{1000}]. This feature is also apparent between PCR_{MOS} and PCR_{PP} fitted on the same variables but to a lesser extent. Other areas of notable improvement include the U.S. Midwest and parts of central and western Canada. Across mountainous areas the performance of PP is similar to MOS suggesting that patterns of circulation and humidity provide no additional information for representing precipitation associated with orographic processes.

During July, strong correlations (greater than 0.7) in MOS-derived estimates are limited to the Pacific Northwest, parts of western Canada, and the eastern and northern shores of Hudson Bay (Fig. 3b). It is unclear in which areas MOS offers consistent improvement over PP. Parts of central United States appear to be better represented by MCAMOS and PCR_{MOS} fitted on [q_{850}] and [Z_{1000}, T_{1000}]. The biases in estimated precipitation totals in these areas has already been shown to be high (Fig. 1) and is likely to be partly due to the large contribution of convective precipitation. It is thus unsurprising that circulation and humidity offer greater predictive power for monthly precipitation of this nature.

The success of various statistical downscaling methods in North America has previously been shown to be highly variable and dependent on the climatology of the study region, and there is little consensus on the most appropriate method/predictor combination (Wilby and Wigley 1997; Wilby et al. 1998). Additionally, RCM simulations used for climate change projections have been shown to provide little additional information on local scales and in some cases estimates for future precipitation lack agreement on the sign of the changes (e.g., Chen et al. 2003). Any improvement that MOS methods can provide may thus prove crucial in areas where projections of precipitation changes are uncertain.

3) AUSTRALIA

Correlation and bias statistics over Australia shown in Fig. 1 suggest that all MOS methods exhibit good skill and that this region merits closer inspection. During the austral summer months, the correlation is as high as 0.7 across much of Australia and the long-term bias is close to zero in the wetter eastern part of the country (not shown). During winter, a definitive pattern of skill is less clear, although the correlation remains higher than 0.7 in the southeast.

More detail of correlation statistics for Australia is given in Fig. 4. In January, while correlations for MOS models are reasonably high across much of the continent, it is only on parts of the eastern and southern coasts that MOS shows an improvement over the most successful PP methods (Fig. 4a). MCAPP exhibits the highest correlations overall when fitted on [RH_{500}] and [q_{850}], suggesting that atmospheric moisture content at higher levels is a good predictor for local precipitation. MCAMOS consistently produces stronger correlations (around 0.2 higher) within the southern half of Australia. Correlations associated with PCR_{MOS} are similar in terms of magnitude and spatial patterns, with the exception of the dry southwest corner of Western Australia and much of the state of Victoria, where correlations are lower than 0.3. In each case, PCR_{MOS} is outperformed by PCR_{PP} fitted with [Z_{1000}, T_{1000}] and, to a lesser extent, [q_{850}]. In the case of the latter area, the lack of skill in MOS is practically relevant given the high population density and agricultural importance of southeast Australia.

During July, precipitation is heaviest along the southern coastlines of Australia (Fig. 4b). MOS estimates yield high correlations (greater than 0.7) in the southeast of Australia, a pattern that is similar in each of the three models. A notable improvement on MOS versus PP correlations is evident in large parts of eastern Australia. In particular, MCA_{MOS} performs substantially better than the strongest MCA_{PP}, fitted on [Z_{1000}, T_{1000}] and [RH_{500}], and PCR_{PP}, fitted on [Z_{1000}, T_{1000}] and [q_{850}] across the Murray–Darling basin and especially away from the coastline (where performance is still higher but to a lesser extent). Conversely, MOS correlations are in general weak across the north of the country. MCA_{MOS} and PCR_{MOS} are clearly outperformed in this region by MCA_{PP} and PCR_{PP}, respectively, fitted on [Z_{1000}, T_{1000}]. MCA_{PP} fitted on [RH_{500}] and PCR_{PP} fitted on [q_{850}] also produce higher correlations than the MOS models, particularly within the continental interior in the case of the latter, although it is known that the station network is poor here and these results may not be robust. In Western Australia, all three MOS models have substantially lower skill. The extratropical influences on precipitation in this region result in greater explained variance from geopotential height and specific humidity at lower atmospheric levels, and both MCA_{PP} and PCR_{PP} based on [Z_{1000}, q_{1000}] marginally outperform MOS.

MOS skill in Australia is clearly highly variable, both spatially and intraseasonally, but in general large parts of eastern Australia, including the Murray–Darling basin, are associated with realistic estimates for much of the year. Previous work has shown that spatial patterns of annual rainfall changes in Australia over the last few decades can generally be split into two regions, with rainfall decreasing in the east and increasing in the west (Taschetto and England 2009). The southeast is particularly important as the factors responsible for long-term drought in this region are not fully understood (Sohn...
4. Discussion and conclusions

This work represents the first attempt to develop nonlocal eventwise MOS for downscaling GCM-simulated precipitation. MOS downscaling models have been developed and compared with PP downscaling models in terms of their skill to estimate local-scale (0.5° × 0.5°) precipitation across the globe. The MOS downscaling models use GCM-simulated precipitation as a predictor variable, and use both local scaling and nonlocal techniques based on one-dimensional MCA and PCR. Both the MCA and PCR techniques derive a link between simulated precipitation across an area of influence and observed precipitation at a given location. PP models were developed using the same MCA and PCR techniques with observed (reanalysis derived) circulation, temperature, and humidity fields as predictors. The skill of each method to estimate local-scale precipitation was assessed using bias and correlation maps following a leave-seven-out cross-validation technique.

MOS downscaling models have been shown to outperform their PP counterparts in many regions of the world. The skill of each MOS model and likewise the improvement over PP methods varies greatly with location and season. Three key continental-scale regions (Europe, North America, and Australia) have been identified where estimates of local precipitation are, in general, skillful. Detailed analysis of each region would be preferable and will be a focus for future work.

All MOS models tend to show similar skill in reproducing temporal variability in the observed record. The local scaling approach is very simple to apply and offers a good alternative to most PP methods. Additionally, in many regions of the midlatitudes (i.e., large parts of Europe) there is often little, if any, improvement gained by employing the far more computationally intensive nonlocal techniques. An exception is across mountainous areas where MCA and (particularly) PCR are able to better capture small-scale variability not represented in local scaling estimates. The success of nonlocal techniques in such regions is a key finding of this work. In principle, the spatial patterns in simulated precipitation identified by MCA and PCR allow for a correction of unrealistic spatial structures of the simulation (Widmann et al. 2003; Marzouk et al. 2010). The poor representation of orography in GCMs, and subsequent difficulties in parameterizing orographic precipitation processes, are a major cause for errors in the spatial structure of the simulated precipitation.

While shown to be skillful in large parts of the midlatitudes, local scaling is not appropriate in tropical and subtropical regions where long-term precipitation means are greatly overestimated. Although local scaling, by construction, completely eliminates the bias in the fitting period, the substantial bias found based on cross validation demonstrates that this method lacks stability in such regions. The dominance of convective processes in rainfall formation in (sub)tropical areas makes a fixed scaling of precipitation unsuitable. The overestimation is particularly prominent within the dry subtropical bands of high pressure and at all times of year.

When applying downscaling methods in a climate change context, the usefulness of the results depends on whether the simulated predictors capture the climate change signal and whether the statistical relationship obtained in the fitting period also holds in a changed climate. With respect to both aspects, MOS can be expected to have advantages compared to PP. First, the simulated precipitation used as the predictor in MOS combines in a physically consistent way the changes in many other atmospheric variables and it seems thus less likely that climate change signals will be missed. Second, the fact that when using MOS most atmospheric processes are still explicitly calculated in the numerical model rather than being replaced by a statistical (and often linear) model makes it less likely that changes in the climate and thus in the statistical relationships between different variables will manifest themselves as instabilities of the statistical model. Whether MOS does indeed provide better estimates of low-frequency precipitation variability remains to be shown in further studies.

As discussed in the introduction, the nudging setup used for calibrating MOS models provides a simulation in which the type 1 and type 2 errors are approximately removed. In application to climate change simulations, both error types would be present and the skill of the methods can thus be expected to be lower. However, as both MOS and PP can be expected to be equally affected by error types 1 and 2, the performance of each approach relative to the other can be expected to be similar in a climate change simulation. Our results can thus be expected to be transferable to the application of PP and MOS downscaling methods to climate change simulations.

Most GCM simulations available for the twentieth and twenty-first centuries [e.g., phase 5 of the Coupled Model Intercomparison Project (CMIP5)] are free running and thus in these cases MOS downscaling of GCM output is limited to a distributionwise setup. The generally better performing and better validated eventwise
MOS downscaling can currently be applied only to a few atmospheric GCMs for which simulations nudged to reanalyses are available [e.g., ECHAM6 and Hadley Centre Coupled Model, version 3 (HadCM3)], but such nudged simulations could likely be conducted without too much effort with other GCMs.

Acknowledgments. ECHAM5 was provided by the Model and Data group at the Max Planck Institute for Meteorology in Hamburg, Germany. ECHAM5 nudging modules were developed by I. Kirchner. GPCC Precipitation data were provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado (http://www.esrl.noaa.gov/psd/). We also thank S. Rast and D. Grawe for assistance in setting up the local running of the ECHAM5 simulations.

REFERENCES


