Vol. 28: 95–107, 2005 | CLIMATE RESEARCH Clim Res | Published March 16

Performance of NCEP–NCAR reanalysis variables in statistical downscaling of daily precipitation

Tereza Cavazos[1,*], Bruce C. Hewitson[2]

[1]Departmento de Oceanografía Física, Centro de Investigacion Cientifica y de Educacion Superior de Ensenada (CICESE), Km 107 Carretera Tijuana-Ensenada, Ensenada, Baja California, 22860, Mexico

[2]Environmental & Geographical Science, University of Cape Town, Private Bag, Rondebosch 7701, South Africa

ABSTRACT: The urgent need for realistic regional climate change scenarios has led to a plethora of empirical downscaling techniques. In many cases, widely differing predictors are used, making comparative evaluation difficult. Additionally, it is not clear that the chosen predictors are always the most important. These limitations and the lack of physics in empirical downscaling highlight the need for a systematic assessment of the performance of physically meaningful predictors and their relevance in surface climate parameters. Accordingly, the objectives of this study are 2-fold: to examine the skill and errors of 29 individual atmospheric predictors of daily precipitation in 15 locations that encompass diverse climate regimes, and to evaluate the best combination of predictors that are able to capture different sources of variation. The predictors utilized are from the National Center for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis. Mid-tropospheric geopotential heights and mid-tropospheric humidity were the 2 most relevant controls of daily precipitation in all the locations and seasons analyzed. The role of the tropospheric thickness, and the surface and 850 hPa meridional wind components appear to be regionally and seasonally dependent. The predictors showed low performance in the near-equatorial and tropical locations analyzed where convective processes dominate and, possibly, where the reanalysis data sets are most deficient. Summer precipitation was characterized by the largest errors, likely also due to the enhanced role of convection and sub-grid scale processes. Nevertheless, the model was able to reproduce the seasonal precipitation and the phase of daily events in the mid-latitude locations analyzed. In general, the proposed downscaling models tended to underestimate (overestimate) large (small) rainfall events, which reveal the sensitivity of the downscaling to the spatial resolution of the predictors.

KEY WORDS: Climate downscaling · Daily precipitation · Skill of predictors · Artificial neural networks

1. INTRODUCTION

Future climate change scenarios are generally derived from projections of climate change undertaken by general circulation models (GCMs). However, the coarse spatial resolution of such projections does not capture the local features needed for regional impact assessments. Statistical downscaling has gained significant acceptance as a pragmatic and computationally efficient approach for adding fine-scale detail to global projections. One of the most extensively tested uses of downscaling is in projecting regional climate change and its impacts on crops, water resources, and terrestrial ecosystems (Leung et al. 2003). Unfortunately, the variety of predictors and downscaling methodologies utilized in different studies (Giorgi et al. 2001) make comparative evaluation difficult. Additionally, it is not clear that the predictors selected are the most relevant in all cases. Some studies (e.g. Zorita et al. 1995, Hewitson & Crane 1996) have pointed out the necessity to include the most physically meaningful predictors into the transfer functions, as this is the first

*Email: tcavazos@cicese.mx

assumption behind the statistical downscaling approach. Historically, large-scale circulation controls such as mean sea level pressure (SLP) and geopotential heights (z) have been the most widely used predictors of local (or regional) temperature and precipitation in statistical climate downscaling studies (see for example Appendix 10.4 in Giorgi et al. 2001). This is not only because the atmospheric circulation accounts for a significant proportion of the local climate variance, but also due to the longer temporal record of these fields, and the relative skill with which GCMs are able to simulate them. Nevertheless, such circulation fields fail to capture key precipitation mechanisms based on thermodynamics and water vapor content of the atmosphere, which are known to be important factors in the character of precipitation (Trenberth et al. 2003). Hence, availability of observed atmospheric data at local scale and lack of realism of GCM predictors at regional scale has restricted the selection of controlling variables and thus, the credibility of the climate change scenarios.

In the last decade, the climate research community has made a strong effort to improve the quality, length, resolution, and availability of climate data sets and climate models. The National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) global reanalysis presents an opportunity to carry out a systematic assessment of the performance of a wide number of atmospheric predictors (of precipitation) at different levels of the atmosphere, which have been assimilated with a temporally consistent scheme (e.g. Kalnay et al. 1996). The objectives of this study are 2-fold: (1) to evaluate the individual performance of 29 NCEP atmospheric variables as predictors of grid cell area-averaged daily precipitation (see Table 1) in 15 locations around the world that encompass a wide variety of climate regimes (see Table 2) and (2) to construct and evaluate a downscaling precipitation model based on the best set of predictors. Artifical neural networks (ANNs) have shown to be particularly effective in deriving empirical relationships between large-scale atmospheric variables and surface climate parameters (e.g. Hewitson & Crane 1994, Trigo & Palutikof 1999, Cavazos 2000, Reusch & Alley 2002, Yuval & Hsieh 2003). Thus, a downscaling technique based on ANNs is used here to derive nonlinear relationships (transfer functions) between each atmospheric control and local precipitation. The top 10 predictors of daily rainfall in each location will be further explored with a P-mode principal component analysis (PCA), which is a useful tool that prevents the selection of redundant predictors. The final downscaling models will be based on the most relevant (significant) predictors of daily rainfall and will be tested at grid point (see Table 2) and local and regional scales (see Table 3). Although a universal downscaling method valid for all variables and all regions is difficult to find (von Storch 1995), the main focus of this study is on the performance of individual atmospheric variables and their integrated physical links with daily precipitation. For the purpose of this study, rainfall mechanisms will be constrained by the spatial resolution of the predictors and predictand utilized. Under this premise it would be reasonable to expect a certain number of predictors to be regionally independent and few others to be regionally and/or seasonally dependent. We aim to determine possible differences in the integrated performance of the NCEP predictors, such as local versus grid point skill and tropical versus extratropical skill of the downscaling models.

This study is organized in the following way. Section 2 describes the data utilized. Section 3 explains the downscaling methodology based on ANN and rotated PCA. Section 4 describes the best predictors of daily precipitation derived from their individual performance. Section 5 presents the integrated downscaling results of the best predictors of summer and winter precipitation. In this section, a comparison between the model performance at grid point and local scale is assessed in 3 different locations, which evidenced possible biases of the downscaling model and data utilized. The downscaling model is also evaluated in space using an observed gridded precipitation data set centered in Iowa State, USA. Section 6 summarizes the conclusions of this study.

2. DATA

The observational records for this analysis consisted of 29 twice daily (00:00 and 12:00 h UTC) gridded atmospheric variables (8 basic variables at different pressure levels; Table 1) from the NCEP–NCAR reanalysis (Kalnay et al. 1996) from 1980 to 1993. This data set is derived from a global spectral model with a resolution of 2.5° latitude by 2.5° longitude. A common gridded daily precipitation data set with a spatial resolution similar to the NCEP–NCAR reanalysis was used initially for all locations. This data set was obtained from the Goddard Space Flight Center (GSFC) assimilation scheme (Schubert et al. 1993) with a resolution of 2° latitude by 2.5° longitude. To evaluate the differences between the downscaling at the GSFC grid point and at local scale, the downscaling analysis was also assessed at individual stations (see Table 3) located within 3 grid cells in Table 2 (Mex = Mexico, Iow = Iowa, and Spa = Spain). For this, data for 3 stations in Sinaloa, Mexico, were obtained from the data set

Table 1. NCEP–NCAR reanalysis list of predictor variables at different levels utilized in this study. Resolution: 2.5° latitude × 2.5° longitude. SLP: sea level pressure; SH: specific humidity; RH: relative humidity

Atmosphere level	Circulation	Humidity	Thickness
Surface	SLP (slp) U and V winds (*u*0, *v*0) Divergence (*d*0) Vorticity (*vo*0)	SH (*q*0) RH (*rh*0)	500–1000 hPa (*th*1)
850 hPa	Geopotential height (*z*8) U and V winds (*u*8, *v*8) Divergence (*d*8) Omega wind (ω8)	SH (*q*8)	500–850 hPa (*th*8)
700 hPa	Geopotential height (*z*7) U and V winds (*u*7, *v*7) Divergence (*d*7)	SH (*q*7) RH (*rh*7)	
500 hPa	Geopotential height (*z*5) U and V winds (*u*5, *v*5) Divergence (*d*5) Vorticity (*vo*5)	SH (*q*5)	
200 hPa	Geopotential height (*z*2) Divergence (*d*2)		

DAT322 from the Instituto Mexicano de Teconología del Agua (IMTA) in Mexico; data for 3 stations in Iowa, USA, were obtained from the National Climatic Data Center (NCDC), and data for a station in Salamanca, Spain, was obtained from the National Institute of Meteorology (INM), Madrid, Spain. The 3 stations selected in Sinaloa, as well as those in Iowa, are characterized by similar mean daily precipitation and standard deviations during the winter and summer of the study period (1980–1993). The precipitation of the 3 stations in each location was averaged at daily time scales to have a better representation of the regional precipitation.

Forty grid points of observed daily precipitation centered in the State of Iowa (see Table 3) were also used in the analysis to spatially validate the proposed downscaling model at regional scale with a realistic precipitation data set. These data are based on rain-gauge measurements and have a resolution of 1° latitude by 1° longitude; they were obtained from the NCEP/Climate Prediction Center 'unified' US Mexico precipitation data set described by Higgins et al. (2000).

The downscaling was assessed for the December–January–February (DJF) and June–July–August (JJA) seasons on daily timescales for a period of 13 yr (1980–1993) of data, and individually tested during the wettest and driest years of the period in selected locations (Mex, Iow, and Spa).

3. DOWNSCALING METHODOLOGY

3.1. Artificial neural networks

The data analysis is based on non-linear ANNs and rotated PCA. ANNs have shown to be particularly effective in deriving empirical relationships between large-scale atmospheric variables and surface climate parameters (e.g. Hewitson & Crane 1994, Trigo & Palutikof 1999, Cavazos 2000, Reusch & Alley 2002, Yuval & Hsieh 2003). The ANN technique has the ability to generalize relationships after being trained in a self-

Table 2. Area-average grid point locations used in the analysis. Lat. and Lon. indicate the center of the grid point (2° latitude × 2.5° longitude. ITCZ: Inter Tropical Convergence Zone)

Location	Code	Lat.	Long.	Description of grid cell location and climate
Argentina	Arg	36° S	65° W	Mid-latitudes; continental, semi-arid, Pampas region
Australia	Aus	34° S	150° E	Mid-latitudes; near SE coast of Australia, Black Mountains
Botswana	Bot	24° S	25° E	Tropics; continental, Kalahari Desert
Zambia	Zam	16° S	22.5° E	Tropics; continental, SE African monsoon, Zambezi River Basin
Brazil	Bra	2° S	60° W	Equatorial; Amazon Basin, near Manaus.
Pacific Ocean	Nin3	Eq	120° W	Equatorial; Niño 3 region
Costa Rica	Cri	10° N	85° W	Near equatorial; ITCZ, NE trades, coastal
Bangladesh	Ban	24° N	90° E	Tropics; SW Asian monsoon, cyclones, mouths of the Ganges/Brahmaputra Rivers
Mexico	Mex	26° N	110° W	Tropics, core of North American monsoon, coastal and mountains, Sinaloa state
China	Chi	30° N	110° E	Subtropics; SW Asian monsoon, continental, mountains
Azores	Azo	36° N	30° W	Subtropics; Mediterranean climate, Azores high, ocean
Spain	Spa	40° N	7.5° W	Mid-latitudes; Mediterranean climate, mountains
USA	Iow	43° N	95° W	Mid-latitudes; continental, low level jet, Iowa, Great Plains
Germany	Ger	50° N	10° E	Mid-latitudes; continental
Siberia	Sib	52° N	110° E	Mid-latitudes; continental, semi-arid, near Lake Baikal

Table 3. Description of local stations (Stn) and grid points of observed precipitation used to validate the downscaling models. Stn-ID is the identification code from the data source, except for the grid points in Iowa. Grid points have 1° latitude × 1° longitude resolution. IMTA: Instituto Mexicano de Teconología de Agua; NCDC: National Climatic Data Center; NCEP/CPC: National Center for Environment Prediction/Climate Prediction Center; INM: National Institute of Meterology (Spain)

Location	Code	Stn-ID	Lat. (°N)	Long. (°W)	Source
Sinaloa	Mex	25023	26.33	108.62	IMTA
	Mex	25017	26.73	108.27	IMTA
	Mex	25110	25.22	107.00	IMTA
Iowa	Iow	137161	42.41	94.62	NCDC
	Iow	137979	42.64	95.15	NCDC
	Iow	137147	43.44	96.17	NCDC
Iowa	Iow	Grid points	40.00–43.44	97.00–90.00	NCEP/CPC
Salamanca	Spa	2867	40.49		INM

organizing learning process through a procedure that minimizes the output error through a back-propagation method (Cavazos 1999). The ANN back-propagation algorithm used here is the feedforward NevProp 3, which allows for randomization of the input data and bootstrapping for model resample validation. The algorithm was obtained from the University of Nevada Center for Biomedical Modeling Research (Goodman 1996) via anonymous ftp.

The general details of the ANN methodology are documented in Hewitson & Crane (1996) and Cavazos (1999, 2000); thus only a short description suffices here. The atmospheric predictor time series in a particular location was pre-processed separately for DJF and JJA using the following approach: (1) the predictor variable contained twice daily information from the target location (Table 2) and its 8 neighboring grid points, and (2) the predictor variable in each grid point was lagged over 36 h to account for current and antecedent conditions. The composite time series of atmospheric predictors are input to the ANN via the input layer that is connected to a hidden layer through a system of hidden connections, which initially contain random weights (Cavazos 1999). The hidden layer then links to a single output, which is compared to the target predictand (daily precipitation) during the training process. The root mean squared error RMSE between the output and the predictand is back-propagated through the ANN to determine the weights for the next training iteration (Cavazos 1999). Training continues until the RMSE of the test data set reaches a global minimum. For each season (DJF and JJA), 75% of the 13 yr of data (966 d) was used to train the network and the remaining 25% (322 d) was used as an independent validation data set. A bootstrapping technique was implemented during the training process to estimate the impact of model variability. Twenty bootstrap runs with 5 random sub-samples were used during training to estimate the global minimum error and to obtain the following measures of accuracy: coefficient of determination (R^2) between observed and simulated (downscaled) precipitation, skill ratio between simulated and observed standard deviation (SD_s/SD_o), mean simulated (P_s) and mean observed precipitation (P_o), mean absolute error ($MAE = N^{-1}\Sigma_i |P_{si} - P_{oi}|$), and root mean squared error ($RMSE = [N^{-1}\Sigma_i(P_{si} - P_{oi})^2]^{1/2}$) (N is the number of observed days). Based on these measures of accuracy, the top 10 predictors of daily precipitation during DJF and JJA at each target location were selected for the next diagnostic analysis.

3.2. Rotated principal component analysis

One question related to the selection of predictors is whether processes throughout the troposphere need to be included in downscaling analyses. To answer this question we explored the unique (orthogonal) sources of variance of the top 10 predictors of daily precipitation with a P-mode rotated PCA (Richman 1986) in each location. The P-mode PCA consists of using different variables associated with a particular process (e.g. precipitation) in a single location; it is a useful tool that prevents the selection of redundant predictors. In this analysis, the sources of the predictors' covariance help to identify groups of atmospheric variables that are physically linked to different precipitation mechanisms. The component loadings were computed from the correlation matrix and loadings >0.6 were considered as significant. Then, the components were rotated orthogonally using the varimax method; the number of components to retain for rotation was determined from sharp changes in the scree plot, which varied in every location and season. An example of the rotated principal component (PC) loadings is given in Section 4.2. The most relevant variable (largest loading) from each rotated PC was selected to conform the final downscaling model in each location.

4. RESULTS

4.1. Best single predictors of daily precipitation

The results from the ANN were evaluated based on the performance of each individual suit of predictors at each location. Individual predictors with the best performance (R^2 > 15% and skill > 15%) were averaged over all locations separately for the corresponding winter and summer seasons, as shown in Table 4. According to these results, the mid-tropospheric circulation (*z*5, *z*7) is the most relevant predictor of daily rainfall. In winter the role of *z*5 in precipitation is likely related to the equatorward migration of the mid-tropospheric flow and associated changes in the location of jet streams and storm tracks. During the summer, the poleward retraction of the mid-tropospheric circulation and warming of the troposphere suggest that precipitation processes may be more directly linked to lower- and upper-tropospheric circulation, as is common in convective and monsoon regimes. This is reflected in the role of the geopotential heights, with *z*7 as the most significant circulation variable (Table 4), on average. The moist condition of the mid-tropospheric air is the second most important factor in precipitation mechanisms. This is not surprising, at least for the warm season, since moist air is associated with vertical motion and convective processes. During winter, the surface meridional wind component (*v*0) and sea level pressure (slp) also appear in the top variables, suggesting an influence from surface meridional synoptic systems on precipitation. The significance of the temperature of the tropospheric layer (*th*1) and low-level zonal wind (*u*8) are most apparent during summer (Table 4) when it is common for an expansion of the troposphere due to monsoonal circulation (e.g. Webster et al. 1998).

4.2. Common sources of variation: precipitation mechanisms

When constructing the final downscaling precipitation model, a few questions arise from Table 4: Is it redundant to use both mid-tropospheric moisture variables, specific humidity (*q*7) and relative humidity (*rh*7), in the winter downscaling model? Is it useful to include *z*7 and *z*2? To answer this, we quantified and compared the mutual relevance of the predictors through a P-mode rotated PCA to avoid utilizing variables that contributed to common sources of variation. A rotated PCA was assessed for winter and summer in each location and the most significant variable from each PC was selected to assemble the final downscaling models. The rotated PCA was initiated with the top 5 variables in each location, and then the other variables were added one by one to see if there was any distinctive change in the variance explained by the rotated components. As an example, Table 5 shows the rotated PC loadings of the top atmospheric variables during winter and summer in Germany (Ger, 50° N, 10° E); the best results were obtained with 9 and 10 variables, respectively. The results in Table 5A show that the geopotential heights associated with the underlying 'tropospheric circulation' in PC1 account for the largest amount of the daily atmospheric variance (50%), followed by the 'mid-tropospheric humidity' component in second place (31.75%). This further supports the primary role of the mid-tropospheric circulation in precipitation processes, as well as the ample use of geopotential heights in climate downscaling. The 2 components associated with the circulation and humidity of the mid-tropospheric layer explain 82% of the daily winter atmospheric variance. The surface meridional wind component (*v*0) is positively correlated with the humidity variables (PC2), possibly due to the role of the meridional wind component in the transport of heat and moisture in synoptic systems. The tropospheric thickness (*th*1) is highly correlated with the geopotential heights in PC1. Mid- and upper-tropospheric

Table 4. Top 10 predictors of daily precipitation averaged over all locations in Table 2 for the corresponding winter (W) and summer (S) seasons. Variable codes as in Table 1

	1	2	3	4	5	6	7	8	9	10
W	*z*5	*q*7	*rh*7	*z*7	*v*0	slp	*th*1	*z*2	*z*8	*th*8
S	*z*7	*q*7	*q*5	*th*1	*z*2	*z*5	*z*8	*u*8	*v*8	*v*0

Table 5. Rotated principal component loadings of the top 10 predictors of daily precipitation for a grid point in Germany (Ger, 50°N, 10°E) during (A) boreal winter (DJF) and (B) boreal summer (JJA). An asterisk represents significant loadings (>0.7) and values <0.6 are not shown. The most relevant variable in each factor is shown in **bold**. The total variance explained by the factors is (A) 82% and (B) 80.6%. Variable codes as in Table 1. Prop. var.: proportional variance

Variable	(A)		(B)		
	PC1	PC2	PC1	PC2	PC3
*rh*7		0.6511		0.7518*	
*q*7		**0.9403***		**0.9310***	
*q*5		0.8321*		0.7224*	
*z*8	0.9095*		0.8532*		
*z*7	0.9742*		**0.9682***		
*z*5	**0.9860***		0.9487*		
*z*2	0.9354*		0.9487*		
*th*1	0.7179*		0.8857*		
*v*0		0.6359			
*u*8					**0.9186***
Prop. var.	0.5005	0.3175	0.4715	0.2283	0.1062

ridges and troughs are linked to the underlying tropospheric temperature, which serves to intensify developing systems through changes in the thickness advection. From the precipitation processes point of view, the first component in Germany (Table 5A) provides the large-scale circulation, while the second provides the vertical motion, and possibly convection, through the mid-tropospheric moisture (*q*7) variable. Finally, the thickness combines with the mid-tropospheric circulation to develop a system, which depends very much on the amount of moisture to precipitate. In mountain locations, such as the grid cell in Spain (Spa, 40° N), *rh*7 also appeared significant in PC2; but in all locations and seasons the most significant variable in the moisture component was *q*7.

The summer results in Germany (Table 5B) are fairly similar to those of winter (Table 5A), except that the best results were obtained adding the low-level wind component (*u*8). It should be noticed, however, that in some locations like the grid cell in Iowa, where the low-level jet can be an important precipitation mechanism during the summer (e.g. Giorgi et al. 1996, Moore et al. 2003), the best results were obtained adding the meridional wind component at the 850 hPa level (*v*8) rather than *u*8. Thus, it appears that the role of the low-level wind component in the mid-latitudes during the summer season is regionally dependent. The significant variables in each factor of Table 5 suggest several combinations of predictors of precipitation. It is recommended, however, to use only 1 significant variable from each factor to avoid redundancy when constructing the final precipitation model. A similar method was assessed in all locations.

4.3. Missing predictors

Kidson & Thompson (1998) in a downscaling study documented that vorticity is a good predictor of precipitation, but their results were much better for monthly than for daily timescales. In our study, vorticity (*vo*0, *vo*5) as well as low- and upper-level divergence (*d*8, *d*2) played a minor role in daily precipitation in all the locations analyzed. Generally, in tropical convective regions low-level convergence is accompanied by upper-level divergence (Webster et al. 1998). The summer results in Table 6 are partially consistent with this premise, as the low-level wind component (*u*8) and the 500–1000 hPa thickness (*th*1) appear as independent factors for tropical rainfall. Unfortunately, the accuracy of tropical divergence derived from global analyses has been suspect (e.g. Trenberth & Olson 1988, Liebmann et al. 1998) due in part to the sparse number of upper-air stations in the Tropics. Additionally, the lack of significance of *d*8, *d*2, and vorticity, may be also linked to errors from the differential calculations from which these fields were obtained. This may explain why these key precipitation predictors are seldom used in statistical downscaling.

Table 6. Proposed downscaling models for winter (W) and summer (S) precipitation and for the Tropics (T) and the mid-latitudes (M) according to the most relevant predictor variables from an artificial neural networks and a rotated principal component analysis assessed in each location of Table 2. Top predictors were averaged over the tropical (30° S–30° N) and midlatitude locations analyzed. Variable codes as in Table 1

	W	S
T	p = *f* (*th*1, *q*7, *z*7, *v*0)	p = *f* (*z*7, *q*7, *th*1, *u*8)
M	p = *f* (*z*5, *q*7, slp) or p = *f* (*z*7, *q*7, slp)	p = *f* (*z*7, *q*7, *u*8) or p = *f* (*z*7, *q*7, *v*8)

FINAL DOWNSCALING MODELS

5.1. Proposed downscaling models

Table 6 shows the precipitation functions of the proposed downscaling models of daily precipitation for winter and summer and for the mid-latitudes and the tropical locations analyzed in this study. The grid point performance of the downscaling models is illustrated in Fig. 1 through different measures of accuracy (see end of Section 3.1). During DJF, with the exception of Costa Rica (Cri) and Siberia (Sib), the R^2 and skill increase poleward from the equator (Fig. 1A), with the largest values (>40% of variance explained) observed mainly in some of the Northern Hemisphere locations (China [Chi], Azores [Azo], Spa, Ger). In DJF the error distribution clearly marks the wet (austral summer) and dry (boreal winter) seasons, with the largest MAE and RMSE (>2 mm d^{-1}) during the austral summer. The lowest R^2 (<35%) during JJA (Fig. 1B) is observed in the tropical locations (from Bot to Mex), but the error distribution is not as distinctive as the one seen in DJF (Fig. 1A). The largest errors in JJA are observed in the wet boreal tropics (Cri and Ban) and small errors are seen in the 2 dry-summer Mediterranean locations (Azo and Spa), as well in some of the Southern Hemisphere locations (Nin3, Botswana [Bot] and Zambia [Zam]) due to the dry season. The correlation varied from 0.3 (R^2 = 9%) in the near-equatorial locations (Nin3, Brazil [Bra]) to 0.80 (R^2 = 64%) in the mid-latitudes (Spa). On average, the worst results were obtained in the tropical locations and in the wet season, likely due to (1) enhanced sub-grid scale processes, such as convection, not captured by the large-scale predictors and (2) deficiencies of the NCEP–NCAR reanalysis data, especially in the Tropics.

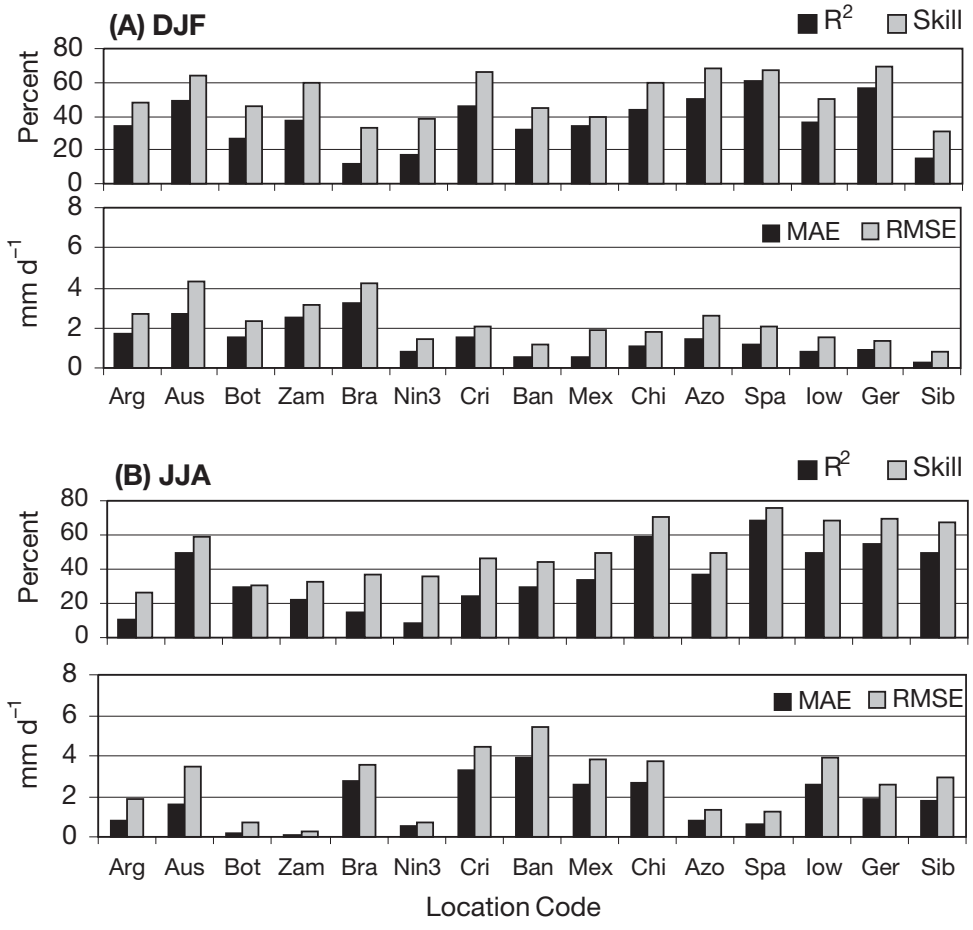

Fig. 1. Measures of performance of the proposed downscaling models shown in Table 6 for each location listed in Table 2. (A) December–January–February (DJF) and (B) June–July–August (JJA). R^2: coefficient of determination; Skill: simulated SD/observed SD; MAE: mean absolute error; RMSE: root mean square error. Locations are displayed from the southernmost (Arg) to the northernmost (Sib) grid cell. The equator is at Nin3

5.2. Downscaling performance at local stations

An important issue in climate downscaling is the difference in the spatial scale of the predictors and predictand. The precipitation functions of the proposed models in Table 6 were obtained from large-scale predictors—predicted at ~2.5° latitude × 2.5° longitude resolution, which could be the scale of a river basin. This means that there still can be large discrepancies between this scale and downscaling precipitation at the local station level. To partially correct for this problem, we averaged the precipitation of 3 stations in Sinaloa and 3 stations in Iowa (Table 3) that have similar mean daily precipitation and standard deviations and performed the downscaling analysis. As expected, the downscaling results were slightly better using 3 stations (regional scale) instead of one (local scale). Table 7 shows the model performance during the wet season in selected grid points and near-by local stations. Sinaloa (Mex), Iowa (Iow), and Salamanca (Spa), respectively, were chosen as examples of poor, medium, and good performances of the downscaling model at grid cell scale (Fig. 1B). There is a general pattern in the 3 cases shown in Table 7: (1) mean daily precipitation (mm d^{-1}) is well simulated at grid point and local scales; and (2) precipitation at station level (Stn) is characterized by larger standard deviation, larger errors, and lower skill than its nearest grid cell. This evidences that the large-scale predictors account for only a fraction of the local precipitation variance. This problem is also reflected in the correlation between daily and monthly precipitation at the station and its nearest GSFC grid cell. Commonly, correlation increases with timescale as seen in the last row of Table 7.

The model showed a poor performance in the tropical stations in Sinaloa (Table 7A). The grid point and the stations utilized are located in the core of the North American monsoon region near the Gulf of California and the Sierra Madre Occidental. Of the predictors

Table 7. Model performance in selected Goddard Space Flight Center (GSFC) grid cells (Mex, Iow, and Spa) and nearby local stations (Table 3) for the corresponding wet season: (A) Mexico (JJA, 26°N, 108°W), (B) Iowa (JJA, 42°N, 95°W), and (C) Spain (DJF, 40°N, 5°W). P_o and P_s are the mean observed and downscaled precipitation (mm d^{-1}), respectively. SD_o and SD_s are observed and simulated standard deviations, respectively. Measures of performance as in Fig. 1. Daily and monthly correlations between observed precipitation at local stations (Stn) and the nearest GSFC grid cell are shown in the last row of each location

	P_o	P_s	SD_o	SD_s	Skill	R^2	MAE	RMSE
(A) Mex	4.73	4.40	4.69	2.31	49.3	33.0	2.61	3.86
Stn (3)	5.46	5.63	9.61	3.54	36.8	17.6	5.66	8.72
Correlation	Daily: 0.53		Monthly: 0.62					
(B) Iow	4.47	4.50	5.39	3.52	65.0	49.0	2.61	3.87
Stn (3)	3.74	3.94	8.17	3.76	46.0	32.5	3.91	6.79
Correlation	Daily: 0.49		Monthly: 0.61					
(C) Spa	1.07	1.15	2.26	1.58	70.0	59.3	0.77	1.45
Stn (3)	1.15	1.10	2.79	1.98	71.0	60.8	0.85	1.75
Correlation	Daily: 0.52		Monthly: 0.83					

utilized in the analysis (Table 6: *z*7, *q*7, *th*1, *u*8), *u*8 and *z*7 were the most significant. The poor performance of the model is possibly due to at least 2 factors: (1) it is known that the NCEP–NCAR reanalysis is able to reproduce the monsoon anticyclone (e.g. Higgins et al. 1997), but it has problems to adequately resolve the low-level moisture flux over the Gulf of California (e.g. Barlow et al. 1998); and (2) sea breezes, local convection and complex topography play an important role in monsoon precipitation over this region (e.g. Berberi 2001). The daily performance of the model at station level during the wettest and driest years of the period (Fig. 2) is consistent with the large errors seen in Table 7A. The model is able to capture the early (15 June 1984) and late (3 July 1982) onsets of the monsoon, but after onset it is unable to identify any of the large precipitation events, possibly due to the poor representation of moisture (Barlow et al. 1998) and local convection. The model underestimated (or overestimated) the seasonal precipitation during the wet (or dry) year.

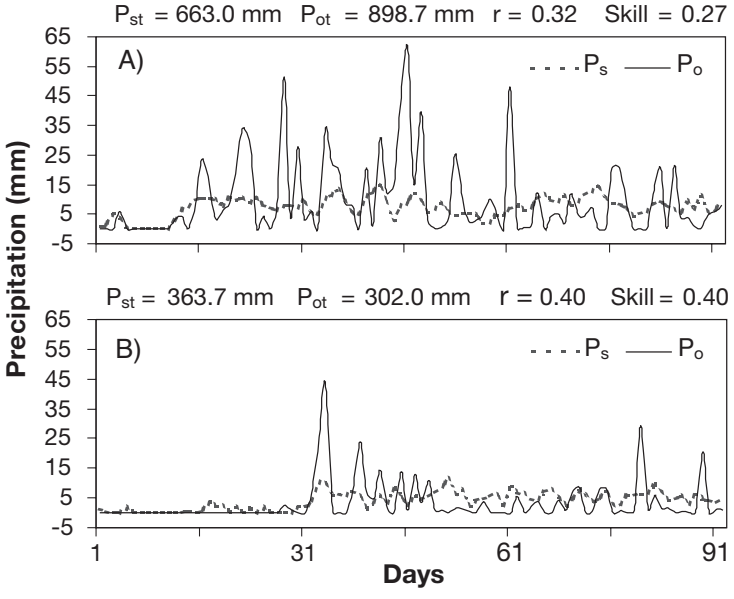

Fig. 2. Local performance of the downscaling model in Sinaloa, Mexico, during the (A) wettest (JJA 1984) and (B) driest (JJA 1982) monsoon years of the 1980–1993 period. P_o and P_{ot}: daily and seasonal observed precipitation, respectively; P_s and P_{st}: daily and seasonal downscaled (simulated) precipitation. r: correlation between P_o and P_s during these 2 years; Skill: proportion of the observed standard deviation explained by the model

The results for Iowa (Table 7B) are better than for Mexico most likely because Iowa is located in the Great Plains, as opposed to the complex terrain of Mexico. The best results were obtained using p = *f* (*z*7, *q*7, *v*8), with *v*8 as the most relevant predictor followed by *q*7. This is possibly associated with the important role of the meridional low-level jet and low-level moisture during the warm-season precipitation in this region (e.g. Giorgi et al. 1996, Moore et al. 2003). Skill (46%) and R^2 (32.5%) at the station level (Table 7B) are still low possibly because summer precipitation in Iowa is associated with intense mesoscale convective systems (e.g. Moore et al. 2003). Nevertheless, at daily timescales (Fig. 3) the model is able to adequately capture the phase of many of the large precipitation events during the 2 extreme years analyzed; as in Mexico, the model tends to underestimate (overestimate) large (small) events, but at seasonal timescale the simulation is good. The wet conditions observed in Iowa during the summer of

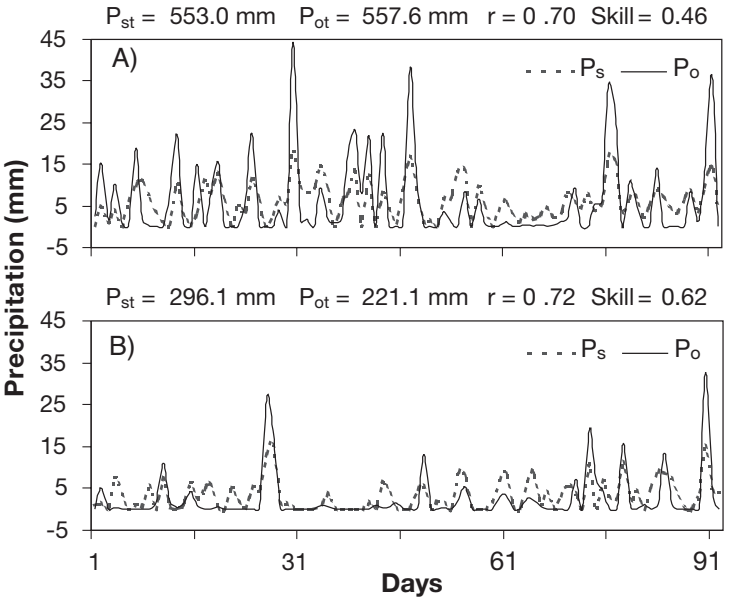

Fig. 3. Local performance of the downscaling model in Iowa, USA, during the (A) wettest (JJA 1993) and (B) driest (JJA 1985) years of the 1980–1993 period. Measures of performance as in Fig. 2

1993 (Fig. 3A) were also observed over the central United States; these regional floods have been extensively documented elsewhere (e.g. Giorgi et al. 1996, Anderson et al. 2003).

The best model results of the proposed downscaling models were obtained in the grid cell in Spain (Spa, Fig. 1), which is characterized by winter precipitation. Salamanca is located within this grid cell and was chosen for the local downscaling. The measures of performance in Table 7C indicate that there are only small differences between the values at the GSFC grid cell and Salamanca, possibly because winter precipitation is mostly controlled by large-scale mechanisms. Correlation between precipitation in Salamanca and its nearest GSFC grid cell considerably increases at monthly timescales. This points out that precipitation at grid cell scale is not necessarily representative of local daily precipitation, which is of general knowledge. However, if the synoptic events observed at grid scale are also observed at regional and local scales, then we can be confident that the downscaling model will be able to capture a large portion of the daily variability associated with the large-scale forcing mechanisms. To illustrate this point Fig. 4 shows the daily performance of the downscaling model during the wet season for 2 winters in Salamanca and its nearest GSFC grid cell (Spa, Fig. 1). Synoptic events in Salamanca closely resemble those at large-scale (GSFC), but with different intensities. The downscaling model reproduced with a high degree of realism the phase of the daily precipitation in Salamanca (Fig. 4A), but again tended to underestimate (overestimate) large (small) events. This bias is not unique to the present analysis; some studies have shown that

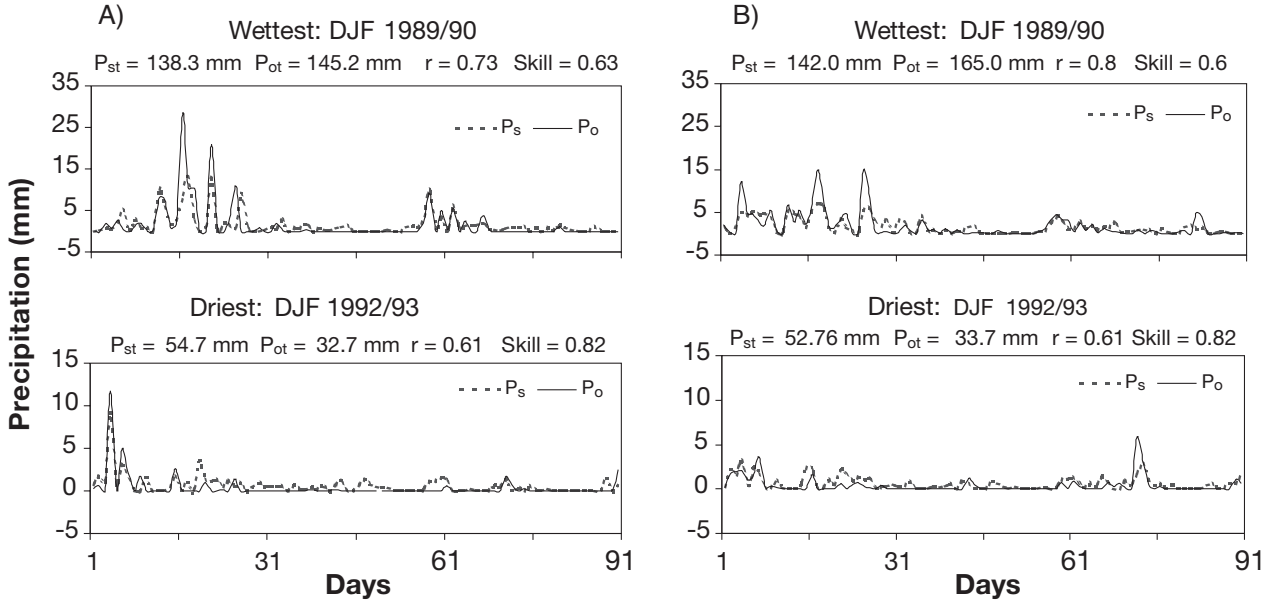

Fig. 4. Daily performance of the proposed downscaling precipitation model for the wet season in (A) Salamanca and (B) the nearest Goddard Space Flight Center (GSFC) grid cell centered at 40° N and 7.5° W for the wettest and driest winters of the 1980–1993 period. Measures of performance as in Fig. 2

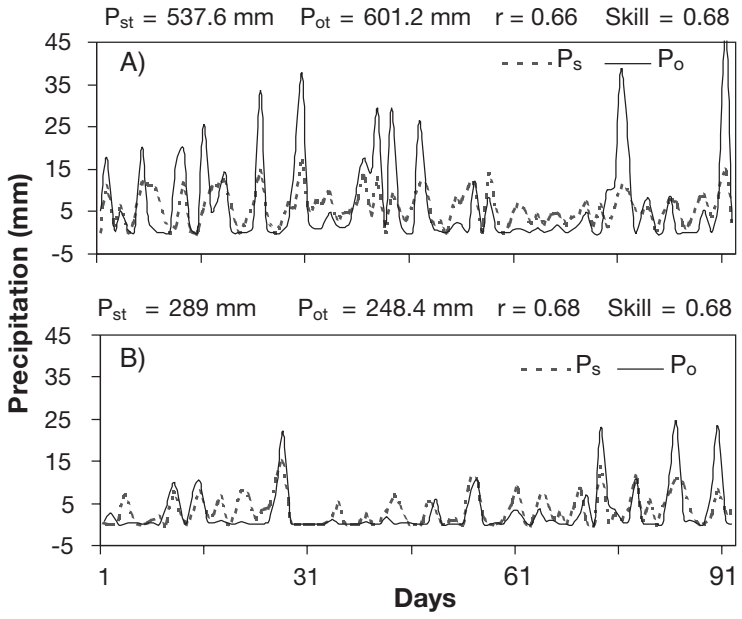

Fig. 5. Model performance from an observed grid point (42° N, 95° W; 1° × 1° resolution) in Iowa during the (A) wettest (JJA 1993) and (B) driest (JJA 1985) years of the 1980–1993 period. Measures of performance as in Fig. 2

nested regional climate models (RCMs) also tend to simulate too many light precipitation events compared with station data (Christensen et al. 1998, Kato et al. 2001, Trenberth et al. 2003) due to cloud parameterization problems (Kain 2004).

5.3. Downscaling performance at regional scale: Iowa

The lack of high-density observational networks has been a major hindrance to the development and evaluation of downscaling techniques (Leung et al. 2003). However, some institutions have produced gridded climatic data sets based on extensive observational records such those for Europe from the Climatic Research Unit (CRU) of the University of East Anglia (New et al. 2000) and those for the United States from the NCEP/CPC (Higgins et al. 2000), among others. In this section we used the NCEP/CPC daily data set at 1° resolution (100 km) to evaluate the downscaling model over the State of Iowa (Table 3).

First, we selected the nearest grid point (42° N, 95° W) to the stations utilized before (Table 3) to test the downscaling model during the boreal summer using p = *f* (*z*7, *q*7, *v*8), and compared the local downscaling results with the downscaling in this grid point during the wettest and driest years of the period. On average, *v*8, *q*7, and *z*7 contributed 74, 20, and 6%, respectively, of the daily precipitation variance, consistent with the results obtained at station level in the last section. The measures of performance at grid point (Fig. 5) and local scale (Fig. 3) are of the same order, except for the skill of the wet year, which is much larger at the grid point. Although the observed seasonal precipitation in these 2 years is a bit larger in the grid point compared to the local stations, there is a very good consistency in the synoptic events that affected both scales. As mentioned in the last section, the wet conditions that dominated the summer of 1993 were observed over most of the Great Plains.

Fig. 6 shows the mean daily regional results of the downscaling analysis for JJA over Iowa. The model slightly overestimated the summer precipitation (~0.4 mm d^{-1}); it is able to explain between 45 and 60% of the daily standard deviation (skill) and between 30 and 42 of the daily variance according to the coefficient of determination (R^2). The spatial patterns of the skill and R^2 are similar but with smaller values of R^2, since this measure is particularly sensitive to outliers (Legates & McCabe 1999).

6. SUMMARY AND CONCLUSIONS

The motivation for this study was the need to derive more realistic regional climate change scenarios in statistical downscaling analyses. Statistical downscaling of precipitation has often focused on the link between local precipitation and atmospheric circulation, but it is questionable whether changes in precipitation resulting from global warming can be derived from changes in atmospheric circulation alone (e.g. Wilby & Wigley 1997, Hewitson 1999, Giorgi et al. 2001, Leung et al. 2003). However, there is little systematic work explicitly evaluating the relative skill of different atmospheric variables as predictors of precipitation (Giorgi et al. 2001). The focus of this study was to clarify this issue by evaluating the relative performance of 29 individual NCEP–NCAR reanalysis variables as predictors of daily precipitation in 15 locations worldwide characterized by diverse climate regimes. The assessment focused first on the relative skill and errors of individual predictors and their physical linkages with precipitation. The objective was to determine a downscaling model (i.e. transfer function) that included the most relevant predictors of daily precipitation in different regions of the world. A downscaling technique based on ANNs was used to diagnose and to evaluate the role of each individual predictor in each target

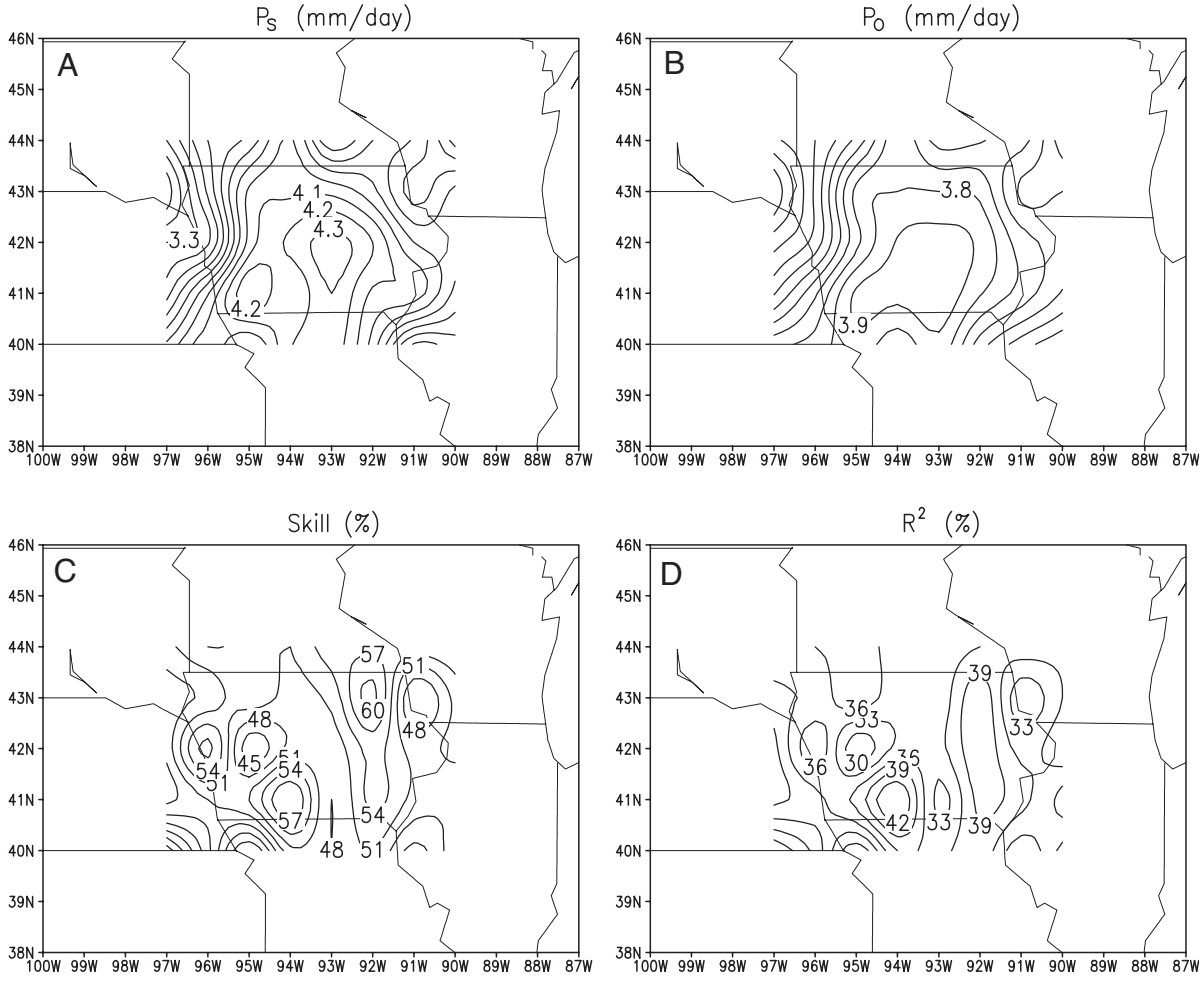

Fig. 6. JJA daily measures of performance of the downscaling precipitation model in 40 observed grid points of 1° latitude × 1° longitude resolution over the State of Iowa. (A) Mean daily downscaled precipitation (P_s), (B) mean daily observed precipitation (P_o), (C) Skill (downscaled standard deviation/observed standard deviation), and (D) coefficient of determination (R^2). Period: 1980–1993

location and a rotated PCA was then used to choose the most relevant predictors that contributed to different sources of variation.

The best performance and the largest number of potentially skillful predictors were obtained in the mid-latitude locations (Table 2) and for the dry/cool season. Mid-tropospheric circulation ($z7$, $z5$) and mid-tropospheric specific humidity ($q7$) were the most relevant predictors of daily precipitation in all the locations and the 2 seasons analyzed (DJF, JJA). This contrasts with most downscaling studies which are mainly based on circulation predictors. Additionally, the tropospheric thickness, and the surface meridional and 850 hPa wind components also emerged as secondary (or even primary) predictors of daily precipitation, but they were regionally and seasonally dependent. The largest errors were obtained during the wet season, indicating that the downscaling models poorly captured sub-grid scale precipitation processes, such as convection. In the tropical locations, where reanalysis data are possibly most deficient, errors were large for both winter and summer. These inconsistencies highlight the need to improve observational records and reanalysis data sets, especially in the Tropics.

Downscaling performance at grid cell (2° latitude × 2.5° longitude) and at local scale were compared in 3 locations (Mexico, Iowa, and Spain). The measures of skill showed significant differences between large- and local-scale downscaling performance, emphasizing the sensitivity of the downscaling models to the spatial and temporal resolution of the predictors. The downscaling model was unable to reproduce the observed monsoon precipitation in Sinaloa, where topography, sea breezes and convection play an important role. Moreover, it is known that the NCEP–NCAR reanalysis has problems to adequately

resolve the low-level moisture flux over the region (Barlow et al. 1998). On average, the model tended to underestimate (overestimate) large (small) precipitation events. However, this bias is not unique to the present analysis; some studies have shown that nested RCMs also tend to simulate too many light precipitation events compared with station data (Christensen et al. 1998, Kato et al. 2001, Trenberth et al. 2003) due to cloud parameterization problems (Kain 2004).

A detailed comparison between local precipitation at 3 stations and observed gridded precipitation (1° resolution) in Iowa showed that the there was a very good consistency between the synoptic events at local and regional scales; the downscaling model was able to adequately capture the seasonal precipitation and the phase of most precipitation events during extreme conditions at both local and regional scales. This underlines the importance of good quality observed networks and gridded data sets to validate downscaling models. The overall results of this study suggest that improved downscaling and more credible climate change scenarios could be obtained by (1) including the most relevant atmospheric predictors (e.g. circulation, humidity-related and temperature variables) at finer spatial resolution (from GCMs, reanalysis, or regional models) and (2) augmenting and improving observational networks.

Acknowledgements. We are thankful to the editor D. Leathers and 3 anonymous reviewers for their insightful comments and suggestions on earlier versions of this manuscript. This work was supported by the Water Research Commission, Pretoria, South Africa and by CONACyT-Mexico (Project # I38853-T).

LITERATURE CITED

Anderson CJ, Arritt RW, Takle ES, Pan Z and 22 others (2003) Hydrological processes in regional climate model simulations of the central United States flood of June-July 1993. J Hydrometeor 4:584–598

Barlow M, Nigam S, Berbery EH (1998) Evolution of the North American monsoon system. J Clim 11:2238–2257

Berberi EH (2001) Mesoscale moisture analysis of the North American monsoon. J Clim 14:121–137

Cavazos T (1999) Large-scale circulation anomalies conducive to extreme events and simulation of daily rainfall in northeastern Mexico and southeastern Texas. J Clim 12: 1506–1523

Cavazos T (2000) Using self-organizing maps to investigate extreme climate events: an application to wintertime precipitation in the Balkans. J Clim 13:1718–1732

Christensen B, Christensen JH, Machenhauser B, Botzet M (1998) Very high resolution regional climate simulations over Scandinavia–Present climate. J Clim 11:3204–3229

Giorgi F, Mearns LO, Shields C, Mayer L (1996) A regional model study of the importance of local versus remote controls of the 1988 drought and the 1993 flood over the central United States. J Clim 9:1150–1162

Giorgi F, Hewitson B, Christensen J, Hulme M and 5 others (2001) Regional climate information: evaluations and projections, Chapter 10. In: Climate change 2001: the scientific basis. Contribution of Working Group I to the Third Assesment Report of IPCC. Cambridge University Press, Cambridge, p 739–768

Goodman PH (1996) NevProp software. Version 3. University of Nevada, Reno, NV

Hewitson BC (1999) Deriving regional precipitation scenarios from general circulation models. Report K751/1/99, Water Research Commission, Pretoria

Hewitson BC, Crane RG (1994) Neural nets: applications in geography. Kluwer Academic Publishers, Dordrecht

Hewitson BC, Crane RG (1996) Climate downscaling: techniques and applications. Clim Res 7:85–95

Higgins RW, Yao Y, Wang XL (1997) Influence of the North American monsoon system on the U.S. precipitation regime. J Clim 10:2600–2622

Higgins RW, Shi W, Yarosh E, Joyce R (2000) Improved United States Precipitation Quality Control System and Analysis. Atlas No 7, NCEP/Climate Prediction Center, Camp Springs, MD

Kain JS (2004) The Kain-Fitsch convective parameterization: an update. J Appl Meteorol 43:170–181

Kalnay E, Kanamitsu M, Kistler R, Collins W and 18 others (1996) The NCEP/NCAR 40-year reanalysis project. Bull Am Meteorol Soc 77:437–471

Kato H, Nishizawa K, Hirakuchi H, Kadokura S, Oshima N, Giorgi F (2001) Performance of RegCM2.5/NCAR-CSM nested system for the simulation of climate change in East Asia caused by global warming. J Meteorol Soc Jpn 79: 99–121

Kidson JW, Thompson CS (1998) Comparison of statistical and model-based downscaling techniques for estimating local climate variations. J Clim 11:735–753

Legates DR, McCabe GJ (1999) Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. Water Resour Res 35:233–241

Leung LR, Mearns LO, Giorgi F, Wilby RL (2003) Regional climate research: needs and opportunities. Bull Am Meteorol Soc 84:89–95

Liebmann B, Marengo JA, Glick JD (1998) A comparison of rainfall, outgoing long-wave radiation, and divergence over the Amazon Basin. J Clim 11:2898–2909

Moore JT, Glass FH, Graves CE, Rochette SM, Singer MJ (2003) The environment of warm-season elevated thunderstorms associated with heavy rainfall over the central United States. Weather Forecasting 18:861–878

New M, Hulme M, Jones P (2000) Representing twentieth-century space-time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. J Clim 13:2217–2238

Reusch DB, Alley RB (2002) Automatic weather stations and artificial neural networks: improving the instrumental record in west Antarctica. Mon Weather Rev 130: 3037–3053

Richman MB (1986) Review article: rotation of principal components. J Climatol 6:293–335

Schubert SD, Rood RB, Pfaendtner J (1993) An assimilated data set for earth science applications. Bull Am Meteorol Soc 74:2331–2342

Trenberth KE, Olson JG (1988) An evaluation and intercomparison of global analyses from the National Meteorological Center and the European Centre for Medium-Range Weather Forecasts. Bull Am Meteorol Soc 69:1047–1057

Trenberth KE, Aiguo D, Rasmussen RM, Parsons DB (2003) The changing character of precipitation. Bull Am Meteorol Soc 84:1205–1217

Trigo RM, Palutikof JP (1999) Simulation of daily temperatures for climate change scenarios over Portugal: a neural network model approach. Clim Res 13:61–75

von Storch (1995) Inconsistencies at the interface of climate impact studies and global climate research. Meteorol Z 4: 72–80

Webster PJ, Magaña VO, Palmer TN, Shukla J, Tomas RA, Yanai M, Yasunari T (1998) Monsoons: processes, predictability, and the prospects for prediction. J Geophys Res 103(C7):14451–14510

Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. Prog Phys Geog 21:530–548

Yuval, Hsieh WW (2003) An adaptive nonlinear MOS scheme for precipitation forecasts using neural networks. Weather Forecasting 18:303–310

Zorita E, Hughes JP, Lettenmaier DP, von Storch H (1995) Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation. J Clim 8: 1023–1042

Editorial responsibility: Daniel Leathers, Newark, Delaware, USA

Submitted: February 17, 2004; Accepted: December 7, 2004
Proofs received from author(s): March 4, 2005