The Analog Method as a Simple Statistical Downscaling Technique: Comparison with More Complicated Methods

EDUARDO ZORITA AND HANS VON STORCH

Institute of Hydrophysics, GKSS Forschungszentrum, Geesthacht, Germany

(Manuscript received 28 April 1997, in final form 21 July 1998)

ABSTRACT

The derivation of local scale information from integrations of coarse-resolution general circulation models (GCM) with the help of statistical models fitted to present observations is generally referred to as statistical downscaling. In this paper a relatively simple analog method is described and applied for downscaling purposes. According to this method the large-scale circulation simulated by a GCM is associated with the local variables observed simultaneously with the most similar large-scale circulation pattern in a pool of historical observations. The similarity of the large-scale circulation patterns is defined in terms of their coordinates in the space spanned by the leading observed empirical orthogonal functions.

The method can be checked by replicating the evolution of the local variables in an independent period. Its performance for monthly and daily winter rainfall in the Iberian Peninsula is compared to more complicated techniques, each belonging to one of the broad families of existing statistical downscaling techniques: a method based on canonical correlation analysis, as representative of linear methods; a method based on classification and regression trees, as representative of a weather generator based on classification methods; and a neural network, as an example of deterministic nonlinear methods.

It is found in these applications that the analog method performs in general as well as the more complicated methods, and it can be applied to both normally and nonnormally distributed local variables. Furthermore, it produces the right level of variability of the local variable and preserves the spatial covariance between local variables. On the other hand linear multivariate methods offer a clearer physical interpretation that supports more strongly its validity in an altered climate. Classification and neural networks are generally more complicated methods and do not directly offer a physical interpretation.

1. Introduction

General Circulation Models (GCMs) are one of the most important tools in the study of climate variability and climate change. These models are state-of-the art numerical coupled models that represent several subsystems of the earth's climate (atmosphere, oceans, sea-ice, land surface processes) that are thought to be capable of simulating the large-scale state of the climate. At planetary scales, GCMs are able to simulate reliably the most important mean features of the global climate, for instance, the intertropical convergence zones, the three-dimensional atmospheric circulation cells, the jet streams, etc. With some limitations, they also simulate reasonably well essential features of the ocean circulation like the western boundary ocean currents and the conveyor belt driven by the thermohaline circulation. Some of the latest GCMs also produce atmosphere–ocean coupled variability in the Pacific basin similar to that linked to the El Niño–Southern Oscillation phenomenon. With respect to the interannual variability, it has been found that some GCMs also reproduce satisfactorily the most important patterns of variability of the atmospheric flow and of the sea surface temperature (SST) at midlatitudes. However, at finer spatial resolutions, with scales of a few grid distances, GCMs have much smaller skill (Grotch and MacCracken 1991). Many examples of the deficiencies of the global GCMs in simulating basic local climatic variables like surface–air temperature and precipitation have been presented, two of which will be mentioned here.

A detailed comparison of the regional performance of several low-resolution GCMs in the Mediterranean Basin can be found in Cubasch et al. (1996). Therein, it was concluded that the skill of these models in simulating the observed climate is much higher for near-surface air temperature than for precipitation, but that even for the former variable clear discrepancies are detected. With respect to climate change, the responses simulated by the models to a doubling of atmospheric CO_2 concentration are not univocal. In some cases two versions of the same atmospheric model coupled to a different ocean model produce temperature change pat-

Corresponding author address: Eduardo Zorita, GKSS-Forschungszentrum, Max-Planck-Str.1, 21502 Geesthacht, Germany.
E-mail: zorita@gkss.de

terns that are negatively correlated to each other. Concerning the changes in simulated precipitation, each model actually predicts patterns that are quite different from one another.

Another example is provided by Risbey and Stone (1996), who analyzed the performance of the Climate Community Model model CCM2 with T42 and T106 resolutions in the Sacramento River basin in California. They found that although the model reproduces the correct mean annual rainfall, its probability distribution differs markedly from the observations: whereas the simulated rainfall occurs mainly in the form of drizzle distributed over many rainy days, the observed rainfall is measured in much stronger precipitation events distributed over far fewer days. There now exists a more recent version of this model (CCM3), but to our knowledge it has not been checked if the regional performance has been improved.

The fact that the models do a credible job on the global scale and fail on the regional scale seems to be a contradiction. However, the global climate is to a great extent the response to the differential solar forcing, the earth rotation, and the large-scale structure of the earth's surface (land–sea distribution, topography). The regional climates, on the other hand, are the response of the global climate to regional details. Therefore, it seems reasonable that GCMs are able to simulate the global climate adequately even though none of the regional climates is simulated skillfully.

There are several reasons for the failure of the models on this regional scale: the spatial resolution provides an inadequate description of the structure of the earth's surface. The land–sea distribution is heavily smeared out and the mountains appear as broad flat hills. For spectral models the truncated representation of the topography is also a source of additional difficulties, which may be severe at the local scale (Lindberg and Broccoli 1996). A clear example is provided by the real annual cycle of precipitation in the Alps: in the northern side a summer rainfall maximum is observed, whereas some hundreds of kilometers southward a winter maximum is apparent (Fliri 1974). It is reasonable to think that it will be quite difficult for the GCMs to simulate properly those small-scale features of the actual climate and therefore the climate change assessment at those scales will have to be considered with care.

Also, the hydrodynamics of the atmosphere are nonlinear and the energy, which is fed into the system at the cyclonic scale, is cascaded through nonlinear interactions to the smallest scales. Because of the numerical truncation this cascade is interrupted and the flow to smallest scales is parameterized. These parameterizations affect the smallest resolved scales most strongly.

Related also to the model resolution is the problem of the representation of the subgrid-scale processes, such as cloud formation, rainfall, infiltration, evaporation, runoff, etc. These have to be parameterized. Obviously, with increasing model resolution more and more processes can be explicitly represented, but many of them occur at too small scales to be realistically modeled in the present and probably next generation of climate models.

These processes are calculated by means of bulk formulas, the parameters of which may not have been fitted for the region of interest. These parameterizations may produce additional errors in the GCM simulations. There are indications (Risbey and Stone 1996; Machenhauer et al. 1996) that this may be the most important source of error of the GCMs, perhaps even more than its inadequate resolution.

However, these subgrid processes are actually those with the greatest ecological or societal impact, since they strongly affect the local climate at the scales of the human and ecological environment. Therefore, there is a broad consensus about the need to simulate the subgrid processes and the local climates properly, perhaps beyond the capabilities of the current GCMs.

a. Strategies to bridge the scale gap

The efforts to improve GCM simulations have been aimed in two directions. In the last years, with increasing computer power, there has been a clear tendency to finer and finer GCM horizontal resolutions. For instance, while in 1990 a T21 resolution (about $5.6^\circ \times 5.6^\circ$) for the atmospheric submodel was considered as state-of-the-art, some of the last integrations with atmospheric models have been carried out with a resolution of T106 (about 1.2°). This resolution is however quite costly, and for climate change estimation the applications so far have been restricted to the ''time slice modus'' (Cubasch et al. 1996). In this modus the atmospheric high-resolution model is forced by the mean boundary conditions simulated in a low-resolution atmosphere–ocean general circulation model. On the other hand the use of so-called limited area models (LAMs) (Giorgi and Mearns 1991) is becoming more frequent and is being also applied to the ocean component of the climate model (Kauker and Oberhuber 1998). These LAMs are sophisticated atmospheric (or oceanic) models of a limited geographical area (of the order of 10^7 km^2) with a resolution of the order of 20–50 km, that use the large-scale fields simulated by the GCMs as boundary conditions, but that take the regional characteristics, such as topography, into account. An increased resolution in the region of interest may improve important aspects of the regional climate simulation. For instance, orographically induced precipitation and cyclonic activity at midlatitudes is better reproduced (Machenauer et al. 1996). It is expected that an increased resolution may lead to better regional simulations (Mearns et al. 1995). However, some problems still remain. For instance, the LAMs developed at the UK Metereological Office (UKMO) and in Meteo-France (based on the UKMO GCM and the ARPAGE GCM, respectively) show systematic errors that are not solved by increasing the res-

olution. These are probably associated with the parameterizations of subgrid processes, which are taken over from the parent GCMs, and with the large-scale errors of the coarse-resolution GCMs themselves (Machenhauer et al. 1996). Therefore, there seems to exist also a need not only for finer resolutions, but also for better subgrid parameterizations (Risbey and Stone 1996).

b. Statistical downscaling

The alternative approach to overcome the scale mismatch between the skill of climate simulations and the needs of ecosystem and sector models is statistical downscaling (Giorgi and Mearns 1991; Hewitson and Crane 1992). This technique is becoming quite popular, due to relative simplicity and lower costs compared to the use of LAMs. A recent comparison between some statistical downscaling methods is given by Wilby and Wigley (1997).

Essentially the idea of the statistical downscaling consists in using the observed relationships between the large-scale circulation and the local climates to set up statistical models that could translate anomalies of the large-scale flow into anomalies of some local climate variable (von Storch 1995).

The statistical downscaling methods found in the literature are also growing in technical complexity. It may seem reasonable to check if the results produced by relatively complicated methods are clearly superior to more simple and more flexible methods. It seems that most statistical downscaling methods can be subjectively classified under the headings of linear methods, weather-type classification methods, or nonlinear deterministic methods. In this paper we have chosen one method belonging to each of these broad families and compared their performance with a simpler analog method. For this purpose the statistical models based on these methods have been fitted to observed historical large-scale data and their output is compared with the observations in an independent period. Within its own family each of these methods may be considered to be technically quite complex. In this sense, due to its simplicity, the analog method represents a benchmark that any other method should be able to surpass.

An important assumption that underlies the statistical approach to climate impact assessment is that the link between the large-scale circulation and the local climate remains unchanged in an altered climate, which is by no means guaranteed. However, if the time series used to tune the statistical model are long enough (ideally of the order of several decades), it is reasonable to assume that they contain many different situations, including those that will be stronger or more probable in an altered climate. If these situations are important for the local climate, the statistical model should be able to identify them in the historical observations and estimate with some skill the probable impact on the local climate. This assertion is only valid if the range of variations of the large-scale variable in the new climate lies roughly within the natural variability of the present climate, which is the information used by the statistical model. However, the fact that the statistical model is able to reproduce reasonably the variability in the past increases the level of confidence in the model but does not strictly imply that it can be used for future conditions, since the statistical relationship may not hold any more. If the range of variations in the new climate is larger than the observed natural variability of the observed climate, the estimation via statistical downscaling may still be useful but it should be considered with care. This drawback is in some sense also present in climate change estimation with GCM experiments, since these models contain many parameterizations that in principle are only valid for the present climate. However, the functional form of these parameterizations schemes are in many cases based on sensible physical reasoning, so that they hopefully will remain more or less valid in an altered climate.

To avoid to some degree the uncertainties of GCM and LAM simulations, the ability of these models to simulate past climates should be tested. The counterpart requirement for statistical downscaling models is that they should be able to reproduce the historical evolution of the local variables when they are driven by the observed large-scale circulation of the past. In the context of climate change, an important point is the replication of the low-frequency (decadal and longer) natural variability of the local variables of interest, that may include mean values but also probability of extreme events. The low-frequency natural variations in the form of trends or oscillations can be considered as *natural* climate changes and a good statistical model should be able to reproduce them.

The paper is divided into seven sections. Section 2 contains the description of the data used in this study. In section 3 the analog method for downscaling is introduced. In the following sections the proposed families of methods are described: linear methods (section 4), classification methods (section 5), and neural networks (section 6), and the results are in each case compared with the output produced by the analog method. The problem of the level of the variability of the simulated local variable is discussed in section 7. The paper is closed by some concluding remarks in section 8.

2. Data description

The daily sea level pressure (SLP) data that represent the observations originate in the National Centers for Environmental Prediction (NCEP, previously the National Meteorological Center) analysis provided by the National Center for Atmospheric Research (Boulder, Colorado). They have been used with a resolution of 5° $\times$ 5°. The daily rainfall data in Spain were kindly supplied by the Instituto Nacional de Meteorología (Madrid, Spain). Monthly rainfall records of 92 stations (mainly in Spain and Portugal, but some of them also

in Southern France, Morocco, and Algeria) were provided by the Universidad Complutense (Madrid, Spain).

3. The analog method

Perhaps the simplest downscaling scheme is the analog method. This method has been essentially applied in the field of weather forecasting (Lorenz 1969; Kruizinga and Murphy 1983), and in short-term climate prediction (Barnett and Preisendorfer 1978; van den Dool 1994). For downscaling purposes the method has seldom been applied (Zorita et al. 1995; Cubasch et al. 1996; Biau et al. 1999), although its idea is simple. The large-scale atmospheric circulation simulated by a GCM is compared to each one of the historical observations and the most similar, in a sense that has to be still defined, is chosen as its analog. The simultaneously observed local weather is then associated to the simulated large-scale pattern.

A relevant problem associated with this method is the need for sufficiently long observations, so that a reasonable analog of the large-scale circulation always can be found. Due to the number of degrees of freedom of the large-scale atmospheric circulation, it has been pointed out (van den Dool 1994) that on a global basis and for prediction purposes several thousand years would be needed. However, many of these degrees of freedom represent just background noise that can be previously filtered out, for instance, by a standard empirical orthogonal function (EOF) analysis, and in downscaling applications the area of interest is not global but normally covers a continent or an ocean basin. Furthermore, for downscaling purposes the analog method is not used in a prediction scheme but rather as a mean to specify a local state coherent with a simultaneous large-scale state. All this reduces the number of degrees of freedom of the problem, and normally useful analogs are indeed found for most of the downscaling applications that we have explored. In this slightly modified form, the anomalies of the atmospheric circulation, for instance, represented by the anomalies of the SLP field f, are described by the few leading EOF patterns:

$$f(i, t) = \sum_{k=1}^{n} x_k(t) g_k(i) + \epsilon(t), \tag{1}$$

where i is a gridpoint index, t is the time, g_k is the kth EOF pattern, $x_k(t)$ is the amplitude of this pattern at time t, n is the number of EOF patterns retained, and ϵ is the part of the variability not described by the leading n patterns, assumed to be small. The analogs are searched only within the space spanned by these n EOF patterns.

Consider an atmospheric anomaly pattern $\Delta f(\mathbf{r})$ (e.g., the long-term mean difference between a doubled CO_2 and control simulation with a GCM or the anomaly between an observed circulation and the long-term mean). This pattern may have coordinates z_k in this EOF space. Its analog is defined as the circulation at time step t that minimizes the distance in EOF space:

$$\sum_{k=1}^{n} [z_k - x_k(t)]^2. \tag{2}$$

The method can be generalized by introducing different weights d_k on the EOF coordinates:

$$\sum_{k=1}^{n} \{d_k [z_k - x_k(t)]^2\}. \tag{3}$$

The weights d_k may be optimized, so that the normalized squared deviation E between the simulated and observed local variable is as small as possible:

$$E = \sum_{j=1}^{N} \sum_{t=1}^{T} [l^o(j, t) - l^s(j, t)]^2 / \sigma_j^2, \tag{4}$$

where $l^o(j, t)$ and $l^s(j, t)$ are the observed and generated local variable at station j, respectively, N is the number of stations, T is the number of time steps, and σ_j^2 is the observed variance at station j.

The optimization procedure can be considerably difficult, since the function E has in general a complicated topography. We will restrict here to the simpler method contained in (2).

This method is illustrated in detail by the following example. We are interested in the December–February (DJF) precipitation over the Iberian Peninsula, given at a number of irregularly distributed meteorological stations. It is assumed that this regional variable is controlled to a great extent by the atmospheric variability in the European–North Atlantic sector. The large-scale variable will be the SLP field, which for this purpose offers some advantages compared to geopotential height data. First, there exist long homogeneous time series of this variable that allow setting up the statistical model in some period and checking it in an independent dataset. Second, in climate change GCM experiments, geopotential heights tend to be much more affected by the global warming, but these changes may be related to changes in the mean atmospheric density and not necessarily to changes of the atmospheric circulation. Therefore, using the geopotential heights as large-scale field may have the effect of including a signal that is not physically related to the atmospheric circulation and therefore to the local variable. The SLP, on the other hand, is much less affected. In this example the leading five EOFs of the SLP field have been retained. They describe 85% (75%) of the monthly (daily) variability.

To check the quality of the analog method we try to reconstruct the time series of the winter (DJF) average precipitation in the period 1901–89 by looking for analogs of the atmospheric circulation in the period 1951–89. In the overlapping period the analog for a particular month is searched in the whole dataset available, but always in a season different from the target pattern.

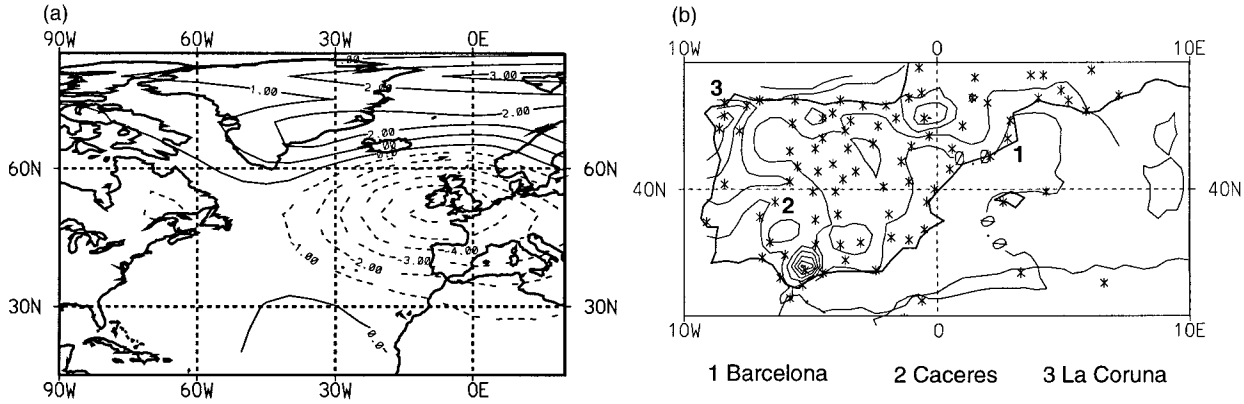

FIG. 1. The first pair of canonical correlation pattern of observed DJF Iberian rainfall (contour interval, 25 mm month^{-1}) and simultaneous SLP anomalies (hPa) in the North Atlantic sector. The correlation between the time coefficients is 0.86. The rainfall pattern explains 50% of the variance. The stars indicate the station positions.

Since the interannual autocorrelation of the atmosphere is negligible, this procedure should amount to searching the analogs in an independent dataset.

To reduce the number of atmospheric degrees of freedom, the EOFs of the winter SLP field are calculated. Then, the winter SLP anomalies are projected onto the leading EOFs. The results obtained by the analog method are subsequently presented in the following sections, compared directly with those achieved by the more complicated techniques.

4. Linear methods

a. Normally distributed variables

Linear models are perhaps the most popular in the downscaling context. They apply the huge battery of already existing linear methods, for instance, from the simple linear regression up to multivariate singular value decomposition, to the concept of *teleconnection.* For instance, one of the first physical teleconnections identified in climate research was the link between the North Atlantic Oscillation (NAO) and the surface–air temperature in Scandinavia. This link can be applied to downscaling by setting up a linear regression between the anomalies of the NAO index (the SLP difference between Azores and Iceland) and the anomalies of the temperature in a Scandinavian station. The changes in the strength of the NAO in a future climate can be then translated to changes in local temperature by means of linear regression, apart from the global temperature increase.

The general idea of the linear methods is the same as in the above example, namely, to link anomalies of the large-scale circulation to anomalies of the local climate. However, the technical complexity of the method can be considerably increased, as shown in this section also dealing with the relationship between the North Atlantic SLP and winter Iberian rainfall. This problem has been already studied by von Storch et al. (1993). The results shown here have been, however, obtained with a much more dense station network.

The linear model has been set up by means of canonical correlation analysis (CCA). Given two random vector time series **X** and **Y**, in this case the monthly rainfall in winter (DJF) at the Iberian stations and the monthly SLP field defined on a grid over the North Atlantic, CCA identifies pairs of patterns whose time evolution is optimally correlated (Barnett and Preisendorfer 1978; Bretherton et al. 1992). These spatial patterns are the eigenvectors of the matrices:

$$\mathbf{M}_X = \mathbf{C}_{XX}^{-1}\mathbf{C}_{XY}\mathbf{C}_{YY}^{-1}\mathbf{C}_{YX}, \qquad \mathbf{M}_Y = \mathbf{C}_{YY}^{-1}\mathbf{C}_{YX}\mathbf{C}_{XX}^{-1}\mathbf{C}_{XY}, \quad (5)$$

where the $\mathbf{C}$'s are the respective cross-covariance matrices of the SLP (X) and local rainfall (Y) time series. It can be shown that both matrices $\mathbf{M}_X$ and $\mathbf{M}_Y$ have the same positive eigenvalues c_k^2, which are the squared correlation coefficients between the time series associated with the kth eigenvector of $\mathbf{M}_X$ and $\mathbf{M}_Y$. Otherwise the time series are pairwise uncorrelated. This does not necessarily mean that the processes represented by the different CCA patterns are physically independent, but they normally represent, at least in a first approximation, different aspects of the variability. For more technical details about CCA and other similar techniques the reader is referred to Bretherton et al. (1992). Prior to the CCA the North Atlantic SLP and Iberian rainfall are filtered by standard EOF analysis to reduce the number of degrees of freedom of both fields. We will compare the results of the CCA with the analog method in the previous section, and therefore the same number of EOFs for the SLP field are retained, so that the same large-scale information enters both methods. For Iberian rainfall the first two EOFs, that describe about 80% of the total variance, are retained for the CCA.

The CCA performed in the period 1951–90 identifies one dominant pair of patterns (Fig. 1). The rainfall pattern is positive at all stations, with highest values near the Atlantic coast and decreasing values toward the

Mediterranean. The associated SLP pattern consists of a low-pressure cell located over the British Isles. The canonical pair of patterns may be physically explained by advection by the large-scale circulation. For instance, the SLP pattern describes variations of the advection of air masses of Atlantic origin to the Iberian Peninsula.

The result of the CCA provides a method for estimating a regional rainfall anomaly $R(j, t)$ at a set of stations from a given large-scale pressure anomaly field $f(i, t)$ in a consistent way. Mathematically, this is accomplished in two steps: if the pairs of CCA patterns for the large-scale and local variable are denoted by p_k and q_k, respectively, the first step is to calculate the amplitude a_k with which the kth large-scale CCA pattern p_k appears in the new SLP anomalies $f(i, t)$. This is achieved by minimizing the sum of squares:

$$E = \sum_{t,i} \left[f(i, t) - \sum_k a_k(t) p_k(i) \right]^2. \quad (6)$$

Equating to zero the derivatives of E with respect to a_k leads to a linear system of equations that can be solved by standard methods. The estimated rainfall anomalies at station j associated with $f(i, t)$ is just the sum of the estimated amplitudes of the local CCA patterns:

$$R(j, t) = \sum_k c_k a_k(t) q_k(j), \quad (7)$$

where c_k are the canonical correlations. This procedure is, of course, capable of describing only the part of rainfall variance that can be traced to the atmospheric circulation. The implicit assumption is that the observed intermonthly SLP–rainfall relationship can be extrapolated to longer timescales, which is reasonable as long as the climate variations are considered small.

The reliability of the suggested statistical relationship is tested by reconstructing the patterns of Iberian monthly rainfall in winter from the beginning of this century. The North Atlantic pressure field has undergone significant changes during this century (Hense et al. 1990; Shabbar et al. 1990) and this variability should have had an effect on rainfall. Note that the data prior to 1951 have not been used to perform the CCA and, thus, represent independent samples.

The area-averaged winter rainfall, obtained indirectly from the pressure fields and from the in situ meteorological observations are displayed in Fig. 2. The two time series, both smoothed with a 5-yr running mean filter, vary coherently, particularly at long timescales. Interestingly, the method is able to reproduce the low-frequency oscillations with a timescale of about 20 yr and the positive rainfall trend. This confirms that winter rainfall in this region is to a large extent controlled by the large-scale circulation, and that the trend and the interdecadal variations are real. This follows since the two datasets, rainfall and pressure, are from independent sources. However, the method overestimates a negative rainfall anomaly around 1975. This mismatch is apparent for both methods, CCA and analog. The reasons for this disagreement are not known. It is possible that in these years other factors than SLP alone were responsible for the rainfall anomalies. Errors in either the SLP or rainfall data could be another reason. This is, however, improbable since both have undergone detailed quality checks.

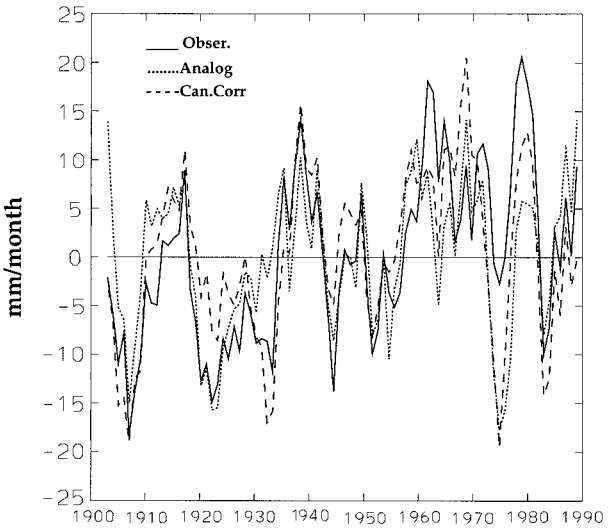

FIG. 2. Five-year-running mean time series of area-averaged DJF rainfall anomalies (mm month^{-1}) as derived from station data in the Iberian Peninsula (solid line); derived from the North Atlantic SLP pressure field: analog method (dotted line) and canonical correlation analysis (CCA, dashed line).

The agreement with the observations is otherwise good for both methods. The correlation between the replicated and observed rainfall is slightly better for the CCA method (actually, this method optimizes the time correlation), and this feature is more apparent in the unfiltered time series (not shown). On the other hand the time series reproduced with the analog method have a larger variance, closer to the observations. This is a common feature of all linear models. This occurs because the linear models filter out the noise inherent in the observations that is not related to the atmospheric circulation. This problem is discussed in more detail in section 7.

A caveat of the linear methods is that they cannot be used directly when the local variables are not normally distributed. Since the variability of the large-scale atmospheric circulation is usually normally distributed (at least it is quite difficult to detect deviations from normality), the output of any linear method is bound to be also normally distributed. There are, however, local variables that strongly deviate from normality, the most important example being perhaps daily precipitation. A solution to this problem may be to transform the local variable so that the distribution of the transformed variable is approximately Gaussian. This is always possible for a random variable with continuous probability density, but it is problematic for daily rainfall, since it gen-

erally contains considerable probability mass at zero values of rainfall. Perhaps more serious is the fact that such a variable transformation would introduce biases in the back-transformed means and variances (Cohn et al. 1989). To correct for these biases, assumptions about the normality of the noise have to be made, which may not be always valid.

In the next subsection we present another possibility to apply linear methods even if the local variables are nonnormally distributed.

b. *Linear methods applied to statistics of nonnormally distributed local variables*

Some climate impact models do not need time series of local forcing functions that are consistent with the simultaneous large-scale fields. Instead, just a few statistical properties of those time series suffice: for instance, for coast protection measures it would be sufficient to have information about changes in the probability of wind speed and direction. The idea is then to establish relationships between the large-scale fields and, for instance, the probability distribution of local wind or the probability distribution of the length of dry periods, etc. The interannual variability of these probabilities is expected to be more normally distributed, since they are estimated usually by suitable time averages. In this case a linear technique may be successful.

In this example we focus on the probability distribution of daily rainfall and the probability of the storm interarrival times of the station Cáceres in the Iberian Peninsula (30.5°N, 6.3°W) in winter (December–February, hereafter DJF). The probability of the storm interarrival time $p_{\text{sit}}(\tau)$ is defined as the probability that the length of dry period exceeds τ days. Note that in general $p_{\text{sit}}(\tau = 1) \neq 1$.

We choose in this case the *seasonally averaged* winter SLP fields as large-scale variables. Two CCA calculations have been performed: one in which the local field for the CCA is the winter daily rainfall distribution at this station; and another one in which the local field is the probability of the length of storm interarrival times. These probability distributions are determined for each winter season from the observations. It is clear, however, that the estimation of the long tails of the probability distributions will be poorer.

The CCA is otherwise analogous to the one in the previous subsection. Since the daily dataset is shorter the analysis is performed in the period 1965–85. The results of both calculations are shown in Fig. 3. The SLP canonical pattern looks similar in both cases to the result of the CCA between SLP and monthly rainfall. This supports the validity of this approach. The canonical patterns of rainfall distribution indicate that the increase of mean rainfall that occurs when the SLP in the North Atlantic is lower than normal (Fig. 1) is mainly caused by an increase of days with weak rainfall at the cost of dry days, whereas the number of days with heavy rainfall (>3 mm day^{-1}) remains almost unchanged. With respect to the storm interarrival times the canonical patterns indicate that the low-pressure cell over the North Atlantic is associated with a reduction of the probability of moderate dry spells (<5 days), where the change in the probability of longer dry spells is smaller.

To validate the statistical model the distribution of interarrival times for the period 1942–89 may be estimated based on the linear model and the SLP data in this period. We show in Fig. 4 the evolution of the 5-day storm-interarrival time (i.e., the probability that a dry period would last longer than 5 days), both from the observations and from the estimation with the CCA method. The analog method has been also applied to the same dataset. In this case the analogs for the whole period are searched *on a daily basis* in the winter months between 1965 and 1985. For the training period the analogs were again searched in a season different from the target day. The agreement between simulations and observations is clear, especially at low frequencies. The time series of the 90% quantile seems to be negatively correlated with the time series of the 5-day interarrival time, also at low frequencies, indicating that periods with higher (lower) rainfall are associated with shorter (longer) dry spans. This relationship is also captured by both statistical methods.

Regarding the relative performance of the linear and analog methods in this case, the same considerations may apply as for the previous subsection: the temporal fluctuations are better captured by the CCA method, but the level of replicated variability is lower than in the analog method.

An important point regarding the multivariate linear methods, such as the one based on CCA, is that it yields spatial patterns that normally allow for a physical interpretation, thus increasing the degree of confidence of the real existence of the link between large-scale circulation and the local variables. This aspect is quite clear in this example (Figs. 4 and 5). Such an interpretation is normally much more difficult for the other methods.

5. Classification methods

The general principle underlying the classification methods is also simple, although the practical implementation can become quite complicated. A classification scheme of the atmospheric circulation in the area of interest is developed and a pool of historical observations is distributed into the defined classes. The classification criteria are then applied to atmospheric circulations simulated by a GCM, so that each circulation can be classified as belonging to one of the classes. To each observed circulation there exists a simultaneous observation of the local variable. The value of the local variable to be associated with the simulated large-scale circulation can be chosen as either the average of all regional observations simultaneous to the elements of that class, or only the regional observations simulta-

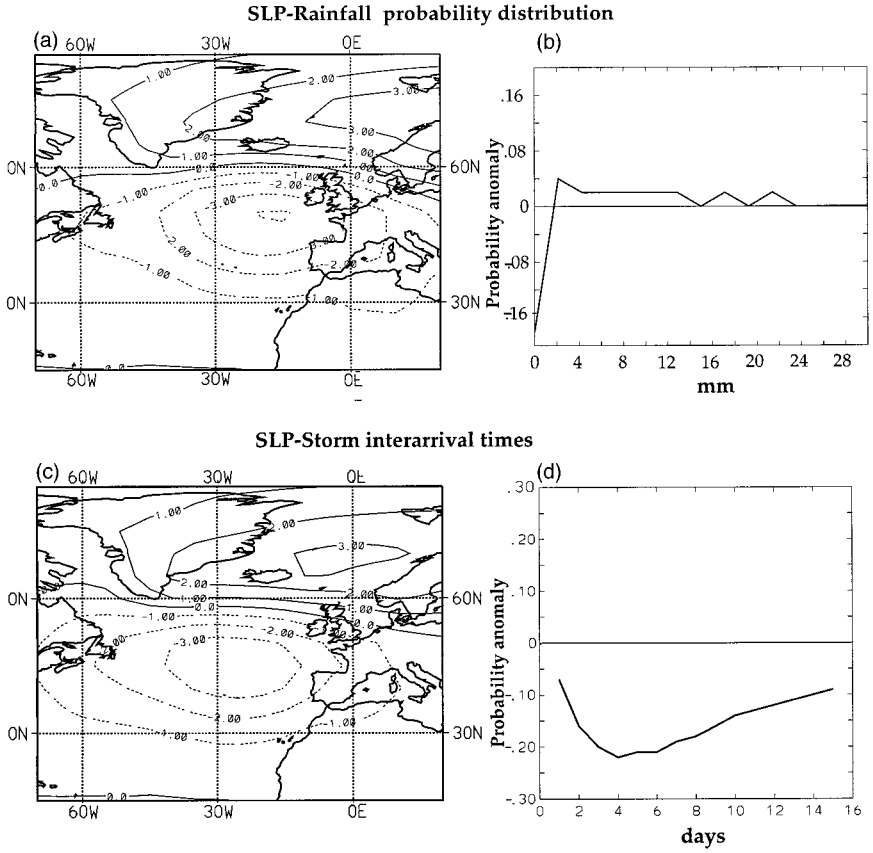

FIG. 3. Results of the two CCAs of seasonal North Atlantic SLP in wintertime and daily rainfall amounts in Cáceres, Spain: (a) and (b) anomalies of SLP (hPa) and of daily rainfall probability distribution (the correlation between the time coefficients is 0.87, the probability pattern explains 30% of the variance); (c) and (d) anomalies SLP (hPa) and of the probability of storm interarrival times (the correlation between the time coefficients is 0.62, the probability pattern explains 38% of the variance).

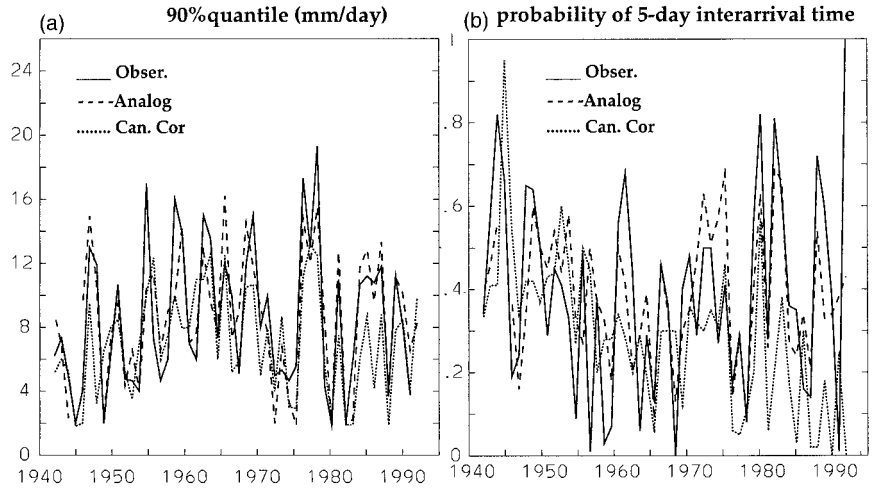

FIG. 4. Time series of 90% quantile of daily rainfall (mm) and probability of 5-day storm interarrival time in Cáceres, Spain in wintertime, as observed and reconstructed from the winter North Atlantic SLP field, by the analog method and the CCA method.

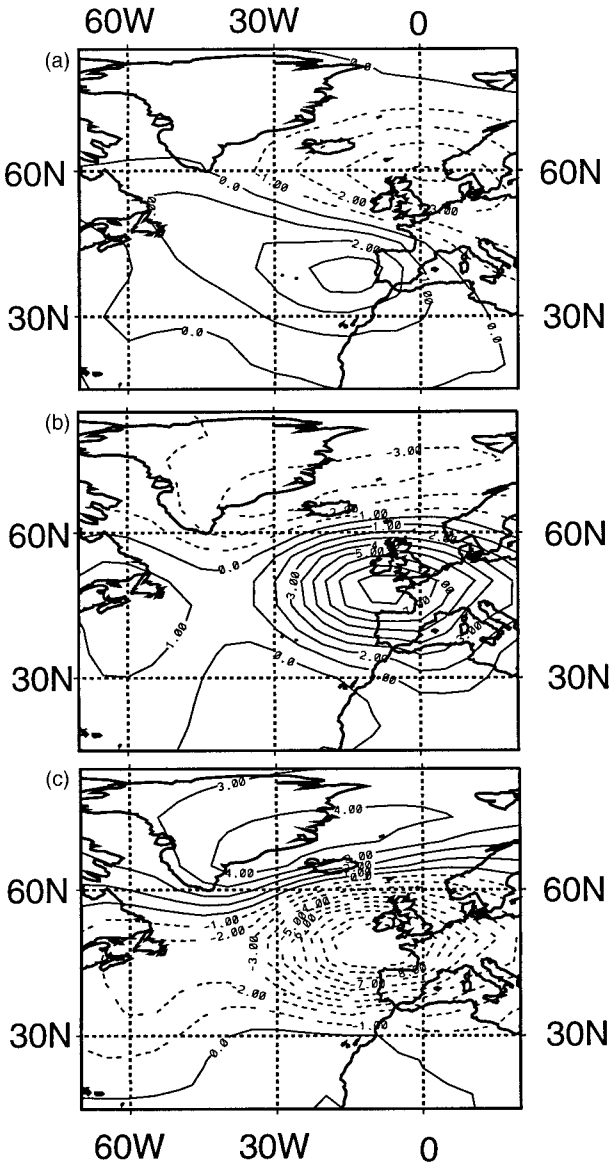

FIG. 5. Average daily SLP anomaly field (hPa) for the three weather states identified by CART analysis for three stations in the Iberian Peninsula: La Coruña (northwest), Barcelona (northeast), and Cáceres (west). (a) State 1 is associated with rain days in La Coruña and dry days in the other two; (b) state 2 is associated with dry days in all three stations; (c) state 3 is associated with rain days in La Coruña and Cáceres and dry days in Barcelona.

neous to *one* element of the class, selected at random. Which of both strategies is best suited depends on the particular problem. For instance, if one is interested in simulating local daily rainfall, averaging over all the elements of a class will in general lead to an underestimation of the local rainfall variance, of extreme events, and in general, to a local rainfall probability distribution different from the observed. On the other hand, averaging over several elements of the class will filter out more effectively measurement errors at that station.

When validating the statistical model one should be aware that a purely random scheme, that is, choosing at random 1 day from sufficiently long historical observations will very probably produce the right mean, variance, and probability distribution of local rainfall. On the other hand the random choice within a class makes it not advisable to compare directly one realization to the actually observed values. Therefore, for these two reasons the assessment of the performance of a classification method should be based generally on other quantities involving the time structure, such as persistence.

A practical problem is to define a method for classifying the large-scale patterns. There are many classification methods. However, it should be pointed out that all classification schemes are to some degree subjective, although some of them, once defined, allow for a programmable classification scheme of circulation patterns. In the most objective schemes only the area over which the circulation data are taken and the number of resulting classes has to be subjectively fixed at some stage of the model design.

The typical example of subjective classification schemes is the traditional Grosswetterlagen classification of the German Weather Service for the Western Europe–North Atlantic sector. This classification has been used for downscaling purposes (Bárdossy and Plate 1992). Weather typing procedures have been developed by many national weather services for their particular regions based on the local expertise. Other well-known subjective schemes are the Schuepp scheme for Switzerland (Schuepp 1953) and the Lamb classification for the British Isles daily weather patterns (Lamb 1972). There exists a programmable procedure (Jones et al. 1993) that takes into account the empirical rules first proposed by Lamb.

More complicated technically is the design of an objective classification scheme. In this respect several types of schemes can be distinguished: classifications that depend only on the large-scale circulation data, classifications that depend on the local variable, and schemes that use information from both data. A typical example of the first and second group is traditional cluster analysis of atmospheric circulation patterns (Cheng and Wallace 1993). In the first case the classes of the large-scale circulation are given directly by the analysis. In the second case the elements of the large-scale circulation class are defined as the simultaneous circulation to each element of the local climate class. This second method has the advantage that the resulting large-scale classes should really correspond to different local situations, which is not necessarily the case in a classification based only on the large-scale patterns.

An example of a quite complicated classification based simultaneously on the large-scale circulation *and* the local variable is represented by classification and regression trees (CART; Breimann et al. 1984) analysis. This method has been mainly applied to the simulation

of local daily rainfall (Hughes et al. 1993; Schnur and Lettenmaier 1998). However, due to its extensive needs of computer time, it has been applied to a limited number of stations and it has been assumed that rainfall is just a two-outcome process, wet or dry. The CART analysis searches recursively for a binary decision tree, whose decision nodes are based on the values of the large-scale atmospheric variables at some key locations, or the values of key large-scale atmospheric indices. Each terminal node of the tree represents a weather state. A weather state resulting from the CART analysis is such that the joint probability distribution of the local variable (including all stations) based on the days belonging to that weather state is, in some sense, maximally different from the probability based on the other weather states.

Since linear models can be normally applied to monthly rainfall, classification methods are preferred for daily rainfall. In this section we present the results of a CART analysis of daily winter rainfall in three stations in the Iberian Peninsula: Barcelona, La Coruña, and Cáceres (see Fig. 1 for the location of these stations) and the daily SLP field over the North Atlantic in wintertime in the period 1970–75. To reduce the dimensionality of the problem the input variables for the CART analysis are again the leading five EOFs of the SLP field. CART is able to identify three ''atmospheric states'' and each daily SLP field can be classified as belonging to one of those states according to the values of its coordinates in the five-dimensional EOF space. The average SLP anomaly field for each state is represented in Fig. 5. State 1 is mainly connected to rain days in one station (La Coruña) and dry days in the other two; state 2 is most probably associated with dry days in all three stations and state 3 with rain days in La Coruña and Cáceres(both western stations) and dry days in Barcelona (located at the Mediterranean sea).

New daily SLP anomaly fields, observed in another period or simulated by a CGM, can be classified into one of these three states and the local rainfall attached to this circulation can be chosen at random from all the days belonging to that particular weather state. In this way daily time series of rainfall consistent with the concurrent SLP field can be produced. To compare the performance of this method to the analog method, the probability of the length of dry periods, also known as storm interarrival times, has been chosen (Fig. 6) in the period 1978–83. It can be seen that in general the CART method underestimates the length of the storm interarrival times as compared with those directly estimated from the observations. The results obtained with the analog method tend to be closer to observations than CART.

The weather states defined by any classification scheme can be also used to validate the performance of a GCM and eventually investigate the reason why a particular GCM may not be simulating properly the local climate in a certain region. Notwithstanding the fact that the weather classes are the result of a more or less subjective definition, and if these classes are not defined in a very restrictive manner, the GCMs in general are grossly able to simulate their probabilities of occurrence, although for some weather states there may be large deviations from observations (Hulme et al. 1993). Much more problematic is the simulation of the local climate associated with the individual large-scale classes, since the local climate is, for instance, quite sensitive to the exact location of anticyclones or deviations of storm tracks from the long-term mean. The best performance is usually found for surface temperature in the winter season, since this local variable is to a great extent determined by large-scale advection. The associated summer temperature and rainfall in both seasons, however, are normally not so satisfactorily reproduced.

If the large-scale circulation is classified on a daily basis, an important aspect of the validation is the dynamical behavior of these classes, for instance, their mean life times. Since the weather typing is not usually defined from a dynamical point of view, this aspect of the validation can give more objective information on the ability of the models to simulate the regional weather. The GCMs do show some skill in reproducing the transition probabilities of the weather types (Zorita et al. 1995), but their performance still has to be considerably improved if realistic local weather time series are to be directly used in climate impact studies. Therefore, some of the deficiencies in the simulation of local weather scenarios lie clearly in the deficiencies of the GCMs to simulate the evolution of the large-scale atmospheric patterns. However, for daily rainfall there exists a more serious problem, which remains still unsolved. Downscaling procedures based on classification schemes and using only the observed large-scale circulation produce daily rainfall time series with less persistence than in the observations, that is, the observed time-clustering of precipitation is not replicated by the downscaling techniques (Hughes et al. 1993). This problem can be partially, but not completely, reduced by simulating daily rainfall not only conditional on the daily weather class, but also on the evolution in the previous few days (Zorita et al. 1995). Therefore, there seems to exist in the real rainfall process some kind of local persistence that presumably cannot be taken into account by large-scale processes alone (Hughes et al. 1993). This fact complicates the statistical downscaling approach for daily rainfall, since it would be desirable not to include information over the local processes themselves, as long as they are potentially not well simulated by GCMs. The solution of including information about generated rainfall in the previous days would be essentially equivalent to using information of the atmospheric circulation in the previous days.

6. Neural networks

Neural networks have found in the last years a wide range of applications. A quite complete review about

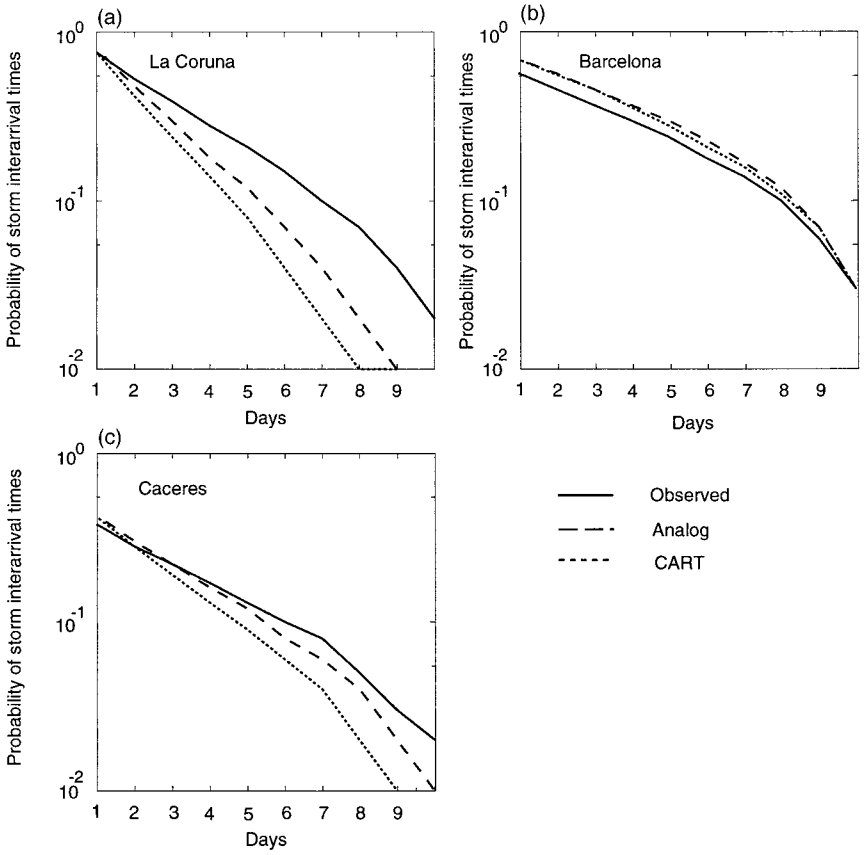

FIG. 6. Probability of the length of the storm-interarrival times in winter in three Iberian stations (La Coruña, Barcelona, Cáceres), as observed (1951–90) (continous line) and simulated with the analog method (dotted lines) and the CART method (dashed lines) using the historical North Atlantic SLP observations in the same period.

this subject can be found in Lau and Widow (1990). For applications in meteorology the reader is referred to Elsner and Tsonis (1992) and the references therein. In climatology recent applications comprise the El Niño–Southern Oscillation phenomenon (Grieger and Latif 1994; Tangang et al. 1997) and Indian monsoon rainfall (Navone and Ceccatto 1994). Neural networks have a great potential in many contexts, but they have been applied to the downscaling problem only in a few cases (Hewitson and Crane 1992; Hewitson 1996).

Only the basic concepts necessary to follow this section will be given here. Very briefly, a neural network is an algorithm that transforms an input vector $\mathbf{x}^{\text{in}}$ into an output vector $\mathbf{x}^{\text{out}}$ by stepwise nonlinear transformations, as illustrated in Fig. 7. Each transformation is carried out in two steps. In a first step each component of the input vector $\mathbf{x}_i^{\text{in}}$ is separately transformed by a nonlinear function f:

$$\mathbf{x}_i^* = f(\mathbf{x}_i^{\text{in}}). \tag{8}$$

In a second step a linear transformation is applied to $\mathbf{x}^*$:

$$\mathbf{x}_j^1 = \sum_i w_{ij}^1 \mathbf{x}_i^*. \tag{9}$$

The resulting vector $\mathbf{x}^1$ is, in turn, the input for the next nonlinear transformation. It is useful to think as if this two-step process is performed by one layer of ''neurons'' (Fig. 7), the whole neural network containing several layers. Finally, the output vector $\mathbf{x}^{\text{out}}$ is the result of the operation on the last layer.

A complete model can be built when the parameters w_{ij}^k are known. This can be achieved by fitting them with a set of known inputs $\mathbf{x}_i^{\text{in}}(t)$ and outputs $\mathbf{y}_i(t)$, by minimizing the squared deviations:

$$\sum_{i=1}^{N} \sum_{t=1}^{T} [\mathbf{x}_i^{\text{out}}(t) - y_i(t)]^2, \tag{10}$$

where N is the number of stations and T the length of the time series.

We have used in this example a neural network of three layers to construct a nonlinear model that links the daily SLP anomalies (as predictor) and daily rainfall amounts (as predictand). The input vector is composed

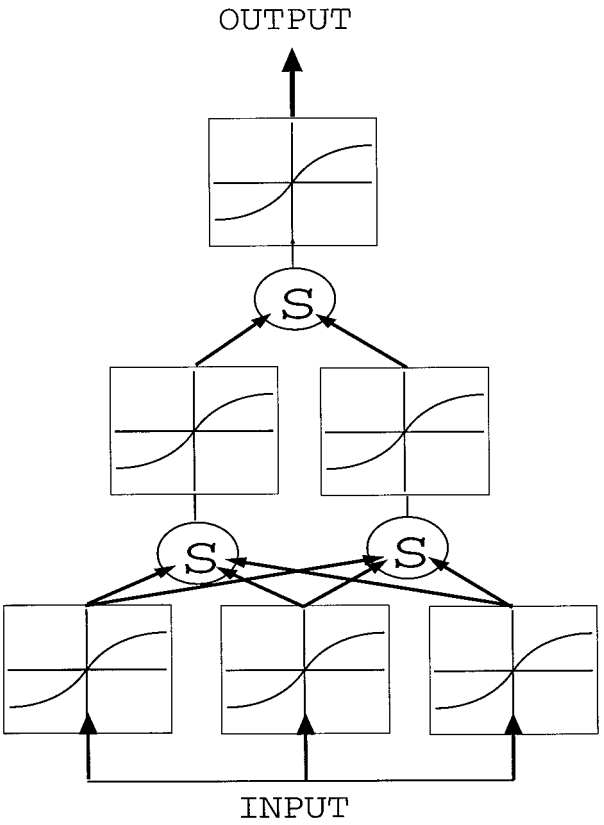

FIG. 7. Schematic structure of an algorithmic neural network.

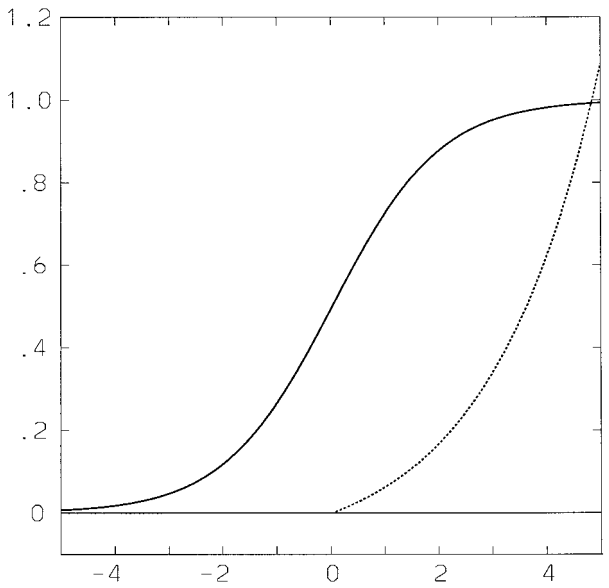

FIG. 8. Two possible nonlinear filter functions that relate the input and output of a neuron in a neural net. Bold line is a classical sigmoidal type; the dashed line represents Eq. (12), more suitable for the simulation of daily rainfall. In this plot $x_c = 0$.

by the principal components associated with the five leading EOFs of the daily SLP field.

The number of elements in the intermediate layer (sometimes called the hidden layer) is somewhat arbitrary but constrained by the following considerations. First, as in any statistical model the number of parameters in w_{ij}^k should be kept to a minimum to avoid overfitting of the noise in the training period. Otherwise the skill of the network falls abruptly when it is applied to a set of predictors in an independent dataset. To understand the second consideration more easily, consider for the moment a linear network (with a linear filter function f) with just a single element in the hidden layer and assume that the desired output time series are normalized ($\mu = 0$, $\sigma = 1$). Then this linear model is comparable with a CCA model with the first pair of canonical patterns given by w_{ij}^1 and w_{ij}^2. If we think of a canonical pair of patterns as representing a physical process, as we did in the Iberian rainfall example, then we should include so many neurons in the hidden layer as many physical situations giving rise to rainfall. As a rule of thumb, this number should be *of the order* of the rainy *Grosswetterlagen* for this region or the number of significant rainfall EOFs. In this example we have included five elements in the hidden layer.

The last question that needs to be solved is the form of the filter function f. For many applications, sigmoidal-type functions of the form

$$f(x) = \frac{1}{1 + e^{x - x_c}} \tag{11}$$

have been used (Fig. 8), but they are not suitable for our downscaling purposes: the net could generate rainfall only in the interval (0, 1), or by rescaling in some a priori finite interval; furthermore it could not generate truly dry days. We have found that the function

$$f(x) = \begin{cases} 0: & x \le x_c \\ e^{r(x - x_c)} - 1: & x > x_c \end{cases} \tag{12}$$

gives reasonable results (Fig. 8), where the cutoff x_c is the value of the input for which the neuron becomes *active.* The fact that it is not strictly differentiable at $x = x_c$ is not a too big practical problem. The parameter x_c can be also optimized with the same algorithm as the parameters w_{ij}^k, with a small mathematical transformation. This choice is dictated by the nature of rainfall and is not universally applicable. For other applications other forms for the nonlinear transformation f could be explored. Of course the form of f for the last layer of neurons strongly influences the probability distribution of the simulated local variable, so that a good choice for f has to take the real distribution into account. Note that in principle one could use different filter functions for the different neurons, so that considerable flexibility is possible.

Following these considerations a neural network has been designed to describe the daily rainfall in wintertime in Cáceres (Spain) in the period 1978–83. These winters

include relatively well-defined wet and dry periods, so that the skill of the net can be better illustrated. The input variables are the coefficients of the leading five SLP EOFs, calculated on a daily basis. The coefficients for days t, $t-1$, $t-2$ are used to estimate the rainfall at day t. Therefore the input layer has 15 neurons. An intermediate layer of five neurons has been also used to allow for nonlinear interactions among the EOF patterns. The output layer consists of a single neuron. The form of filter functions is the same for all neurons (see Fig. 8, dashed line), but for the output neuron the exponential parameter is about five times larger. However, the value of this parameter is not very critical since it can be chosen a priori within a wide range and the value of the weights can be also fitted accordingly. This fact also points to the difficulty in ascribing physical meaning to these weights. The weights w_{ij}^k and the values of the cutoffs x_c (different for each neuron) of the network have been fitted with daily data in the winters between 1970 and 1975 by the *back propagation error* algorithm (Rummelhart et al. 1986).

The winter daily SLP data in the period 1978–83 are used to simulate rainfall at Cáceres. The results with the neural net are also compared with those obtained with the analog method, by looking for 3-day evolution analogs in the whole dataset available except in the winter months in 1978–83. Figure 9a shows the observed rainfall and the rainfall estimated by these two methods. Figure 9b shows the same time series but smoothed with a 5-day running mean filter.

It can be seen in Fig. 9 that the analog method reproduces better the time evolution of daily rainfall, although the agreement with observations is not optimal, in particular on a day-to-day basis. The level of variability is also more realistic with the analog method. The neural net shares a common drawback with other ''deterministic models:'' they produce time series with less variance than in the observations and therefore they tend to underestimate the frequency or intensity of heavy rainfall and the frequency of dry days. The agreement with the observations is clearer for both methods in the smoothed data time series, indicating that there exists high-frequency variability that is not captured by both methods. Note that only information from the SLP has been used and that incorporation of geopotential heights or temperature in upper-atmospheric layers (also large-scale fields) is likely to improve the results in the validation period. However, the use of the geopotential height for downscaling GCM information is more problematic, as explained in section 3. It is also suspected (Hughes et al. 1993; Zorita et al. 1995) that the daily rainfall process is not only conditioned by the large-scale metereological fields but also by the local rainfall in previous days, and this information is only indirectly and not completely available to the net through the SLP field in the previous days. But perhaps the biggest drawback is the difficulty to assign a physical interpretation to the weights.

7. The problem of the level of simulated variability

Some models presented here, for instance, the linear models and neural networks, describe in general a partial relationship between independent variables representing the large-scale climate variability and dependent local variables. The part of the local variables that remains undescribed by the independent variables is normally referred to as noise. From this point of view the observed local variable at time $tR(t)$ is one realization of a stochastic variable $\tilde{R}(t)$, with a probability distribution $P[F(t)]$ that depends on the simultaneous large-scale forcing $F(t)$.

The parameters of these models are usually fitted so that in the mean the error between the estimated and observed values is minimized, which means that the fitted model yields the best estimation of the mean of the probability distribution, $E[P(F(t)] = \overline{P(F)}$, (in the sense of smallest variance of the errors). We denote this optimal fitted model by M and the best estimation of $\overline{P(F)}$ by $\hat{R}(t)$. With this notation:

$$\hat{R}(t) = M[F(t)]. \tag{13}$$

However, the fitted model M is not optimized with respect to the the variance of $\tilde{R}(t)$. There are two contributions to the variance of $\tilde{R}(t)$: a local one,

$$\mathrm{Var}_{\mathrm{local}} = E\{[P(F) - \overline{P(F)}]^2\}, \tag{14}$$

caused generally by local processes and measurements errors, and the variance forced externally by F:

$$\mathrm{Var}_{\mathrm{external}} \propto E[(F - \overline{F})^2], \tag{15}$$

where we have assumed that the variability of the external forcing is independent of the internal variability.

When the statistical model is applied to an external forcing simulated by a GCM or taken from the observations in a verification period, the variance of the simulated output is less than the observed variance. This occurs because the variance of the simulated local variable is caused only by the variance of the external forcing $\mathrm{Var}_{\mathrm{external}}$, and does not contain the *internal* contribution to the variance of $\tilde{R}(t)$. This is not important if the aim of the model is just the estimation of changes in the mean local climate, but it is really important if the output of the statistical model is used to drive an ecosystem or sector model. In this case the level and structure of this noise may need to be addressed.

Some authors (Karl et al. 1990) have used inflated regression coefficients in linear models to increase the variance of the simulated output. However, in doing so one is artificially enlarging the part of the variability of the local variable that is driven by the large-scale forcing. Another approach to this problem has been recently proposed by Bürger (1996). According to this author the step in the design of the downscaling model, namely, the estimation of the model parameters by minimizing the differences between the model response and the ob-

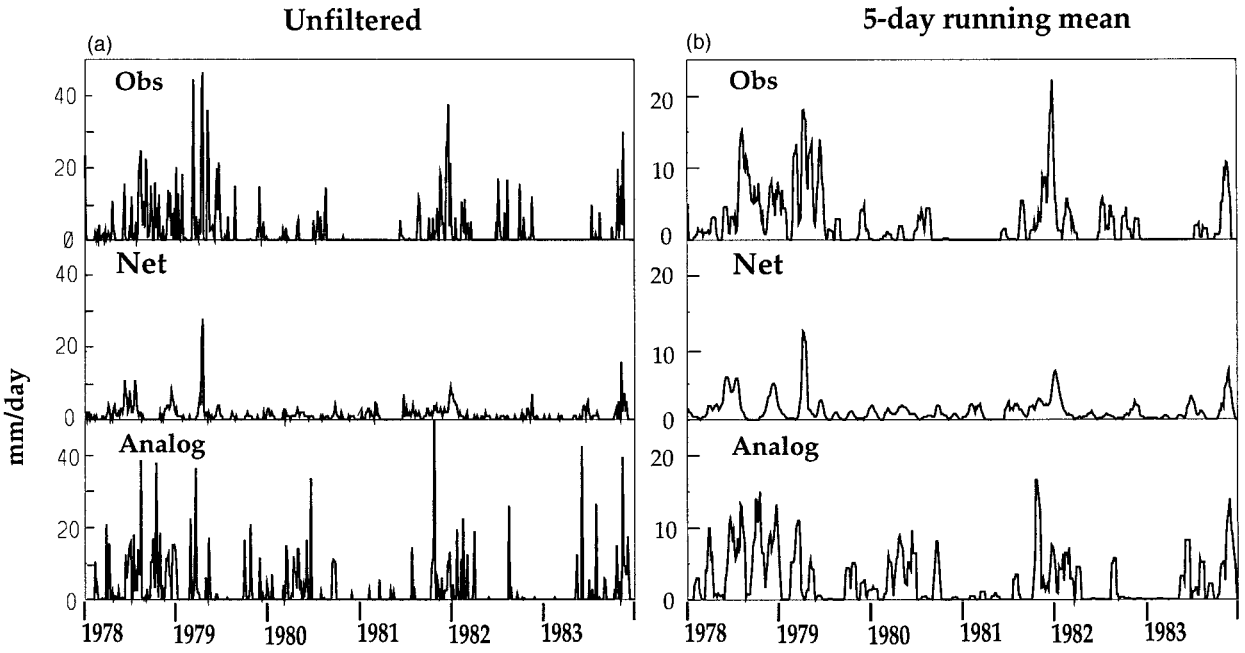

FIG. 9. Daily rainfall (mm day^{-1}) time series at Cáceres, Spain in DJF in the years 1978–83 (first point is 1 Jan 1978, last 31 Dec 1983), observed and simulated by the neural network and the analog method. The input for both models are the coefficients of the five leading SLP EOFS in the current and two previous days. The coefficients of the neural net have been fitted in the winters between 1970 and 1975.

servations, is replaced by a constrained minimization procedure. The simulated local variables are forced to have the same covariance structure, and therefore the same individual variances, as the local observations. The price that has to be paid is that the fitting between simulations and observations in the training period is not as good as with an unconstrained minimization. Therefore, the statistical model produces a simulated output with the right level of local variability, but it is less consistent with the large-scale forcing. One has to find a compromise between both requirements that surely will depend on the particular application.

A more consistent way would be to acknowledge our ignorance about the origin of this unexplained part of the local variability and try to take it into account as an additional and independent random component. Therefore, for the purposes of obtaining estimates of the local variable, for instance, to be subsequently used as forcing for an ecosystem model, the estimation of (13) could be replaced by a more useful estimation of the local variable conditional on the external forcing:

$$\hat{R}(t) = M[F(t)] + \epsilon, \tag{16}$$

where ϵ is a random noise. The level of variance of ϵ should be chosen in a way that the variance of the observed and simulated local variables is the same. For applications to perturbed climates, it has be to be assumed that the local noise remains unchanged.

The design of the local noise ϵ is straightforward for a linear model with normally distributed variables. In this case ϵ can be simply a Gaussian noise with zero mean and variance given by:

$$\mathrm{Var}(\epsilon) = \mathrm{Var}(R) - \mathrm{Var}[M(F)]. \tag{17}$$

In other words, the variance of ϵ is the difference between the variance of the observed local variable and the variance of the simulated response. For more complicated cases, for instance, daily rainfall, it has to be guaranteed that the simulated local rainfall is nonnegative. Therefore ϵ may depend in general on the estimation $M[F(t)]$.

These considerations do not apply to the methods based on some sampling of a pool of observations, such as the analog method or in general the classifications methods. These techniques incorporate implicitly the part of the variance that is not due to the large-scale forcing, by choosing one observation that may be consistent with the forcing, but that also contains one realization of the local noise. Therefore the methods based on sampling will yield in general a more uncertain estimation of the mean local variable (in the sense of larger estimation variance), but provide at the same time a more realistic variability.

8. Concluding remarks

A relatively simple statistical downscaling technique based on the analog method has been presented and applied to daily and monthly winter rainfall in the Iberian Peninsula. The relationship between the large-scale

SLP field in the North Atlantic and Iberian rainfall was already known (Lamb and Peppler 1987; von Storch et al. 1993) and it represents a good example to test the relative performance of statistical downscaling methods. The use of monthly and daily rainfall allows for the testing of the method in two different conditions, one where the local variables are more or less normally distributed and one where they are clearly nonnormally distributed. The analog method has been compared in these two situations with more complicated statistical downscaling techniques. These were taken to be relatively complicated methods representing three different families: linear methods (canonical correlation classification methods (classification and regression trees) and neural networks.

It has been found that the analog method performs quite satisfactorily when compared to the other methods. When compared to the linear method based on canonical correlation, it can be said that the performance of both methods is very similar when applied to monthly rainfall, and also when both are applied to the statistics of daily rainfall. The linear methods, however, offer the advantage of a straightforward physical interpretation of the spatial patterns obtained, thus increasing the degree of confidence on the relationship between large-scale and local variable. On the other hand the analog method produces the right level of variability of the local variable when driven by historical large-scale observations, and this point, which may be quite important if the downscaled climate changes are used to drive an ecosystem or hydrological model, is not guaranteed by a linear method.

The analog method and the classification methods in general automatically produce the right mean and variance of the local variable (a purely random resampling of the pool of historical observations would also be succesful). The classification methods are not intended to replicate exactly the observed realization of the local variable, but just their statistical properties. Therefore a comparison of both methods should be based in time-evolution statistics, such as autocorrelation, or length of dry periods. In this sense the CART method underestimates the persistence of dry days, also in other locations (Zorita et al. 1995), and the analog method behaves somewhat better, although still underestimating this persistence.

Neural networks are designed to be able to replicate the observed local variables when they are driven by the observed large-scale fields, once their internal parameters have been estimated. Therefore the comparison with the analog method can be made on the basis of daily rainfall time series, observed in an independent period. In our example the neural network performed worse than the analog method both in terms of coherent evolution of observations and replication and in terms of the level of replicated variability. This is somewhat surprising since the analog method is technically by far more simple than the neural network and it needs certainly much less computing time. Other more successful applications of neural nets for the specification of local rainfall from the large-scale circulation have been carried out for monthly values (Hewitson and Crane 1992), where a linear model may also produce good results.

The results of this study suggest that in the cases of normally distributed variables, where linear methods can be applied, they can provide a direct physical interpretation and their output can be augmented to produce the right level of variability by adding stochastic noise. In the nonlinear cases we have found that the analog method is in general preferable to the others, due to its technical simplicity and comparable or better performance.

Acknowledgments. We thank María Casado from the Instituto Nacional de Meteorología (Madrid), who kindly supplied the Spanish daily rainfall data, and Fidel González from the Universidad Complutense (Madrid, Spain), who provided the monthly rainfall records. Frank Kauker, Stefan Güss, Hauke Heyen, Marisa Montoya, and Ralf Weisse helped us with valuable comments to correct and improve earlier versions of this manuscript.

REFERENCES

Bárdossy, A., and E. J. Plate, 1992: Space–time model for daily rainfall using atmospheric circulation patterns. *Water Resour. Res.,* **28,** 1247–1259.

Barnett, T., and R. Preisendorfer, 1978: Multifield analog prediction of short-term climate fluctuations using a climate state vector. *J. Atmos. Sci.,* **35,** 1771–1787.

Biau, G., E. Zorita, H. von Storch, and H. Wackernagel, 1999: Estimation of precipitation by kriging in the EOF space of the sea level pressure field. *J. Climate,* **12,** 1070–1085.

Breimann, L., J. H. Friedmann, R. A. Olsen, and J. C. Stone, 1984: *Classifications and Regression Trees.* Wadsworth, 391 pp.

Bretherton, C., C. Smith, and J. Wallace, 1992: An intercomparison of methods to find coupled patterns in climate data. *J. Climate,* **5,** 541–560.

Bürger, G., 1996: Expanded downscaling for generating local weather scenarios. *Climate Res.,* **7,** 111–128.

Cheng, X., and J. M. Wallace, 1993: Cluster analysis of the Northern Hemisphere wintertime 500-hPa height fields: Spatial patterns. *J. Atmos. Sci.,* **50,** 2647–2696.

Cohn, T. A., L. L. DeLong, E. J. Gilroy, R. M. Hirsch, and D. K. Wells, 1989: Estimating constituent loads. *Water Resour. Res.,* **25,** 937–942.

Cubasch, U., H. von Storch, J. Waszkewitz, and E. Zorita, 1996: Estimates of climate changes in southern Europe using different downscaling techniques. *Climate Res.,* **7,** 129–149.

Elsner, J. B., and A. A. Tsonis, 1992: Nonlinear predictions, chaos, and noise. *Bull. Amer. Meteor. Soc.,* **73,** 49–60.

Fliri, F., 1974: Niederschlag und Lufttemperatur im Alpenraum. *Wiss. Alpenv.,* **24,** 111–125.

Giorgi, F., and L. O. Mearns, 1991: Approaches to the simulations of regional climate change: A review. *Rev. Geophys.,* **29,** 191–216.

Grieger, B., and M. Latif, 1994: Reconstructing the El Niño attractor with neural networks. *Climate Dyn.,* **10,** 267–276.

Grotch, S., and M. MacCracken, 1991: The use of general circulation models to predict regional climate change. *J. Climate,* **4,** 286–303.

Hense, A., R. Glowienka-Hense, H. von Storch, and U. Stahler, 1990: Northern Hemisphere atmospheric response to changes of Atlantic Ocean SST on decadal time scales: A GCM experiment. *Climate Dyn.,* **4,** 157–174.

Hewitson, B., 1996: Climate downscaling: Techniques and application. *Climate Res.,* **7,** 85–95.

——, and R. G. Crane, 1992: Large-scale atmospheric controls on local precipitation in tropical Mexico. *Geophys. Res. Lett.,* **19,** 1835–1838.

Hughes, J., D. Lettenmaier, and P. Guttorp, 1993: A stochastic approach for assessing the effect of changes in regional circulation patterns on local precipitation. *Water Resour. Res.,* **29,** 3303–3315.

Hulme, M., K. Briffa, P. Jones, and C. Senior, 1993: Validation of GCM control simulations using indices of daily airflow types over the British Isles. *Climate Dyn.,* **9,** 95–105.

Jones, P., M. Hulme, and K. Briffa, 1993: A comparison of the Lamb circulation types with an objective classification derived from grid-point mean-sea-level pressure data. *Int. J. Climatol.,* **13,** 655–665.

Karl, T. M., W.-C. Wang, M. E. Schlesinger, D. E. Knight, and D. Portman, 1990: A method of relating General Circulation Model simulated climate to the observed local climate. Part I: Seasonal statistics. *J. Climate,* **3,** 1053–1079.

Kauker, F., and J. Oberhuber, 1998: A regional version of the ocean general circulation model OPYC with open boundaries and tides. *Tellus,* in press.

Kruizinga, S., and A. Murphy, 1983: Use of an analogue procedure to formulate objective probabilistic temperature forecasts in the Netherlands. *Mon. Wea. Rev.,* **111,** 2244–2254.

Lamb, H. H., 1972: British Isles weather types and a register of daily sequences of circulation patterns, 1861–1971. Geopysical Memoir 116, HMSO, London, 85 pp. [Available from HMSO, P.O. Box 276, London SW8 5DT, United Kingdom.]

Lamb, P., and R. Peppler, 1987: The North Atlantic oscillation: Concept and an application. *Bull. Amer. Meteor. Soc.,* **68,** 1218–1225.

Lau, C., and B. Widrow, Eds., 1990: Special issue on neural networks. *Proc. IEEE,* **78,** 1411–1414.

Lindberg, C., and A. Broccoli, 1996: Representation of topography in spectral climate models and its effect on simulated precipitation. *J. Climate,* **9,** 2641–2659.

Lorenz, E. N., 1969: Atmospheric predictability as revealed by naturally occurring analogs. *J. Atmos. Sci.,* **26,** 639–646.

Machenhauer, B., M. Windelbrand, M. Bozet, and M. Deque, 1996: Validation of present-day regional climate simulations over Europe: Nested LAM and variable resolution global model simulations with observed or mixed ocean boundary conditions. Max-Planck-Institut Rep. 191, 60 pp. [Available from Max-Planck-Institut für Meteorologie, Bundesstr. 55, D-20146 Hamburg, Germany.]

Mearns, L. O., F. Giorgi, L. McDaniel, and C. Shields, 1995: Analysis of daily variability of precipitation in a nested regional climate model: Comparisons with observations and doubled CO_2 results. *Global Planet. Change,* **10,** 55–78.

Navone, D., and H. A. Ceccatto, 1994: Predicting Indian monsoon rainfall—A neural network approach. *Climate Dyn.,* **10,** 305–312.

Risbey, J., and P. Stone, 1996: A case study of the adequacy of GCM simulations for input to regional climate change. *J. Climate,* **9,** 1441–1467.

Rummelhart, D. E., G. E. Hinton, and R. J. Williams, 1986: Learning representations by back-propagating errors. *Nature,* **323,** 533–536.

Schnur, R., and D. P. Lettenmaier, 1998: A case study of statistical downscaling in Australia using weather classification by recursive partitioning. *J. Hydrology,* **213,** 362–379.

Schuepp, M., 1953: Die Klassifikation der Wetterlagen im Alpengebiet. *Geofis. Pura Appl.,* **44,** 242–248.

Shabbar, A., K. Kiguchi, and K. Knox, 1990: Regional analysis of the Northern Hemisphere 50 kPa geopotential height from 1946 to 1985. *J. Climate,* **3,** 543–557.

Tangang, F. T, W. W. Hsieh, and B. Tang, 1997: Forecasting the equatorial Pacific sea surface temperatures by neural networks models. *Climate Dyn.,* **13,** 135–147.

van den Dool, H., 1994: Searching for analogs, how long must we wait? *Tellus,* **46A,** 314–324.

von Storch, H., 1995: Inconsistencies at the interface between climate research and climate impact studies. *Meteor. Z.,* **4,** 72–80.

——, E. Zorita, and U. Cubasch, 1993: Downscaling of climate change estimates to regional scales: An application to Iberian winter rainfall in winter time. *J. Climate,* **6,** 1161–1171.

Wilby, R. L., and T. M. L. Wigley, 1997: Downscaling general circulation model output: A review of methods and limitations. *Prog. Phys. Geogr.,* **21,** 530–548.

Zorita, E., J. Hughes, D. Lettenmaier, and H. von Storch, 1995: Stochastic downscaling of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *J. Climate,* **8,** 1023–1042.