Using Self-Organizing Maps to Investigate Extreme Climate Events: An Application to Wintertime Precipitation in the Balkans

TEREZA CAVAZOS

Center for Integrated Regional Assessments, The Pennsylvania State University, University Park, Pennsylvania, and Department of Environmental and Geographical Science, University of Cape Town, Rondebosch, South Africa

(Manuscript received 25 February 1999, in final form 11 June 1999)

ABSTRACT

This paper examines some of the physical mechanisms and remote linkages associated with extreme wintertime precipitation in the Balkans. The analysis is assessed on daily timescales to determine the role of the circulation and atmospheric moisture on extreme events, and also at intraseasonal and interannual timescales to find possible linkages with the North Atlantic Oscillation (NAO) and the Arctic Oscillation (AO) patterns. A nonlinear classification known as the self-organizing map (SOM) is employed to obtain the climate modes and anomalies that dominated during the 1980–93 period. An artificial neural network (ANN) is also used to derive daily precipitation at gridpoint scale and at local scale in Bucharest, Romania. Of the predictors used, 500–1000-hPa thickness, 700-hPa geopotential heights, and 700-hPa moisture are the most important controls of daily precipitation. These results are substantiated with the climate states from the SOM classification, which show strong meridional flow over central and eastern Europe coupled to increased winter disturbances in the central Mediterranean and a tongue of moisture at the 700-hPa level from the eastern Mediterranean and the Black Sea during anomalously wet events in the Bulgarian region. Dry events are almost an inverse of these conditions. Extreme events are further modulated by changes in the circulation associated with the AO. In contrast, the NAO does not play a role on wintertime precipitation over the region. The ANN captures well synoptic events and dry spells, but tends to overestimate (underestimate) small (large) events. This suggests a problem for area-averaged precipitation, which is already biased by its spatial resolution. However, comparison between precipitation at Bucharest station and at its nearest grid point shows that the performance of the ANN is slightly better at gridpoint scale.

1. Introduction

Society is accustomed to dealing with climate variations. Agriculture, water management, and many other human activities are already organized to take into account the seasonal march. However, as population grows, society is increasingly vulnerable to climate variations at different spatial and temporal scales, but particularly important are those associated with extreme events. Knight et al. (1995) and Velev (1996) document that by the end of 1993 some Bulgarian stations had periods as long as 11 consecutive years of below normal annual precipitation, and that the reasons for frequent drying conditions are changes in the atmospheric circulation. Similarly, the entire Balkan Peninsula was characterized by decreases in the annual precipitation during the 1982–91 decade; in particular, winter precipitation was about 20% below normal (Sahsamanoglou et al. 1997). Winter droughts in the 1980s and the beginning of the 1990s were not exclusively localized over the Balkans. Hurrell and van Loon (1997) report that since the early 1980s conditions have been anomalously dry over southern Europe and the Mediterranean and wetter than normal over northern Europe and parts of Scandinavia. This is associated with the positive phase of the North Atlantic Oscillation (NAO) pattern documented by Barnston and Livezey (1987). Opposite patterns of precipitation anomalies are typically observed during strong negative phases of the NAO. Recurring positive phases of the NAO, characterized by strong zonal flow over the North Atlantic and Europe, occurred during the 1979/80 winter with the atmosphere locked into this mode through the 1994/95 winter season (Halpert and Bell 1997). During this 15-yr interval, a substantive negative phase of the NAO appeared only twice, in the winters of 1984/85 and 1985/86. However, seasonal winter precipitation in the Bulgarian region (Fig. 1) is negatively, but not significantly, correlated with the NAO index. This suggests that winter precipitation variability over this part of Europe is associated with other remote and mesoscale forcings.

Winter in the central and eastern Mediterranean re-

Corresponding author address: Tereza Cavazos, Dept. of Geography and Regional Development, The University of Arizona, Tucson, AZ 85721.
E-mail: cavazos@geog.climate.arizona.edu

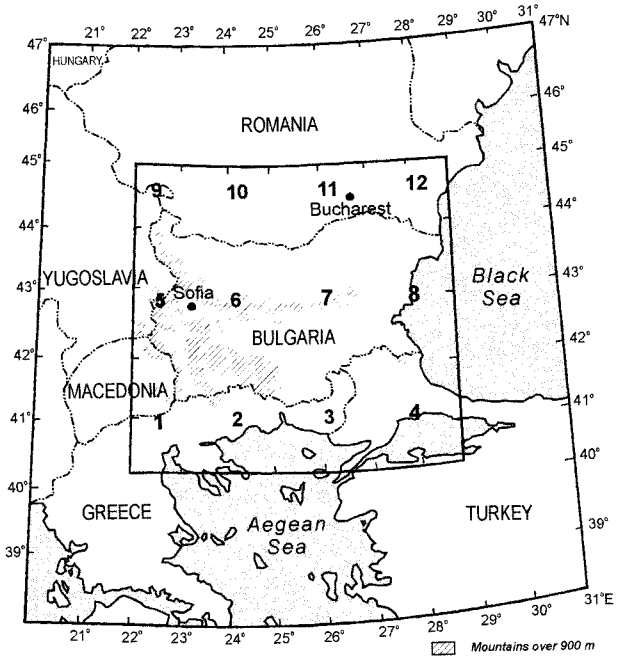

FIG. 1. Study area: the Balkan region. The small frame shows the central location of the 12 NCEP–NCAR target grid points used in this analysis. Some atmospheric controls are extracted for the larger area.

gion is characterized by cyclonic disturbances and low pressure in the Mediterranean Sea, with higher pressure to the east associated with the Siberian high (Palutikof et al. 1996). Wigley and Farmer (1982) document that the Gulf of Genoa is the most important point of origin of cyclonic disturbances within the Mediterranean Basin. These precipitation-producing depressions form in the lee of the Alps, where orographic effects, the Gulf of Genoa and the Adriatic Sea play significant roles (e.g., Wigley and Farmer 1982; Furlan 1977). Velev (1996) reports that Mediterranean cyclones moving from the Adriatic to the Aegean Sea have become less frequent and, consequently, pronounced January droughts and warming have been observed in Bulgaria during the 1982–93 period. Other studies also document a warming trend during winter over parts of Eurasia over the last 30 yr, which has been linked to a prevalence of the high index phase of the NAO (Hurrell 1995) and to the deepening of the polar vortex from the surface to the lower stratosphere (Thompson and Wallace 1998, hereafter TW98). TW98 document that the Arctic Oscillation (AO) pattern is more strongly coupled to surface air temperature fluctuations over Eurasia than the NAO, and that the AO can be interpreted as the surface signature of modulations in the strength of the polar vortex. Although the present study is not concerned with surface temperatures at the moment, the circulation anomalies over the Arctic and Europe associated with these teleconnection patterns could be possibly linked to the changes in the winter disturbances in the Mediterranean and, hence, to the winter droughts and warming in Bulgaria reported by Velev (1996). Thus, one of the aims of this study is to improve our understanding of the remote linkages associated with extreme precipitation events in the Balkan region.

This paper explores the influence of the circulation, humidity, and remote (teleconnections) controls on extreme precipitation events during winter in the Bulgarian region by means of a nonlinear classification approach known as a self-organizing map (SOM). In addition, an empirical downscaling technique based on artificial neural networks (ANNs) is used as a diagnostic tool to derive daily precipitation (e.g., transfer functions) in the 12 grid points shown in Fig. 1. The grid points represent an approximate area of 2° lat × 2° long. A comparison between daily local precipitation at Bucharest, Romania, and the nearest grid point will be also assessed. Elucidating these issues and finding climate anomalies and relationships between various atmospheric controls and teleconnection patterns is crucial to better understand the physical mechanisms and remote linkages that trigger extreme precipitation events over southeastern Europe. This, in turn, may lead to improved diagnostic and downscaling models for regional climate variability and climate change impact studies.

2. Data

The winter (December–February) records for this analysis consist of twice-daily gridded precipitation rate from the National Centers for Environmental Protection–National Center for Atmospheric Research (NCEP–NCAR) global reanalysis project (Kalnay et al. 1996) and monthly gridded precipitation from the Climatic Research Unit (CRU) of the University of East Anglia. The NCEP–NCAR reanalysis data is derived from a global spectral model with an approximate horizontal resolution of 2° lat × 2° long grid. Precipitation observations are not assimilated into the model, thus the precipitation fields that are generated by the reanalysis system are short-range forecast accumulations (Janowiak et al. 1998). The monthly European dataset from CRU has a resolution of 0.5° lat × 0.5° long and spans from 1901 to 1995 (New et al. 2000; hereafter CRU05 dataset). The NCEP–NCAR precipitation rate (kg m^{-2} s^{-1}) was transformed to daily precipitation (mm) and extracted for 12 grid points of the study area (3° lat × 4° long cells in Fig. 1) from 1973 to 1997. Monthly winter precipitation data from several Bulgarian stations were also employed for comparison; these data were obtained from Velev (1996; Table 2). Daily precipitation from Bucharest station, Romania, was obtained from January 1981 to February 1993 from the National Institute of Meteorology, Hydrology and Water Management (INMH), Bucharest.

The atmospheric controls utilized in this analysis are the following: The NAO index obtained from the National Oceanic and Atmospheric Administration

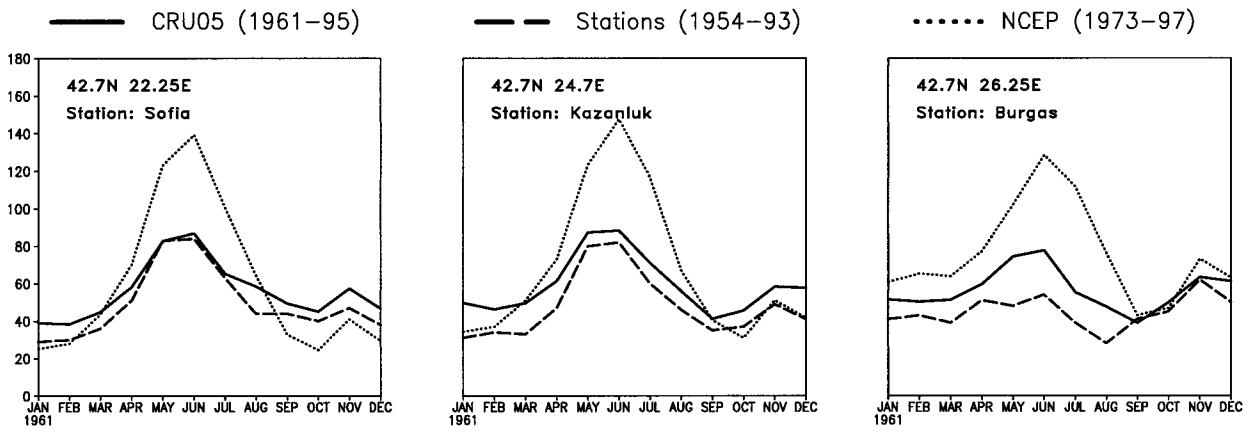

FIG. 2. Zonal comparison (along the 42.7°N lat line) of monthly precipitation (mm) from two gridded datasets (CRU05 and NCEP–NCAR reanalysis) and data from meteorological stations close to the grid points. Station data are from Velev (1996). CRU05 monthly values represent area-average precipitation in a 0.5° lat × 0.5° long grid centered in the grid point, while the NCEP–NCAR data are for an approximate 2° lat × 2° long area. The NCEP–NCAR precipitation corresponds to grid points 5, 6, and 7 in Fig. 1.

(NOAA) teleconnections web site is used. It is based on Barnston and Livezey (1987) rotated principal component technique but using monthly mean 700-hPa geopotential height anomalies for the 1964–94 period. Twice-daily specific humidity at the earth's surface (q0) from the NCEP–NCAR reanalysis for the 12 target grid points in Fig. 1. Twice-daily specific humidity at the 700-hPa level (q7), 700-hPa geopotential heights (Z7), and 500–1000-hPa thickness (Th5_1) from the Goddard Space Flight Center (GSFC) reanalysis project[1] that spans from 1980 to 1993 on a 2° lat × 2.5° long grid. The Th5_1 field is based on the 500-hPa geopotential heights (Z5) and the sea level pressure (SLP) with the latter converted to 1000-hPa heights (Z1) using $Z1 \cong 8(SLP - 1000)$. The Z7, Th5_1, and q7 were regridded to match the precipitation data (NCEP–NCAR grid). The Z7 and Th5_1 were extracted for an area that is one grid larger than the precipitation grid (large area in Fig. 1) and q7 was extracted for the 12 target grid points in Fig. 1. Specific humidity at the earth's surface was obtained from the NCEP–NCAR reanalysis because the q0 from GSFC had many gaps. The analysis is constrained to the 1980–93 period, which is the record of the atmospheric variables from the GSFC reanalysis dataset.

Other atmospheric variables and teleconnection indices employed in this study are: SLP, Z5, and q7, which were extracted for a large-scale geographical window (60°W–60°E, 20°–90°N) from the NCEP–NCAR reanalysis for selected periods. The AO of TW98 is available via anonymous FTP from the Joint Institute for the Atmosphere and Ocean (JISAO) of the University of Washington.[2] The AO index is based on the leading principal component of the wintertime (November–April) monthly SLP anomaly field over the domain poleward of 20°N (Thompson and Wallace 1998), and the original time series spans from 1898 to 1997. The winter climatologies of the SLP, Z7, and Z5 fields are similar for both the GSFC and the NCEP–NCAR reanalyses over the North Atlantic and Europe during the analyzed period.

3. Intercomparison of precipitation data

As indicated in the last section, the NCEP–NCAR reanalysis precipitation is derived from a global spectral model. Thus, the gridpoint precipitation may differ from local observations as the reanalysis precipitation is highly model dependent. Because of this reason, a monthly comparison with two other datasets was assessed. The NCEP–NCAR daily precipitation was transformed to monthly totals to compare them with the observed-based CRU05 monthly dataset and with monthly observed precipitation data from several meteorological stations in Bulgaria. Figure 2 shows a zonal sample of the annual cycle of precipitation for three grid points and three nearby meteorological stations along the latitude line of 42.7°N. There is a good agreement among the three datasets during the winter and fall seasons. The CRU05 captures well the annual cycle of precipitation, while the NCEP–NCAR reanalysis greatly overestimates spring and summer precipitation. Janowiak et al. (1998) states that deficiencies in the NCEP–NCAR precipitation reanalysis are not unexpected because the reanalysis is highly model dependent, particularly in the Tropics

[1] Available online at http://dao.gsfc.nasa.gov/.

[2] Available online at ftp://ftp.atmos.washington.edu/jisao/davet.

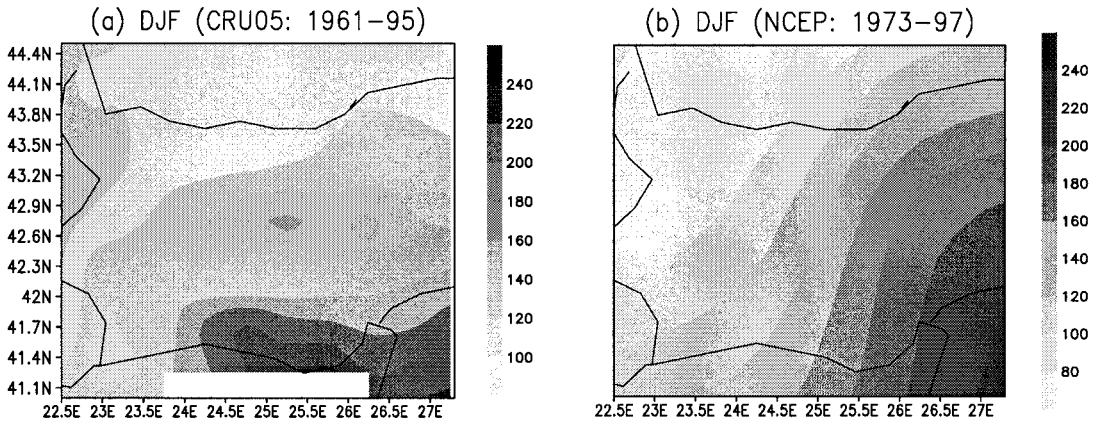

FIG. 3. Spatial distribution of mean total winter (Dec–Feb) precipitation (mm) in the Bulgarian region from (a) the CRU05 gridded dataset and (b) the NCEP–NCAR reanalysis. Grid resolution as in Fig. 2. White area in the CRU05 map indicates lack of data over the Aegean Sea.

and the warm season extratropics where convective precipitation dominates. The CRU05 overestimates the winter precipitation along the 42.7°N lat line, but in general the spatial distribution of winter precipitation from this dataset (Fig. 3a) agrees well with the precipitation reported by Velev (1996) for other Bulgarian stations, and also with the distribution documented by a Bulgarian atlas (Bulgarian Academy of Sciences 1973, p. 55). The NCEP–NCAR dataset reproduces well the winter precipitation over the central and western regions (Fig. 3b), but overestimates it by about 100 mm in the southeastern part of the study area.

An additional comparison was assessed between daily, monthly, and seasonal precipitation at Bucharest station and the nearest NCEP–NCAR gridpoint precipitation (gridpoint 11 in Fig. 1). The results of this comparison indicate that the daily correlation between the two datasets is 0.62 with a mean absolute error of 1.21 mm day^{-1}. Table 1 shows that gridpoint 11 underestimates the mean daily precipitation by 16% and the mean seasonal precipitation by 10%. The standard deviations indicate that local-scale precipitation at Bucharest has 32% more daily variation than the grid point does. At seasonal scales the discrepancy is smaller. It will be shown later that grid point 11 captures well most of the synoptic events observed in Bucharest, but the amplitude of large precipitation events is underestimated. This reflects a typical bias of area-averaged precipitation datasets, which tend to smooth out large events. The consistency between the monthly datasets during winter and the relative consistency at seasonal and synoptic timescales between gridpoint 11 and Bucharest station reveal that the NCEP–NCAR winter precipitation data can be a useful source for diagnostic analyses over the study area.

TABLE 1. Daily and seasonal comparison of winter (Dec–Feb) precipitation (P) and standard deviation (σ) between Bucharest station and the nearest NCEP–NCAR grid point 11. The daily correlation between the two datasets is 0.62 and the mean absolute error is 1.21 mm day^{-1}. Geographical locations of Bucharest and grid point 11 are shown in Fig. 1. Data are from January 1981 to February 1993; P and σ are in mm.

	P_{daily}	σ_{daily}	P_{DJF}	σ_{DJF}
Bucharest	1.29	4.21	113.10	37.17
Grid 11	1.08	2.81	101.60	39.78

4. Neural network methodology

The statistical downscaling methodology described in this section is used as a diagnostic tool to derive anomalous atmospheric patterns characteristic of extreme precipitation events and to derive daily precipitation at gridpoint and local scales. The methodology consists of three main steps: 1) classification of the atmospheric controls into different climate modes (i.e., weather types), 2) derivation of large-scale climate anomalies associated with extreme precipitation events, and 3) derivation of empirical transfer functions between the atmospheric controls and daily precipitation at gridpoint and local scales. The details of this methodology are discussed in Cavazos (1999) and only a short description suffices here. It is based on two types of ANNs: SOMs and feed-forward ANNs. The SOM technique is used for classification and pattern recognition. Detailed documentation of the theoretical framework of SOMs can be found in Kohonen (1995) and the SOM algorithm (Kohonen et al. 1995). SOM is partially analogous to cluster analysis, but one of the advantages of the SOM is that it generates a nonlinear classification based on an iterative process, which rewards the winning class and its closest neighbors, while inhibits those farther away. The unsupervised learning algorithm of the SOM seeks to discover patterns; in this case, the SOM is supposed to find significant features that characterize the daily atmospheric circulation and humidity fields

(e.g., the atmospheric controls) during winter over the study area. The classification is carried out at each of the 12 target grid points in Fig. 1 using gridpoint, spatial, and remote atmospheric controls. Gridpoint variables, such as q0 and q7, contain information from the target grid point only. Spatial variables (Z7 and Th5_1) include information from a 3×3 matrix of cells centered in the target grid point. The NAO index is used as a remote control. All variables, except the NAO, are lagged over 24 h with a 12-hourly resolution and 12 h lead (i.e., three lagged times). Thus, the SOM classes at each grid point are a function of a time series of atmospheric controls, **X**, composed of *m* daily vectors of 61 atmospheric variables:

$$\text{SOM classes} = f({}_m\mathbf{X}_{61}) = f(\text{q0}_3, \text{q7}_3, \text{NAO}, \text{Th5\_1}_{27}, \text{Z7}_{27}).$$

Where *m* spans over 1152 winter days; the subscripts in the right-hand side represent the number of variables used in each group according to the number of grids and lag times utilized: 3 ($=1$ grid $\times$ 3 time lags) for grid point and 27 ($=9 \times 3$) for spatial variables. The SOM algorithm defines a mapping from the normalized input vector **X** onto a two-dimensional array of nodes or classes; the mapping is done through a training process presenting a single daily vector at a time. There is not a rule to determine the ''significant'' size of an SOM map: the larger the map the better the continuity of atmospheric events. For simplicity, nine classes were chosen in this study, which means a map of 3×3 nodes with an average of 128 winter days per class. Every node has an associated parametric reference (weight) vector that has the same dimension as the input vector **X**; the initial reference weights are selected at random. Most SOM applications use either the smallest of the Euclidian distances or the maximum dot product (maximum correlation) in defining the best matching between the input vector and the reference vector. The best-matching node in the map is rewarded by slightly modifying (i.e., minimizing) the reference vector component (Kohonen 1982) and its nearest neighbors. Then, the reference weights of the map are updated and training continues until convergence is reached. The central result in self-organization is a nonlinear projection of the probability density function of the multidimensional input vector onto the two-dimensional map (Kohonen 1995). After training, the best map is expected to yield the smallest average quantization error. The quantization error is the average of all the minimum Euclidian distances (or maximum correlations if dot product is chosen) from the daily records of the input vector **X**. Therefore, an appreciable number of random initializations (>10) ought to be tried to avoid falling in a local minimum. The SOM package has an automatic procedure to calculate the error.

After training the SOM algorithm, the time series of atmospheric controls are composited according to the SOM classes. The median precipitation associated to each SOM class is obtained in each grid point; this helps categorize the SOMs from -1 (the driest SOM class) to $+1$ (the wettest SOM class) in increments of 0.25. This arbitrary linear categorization ensures that similar SOMs, and thus similar days, are close together and that the distance between SOM classes conducive to extreme precipitation events is maximum. The SOM categorical variable (SOM_{CAT}) is added to the input time series of atmospheric controls to condition each day according to the nine SOM classes. This makes a total of 62 predictor variables to derive daily precipitation (*P*) at each grid point and location with the feed-forward ANN:

$$P = f({}_m\mathbf{X}_{62}) = f(\text{q0}_3, \text{q7}_3, \text{NAO}, \text{Th5\_1}_{27}, \text{Z7}_{27}, \text{SOM}_{\text{CAT}}).$$

The ANN used is the NevProp Version 3, which is available online (Goodman 1996). It consists of $N = 62$ input nodes (or predictor variables), five hidden nodes, and an output node representing the response or dependent variable (i.e., precipitation). The neural network is trained separately for each of the 12 grid points of the study area and also for Bucharest station for which controlling variables from the nearest NCEP–NCAR grid point (#11) are used. For each grid point 75% of the 1152 winter days are used to train the net and the remaining 25% are used to validate the net. During training with the feed-forward ANN, a bootstrapping technique is used to estimate the impact of model variability. One hundred bootstrap runs with five random subsamples with replacement are used during training to obtain different measures of accuracy. By running one hundred bootstrap ensembles it is attempted to find the global minimum of the error surface. Bootstrapping produces more stable models, since the averaging of all the random subsamples reduces the effect of individual local minimum of each model in the ensemble (Tangang et al. 1998). At the end, each grid point has a diagnostic nonlinear relationship (i.e., a transfer function) from the feed-forward model that indicates the degree to which the observed atmospheric predictors are able to explain the daily precipitation variance. These transfer functions can be further used to downscale GCM outputs for climate change impact analyses provided that the GCM circulation and humidity fields (e.g., the predictors) have a good agreement with the observed fields over the study area (not done in this analysis). Parts of the methodology employed in this analysis and in past works (e.g., Cavazos 1999) are based on a larger statistical downscaling software developed by Hewitson and Crane for climate change impact studies (see, e.g., Crane and Hewitson 1998; Hewitson 1998; Hewitson and Crane 1999, submitted to *Geophys. Res. Lett.*).

There are three differences between the current methodology and the one employed by Cavazos (1999). First, nine SOM patterns (as opposed to only four) are selected here to develop a more realistic classification of at-

TABLE 2. Interannual grid point correlations between wintertime (Dec–Feb) NCEP–NCAR precipitation (P) in the 12 grid points shown in Fig. 1 and the standardized values of the North Atlantic Oscillation (NAO), as defined in the NOAA teleconnections Web site (http://nic/fb4.noaa.gov:80/data/teledoc/telecontents.html), and the AO of TW98 (see footnote 2). The third value is the correlation between the CRU05 Dec–Feb precipitation averaged over the NCEP–NCAR grid area and the AO index. Correlations are in hundreds. One and two asterisks indicate statistical significance at the 5% and 1% levels, respectively. Period analyzed: 1974–95.

Correlations: $P_{\text{-NCEP-NCAR}}$ & NAO ($P_{\text{-NCEP-NCAR}}$ & AO) $P_{\text{-CRU05}}$ & AO			
−26 (−56**) −51*	−11 (−53*) −42	−4 (−40) −33	9 (−40) −23
−7 (−60**) −53*	0 (−53*) −48*	−6 (−49*) −40	−13 (−34) −37
−4 (−43*) −42	+22 (−13) −41	+17 (−1) −42	−14 (−15) −54*

mospheric conditions conducive to different types of climate events; second, a slightly different ensemble of atmospheric predictors is employed in this analysis; and third, a single diagnostic transfer function per grid point is derived from the feed-forward ANN, as opposed to one for each SOM class, by adding the SOM categorical variable to the input predictors.

Selection of controlling variables

Derivation of precipitation for hydrological and climate change applications requires special consideration of the precipitation mechanisms in a particular region at different timescales. A large number of possible combinations of predictors and methodologies to simulate precipitation is found in the literature. Several studies have noted significant improvements in the performance of downscaling models using nonlinear approaches, such as ANNs, over linear techniques (e.g., McGinnis 1994; Weichert and Bürger 1998; Trigo and Palutikof 1999), while others have seen little difference (e.g., Winkler et al. 1997), or no improvement (at least at daily timescales) over other approaches (Wilby et al. 1998). In this analysis, the SOM-ANN approach is chosen as it has produced physically consistent diagnostic (Cavazos 1999) and downscaling (Hewitson 1998) results. Concerning the predictor variables, the selection is very much constrained by data availability and the GCM output used if downscaling is the objective of the analysis. Winter precipitation in the Balkans, for example, is strongly modulated by the intensity of the zonal flow, the type of airmass and its direction of movement, moisture structure, and by differential heating induced by landscape characteristics and the bodies of water—the Mediterranean, the Aegean, and the Black Seas. The mean SLP field at monthly timescales has been extensively used to explain winter precipitation variability over southern Europe and the Mediterranean (e.g., Zorita et al. 1992; von Storch et al. 1993; Corte-Real et al. 1995; Busuioc and von Storch 1996; Busuioc et al. 1999), as well as the 500-hPa geopotential height mean field (e.g., Corte-Real et al. 1995; Palutikof et al. 1996). Mean large-scale circulation fields account for the circulation dynamics, but would fail to capture the thermodynamic and water vapor processes among other local forcings. Additionally, extreme events would be smoothed out at monthly timescales. In this analysis, the selection of predictors for daily precipitation is based on this simple rationale: dynamical, thermal, and moisture components. The large-scale circulation is partly captured by the Z7 and the Th5_1 fields, and the NAO and the SOM indices. Baroclinic systems such as frontal and cyclonic disturbances are associated with steep temperature gradients of the atmospheric layer, thus the use of 500–1000-hPa thickness. Finally, specific humidities at the earth's surface and at midtropospheric levels (700-hPa) are used to account for frontal moisture and convection. Time-lagged was also added to the atmospheric controls to carry information from the large-scale state into the local scale. There are clearly other predictors (and methodologies) that can be used; data availability and redundancy of predictors played an important role in the final selection of the current study. In the next section, for example, it is demonstrated that even though Z5 and SLP were not explicitly used as predictors, the synoptic composites of these two fields, as derived from the SOM climate modes representative of extreme events, are consistent with past studies of wintertime precipitation variability over southern Europe.

5. Results from the SOM classification

a. Circulation anomalies

The purpose of this section is to improve our understanding of the physical and remote linkages associated with extreme precipitation events over the Balkan region. This is accomplished by exploring the climate anomalies characteristic of extreme events according to the SOM classification. As explained in the last section, the SOM classes of each grid point are used to composite the precipitation associated to each climate mode. The SOM classes representative of extreme and average precipitation events are selected according to the median precipitation. The wet and dry SOM classes for the nearest grid point to Sofia, Bulgaria (gridpoint 5 in Fig. 1), are used to composite the large-scale Z5 and SLP fields and midtropospheric moisture (q7) patterns. These mean fields and their anomalies are chosen to explore the physical characteristics of the atmosphere during extreme precipitation events in the Bulgarian region during the 1980–93 period.

Figures 4a,b show the mean 500-hPa geopotential heights characteristic of wet and dry conditions, respectively. Figure 4a indicates that anomalously wet events

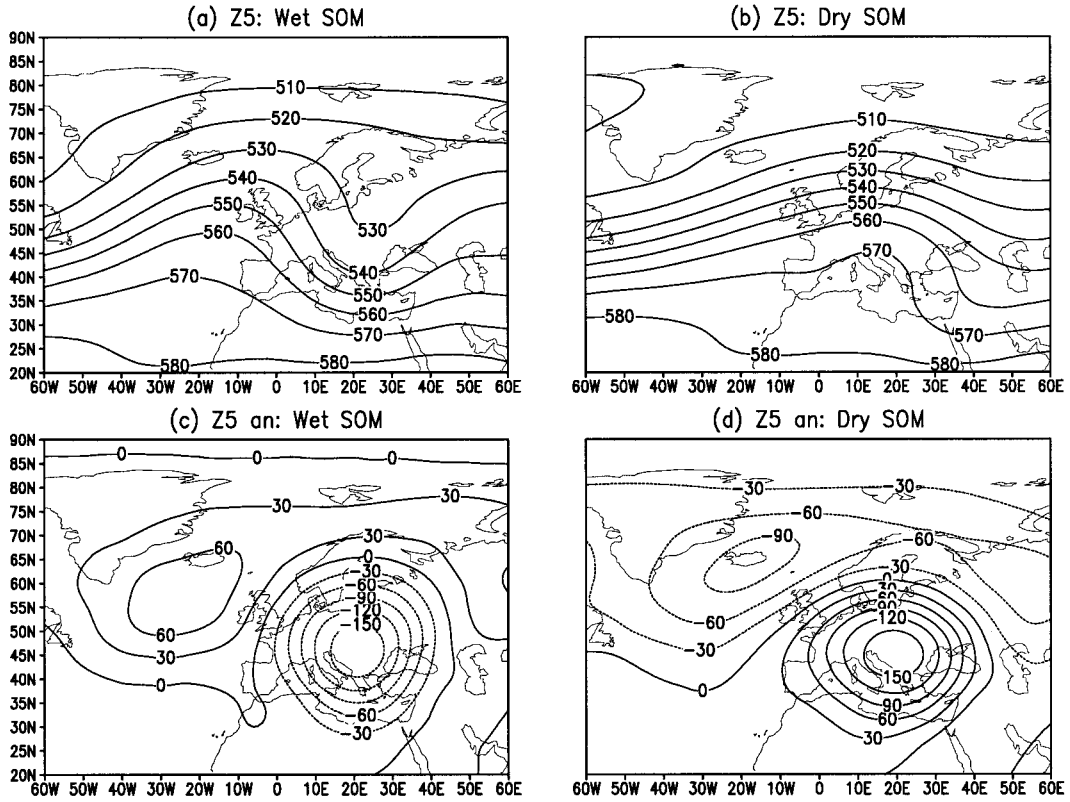

FIG. 4. SOM composites of wintertime (Dec–Feb) mean 500-hPa geopotential heights (Z5) and anomalies (Z5 an). Composites are based on: (a) and (c) wet SOM and (b) and (d) dry SOM climate modes for grid point 5 in Fig. 1. Contour intervals are 10 dam for (a) and (b), and 30 m for (c) and (d). Data are from the NCEP–NCAR reanalysis. Anomalies are computed as departures from the 1980–93 base-period means.

over the Bulgarian region are associated with a strong midtropospheric meridional circulation over central and eastern Europe. Split flow over the subtropical Atlantic and the trough over the Mediterranean suggest a strong subtropical jet entrance region over northern Africa and the eastern Mediterranean during synoptic conditions conducive to wet events. In contrast, anomalously dry events are associated with an enhanced zonal circulation over the eastern North Atlantic and Europe and a ridge that is shifted toward eastern Europe (Fig. 4b). The circulation anomalies during wet and dry events are almost an inverse of one another in central Europe (Figs. 4c,d). An expansion of the circumpolar vortex is evident during dry events (Fig. 4d). The midtrospheric meridional flow during wet events (Fig. 4a) favors the advection of cold air toward the central and eastern Mediterranean where interaction with the warm seas generates strong cyclonic activity. This is further substantiated with the surface pattern (Fig. 5a), which shows a low pressure system over eastern Europe bounded by the Azores anticyclone and the cold Siberian high. The relatively low pressure over Italy and the Balkans is consistent with the axis of maximum frequency of cold fronts in the central Mediterranean Basin during winter, which is found south of Italy where depressions are steered over the Black Sea (Flocas 1984). In contrast, dry events are characterized by an extended Azores high over Southern Europe and the Mediterranean (Fig. 5b). The relatively strong midtropospheric ridge (Fig. 4b) and anticyclone during dry events are indicative of a northward displaced North Atlantic jet. Thus, extreme precipitation events seem to be regulated by changes in the location of the jet stream and associated storm tracks. As with the midtropospheric anomalies, the surface anomalies during extreme events (Figs. 5c,d) are also negatively correlated with each other in central Europe. The expansion of the circumpolar vortex is also evident in the SLP anomalies of the dry events (Fig. 5d).

The signature of the large-scale circulation anomalies and correspondent precipitation response over the Bulgarian region during extreme events appear as a NAO-like response [see, e.g., the December and February NAO patterns documented by Barnston and Livezey (1987)]. Barnston and Livezey (1987) found that the NAO is clearly recognizable in winter studies using different methodologies and timescales, and for several levels in the atmosphere. However, the close anticyclonic pattern evident in the Z5 anomaly field of the dry SOM class (Fig. 4d) bears a much stronger resemblance to the 500-hPa geopotential height anomalies associated with the AO documented by TW98 (see their

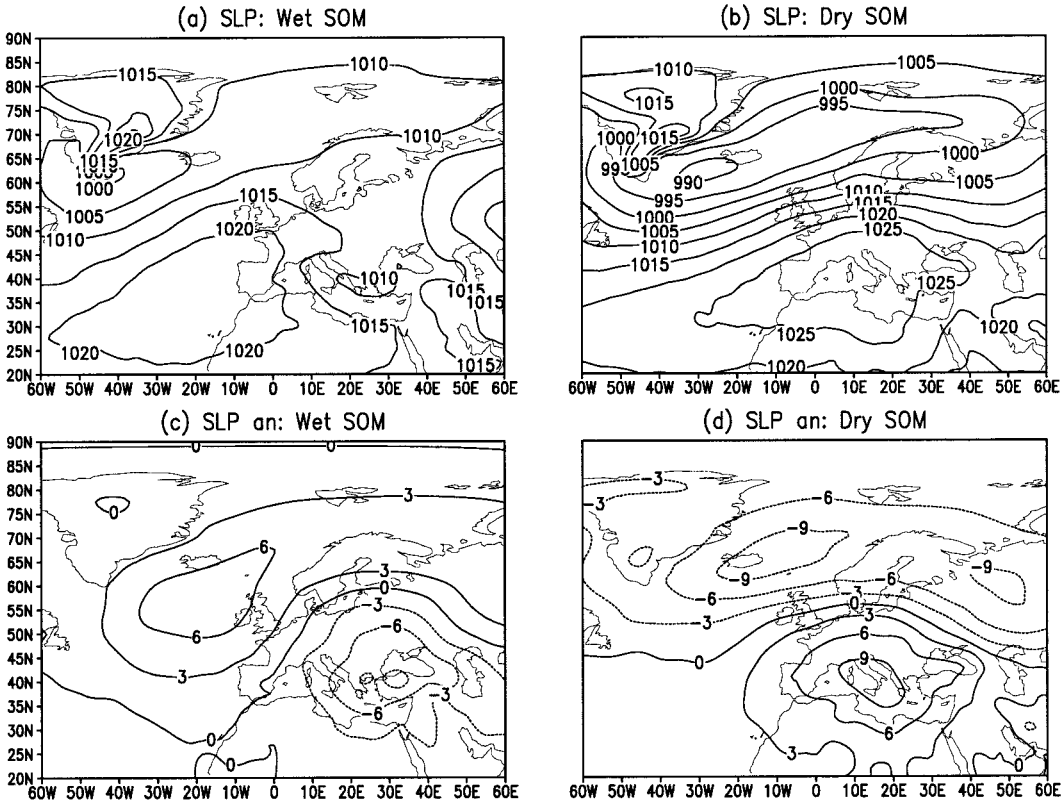

FIG. 5. Same as Fig. 4, but for the mean SLP field and anomalies. Contour intervals are 5 mb (a) and (b), and 3 mb for (c) and (d).

Fig. 1) than to the typical NAO pattern. The high phase of the AO is associated with the deepening of the polar vortex, in combination with above-normal heights over much of Europe and eastern Asia. TW98 found that the AO resembles the NAO in many respects, but its primary center of action is the Arctic. They also hypothesize that under certain conditions, dynamical processes at stratospheric levels can affect the strength of the polar vortex all the way down to the earth's surface. The visual inspection of the climate anomalies associated with extreme events reveals that there is a possible stronger link between extreme precipitation events over the Balkans and the AO pattern than with the NAO.

The composited wet and dry SOM climate modes (Figs. 4, 5) are also consistent with other studies (e.g., Corte-Real et al. 1995; Busuioc and von Storch 1996) based on empirical orthogonal functions (EOF) and canonical correlation analysis (CCA) between large-scale SLP and regional precipitation in Europe. Busuioc and von Storch (1996, hereafter BvS96), for example, related winter precipitation in Romania to the European-scale SLP. Although BvS96 do not specify the canonical modes representative of extreme precipitation events, their first and second modes (their Figs. 7, 8) agree with the sign of the precipitation and the SLP anomaly patterns of the wet and dry SOMs, respectively (Figs. 5c,d). Interestingly, BvS96' second canonical mode (similar to the dry SOM in Fig. 5d, but with an anticyclone slightly shifted to the southwest) is associated with below-normal precipitation only in southern Romania due to the southward-displaced anticyclone, while positive anomalies dominate most of the rest of Romania. Notice that the study area in Fig. 1 also includes the southern portion of Romania. It appears that the spatial extent and persistence of the anticyclonic (cyclonic) circulation over eastern Europe roughly determines the boundary of dry (wet) regions during the winter season. This is interesting because the high phases of the NAO and the AO are both associated with anticyclonic circulation over Europe, as noticed above. However, the anticyclone during the high phase of NAO is weak in southeastern Europe as it lies at the boundary of its climate response, while the high index of the AO is coupled to a relatively stronger anticyclone over southeastern Europe. Therefore, from a teleconnections point of view, the AO may play a stronger role on precipitation events in southeastern Europe than the NAO.

b. The role of midtropospheric moisture during extreme events

As mentioned before, a midtropospheric trough in central and eastern Europe would favor the advection of cold air toward the eastern Mediterranean where in-

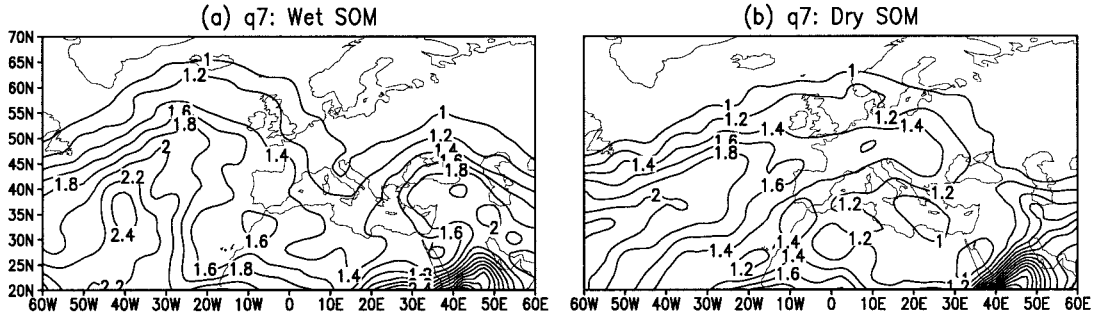

FIG. 6. Winter composite of specific humidity at the 700-hPa level (q7) for the (a) wet and (b) dry SOM patterns of grid point 5 in Fig. 1. Only specific humidities >1 g kg^{-1} are plotted. Contour interval is 0.2 g kg^{-1}. Composites are from the 1980–93 period.

stability of the boundary layer would favor cyclonic activity and transport of moisture toward eastern Europe. Hurrell and van Loon (1997) have demonstrated that changes in the mean and transient flows affect the transport and convergence of moisture and, therefore, can be tied to changes in regional precipitation. In particular, moisture transport fields have been used to explain changes in the NAO pattern over Europe, and therefore, in precipitation fields (e.g., Hurrell 1995; Hurrell and van Loon 1997).

Analyses of several synoptic cases belonging to the wet and dry SOM composites showed similar surface specific humidities (q0) in the eastern Mediterranean Basin and over the study area during both wet and dry events, but striking differences were observed in the humidity patterns at the 700-hPa level (q7). Although, specific humidity at the surface level may play an important role in destabilizing the boundary layer, the humidity at higher levels could be a more discriminating factor during large precipitation events. It seems that in order to break the subsidence circulation typical of dry events it is not only necessary a thermal forcing at lower levels (e.g., moisture), but also at upper levels combined with a dynamical control to lift the moisture. Meridional flow and convergence from the surface to the midtroposphere and a tongue of moisture at the Z7 level, that was advected from the eastern Mediterranean and the Black Sea toward the study area, marked the difference between the wet and dry synoptic precipitation events.

This evidence is further confirmed by the average q7 patterns of the wet and dry SOM climate modes (Fig. 6), which also show that wet events in the Bulgarian region are associated with a south-southeasterly moisture tongue from the eastern Mediterranean and the Black Sea. A contrast in the spatial distribution of moisture is seen in the dry SOM (Fig. 6b), which shows larger moisture amounts over the northern portion of the study area, but also significant reductions of moisture to the south and east, over the Mediterranean and Black Seas, respectively. This pattern, in combination with an enhanced midtropospheric zonal flow and a strong anticyclone over central and eastern Europe (Figs. 4b, 5b), result in dry conditions over the study area. The persistence of the synoptic conditions associated with extreme events, similar to those shown in Figs. 4, 5, 6 are evident at seasonal timescales in the images over eastern Europe provided by the NOAA–CIRES Climate Diagnostic Center web site for the 1980/81 (wet) and 1989/90 (dry) winters.[3] Interestingly, an increase (decrease) in the moisture transport over the Mediterranean has been documented for the low (high) phase of NAO pattern (e.g., Hurrell 1995).

6. Large-scale teleconnections

The climate anomalies characteristic of extreme events presented in the last section were qualitatively associated in several occasions with the typical climate responses of the NAO phases. This is not only because the NAO plays some role on the precipitation variability in southeastern Europe (e.g., Hurrell and van Loon 1997), but also because the European climate-related anomalies associated with the NAO are the most well documented of all the teleconnection patterns that affect Europe. In this section it is shown that the AO pattern is yet more relevant in explaining winter precipitation variability over the Balkans than the NAO.

a. The role of the NAO and the AO patterns in wintertime precipitation

Table 2 shows gridpoint correlations between wintertime precipitation and the NAO and the AO indices at interannual timescales. Correlations with the AO index are all negative and statistically significant over a large portion of the study area for both the NCEP–NCAR and the CRU05 precipitation datasets, while the correlations with NAO are small. The CRU05 precipitation data exhibit a similar spatial distribution of cor-

[3] Online at http://www.cdc.noaa.gov/cgi-bin/DataMenus.pl?stat=mon.mean&dataset=NCEP.

TABLE 3. Wintertime (Dec–Feb) correlations between the wet and dry SOM indices representative of extreme events and (a) the AO index at monthly timescales and (b) the AO and the NAO indices at interannual timescales for the 1980–93 period. One and two asterisks indicate statistical significance at the 5% and 1% levels, respectively.

	Dry SOM	Wet SOM
(a) Intraseasonal		
AO	0.60**	−0.34*
Dry SOM		−0.57**
	AO	Dry SOM	Wet SOM
(b) Interannual			
NAO	+0.40	+0.22	−0.25
AO		+0.75**	−0.60*
Dry SOM			−0.63**

relations to that of NCEP–NCAR, with the exception of the southeastern corner of the study area. This discrepancy may be due to overestimation of winter precipitation over this region by the NCEP–NCAR data, as noticed in section 3. The correlations in Table 2 suggest that the AO pattern play a significant role on the interannual variability of winter precipitation over large portions of the study area, while the role of NAO is minor.

b. Interannual variability of extreme events

To investigate the interannual variability of wintertime precipitation, the study area was divided into regions according to a rotated principal component analysis (RPCA) of the time series of the NCEP–NCAR precipitation for the 1973–96 period. The grid points that loaded high in each component were averaged to obtain regional standardized anomalies of precipitation. Multivariate SOM indices based on the frequency of days of the wet and dry SOM classes were also devised for the same grid point used to composite Figs. 4, 5, and 6 (grid point 5 in Fig. 1). The monthly and winter frequency of days associated with wet and dry conditions were standardized to obtain an SOM index for the 1980–93 period. At intraseasonal timescales, only January shows a clear increasing (decreasing) trend in the frequency of the atmospheric conditions conducive to dry (wet) conditions during the study period. The frequency of the dry SOM index is second in the December months, but without any particular trend. In contrast, the frequency of the dry SOM index is almost absent during the February months of the period. From 1983 to 1988 February was also characterized by the lowest values of the AO (i.e., retracted polar vortex and meridional flow over central Europe). Similarly, Velev (1996) documents that cyclones moving from the Adriatic to the Aegean Sea have become less frequent in January during the 1982–93 period, but with no change in December and February. However, according to the current analysis, the SOM index also suggests reduced winter disturbances in December, though in a lesser degree than in January. Significant correlations in Table 3 suggest that changes in the large-scale circulation associated with the AO strongly modulate the atmospheric conditions conducive to extreme precipitation events in the Balkan region, especially dry events at intraseasonal and interannual timescales. The circulation anomalies obtained from the SOM composites (Figs. 4, 5) are consistent with these results. The interannual relationship between the AO, the NAO, and the SOM indices is further illustrated in Fig. 7. Changes in the NAO and the AO indices (Figs. 7a,b) parallel to similar changes in the dry SOM index (Fig. 7c) during most of the study period. Although the NAO reached some historical values during the 1980–93 period, the wet and dry SOM indices are more strongly coupled to the AO. Interestingly, the largest positive anomalies of the wet SOM index at the beginning of the 1980s (Fig. 7d), which are reflected in positive precipitation anomalies in the northeastern corner of the study area ($P_{\_NE}$ in Fig. 7e), correspond to small values of the AO, but not negative. Although this is consistent with the sign of the correlation in Table 2, the weak sign of the AO (e.g., weak meridional flow) as compared to the strong wet SOM index in Fig. 7d, suggests that some anomalous wet

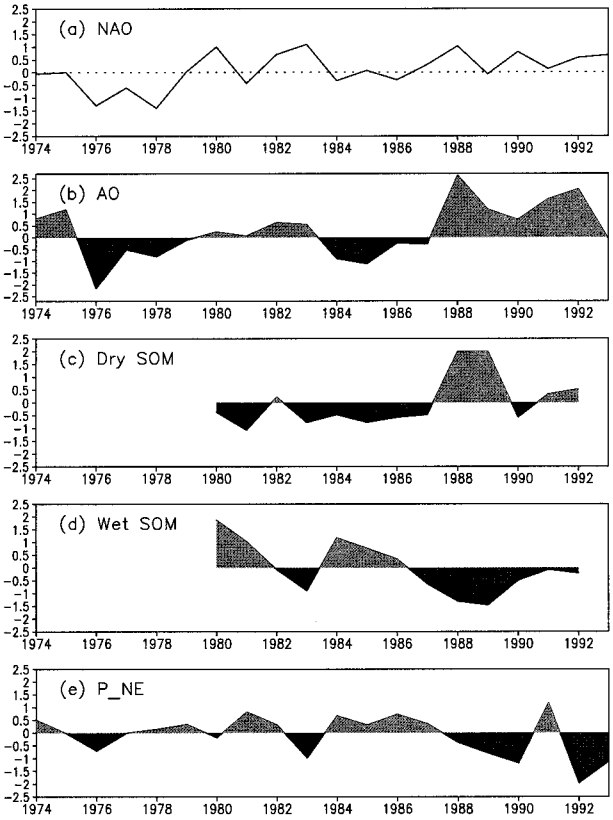

FIG. 7. Time series of wintertime (Dec–Feb) standardized values of (a) NAO index, (b) AO index, (c) dry SOM multivariate index for grid point 5, (d) wet SOM index for grid point 5, and (e) precipitation (P_{NE}) in the northwestern corner of the study area (averaged over grid points 5, 9, and 10 according to an RPCA).

TABLE 4. Average contribution (%) of the atmospheric predictors to explaining daily precipitation from a feed-forward ANN model. In the first row, the contribution is averaged over the 12 grid points in Fig. 1. The contribution of the predictors in grid point 11 and Bucharest are shown for comparison. Predictors are: 500–1000-hPa thickness (Th5_1), 700-hPa geopotential heights (Z7), 700-hPa specific humidity (q7), Earth's surface specific humidity (qo), SOM category (SOM_{CAT}), and the NAO index.

Predictor	Th5_1	Z7	q7	qo	SOM_{CAT}	NAO
All grids	34.8	32.7	24.04	4.00	2.53	1.93
Grid 11	40.20	30.58	20.95	3.95	3.20	1.14
Bucharest	48.86	27.10	17.59	2.26	2.03	2.15

conditions may need just a little relaxation of the large-scale zonal flow to enhance the synoptic (e.g., cyclonic disturbances) and local (e.g., moisture) forcings around the study area. A closer look at the large-scale atmospheric conditions during the 1981/82 winter (see, e.g., the NOAA–CIRES Climate Diagnostic Center web site) reveals that central Europe was dominated by weak meridional flow from the midtroposphere to the stratosphere, consistent with the sign of the teleconnection indices (Figs. 7a,b). However, at lower levels of the atmosphere there was a strong meridional flow and a clear split flow over central Europe from the Pyrenees to the Alps favoring the advection of moisture from the North Atlantic (through the Gulf of Biscay) to the Mediterranean. A low pressure system over the central Mediterranean also suggests the passage of shallow fronts over the Adriatic. These conditions are particularly notorious during January and February 1982. Accordingly, atmospheric conditions conducive to anomalous large precipitation events can be coupled to remote teleconnections (e.g., from 1984 to 1986); but there are other cases, such as during the 1981/82 winter, when the lower-tropospheric circulation and mesoscale forcings (e.g., moisture, shallow fronts) are the main driving controls of local precipitation. Thus, the SOM indices associated with extreme events are able to capture significant interannual differences, which confirms the adequate selection of controlling variables.

7. Derivation of daily precipitation from the ANNs

In this section, the performance of the feed-forward ANN model is shown in a very general form (i.e., averaged over all grid points). Daily results for Bucharest station and the nearest NCEP–NCAR grid point 11 (see Fig. 1) are also shown for comparison. In the last two sections, we demonstrated the usefulness of the SOM classification and the SOM multivariate indices for understanding the variability of extreme precipitation events at synoptic, intraseasonal and interannual timescales. The results from the ANN show that, of the atmospheric controls used, the most important predictors of daily precipitation are Th5_1, Z7, and q7, at both gridpoint and local scales, as indicated in Table 4. Although q0, the SOM_{CAT}, and the NAO index play minor roles in explaining daily precipitation, the physical consistency of the climate anomalies during extreme events, as well as their agreement with past works, proves the robustness of the other three predictors utilized (Table 4). The difference in the contribution of the specific humidities further supports the significant role of the midtropospheric moisture in precipitation, as discussed in section 5b.

Table 5 shows the feed-forward ANN model performance overall grid points. The empirical model explains in average 52% of the daily precipitation variability over the study area with a $\pm 6\%$ error ($\sigma - R^2$) due to the impact of model variability (i.e., from the bootstraping analysis). The model reproduces 75% of the daily variance (i.e., skill ratio between simulated and observed standard deviations). Looking only at the number of wet days (precipitation > 0.1 mm) correctly simulated as wet days by the model (first column in Table 6), it would be concluded that the simulation is good, as the model only misses one precipitation day. However, a different story is to capture the correct number of events in different precipitation ranges, and more difficult yet is to simulate the daily variance, especially that of extreme events. In this analysis, the model correctly simulates 6.1% of the top 10th percentile events as large events (e.g., above the 3.6 mm day^{-1} threshold in Table 5). A proportion of these events (>5 mm) is underestimated, as indicated by the negative percentage errors in Table 6. The increased underestimation toward the right-hand side tail of the distribution and overestimation of rainy days in the median range (1–5 mm) are a typical problem in precipitation downscaling. Large precipitation events are difficult to simulate not only because of the complexity and the variability of the physical processes involved, but also because of the scarcity of these events on a daily time frame. This is evident in Table 6, where the number of observed large events (>10 mm) are 60% less frequent than the median events (1–5 mm). Thus, one must realize that if the model has a sound set of predictors, then the problem at simulating large precip-

TABLE 5. Average model performance over all grid points: P_o and P_s are the mean observed and simulated precipitation (mm day^{-1}); adj R^2 is the adjusted explained variance due to bias from resampling; $\sigma - R^2$ is the standard error of R^2, rmse is the root-mean-square error in mm day^{-1}, skill ($=\sigma_s/\sigma_o$) is the ratio of skill between simulated and observed precipitation, Top10 (%) is the average number of rainy days in the top 10th percentile group correctly simulated by the model as large events (i.e., above the minimum threshold (mm day^{-1}) in this percentile range).

P_o	P_s	adj R^2	$\sigma - R^2$	Rmse	Skill	Top10	Threshold
1.46	1.48	0.52	0.06	1.77	0.75	6.1	3.6

TABLE 6. Average frequency distribution of wet days (larger > 0.1 mm) for simulated ($Freq_s$) and observed ($Freq_o$) precipitation events over all grid points. Precipitation range (P_{range}) is in mm. In parenthesis is the relative frequency of wet days (%) in each precipitation range. Diff is the residual between simulated ($Freq_s$) and observed ($Freq_o$) number of wet days. The residual error (%) indicates the percentage of days that are overestimated (+) and underestimated (−) by the model in each precipitation range. Total number of winter days is 1152. Median of wet days is 1 mm.

P_{range}	>0.1	[0.1–1)	[1–5)	[5–10)	[10–15)	[15–20)	≥20
$Freq_s$	183 (16)	28 (15.3)	98 (53.6)	37 (20.2)	12 (6.5)	5 (2.7)	3 (1.6)
$Freq_o$	184 (16)	24 (13.0)	80 (43.5)	49 (26.6)	18 (9.7)	8 (4.3)	7 (3.8)
Diff	−1	+4	+18	−12	−6	−3	−4
Error (%)	0.5	+16	+22.5	−24.5	−33.3	−37.5	−57

itation events is not only the lack of other local forcings (e.g., land surface processes, convection) in the model, but also the weak representation and large variability of these events, which in many cases are due to random processes. At the other end of the tail, though the model tends to overestimate the number of precipitation days in the 0.1–5 mm ranges, Fig. 8 illustrates that the simulation of dry cases is well captured.

As in Cavazos (1999), where a selection of circulation and humidity predictors contributed to derive a distinct gridpoint response in the wet El Niño and dry La Niña winters in northeastern Mexico and southeastern Texas, in the Balkan region the neural network model also discriminates dry from wet periods. This is illustrated in Fig. 8 with a comparison between simulated (P_s) and observed (P_o) precipitation in Bucharest station and in the nearest NCEP–NCAR grid point 11 for 3 yr of the 1981–93 period. In general, the synoptic events observed in Bucharest station (Fig. 8a) are well captured at gridpoint scale (Fig. 8b), but the amplitude of large events is much smaller in the grid point. This is reflected in the wettest winter in Bucharest (1983/84), which shows that observed precipitation in grid point 11 is 26 mm season^{-1} less than in Bucharest. The winter precipitation in an average winter (e.g., 1987/88) is similar at both local and gridpoint scales. In the driest year (1988/89) grid point 11 has 15 mm season^{-1} more precipitation than Bucharest. The neural network also tends to underestimate (overestimate) small (large) events as illustrated in the simulated results shown in Fig. 8. This may represent a major problem for area-averaged precipitation data, which is already biased by its spatial resolution. Nevertheless, the performance of the ANN is slightly better in grid point 11 than in Bucharest as indicated by the average measures of skill shown in Table 7, and in Fig. 8 for particular winters.

8. Summary and conclusions

This paper examines local and remote sources of variation associated with wintertime precipitation in the Balkan region with special emphasis of extreme events. By applying a SOM nonlinear classification technique and a feed-forward ANN to simulate daily precipitation over the study area, the usefulness of the SOM–ANN approach for identifying spatial and temporal atmospheric structures that contribute significantly to precipitation variability is demonstrated. Some of the climate anomalies characteristic of extreme conditions derived from the SOMs are physically consistent with past studies based on EOF and CCA techniques, and new results are documented here concerning the role of midtropospheric moisture during extreme events in the Balkan region. This work has attempted to determine the relevance of the circulation, and the thermal and moisture components on daily precipitation, as well as the role of the NAO and the AO indices on winter precipitation at intraseasonal and interannual timescales. Nine SOM climate modes are derived using the NAO index and an ensemble of time-lagged circulation (Z7), 500–1000-hPa thickness (Th5_1), and specific humidity (q0 and q7) controls. After the SOM classification, a categorical variable (SOM_{CAT}) is devised to condition each day according to the nine classes. SOM_{CAT} is then included as an additional variable in the input atmospheric predictors, which are used to simulate daily precipitation with the feed-forward ANN.

The climate anomalies derived from the SOM classification reveal that extreme precipitation events in the Balkans appear to be regulated by 1) changes in the intensity of the zonal flow and the polar vortex (e.g., AO), 2) changes in the location of the jet stream and associated storm tracks, and 3) changes in the midtropospheric moisture over the eastern Mediterranean. Past studies also document that atmospheric moisture plays an important role in the reinforcement and suppression of droughts and flood conditions in other regions (Giorgi et al. 1996; Seth and Giorgi 1998). For the Mediterranean in particular, Hurrell (1995) reports a decrease (increase) in the moisture transport during the high (low) phase of the NAO.

The synoptic signatures characteristic of extreme conditions over the Balkans are linked to the large-scale circulation anomalies associated with remote controls through an SOM multivariate index based on the frequency of days with atmospheric conditions conducive to wet and dry events. The SOM index reveals an increasing trend of dry conditions during January and, in a lesser degree, in December of the 1980–93 period; this trend seems to be modulated by changes in the zonal circulation associated with the AO. Dry conditions dominated almost half of the period, from 1988 to 1992,

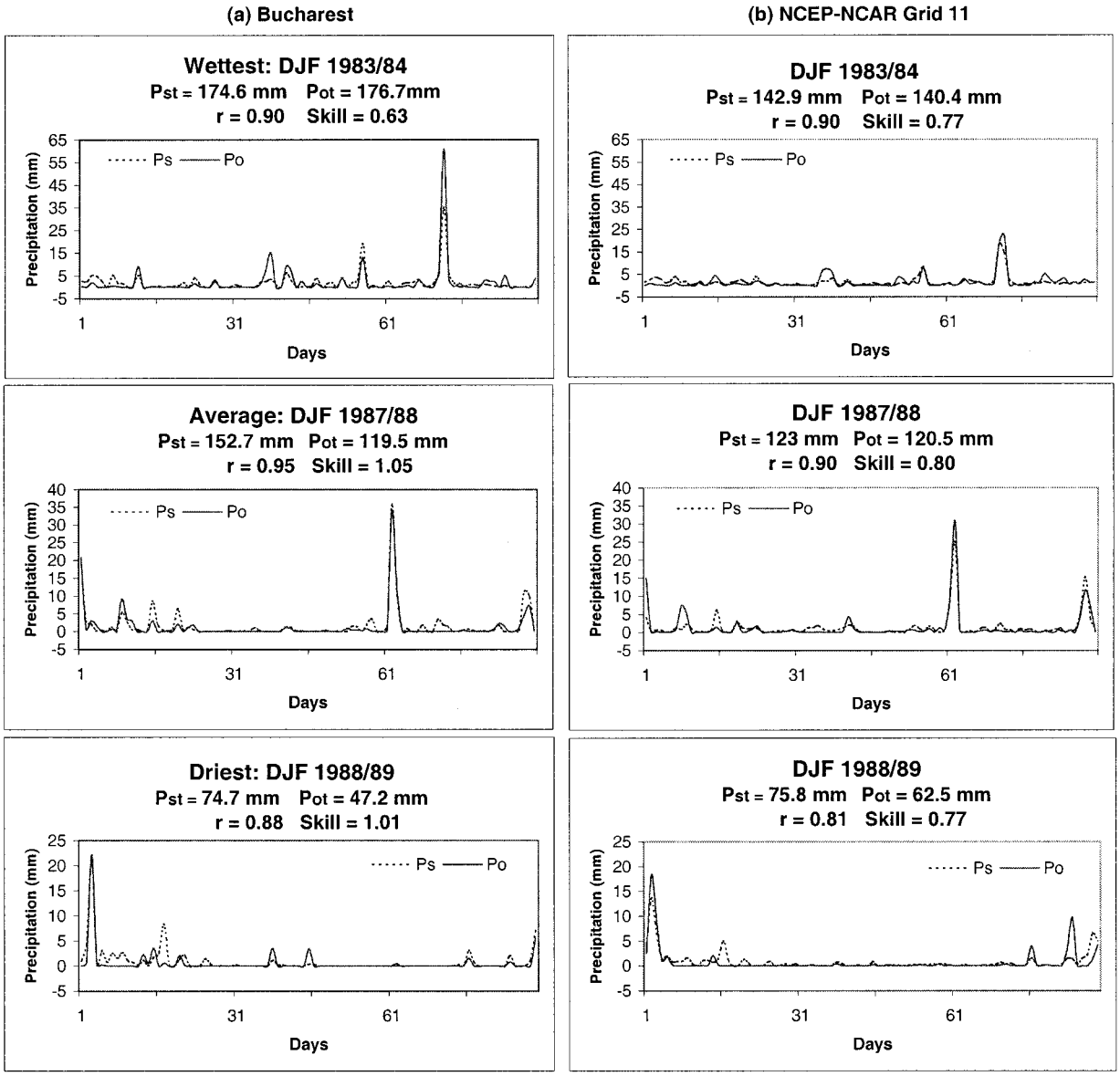

FIG. 8. Daily simulated (P_s) and observed (P_o) precipitation (mm) for (a) Bucharest station and (b) nearest NCEP–NCAR grid point 11 for the wettest, average, and driest winter in Bucharest during the 1981–93 period. Total simulated (P_{st}) and total observed (P_{ot}) winter precipitation is also shown, as well as correlation (r), and ratio of simulated and observed standard deviations (skill). Skill >1 means that simulated standard deviation $>$ observed standard deviation. Precipitation is derived from the full ANN model.

when the NAO, the AO, and the dry-SOM indices were persistently positive; but the NAO plays little role in the wintertime precipitation in the Bulgarian region. The eastward-shifted anomalous ridge over central and eastern Europe and the deepened circumpolar vortex seen in the climate anomalies of the dry SOM (Fig. 4b) bear a strong resemblance to the 500-hPa geopotential height anomaly field characteristic of the high phase of the AO documented by TW98. In a most recent paper, Thompson et al. (2000), also report a negative trend in the wintertime precipitation in the Balkans during the 1968–96 period and indicate that it is associated with the high index polarity of the AO. Thompson et al. argue that the NAO is a regional response of the much larger-scale AO pattern. Anomalous wet conditions are less strongly associated with the AO than dry events; under certain

TABLE 7. ANN model performance for Bucharest station and the nearest NCEP–NCAR grid point 11. Measures of performance as in Table 5.

	P_o	P_s	R^2	adjR^2	$\sigma - R^2$	Rmse	Skill
Bucharest	1.29	1.31	0.64	0.52	0.11	2.25	0.72
Grid point 11	1.08	1.12	0.67	0.54	0.04	1.67	0.77

conditions, low-level synoptic disturbances occur when there is a little relaxation of the large-scale midtropospheric zonal flow.

In spite of the subjective selection of the number of SOM modes, this study has demonstrated that the SOM climate anomalies typical of extreme conditions are physically consistent with local and remote controls. Moreover, the SOM technique offers the advantage of disclosing multiple sources of variation at different timescales. Even 5-day or 10-day means may mix extreme with nonextreme conditions, and thus, smooth the atmospheric response. Instead, the SOM classifies the days in a nonlinear form maximizing different atmospheric conditions in each mode. Then, it is possible to run a sequence of synoptic pictures with time frames of different lengths according to the states of the atmosphere in a particular season, which may help to disclose the development and possible causes of extreme events.

The results from the ANN show that, of the variables used, the most important controls of daily precipitation at gridpoint and local scales are Th5_1, Z7, and q7 (Tables 2, 5). Interestingly, the SOM_{CAT} does not play a significant role in explaining daily precipitation variance. It seems that the thickness and the midtropospheric circulation and moisture carry much more information than the simple categorical variable associated with the climate modes. The ANN model is able to capture synoptic events at both local and gridpoint scales and differentiates well dry from wet periods, but tends to overestimate (underestimate) small (large) events (Fig. 8). The tendency of the ANN to better resolve synoptic and low-frequency events than daily events has been documented before (e.g., Wilby et al. 1998; Cavazos 1999). By comparison, Bucharest and its nearest NCEP–NCAR gridpoint exhibit similar amounts of observed precipitation and variance at monthly and seasonal timescales, but larger values are seen in Bucharest at daily timescales. At interannual timescales, the gridpoint overestimates (underestimates) dry (wet) winters. This can be a major disadvantage for area-averaged precipitation data because precipitation is already biased by its spatial resolution, then the ANN adds an extra bias. Nevertheless, the performance of the ANN is slightly better at gridpoint scale. Thus, analyses employing area-averaged precipitation are still useful for diagnostic studies and for regional and river-basin applications, and even for downscaling purposes if the synoptic events and a large proportion of the variance are captured by the model. For daily timescales and large events results could be possibly improved by: 1) using longer time series of predictors and finer resolution atmospheric controls and 2) including a stochastic element into the model. If the model has a sound set of predictors, then the problem at simulating large events is not only the lack of other local forcings (e.g., land surface processes, convection), but also the weak representation and large variability of these events, which in some cases are due to random processes. Hewitson (1998) suggests that under such circumstances, the residual variance may be considered decoupled from the mesoscale forcing, and represented by stochastic procedures. Hewitson (1998) employed the residual variance distribution associated with several SOM climate states to account exclusively for the random precipitation processes not captured by the predictors in an ANN model. Many studies have employed stochastic techniques to solve the problem of the unexplained variance (e.g., Katz and Parlange 1996; Hewitson 1998; Wilby et al. 1998; Crane and Hewitson 1998) and the variability of extreme events (e.g., Katz and Brown 1992; Mearns et al. 1997; Katz 1998). Stochastic modeling can account for the essential features of the precipitation process; namely, its intermittency, the tendency of wet and dry spells to persist, and the positively skewed distribution of intensity (Katz 1998). Nevertheless, stochastic techniques have their own limitations. Most weather generators do not reproduce the autocorrelation structure of the variables on interannual timescales nor the change in the persistence of extreme events (e.g., Mearns et al. 1997). Mearns et al. (1997) suggest that one way of capturing these changes is to condition the weather generator on different climate states. The results documented in this analysis suggest that the SOM climate states could be used to condition stochastic simulations. Accordingly, even though the SOM_{CAT} is not a significant predictor of daily precipitation, the use of the SOM technique is strongly justified in diagnostic and downscaling studies as it helps to identify significant spatial and temporal structures in the data and, as Hewitson (1998) notes, can be used in combination with stochastic procedures to account for the unexplained variance.

Acknowledgments. I am thankful to B. C. Hewitson, G. C. Knight, and R. G. Crane for their comments and suggestions, and to T. P. Mitchell for pointing out the TW98 paper and the source of the AO data at JISAO. I am also grateful to F. Zwiers and two other reviewers whose suggestions and critiques helped clarify a number of points. This research was partially supported by the Center for Integrated Regional Assessment (CIRA) of the Pennsylvania State University through an NSF grant to G. C. Knight, and by the University of Cape Town through a grant from the Foundation for Research Development to B. C. Hewitson. The CRU05 dataset was provided by David Viner and the Climate Impacts LINK Project of the Climatic Research Center of the University of East Anglia to the Pennsylvania State University/University of Arizona/Bulgarian Academy of Science consortium. Precipitation from Bucharest station was provided by M. Caraman of the National Institute of Meteorology, Hydrology and Water Management, Bucharest through A. Busuioc. I am grateful to D. P. Brown for arranging the CRU05 dataset to my particular needs.

REFERENCES

Barnston, A. G., and R. E. Livezey, 1987: Classification, seasonality, and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.*, **115,** 1083–1126.

Bulgarian Academy of Sciences, 1973: *Atlas Narodna Republika Bulgaria* (*Atlas of the People's Republic of Bulgaria*). Institute of Geography, Bulgarian Academy of Sciences, 168 pp.

Busuioc, A., and H. von Storch, 1996: Changes in the winter precipitation in Romania and its relation to the large-scale circulation. *Tellus,* **48A,** 538–552.

——, ——, and R. Schnur, 1999: Verification of GCM-generated regional seasonal precipitation for current climate and of statistical downscaling estimates under changing climate conditions. *J. Climate,* **12,** 258–272.

Cavazos, T., 1999: Large-scale circulation anomalies conducive to extreme events and derivation of daily rainfall in northeastern Mexico and southeastern Texas. *J. Climate,* **12,** 1506–1523.

Corte-Real, J., X. Zhang, and X. Wang, 1995: Large-scale regimes and surface climatic anomalies over the Mediterranean. *Int. J. Climatol.,* **15,** 1135–1150.

Crane, R. G., and B. C. Hewitson, 1998: Doubled CO_2 precipitation changes for the Susquehanna Basin: Downscaling from the GENESIS General Circulation Model. *Int. J. Climatol.,* **18,** 65–76.

Flocas, A. A., 1984: The annual and seasonal distribution of fronts over central-southern Europe and the Mediterranean. *Int. J. Climatol.,* **4,** 255–267.

Furlan, D., 1977: The climate of southeast Europe. *Climates of Central and Eastern Europe,* Vol. 6, *World Survey of Climatology,* C. C. Wallén, Ed., Elsevier, 185–223.

Giorgi, F., L. O. Mearns, C. Shields, and L. Mayer, 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. Climate,* **9,** 1150–1162.

Goodman, P. H., 1996: NevProp software Version 3. University of Nevada.

Halpert, M. S., and G. D. Bell, 1997: Climate assessment for 1996. *Bull. Amer. Meteor. Soc.,* **78,** S1–S49.

Hewitson, B. C., 1998: Deriving regional climate change scenarios from GCMs. Water Research Commission Report K5/751, Pretoria, South Africa, 19 pp. [Available online at http://www.wrc.org.za]

Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation regional temperatures and precipitation. *Science,* **269,** 676–679.

——, and H. van Loon, 1997: Decadal variations associated with the North Atlantic Oscillation. *Climatic Change,* **36,** 301–326.

Janowiak, J. E., G. Arnold, C. R. Knodragunta, R. E. Livezey, and G. J. Huffman, 1998: A comparison of the NCEP–NCART reanalysis precipitation and the GPCP rain gauge–satellite combined dataset with observational error considerations. *J. Climate,* **11,** 2960–2979.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Katz, R. W., 1998: Extreme value theory for precipitation: Implications for climate change. *Ninth Symp. on Global Change Studies, Joint Session, 14th Conf. on Probability and Statistics in the Atmospheric Sciences,* Phoenix, AZ, Amer. Meteor. Soc., J11–J15.

——, and M. B. Parlange, 1996: Mixtures of stochastic processes: Applications to statistical downscaling. *Climate Res.,* **7,** 185–193.

——, and B. G. Brown, 1992: Extreme events in a changing climate: Variability is more important than averages. *Climatic Change,* **21,** 289–302.

Knight, C. G., S. B. Velev, and M. P. Staneva, 1995: The emerging water crisis in Bulgaria. *Geo. J.,* **35.4,** 415–423.

Kohonen, T., 1982: Self-organized formation of topologically correct features maps. *Bio. Cyber.,* **43,** 59–69.

——, 1995: *Self-Organizing Maps.* Springer Series in Information Sciences, Vol. 30, Springer-Verlag, 362 pp.

——, J. Hynninen, J. Kangas, and J. Laaksonen, 1995: SOM_PAK, The self-organizing map program package Version 3.1. Helsinki University of Technology, Laboratory of Computer and Information Science. [Available online at: http://www.cis.hut.fi/nnrc/ and ftp://cochlea.fi.]

McGinnis, D. L., 1994: Predicting snowfall from synoptic circulation: A comparison of linear regression and neural networks. *Neural Nets: Applications in Geography,* B. Hewitson and R. G. Crane, Eds., Kluwer Academic Publishers, 79–99.

Mearns, L. O., C. Rosenzweig, and R. Goldberg, 1997: Mean and variance change in climate scenarios: Methods, agricultural applications, and measures of uncertainty. *Climatic Change,* **35,** 367–396.

New, M. G., and Coauthors, 2000: Representing twentieth-century space–time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *J. Climate,* in press.

Palutikof, J. P., M. Conte, J. C. Mendes, C. M. Goodess, and F. E. Santo, 1996: Climate and climatic change. *Mediterranean Desertification and Land Use,* J. Brandt and J. B. Thornes, Eds., Wiley, 43–86.

Sahsamanoglou, H., T. Makrogiannis, N. Hatzianastasiou, and N. Rammos, 1997: Long term change of precipitation over the Balkan Peninsula. *Eastern Europe and Global Change,* Ghazi et al., Eds., European Commission, 111–124.

Seth, A., and F. Giorgi, 1998: The effects of domain choice on summer precipitation simulation and sensitivity in a regional climate model. *J. Climate,* **11,** 2698–2712.

Tangang, F. T., G. M. Tang, A. H. Monahan, and W. H. Hsieh, 1998: Forecasting ENSO events: A neural network-extended EOF approach. *J. Climate,* **11,** 29–41.

Thompson, D. W. J., and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.,* **25** (9), 1297–1300.

——, ——, and G. C. Hegerl, 2000: Annular modes in the extratropical circulation. Part II: Trends. *J. Climate,* **13,** 1018–1036.

Trigo, R. M., and J. J. Palutikof, 1999: Simulation of daily temperatures for climate change scenarios over Portugal: A neural network model approach. *Climate Res.,* **13,** 45–59.

Velev, S. B., 1996: Is Bulgaria becoming warmer and drier? *Geo. J.,* **40.4,** 363–370.

von Storch, H., E. Zorita, and U. Cubash, 1993: Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *J. Climate,* **6,** 1161–1171.

Weichert, A., and G. Bürger, 1998: Linear versus non-linear techniques in downscaling. *Climate Res.,* **10,** 83–93.

Wigley, T. M. L., and G. Farmer, 1982: Climates of eastern and Near East. *Paleoclimates, Paleoenvironments and Human Communities in the Eastern Mediterranean Region in Later Prehistory,* BAR International Series, Vol. 133, British Archeological Reports, 3–37.

Wilby, R. L., T. M. L. Wigley, D. Conway, P. D. Jones, B. C. Hewitson, J. Main, and D. S. Wilks, 1998: Statistical downscaling of general circulation model output: A comparison of methods. *Water Resour. Res.,* **34,** 2995–3008.

Winkler, J. A., J. P. Palutikof, J. A. Andresen, and C. M. Goodess, 1997: The simulation of daily temperature time series from GCM output. Part II: Sensitivity analysis of an empirical transfer function methodology. *J. Climate,* **10,** 2514–2535.

Zorita, E., V. Kharin, and H. von Storch, 1992: The atmospheric circulation and sea surface temperature in the North Atlantic area in winter: Their implication and relevance for Iberian precipitation. *J. Climate,* **5,** 1097–1108.