Generating Rainfall and Temperature Scenarios at Multiple Sites: Examples from the Mediterranean

J. P. PALUTIKOF, C. M. GOODESS, S. J. WATKINS, AND T. HOLT

Climatic Research Unit, University of East Anglia, Norwich, United Kingdom

(Manuscript received 21 May 2001, in final form 26 April 2002)

ABSTRACT

A statistical downscaling methodology was implemented to generate daily time series of temperature and rainfall for point locations within a catchment, based on the output from general circulation models. The rainfall scenarios were constructed by a two-stage process. First, for a single station, a conditional first-order Markov chain was used to generate wet and dry day successions. Then, the multisite scenarios were constructed by sampling from a benchmark file containing a daily time series of multiple-site observations, classified by season, circulation weather type, and whether the day is wet or dry at the reference station. The temperature scenarios were constructed using deterministic transfer functions initialized by free atmosphere variables. The relationship between the temperature and rainfall scenarios is established in two ways. First, sea level pressure fields define the circulation weather types underpinning the rainfall scenarios and are used to construct predictor variables in the temperature scenarios. Second, separate temperature transfer functions are developed for wet and dry days.

The methods were evaluated in two Mediterranean catchments. The rainfall scenarios were always too dry, despite the application of Monte Carlo techniques in an attempt to overcome the problem. The temperature scenarios were generally too cool. The scenarios were used to explore the occurrence of extreme events, and the changes predicted in response to climate change, taking the example of temperature. The nonlinear relationship between changes in the mean and changes at the extremes was clearly demonstrated.

1. Introduction

The principal source of information for climate change impacts analyses is the general circulation model (or GCM). Using data taken directly from GCMs is generally unsatisfactory in this context. First, the model generates output at the coordinates of a coarse three-dimensional grid, with a typical resolution of 2.5° lat by 3.75° lon (the resolution, e.g., of the U.K. Hadley Centre Unified Model). Second, the models may only be able to satisfactorily reproduce the characteristics of surface meterological variables such as temperature or rainfall at the hemispheric or global scale rather than the required regional or local scale (Palutikof et al. 1997; Goodess and Palutikof 1998).

There is therefore a scale mismatch between the requirements of the impacts analyst and the availability of useful information from GCMs. This may be overcome by downscaling. Dynamical downscaling nests a higher resolution regional climate model within a GCM (McGregor 1997; Giorgi and Mearns 1999). Statistical downscaling builds statistical models to link mesoscale free atmosphere predictors (most commonly, 500-hPa geopotential height and sea level pressure) to local/regional-scale surface meteorological variables such as temperature and precipitation. The underlying assumption is that GCMs are better able to reproduce these mesoscale features than surface variables (Winkler et al. 1997). The two approaches are compared by Mearns and Giorgi (1999) and Murphy (2000).

Statistical downscaling is used here [see reviews by Wilby et al. (1998) and Wilks and Wilby (1999)]. The success of a downscaling technique is usually judged in two stages by its ability to simulate the predictand variable(s), first, when initialized with an independent set of observed data and, second, when initialized with climate model data for the present. Such criteria consider technical performance alone. Here, we approach the task of downscaling within the context of the problem to be addressed, thus adding a further measure of success: the extent to which the needs of the impacts analyst have been met.

2. The problem

Mediterranean Desertification and Land Use (MEDALUS) was an interdisciplinary project funded by the European Commission to investigate the time-depen-

Corresponding author address: Dr. J. P. Palutikof, Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, United Kingdom.
E-mail: j.palutikof@uea.ac.uk

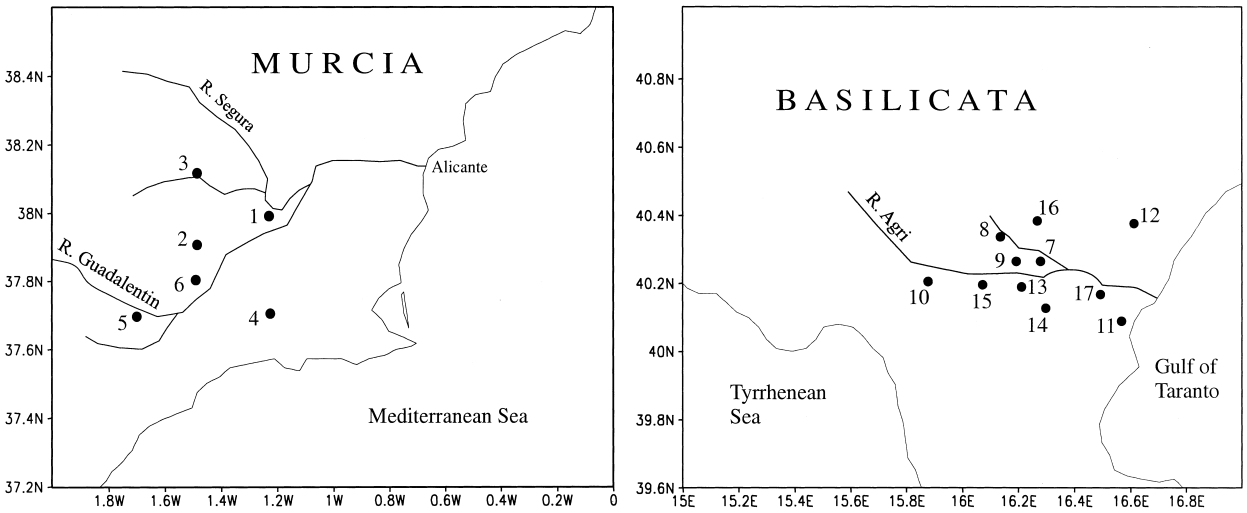

FIG. 1. Location of the stations used in downscaling. (left) Guadalentin Basin, Spain. (right) Agri Basin, Italy.

dent evolution of desertification processes in Mediterranean countries of the European Union under changing socioeconomic and climatic conditions. A principal role of the climatologists was to generate scenarios of climate change at high temporal and spatial resolutions for use by hydrological modelers. Although no formal specification was prepared, the requirements may be summarized as follows.

1) Within a catchment, scenarios should be generated for a well-distributed network of around six to seven sites for rainfall.
2) For at least one of these sites, scenarios of maximum and minimum temperature should also be generated.
3) Scenarios should be at the daily scale.
4) The relationships between the rainfall scenarios at the different sites, and between the rainfall and temperature scenarios at the single site, should be consistent and coherent at the daily scale.
5) The scenarios should be explicitly evaluated in terms of their ability to simulate the occurrence of extreme events.

Here, we present and test in two Mediterranean catchments a family of scenario construction methods that meet these five criteria.

3. The catchments

Ideally, a statistical downscaling technique should be portable between areas. This can only be demonstrated by application in at least two, preferably geographically distinct, locations. Here, scenario sets are developed for two catchments. The first is the Guadalentin Basin in southeast Spain (see Fig. 1 and Table 1). Rainfall scenarios were developed for six sites, and temperature scenarios for three of these. Table 1 demonstrates the geographical diversity of the sites, with altitudes ranging from 75 m at Alcantarilla to 760 m at Alhama de Murcia. The maximum separating distance is around 55 km (Lorca to Alcantarilla). The second catchment is the Agri Basin in southern Italy. Eleven sites were used for rainfall scenario construction, as shown in Table 1, and temperature scenarios were constructed for three of these. Heights range between 909 m, at Stigliano on the northern slopes of the middle catchment, and sea level at Nova Siri Scalo. The maximum separating distance (Moliterno to Nova Siri Scalo) is around 70 km.

The Guadalentin is situated in southeast Spain. Westerly flow from the Atlantic may extend across the Iberian Peninsula to influence weather conditions over the catchment, although intense rainfall events, particularly in autumn, are more likely to be associated with easterly flow over the warm Mediterranean Sea (Linés Escardó 1970; Goodess and Palutikof 1998). The location of the Agri Basin, in southern Italy, is such that cyclonic disturbances tracking from the Gulf of Genoa are the dominant rainfall-producing mechanism (Trigo et al. 1999). From Table 1, annual rainfall in the Agri Basin is more than double that in the Guadalentin. Whereas in the Guadalentin the average number of wet days per year seldom exceeds 40, in the Agri all sites have an average of at least 80 wet days per year.

Thus, applying the downscaling techniques in the Guadalentin and the Agri permits evaluation of performance under different meterological conditions, while topography is, to a first approximation, held constant. This is important in the Mediterranean region where the complex orography plays a major role in determining climate variations with height and exposure (Agnew and Palutikof 2000).

4. The data

The observations for the Guadalentin have been extensively quality controlled, and are of high quality for all sites over the 30-yr period 1958–87 used here (Brandt

TABLE 1. Characteristics of the stations used in the downscaling. Stations marked with an asterisk were used for temperature and rainfall scenario development, otherwise rainfall only.

Fig. 1 number	Station names	Abbrev.	Lat	Lon	Alt (m)	Annual rain days	Annual rainfall (mm)	Mean annual TMAX (°C)	Mean annual TMIN (°C)
Guadalentin Basin									
1	Alcantarilla*	ALC	38.0°N	1.2°W	75	48	289	23.6	11.0
2	Alhama de Murcia*	ADM	37.9°N	1.5°W	760	44	418	19.1	8.9
3	Embalse de Cierva	EDC	38.1°N	1.5°W	400	31	266		
4	Fuente Alamo	FA	37.7°N	1.2°W	200	34	272		
5	Lorca*	LOR	37.7°N	1.7°W	335	38	234	21.3	11.3
6	Totana	TOT	37.8°N	1.5°W	200	28	293		
Agri Basin									
7	Aliano	ALI	40.3°N	16.2°E	497	83	748		
8	Corleto Perticara	COR	40.4°N	16.0°E	746	126	803		
9	Missanello	MIS	40.3°N	16.3°E	566	88	804		
10	Moliterno*	MOL	40.2°N	15.9°E	879	137	1134	16.6	8.2
11	Nova Siri Scalo*	NSS	40.1°N	16.6°E	2	95	550	21.2	11.7
12	Pisticci	PIS	40.4°N	16.6°E	364	90	604		
13	Roccanova	ROC	40.2°N	16.2°E	654	88	725		
14	Senise	SEN	40.2°N	16.3°E	330	122	744		
15	S. Martino Agri	SMA	40.2°N	16.1°E	661	99	791		
16	Stigliano*	STI	40.4°N	16.2°E	909	87	811	16.1	8.3
17	Tursi	TUR	40.3°N	16.5°E	348	90	730		

and Thornes 1993). In the Agri, the common period for which both temperature and rainfall data are available with few missing values is 1956–76. This much shorter period is a problem for the multisite scenario construction, as will be seen. The data have been checked for major inconsistencies although the checks were less rigorous than for the Guadalentin.

The rainfall scenarios require sea level pressure data as a basis for circulation typing, and the temperature-scenario transfer functions use predictor variables based on sea level pressure and 500-hPa geopotential height. These are taken from the National Meteorological Center (NMC, now known as the National Centers for Environmental Prediction) operational analyses, available twice daily (0000 and 1200 UTC) at equally spaced points on a polar stereographic projection with a horizontal resolution of approximately 380 km. Discontinuities in the data have been identified (Trenberth and Olson 1988), and the years used for calibration and validation of the scenarios are selected with the need for homogeneity in mind. Thus, 1961–64 and 1977 are omitted from all NMC-based analyses (see Table 2) be-

TABLE 2. Details of observation years used for scenario construction and validation.

	Guadalentin	Agri
Rainfall		
Reference scenarios		
a. Calibration	1958–60; 1965–76; 1978–87	1956–60; 1965–76; 1978–88
b. Validation of method	Cross*	Cross*
c. Validation of GCM-generated scenarios	Cross*	Cross*
Multisite scenarios		
a. Calibration	1958–87	1956–76
b. Validation of method	Highly constrained—no validation required	Highly constrained—no validation required
c. Validation of GCM-generated scenarios	By comparison of time series characteristics with 1958–87 observed	By comparison of time series characteristics with 1956–76 (56–88) observed
Temperature		
a. Calibration	1975–84	
b. Validation of method	Simple**: 1973–74 and 1985–87	
c. Validation of GCM-generated scenario	Simple**: 1975–84	

* Cross validation = multiple calibration runs that successively exclude 1 yr from the scenario development and validate on that year (see Trigo and Palutikof 1999, for an explanation).

** Simple validation = one calibration period and one validation period.

TABLE 3. General structure of the multisite, multivariable downscaling approach.

	Reference rainfall scenario	Multisite rainfall scenario	Temperature scenario
Input	Season Weather type	Season Weather type	Season Predictor variables based on SLP, Z500, and THK
	Precipitation occurrence on day $i - 1$	Precipitation occurrence on day i at reference station	Precipitation occurrence on day i at reference station
Model type	Weather generator, conditional on season, weather type, and precipitation occurrence on day $i - 1$	Nonparametric, based on uniform random sampling (with replacement) from benchmark file of potential scenario days, classified by season, weather type, and precipitation occurrence	One of 20 linear transfer functions, depending on season (or all data), precipitation occurrence, and predictand variable
Output	Precipitation occurrence on day i at reference station (input to all scenario types)	Ten years of multisite daily rainfall for three scenario decades: 1970–79 2030–39 2090–99	Ten years of daily TMAX and TMIN for three scenario decades: 1970–79 2030–39 2090–99

SLP = sea level pressure.
Z500 = 500-hPa geopotential height.
THK = 1000–500 hPa geopotential height difference.

cause of the large number of missing data in those years. In addition, in the NMC 500-hPa geopotential height data used for the Guadalentin, we found a major discontinuity between 1972 and 1973, which we suspect may reflect a data problem. The decade 1965–74 had originally been selected for validation of the temperature scenarios, but following recognition of this problem the five years 1973–74 and 1985–87 were used instead.

The GCM employed here is the second Hadley Centre Coupled Model (HadCM2). It is a coupled ocean-atmosphere model. Three runs were performed: control, greenhouse gas forcing only, and greenhouse gas + sulfate aerosol forcing (Johns et al. 1997). The latter experiment is used here. It is forced between 1861 and 1990 with historically recorded atmospheric concentrations of greenhouse gases and sulfate aerosols, and between 1991 and 2099 with a constant increase of 1% yr^{-1} in effective atmospheric carbon dioxide and a representation of the cooling effect of sulfate aerosols. Here, we construct scenarios for the decades 1970–79 (for validation by comparison with observations and as a present-day baseline), 2030–39 (within the planning horizon, e.g., of utility companies) and 2090–99, the last decade of the experiment.

5. The downscaling models

The complete downscaling methodology consists of three individual models, as shown in Table 3. The first generates a time series of precipitation occurrence, known as the reference rain day scenario, using a conditional weather generator. Wet and dry day occurrence then forms an input variable to the downscaling models for both the multisite rainfall scenarios and the temperature scenarios, helping to ensure consistency between them. The multisite rainfall scenarios are constructed by conditionally sampling (according to season, circulation weather type, and precipitation occurrence) from a benchmark file of rainfall observations. The temperature downscaling models are linear transfer functions in which the predictor variables are derived from sea level pressure and 500-hPa geopotential height fields. Using sea level pressure as a predictor variable for both the rainfall and temperature scenarios should again help to ensure stability of interscenario relationships.

a. The reference rain day scenario

The first step is to define a reference station around which to focus the downscaling. No definitive criteria exist for the choice of reference station. In the relatively dry Guadalentin, the site with the largest number of rain days was selected, Alcantarilla. In the Agri, a site from the middle catchment was chosen, Missanello.

At the reference sites, scenarios of rainday occurrence are developed based on a circulation-typing approach combined with a conditional first-order Markov chain to describe wet day/dry day probabilities. Using a first-order Markov chain apparently limits the memory of the rain day occurrence model to 1 day. Without circulation typing, Wilks (1999b) and Hayhoe (2000) use higher-order Markov chains to increase model memory. However, using a circulation-typing approach should make this unnecessary, because the circulation weather types (CWTs) themselves possess persistence (Conway and Jones 1998).

Fourteen CWTs were predefined from flow and vorticity parameters calculated from sea level pressure data (Goodess and Palutikof 1998). On inspection, combi-

TABLE 4. The circulation weather type (CWT) groups for the Guadalentin and Agri.

The 14 original CWTs	Abbrev.	CWT groups	
		Guadalentin	Agri
Rotational types:			
1. Cyclonic	C	C+HYC	C+HYC
2. Hybrid cyclonic	HYC		
3. Unclassified cyclonic	UC	UC	UC
4. Anticyclonic	A	A+HYA	A+HYA
5. Hybrid anticyclonic	HYA		
6. Unclassified anticyclonic	UA	UA	UA
Directional types:			
7. North	N	N	N
8. Northeast	NE	NE	NE
9. East	E	E+SE	E+SE
10. Southeast	SE		
11. South	S	S+SW	S+SW
12. Southwest	SW		
13. West	W	W+NW	W
14. Northwest	NW		NW

nations could be made on the basis of similarities in the underlying sea level pressure patterns and rain day occurrence (Goodess 2000), giving nine CWT groups for the Guadalentin and ten for the Agri (see Table 4).

The probability of day t being wet or dry depends on (a) the CWT of day t and (b) whether day $t - 1$ was wet or dry (following Conway et al. 1996; Conway and Jones 1998; the B-Circ method of Wilby et al. 1998). For n CWTs, to calculate the probability $\Pr\{A_w\}$ that day t is wet when day $t - 1$ is wet,

$$\Pr(A_w) = \sum_{i=1}^{i=n} \Pr(A_w | w\mathrm{CWT}_i)\Pr(w\mathrm{CWT}_i), \quad (1)$$

and when day $t - 1$ is dry,

$$\Pr(A_w) = \sum_{i=1}^{i=n} \Pr(A_w | d\mathrm{CWT}_i)\Pr(d\mathrm{CWT}_i), \quad (2)$$

where $d\mathrm{CWT}_i$ is the occurrence of a day of CWT type i preceded by a dry day, and $w\mathrm{CWT}_i$ is the occurrence of a day of CWT type i preceded by a wet day. This structure is different from that described by Hay et al. (1991), Wilby et al. (1994), and Goodess and Palutikof (1998), in which the CWT of day t is conditional upon the CWT of day $t - 1$, and the probability of rain on day t is conditional on the predicted CWT:

$$\Pr(\mathrm{CWT}_{i,t}) = \sum_{i=1}^{i=n} \Pr(\mathrm{CWT}_{i,t} | \mathrm{CWT}_{i,t-1}) \times \Pr(\mathrm{CWT}_{i,t-1}) \quad \text{and} \quad (3)$$

$$\Pr(A_w) = \Pr(A_w | \mathrm{CWT}_{i,t})\Pr(\mathrm{CWT}_{i,t}). \quad (4)$$

In the application stage, these authors construct the circulation-type transitional probability matrix (TPM) from GCM output, and take the rainfall probabilities from observations. An advantage is that, once the TPM has been constructed, the downscaling model can be run for very long periods of time, without reference to the GCM simulation. However, the need to tie together the rainfall and temperature scenarios, through the sea level pressure fields common to the construction of both, precludes the use of this approach here. Hence, the length of the scenarios that can be generated is constrained by the GCM experiment. A relatively stable snapshot of the climate can be obtained if the scenario is constructed using 10 yr of GCM data to initialize the downscaling models. A longer period will introduce trends in time caused by increasing atmospheric carbon dioxide concentrations.

We use NMC sea level pressure data in the calibration stages to carry out the circulation typing, and HadCM2 output in the application stages. While NMC data are available twice daily at 0000 and 1200 UTC, HadCM2 output was archived on a daily basis, where the daily value is the mean of all 10-min values generated on that day. There is no perfect solution to this mismatch. Here, we used the midday sea level pressure from the NMC reanalyses to establish circulation type/rain day occurrence relationships. This ensures that CWTs in the observations at least will be accurately determined. Future GCM simulations are likely to archive data at a higher temporal resolution, which will make it possible to determine modeled CWTs with respect to a single point in time.

b. The multisite rainfall scenarios

To construct the rainfall scenarios, daily observations at all the required stations are merged into a single file, known as the benchmark file. Each day in the benchmark file is a potential scenario day. The weather generator is conditional upon season (because circulation type–rain day occurrence relationships vary by season), CWT, and whether the day at the reference site is wet or dry.

So, each potential scenario day is categorized according to this threefold classification. Then, to construct a scenario, we take the reference rain day scenario, and step through it day by day. Each day can be assigned to a class based on season, CWT, and whether it is wet or dry at the reference station. Once the class is determined, a uniform random number generator is used (with replacement) to select one potential scenario day from the same class in the benchmark file (Conway and Jones 1998; Conway et al. 1996). By repeating this process for each day, a multisite scenario is constructed. The scenarios are 10 yr in length and are constructed from 30 yr of data in the Guadalentin, and 21 yr in the Agri. It is implicitly assumed that the 10-yr scenarios adequately sample the observed time series, which in turn are sufficiently long to capture the full potential range of variability.

A number of authors have taken the nonparametric route of using historical observations directly as a basis for creating multisite scenarios with correct spatial dependencies. Zorita et al. (1995) and Cubasch et al. (1996) used an analog approach, in which generated rainfall was set equal to that observed on the day in the historical record with the most similar large-scale atmospheric circulation characteristics. This is essentially deterministic. Others have sought to increase the probabilistic element (Brandsma and Buishand 1998; Rajagopalan and Lall 1999) by creating a set of nearest neighbors. Given the multisite precipitation characteristics on day $t - 1$, they find the k nearest neighbors (where k is typically 20), and then resample from the successors to these nearest neighbors in order to obtain the precipitation (and other weather variable) characteristics for day t. As here, Brandsma and Buishand (1998) take into account circulation characteristics, thereby substantially improving their results, whereas Rajagopalan and Lall (1999) do not. The major difference between these techniques and our approach to multisite simulation is that we use separate scenario generators for temperature and rainfall, linked by conditioning temperature on whether the day is wet or dry. This allows us to explicitly incorporate a wide range of atmospheric circulation predictors for temperature, as described below.

Other authors take a parametric approach, fitting a probability distribution such as the gamma or mixed exponential (Wilks 1999a) to the observations and then sampling for the rainfall amount. However, this provides no mechanism for preserving intersite relationships in multisite scenarios. Other approaches to downscaling for multiple sites have been implemented (Charles et al. 1999; Hughes et al. 1999; Wilks 1998, 1999b), but none are problem free. Wilks (1998, 1999b) downscales using a WGEN-type approach (Richardson 1981) in which each of a collection of single-site models is driven with temporally independent but spatially correlated random numbers. When the number of sites (and variables) becomes large, internal inconsistencies can develop in the correlation matrices, and Wilks had to introduce consistency adjustments. The method of Hughes et al. (1999), which uses a nonhomogeneous hidden Markov model to downscale for multisite precipitation occurrence (Charles et al. 1999), has a very large computational requirement.

c. The temperature scenarios

Temperature is downscaled using a deterministic approach based on multilinear regression analysis (Winkler et al. 1997), with independent variables based on sea level pressure (SLP) and 500-hPa geopotential height (Z500). We have already noted the mismatch between the twice-daily NMC reanalyses and the single-daily means available for HadCM2. For the temperature scenarios, from the NMC reanalyses we calculated a daily mean from the 0000 and following 1200 UTC values and used this mean to construct the transfer functions, which, in the application stage, are initialized with the HadCM2 daily mean values.

From daily SLP and Z500, 1000–500 hPa geopotential thickness (THK) can be derived, after first calculating 1000-hPa height from the hydrostatic balance equation (Peixoto and Oort 1992). From the gridded fields, daily point values for SLP, Z500, and THK are interpolated using a 16-point Bessel interpolation scheme, centered over the location of the predictand site. Then, north–south and east–west gradients are calculated (over 4° of latitude and longitude, respectively) in order to approximate the strength and orientation of the lower- and upper-level flow, as are backward and forward tendencies of the point and gradient values (over 24 h) in order to incorporate persistence. This gives 27 potential predictors.

These predictors were entered into a stepwise multiple regression analysis, using a probability for F-to-enter of 0.05 as the criterion to admit a variable, and a maximum probability for F-to-remove of 0.10 as the criterion for removal. A linear model was used in preference to the complexity of a multilinear artificial neural network approach (Weichert and Bürger 1998; Trigo and Palutikof 1999). Absolute rather than standarized values of both predictors and predictands were used, in order to maximize the potential for change between the present day and the future (Winkler et al. 1997). Separate equations were calculated for maximum and minimum temperature (TMAX and TMIN), for wet and dry days, and for the standard seasons and annually (i.e., all data). This gives a total of 20 equations for each site.

Selection of the transfer function to predict temperature is conditional upon whether the day is wet or dry, and hence the temperature scenarios are related to the multisite rainfall scenarios. Richardson (1981) developed the more conventional weather generator approach, which uses a first-order autoregressive model to construct scenarios of temperature and radiation based on generated rainfall. Wilks (1999b) has shown

that this approach can be successfully modified for the multisite multivariable case. However, the Richardson model is based on a set of parameters that were originally chosen for the climatology of the United States. Hayhoe (1998, 2000) showed that these parameters were unsuited to the Canadian case and, by recalculating them, obtained a substantial improvement in performance. An alternative to the Richardson approach, developed for European application (Semenov et al. 1998), has not been tested in the multisite context.

6. Application of the downscaling models

a. The reference rain day scenario

Underpinning the construction of the reference rain day scenarios is the need to demonstrate that

1) the CWT classification scheme is physically valid;
2) there are consistent and distinct relationships between CWT and rain day occurrence; and,
3) the GCM accurately simulates CWTs.

These criteria have been evaluated for the two regions by Goodess (2000), who concluded that the first two are clearly met. However, systematic errors in the GCM can be traced through to the downscaled series. Nonetheless, the GCM performance is considered adequate for the purposes of developing the statistical downscaling methodology, particularly because downscaled time series, rather than observed series, are used to provide a baseline for the climate change scenarios, on the assumption that the errors are consistent throughout the GCM run. The main findings from the evaluation, as these relate to the scenario construction method, are summarized in Tables 5 (Guadalentin) and 6 (Agri).

In the Guadalentin, the frequent CWTs are the four high-vorticity types, C+HYC, UC, A+HYA, and UA. The two cyclonic categories are associated with a high number of rain days, whereas the anticyclonic categories are generally dry. The GCM substantially underestimates the number of C+HYC and UC days, and overestimates the number of A+HYA days. We should therefore expect too few wet days in the model scenarios. The seasonal distribution of errors is such that the summer scenarios should be the least successful. The modeled CWT changes between 1970–79 and 2030–39 indicate a possible, but small, reduction in rain day occurrence, as the C+HYC and UC types become less common and the A+HYA and UA types increase in importance. Between 1970–79 and 2090–99, there is a much stronger indication of an increase in wet day occurrence throughout the year, especially marked in winter, and associated with an increased frequency of C+HYC, UC, E+SE, and S+SW groups, and reduced occurrence of A+HYA and N groups. Thus, the changes in CWT frequencies are not uniform over time, which might be related to the fact that, although the greenhouse gas forcing increases monotonically, sulfate emissions

TABLE 5. CWT characteristics underlying the reference rain day scenario generator for the Guadalentin.

CWT	Observed frequency, all days (1956–89)	Observed rain day occurrence frequency Alcantarilla (1958–87)	Frequency of CWTs in HadCM2 relative to observed (1956–89)	HadCM2 CWT change 2030–39 cf. 1970–79	HadCM2 CWT change 2090–99 cf. 1970–79
C+HYC	High (summer)	High	Very low (spring, summer)	↓ All year	↑ Winter
UC	High (spring, summer, autumn)	High (winter) Low (summer)	Very low (summer) Winter accurate	↓ Summer	↑ Winter
A+HYA	High (winter)	Low	Very high	↑ Winter, summer	↑ Spring ↓ Winter, summer
UA	High (winter)	Low	High (winter, summer)	↑ Summer, autumn ↓ Spring	↑ Autumn ↓ Spring
N	Average (winter high)	Low	Low (winter) High (spring, summer)	↓ Winter	↓↓
NE	Low	High (autumn)	Low (winter, spring) High (summer)	Little change	↓ Winter, autumn
E+SE	Low	High	High (summer)	↑ Winter	↑
S+SW	Low	High (winter and autumn)	High	↑ All year	↑↑
W+NW	Average	Low	Accurate	↑ Spring ↓ Autumn	↑ Spring (small) ↓ Winter, autumn

↑ = Moderate change in direction of arrow; ↑↑ = large change.

TABLE 6. CWT characteristics underlying the reference rain day scenario generator for the Agri.

CWT	Observed frequency, all days (1956–89)	Observed rain day occurrence frequency Missanello (1956–88)	Frequency of CWTs in HadCM2 relative to observed (1956–89)	HadCM2 CWT change 2030–39 cf. 1970–79	HadCM2 CWT change 2090–99 cf. 1970–79
C+HYC	High (winter)	High (all year)	Low (winter)	↓ Winter	↓↓ Winter, autumn
UC	High (spring, summer, autumn)	High (summer) Low (winter, autumn)	Very low	Little change	↓ Winter, spring
A+HYA	Low	Low	Very high	↑ Winter ↓ Summer	↑↑ Winter ↑ Autumn ↓ Summer
UA	High (spring, summer, autumn)	Low	High (winter) Low (spring, summer, autumn)	↑ Winter	↑ Autumn
N	Low	Low	High (winter, autumn) Low (summer)	↓ Spring, summer	↓ All year
NE	Low	High (summer) Low (winter)	High (spring, autumn)	↑ Autumn ↓ Winter, summer	↑ Summer, autumn ↓ Winter
E+SE	Low (summer)	High	High (spring, summer) Low (winter)	↓ Spring	Little change
S+SW	Average (low summer)	High (autumn) Low (spring)	High (summer)	↓ Autumn	↑ Winter ↓ Autumn
W	Low (summer)	High (autumn) Low (summer)	High (summer)	↑ Spring ↓ Winter	Little change
NW	Low (autumn)	High (autumn) Low (winter, spring, summer)	High	↑ Summer	Little change

↑ = Moderate change in direction of arrow; ↑↑ = large change.

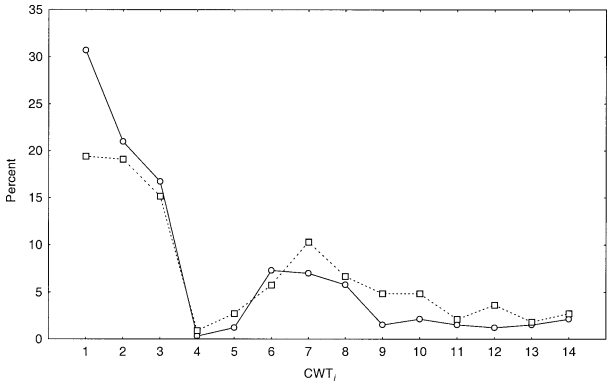

FIG. 2. Transition from CWT type 1 on day $t - 1$ to CWT type i on day t for the Agri. Solid line observed, dashed line GCM, all data for 1970–79. For CWT names, see Table 4.

stabilize in the model after about 2050 after a very rapid increase in the early decades.

In the Agri, the relationships between CWTs and rain day occurrence are broadly the same as for the Guadalentin. However, the A+HYA types are much less common, and are replaced by a higher frequency of wet C+HYC and dry UA types. Comparing the HadCM2 CWTs for 1956–89 with observations, the GCM underestimates the occurrence of wetter C+HYC and UC types, and overestimates drier A+HYA types, so that again the GCM-based scenario is expected to be too dry, especially in summer. Both the 2030–39 scenario and the 2090–99 scenario are expected to be drier than the 1970–79 baseline, particularly in winter and autumn.

Not only should the GCMs successfully reproduce the occurrence of CWTs, but also the transitions between days [although this does not affect the performance of the weather generator used here, see Eq. (1)]. As an example, Fig. 2 shows for the Agri the percentage of transitions from CWT type 1 (cyclonic) on day $t - 1$ to CWT type i on day t. It can be seen that the GCM successfully reproduces the broad pattern, but notably underestimates the autocorrelation.

To maximize the potential for change, a Monte Carlo approach to the generation of the reference rain day scenario was adopted. Although the sequence of CWTs is fixed (being based either on observations or HadCM2), the precipitation outcome can be varied by making multiple runs of the rain day occurrence generator, each with a different starting value for the seed of the random number generator.

The first step in the validation is to test the rain day occurrence generator on independent observations, using a cross-validation procedure. Taking the case of Alcantarilla, the rainfall record contains 25 complete years with CWT data. Sequentially, one year was removed from the 25, and the remainder used to calculate the probability of rain day occurrence depending on CWT and whether the previous day was wet or dry. Then, the resulting model was run 1000 times for the removed year. The same procedure was used for Missanello, based on 33 yr, 1956–88. The results for winter are shown in the first half of Table 7. At both sites, the means (over 1000 simulations) of all the test parameters are too low compared to observations. We only show

TABLE 7. Validation of reference rain day scenario generator performance in winter. NRD = number of rain days; SD = standard deviation of seasonal totals; LW = length of the longest wet spell; LD = length of the longest dry spell.

	Observed 1958–87	Simulated (mean of 1000 values)	Maximum simulated value	Minimum simulated value	No. simulated runs significantly > observed*	No. simulated runs significantly < observed*
Cross validation on observations—Alcantarilla						
Mean NRD	14.8	13.3	16.0	10.4	0	33
SD NRD	6.5	4.5	7.0	2.2	0	580
LW	9	6.4	14	4		
LD	59	46.0	79	29		
Cross validation on observations—Missanello						
Mean NRD	29.2	27.9	31.2	20.4	0	40
SD NRD	6.4	5.9	10.5	3.7	18	66
LW	16	9.8	22	6		
LD	31	25.3	50	15		
Cross validation on GCM data—Alcantarilla						
Mean NRD	14.8	12.4	15.6	9.7	0	121
SD NRD	6.5	4.6	6.9	2.1	0	415
LW	9	6.0	12	3		
LD	59	48.5	88	29		
Cross validation on GCM data—Missanello						
Mean	29.2	23.1	26.6	18.3	0	974
SD NRD	6.4	6.2	10.7	3.5	16	26
LW	16	8.4	16	5		
LD	31	29.1	66	17		

* T test and 5% significance level for mean NRD; F test and 10% significance level for SD NRD.

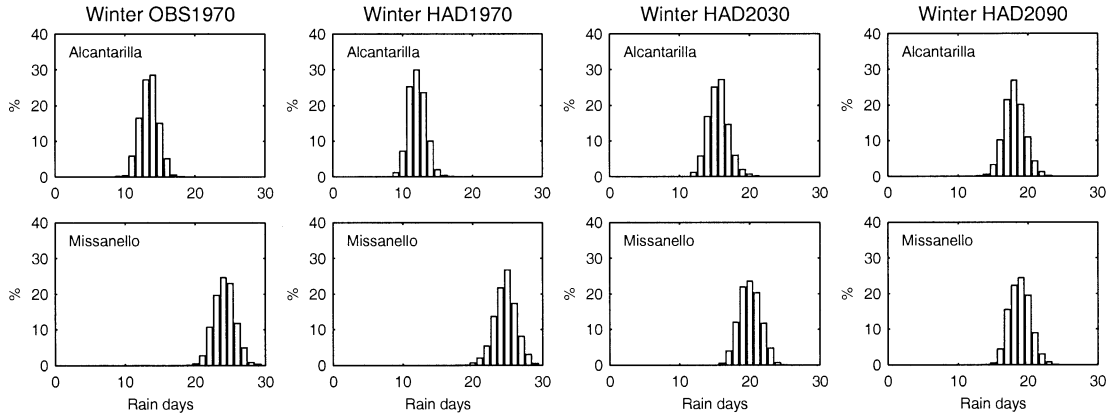

FIG. 3. Frequency distributions of the number of rain days generated by 1000 simulations with the reference rain day generator for winter. Circulation weather types for the simulations are derived from observations for 1970–79 (OBS1970) and from three decades in the HadCM2 GCM: 1970–79 (HAD1970), 2030–39 (HAD2030), and 2090–99 (HAD2090).

the winter data, but the same would also be true for spring. Performance is better in autumn and, particularly, in summer. However, the summer result must be judged in the context of the very dry Mediterranean summer climate, with very few rain days. The simulated value of LW and LD is shorter than observed in most cases, a result also observed, for example, by Wilks and Wilby (1999).

The lower part of Table 7 shows the performance of the rain day occurrence generator when initialized with HadCM2 data. Again, a cross-validation procedure was used. The probabilities of rain day occurrence were calculated from the observed record less 1 yr, and the resulting model was applied to CWTs calculated from the removed-year data in HadCM2. Again, the generator was run 1000 times for each year. The results indicate that substituting modeled CWTs for observed CWTs further degrades the performance of the rain day occurrence generator in terms of the mean, although not in terms of the variance characteristics.

The final step is to apply the rain day occurrence generator using HadCM2 CWTs for the decades 1970–79, 2030–39, and 2090–99, performing 1000 runs for each decade. Figure 3 shows the frequency distributions of the number of rain days in the 1000 runs, taking winter as the example. Comparing the simulations based on GCM CWTs for 1970–79 with those based on observed CWTs for the same period, at Alcantarilla the model-based simulations produce too few rain days, with a mode of 12 as opposed to a mode of 14 in the observation-based runs (which themselves already contain characteristics different from the observed record of rain days). At Missanello performance is better, with the mode agreeing with that of the observation-based runs. Nevertheless, taking the results for all seasons, it is the case that the model-based simulations have too few rain days. To minimize the effect of this problem, we extracted from the 1000 model-based runs for 1970–79 all those runs for which the number of rain days in every season falls within the observed decadal range (calculated from overlapping decades in the full record of 1958–87 for Alcantarilla and 1956–88 for Missanello). Then, one simulation was selected randomly from these subsets, which had just 63 members in the Guadalentin and 33 in the Agri. For Alcantarilla, this was the 989th run, in a ranking from the fewest to the most rain days, and for Missanello the 979th, indicating clearly the tendency of the GCM-based simulations to be too dry. The scenarios with the same ranked order were selected from the 2030–39 and 2090–99 runs. These form the reference scenarios from which the multisite rainfall scenarios are constructed. Note that alternative selection criteria could be employed, for example to maximize the between-decade change in rain day number, and would be plausible visualizations of possible futures.

b. The multisite rainfall scenarios

The multisite rainfall scenarios are created by randomly sampling from the benchmark file, with replacement. There is therefore a risk of oversampling, which may be exacerbated by the systematic errors in the HadCM2 simulation. We calculated the number of times that each bin is sampled during scenario construction as a ratio of the number of observations in that bin. For the 1970–79 GCM-based Guadalentin scenario, for example, the greatest problems arise in the A+HYA, the UA, and the E+SE CWT bins, because the model overestimates their occurrence (see Table 5). The largest errors occur in summer. If we take a cutoff of 0.67 (double the expected perfect ratio) to separate properly sampled from excessively sampled bins, we find that 15 out of 72 bins (four seasons by two precipitation states by nine CWTs) are excessively sampled in the scenario decade 1970–79, compared to 13 in the 2030–39 decade and 19 in the 2090–99 decade. Therefore, future chang-

TABLE 8. Characteristics of the multisite rainfall scenarios for the Guadalentin compared to observations. The wettest decade in the observations is 1971–80 and the driest is 1958–67, selected on the basis of the number of rain days at Alcantarilla. Full site names are given in Table 1.

Abbreviated name	ALC	ADM	EDC	FA	LOR	TOT
Mean rainfall (mm day^{-1}):						
Wettest observed decade	0.94	1.36	0.92	0.73	1.01	0.80
Driest observed decade	0.70	1.06	0.68	0.65	0.76	0.78
Scenario 1970–79	0.67	0.99	0.68	0.49	0.63	0.56
Scenario 2030–39	0.74	1.06	0.79	0.58	0.73	0.66
Scenario 2090–99	0.84	1.16	0.77	0.70	0.76	0.83
Standard deviation (daily data):						
Wettest observed decade	4.62	5.58	4.63	3.65	5.07	4.37
Driest observed decade	3.67	4.60	3.80	3.48	4.10	4.10
Scenario 1970–79	3.61	4.68	3.96	2.55	3.85	3.38
Scenario 2030–39	4.02	5.00	4.48	3.09	3.73	4.48
Scenario 2090–99	4.25	5.15	4.63	3.62	3.82	5.14
Number of rain days:						
Wettest observed decade	539	514	381	424	277	289
Driest observed decade	425	412	323	390	302	310
Scenario 1970–79	431	383	303	313	242	261
Scenario 2030–39	467	402	328	348	280	273
Scenario 2090–99	539	466	335	393	295	301
Longest dry day run:						
Observed, 1958–87	118	140	178	156	141	145
Scenario, 1970–79	88	72	90	86	146	120
Scenario, 2030–39	60	67	73	63	98	96
Scenario, 2090–99	84	84	74	84	90	94
Longest wet day run:						
Observed, 1958–87	9	10	5	6	6	7
Scenario, 1970–79	9	3	3	4	4	8
Scenario, 2030–39	6	5	5	5	4	5
Scenario, 2090–99	9	5	5	5	8	4

es in the GCM appear to have little effect on sampling performance.

The characteristics of the multisite rainfall scenarios, compared to observations, are shown in Tables 8 (Guadalentin) and 9 (Agri). Despite all efforts, the rainfall scenarios remain too dry. Here, the comparison is made with observed extreme decades, selected on the basis of the number of rain days at one site. However, using different criteria to select the decades, for example, rainfall amount, leads to the same conclusion. In the Guadalentin, only one site has a rainfall amount in the 1970–79 scenario that lies between the two observed extreme decades. The comparable figures for the standard deviation of daily rainfall and the number of rain days are two and one sites, respectively. In the wetter Agri (Table 9), performance is better: the station numbers are two (i.e., 18% of sites) for rainfall amount, five (45%) for the standard deviation of daily rainfall, and nine (82%) for the number of rain days. We have not attempted to estimate the sampling uncertainty associated with the multisite rainfall scenarios. This would be possible if an ensemble of GCM simulations were available, and identical procedures for constructing the multisite scenarios were carried out for each.

Comparison of the scenarios for the present day and future shows different trends in the two regions. In the Guadalentin (Table 8), all three measures (amount, standard deviation, and rain day number) increase with time. The increases between 1970–79 and 2030–39 are generally of the same order as those between 2030–39 and 2090–99. In the Agri (Table 9), the scenarios indicate drier conditions in the future. The larger changes are between 2030–39 and 2090–99, with more modest reductions between 1970–79 and 2030–39. These changes are consistent with the changes in CWTs noted in section 6a. The size of the changes in the Guadalentin is substantial when compared to the interannual variability of the present-day rainfall regime. Thus, at three sites the scenario change in the number of rain days between 1970–79 and 2090–99 is greater than the natural variability expressed as the difference in rain day number between the wettest and driest decades in the observations. At two sites, Lorca and Totana, the ratio of scenario change to natural variability exceeds 2. In the Agri the relative changes are more modest. Only at two sites does the scenario change exceed natural variability, and in general the ratio of the scenario change to natural variability is less than 0.50.

Two potential shortcomings in the complete methodology for generating multisite rainfall scenarios can be identified. First, rainfall patterns can only change in response to differences in the frequency of occurrence

TABLE 9. Characteristics of the multisite rainfall scenarios for the Agri compared to observations. The wettest decade in the observations is 1957–66 and the driest is 1968–77, selected on the basis of the number of rain days at Missanello. Full site names are given in Table 1.

Abbreviated name	ALI	COR	MIS	MOL	NSS	PIS	ROC	SEN	SMA	STI	TUR
Mean rainfall (mm day^{-1}):											
Wettest observed decade	2.11	2.30	2.18	3.23	1.69	1.93	2.11	2.20	2.56	2.21	2.10
Driest observed decade	2.13*	1.88	2.33*	2.98	1.65	1.70	1.99	1.93	2.09	2.09	1.93
Scenario 1970–79	1.90	1.86	2.00	2.82	1.56	1.64	1.85	1.88	1.96	1.9	1.99
Scenario 2030–39	1.93	1.76	1.93	2.65	1.38	1.65	1.72	1.76	1.89	1.85	1.74
Scenario 2090–99	1.63	1.53	1.77	2.39	1.28	1.35	1.55	1.46	1.67	1.61	1.44
Standard deviation (daily data):											
Wettest observed decade	6.30	5.84	5.95	7.83	7.13	8.49	6.13	6.58	6.64	7.16	7.51
Driest observed decade	6.71	5.53	7.39	7.87	7.92	6.48	6.50	6.4	6.43	9.12	9.04
Scenario 1970–79	6.59	5.52	6.31	7.52	8.30	8.19	6.17	6.36	5.87	8.03	9.63
Scenario 2030–39	6.78	5.18	6.07	7.06	8.35	10.48	5.73	6.46	5.72	8.25	8.28
Scenario 2090–99	5.50	4.50	5.38	6.65	6.87	7.06	5.12	5.16	5.28	6.33	5.96
Number of rain days:											
Wettest observed decade	824	1258	1070	1388	1027	742	949	1379	1196	919	1122
Driest observed decade	858	1102	828	1331	906	995	885	1096	920	808	915
Scenario 1970–79	765	1113	834	1287	920	815	792	1130	932	815	946
Scenario 2030–39	765	1074	808	1225	848	761	781	1100	934	771	920
Scenario 2090–99	699	1049	780	1201	835	765	729	1040	837	720	865
Longest dry day run:											
Observed 1958–87	68	50	73	43	57	64	49	63	59	81	63
Scenario 1970–79	36	23	49	28	42	46	46	29	39	47	42
Scenario 2030–39	35	33	54	19	35	43	35	37	56	40	24
Scenario 2090–99	47	33	47	23	28	54	45	36	33	43	30
Longest wet day run:											
Observed 1958–87	12	14	16	25	11	13	11	19	16	11	13
Scenario 1970–79	7	8	7	10	8	5	7	8	8	8	8
Scenario 2030–39	6	11	10	15	10	6	13	15	13	8	10
Scenario 2090–99	8	8	8	9	7	8	8	9	7	6	8

* Because wettest and driest decades are selected on the basis of rain day number, the rainfall amount per day may be greater in the driest decade than in the wettest.

TABLE 10. Mean seasonal rainfall from the raw GCM output and the downscaled scenarios. All ratios are calculated with respect to the 1970–79 decade.

		Winter	Spring	Summer	Autumn
Agri					
1970–79 (mm)	Raw GCM	148	121	105	148
	Scenario	242	139	68	188
Ratio 2030–39	Raw GCM	0.86	1.11	1.20	0.92
	Scenario	0.65	1.10	1.91	0.84
Ratio 2090–99	Raw GCM	1.15	1.12	0.83	0.87
	Scenario	0.47	1.07	1.88	0.71
Guadalentin					
1970–79 (mm)	Raw GCM	62	54	31	40
	Scenario	61	65	79	81
Ratio 2030–39	Raw GCM	0.99	1.27	0.92	0.93
	Scenario	1.02	1.67	0.76	0.96
Ratio 2090–99	Raw GCM	1.43	1.25	0.79	0.82
	Scenario	0.91	1.28	1.01	1.21

of CWTs. Analyses of observed rainfall and CWTs have reached mixed conclusions. Bárdossy and Caspary (1990) and Caspary (1996) found strong links between changes in the atmospheric circulation and in observed rainfall across Europe. Others have found, however, that changes in rainfall cannot be adequately explained by CWT changes (Widmann and Schär 1997; Frei et al. 1998) and that atmospheric temperature and humidity may also play a role (Wilby et al. 1998; Buishand and Brandsma 1999). In Table 10 the daily rainfall scenarios have been summed across seasons and stations to generate seasonal catchment means for the Guadalentin and Agri. These are compared with raw rainfall data from HadCM2 for the nearest land grid box. Broad agreement between the two would imply that the main cause of rainfall change in the model is a change in CWT occurrence. Table 10 shows very good agreement between raw and downscaled information for the decade 2030–39. In the Agri, the agreement remains strong in the spring and autumn of 2090–99, but is poor in the winter and summer of this decade. In the Guadalentin, good agreement in the 2090–99 decade is found only in spring.

Second, there is no mechanism for generating a future change in the relationship between CWT and rainfall occurrence. Wilby et al. (1998) identified this shortcoming as responsible for the small changes obtained by vorticity-based downscaling methods when compared to the changes achieved using artificial neural networks. We examined the relationship between CWTs and rainfall in two HadCM2 grid points located close to the Guadalentin and the Agri for the three scenario decades: 1970–79, 2030–39, and 2090–99. The results are shown in Fig. 4. Two ratios were used: first, mean rainfall per wet type-day to mean rainfall on all wet days ($PREC_{ct}/PREC_{tot}$) and, second, the probability that a type-day will be wet to the probability that any day will be wet ($PROB_{ct}/PROB_{tot}$). Figure 4 suggests that, with few exceptions, the ratios remain stable between 1970–79 and 2030–39. Greater instability is manifest between 2030–39 and 2090–99, particularly in the drier Guadalentin. Where changes do occur, they tend to be in the relatively rare directional types and, for the Guadalentin, cyclonic types. However, at least for the changes between 1970–79 and 2030–39, the assumption that rainfall–CWT relationships are stable is reasonable.

c. Temperature scenarios

The temperature transfer functions are developed separately for wet and dry days, in an effort to impose stability between rainfall and temperature relationships in the scenarios. This step is only justified if there are distinct differences between wet and dry day temperatures. Table 11 shows, for a 10-yr period at Alcantarilla, the observed means and standard deviations for wet (defined as any precipitation) and dry days. Significance levels are set at 5% for a *t* test (assuming unequal variance) of the difference of the means, and 10% for Bartlett's *F* test for differences in the standard deviations. Throughout this paper, and to avoid problems due to autocorrelation, significance levels are calculated on a sample of temperature data (usually 10%), drawn randomly. For maximum temperatures, the means are significantly lower on wet days, a result of increased cloudiness and changes in latent and sensible heat fluxes. For minimum temperatures, the means are generally (with the exception of summer) warmer on wet days due to the reduction in radiative heat loss caused by enhanced cloudiness, and the differences are statistically significant in winter and autumn. In spring, the TMIN means are not significantly different, but the standard deviations are. Thus, in every season with a substantial number of wet days (i.e., excluding summer), there are significant differences between either the means or the standard deviations of TMAX and TMIN calculated for wet and dry days separately. This is sufficient justification to develop separate transfer functions.

The wet and dry day transfer functions were constructed following the procedure outlined in section 5c.

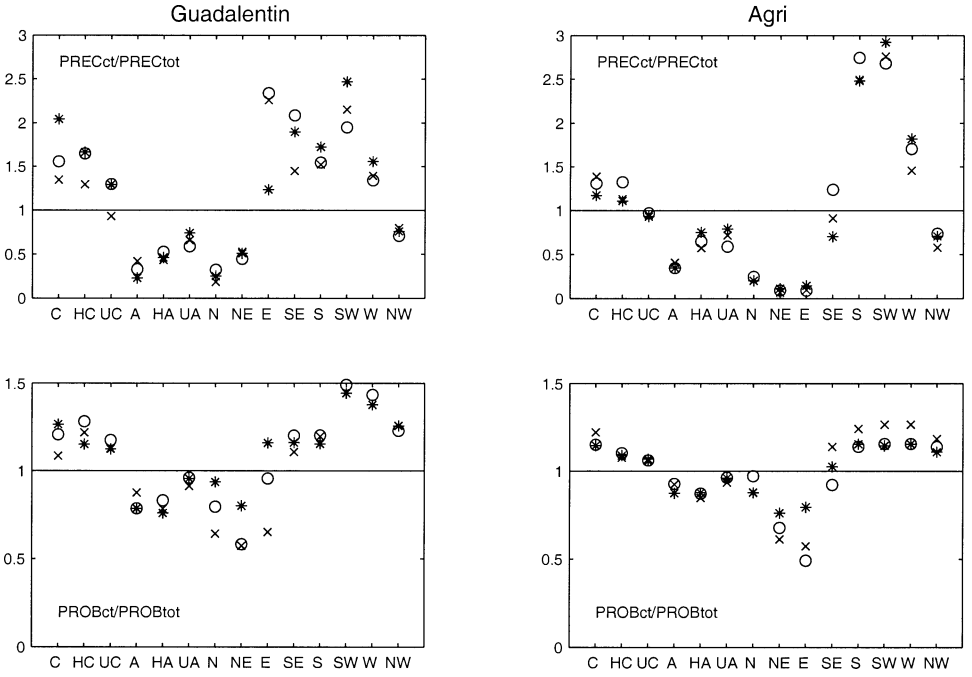

FIG. 4. Relationship between CWTs and rainfall at the two HadCM2 land grid boxes located closest to the Guadalentin and the Agri for the three scenario decades: 1970–79 (*), 2030–39 (○), and 2090–99 (×). $PREC_{ct}/PREC_{tot}$ = ratio of mean rainfall per wet type day to mean rainfall on all wet days. $PROB_{ct}/PROB_{tot}$ = ratio of the probability that a type day will be wet to the probability that any day will be wet.

For the annual regression equations, the adjusted r^2 values are always greater than 0.75. The seasonal equations have lower r^2 values, especially for TMIN, falling as low as 0.37 for dry days at Alcantarilla in winter and Nova Siri Scala in summer.

Ideally, we would hope that SLP would appear as an important predictor in the temperature transfer functions. This would strengthen the links between the rainfall and temperature scenarios, since the CWTs that condition the rainfall scenarios are calculated from SLP. Inspection shows that without exception at both sites it is the 1000–500-hPa geopotential thickness that is selected first by the stepwise procedure. However, at the second step an SLP predictor is generally selected, in the Guadalentin most commonly the north–south pressure gradient and in the Agri the absolute SLP.

TABLE 11. Observed wet and dry day TMAX and TMIN (°C) characteristics for Alcantarilla, 1975–84. An asterisk indicates significant differences between wet and dry day values (see text for explanation).

		Means		Standard deviation	
		Wet days	Dry days	Wet days	Dry days
Annual	TMAX	19.00*	24.29*	5.65*	6.66*
	TMIN	10.62	11.10	4.86	6.21
Winter	TMAX	14.35*	17.15*	3.05	3.36
	TMIN	6.62*	4.74*	3.29	3.48
Spring	TMAX	18.61*	22.46*	3.91	4.24
	TMIN	9.81	8.57	3.04*	3.79*
Summer	TMAX	26.91*	31.73*	3.59	3.28
	TMIN	17.19	18.04	2.56	2.69
Autumn	TMAX	21.40*	24.70*	5.11	5.09
	TMIN	13.32*	11.98*	4.00	4.83

The transfer functions were validated on independent observations. In Table 12, the root-mean-square errors are generally lower, and the correlation coefficients (r) higher, for the seasonal equations. In the following discussions, only the seasonal equations are used. We also explored the autocorrelation properties of the transfer-function-generated temperatures. Figure 5 shows, for standardized monthly anomalies, the lag 1 to lag 5 autocorrelation coefficients for 1985–87 at Alcantarilla from observations, transfer functions initialized with NMC data, and transfer functions initialized with GCM data. For TMIN, the GCM-based temperatures reproduce successfully the main observed characteristics, although with a systematic underestimate. The TMAX results are less satisfactory.

Table 13 compares the performance of the transfer function-derived scenarios with raw GCM data, interpolated to the site using a 16-point Bessel scheme. The example of Alcantarilla is given, for 1970–79. As found in studies based on other GCMs (e.g., Palutikof et al. 1997, using the Canadian Climate Centre model), the raw TMAX means are always too cool and the raw TMIN means are always too warm, while the modeled standard deviations of the daily data are too low. With few exceptions, the differences are statistically significant. The downscaled means are all closer to the observations than the GCM data, with many fewer (2 as

TABLE 12. Validation of the temperature transfer functions. Validation years at each site are shown in row 1 of header. In row 2, annual = single transfer function developed from all data in the construction decade, seasonal = four transfer functions developed from standard seasons. In row 3, rmse = root-mean-square error, r = correlation coefficient.

		Alcantarilla (1985–87)				Nova Siri Scala (1979–88)			
		Annual		Seasonal		Annual		Seasonal	
		rmse	r	rmse	r	rmse	r	rmse	r
Annual	TMAX	2.38	0.94	1.98	0.96	3.16	0.92	2.66	0.94
	TMIN	2.60	0.92	2.44	0.93	2.43	0.93	2.18	0.94
Winter	TMAX	2.90	0.73	2.18	0.76	4.21	0.47	2.88	0.60
	TMIN	3.33	0.59	2.97	0.56	3.22	0.59	2.47	0.60
Spring	TMAX	2.48	0.89	2.07	0.90	2.59	0.84	2.52	0.84
	TMIN	2.45	0.82	2.36	0.82	2.09	0.86	2.15	0.85
Summer	TMAX	1.89	0.84	1.55	0.89	2.30	0.76	2.32	0.80
	TMIN	1.88	0.67	1.67	0.69	1.92	0.73	1.75	0.75
Autumn	TMAX	2.15	0.94	2.06	0.94	3.14	0.90	2.90	0.89
	TMIN	2.60	0.86	2.61	0.85	2.23	0.89	2.30	0.86

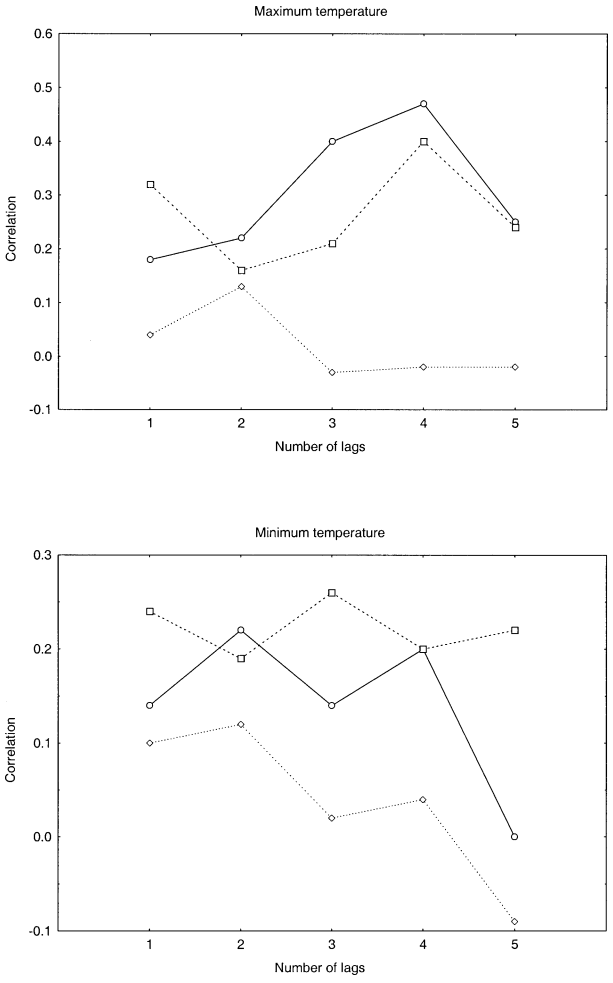

FIG. 5. Lag 1 to lag 5 autocorrelation coefficients for standardized monthly temperature anomalies at Alcantarilla (1985–87) from observations (solid line), transfer functions initialized with NMC data (dashed line), and transfer functions initialized with GCM data (dotted line).

opposed to 9) statistically significant differences. Except in spring, they are always too low. Whereas the differences are generally small at Alcantarilla, at Nova Siri Scala the transfer-function-predicted means are sometimes more than 2°C below observed and the standard deviations are, with the exception of winter TMAX, significantly too low. The Agri temperature scenarios are not considered further here. The discrepancies in observed and simulated mean values arise from the GCM predictor variables. When the transfer functions are tested with NMC-derived predictors for an independent period, mean TMAX and TMIN agree closely (see Table 14 for the example of Alcantarilla).

Finally, Table 15 compares downscaled and GCM-based temperatures for the future at Alcantarilla. There are clear seasonal differences between the raw GCM data and the transfer-function-derived scenario. In winter, the warming is greatest in the GCM for both TMAX and TMIN, whereas in spring and autumn the downscaled scenarios warm the most. The result, for the year as a whole, is that the warming is almost identical in the GCM and in the downscaled scenario.

7. Between-site and between-variable consistency in the scenarios

In section 2 it was stated that the downscaled scenarios should maintain stable relationships between sites and between variables. Two tests are described here that explore the extent to which the scenario generation methods have been successful in this regard. Designing these tests was not straightforward, and there is a lack of precedent in the literature. We opted to use scatterplots to test between-site relationships in the temperature scenarios. As an example test of between-variable relationship stability, temperature data were partitioned into wet and dry days, and box-and-whisker plots created for each partition.

TABLE 13. TMAX and TMIN (°C) at Alcantarilla for 1970–79 predicted by the wet/dry day seasonal regression equations initialized with GCM data, compared with observations and with GCM interpolated values for the same period. An asterisk indicates significant difference from observed (see text for explanation). SD = standard deviation.

		Observed		GCM (interpolated)		Regression derived	
		Mean	SD	Mean	SD	Mean	SD
Annual	TMAX	23.6	6.8	18.6*	6.7	22.5	6.4
	TMIN	11.0	6.0	12.8*	5.5*	10.6	6.2
Winter	TMAX	16.7	3.5	11.6*	2.6*	16.2	3.1
	TMIN	5.1	3.5	7.4*	3.2*	3.9	2.8*
Spring	TMAX	21.8	4.4	16.8*	3.8	21.8	4.1
	TMIN	8.8	3.7	10.9*	3.1	9.3	3.6
Summer	TMAX	31.4	3.5	26.3*	3.5	29.8*	3.7
	TMIN	18.0	2.7	18.9*	2.7	17.5	2.6
Autumn	TMAX	24.2	5.2	19.9*	5.7	22.2*	5.5
	TMIN	12.2	4.7	13.9	4.6	11.8	5.7*

a. Between-site relationships

Figure 6 shows scatterplots of present-day TMAX at Alcantarilla and Lorca in the Guadalentin, based on observations (above) and GCM-derived downscaled scenarios (below). If the downscaling models succeed in preserving between-site relationships, strong similarities should exist between the observed and downscaled scatterplots for the same site pairs. (Note that, for rainfall, the multisite scenarios are constructed in such a way that these relationships are necessarily preserved.) The linear best-fit lines and their equations indicate that, in terms of both the intercept and the gradient terms, agreement is close. This is also true for TMIN, and for the other two site pairs in the catchment (not shown). In general (although not in Fig. 6), the correlation coefficients are higher in the predicted site pairs, which is to be expected given the deterministic transfer-function-based methodology used.

The differences seen in Fig. 6 are small enough to support the conclusion that between-site relationships have been successfully preserved by the downscaling strategy for temperature. However, Fig. 6 is based on all data, so that a large part of the agreement can be expected to be due to the seasonal cycle. We transformed the monthly mean temperatures into standardized anomalies, and recalculated the correlations, as shown in Table 16. This table shows intersite correlations for the observed data, for transfer-function-generated temperatures based on NMC-derived (observed) predictors, and for transfer-function-generated temperatures based on GCM data. It is clear that the downscaling method substantially exaggerates the intersite correlations, and that this is a function of the method itself, and not of any shortcomings in the GCM data. The intersite correlations based on the GCM-based scenarios successfully reproduce those based on NMC-derived predictors.

TABLE 14. TMAX and TMIN means at Alcantarilla calculated over the years 1973–74 and 1985–87. Transfer functions are initialized with NMC data.

		Observed	Transfer function predicted
Winter	TMAX	17.0	17.0
	TMIN	5.3	4.7
Spring	TMAX	22.7	22.6
	TMIN	8.7	9.3
Summer	TMAX	32.1	32.4
	TMIN	18.2	18.3
Autumn	TMAX	25.1	25.1
	TMIN	12.7	12.7

b. Between-variable relationships

For Alcantarilla in the Guadalentin, and Nova Siri Scala in the Agri, observed and downscaled temperature data for 1970–79 were partitioned into dry and wet day samples, and box-and-whisker plots were produced, as shown in Fig. 7. The principal characteristics of the observed plots (left-hand column) are, first, that the wet day median temperature is generally lower than the dry day median and, second, that the interquartile range is lower for the wet day sample than it is for the dry day sample. The Nova Siri Scala downscaled scenarios are generally successful at capturing these characteristics, both the lower median temperatures and the smaller interquartile range on wet days. However, the Alcantarilla plots for the simulated data largely fail to capture these two characteristics.

8. Analysis of the scenarios with respect to extreme event occurrence

The rainfall and temperature scenarios can be used to study the impact of climate change on the occurrence of extreme events. Changes in the occurrence of extremes may have more relevance for impact analysis than changes in the mean (Lerchl 1998; Parmesan et al. 2000). A simple example of the analyses that could be performed is given below, based on the temperature scenarios for Alcantarilla, although many other possibilities exist.

Table 17 shows counts of extreme hot and cold days, for Alcantarilla. The TMAX scenarios are used to an-

TABLE 15. Scenarios of mean TMAX and TMIN (°C) for Alcantarilla for 2090–99 constructed from the wet/dry day seasonal regression equations. GCM-interpolated values for the same period are shown for comparison. Differences with respect to 1970–79 are for the same scenario type in the earlier period.

		Transfer function mean	Difference (−1970/79)	GCM-interpolated mean	Difference (−1970/79)
Annual	TMAX	26.78	4.31	22.75	4.12
	TMIN	14.44	3.94	16.90	4.14
Winter	TMAX	19.41	3.17	16.26	4.70
	TMIN	7.55	3.63	12.19	4.79
Spring	TMAX	25.06	3.34	19.82	2.99
	TMIN	12.65	3.35	14.05	3.17
Summer	TMAX	33.61	3.87	29.89	3.62
	TMIN	20.43	2.89	22.65	3.79
Autumn	TMAX	29.01	6.85	25.02	5.16
	TMIN	17.12	5.89	18.72	4.80

alyze changes in the number of hot days (≥35°C) and in the number of degree-days above a 35°C threshold. The scenario for 1970–79 comes much closer to simulating correctly the observed occurrence of hot days than the raw GCM data. However, as already noted, the TMAX scenarios are still too cool, and this has an effect on the simulation of extremes. Between 1970–79 and 2090–99, while downscaled mean summer TMAX increases by 3.9°C, the average number of hot days rises by 388, representing a fivefold increase on the 1970–79 figure. In the decade 2090–99, the number of hot days is equivalent to more than 1 month with maximum daytime temperatures continually at or above 35°C. Such a result has clear significance for human health and comfort. Cold extremes were accurately simulated, while those based on raw GCM data are substantial underestimates. The number of low-temperature days in the decade falls from 91 in the 1970–79 scenario, to only 1 in 2090–99. Underpinning this change is an increase in mean winter TMIN from 3.9°C in 1970–79 to 7.5°C in 2090–99.

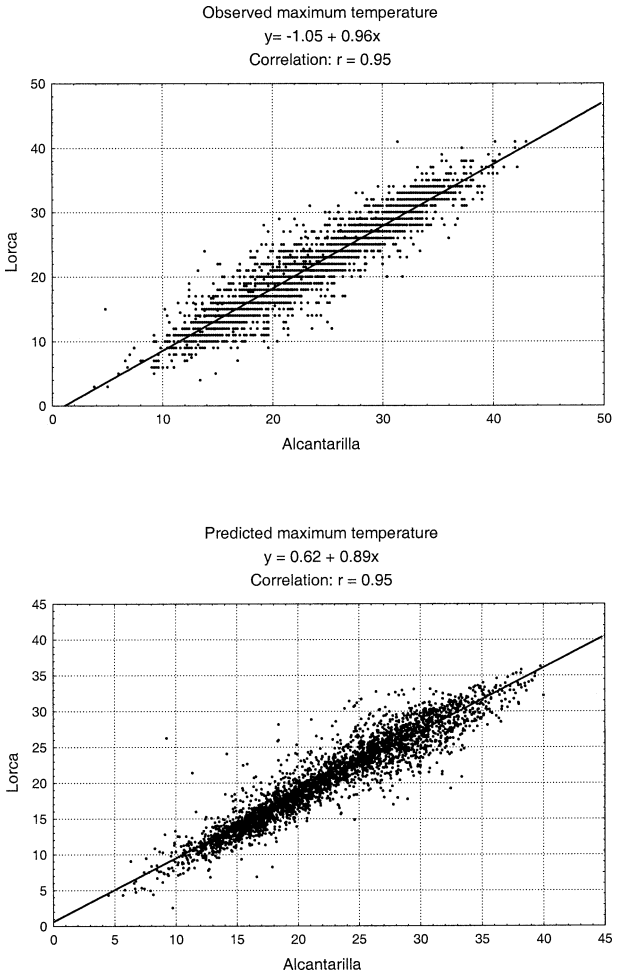

FIG. 6. Scatterplots of (top) observed and (bottom) downscaled (from HadCM2) maximum temperatures for Alcantarilla for 1970–79.

9. Conclusions

A suite of methods has been presented to generate consistent daily multisite and multivariable (rainfall and temperature) scenarios of future climates. The methods

TABLE 16. Intersite correlations of standardized anomalies of monthly temperatures for the Guadalentin. For site names, see Table 1.

	Observed 1970–79 n = 120	Transfer function predicted, from NMC data 1973–74; 1985–87 n = 60	Transfer function predicted, from GCM data 1970–79 n = 120
TMAX			
LOR-ADM	0.58	0.94	0.98
LOR-ALC	0.73	0.97	0.95
ADM-ALC	0.82	0.95	0.96
TMIN			
LOR-ADM	0.73	0.97	0.98
LOR-ALC	0.61	0.94	0.95
ADM-ALC	0.59	0.88	0.94

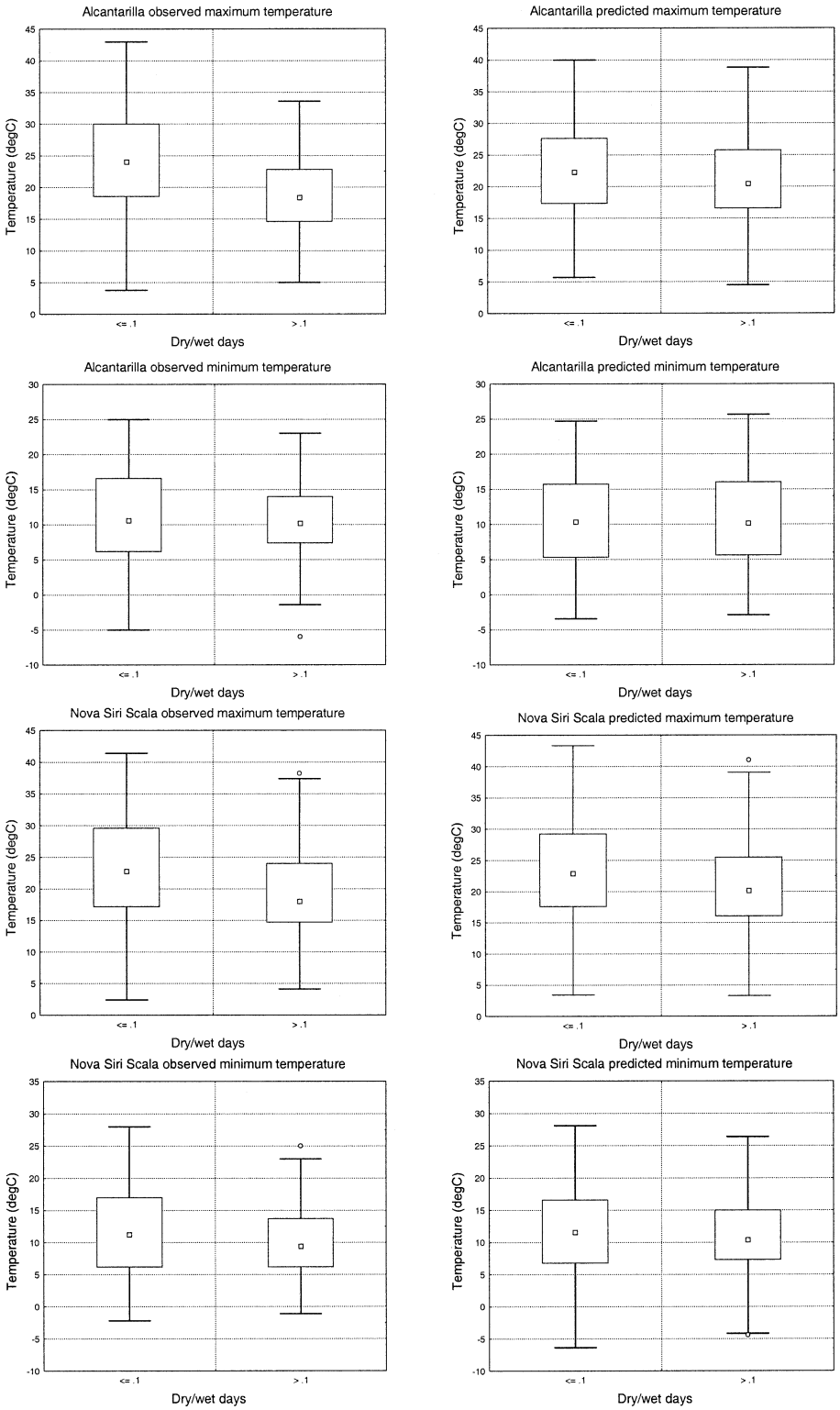

FIG. 7. Box-and-whisker plots of (left) observed and (right) downscaled (from HadCM2) temperatures for 1970–79, classified according to rainfall occurrence (dry = <0.1 mm rain; wet = ≥0.1 mm rain). Inner box = median; outer box = 25th and 75th percentiles; whiskers = maximum and minimum values.

TABLE 17. Behavior of temperature extremes at Alcantarilla, expressed as 10-yr totals. NoHD = number of hot days ≥35°C; HDG = degree-days above a threshold of 35°C; NoFD = number of cold days ≤0°C; FDG = degree-days below a threshold of 0°C.

	Observations 1970–79	GCM data 1970–79	Scenarios for		
			1970–79	2030–39	2090–99
Hot day measures calculated from TMAX:					
NoHD	146	7	77	134	465
HDG	272	9	131	211	1197
Cold day measures calculated from TMIN:					
NoFD	84	0	91	34	1
FDG	94	0	111	40	2

have been tested in two catchments: the Agri in Italy and the Guadalentin in Spain.

The daily rainfall scenarios are constructed in a two-stage procedure. First, for a single site, a reference rain day scenario is built using a circulation-typing approach combined with a conditional first-order Markov chain to describe wet day/dry day probabilities. The resulting scenarios for the present day are realistic, but tend always to be too dry, and this is particularly the case in the drier of the two catchments, the Guadalentin. We have attempted to overcome this problem by using a Monte Carlo approach to generate 1000 reference scenarios from which we select just one scenario that compares well with observations. In both the Agri and the Guadelentin, the scenarios selected by this method are at the very (wet) extreme of the distribution of simulated occurrence sequences, a problem that clearly needs to be addressed in future developments of the method.

The generation of reference rain day scenarios is the least successful aspect of the downscaling approach presented here. In common with other weather generators (Wilby et al. 1998; Wilks and Wilby 1999), the conditional weather generator tends to underestimate variability and persistence. This problem may be related to the omission of low-frequency forcing. A number of additional, low-frequency variables that could be used to condition the parameters of downscaling models have been proposed, including atmospheric humidity and temperature (Buishand and Brandsma 1999; Wilks and Wilby 1999; Goodess 2000). Incorporation of such variables may also help to address the problem that circulation changes may not be the only forcing factor for rainfall changes. However, it would require the GCM to accurately simulate these variables as well as the CWTs.

The multisite scenarios are constructed by sampling days from a benchmark file of multisite observations classified by season, weather type, and whether the day is wet or dry at the reference station. Despite using a selective procedure to identify a single reference scenario from the full Monte Carlo simulation set, the problem of overaridity persists. Overall, it is likely that the methodology for constructing the rainfall scenarios would work better in regions where weather types are better simulated. A possible candidate is maritime northern Europe, with strong westerly flow and strong cyclonic vorticity, both highly correlated with the occurrence of rainfall.

The temperature downscaling uses deterministic transfer functions to predict daily maximum or minimum temperature from free atmosphere variables. The connection between the temperature and rainfall scenarios is made in two ways. First, the sea level pressure data used to define the circulation weather types underpinning the rainfall scenarios are also used to construct predictor variables for the temperature transfer functions. Second, separate temperature transfer functions are developed for wet and dry days. When applying the downscaling methodologies, the rainfall scenarios are built first, and then used to determine on a day-by-day basis which transfer function should be used to predict temperature. The temperature scenarios proved to be a great improvement over raw GCM data, when compared to present-day observations. However, a cold bias was introduced in the downscaling methodology which in the Agri Basin was sufficiently large that temperatures in the 2030–39 scenario were lower than present-day observations.

In order to demonstrate the potential of the multisite, multivariable scenarios, we looked at the change in the occurrence of extreme events over the three scenario decades, using temperature as an example. We were able to demonstrate for extreme high and low temperatures that the changes are disproportionately large when compared to the change in the mean. These results showed clearly that we cannot assume the extremes of the parent distribution will linearly track the behavior of the mean. Rather, independent analyses of the occurrence of extremes should always be carried out.

Acknowledgments. This work was funded by the Commission of the European Union under Contract ENV4-CT95-0121 Mediterranean Desertification and Land Use Project 3: Regional indicators. The HadCM2 model data were supplied by the Climate Impacts LINK project (funded by the U.K. Department of Environment, Food and Rural Affairs).

REFERENCES

Agnew, M. D., and J. P. Palutikof, 2000: GIS-based construction of baseline climatologies for the Mediterranean using terrain variables. *Climate Res.,* **14,** 115–127.

Bárdossy, A., and H. J. Caspary, 1990: Detection of climate change in Europe by analysing European circulation patterns from 1881 to 1989. *Theor. Appl. Climatol.,* **42,** 155–167.

Brandsma, T., and T. A. Buishand, 1998: Simulation of extreme precipitation in the Rhine Basin by nearest-neighbour resampling. *Hydrol. Earth Syst. Sci.,* **2,** 195–209.

Brandt, C. J., and J. B. Thornes, 1993: Climatic variability in semiarid environments in Spain and Portugal. Final Rep. to the European Commission, Project EV4C.0091.UK(H), 156 pp. [Available from School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, United Kingdom.]

Buishand, T. A., and T. Brandsma, 1999: Dependence of precipitation on temperature at Florence and Livorno (Italy). *Climate Res.,* **12,** 53–63.

Caspary, H. J., 1996: Recent winter floods in Germany caused by changes in the atmospheric circulation across Europe. *Phys. Chem. Earth,* **20,** 459–462.

Charles, S. P., B. C. Bates, and J. P. Hughes, 1999: A spatiotemporal model for downscaling precipitation occurrence and amounts. *J. Geophys. Res.,* **104,** 31 657–31 669.

Conway, D., and P. D. Jones, 1998: The use of weather types and air flow indices for GCM downscaling. *J. Hydrol.,* **212–213,** 348–361.

——, R. L. Wilby, and P. D. Jones, 1996: Precipitation and air flow indices over the British Isles. *Climate Res.,* **7,** 169–183.

Cubasch, U., H. von Storch, J. Waszkewitz, and E. Zorita, 1996: Estimates of climate change in Southern Europe derived from dynamical climate model output. *Climate Res.,* **7,** 129–149.

Frei, C., C. Schär, D. Lüthi, and H. C. Davies, 1998: Heavy precipitation processes in a warmer climate. *Geophys. Res. Lett.,* **25,** 1431–1434.

Giorgi, F., and L. O. Mearns, 1999: Regional climate modeling revisited. *J. Geophys. Res.,* **104,** 6335–6352.

Goodess, C. M., 2000: The construction of daily rainfall scenarios for Mediterranean sites using a circulation-type approach to downscaling. Ph.D. thesis, University of East Anglia, Norwich, United Kingdom, 316 pp.

——, and J. P. Palutikof, 1998: Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling. *Int. J. Climatol.,* **18,** 1051–1083.

Hay, L. E., G. J. McCabe, D. M. Wolcock, and M. A. Ayers, 1991: Simulation of precipitation by weather type analysis. *Water Resour. Res.,* **27,** 493–501.

Hayhoe, H. N., 1998: Relationship between weather variables in observed and WXGEN generated data series. *Agric. For. Meteor.,* **90,** 203–214.

——, 2000: Improvements of stochastic weather data generators for diverse climates. *Climate Res.,* **14,** 75–87.

Hughes, J. P., P. Guttorp, and S. P. Charles, 1999: A non-homogeneous hidden Markov model for precipitation occurrence. *Appl. Stat.,* **48,** 15–30.

Johns, T. C., R. E. Carnell, J. F. Crossley, J. M. Gregory, J. F. B. Mitchell, C. A. Senior, S. F. B. Tett, and R. A. Wood, 1997: The second Hadley Centre coupled ocean–atmosphere GCM: Model description, spinup and validation. *Climate Dyn.,* **13,** 103–134.

Lerchl, A., 1998: Changes in the seasonality of mortality in Germany from 1946 to 1995: The role of temperature. *Int. J. Biometeor.,* **42,** 84–88.

Linés Escardó, A., 1970: The climate of the Iberian Peninsula. *Climates of Northern and Western Europe,* C. C. Wallén, Ed., Vol. 5, *World Survey of Climatology,* Elsevier, 195–226.

McGregor, J. L., 1997: Regional climate modelling. *Meteor. Atmos. Phys.,* **63,** 105–117.

Mearns, L. O., and F. Giorgi, 1999: Comparison of climate change scenarios generated from regional climate model experiments and statistical downscaling. *J. Geophys. Res.,* **104,** 6603–6621.

Murphy, J., 2000: Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *Int. J. Climatol.,* **20,** 489–501.

Palutikof, J. P., J. A. Winkler, C. M. Goodess, and J. A. Andresen, 1997: The simulation of daily temperature time series from GCM output. Part I: Comparison of model data with observations. *J. Climate,* **10,** 2497–2513.

Parmesan, C., T. L. Root, and M. R. Willig, 2000: Impacts of extreme weather and climate on terrestrial biota. *Bull. Amer. Meteor. Soc.,* **81,** 443–450.

Peixoto, J. P., and A. H. Oort, 1992: *Physics of Climate.* American Institute of Physics, 520 pp.

Rajagopalan, B., and U. Lall, 1999: A *k*-nearest neighbour simulator for daily precipitation and other weather variables. *Water Resour. Res.,* **35,** 3089–3101.

Richardson, C. W., 1981: Stochastic simulation of daily precipitation, temperature and solar radiation. *Water Resour. Res.,* **17,** 182–190.

Semenov, M. A., R. J. Brooks, E. Barrow, and C. W. Richardson, 1998: Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Climate Res.,* **10,** 95–107.

Trenberth, K. E., and J. G. Olson, 1988: An evaluation and comparison of global analyses from the National Meteorological Center and the European Centre for Medium Range Weather Forecasts. *Bull. Amer. Meteor. Soc.,* **69,** 1047–1057.

Trigo, I. F., T. D. Davies, and G. R. Bigg, 1999: Objective climatology of cyclones in the Mediterranean region. *J. Climate,* **12,** 1685–1696.

Trigo, R. M., and J. P. Palutikof, 1999: Simulation of daily maximum and minimum temperature over Portugal: A neural network approach. *Climate Res.,* **13,** 45–59.

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

Widmann, M., and C. Schär, 1997: A principal component and long-term analysis of daily precipitation in Switzerland. *Int. J. Climatol.,* **17,** 1333–1356.

Wilby, R. L., B. Greenfield, and C. Glenny, 1994: A coupled synoptic–hydrological model for climate change impact assessment. *J. Hydrol.,* **153,** 265–290.

——, 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.

Wilks, D. S., 1998: Multisite generalization of a daily stochastic precipitation generation model. *J. Hydrol.,* **210,** 178–191.

——, 1999a: Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agric. For. Meteor.,* **93,** 153–169.

——, 1999b: Simultaneous stochastic simulation of daily precipitation, temperature and solar radiation at multiple sites in complex terrain. *Agric. For. Meteor.,* **96,** 85–101.

——, and R. L. Wilby, 1999: The weather generation game: A review of stochastic weather models. *Prog. Phys. Geogr.,* **23,** 329–357.

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,** 2514–2532.

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