A simple statistical-dynamical downscaling scheme based on weather types and conditional resampling

J. Boé,[1] L. Terray,[1] F. Habets,[2] and E. Martin[2]

Received 15 November 2005; revised 12 June 2006; accepted 14 August 2006; published 15 December 2006.

[1] A multivariate statistical downscaling methodology is implemented to generate local precipitation and temperature series at different sites based on the results from a variable resolution general circulation model. It starts from regional climate properties to establish discriminating weather types for the chosen local variable, precipitation in this case. Intratype variations of the relevant forcing parameters are then taken into account by multivariate regression using the distances of a given day to the different weather types as predictors. The final step consists of conditional resampling. The methodology is evaluated in the Seine basin in France. Using reanalysis fields as predictors, satisfying results are obtained at daily timescale and concerning low-frequency variations, both for temperature and precipitation. The use of model results as predictors gives a realistic representation of regional climate properties. Nevertheless, as the validation of a statistical downscaling algorithm for present day climate conditions does not necessarily imply the validity of its climate change projections, the plausibility of the downscaled climate projections is assessed by verifying the consistency between spatially averaged downscaled results and direct model outputs for two climate change scenarios. Despite some discrepancies for precipitation with the more extreme scenario, the consistency is good for both local variables. This result reinforces the confidence in the use of the downscaling scheme in altered climates. Finally, it is shown that the intertype variations of the atmospheric circulation represent only a fraction of the climate change signal for the local variables. Thus a downscaling methodology based on weather typing should incorporate information concerning intratype modifications.

Citation: Boé, J., L. Terray, F. Habets, and E. Martin (2006), A simple statistical-dynamical downscaling scheme based on weather types and conditional resampling, *J. Geophys. Res.*, *111*, D23106, doi:10.1029/2005JD006889.

1. Introduction

[2] Water plays a central role in the behavior of the Earth system, and human activities are very dependent on water resources. The question of water cycle modifications under climate change conditions appears crucial, both to understand anthropogenic influence on climate and assess its impacts. Global modifications of precipitation are expected to be important in terms of mean but also in terms of statistical distribution [*Allen and Ingram*, 2002; *Trenberth et al.*, 2003]. To quantify the impacts of hydrological cycle modifications at watershed scale, a solution would consist in using an hydrometeorological model forced by the results of a Coupled Atmospheric-Oceanic General Circulation Model (AOGCM). A major difficulty however exists following this approach: hydrometeorological models need most of the time very high resolution forcings that AOGCM are unable to provide. At present, a typical resolution for an AOGCM is 300 km, whereas hydrometeorological models often need data with a resolution lower than 10 km. As a preliminary step, methodologies must be consequently used to derive the high-resolution forcings from the AOGCMs coarse resolution results: this is the downscaling issue.

[3] Several studies interested by the modification of different hydrological variables on French watersheds used a scale factor adjustment to obtain the high-resolution forcings required for the hydrometeorological models [*Etchevers et al.*, 2002]. Coarse-scale climate change projections were applied to a high-resolution observed climate baseline using the monthly anomalies between present and future climate simulation to modify current climate meteorological parameters. This methodology eliminates the mean biases due to the climate simulation upon the hydrometeorological forcings but the modifications that occur at the submonthly level (concerning dry and wet spells or daily extreme events for example) are not captured. To go further, other methodologies to bridge the scale gap between AOGCM and hydrometeorological models must be used.

[4] Two main families of downscaling techniques can be distinguished [*Mearns et al.*, 1999]. A first approach, or dynamical downscaling, is a model-based methodology

[1]Climate Modelling and Global Change Team, Centre Européen de Recherche et de Formation Avanceé en Calcul Scientifique, Centre National de la Recherche Scientifique, Toulouse, France.
[2]Centre National de Recherches Météorologiques, Météo-France, Toulouse, France.

0148-0227/06/2005JD006889$09.00

leading to sub-AOGCM grid-scale features by most of the time nesting a finer-scale Limited Area Model (LAM) within a GCM [*Giorgi et al.*, 1990]. The current generation of LAMs have a typical resolution of about 50 km. The second approach, statistical downscaling (SD), is based on the idea that regional climate is conditioned by two factors: the large-scale circulation (LSC) that is well resolved by the models, and small-scale features (e.g., land use, topography, land-sea distribution) that are not adequately described in GCMs [*von Storch*, 1999]. Thus an empirical relationship linking large-scale information ("predictors") and local or regional variables ("predictands") is first established for current climate. Then, applying this empirical relationship, the local variables for future climate are derived from the LSC simulated by an AOGCM. This approach is based on the strong hypothesis that the empirical relationship established for present climate is still valid under altered climate conditions [*Wilby et al.*, 2004]. This "stationarity assumption" is the major theoretical weakness of SD as it is not verifiable (note that this limitation also exists for dynamical model concerning the physical parameterizations).

[5] The goal of this paper is to describe and validate a new downscaling procedure intended to provide the high-resolution variables necessary to force the SAFRAN-ISBA-MODCOU (SIM) hydrometeorological model developed at Météo-France [*Habets et al.*, 1999] in order to investigate the impacts of climate change on the Seine basin hydrology in a future study. In this context, the downscaling methodology must deal with multiple variables at an 8 km horizontal resolution. It is based on an hybrid dynamical/statistical approach. *Terray et al.* [2004] studied the response to climate change in terms of wintertime North Atlantic weather regimes and suggest that improved model representation of the atmospheric circulation at regional scale is needed to achieve more reliable projections for anthropogenic climate change on European climate. Moreover, the quality of the LSC simulated by the model is a crucial point for statistical downscaling. For these reasons, a variable resolution GCM of the atmosphere with higher horizontal resolution over Europe is used to provide the predictors needed by the Statistical Downscaling Model (SDM).

[6] This paper is divided into seven sections. Section 2 is devoted to the description of the data and models used in this study. Section 3 deals with the construction of the statistical downscaling methodology and section 4 presents its validation for current climate. In section 5 the performances of the SDM using GCM outputs for current climate are assessed. In section 6 analyses based on climate change projection are described. The conclusions of our study are presented in section 7.

2. Data Sets and Model

[7] The need for a downscaling procedure comes from the objective to study the impacts of climate change on the hydrological cycle of the Seine basin using the SIM hydrometeorological coupled system. In this system, SAFRAN [*Durand et al.*, 1993] analyses the low-level and surface atmospheric variables needed by the surface scheme ISBA [*Noilhan and Planton*, 1989] such as precipitation, incoming longwave and shortwave radiation fluxes, wind speed, air temperature and humidity. ISBA is coupled with the distributed hydrological model MODCOU [*Ledoux et al.*, 1984]. The SAFRAN analysis takes into account all the available observations collected by Météo-France, as well as the operational analyses of the weather prediction model of Météo-France. Over the Seine basin, over 1000 rain gauges and 200 synoptic stations are available. The SAFRAN analyses are performed in homogeneous climatic zones, and then interpolated onto a regular 8-km grid given the altitude of the grid points (a vertical gradient of the variables is assumed). In this study, the predictands are daily SAFRAN precipitation and temperature, available from August 1985 to July 2003 on a domain that encompasses the entire Seine watershed and represents 2497 points (see domain D2 on Figure 1).

[8] For the construction of the SDM, the 500 hPa geopotential height (Z500) used as LSC predictor come from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA40 reanalysis. 0000 UTC, 0600 UTC, 1200 UTC, 1800 UTC values available on a $2.5^\circ \times 2.5^\circ$ resolution grid from September 1957 to August 2002 are daily averaged. The domain used for z500, hereafter D1, is presented on Figure 1 (see section 3.3 for a discussion about the sensitivity to the choice of the domain).

[9] As the overlapping period between ERA40 reanalysis and the SAFRAN data set (August 1985 to August 2002) is limited an alternative predictands data set is needed for a comprehensive validation of the SDM. To test the downscaling methodology over an independent period and to assess the realism of the low-frequency variations of predicted variables daily maximum and minimum temperatures and precipitation series from several meteorological stations within D2 (Figure 1) are extracted from the SQR (Série Quotidienne de Reference) Météo-France data set [*Moisselin et al.*, 2002] for the entire ERA40 period (September 1957 to August 2002).

[10] The global GCM used in this study is the variable resolution new version of the Météo-France Action de Recherche Petite Echelle Grande Echelle (ARPEGE) atmospheric model [*Gibelin and Déqué*, 2003]. The model uses semi-Lagrangian advection and a two time level discretization. Vertical discretization uses hybrid coordinates with 31 vertical levels. It has a T106 spectral truncation. The variable resolution allows one to increase the spectral and gridpoint resolution over a given region of interest. In the present case, the center of the high-resolution region is located in the middle of the Mediterranean basin. The highest horizontal resolution is about 0.5° and remains fairly high over the entire North Atlantic–European sector because of a weak resolution gradient. To test the downscaling scheme a simulation has been performed for the current climate (1950–1999) where the model is forced by monthly mean observed sea surface temperature (SST), historical greenhouse gas (GHG) and sulfate aerosols concentrations. For future climate two atmospheric simulations realized within the PRUDENCE (Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects) project have been employed using SST forcing and Intergovernmental Panel on Climate Change (IPCC) A2 and B2 Special Report on Emissions Scenarios (SRES) of future GHG and sulphur emissions. The SST boundary forcings are combination of observed

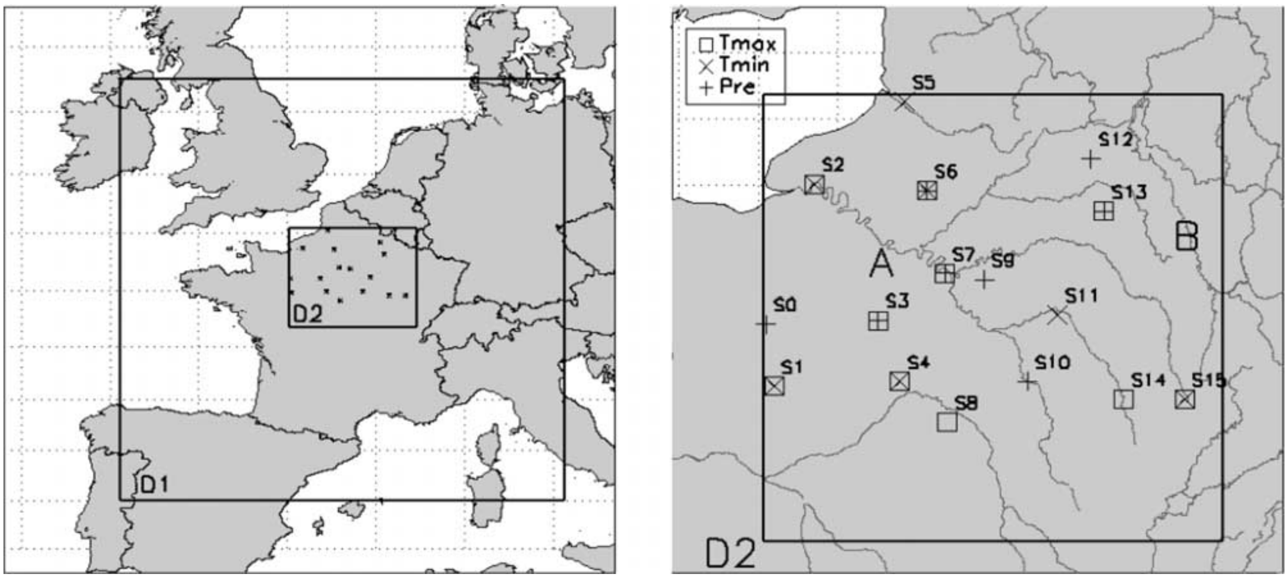

Figure 1. Location of the study area. (left) Domain used for the atmospheric predictors (D1) and domain of SAFRAN predictands (D2). (right) Zoom on D2 and location of the meteorological stations used as alternative predictands data set. Tmax and Tmin are daily maximum and minimum temperature, respectively, and Pre is precipitation.

SSTs and mean SSTs changes derived from transient simulations with the ARPEGE-Ocean Parallelisé (OPA [*Royer et al.*, 2002]) AOGCM for the B2 simulation and the Third Hadley Center Coupled Ocean-Atmosphere General Circulation Model (HAD CM3 [*Jones et al.*, 2003]) for the A2 simulation. As the present day simulation described above is based on a slightly different version of the ARPEGE model, a control simulation with the same version as the future climate simulations is also used. It is forced by monthly mean observed SST and historical GHG concentrations for the period 1960–2000.

3. Downscaling Methodology

3.1. Concepts

[11] The most intuitive statistical downscaling approach is probably the analog method. This SDM is based on the idea that the same causes (here the LSC as predictor) produce the same effects (for the predictands, i.e., the regional climate). To obtain the local variables for a particular day, the one with the most similar LSC pattern is searched on the past observations given a measure of distance. Lorenz introduced the analog approach in the field of weather forecasting in 1969 but its use for downscaling purposes is more recent [*Zorita et al.*, 1995; *Martin et al.*, 1997; *Timbal*, 2004]. This method which allows to deal with spatial and multivariate problems in a quite easy way gives satisfactory results. It often favorably compares with more sophisticated techniques [*Zorita and von Storch*, 1999] and can thus be considered as a natural "benchmark" method when developing a SDM. However, two difficulties arise with the analog approach. First, the main weakness of all the empirical downscaling methods is that their basic assumption (i.e., that the statistical relationship established for present climate is still valid under altered climate) is not verifiable. In order to weaken this stationarity hypothesis it is preferable to build a SDM that yields physically interpretable linkage between LSC predictors and regional climate [*Wilby et al.*, 2004]. It is not the case for the analog method.

[12] In addition, considering that the same causes in terms of LSC give the same effects for the regional climate is an important approximation. It is better to see the regional climate as a random process conditioned upon a driving LSC [*von Storch*, 1999]. As the regional climate is not completely determined by the LSC, quasi-identical LSC patterns can have large different effects in terms of regional climate [*Roebber and Bosart*, 1998]. Following this view, instead of searching only for the day with the nearest LSC pattern it could be preferable to search for an ensemble of days with similar LSC patterns and to consider the statistical distribution of the regional variables for these days.

[13] The classical analog approach can be generalized in a k-nearest neighbors analog method based on an ensemble of analog days [*Gutiérrez et al.*, 2004], but the first drawback still remains. Another possibility consists in using a small number of weather types. Each day is classified in a weather type and local variables are attributed depending on this type. Two main issues are to address. The weather types, defined in terms of LSC similarity, should bring enough information concerning the regional climate and a procedure to link local variables and weather types is necessary. Moreover, to be really more attractive than the analog method, the weather typing approach should be based on a small number of weather types (as the analog method can be seen as a limit case of weather typing where each day defines a weather type), in order to examine the physical mechanisms that support the statistical model.

[14] During the last decades, because of the development of high-speed computers, objective automatic classification algorithms have been developed to complement older subjective schemes. In particular, the k-means algorithm

[*Wilson et al.*, 1992] is now widely used in climate research. Given a prescribed number of clusters k and a measure of similarity it searches to produce k clusters of greatest possible distinction. The classifications in weather regimes by automatic classification algorithm can be useful when it comes to explore the links between LSC and regional climate [*Zhang et al.*, 1997; *Corte-Real et al.*, 1999; *Cassou et al.*, 2005]. Nevertheless, *Plaut et al.* [2001] showed that the classical North Atlantic weather regimes are not very discriminating for the French Alps precipitation. In order to obtain discriminating LSC patterns, they perform a preselection of the days characterized by heavy precipitation events (defined in terms of fixed precipitation threshold) and then classify the corresponding LSC patterns with the k-means algorithm. This procedure leads to highly discriminating weather types for Alpine precipitation. This idea of a bottom-up approach, which starts from local variable properties to define discriminating weather types is retained in this study. A data-driven method more adapted to the context and goal of this study is used to establish the precipitation classes: the k-means algorithm is first used to separate the precipitations within the domain D2 in a few characteristic types. The LSC pattern of the days belonging to each precipitation types is then classified with the k-means algorithm. The priority is given to precipitation when building the weather types as it is the most important variable for hydrological application but it will be tested if the weather types are also discriminating for temperature.

[15] The choice of predictor(s) is a major issue for statistical downscaling. Predictor(s) must have strong predictive skills for the predictands in present climate, but also be sensitive to climate change signal [*Wilby et al.*, 1998]. Conversely, a predictor that does not seem to be important for present climate could become essential under perturbed climate conditions. For instance, it is suspected that future changes in surface temperature will be dominated by changes in the radiative properties of the atmosphere, rather than by circulation changes [*Schubert*, 1998]. Thus the use of a single dynamical predictor could be problematic when one is interested in temperature. Another example concerns precipitation: the changes in the water-holding capacity of the atmosphere (linked with temperature change through the Clausius-Clapeyron equation) is supposed to play an important role on precipitation changes [*Allen and Ingram*, 2002]. Coherently, some studies show that the inclusion of humidity as a predictor can have significant impacts on the results of a SDM [*Charles et al.*, 1999]. In addition, the predictors should also be realistically simulated by the GCMs and this condition might be a major constraint for the choice of predictors. In general, more confidence is given to the simulated large-scale dynamical variables. The use of thermodynamical variables like surface temperature or humidity, for which the local and surface effects can be important is more questionable. This issue will be addressed in the following. A downscaling procedure which only uses a LSC predictor will be compared to one that incorporates in addition temperature or humidity information from the model.

3.2. Application: Weather Typing Approach

3.2.1. Derivation of the Weather Types

[16] The first step of the downscaling process consists in establishing the weather types (step 1 on Figure 2). A k-means classification of SAFRAN precipitations is realized on D2 for the learning period (August 1985 to August 2002). Three seasons are considered. Traditional summer (June to August, JJA) and winter (December to February, DJF) seasons are kept and spring and autumn days are gathered to constitute the third season for the classification. Only configurations with two or three clusters per season are tested as the spatial variations of precipitations are quite homogeneous on the domain.

[17] The days belonging to each precipitation cluster are then classified depending upon their LSC. Daily maps of 500 hPa geopotential height are classified within each precipitation cluster over the domain D1 with the k-means algorithm in the subspace of the first ten Principal Components (PC), accounting for more than 95% of the explained variance. This number of PCs is considered sufficient to capture the main features of Z500 variability. Here, the PCs have not been scaled with their eigen values to have unit length. Nevertheless, note that when the PC are scaled or when the number of PCs retained is modified, the results of the classification are essentially the same. The major drawback of the k-means algorithm is that the number k of clusters must be chosen a priori. Different approaches are used to determine the optimal value of k, but no consensus exists. Here, two tests are employed: the test based on a classifiabilty index described by *Michelangeli et al.* [1995], and the test based on a inter/intra cluster variance ratio used by *Straus and Molteni* [2004]. Following those two approaches, k varying between two and five are tested for each precipitation cluster. Only solutions for which the two tests are coherent are kept. When there is no valid solution, a composite including all the days within the precipitation cluster is alternatively used. As several valid combinations still remain after the two tests, the most discriminating combination of weather types in terms of precipitation is searched. An inter/intra cluster variance ratio is computed for precipitation for each valid combination of weather types and the configuration with the highest ratio is chosen.

[18] The precipitation clusters obtained mainly differ by their global intensity on D2. For the sake of simplicity they will be thus hereafter designated given their observed characteristics (not shown) as "dry," "wet" or "very wet." Table 1 shows the number of weather types obtained for each season and associated with the three precipitation clusters.

[19] The procedure described above based on precipitation clustering is only a preliminary step to establish the LSC weather types. Once the weather types have been obtained, each day is associated to a weather type depending only on it LSC. The nearest weather types in terms of Euclidian distance between the daily map of Z500 and the weather types is searched (step 2 on Figure 2). Figure 3 shows the Z500 anomalies that correspond to each winter weather type and Table 2 presents spatially averaged characteristics of rainfall and temperature for each winter weather type. WT0, related to the "very" wet precipitation cluster, is characterized by a dipole pressure pattern with a very strong Z500 negative anomaly centered over the British isles, and a weaker positive anomaly over Algeria. The pronounced zonal flow over the Seine basin leads to very warm and wet conditions on D2 (Table 2). For this weather type, 98% (96%) of the days have greater rainfall

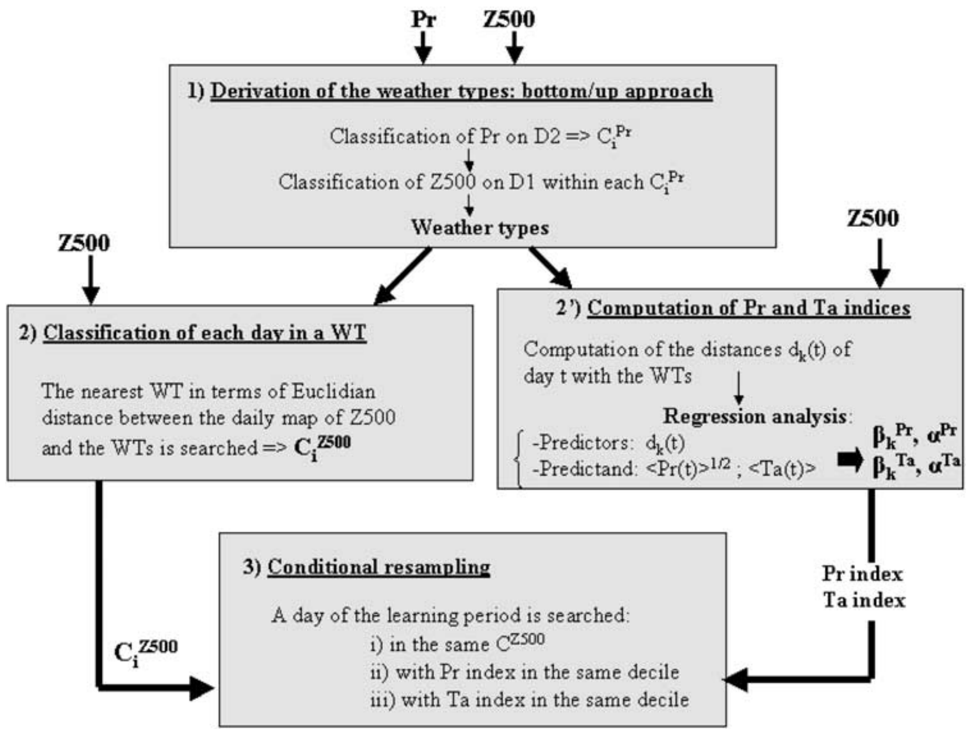

Figure 2. Flowchart presenting the main step of the downscaling algorithm "weather typing 1" for a season. *Pr* is precipitation, *Ta* is temperature and Z500 is 500 hPa geopotential height. *WT* is weather type. C_i^{Pr} is the ensemble of the days that belong to the precipitation cluster i. C_i^{Z500} is the ensemble of the days that belong to the weather type i. The β_k are the regression coefficients and α the constant of the regression equations (index Ta for temperature and index Pr for precipitation).

(temperature) than the global winter median. WT1, obtained from the "wet" precipitation cluster, is quite spatially similar to WT0, although less intense. WT1 is also related to warm and wet anomalies over D2, but less intense than with WT0. The last three weather types are derived from the "dry" rainfall cluster. WT2 features a strong anomalous high over the North Sea and corresponds to very dry condition on the Seine basin. WT3 is close to the mean state both in terms of rainfall and LSC pattern (note that this weather type occurs only 26% of the time). With a negative Z500 anomaly centered on the study area, WT4 circulation is characterized by anomalous northeasterlies leading to advection of cold air and snow type events. 64% of snowfall amounts occur within this weather type (which represents only 0.10 mm/day as snowfall on the domain D2 is not very large).

3.2.2. Downscaling Algorithms

[20] Weather is a continuous process and discretizing it into discrete states can rise issues in downscaling context as noted by *Mehrotra and Sharma* [2005] who present an alternative approach based on continuous weather states. In particular, intratype variations concerning the regional climate may not be captured [*Wilby et al.*, 2004]. The methodology described above leads to weather types that are really discriminating for precipitation on D2. They also prove their efficiency in discriminating temperature events even if they have not been explicitly constructed for the latter. Nevertheless, when looking at the standard deviation of rainfall amounts and temperature for each weather type in Table 2 it is also seen that intracluster variability in terms of regional climate is not negligible. Intracluster regional climate variability has two causes. A spread of LSC patterns within a given weather type exists and the regional climate is not entirely determined by the LSC. Considering the similarity between the circulation of a particular day and the weather types is a way to deal with the LSC intracluster variability. In particular, *Plaut* [2004] showed that strong relationships exist between precipitation in the Alpes Maritimes (France) and the similarity to some particular circulation patterns. Here, the distances to the weather types are used to better capture intracluster variability. The distance $d_k(t)$ between the LSC of day t and the weather type k is measured by an Euclidean distance computed over the ten first principal components of Z500 (step 2′ on Figure 2):

$$d_k^2(t) = \sum_{i=1}^{10} \left[a_i(t) - A_i^k\right]^2 \quad (1)$$

where a_i refers to Z500 PC scores and A_i^k to the coordinates of the weather types k in the PC-space. PC have been scaled to have unit variances.

Table 1. Number of Weather Types Retained for Each Season Depending on the Precipitation Class

Precipitation Cluster	Winter	Summer	Spring/Autumn
Very wet	1	1	1
Wet	1	1	1
Dry	3	4	4

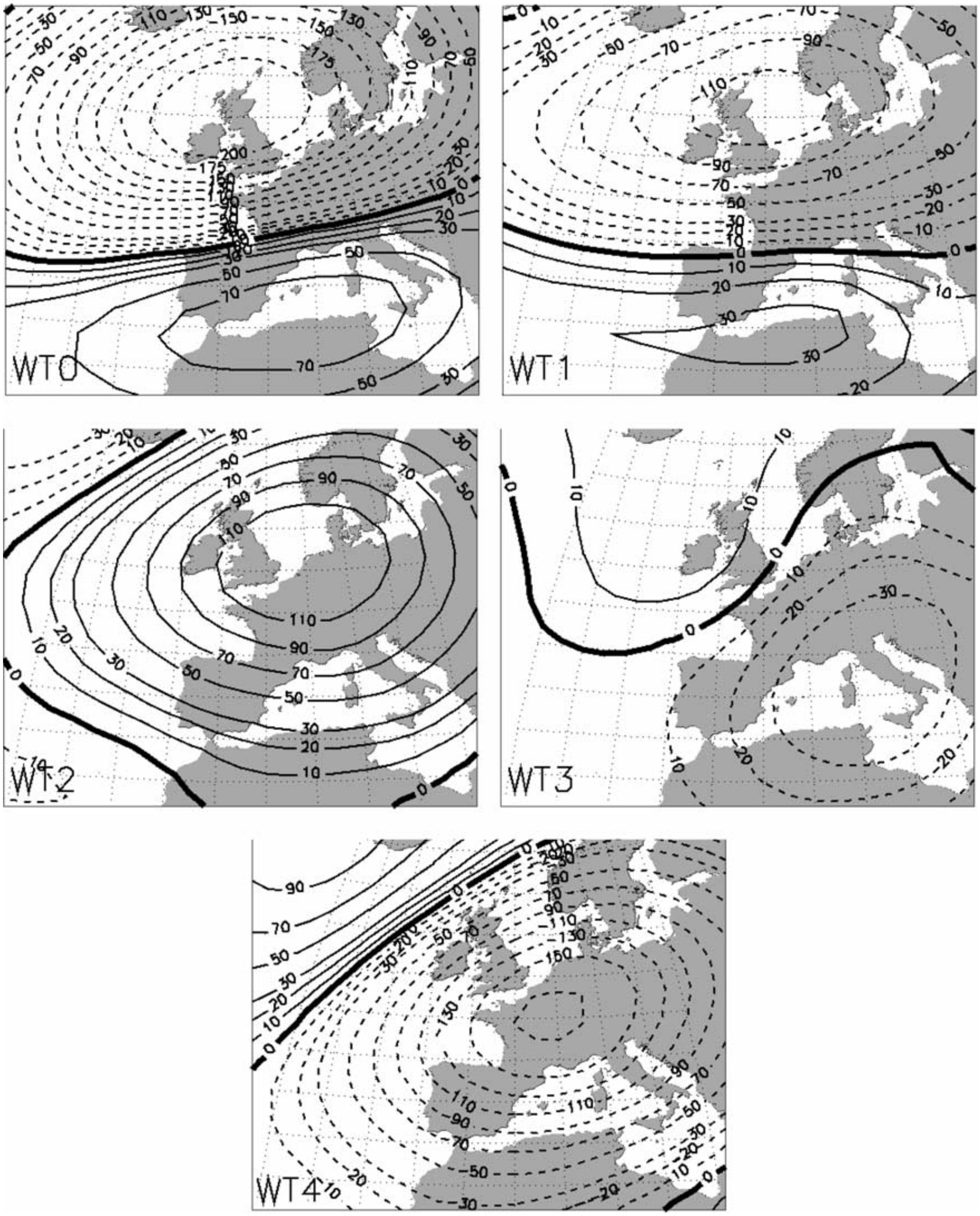

Figure 3. Winter weather types: Z500 anomalies composite (gpm). WT0 corresponds to the very wet precipitation cluster, WT1 to the wet precipitation cluster and WT2, WT3 and WT4 to the dry precipitation cluster.

[21] The N daily series of distances corresponding to the N weather types of the season are used as predictors in a multiple linear regression. The predictand is the spatial mean of precipitation on D2, $\langle Pr(t)\rangle$ (here and hereafter the symbol $\langle . \rangle$ refers to the spatial mean on D2). Because of the skewed character of daily precipitation, the precipitation series is transformed before solving the regression equation (note also that only the spatial average, less skewed, is used here). A square root transformation is applied. For the learning period the following regression equation is thus solved:

$$\langle Pr(t)\rangle^{1/2} = \sum_{k=1}^{N} [\beta_k . d_k(t)] + \alpha + e(t) \qquad (2)$$

where the β_k are the regression coefficients, α a constant and $e(t)$ the residuals. The high multiple linear correlation

Table 2. Area-Averaged (D2) Rainfall and Temperature Characteristics Depending Upon the Weather Type for Winter (1985–2002)[a]

	WT0	WT1	WT2	WT3	WT4	All Days
Occurrence	9.1%	8.9%	40.5%	23.6%	17.7%	
Percentage of days with rainfall > global median	98%	88%	27%	55%	53%	50%
Rainfall amounts, mm/day	7.50 (5.27)	4.63 (3.93)	0.79 (1.59)	2.06 (2.99)	2.02 (2.69)	2.27 (3.49)
95th percentile of rainfall amounts, mm/day	19.47	12.62	4.36	7.82	7.13	9.26
Percentage of days with temperature > global median	96%	81%	49%	43%	21%	50%
Mean temperature, K	281.1 (2.23)	279.2 (2.98)	277.5 (3.71)	276.4 (3.71)	274.2 (4.04)	277.1 (4.06)
95th percentile of temperature, K	284.6	282.8	283.5	281.8	279.6	283.3

[a]Standard deviation is shown in brackets.

coefficients obtained (between 0.64 and 0.71 according to the season) indicates that a strong physical link actually exists between the distances to the weather types and the precipitation on D2 at the daily level. During the downscaling procedure, the regression equation is used to compute a precipitation index that only depend upon LSC (Z500). As the temporal variance of the precipitation index computed by linear regression is underestimated, its direct values are not used in the algorithm. Instead, the values of the index are classified into ten equally populated categories and only these categories are considered. This procedure is similar to the one described in *Plaut* [2004].

[22] To take into account temperature, two different approaches are tested. In the first one, a temperature index is computed using a multiple linear regression with the distances to the weather types as predictors as in the precipitation case. A high multiple linear correlation coefficient is also obtained (between 0.66 and 0.77). This method, named hereafter "weather typing 1," only uses a large-scale dynamical predictor (Z500). Note that this scheme does not directly account for the radiative component associated with global temperature increase. This scheme is thus likely to be partly insensitive to climate change signal, in particular for temperature. To overcome this intrinsic drawback and to assess its potential associated improvement an alternative procedure is proposed. In this second procedure the temperature index based on LSC described above is not used. Instead, the direct temperature from the model for which the downscaling is needed (or from ERA40 in case of validation) is averaged on D2 and used as temperature index. This downscaling algorithm is hereafter named weather typing 2.

[23] In the two approaches the final step of the downscaling algorithm is identical. It consists of conditional resampling from the historical record (step 3 on Figure 2).

[24] To recapitulate, once the weather types have been derived, three steps are necessary to downscale GCM outputs or reanalysis data for a particular day (step 2,2′, and 3 on Figure 2). First, the day is classified in the nearest weather type. Next, using the regression coefficients estimated for the learning period, the value of the precipitation index for this day is computed. For temperature, with the weather typing 1 method, the temperature index is computed in the same way as for precipitation. With weather typing 2, the temperature from the GCM averaged on D2 is used as temperature index. Finally, a day from the learning period that belongs to the same weather type, and with precipitation and temperature indices belonging to the same deciles is randomly chosen. If there is no overlapping day between the three conditions in the learning period, those concerning the weather type and the precipitation index remain, and the neighboring quantiles are searched for temperature.

[25] In the following, the results obtained with the weather typing approaches are compared with those obtained by the analog method. A detailed description of the analog method is presented in *Zorita and von Storch* [1999]. In our case, the distance between the circulation of two days necessary in the analog algorithm is computed in the space spanned by the ten first principal components of the predictor. The LSC domain (D1) and predictor (Z500 only) are the same as for weather typing 1 in order to allow the comparison of the two methods.

3.3. Sensitivity to Different Parameters

[26] Several choices have been required during the construction of the SDMs. Mainly, the domain for the LSC predictor, the LSC variable used as main predictor, and the seasonal stratification during the construction of the weather types. To evaluate the impacts of these choices, different sensitivity tests have been performed. Different domains for LSC predictors have been tested: from a small domain centered on France to the entire North Atlantic sector. A large domain sometimes gives an improved representation of spell properties, in particular for the analog method (as the LSC on the North Atlantic sector for a particular day provides also information about the LSC for Europe for the forthcoming days). Nevertheless, the correlations between original and downscaled series with predictors from ERA40 reanalysis quickly decrease when the domain is enlarged, in particular for precipitation. To account for a realistic simulation of the LSC by the GCM, the domain must also not be too small. The domain D1 finally used here appeared as the best compromise. Mean sea level pressure and 700 hPa geopotential height have been tested instead of Z500. The overall performances of the SDMs used in this study are rather insensitive to the choice of the LSC predictor. For the weather classification three seasons have been chosen (winter (DJF), summer (JJA), and the rest of the months). It appeared important to keep only DJF months for winter and JJA months for summer. When these seasons are extended a deterioration of the downscaling results is seen. Conversely, separate spring and autumn days for the classification does not improve the results. With the weather typing methods, at the conditional resampling stage, ten categories are used for the temperature and precipitation indices. This is a conservative choice allowing for large resampling sets for the majority of the days. More categories can improve the results of the two downscaling schemes

Table 3. Skill Achieved by the Different Downscaling Models for Daily Temperature and Precipitation (1985–2002)

Variable	Analog	Weather Typing 1	Weather Typing 2
		Precipitation	
Correlation	0.17	0.30	0.31
Brier score	18	29	30
Ind	25	35	35
		Temperature	
Correlation	0.85	0.85	0.93
Brier score	85	86	93

with the ERA40 predictors, but with the ARPEGE ones, the picture is less clear.

4. Validation of the Downscaling Algorithms

4.1. SAFRAN Predictands

[27] The two weather typing methods are applied and the results are compared with those obtained by the analog method. In this section, Z500 comes from the ERA40 reanalysis and predictands are SAFRAN daily precipitation and temperature. Downscaling results are evaluated for the entire learning period (August 1985 to August 2002). Note that when searching the analog of a specific day, the latter and its ten predecessors and successors are rejected from the possible resampling set in order to avoid artificial skill. The predictand time series are reconstructed day by day and the relative performances of the three SDMs are assessed by first comparing different skill scores already used in *Timbal et al.* [2003]. Linear correlations are computed between observed and reconstructed series (unless specified otherwise, here and hereafter all the correlation coefficients given are significant at the 0.05 level, using the "random-phase" test accounting for autocorrelation described in *Ebisuzaki* [1997]). Brier score (BS) is defined as:

$$BS = 100 \times \left[1 - \frac{MSE}{MSE_{ref}}\right] \quad (3)$$

where MSE is the Mean Square Error for downscaled series and MSE_{ref} the one for a reference scheme: here, a purely random choice of analogs. BS varies between 100% (perfect forecast) and 0% (random choice of analogs). For precipitation, an additional score (Ind) quantifying the skill of the SDM in reproducing rain occurrence is defined as:

$$Ind = 100 \times \left[1 - \frac{m}{w+m} - \frac{f}{d+f}\right] \quad (4)$$

The letter w stands for a wet day both forecast and observed, d to a dry day both forecast and observed, m to a wet day missed by the forecast and f to a dry day missed by the forecast (here and hereafter a dry day is defined as a day with no rain). For a perfect forecast Ind equals 100%. Scores are computed for each grid point and then averaged over the entire domain.

[28] For temperature, as expected since ERA40 temperature information is incorporated through the downscaling process, the skill scores are better for weather typing 2 than others (Table 3). Weather typing 1 and analogs give very similar results. Concerning precipitation, the two weather typing schemes perform better than the analog method. Brier scores and correlations are weaker than for temperature (precipitation is well known to be a difficult variable to downscale), but skill exists. Note that at the monthly level the correlations between reconstructed and original precipitation series is much higher (between 0.63 and 0.75 on D2 with weather typing 1).

[29] Correlation between downscaled and original series can be a good indicator of the relative performances of different SDMs, but it might be misleading. To assess the hydrological impacts of climate change, it is important that statistical properties like first and second moments, persistence, extremes are well represented. However, in the empirical downscaling context, the objective to obtain the best correlation and the objective to obtain the best representation of some statistical properties, like daily variance, are often conflicting. For example, regression-based methods are known to give good time correlation but also to tend to strongly underestimate temporal variance [*Zorita and von Storch*, 1999]. Here, with an analog based approach, a better correlation could also be achieved. If instead of considering only a single analog, the means of local variables are computed for the five days with the most similar LSC patterns, the correlation for precipitation practically doubles but the daily variance becomes greatly underestimated.

[30] Table 4 introduces several diagnostics of daily precipitation concerning first and second moments, extremes, persistence, used by *Wilby et al.* [1998]. These diagnostics are computed for original and downscaled precipitations (Table 5). Each diagnostic is computed on each point of D2 for the four seasons. Mean and RMSE are then computed on these values (4 seasons by 2497 points values). Globally, the results from the different SDMs are close to the observations. Mean, SD wet, 95% wet and p_d are well represented (but a purely random resampling of the learning period days would also be successful). Persistence properties (p_{dd}, p_{ww} and spell diagnostics) are also well reproduced by the weather typing methods. These latter properties are very important for hydrological applications and more difficult to achieve. The major differences between weather typing and analog methods concern p_{dd}, p_{ww} and L_d as the analog method underestimates these properties.

[31] Mean, standard deviation, 95th percentile were also computed for downscaled and original temperature series. These properties are very well reproduced by all the methods (not shown).

Table 4. Standard Precipitation Diagnostics

Daily Precipitation Diagnostics	
Mean	mean wet-day amount, mm
SD wet	standard deviation of wet day amount, mm
95% wet	95th percentile of wet day amount, mm
p_d	unconditional probability of a dry day
p_{dd}	probability of a dry day conditional on the previous day being dry
p_{ww}	probability of a wet day conditional on the previous day being wet
Spell Diagnostics	
L_w	mean wet spell length in days
L_d	mean dry spell length in days

Table 5. Spatial Mean and Root Mean Square Errors (RMSE) for Precipitation Diagnostics With Different SDMs[a]

		Analog		Weather Typing 1		Weather Typing 2	
Diagnostic	Observations, Mean	Mean	RMSE	Mean	RMSE	Mean	RMSE
Mean	4.51	4.53	(0.18)	4.60	(0.26)	4.56	(0.20)
SD w.	5.49	5.45	(0.35)	5.63	(0.48)	5.59	(0.36)
95% w.	15.6	15.6	(0.91)	16.0	(1.60)	15.9	(1.26)
p_d	0.51	0.50	(0.011)	0.50	(0.011)	0.50	(0.011)
p_{dd}	0.67	0.63	(0.049)	0.67	(0.023)	0.67	(0.020)
p_{ww}	0.66	0.62	(0.044)	0.66	(0.026)	0.66	(0.024)
L_d	4.80	4.07	(0.79)	4.46	(0.47)	4.52	(0.42)
L_w	4.15	3.91	(0.36)	4.43	(0.43)	4.40	(0.40)

[a]Mean and RMSE are computed on 9988 values (2497 grid points by four seasons) for the 1985–2002 period.

4.2. Alternative Predictands Data Set

[32] When validating a SDM, it is particularly important to test its ability to reproduce low-frequency variations of past climate, like trends or oscillations, as it can be considered as a sort of "natural" climate change [*Zorita and von Storch*, 1999]. Several downscaling methods, in particular stochastic weather generators and also some circulation-based methods, often greatly underestimate low-frequency variability [*Wilby et al.*, 1998]. As the length of the SAFRAN data set is quite short, the validation of the low-frequency variability of downscaled variables is problematic. Moreover, the previous diagnostics were computed for a period identical to the learning period, which can hamper the conclusions. To address these two problems, several meteorological station observations are now used as alternative predictands (section 2, Figure 1). Eight precipitation, ten maximum temperature (Tmax) and seven minimum temperature (Tmin) stations are available with no missing values for the complete ERA40 reanalysis period (1957–2002) on D2. Five stations provide both Tmax and Tmin, and hereafter temperature indicates the mean between Tmax and Tmin for these stations. The three SDMs are applied with these new predictands. The local variables are reconstructed for all the ERA40 period (1957–2002) with the same learning period as previously (1985–2002).

[33] RMSE and mean correlation between reconstructed and original local variables series for the learning period, the independent period (September 1957 to August 1985) and the entire period are given in Table 6. The results obtained for the learning period and for the independent period are very similar. Here, using the same period for the construction and the validation of the SDMs does not lead to artificial skill.

[34] The interannual variability of original and downscaled series is now explored. Table 7 shows the mean correlations between seasonally averaged series of temperature and precipitation. Concerning precipitation, the two weather typing approaches clearly outperform the analog method for all the seasons. The best results are obtained for winter and the worst for summer independently of the SDMs. The convective nature of precipitation in summer probably explains the weaker link between LSC and precipitation for this particular season.

[35] As expected by construction, for temperature, weather typing 2 is the best SDM in all seasons. For summer and winter weather typing 1 gives also very good results whereas for spring and autumn correlations are weaker. The worst results are obtained with the analog method whatever the seasons.

[36] To further pursue the study of low-frequency variations at different timescales, downscaled and original daily series of temperature and precipitation are filtered using different Hanning window width, and correlations between observed and downscaled series are computed. Results are plotted on Figure 4.

[37] In order to see if the use of humidity predictor from ERA40 reanalysis allows to better capture time variability of precipitation, an additional weather typing method is also used. It corresponds to a modified version of weather typing 2 where temperature as secondary predictor is replaced by ERA40 1000 hPa humidity (this method is named weather typing 2bis). With this alternative secondary predictor, at daily timescale, the results for the diagnostics described in section 4.1 are very similar with those obtained with temperature predictor (not shown). For temperature, up to 365 days, all the methods give high correlations. Once the annual cycle is filtered only weather typing 2 is able to well reproduce the temporal variations while the other methods lead to an important decrease of correlation. The difference between weather typing 1 and weather typing 2 results seems to indicate that a part of temperature low-frequency variability is not driven by LSC as the use of temperature as secondary predictor seems to be required to

Table 6. Mean Correlation and RMSE (in Brackets) Between Reconstructed and Original Daily Precipitation and Temperature Series for Different Subperiods

Variable	Analog	Weather Typing 1	Weather Typing 2
Precipitation 1957–1985	0.17 (5.24)	0.22 (5.00)	0.23 (5.00)
Precipitation 1985–2002	0.15 (5.42)	0.25 (5.11)	0.29 (4.91)
Precipitation 1957–2002	0.16 (5.31)	0.23 (5.05)	0.25 (4.97)
Temperature 1957–1985	0.84 (3.59)	0.84 (3.65)	0.91 (2.62)
Temperature 1985–2002	0.84 (3.54)	0.84 (3.62)	0.92 (2.51)
Temperature 1957–2002	0.84 (3.57)	0.84 (3.64)	0.92 (2.56)

Table 7. Mean Correlation Between Reconstructed and Original Seasonal Precipitation and Temperature Series (1957–2002)

Variable	Analog	Weather Typing 1	Weather Typing 2
Precipitation winter	0.61	0.69	0.70
Precipitation spring	0.44	0.57	0.61
Precipitation summer	0.38	0.50	0.46
Precipitation autumn	0.51	0.66	0.69
Temperature winter	0.75	0.83	0.94
Temperature spring	0.60	0.69	0.91
Temperature summer	0.65	0.79	0.91
Temperature autumn	0.48	0.52	0.81

capture this signal. The increase of GHG concentration and its associated direct radiative signal on temperature that is not captured by LSC changes is hypothesized. After 365 days, weather typing 1 clearly outperforms analog methods while the performances of weather typing 2bis continue deteriorating. For precipitation, all the weather typing methods perform greatly better than analogs especially after 100 days. Even if the correlations between original and reconstructed precipitation series at the daily level are limited (section 4.1, Table 3) it is seen that with the weather typing methods much higher correlations are obtained at lower frequency scales, indicating an actual strong physical link between LSC and precipitation on D2. The use of humidity as secondary predictor is not conclusive as it does not allow to better capture precipitation variations and greatly deteriorates temperature results. A problem concerning the quality of the humidity field in ERA40 could be hypothesized (for example, in a different context, *Bengtsson et al.* [2004] judge the trend of integrated water vapor for 1958–2001 too large and assume that the overestimation is due to changes in the observing system).

[38] Regarding some studies, significant linear trends can be detected during the 20th century on some climate extremes indices [*Klein Tank and Können*, 2003]. The ability of a downscaling algorithm to reproduce this type of trends would reinforce our confidence in its use in climate change studies. The 90th percentile of Tmin for each summer and the 90th percentile of Tmax for each winter are computed for observed and downscaled series and trends are analyzed on the 1957–2002 period. The learning period for downscaling is the same as previously (1985–2002). The majority of stations shows a linear trend significant at the 0.05 level.

[39] For Tmin (Table 8), the analog method is globally unable to reproduce the trend of summer 90th percentile, whereas the weather typing 2 method reproduces them fairly well, despite a slight underestimation. For Tmax (Table 9), all the downscaling methods reproduce the trends. Analogs and weather typing 2 slightly overestimate the value for a majority of stations. It is worth noting that the analog method, which only uses a dynamical predictor is able to well reproduce the trends of the 90th percentile of Tmax and Tmin (not shown) in winter. The warm extremes modifications seem to be linked to atmospheric circulation changes in winter. Conversely, in summer, the analog method is unable to reproduce the trends of Tmin extremes, and poorly reproduce the trends of Tmax extremes (underestimation by a factor three, not shown). It suggests that the direct radiative effect of enhanced greenhouse effect plays a major role in temperature extreme changes for this season. If a downscaling algorithm is not able to reproduce long-term variations for the present climate, it is doubtful that it can well represent anthropogenic climate change. Our analyses indicate that the two weather typing methods we described in this paper globally capture both climate oscillations and trends, whereas the success of the analog method is more limited. Considering, in addition, skill scores and daily statistics computed in the previous section, it appears that the two weather typing approaches clearly outperform the analog method.

5. Application to Present Climate Simulation

5.1. Model Validation

[40] Results shown in section 4 indicate that the weather typing methods are able to well represent regional climate

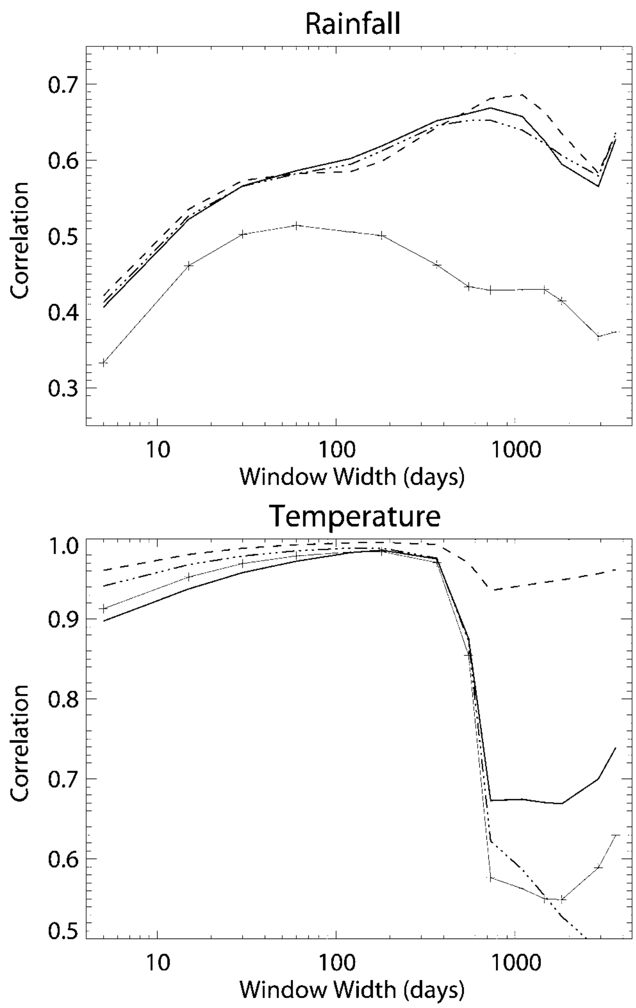

Figure 4. Mean correlation between downscaled and observed series (1957–2002) as a function of the width of the Hanning window filtering applied to the series. Thin line with crosses indicates analog. Thick line indicates weather typing, solid line indicates weather typing 1, dashed line indicates weather typing 2, and dotted-dashed line indicates weather typing 2bis (temperature is replaced by humidity as predictor).

Table 8. Linear Trend for the 90th Percentile of Minimum Temperature in Summer (K by Decade), 1957–2002[a]

Station	Observations	Analog	Weather Typing 1	Weather Typing 2
S11	**0.28**	0.02	**0.10**	**0.19**
S4	**0.41**	0.05	**0.20**	**0.26**
S15	**0.31**	0.05	**0.20**	**0.33**
S6	**0.36**	−0.07	**0.20**	**0.28**
S1	**0.42**	0.04	**0.17**	**0.31**
S2	**0.47**	**0.16**	**0.16**	**0.29**
S5	**0.33**	−0.04	**0.11**	**0.21**

[a]The values significant at the 0.05 level are presented in bold.

when using predictors from ERA40 reanalysis. The ability of the SDMs must now be evaluated using GCM results as predictors. Here, the variable resolution version of the ARPEGE model with high resolution on the study area (around 60 km on D1) is used to provide suitable predictor(s). First, a brief validation of the GCM results in the downscaling context is carried out.

[41] Figure 5 shows the comparison of Z500 in winter and summer between the ARPEGE model and the ERA40 reanalysis. Mean and standard deviations for the 1958–1999 period are given. The model well reproduces the main spatial features of the mean Z500 but biases in variability exist. In particular, in the model, an underestimation of the winter variability in the northeast is seen. The ARPEGE model has realistic modes of variability. The first four Z500 Empirical Orthogonal Functions (EOF) of the model computed on the domain D1 are spatially well correlated with their ERA40 counterparts (the absolute value of correlation coefficients is always greater than 0.85. Not shown). For the downscaling procedure, the daily Z500 anomalies from the model are projected on the first ten EOFs from ERA40 reanalysis. Then the results of the projection can be directly classified in the weather types established for the learning period with ERA40 reanalysis. As the weather types are imposed to the model, the validation will be focused on two aspects: the probability of occurrence of each weather types, and their persistence properties.

[42] Figure 6 gives the probability of occurrence of each weather type from ERA40 and model simulation for the different seasons. The model results are globally close to the occurrences estimated from ERA40 reanalysis with some discrepancies particularly in winter and summer. This could have an impact on downscaling results. For example, in winter, the weather types WT0 and WT1 that correspond to wet conditions are too frequent and the occurrence of the driest weather type (WT3) is underestimated. The persistence properties of the weather types are also very important in the downscaling context, in particular to reproduce wet and dry spell properties. Figure 7 depicts the probability that a weather type lasts at least N consecutive days, derived from ERA40 reanalysis and model simulation. Model results are similar to those from ERA40 in all seasons.

5.2. Downscaling

[43] In all this section, the two weather typing methods are used to downscale the ARPEGE model present climate simulation.

[44] Figure 8 depicts the mean annual cycle of temperature and rainfall over the domain D2 for the 1985–1999 period. Direct model signals are compared with those obtained by downscaling (weather typing 2). Large overestimation of modelled rainfall occurs for all the months except August, September and October. A cold bias is seen in all seasons, and is particularly marked in summer. Despite the use of the modelled temperature as secondary predictor with weather typing 2 there is no bias in mean downscaled temperature. For rainfall, downscaling results are far better than the direct model values, and are very close to observations. Even with the variable resolution version of ARPEGE allowing for higher resolution over the zone of interest important biases exist for mean regional climate of the Seine basin. Here, the dynamical downscaling is not sufficient to obtain a realistic representation of regional climate. By contrast, statistical downscaling corrects the raw model errors concerning local climate, mainly because of its good representation of LSC.

[45] The diagnostics described in Table 4 are now computed for downscaled precipitation (Table 10). The period considered is August 1985 to December 1999. The results are close to the observations for both methods. Moreover, they are very similar to those previously obtained using ERA40 predictors (Table 4). In particular, persistence properties like p_{dd} and p_{ww} are quite well reproduced.

[46] For hydrometeorological studies, a realistic representation of the spatial variability of the regional climate is necessary. Here, the SDMs are based on resampling strategies, which are known to provide a good way to deal with spatial variability issues. As anticipated, the downscaled ARPEGE simulation with the weather typing 2 method presents a realistic representation of the spatial variability of mean rainfall amounts both in winter and summer as shown on Figure 9. The spatial variability of the 95th percentile of daily downscaled rainfall in winter and summer is also spatially close to the observations (Figure 10).

[47] Figure 11 shows the Probability Density Function (PDF) of rainfall and temperature as observed and downscaled. Two grid points are considered, corresponding to different rainfall characteristics. The point A, characterized by wet conditions, is located in the east of D2. The point B is situated in the west of the domain and is characterized by driest condition (see Figure 1 for the exact location of the points). Some minor discrepancies are seen. For the point A the model slightly underestimates the probability to have dry conditions whereas it is the inverse for the point B. Concerning temperature, the shape of the PDF is well

Table 9. Linear Trend for the 90th Percentile of Maximum Temperature in Winter (K by Decade), 1957–2002[a]

Station	Observations	Analog	Weather Typing 1	Weather Typing 2
S8	**0.24**	**0.43**	**0.33**	**0.47**
S14	**0.30**	**0.38**	**0.35**	**0.46**
S3	**0.26**	**0.40**	**0.25**	**0.39**
S4	**0.21**	**0.42**	**0.27**	**0.43**
S13	**0.23**	**0.42**	**0.26**	**0.47**
S15	**0.35**	**0.36**	**0.30**	**0.33**
S6	**0.31**	**0.45**	**0.25**	**0.48**
S1	**0.27**	**0.34**	**0.22**	**0.33**
S7	**0.32**	**0.45**	**0.28**	**0.44**
S2	**0.30**	**0.22**	**0.23**	**0.30**

[a]The values significant at the 0.05 level are presented in bold.

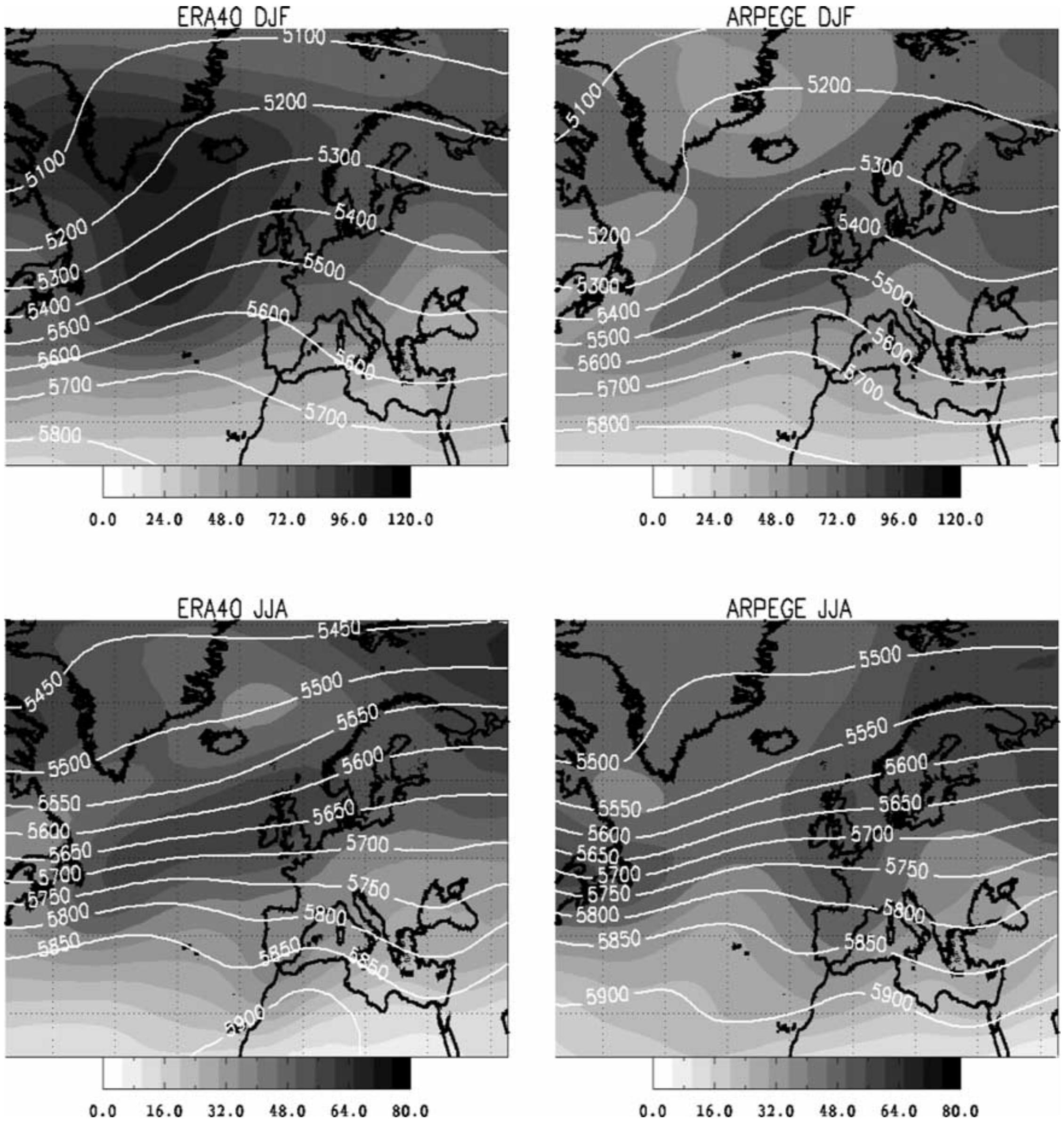

Figure 5. Mean and standard deviation of Z500 in (top) winter (December to February) and (bottom) summer (June to August) over the 1958–1999 period for (left) ERA40 and (right) the ARPEGE Model. Contour lines indicate mean (m). Shading indicates standard deviation (m).

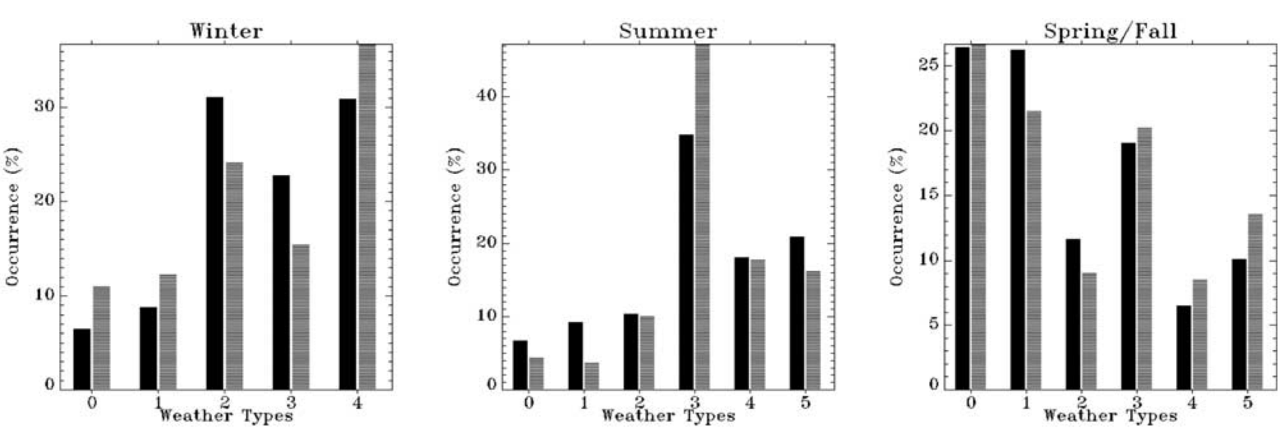

Figure 6. Probability of occurrence of the weather types for each season (1961–1990): ERA40 (solid) versus ARPEGE (shaded).

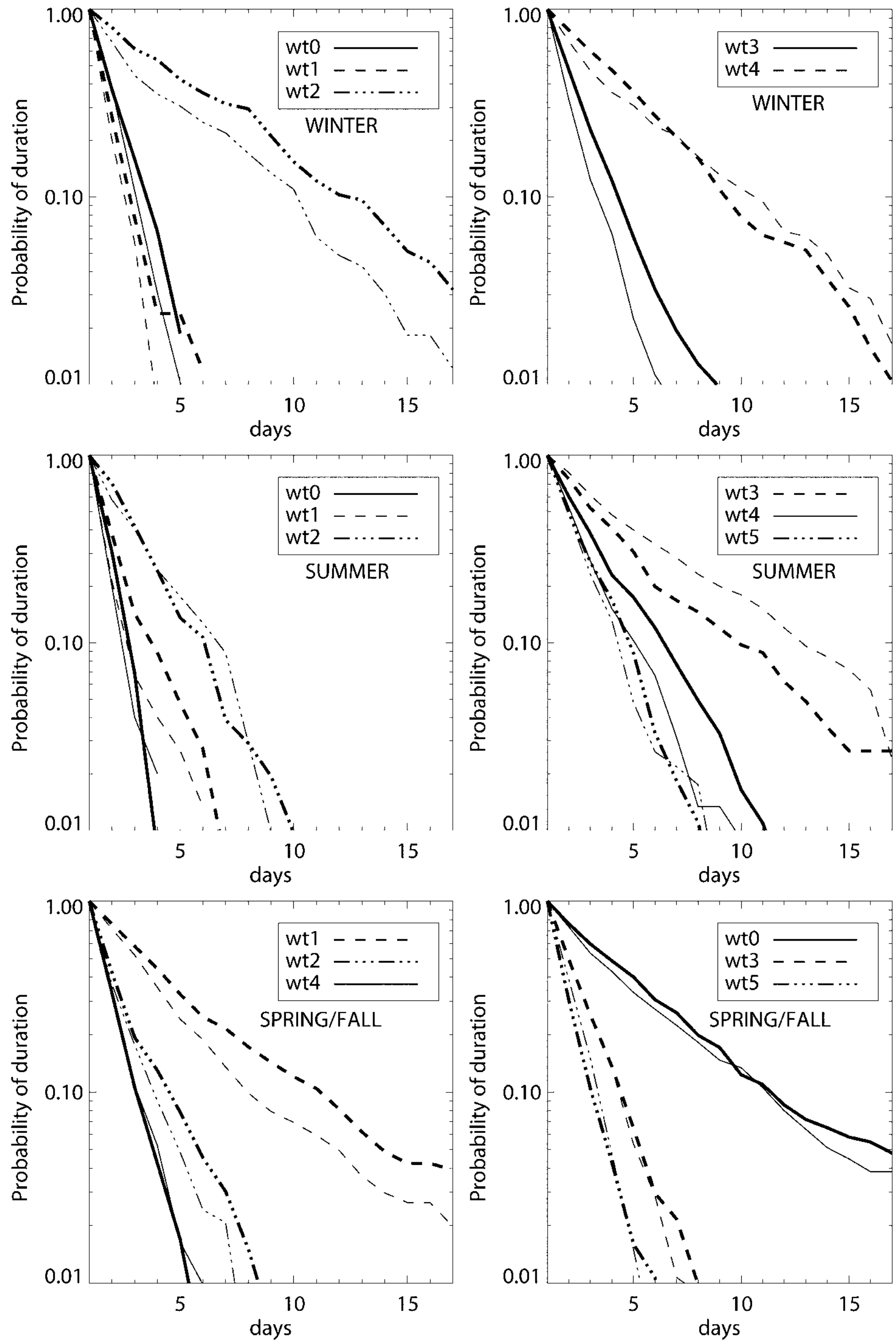

Figure 7. Persistence of the weather types for each season (1961–1990). The curves show the probability to have at least n consecutive days belonging to the same weather type. Thick lines indicate ERA40. Thin lines indicate ARPEGE.

reproduced. In particular, the bimodality is captured. Globally, the model downscaling give a good representation of the distribution of the variables.

[48] The use of the SDMs with ARPEGE results as predictors reproduces daily precipitation and temperature statistics and the spatial variability of regional climate. The impact of the possible problems noted during the validation of the model (section 5.1) on the downscaling results is limited. Only a small overestimation of winter rainfall is seen (Figure 8) coherently with the overestimation of the occurrence of the two wet weather types (Figure 6) in the model. Using ARPEGE surface temperature as secondary predictor (spatially averaged and using only quantile values) is not responsible for a degradation of the results. It is thus

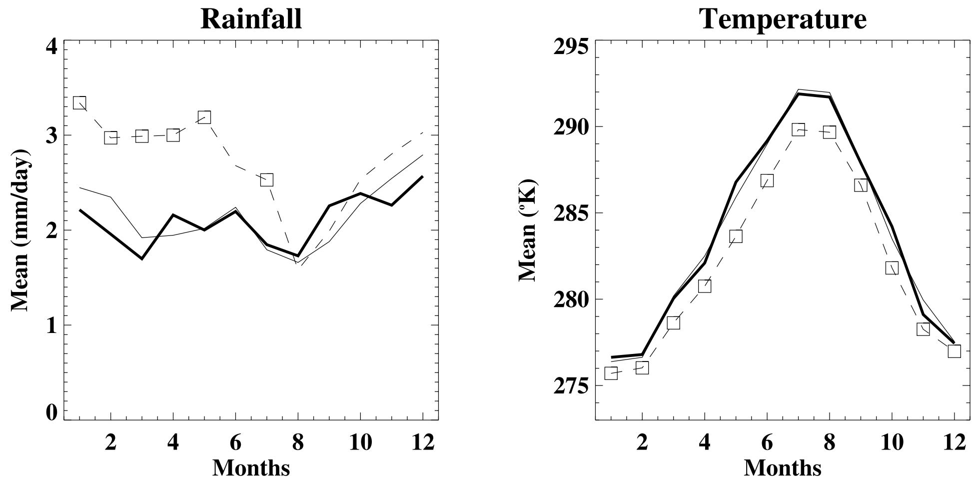

Figure 8. Mean seasonal cycle of rainfall and temperature over D2 (1985–1999) as observed (thick line), simulated (dotted line) and downscaled (thin line). The squares indicate the months where the difference between simulated values and observations are significant at the 0.05 level.

concluded that modelled temperature can be used as secondary predictor.

6. Downscaled ARPEGE A2 and B2 Scenarios

6.1. Mean Seasonal Changes

[49] As noticed by *Charles et al.* [1999], the validation of a statistical downscaling algorithm for present day conditions does not necessarily imply its validity for climate change projections. Indeed, as statistical downscaling uses only partial relationships that link large-scale information and regional climate, processes that are of secondary importance in present climate may play a primary role in perturbed climate. As global climate models account for many more processes and their complex interactions, *Busuioc et al.* [1999] attempt to test the validity of their downscaling scheme in future climate by verifying the consistency between downscaled and direct model results. The consistency adds confidence that the downscaling relationship remains valid in altered climate.

[50] Here, a control and two climate change simulations following the A2 and B2 scenarios are downscaled with weather typing 1 and weather typing 2. Mean seasonal downscaled precipitation and temperature anomalies between future and present climate averaged on the domain D2 are compared with those obtained directly from the model.

[51] In autumn and spring, for the A2 scenario (Figure 12), the anomalies obtained from weather typing 1 and model are quite similar, whereas in summer and winter the changes in downscaled temperature are very small compared to the direct model. For the B2 scenario (Figure 13) an underestimation is seen for all the seasons. The only use of Z500 as predictor thus generally does not reproduce the seasonal temperature changes of the model both for the A2 and B2 simulations, even if the differences are dependent on the season and scenario. Conversely, the use of temperature as secondary predictor in weather typing 2 leads to temperature changes that are very similar to those obtained directly with the model. Concerning precipitation, for the B2 scenario the results obtained with weather typing 2 and the model are quasi-identical. For the A2 scenario the downscaling and model results are also consistent except in autumn. Note that for the period 2070–2099, modelled and downscaled autumn precipitation nevertheless exhibits coherent decreasing trends (not shown).

[52] Another indication reinforcing our confidence that the empirical relationship established for the present climate with the weather typing 2 method is still valid in altered climates is given in Table 11, showing the correlations of spatially averaged downscaled precipitation and tempera-

Table 10. Spatial Mean and Root Mean Square Errors (RMSE) for Precipitation Diagnostics With Different SDMs Using ARPEGE Predictor(s) for the 1985–1999 Period[a]

		Weather Typing 1		Weather Typing 2	
Diagnostic	Observations, Mean	Mean	RMSE	Mean	RMSE
Mean	4.47	4.40	(0.27)	4.57	(0.41)
SD wet	5.45	5.27	(0.53)	5.56	(0.64)
95% wet	15.54	15.07	(1.49)	15.99	(2.25)
p_d	0.52	0.52	(0.018)	0.52	(0.024)
p_{dd}	0.68	0.67	(0.024)	0.68	(0.025)
p_{ww}	0.65	0.65	(0.027)	0.65	(0.036)
L_d	4.91	4.57	(0.49)	4.64	(0.51)
L_w	4.11	4.23	(0.39)	4.31	(0.56)

[a]Mean and RMSE are computed on 9988 values (2497 grid points by four seasons).

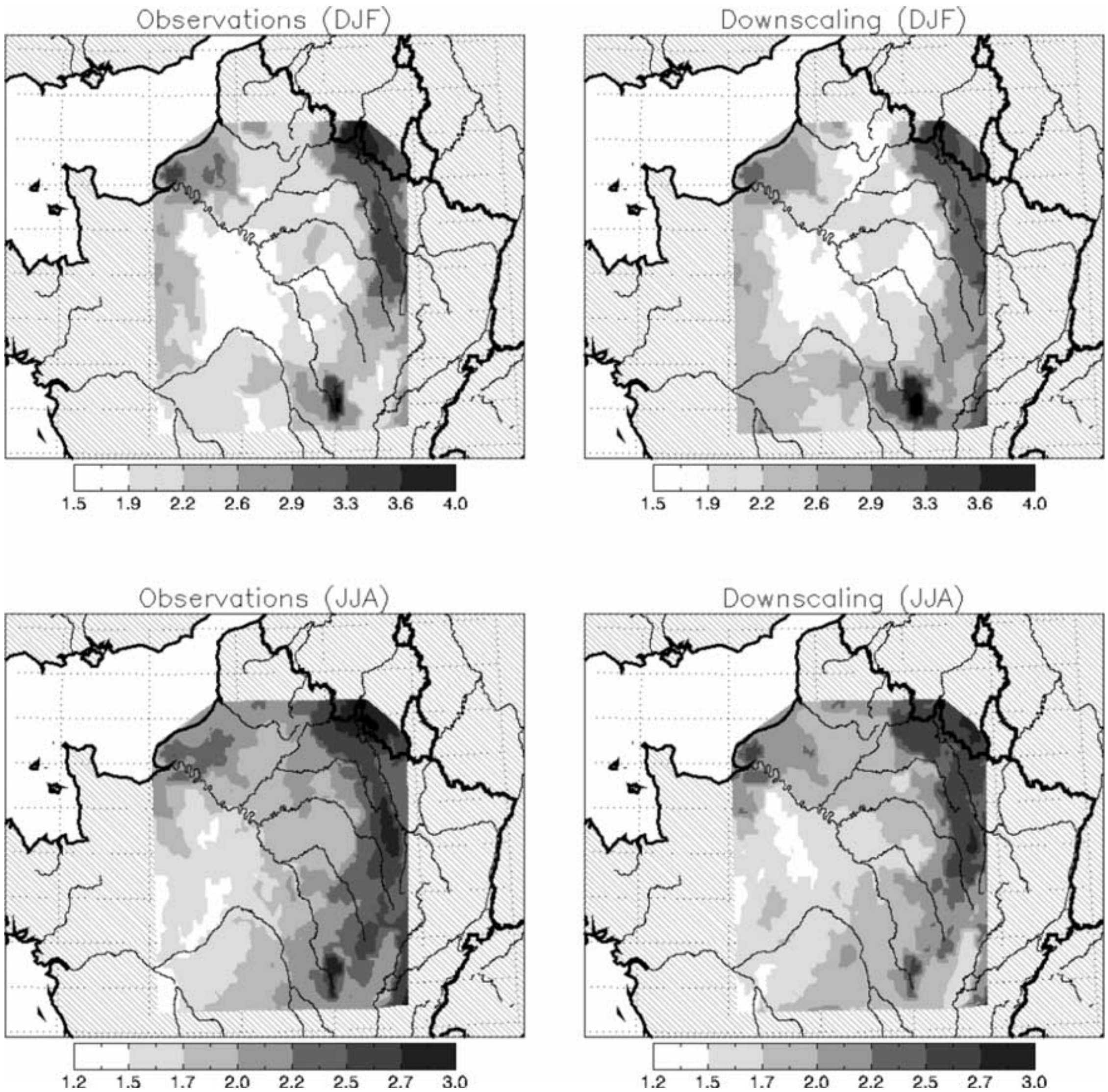

Figure 9. Rainfall climatologies (1985–1999) as (left) observed and (right) downscaled with weather typing 2 (mm/day). (top) Winter (December to February). (bottom) Summer (June to August). The spatial correlation between downscaled and observed fields is 0.91 in winter and 0.90 in summer.

ture series with the direct model results. Even if a GCM fails to totally reproduce the small scale features of regional climate by construction, it should nevertheless partially capture the links between the large-scale climatic state and the regional climate. Consequently, for present climate, a significant correlation is found between the downscaled series and the direct model results averaged on D2 for temperature and precipitation in all seasons. It is then interesting to note that the strength of this statistical link is unchanged in future climate simulations for both scenarios with the weather typing 2 method. If a major modification (captured by the model) of the link between LSC patterns and regional climate has occurred with climate change, a net degradation of the correlation between downscaled and direct model results would have been be noted. It is not the case for weather typing 2. For weather typing 1, an important diminution of the correlation for temperature is noted (not shown).

[53] In summary, our results give confidence in the use of the weather typing 2 SDM for climate change studies. Conversely it is concluded that weather typing 1 is not applicable to altered climates. Downscaling based only upon LSC predictors does not correctly represent temperature changes and even, to a lesser extent precipitation modifications. Differences are noticed between weather typing 1 and weather typing 2 results concerning precipitation. As these two methods only differ by the use of model temperature as secondary predictor in weather typing 2, this might underline the role of temperature changes in precipitation modifications. The relative change in water-holding capacity of the atmosphere is governed by the Clausius-Clapeyron equation and is therefore approximately proportional to temperature change. As the models suggest that the changes in relative humidity are small, the moisture content should vary proportionally to temperature changes [*Trenberth et al.*, 2003]. Even if the interactions between humidity modifications and precipitation are complex, the misrepresentation of temperature changes through

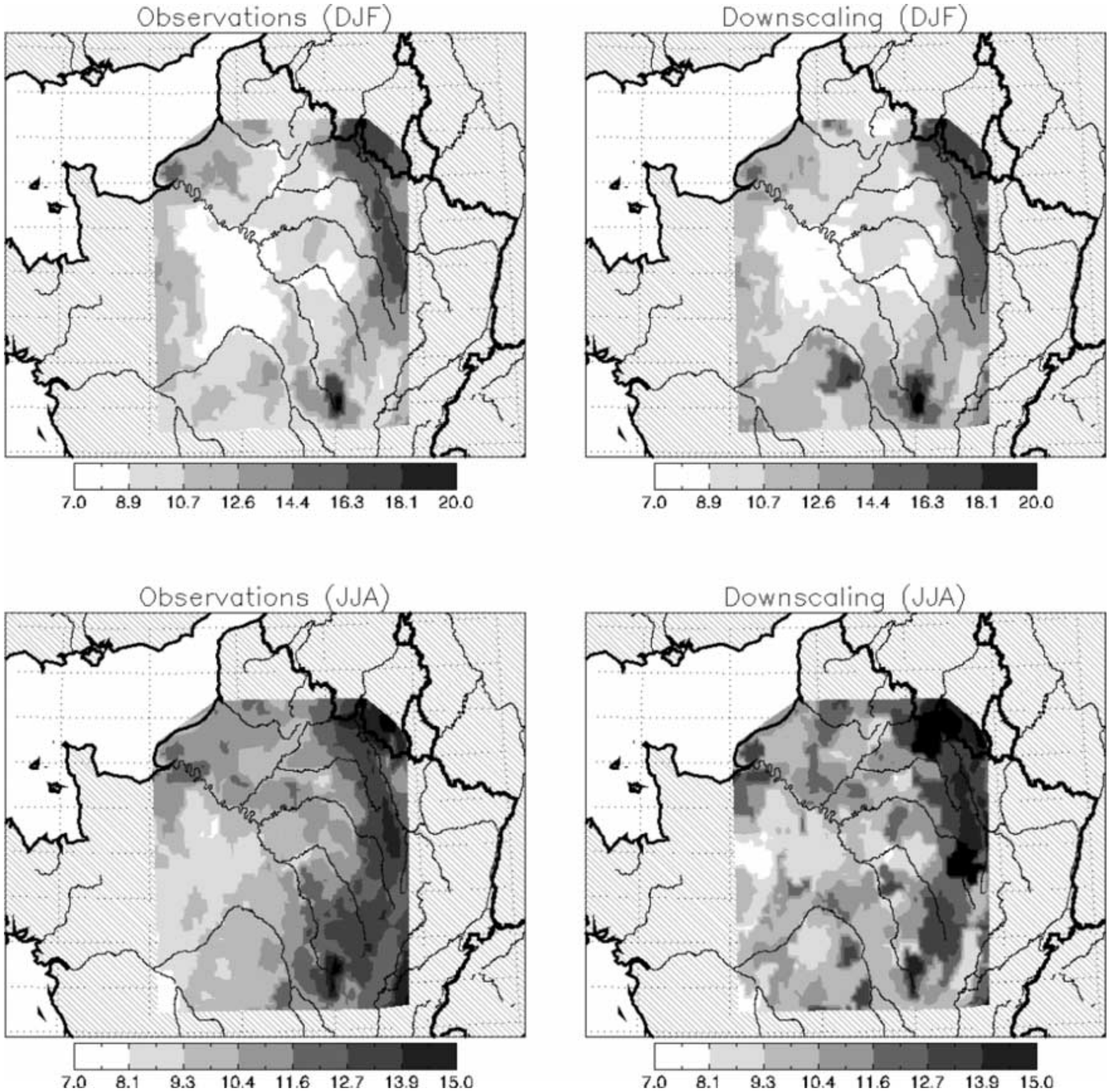

Figure 10. The 95th percentile of daily rainfall (mm/day) for the period 1985–1999 as (left) observed and (right) downscaled with weather typing 2. (top) Winter (December to February). (bottom) Summer (June to August). The spatial correlation between downscaled and observed values is 0.90 in winter and 0.74 in summer.

downscaling process has an impact on precipitation results as seen in weather typing 1 method.

[54] Whereas downscaling results are in very good agreement with the model for the B2 scenario (2040–2069) some discrepancies can be seen for the A2 scenario (2070–2099). It is intuitively consistent with the hypothesis that the more different the future climate from the present climate is, the more questionable the use of a statistical downscaling relationship established for present climate is. Nevertheless, even for the A2 scenario at the end of the century, in our case for weather typing 2, the discrepancies remain very limited.

6.2. Intratypes and Intertypes Modifications

[55] *Wilby et al.* [2004] noticed that weather classification schemes can be insensitive to future climate forcing. Indeed, climate change can influence the probability of occurrence of the weather types but also the links between a weather type and the regional climate. The simplest weather typing downscaling procedure that would consist given a set of weather types, in randomly choosing a day of the observation data set that belongs to the same weather type only takes into account the modification of the probabilities of occurrence. A simple mathematical decomposition is used to separate linearly the effects of occurrence changes and the effects of the modification of the links between weather types and regional climate (R. Vautard, Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit, submitted to *Geophysical Research Letters*, 2006) that may occur in the climate scenario.

[56] Let X_i^p the mean conditional value of a regional variable for the weather type i in present climate and X_i^f the value in future climate.

[57] The total anomaly ΔX of the variable between future and present climate can be written as:

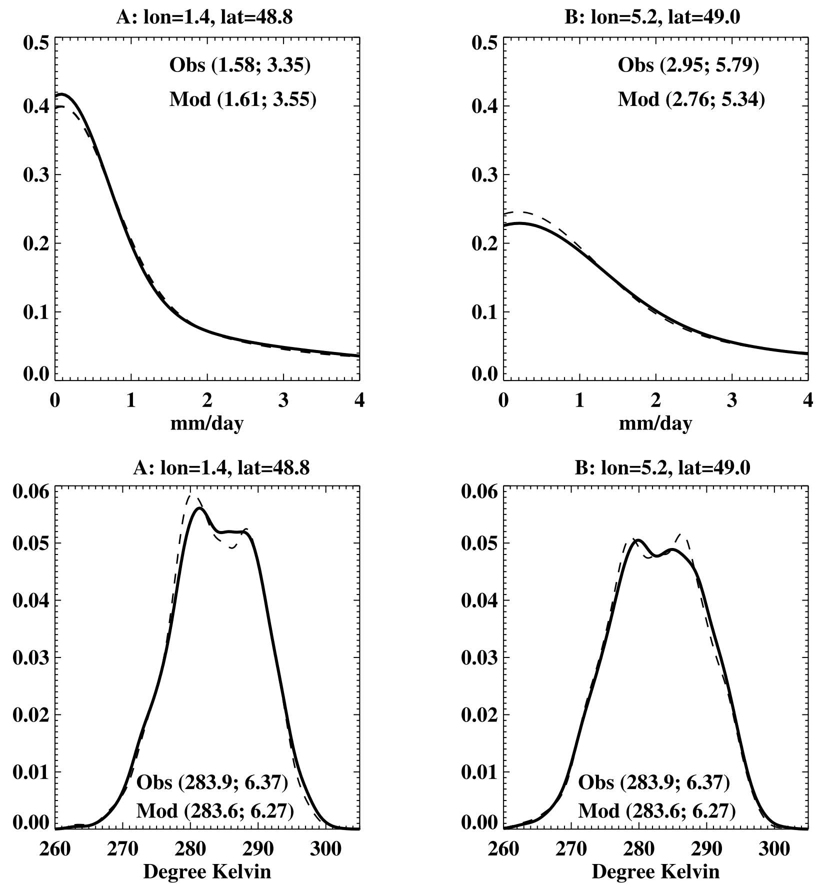

Figure 11. Probability density function of (top) rainfall and (bottom) temperature for two different points of the domain ((left) A and (right) B). See Figure 1 for the location of the points. The PDF is plotted for the observations (Obs., thick line) and for the downscaling of ARPEGE (Mod., dotted line). The mean and the standard deviation of the variables is indicated in brackets.

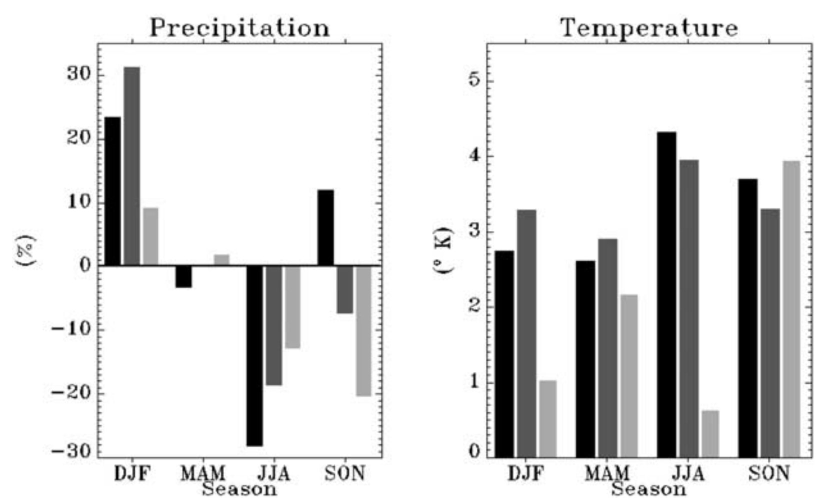

Figure 12. Seasonal anomalies of (left) precipitation (percent) and (right) temperature (K) between 2070–2100 and 1960–2000 with the A2 scenarios in average over D2. Solid indicates model, heavy shading indicates weather typing 2, and light shading indicates weather typing 1.

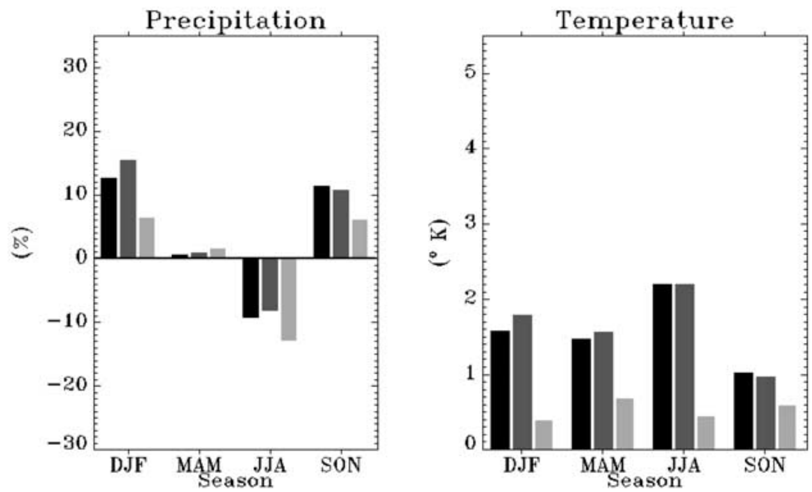

Figure 13. Seasonal anomalies of (left) precipitation (percent) and (right) temperature (K) between 2040–2070 and 1960–2000 with the B2 scenarios in average over D2. Solid indicates model, heavy shading indicates weather typing 2, and light shading indicates weather typing 1.

$$\Delta X = \sum_{i=1}^{k} \left(f_i^f X_i^f - f_i^p X_i^p \right) \quad (5)$$

where f_i^f and f_i^p are the frequency of occurrence of each weather type in future and present climate respectively.

[58] Equation (5) can be rewritten as:

$$\Delta X = \sum_{i=1}^{k} \left(f_i^f \left(X_i^f - X_i^p \right) \right) + \sum_{i=1}^{k} \left(\left(f_i^f - f_i^p \right) X_i^p \right) \quad (6)$$

The first term in equation (6) is the "intratypes" anomaly that is due to the modification of the link between the weather types and the regional climate. The second term is the anomaly that is due to the modification of the frequencies of occurrence (the "intertypes" anomaly). This decomposition is applied to the anomalies between the B2 simulation (2040–2060) and the control run (1960–2000) for direct and downscaled temperature and precipitation, averaged on the domain D2 (Table 12).

[59] Note that in the first term of equation (6) the modifications of the probabilities of occurrence play a role. Indeed this term can be decomposed as:

$$\sum_{i=1}^{k} \left(f_i^p \left(X_i^f - X_i^p \right) \right) + \sum_{i=1}^{k} \left(\Delta f_i \left(X_i^f - X_i^p \right) \right) \quad (7)$$

where $\Delta f_i = f_i^f - f_i^p$ is the modification of the frequencies of occurrence. Here, the term in equation (7) that is due to occurrence changes (second term) is much smaller than the first term both for temperature and precipitation (not shown).

[60] The results obtained with direct model variables and downscaled outputs are coherent both for temperature and precipitation with weather typing 2. This downscaling algorithm captures both intertypes and intratypes changes, which reinforces our confidence in downscaled results in climate change context. Table 12 indicates that the changes within the weather types are at least as much important as the changes due to modification of the frequency of occurrence. For precipitation the two anomalies are even most of the time of opposite signs. Note that the intratypes changes should not be interpreted as pure physical changes in comparison with the dynamical changes represented by intertypes changes. For precipitation, intratypes changes with weather typing 1 that uses only a dynamical predictor are very similar with those obtained with weather typing 2 in summer and spring, and smaller in winter and autumn. The dynamical predictor is thus able to capture an important part of intratype changes for precipitation. Conversely, for temperature it is not the case: weather typing 1 is totally unable to capture intratypes changes. These results highlight the drawback to use only the modifications of the frequencies of occurrence in weather typing approaches, as it represents only a part of climate change signal.

7. Summary and Conclusion

[61] Among different statistical downscaling methods, the weather typing approaches have strengths and weaknesses as summarized by *Wilby et al.* [2004]. In particular, these

Table 11. Correlation of Spatially Averaged Precipitation and Temperature Series Between Downscaled (Using Weather Typing 2) and Direct Model Results for Control, A2 and B2 Simulations

	Present (1960–2000)	B2 (2040–2060)	A2 (2070–2100)
		Precipitation	
Winter	0.52	0.52	0.53
Spring	0.52	0.54	0.54
Summer	0.42	0.37	0.36
Autumn	0.50	0.48	0.47
		Temperature	
Winter	0.92	0.95	0.95
Spring	0.91	0.95	0.95
Summer	0.92	0.96	0.94
Autumn	0.92	0.94	0.93

Table 12. Decomposition of the Temperature and Precipitation Seasonal Anomalies Averaged on the Domain D2 Between the B2 Simulation (2040–2060) and the Control Run (1960–2000) for Direct Model and Downscaled Results With Weather Typing 2[a]

	Model		Downscaling	
	Intertypes	Intratypes	Intertypes	Intratypes
		Precipitation, mm/day		
DJF	−0.03	0.38	0.06	0.26 (0.10)
MAM	−0.31	0.37	−0.34	0.38 (0.39)
JJA	−0.41	0.21	−0.72	0.45 (0.44)
SON	−0.17	0.38	−0.24	0.47 (0.36)
		Temperature, K		
DJF	0.16	1.37	0.24	1.54 (0.11)
MAM	0.73	0.67	0.86	0.72 (−0.32)
JJA	0.84	1.32	0.89	1.32 (−0.49)
SON	0.45	0.41	0.44	0.52 (0.12)

[a]For intratypes changes results obtained with weather typing 1 are also shown in brackets (the intertypes changes are quasi-identical with the two methods). See text for calculation explanation.

methods yield physically interpretable linkages to surface climate and can be applied to a wide range of problems. Nevertheless they require additional task of weather classification and can be insensitive to future climate forcing. In particular they may not capture intratype variations in surface climate. In this study, a simple yet efficient weather typing-based methodology intended to overcome these weaknesses is developed.

[62] A bottom-up approach starting from regional climate properties to establish discriminating weather types for local variables is applied. To take into account the intratype variations in surface climate, the distances to the weather types are then used in the downscaling process. The performance of the weather typing methods is tested against a standard analog approach, considered as a natural "benchmark" method, as it has proven to be both simple and skillful (*Zorita and von Storch* [1999]). The major advantages of weather typing compared to analogs is the possible study of the physical mechanisms on which the statistical downscaling is based.

[63] First, using ERA40 reanalysis fields as predictors (z500 for weather typing 1, z500 and surface temperature for weather typing 2), it was shown that the weather typing method developed in this paper is superior to the standard analog method and successfully reproduces daily statistic properties of precipitation. Moreover, low-frequency variations of downscaled temperature and precipitation are well simulated and close to the observations. Next, the weather typing methods were tested using as predictors the results of a global GCM simulation with high resolution over the study area thanks to a variable resolution. As the atmospheric circulation is globally well reproduced in the GCM, when applying the SDMs, a good representation of regional climate properties is obtained. The model temperature can be used as secondary predictor without loss of performance, even if temperature is globally underestimated by the model because of the way it is used within the downscaling process. It is worth noting that even with high resolution over the study area the model has important biases for mean temperature and precipitation. Using statistical downscaling provides an important improvement of the results.

[64] The consistency of downscaling and raw model results are tested for two future climate simulations. The weather typing method that uses temperature as a secondary predictor yields results that are consistent with model changes both for temperature and precipitation with the B2 scenario. With the A2 scenario the results are also consistent except in autumn for precipitation. These results give confidence in the use of this SDM for climate change applications. Conversely, if the temperature is not used as secondary predictor, the SDM fails to reproduce the temperature changes and in a lesser extent the precipitation changes. The latter is considered not applicable in altered climates. By contrast, this SDM correctly reproduces present day regional climate. As in *Charles et al.* [1999], these results highlight the importance of evaluating the plausibility of climate change projections on the basis of statistical downscaling.

[65] The SDMs used in this study are based on a resampling strategy. The major weakness of this approach is that resampling is incapable of giving values outside the already observed range of values. This drawback is common to many SDM, as resampling strategies are often chosen when dealing with spatial and multivariate problems. It is the case for the weather typing and analog methods, and also when single-site results are extended to multiple sites from regression-based methods [*Wilby et al.*, 2003] or weather generators [*Palutikof et al.*, 2002]. For hydrological applications, as the greatest present day cumulated amount of precipitation over N days can be exceeded, the greatest streamflow of present climate can also be exceeded in future climate. Depending upon the characteristics of the watershed, the limitation of resampling strategies concerning the extreme values is more or less important. It must also be noted that in the A2 scenario for the whole 2070–2100 period, the greatest value of mean precipitation on the domain D2 in the control run (1960–2000) is only exceeded for 16 days. This thus reinforces our confidence in the use of a resampling strategy for the downscaling of the Seine basin precipitation. Nevertheless, generally, when analyzing the results of climate change impact assessments obtained by SDM on the basis of resampling this point should be carefully considered.

[66] Finally it is concluded that the downscaling methodology developed in this paper can be applied to the study of the impacts of climate change on the Seine basin hydrology, which will be the object of a future study. This method may also find other applications, as for example seasonal hydrological forecast.

[67] **Acknowledgments.** This work was supported by the Programme National de Recherche en Hydrologie (PNRH) under the RIVAGES project and by the European Community via the sixth framework program ENSEMBLES project under contract GOCE-CT-2003-505539. ECMWF ERA-40 data used in this study have been obtained from the ECMWF data server. The authors are grateful to the Division de la Climatologie (DCLIM) at Météo-France for providing the SQR data set. They would like to thank M. Déqué for providing the PRUDENCE simulations and C. Cassou for stimulating discussion about this work.

References

Allen, M. R., and W. J. Ingram (2002), Constraints on future changes in climate and the hydrologic cycle, *Nature*, *419*, 224–232.

Bengtsson, L., S. Hagemann, and K. I. Hodges (2004), Can climate trends be calculated from reanalysis data?, *J. Geophys. Res.*, *109*, D11111, doi:10.1029/2004JD004536.

Busuioc, A., H. von Storch, and R. Schnur (1999), Verification of GCM-generated regional seasonal precipitation for current climate and of statistical downscaling estimates under changing climate conditions, *J. Clim.*, *12*(1), 258–272.

Cassou, C., L. Terray, and A. S. Phillips (2005), Tropical Atlantic influence on European heatwaves, *J. Clim.*, *18*(15), 2805–2811.

Charles, S. P., B. C. Bates, P. H. Whetton, and J. P. Hugues (1999), Validation of downscaling models for changes climate conditions: Case study of southwestern Australia, *Clim. Res.*, *12*, 1–14.

Corte-Real, J., B. Qian, and H. Xu (1999), Circulation patterns, daily precipitation in Portugal and implications for climate change simulated by the second Hadley Centre GCM, *Clim. Dyn.*, *15*, 921–935.

Durand, Y., E. Brun, L. Merindol, G. Guyomarc'h, B. Lesaffre, and E. Martin (1993), A meteorological estimation of relevant parameters for snow schemes used with atmospheric models, *Ann. Glaciol.*, *18*, 65–71.

Ebisuzaki, W. (1997), A method to estimate the statistical significance of a correlation when the data are serially correlated, *J. Clim.*, *10*(9), 2147–2153.

Etchevers, P., C. Golaz, F. Habets, and J. Noilhan (2002), Impact of a climate change on the Rhone river catchment hydrology, *J. Geophys. Res.*, *107*(D16), 4293, doi:10.1029/2001JD000490.

Gibelin, A.-L., and M. Déqué (2003), Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model, *Clim. Dyn.*, *20*, 327–339.

Giorgi, F., M. R. Marinucci, and G. Visconti (1990), Use of a limited-area model nested in a general circulation model for regional climate simulation over Europe, *J. Geophys. Res.*, *95*(D11), 18,413–18,431.

Gutiérrez, J. M., A. S. Cofiño, R. Cano, and M. A. Rodríguez (2004), Clustering methods for statistical downscaling in short-range weather forecast, *Mon. Weather Rev.*, *132*(9), 2169–2183.

Habets, F., P. Etchevers, C. Golaz, E. Leblois, E. Ledoux, E. Martin, J. Noilhan, and C. Ottlé (1999), Simulation of the water budget and the river flows of the Rhone basin, *J. Geophys. Res.*, *104*(D24), 31,145–31,172.

Jones, T. C., et al. (2003), Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emission scenarios, *Clim. Dyn.*, *20*, 583–612.

Klein Tank, A. M. G., and G. P. Können (2003), Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99, *J. Clim.*, *16*(22), 3665–3680.

Ledoux, E., G. Girard, and J. P. Villeneuve (1984), Proposition d'un modèle couplé pour la simulation conjointe des écoulements de surface et des écoulements souterrains sur un bassin hydrologique, *Houille Blanche*, *1/2-1984*, 101–110.

Martin, E., B. Timbal, and E. Brun (1997), Downscaling of general circulation model outputs: Simulation of the snow climatology of the French Alps and sensitivity to climate change, *Clim. Dyn.*, *13*, 45–56.

Mearns, L. O., I. Bogardi, F. Giorgi, I. Matyasovszky, and M. Palecki (1999), Comparison of climate change scenarios generated from regional climate model experiments and statistical downscaling, *J. Geophys. Res.*, *104*(D6), 6603–6621.

Mehrotra, R., and A. Sharma (2005), A nonparametric nonhomogeneous hidden Markov model for downscaling of multisite daily rainfall occurences, *J. Geophys. Res.*, *110*, D16108, doi:10.1029/2004JD005677.

Michelangeli, P.-A., R. Vautard, and B. Legras (1995), Weather regimes: Recurrence and quasi stationarity, *J. Atmos. Sci.*, *52*(8), 1237–1256.

Moisselin, J. M., M. Schneider, C. Canellas, and O. Mestre (2002), Changements climatiques en France au 20ème siècle. Etude des longues séries de données homogénéisés françaises de précipitations et températures, *Meteorologie*, *38*, 45–46.

Noilhan, J., and S. Planton (1989), A simple parametrization of land surface processes for meteorological models, *Mon. Weather Rev.*, *117*, 536–549.

Palutikof, J. P., C. M. Goodess, S. J. Watkins, and T. Holt (2002), Generating rainfall and temperature scenarios at multiple sites: Examples from the Mediterranean, *J. Clim.*, *15*(24), 3529–3548.

Plaut, G. (2004), Building downscaling algorithms for precipitation: The "Potential Precipitation Circulation Index"or "ppci", an adaptive predictor to downscale extreme precipitation on the french Alpes Maritimes, STARDEX project partner 5 (CNRS-INLN) report for deliverable D10, Clim. Res. Unit, Univ. of East Anglia, Norwich, U.K. (Available http://www.cru.uea.ac.uk/cru/projects/stardex/deliverables/D10/)

Plaut, G., E. Schuepbach, and M. Doctor (2001), Heavy precipitation events over a few Alpine sub-regions and the links with large-scale circulation, 1971–1995, *Clim. Res.*, *17*, 285–302.

Roebber, P. J., and L. F. Bosart (1998), The sensitivity of precipitation to circulation details. Part I: An analysis of regional analogs, *Mon. Weather Rev.*, *126*, 437–455.

Royer, J. F., et al. (2002), Simulation of climate changes during the 21st century including stratospheric ozone, *C. R. Geosci.*, *334*, 147–154.

Schubert, S. (1998), Downscaling local extreme temperature changes in south-eastern Australia from the CSIRO MARK2 GCM, *Int. J. Climatol.*, *18*, 1419–1438.

Straus, D. M., and F. Molteni (2004), Circulation regimes and SST forcings: Results from large GCM ensembles, *J. Clim.*, *17*(8), 1641–1656.

Terray, L., M.-E. Demory, M. Déqué, G. de Coetlogon, and E. Maisonnave (2004), Simulation of late-twenty-first-century changes in wintertime atmospheric circulation over Europe due to anthropogenic causes, *J. Clim.*, *17*(24), 4630–4635.

Timbal, B. (2004), Southwest Australia past and future rainfall trends, *Clim. Res.*, *26*(3), 233–249.

Timbal, B., A. Dufour, and B. McAvaney (2003), An estimate of future climate change for western France using a statistical downscaling technique, *Clim. Dyn.*, *20*, 807–823.

Trenberth, K. E., A. Dai, R. M. Rasmussen, and D. B. Parsons (2003), The changing character of precipitation, *Bull. Am. Meteorol. Soc*, *84*, 1205–1217.

von Storch, H. (1999), On the use of "inflation" in statistical downscaling, *J. Clim.*, *12*(12), 3505–3506.

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*(11), 2995–3008.

Wilby, R. L., O. J. Tomlinson, and C. W. Dawson (2003), Multi-site simulation of precipitation by conditional resampling, *Clim. Res.*, *23*, 183–194.

Wilby, R. L., S. P. Charles, E. Zorita, B. Timbal, P. Whetton, and L. O. Mearns (2004), Guidelines for use of climate scenarios developed from statistical downscaling methods, Data Distrib. Cent. of the Intergovt. Panel on Clim. Change, Geneva, Switzerland.

Wilson, L. L., D. P. Lettenmaier, and E. Skyllingstad (1992), A hierarchical stochastic model of large-scale atmospheric circulation patterns and multiple station daily precipitation, *J. Geophys. Res.*, *97*(D3), 2791–2809.

Zhang, X., X. Wang, and J. Corte-Real (1997), On the relationships between daily circulation patterns and precipitation in Portugal, *J. Geophys. Res.*, *102*(D12), 13,495–13,507.

Zorita, E., and H. von Storch (1999), The analog method as a simple statistical downscaling technique: comparison with more complicated methods, *J. Clim.*, *12*(8), 2474–2489.

Zorita, E., J. P. Hughes, D. P. Lettemaier, and H. von Storch (1995), Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation, *J. Clim.*, *8*(5), 1023–1042.

J. Boé and L. Terray, Climate Modelling and Global Change Team, CERFACS/CNRS, SUC URA1875, 42 Avenue Gaspard Coriolis, F-31057, Toulouse Cedex 1, France. (boe@cerfacs.fr)

F. Habets and E. Martin, Centre National de Recherches Météorologiques, Météo-France, F-31057 Toulouse, France.