

B. Timbal · A. Dufour · B. McAvaney

An estimate of future climate change for western France using a statistical downscaling technique

Received: 17 April 2002 / Accepted: 8 November 2002 / Published online: 8 February 2003
© Springer-Verlag 2003

Abstract A statistical downscaling procedure based on an analogue technique is used to determine projections for future climate change in western France. Three ocean and atmosphere coupled models are used as the starting point of the regionalization technique. Models' climatology and day to day variability are found to reproduce the broad main characteristics seen in the reanalyses. The response of the coupled models to a similar CO₂ increase scenario exhibit marked differences for mean sea-level pressure; precipitable water and temperature show arguably less spread. Using the reanalysis fields as predictors, the statistical model parameters are set for daily extreme temperatures and rain occurrences for seventeen stations in western France. The technique shows some amount of skill for all three predictands and across all seasons but failed to give reliable estimates of rainfall amounts. The quality of both local observations and large-scale predictors has an impact on the statistical model skill. The technique is partially able to reproduce the observed climatic trends and inter annual variability, showing the sensitivity of the analogue approach to changed climatic conditions albeit an incomplete explained variance by the statistical technique. The model is applied to the coupled model control simulations and the gain compared with direct model grid-average outputs is shown to be substantial at station level. The method is then applied to altered climate conditions; the impact of large-scale model uncertain responses and model sensitivities are quantified using the three coupled models. The warming in the downscaled projections are reduced compared with their global model counterparts.

B. Timbal (✉) · B. McAvaney
Bureau of Meteorology Research Centre,
GPO Box 1289k, Melbourne,
VIC 3001, Australia
E-mail: b.timbal@bom.gov.au

Current affiliation: A. Dufour
Centre National de la Météorologie Nationale,
Météo-France, Toulouse, France

1 Introduction

Since anthropogenic climate change has become an important issue, the need to provide regional climate change information has increased, both for impact assessment studies and policy making (Mearns et al. 2001). However, the available tools have directed research toward understanding the climate system as a whole. A regional climate is determined by interactions at large, regional and local scales. Coupled GCMs are run at too coarse resolution to permit accurate description of these regional and local interactions. So far, they have been unable to provide consistent estimates of climate change on a local scale (Kattenberg et al. 1996; Giorgi et al. 2001). Several regionalization techniques have been developed to bridge the gap between the large-scale information provided by coupled models and fine spatial scales required for regional and environmental impact studies. This “downscaling” process is either dynamical (Giorgi et al. 1990) or statistical (Hewitson and Crane 1996). Statistical downscaling is based on the view that regional climate may be seen to be conditioned by two factors: large-scale climatic state and regional and local features. Local climate information is derived by first developing a statistical model which relates large-scale variables or “predictors” for which GCMs are trustable to regional or local surface “predictands” for which models are less skillful (McAvaney et al. 2001). The main advantage of these techniques is that they are computationally inexpensive, and can be applied to outputs from different GCM experiments. Statistical techniques used to perform such studies are numerous (see Giorgi et al. 2001 for a complete review). Several authors used such approaches for Europe (Murphy 1999) or smaller regions in Europe e.g. the British Isles (Conoway et al. 1996), the Iberian Peninsula (von Storch et al. 1993) or the Alps (Gyalistras et al. 1994). Previous downscaling applications for France examined mountainous areas in the southeast: the Alps (Martin et al.

1997) or the mountains surrounding the Mediterranean Sea (Guilbaud and Obled 1998). A novelty of this study is to assess the use of such techniques for the western half of France which is strongly affected by oceanic influence and large-scale circulation.

To understand the benefit of using a downscaling approach for impact studies, requires us to be aware of the different levels of uncertainties involved in generating regional climate information. The first level is associated with emission scenarios. All model transient simulations used in this study are based on a 1% per year increase in CO_2 . This is roughly equivalent to the IPCC scenario a (Leggett et al. 1992) and referred to in the third IPCC report (Cubasch et al. 2001) as CMIP2 since it is the recommended scenarios for the Coupled Model Intercomparison Project (Meehl et al. 2000). This scenario does not reflect the uncertainty about the real future rate of increase in CO_2 and other well-mixed greenhouse gases (in turn due to uncertainties in future emissions), nor does it include the effect of sulphate aerosols. No attempt is made in this article to discuss this level of uncertainties. A second level of uncertainty is due to the simulation of the transient climate response by coupled GCMs for a given forcing scenario. Uncertainties arise due to imperfect knowledge and representation of physical processes, over simplifications and assumptions in physical parametrizations. These uncertainties are seen in the atmospheric part of the climate model as well as from the coupled system due to ocean mechanisms and coupled exchanges between the two mediums. The inter-model differences in simulating a response to a given forcing has been documented throughout the history of atmospheric model development (Cess et al. 1990, 1996; Colman et al. 2001; Colman 2002). They remain critical in understanding the large spread amongst climate change projections as seen in the third IPCC report (Cubasch et al. 2001). To investigate the uncertainties attached to climate change simulation, several projections provided by state-of-the-art GCMs of different origins were used. The models used cover a large spectrum of the observed model sensitivity spread (Colman 2002) and future climate projections (Meehl et al. 2000; Cubasch et al. 2001). A third level of uncertainty lies in the regionalization tool itself. The comparison of techniques (Wilby et al. 1998; Zorita and von Storch 1999) have illustrated how regionalization tools could yield different results. Finally, climate observations (in situ data as well as model generated analyses) are also subject to uncertainty. The relevance of these uncertainties in the development and the application of a regionalization tool is illustrated in this study.

The analogue method used here was first described by Lorenz (1969). Recently, analogue techniques have been successfully applied to climate simulations at mid-latitude (Zorita et al. 1995; Martin et al. 1997; Timbal and McAvaney 2001). They compare well with more sophisticated methods (Zorita and von Storch 1999). Three major assumptions govern any statistical downscaling approach in general and the analogue technique

in particular. First, the predictors on which the method rely must be properly represented by the global models (Hulme et al. 1993). Otherwise, the lack of reasonable estimates of the predictors dooms the statistical approach to failure. Second, analyses of the downscaled GCM outputs must demonstrate the improvement compared with direct model outputs (Palutikof et al. 1997). Thirdly, the statistical relationship between large-scale predictors and local predictands must remain valid under altered climatic conditions. In other words all expected future realizations of the predictors must be contained in the observational record. It can be argued that daily synoptic situations are highly variable and, provided a sufficiently long training period is used, most of the situations expected in a different climate will be included. The validity of these assumptions is discussed in the results sections.

After a description of the datasets (surface observations, analyses used to train the statistical model and GCM outputs), the statistical model developed is presented and validated. Finally, downscaled projections are deduced by applying the technique to transient runs from the coupled models. The downscaled climate change projections are compared and put in perspective with those provided directly by the GCMs.

2 The datasets

Daily observations used include surface air temperature (T_{max} and T_{min}) and rainfall at 17 high quality synoptic stations in the western half of France (Fig. 1). These data are highly reliable over the period of interest (1958–1998). No missing data are reported for temperatures, and less than 0.05% for rainfall in all cases (the largest amount of missing data being Angoulême with 0.048%). Part of the quality check on these stations was a homogenization process (Mestre 2000) conducted on monthly values. It detects and corrects in long-term climate time-series multiple aberrant points and ensures relative homogeneity. Correction coefficients applied are

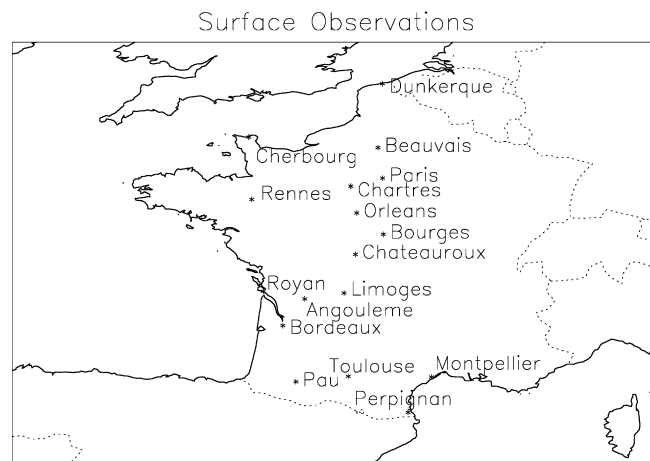


Fig. 1 The surface observations used in western France

for monthly means, whereas in this study the datasets for the statistical model are daily. This may seem a limitation of the correction but it is expected to remove most, if not all, of the shifts due to altered location of the stations or instrument changes. The amplitudes of such coefficients are small, they are usually larger in summer rather than in winter. The impact of applying such homogenization on daily values is evaluated later on.

Large-scale “observed” predictors are derived from the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) collaborative Re-analyses (hereafter, NNR). This is a 41-year record of global analyses of atmospheric fields (Kalnay et al. 1996) produced using a frozen global data assimilation system. Most of variables used in this study are strongly influenced by observed data, and hence, are the most reliable. Although, precipitable water, a variable used, is strongly influenced by the model.

The downscaling technique is applied to three atmosphere ocean coupled GCMs. The global performances and climate change projection of these models, amongst others, were assessed in the Coupled Model Intercomparison Project (CMIP: Meehl et al. 2000; Lambert and Boer 2001) and in the third IPCC scientific assessment (Houghton et al. 2001). The models are labelled according to their respective modelling centres: Bureau of Meteorology Research Centre (BMRC, McAvaney and Colman 1993; Power et al. 1993), Commonwealth Scientific and Industrial Research Organization (CSIRO, Gordon and O’Farrel 1997), and Laboratoire de Météorologie Dynamique (LMD, Braconnot et al. 1997). These models used similar low horizontal resolution of the order of 500 km squared grid but with major differences in both the numerical scheme employed to solve the dynamical equations and the physical parametrization packages used. While the CSIRO and BMRC models are flux corrected to limit control model drift away from observed climatology, the LMD model is not. Daily fields were extracted over 20 years in a transient experiment at the time the CO₂ concentration reached double present values. The LMD scenario was somewhat different since the 20 years were extracted after the coupled model remained at constant 2 * CO₂ for roughly 50 years. A selection of 20 years were also extracted from each coupled model control run from corresponding model years. By doing so, similar model drifts are discarded in both experiments. This is a common practice to retain the climate change signal, although, this method does not take into account possible non linearity in the development of model drifts (Raper and Cubasch 1996).

Since the statistical model applicability is highly dependent on the quality of the predictors, the coupled model control run climatologies were first compared with the NNR. All model fields were interpolated onto the 2.5° × 2.5° NNR grid. Comparisons were carried out for seasonal means over the 20 model years, for a 50°E to 50°W and 20°N to 70°N domain covering most of the

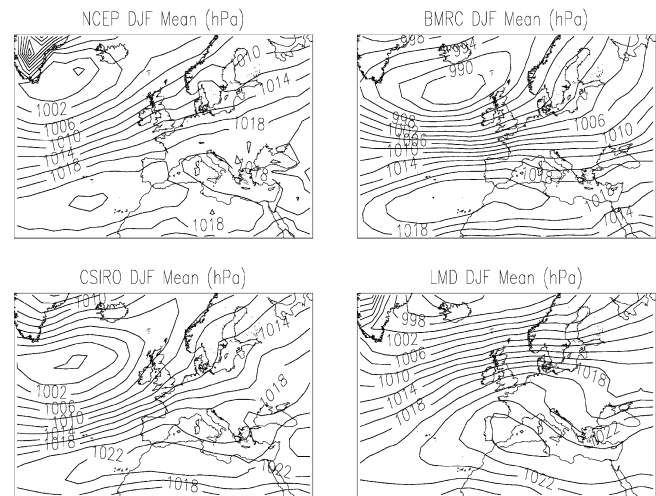


Fig. 2 Winter (December to February) climatologies of *MSLP*: NNR (top left) and the three models (BMRC, CSIRO and LMD)

North Atlantic, Europe and North Africa in order to track most of the circulation patterns that affect western France (e.g. the Atlantic storm-track, the Azores High and Mediterranean lows).

Amongst the atmospheric predictors tested, mean sea level pressure (*MSLP*), temperature at 850 hPa (T_{850}) and precipitable water (*PWTR*) were found to have the most skill. Only these modelled variables are discussed here. Winter seasonal means of *MSLP* (Fig. 2) are compared with NNR. The models reproduce the main features but with some large discrepancies in the positioning and intensity of the centres of activity (Icelandic Low and Azores High). The flow affecting Western Europe is too zonal in all models. The BMRC model has, for example, an Icelandic Low too deep and too close to Europe, while its position is too far south in the CSIRO model and by contrast displaced too far north in the LMD model. The latter also exhibits an Azores High too far north and east. In summer similar displacements of large-scale centres of activities are observed in model climatologies (not shown). For example, the trough over the Mediterranean Sea is over estimated in both the BMRC and the LMD models. The extension of a ridge of high pressure over Western Europe from the Azores High is too strong in the CSIRO model and displaced over the northern Atlantic in the LMD model.

Differences between model and NNR are summarized in Table 1. For each variable the mean difference between each model and the NNR is calculated over the entire domain used; the maximal absolute difference over the same domain is also shown. Results show a coherent pattern over the four seasonal means. The BMRC model has the largest mean and maximal errors for *MSLP* due to a general bias of the model toward low surface pressure over the northern Atlantic. Of the two other models, the CSIRO model is closer to the NNR throughout the year.

Table 1 Seasonal mean differences and largest differences (Maximum) between coupled models and the NNR over the entire domain of interest for *MSLP* (hPa), T_{850} (°C) and *PWTR* (mm)

		Winter		Spring		Summer		Autumn	
		Mean	Maximum	Mean	Maximum	Mean	Maximum	Mean	Maximum
BMRC	<i>MSLP</i>	5.6	19.6	4.8	11.9	6.0	16.0	3.7	16.5
	T_{850}	1.2	7.6	1.7	9.4	2.4	14.5	1.5	8.4
	<i>PWTR</i>	2.0	10.3	1.7	10.7	2.6	13.5	2.2	12.1
CSIRO	<i>MSLP</i>	2.7	10.2	1.8	8.7	2.5	10.8	1.8	6.6
	T_{850}	0.8	4.4	1.0	5.2	1.6	5.9	0.9	3.8
	<i>PWTR</i>	1.4	9.9	1.4	7.1	2.3	11.4	1.6	13.3
LMD	<i>MSLP</i>	4.1	19.1	2.8	14.3	4.2	11.4	3.5	15.5
	T_{850}	4.2	22.3	3.9	13.7	3.6	8.2	3.6	12.9
	<i>PWTR</i>	3.0	18.0	2.5	12.5	3.5	12.1	3.0	11.5

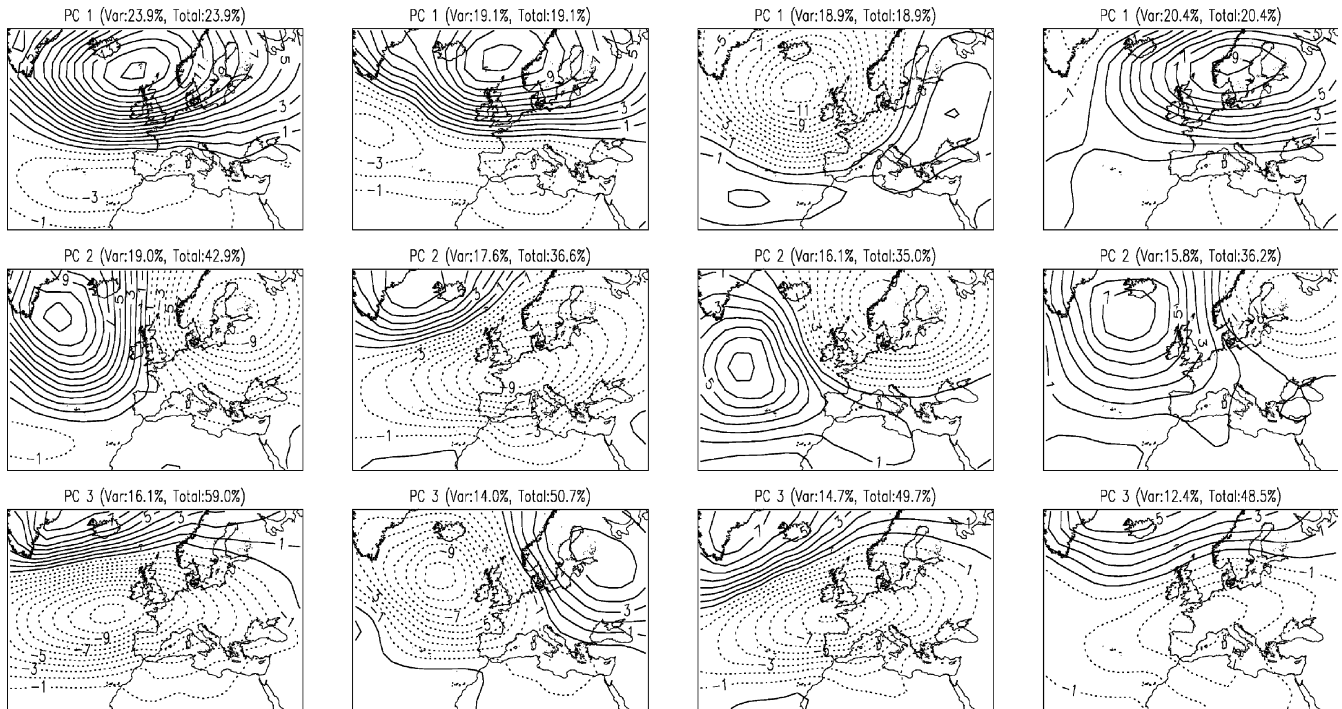


Fig. 3 First three PCs for winter *MSLP* (in hPa): NNR (left), BMRC (second column), CSIRO (third column) and LMD (right) models. The percentage of explained variance is shown for each PC

Mean errors for T_{850} are below 2.5 °C in all seasons for both the CSIRO and the BMRC model. Largest errors are usually below 10 °C, apart from summer for the BMRC model, where a very large but localized maximal error is obtained. The only non flux-corrected model in this study, the LMD model, shows the largest temperature drift, with mean errors from 3.6 to 4.2 °C. The largest errors are located over regions remote from Western Europe: Greenland, the Central Atlantic, North Africa and the Arabian Peninsula. This is encouraging, since Timbal and McAvaney (2001) have shown that T_{850} is strongly related to daily maximum temperature on a small domain around the observation, in this case where the models exhibit the smallest errors. *PWTR* mean errors are below 3.5 mm indicating a realistic modelled moisture content. However, large maximal errors once again indicate strong local mismatches between the models' climatologies and the NNR. Consistent with

those of T_{850} , the largest *PWTR* mean errors are observed in the non flux-corrected LMD model.

Although model systematic errors are likely to affect local climatology over Western Europe, the model's reliability in reproducing the daily synoptic scale variability is the main concern, since model biases are removed by the statistical downscaling. Accordingly, we turn now to the principal components (PCs) calculated using daily NNR on the regular 2.5 by 2.5 grid, based on the covariance matrix. This method allows one to study the main modes of daily variability in a concise manner (Preisendorfer and Mobley 1988; Jolliffe 1989). The first three PCs of *MSLP* are plotted in winter (Fig. 3) and compared with the similar components from the models. The first PC consists of a very active centre where the Icelandic Low is located with a weaker centre of opposite sign near the Azores High. The second PC exhibits two centres of activity of equal intensity. Their northern

position (eastern Baltic Sea and southeast of Greenland) indicates higher variability at high latitude. The third PC has a main centre of activity to the SW of the British Isles. These PCs do not exhibit patterns of ill-defined PCs (Richman 1986; Jolliffe 1989). The size of the chosen domain is sufficient to allow enough degrees of freedom and ensures that main centres (two in each PCs) of daily variability are identified.

Although all three models are affected by large biases in their winter means, the models exhibit realistic modes of variability with close to observed amount of variance (around 50% for the first three PCs compared with 58% in the NNR). Pattern correlations between models and NNR PCs (Table 2) are somewhat higher for the BMRC model (the second and third PCs are ranked in inverse order but since they explain a relatively similar amount of variance, this is not significant) than the other models. The BMRC model exhibits the largest bias for the mean climatology, demonstrating that there is no clear impact of model bias on a model's ability to reproduce daily variance.

T_{850} is another important predictor (Timbal and McAvaney 2001), although its first three winter PCs (Fig. 4) explain only 35% of the variance (compared with 58% for *MSLP*). T_{850} has less spatial coherence than surface pressure as more variance is explained by smaller local features, and hence is less likely to be well captured by the coarse resolution coupled models. Indeed, the first three PCs from the modelled climatologies show large discrepancies compared with the NNR and explain more variance than observed. Large coherent modes of variability are favoured by the models. The first PCs, for example, are close to the observed, but with a displacement of the main centre of variability and large differences in the intensity. In the CSIRO model, the first and second PCs are inverted and the model tends to combine these two modes of variability. In the LMD model, both the second and third PCs are closely related to the second observed PC.

There is little spatial coherence in the daily variability of atmospheric moisture (*PWTR*). The first three PCs explained only 20% of variance (not shown). Not surprisingly, therefore, the models do a poor job in

handling these modes, with very low pattern correlations (Table 2). Pattern correlations are not given when model's modes do not resemble any observed modes.

Results for the same variables are summarized for the other seasons (Table 2) under the form of pattern correlations and explained variance. PCs which are ranked in a different order, in terms of explained variance are denoted with an asterisk. No figures are given when the correlation achieved was below 0.4 or when the modelled pattern exhibit similar correlation (within 0.05) with two observed patterns, suggesting an ill-defined pattern. *MSLP* daily variability is, all year around, the most realistically reproduced, while *PWTR* is generally the worst. No single model stands out as consistently best at reproducing the observed PCs. No particular link could be established during the validation of the models between the size of the model systematic errors and its ability to reproduce daily synoptic variability, although the number of models considered is not large enough to draw any strong conclusion.

The main predictor changes, due to CO_2 increases, are now analyzed. They are defined as seasonal anomalies between the transient run and the control run. Anomalies are remarkably different between models on the continental scale. In winter, for example, little agreement is found amongst the models' patterns for *MSLP* anomalies (first row in Fig. 5). A pressure increase centred over Western Europe indicates a shift further north of the main features (Azores High and Icelandic Low) in the BMRC model; while surface pressure rises also but over northern Europe in the CSIRO model. Both CSIRO and LMD models show an increased cyclonicity over the North Atlantic near Europe. All models agree on a strong warming trend for upper-air temperature, but the magnitude of the warming is about twice as large in the CSIRO model (ranging between 3 to 4 °C over France) as in the BMRC model (1° to 3 °C). The LMD model lies in between with a warming of around 3 °C. This is coherent with the model global responses (Meehl et al. 2000) and their climate sensitivity (Colman 2002). In all three models, maximum warming tends to occur at high and low latitudes with minimum values over the northern parts

Table 2 Correlation between model and NNR PCs for the main predictors (*MSLP*, T_{850} and *PWTR*), for the four calendar seasons and the three GCMs. PCs which are ranked differently are noted with an asterisk. Missing correlations indicate a modelled PC which is not similar to any of the first three PCs from the NNR

		Winter			Spring			Summer			Autumn		
		PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3	PC1	PC2	PC3
BMRC	<i>MSLP</i>	0.91	0.75 *3	0.87 *2	0.96	0.94	0.86	0.95 *2	0.80 *1	0.85	0.94 *2	0.96 *1	0.79
	T_{850}	0.80	0.49	0.58	0.75	0.86			0.51 *3	0.54 *2	0.85	0.89 *3	
	<i>PWTR</i>	0.88	0.86		0.51	0.79	0.62		0.56	0.54	0.58	0.88	0.58
CSIRO	<i>MSLP</i>	0.73	0.72	0.88	0.61	0.58	0.79	0.67	0.82 *3	0.97 *2	0.75	0.86	0.79
	T_{850}	0.68 *2	0.71 *1	0.64	0.88	0.91	0.85	0.74	0.90 *3	0.57 *2	0.89	0.95	0.89
	<i>PWTR</i>	0.81	0.67		0.83	0.83	0.69	0.41	0.61		0.70	0.69	0.61
LMD	<i>MSLP</i>	0.73	0.84	0.74	0.82	0.56	0.79 *2	0.77	0.91 *3	0.75 *2	0.57	0.67	0.53
	T_{850}	0.83	0.41 *2	0.50 *2	0.89	0.75		0.60	0.56 *3	0.54	0.89	0.50	
	<i>PWTR</i>		0.62	0.62	0.47				0.43				

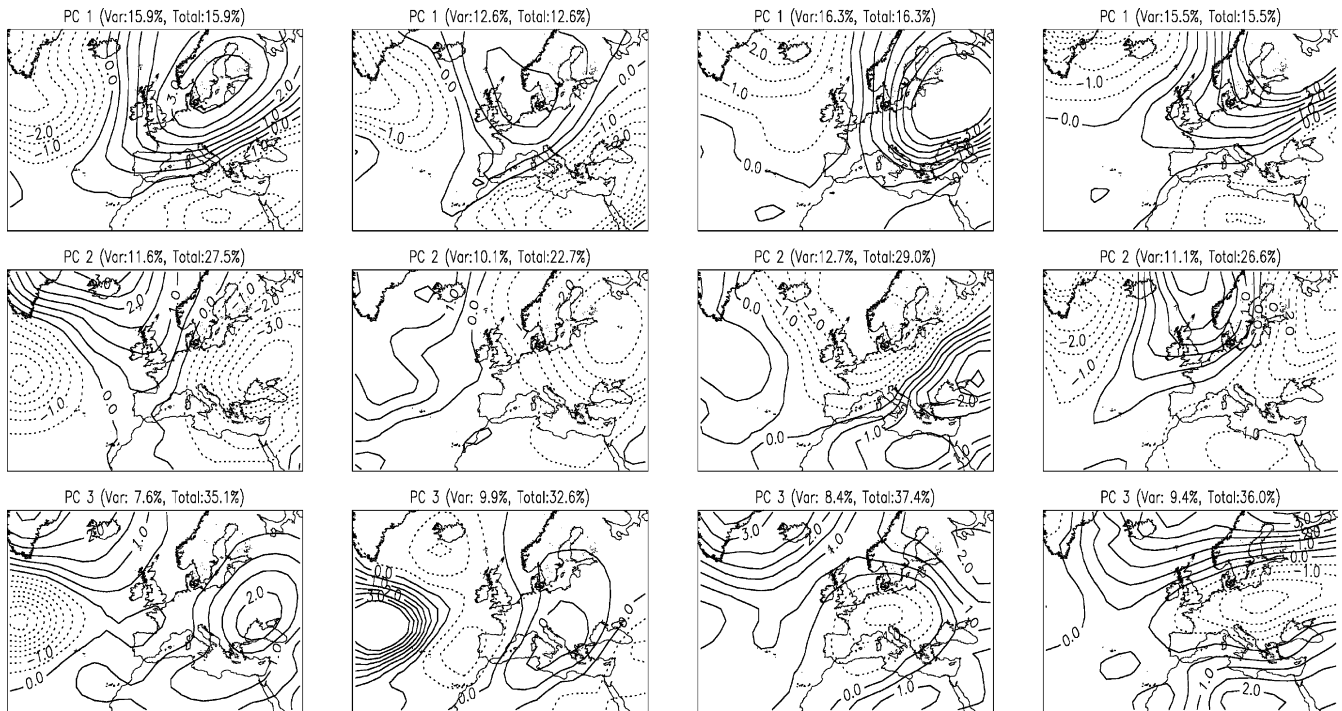
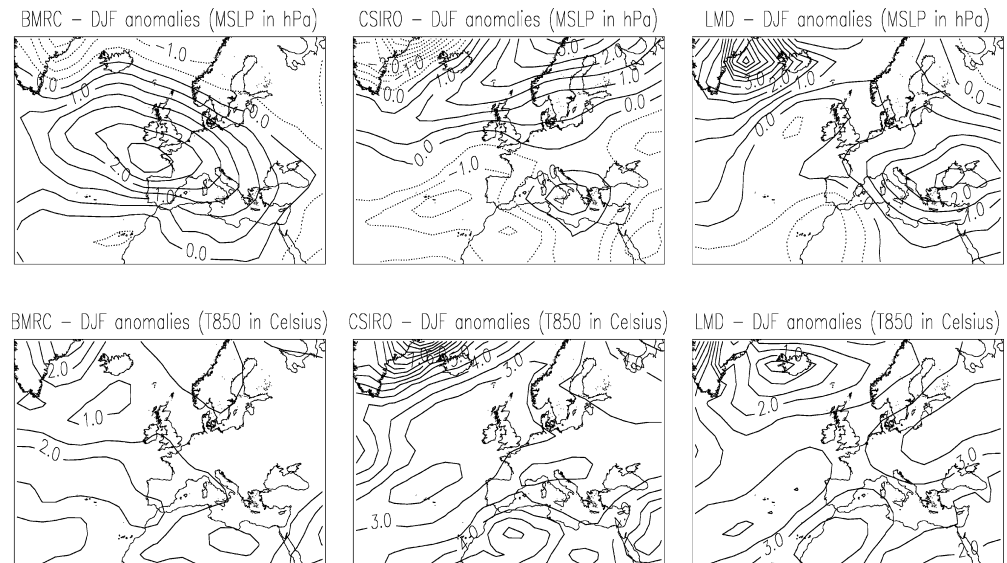


Fig. 4 Same as Fig. 3 but for T_{850} (in °C): NNR (left), BMRC (second column), CSIRO (third column) and LMD (right)

Fig. 5 Winter climate change scenarios provided by the coupled models (BMRC, CSIRO and LMD from right to left) for the three main predictors ($MSLP$, T_{850} and $PWTR$)



of the Atlantic Ocean. Local patterns are clearly different. Mean changes in $PWTR$ indicates a general increase of atmospheric moisture due to the temperature changes. Geographical patterns differ due to dynamical factors. The CSIRO model exhibits the largest signal with a maximum of increased $PWTR$ extending from the tropical Atlantic to the Iberian Peninsula. Similar geographical features are seen in both the LMD and BMRC models with the ridge of maximum increase being displaced north of the Azores Islands.

For the sake of brevity, the other seasons will not be shown. The features seen amongst the models in winter vary greatly throughout the year. For example, the models agree on positive $MSLP$ anomalies directly west of France, with a decrease almost everywhere else, in spring. This suggests increased blocking over Western Europe and this in turn should affect surface predictands. In summer, a similar but further north (west of Scotland), pressure increase is indicated by the three models with, again, a general pressure decrease else-

where (most notably around the Mediterranean Sea). In autumn as in winter, models generally disagree on the pattern of *MSLP* anomalies.

Some features for T_{850} , such as a stronger atmospheric warming over continental Europe than over the Atlantic Ocean are seen year round. The contrast is at its strongest in summer. The largest warming signal is seen in the LMD model (3 to 4 °C in all seasons, winter excepted). This is similar to the CSIRO model (2.5 to 4 °C), while the BMRC model indicates a more moderate warming of around 2 °C. *PWTR* anomalies range, over France and all year round, from 2 to 7 mm. The BMRC model consistently gives the smallest signal with the other extreme is, depending on the season, either the LMD or the CSIRO model.

These results exhibit the spread of responses, to a common CO_2 forcing, as far as the three main predictors, used by the statistical technique, are concern. Therefore the use of these GCMs enables us to explore the uncertainties associated with large-scale GCM projections. This choice of GCMs covers a large range of predicted climate changes under the chosen scenario (Cubasch et al. 2001).

3 The downscaling method

The statistical model (SM) contains numerous parameters discussed in details in Timbal and McAvaney (2001). The particular set up of the technique for Western Europe is presented here. A pool permutation technique was chosen instead of splitting the dataset in two, to ensure the largest possible set of data (Preisendorfer and Barnett 1983). Day by day analogues were chosen in a different calendar year to avoid any artificial skill due to inter-annual variability. A analogue for any particular day was searched for over all the days included in the season amongst the other 40 years available in the NNR (3600 situations). The skill obtained for the SM was compared with a fully cross-validated approach, splitting the time series in two. In one case the SM was optimized over half of the dataset and applied to the other one. In the other case the SM was optimized and applied over the entire dataset. Due to the limited extent of the SM optimization, no artificial skill appears in the results.

Several domains were tested, ranging from a minimal size just encompassing the domain of interest, to the entire grid used for the model field validation (Fig. 6), in all cases both raw grid data and leading PCs were tested. On the one hand, a larger domain requires a longer database to find a suitable analogue (Van Del Dool 1989). On the other hand, a smaller domain does not completely capture the synoptic signal affecting the local station. A medium size was found to be the best compromise and for this particular domain the use of raw data yielded better results than using PCs. A further refinement was added by using different domains for different stations. Northern stations (Dunkerque, Cherbourg, Beauvais and Rennes) perform better when

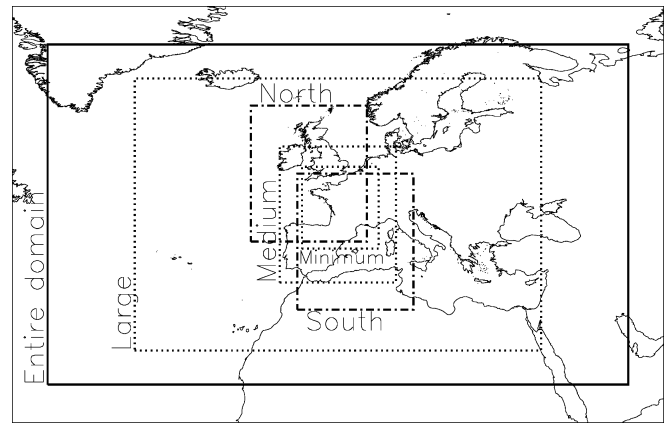


Fig. 6 The various domains tested during the SM development

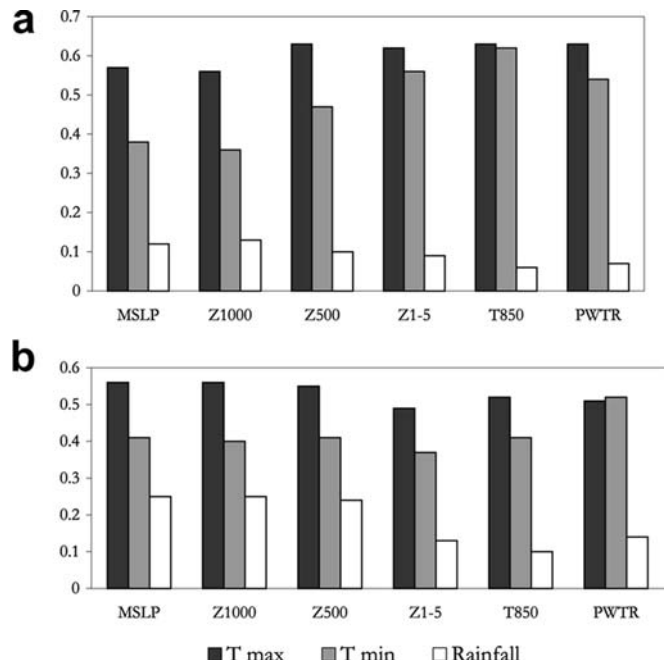


Fig. 7 Correlation between observed and reconstructed series for the predictands (T_{max} , T_{min} and rainfall) in a summer and b winter using several predictors

the domain is towards north and west, while for Mediterranean stations (Montpellier, Perpignan) domains towards the south and east are more effective. Similar dependence using an analogue approach was found by Guilbaud and Obled (1998). It shows that this technique is strongly driven by dominant meteorological processes.

The predictive skill of atmospheric predictors daily departure from seasonal means has been assessed for each season, each station and each predictand using temporal correlation between daily values of the observed and the reconstructed series (Fig. 7). Statistical significance of the results was calculated using a Monte-Carlo approach: a 100 random selection of analogues were made, statistics were calculated from this ensemble for the 90 and 99% level. The dynamic fields (*MSLP*,

Z_{1000} , Z_{500}) perform better for maximum temperatures in winter. Wintertime weather is heavily influenced by cyclonic systems. In other cases, the thermal information given by T_{850} , $Z_{1000} - Z_{500}$ (the thickness difference between 1000 and 500 hPa) and moisture availability ($PWTR$) are more relevant. For rainfall, skill is limited to the dynamic predictors, especially in winter and to some extent to $PWTR$ in winter too. These findings are coherent with several previous studies (see Rummukainen 1997 for an overview). At station level, predictors usually rank in the same order. Apart for T_{max} , T_{850} performs better than $MSLP$ for southern stations (Angoulême, Toulouse, Bordeaux, Montpellier, Perpignan, Limoges and Pau), while for northern stations (Dunkerque, Beauvais, Cherbourg and Rennes) this is reversed. This behaviour is coherent over many stations, showing that the winter time weather type encountered by northern stations does not affect southern stations.

Combinations (from 1 to 6) of predictors are then tested, the mean of all possible combination and the spread between the best and worse combination are shown (Fig. 8). In all cases, two predictors give better results than the individual predictor alone. The mean (stars in Fig. 8) of all possible combinations show increased correlations when up to four predictors are combined. However, the most skillful combination (top of the vertical lines) remains steady and even tends to drop if too many predictors are considered. When more predictors are used, a larger pool of historical analyses must be searched to find an equally good match. The useful combinations were found to be $MSLP$ combined with T_{850} for T_{max} and $MSLP$, T_{850} and $PWTR$ for T_{min} and rainfall.

This completes the optimization of the SM. Overall six different models were used depending on the target variable. The 17 observed stations are separated in three groups from north to south and use different domains; different predictors for T_{max} and for T_{min} or rain occurrence were used. All other parameters are identical in all six SMs. No model has any seasonal dependence.

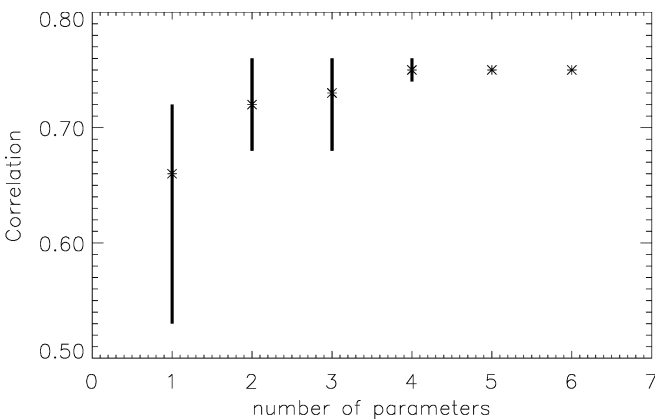


Fig. 8 Range of correlations between observed and reconstructed series for T_{max} , using a combination of several predictors ($MSLP$, T_{850} , $PWTR$, Z_{500} , Z_{1000} and $Z_{1000}-Z_{500}$). The star indicates the mean of all possible combinations

Once the SM is optimized, an analogue is found for each day and the surface predictands observed on the same day are determined. The surface predictand time series are, then, reconstructed day by day and compared to the observed values. The main statistical tools used for this comparison are the correlation and the root mean square errors (RMSE) between the two series. Spatially, correlations and RMSE are very homogeneous (Fig. 9 for T_{max}). The analogue technique preserves the spatial correlation provided by the observed predictands on the chosen date of the analogue. Correlations are lower for coastal stations such as Dunkerque and Cherbourg compared to inland stations; RMSE is also among the lowest for these stations. This is due to a strong oceanic influence which generates low day-to-day variability.

A skillscore is developed to assess the skill of the statistical model and compare it with simpler approaches such as persistence and climatology. This skillscore is similar to that deduced from the Brier score for probabilistic forecasts by Wilks (1995):

$$BSS = \left[1 - \frac{MSE}{MSE_{ref}} \right] \times 100$$

MSE is the mean square error and the reference is a random choice of analogue. A perfect forecast gives a score of 100% while a score of 0 is obtained if analogues are randomly chosen. Negative values are obtained if the method has less skill than a random choice of analogue. This is a direct measure of the skill of the SM in identifying a suitable analogue. The analogue statistical models show skill from 60 to 70% all year round for T_{max} and T_{min} (Fig. 10), indicating a skillful choice of analogue compared to a random choice. It performs better than persistence (in six out of eight cases) and markedly better than climatology (in all cases). Persistence is very high for T_{max} in winter and to a lesser degree in autumn (higher than for T_{min}) making persistence difficult to beat in winter.

The technique was first tested for rainfall amount (Fig. 7). However, correlations with individual predic-

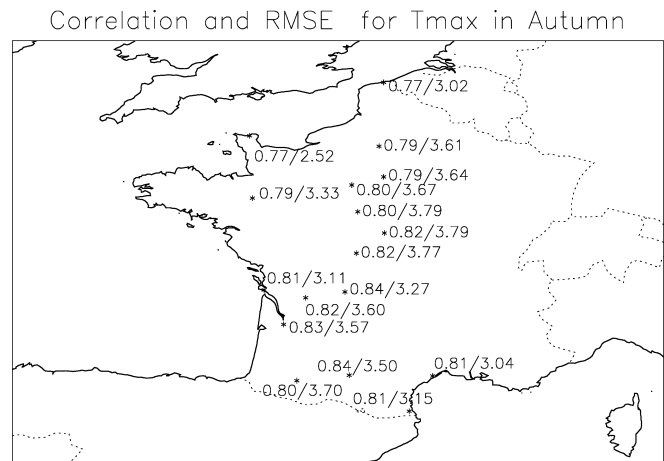


Fig. 9 Correlation and RMSE between analogue reconstructed and observed T_{max} series in autumn for each station

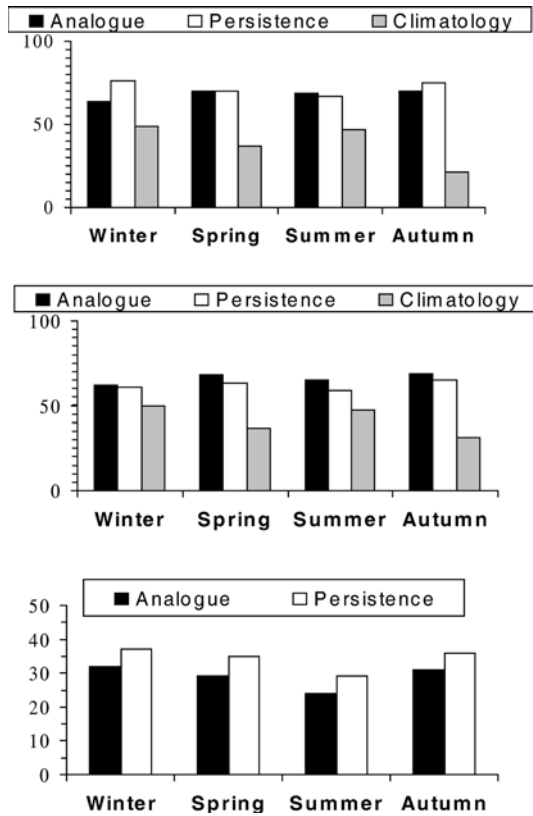


Fig. 10 Skill (in percent) achieved by the analogue technique for T_{max} (top), for T_{min} (middle) and rain days (bottom) and compared with references. See discussion in the text regarding how the skillscores are calculated

tors were below 0.3 in all cases and the reconstructed time series, reproduce the observed probability distribution functions (PDFs) with a large bias towards lower rainfall. Therefore validation was focused on the rain occurrence instead of rain amounts. The ratio between the number of wet days in the analogue series and the observed one was between 0.96 and 1.02 all year around (first column in Table 8) indicating an unbiased method for rain events. The skill of the SM in reproducing rain days was also assessed using an index I defined as:

$$I = 100 \times \left(1 - \frac{m}{w+m} - \frac{f}{d+f} \right)$$

The letters refer to w : a wet day forecast and observed, d : a dry day forecast and observed, m : a wet day missed by the forecast and f : a dry day falsely forecast as wet. This index gives 100% for a perfect forecast ($m = f = 0$); it would give a value of 0 for a random choice of uncorrelated days ($w = m$ and $f = d$). This index takes into account the asymmetrical partition between dry and wet days. A simple scheme of extreme persistence, which would assume either rain (or no rain) every day would have a score of 0, unless such extreme persistence is observed, giving a value of 100%. Results are consistent across all stations, ranging from 0.30 to 0.37 (Fig. 11 for autumn), but lower around the Mediterranean Sea. This reflects the convective nature (i.e. smaller scale) of

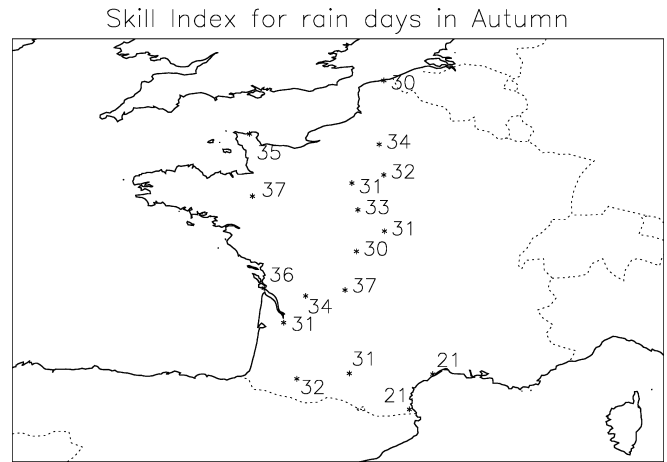


Fig. 11 Index I between analogue reconstructed and observed rain day series in autumn at each station

rainfall for that part of France in autumn. The index I averaged over all the stations used and expressed in percent is around 30% all year around (lowest diagram in Fig. 10), the lowest values are seen in summer. This skillscore is not directly comparable with the score used for temperature. However, for both predictands, the technique can be compared with persistence. For rainfall occurrence, the analogue model is lower all year round than persistence, this was not the case for temperature.

An important effect on the SM skill is the quantity of data used in the optimization of the SM. The impact of the number of years available for the choice of analogue was assessed by looking at RMSE of the reconstructed series when using only part of the NNR dataset. This choice ranged from one year to the entire 41 years and starting from either 1958 and going forward in time or 1998 and going backward in time (Fig. 12). When only a few years are used, RMSE values show large changes when one more year is added. However these jumps are not statistically significant. Instead a logarithm fit to the data is a more useful piece of information. The RMSE decreases rapidly as more years are used; the rate of improvement slows around 25 years when starting in 1998. When further years are added, prior to the 1970s, no advantage is gained using a larger pool to draw analogues from or is counter-balanced by the decrease in data reliability. Since NNR depends on the amount of data available, the latest period for which more observation were available, is bound to be more reliable. Indeed, a clear cut difference appears between the two curves showing that if only 20 years of the NNR dataset were used, the second half of the NNR dataset provides more consistent information leading to smaller RMSE.

From this figure however, it is not clear which part of this reduction can be attributed to errors in the predictors (NNR) or in the surface predictands. As stated earlier, the daily surface observations were found to include temporal inhomogeneities (Mestre 1996). A monthly calibration procedure was applied to remove historical jumps. To test the utility of such homogeni-

zation procedures at the daily time scale, the SM is used with both raw observed series and homogenized ones. Since the statistical model does not depend on the targeted variable, it is used as a tool to measure the relevance and effectiveness of series corrections. Stations are classified into three groups, according to the size of the correction factor. A general but small decrease of RMSE is noted, consistent for all stations (Table 3 shows in autumn for T_{min} , an example for each group of coefficient). The a priori classification into three groups appears to be reasonable. Even at a daily time scale, this homogenization method has improved the observations. Improvements are only significant at the 5% level for the largest corrections. The RMSE difference of the order of 0.1 to 0.3 cannot explain the larger gap between the two curves in Fig. 12 (0.5 around 10 years). Therefore the improved reliability of the NNR data over the period 1958 to 1998 is assumed to explain a large part of the greater skill shown by the SM when using the latter part of the NNR dataset.

After testing the impact of data quality and to maximize the skill of the SM, the homogenized surface temperature are used with the entire NNR from 1958 to 1998. Once optimized, parameters are set and the SM is applied to control run GCM predictors. This is an important step towards downscaling of climate change

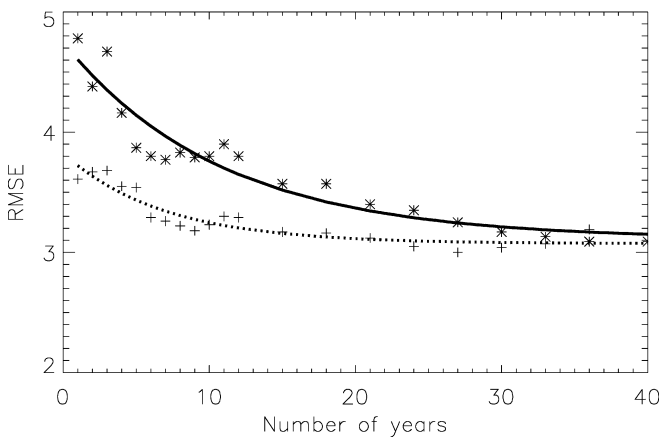


Fig. 12 SM performance measured as RMSE between observed and reconstructed T_{max} series as a function of the number of years used to refine analogues. The *star* and *fitted curve (solid line)* commence in 1958 on wards; the *cross* and *dotted line* start in 1998 and go back in time

Table 3 RMSE between the observed and reconstructed series of T_{min} in autumn using the raw (second column) and homogenized (third column) dataset. Differences significant at the 95% level are shown in bold. The first column indicates the maximum size of the coefficient applied to the series over the 1958 to 1998 period, averaged over the three autumn months, for the particular station

	Maximum Coefficient	Raw series	Homogenized
Bourges	0.4	3.22	3.21
Pau	0.8	3.33	3.19
Cherbourg	1.2	2.81	2.45

scenarios. Mean biases between control runs GCM predictors and NNR are removed. Analogues are chosen from the entire dataset of NNR 1958–1998, and associated with surface predictands observed the same day. To measure the benefit of downscaling GCM large-scale predictors, raw time series provided by the nearest GCM grid points over land are calculated for each station. This is a coarse evaluation since it is generally accepted that GCMs provide information averaged over an entire grid box (Skelly and Henderson-Sellers 1996). (LMD is not used for the direct use of GCM information, since it does not include a diurnal cycle and therefore does not represent daily extremes of temperature.)

For each station, seasonal means of the reconstructed temperature series are compared with the observed ones (Table 4 and 5 for T_{max} and T_{min}). There is no overall tendency towards positive or negative bias. The means were not significantly different, at the 95% confidence level, at any station in any season. Differences are relatively small and without any particular seasonal or geographical trend. This was expected since the technique used anomalies for the predictors, thereby removing model biases and is unbiased as seen when using the NNR (first column) during the validation of the SM. By contrast, direct GCM temperatures (fourth and fifth columns) show biases of up to 5 °C, with a marked tendency toward a smaller than observed amplitude of the diurnal cycle (i.e. cooler T_{max} and warmer T_{min}). Thus, the downscaling technique, by providing unbiased estimates of surface locally observed predictands, is a marked improvement over the raw GCM grid-average values. The variance of the time series was also verified (Table 6 and 7 for T_{max} and T_{min}). Reconstructed series using analogues have smaller variances than observed. This reduction is partly due to the downscaling method, which underestimates the natural variability. This is seen

Table 4 Anomalies (in °C) for seasonal mean of T_{max} between observed and reconstructed series: using the NNR (first column), applying the SM to GCM outputs (BMRC and CSIRO model, 2nd and 3rd columns) and using T_{max} as modelled by the same GCMs using the nearest grid point value (4th and 5th columns)

	NNR	BMRC SM	CSIRO SM	BMRC GCM	CSIRO GCM
DJF	0.07	0.18	0.16	0.39	1.21
MAM	0.30	0.03	0.20	-3.92	-1.78
JJA	0.20	-0.24	0.14	-4.94	-2.76
SON	0.33	0.09	0.30	-1.70	-0.81

Table 5 Same as Table 4 but for T_{min}

	NNR	BMRC SM	CSIRO SM	BMRC GCM	CSIRO GCM
DJF	0.06	0.03	0.12	2.70	2.65
MAM	0.12	-0.03	0.15	0.30	2.48
JJA	0.07	-0.30	-0.12	-0.83	1.29
SON	0.09	-0.07	0.07	1.56	1.41

Table 6 Same as Table 4 but for variance (in °C²) of T_{max}

	NNR	BMRC SM	CSIRO SM	BMRC GCM	CSIRO GCM
DJF	-1.4	-5.1	-2.9	-9.0	-7.8
MAM	-1.6	-6.7	-7.1	-12.0	-14.3
JJA	-0.9	-4.5	-4.9	-5.8	-8.8
SON	-2.2	-8.6	-9.7	-12.3	-15.6

Table 7 Same as Table 6 but for T_{min}

	NNR	BMRC SM	CSIRO SM	BMRC GCM	CSIRO GCM
DJF	-1.8	-3.8	-0.7	-2.8	-0.6
MAM	-1.4	-2.2	-1.3	-3.6	-2.9
JJA	-0.8	-1.1	-1.7	-1.6	-1.3
SON	-1.3	-2.7	-2.4	-5.3	-5.9

Table 8 Ratio of total wet days between reconstructed and observed series: using NNR predictors during the validation phase (1st column), applying the technique to BMRC and CSIRO outputs (2nd and 3rd columns) and using modelled rainfall (4th and 5th columns)

	NNR	BMRC SM	CSIRO SM	BMRC GCM	CSIRO GCM
DJF	0.96	1.07	1.01	1.90	1.81
MAM	0.98	1.00	0.98	1.64	1.86
JJA	1.02	1.09	1.03	1.47	2.40
SON	0.98	1.04	1.04	1.53	1.84

in the first column during the validation phase using NNR, with the reduction of variance varying from 0.8–0.9 °C² in summer to 1.5–2.2 °C² in autumn and winter. This reduction of variance is seen in most downscaling techniques (von Storch 1999). In the case of the analogue technique it is a consequence of the limited pool of data to choose analogues from, which therefore prevents finding the perfect match (Van Del Dool 1994). When GCM predictors are used, there is a tendency toward lower variance, rather small for T_{min} but much larger and more consistent for T_{max} (up to 10 °C²). GCM raw data show arguably as large differences (up to 16 °C²), which are mostly negative with one notable exception in summer. Overall, although biases tend toward lower variance, the SM shows a marked improvement compared with raw GCM outputs.

For rainfall, the numbers of wet days in the reconstructed series are compared with observations (Table 8). Values close to one obtained with the analogue during validation (with NNR) prove that the technique is unbiased. When applied to GCM outputs, the total rain days are overestimated by less than 10%. Compared with this, rainfall modelled by GCMs shows a very large tendency towards more wet days by a factor ranging from 1.4 to 2.4.

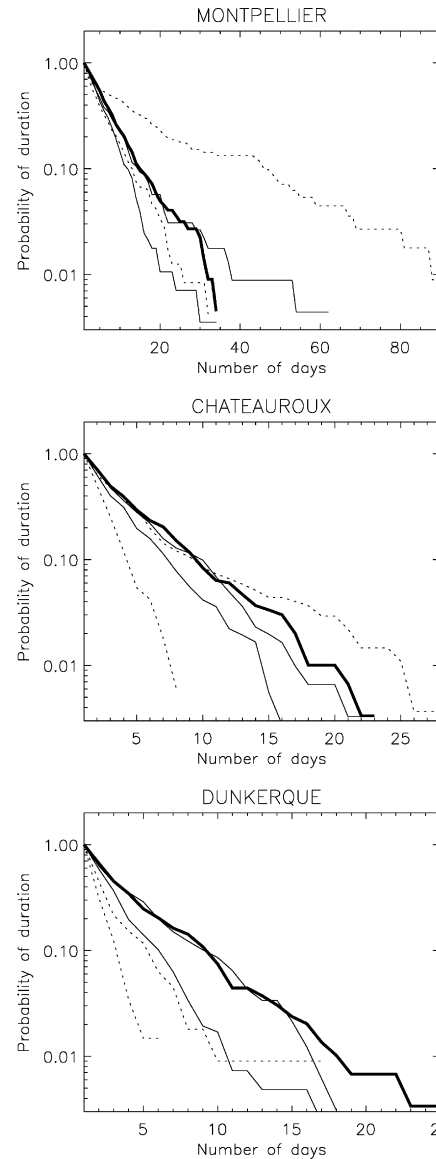


Fig. 13 Observed (*thick line*) dry spell duration in summer in three locations; analogue reconstructed series using BMRC and CSIRO predictors (*thin lines*) and direct GCM rainfall (*dashed lines*)

The temporal structure of the rain occurrence series or conditional probabilities (the four possible combination of wet and dry days following each other) is of particular interest for impact studies (e.g. in agrometeorology). They are examined in the form of the probably most important critical occurrence: dry and wet spells. Dry spell duration (DSD) is shown in summer for three locations generally depicting the range of climate encountered in the western part of France (Fig. 13). Dunkerque (bottom diagram) is representative of the mild oceanic influence which affects most of northwest France, showing a linear relation between the logarithm of probability and spell length. The analogue technique reproduces most of this tendency with a slight underestimation especially for rare events (i.e. those occurring with a probability less than 1%) while the direct GCM

Table 9 Difference (in °C) in Pau and Cherbourg of GEV return value for T_{max} in summer between observed and reconstructed series: using the NNR (1st column), applying the SM to BMRC and CSIRO outputs (2nd and 3rd columns); and using T_{max} as modelled by the same GCMs using the nearest grid point value (4th and 5th columns)

		BMRC SM	CSIRO SM	BMRC GCM	CSIRO GCM
Pau	10	-0.2	-0.5	-13.1	-4.9
	25	-0.3	-0.6	-8.5	-5.1
	50	-0.3	-0.5	-4.0	-5.2
	100	-0.1	-0.6	4.3	-5.2
Cherbourg	10	0.1	-0.2	-6.2	-7.7
	25	0.1	-0.1	-5.8	-8.2
	50	0.2	-0.2	-4.7	-8.6
	100	-0.1	0.5	-3.6	-8.8

control climate show very few long dry spells. In Chateauroux (middle diagram), as in most central parts of France, observed DSDs are similar to those further north, but large disagreements in some cases and are generally poor matches to the observed PDFs.

One of the key issues in the climate change debate is how extreme events will be affected. Here, we apply the generalized extreme value (GEV) distribution described by Zwiers et al. (1998) to calculate the 10, 25, 50 and 100 year return values of T_{max} . Return values for two very different climates (Pau and Cherbourg) are shown in Table 9. During the validation phase (using NNR), the analogue values match the observed. This remains valid when the SM is applied to GCM outputs in the return values in most cases only slightly under-evaluated (within 0.5 °C of value calculated from observation). However most differences with GCM control run raw data are around 4 to 8 °C.

4 Downscaling of GCM climate change scenarios

The observations recorded from 1958 to 1998 show a significant warming of 1.1° (0.9 °C) for T_{max} (T_{min}) in the annual mean. This was calculated by applying a linear regression fitted to each individual station and then averaged over the 17 locations (Table 10). The reconstructed series using the analogue technique reproduces between 70% and 90% of this trend for T_{max} . The difference between a large warming in summer and a smaller one in autumn is also reproduced. A smaller part of the trend on T_{min} is reproduced by the analogue technique. These results were obtained using $MSLP$ and T_{850} as predictors; when $PWTR$ is also used all percentages of reproduced trend drop by approximately the analogue technique is more successful in reproducing these probabilities. Note that a large spread exists between the two GCMs used. This spread is, however, smaller than with direct GCM rainfall. Finally, Montpellier (top diagram) is typical of a Mediterranean climate with long dry spell in summer: spells longer than 20 days are seen 6% of the time, 10 times more than in

Table 10 Linear warming trend (°C) deduced over the 1958 to 1998 period for both T_{max} and T_{min} observation; percentage of this trend reproduced by the analogue technique

	T_{max}		T_{min}	
	Observation	SM	Observation	SM
DJF	1.15	77%	1.04	42%
MAM	1.15	69%	0.82	26%
JJA	1.73	89%	1.52	47%
SON	0.39	74%	0.52	23%

other locations. The analogue technique reproduces this behaviour but with a large spread amongst models. Raw GCM outputs are even more spread and far from observations. If anything, DSDs are rather underestimated in all cases, Montpellier excepted. This illustrates for extremes cases the underestimation of the conditional dry-dry probability $\widehat{P}_{d,d}$ which varies from north to south. In Dunkerque observed $\widehat{P}_{d,d}$ is 0.75 but the analogue technique only gives a value of 0.72 while in Montpellier the analogue technique reproduces the higher observed value: 0.87.

The strength of the analogue technique is its ability to reproduce not only the mean state of a variable but also the details of its probability distribution function (PDF). The statistical model successfully reproduces the varying shapes from one station to another (Fig. 14 for T_{max} in summer in four locations). Royan, Dunkerque and Perpignan exhibit a rather narrow peak of maximum probability around the mean (the mean rising from north to south) typical of a maritime influence, and asymmetrical tails toward extreme temperatures. Bourges exhibits a more continental climate with a broader peak and symmetrical tails. Downscaled GCMs T_{max} (light grey shapes) show differences with observations as they tend to concentrate on the central values, leading to fewer extreme cases. This is consistent with the previously noted reduction of variance for T_{max} . However, results are close to observed PDFs and the differences between using BMRC and CSIRO control runs are small. In contrast, raw GCM outputs (dark grey shapes) show 20%–30%. This must be kept in mind when analyzing the sensitivity of the SM model to climate change for T_{min} in particular for which $PWTR$ was found to be a useful predictor (see earlier discussion).

The interannual variability of the observed time series is also of interest. The correlations between reconstructed series using the analogue SM and observed series are, for all seasons and stations, above 0.8 for T_{max} , 0.7 for T_{min} and 0.6 for rain days (Table 11). Although generally better for T_{max} and in winter, the SM shows coherent inter-annual skill, across all the predictands and throughout the year. At station level (Fig 15: Beauvais in winter), year to year variations are much larger than anticipated climate change: about 7 °C for T_{max} , 5 °C for T_{min} and about a factor 2.5 for total rain days per season. The reconstructed series reproduce a large part of these extremes, which reinforces our con-

Fig. 14 Probability distribution functions for T_{max} in summer in four different locations. The black line is the observed PDF, the dark grey shape is derived from two GCM surface temperatures and the light grey shape is obtained by applying the downscaling technique to the same two GCMs. Both shapes are an envelope of the 2 GCM responses

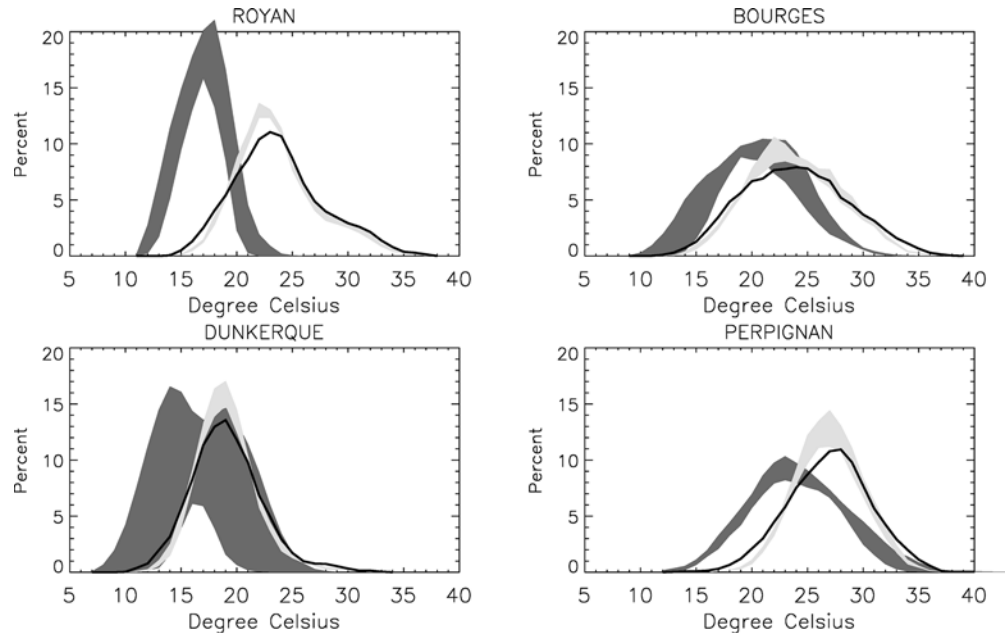


Table 11 1958 to 1998 correlation between reconstructed series using the analogue technique and observations for T_{max} , T_{min} and Rain days

	T_{max}	T_{min}	Rain days
DJF	0.89	0.85	0.75
MAM	0.88	0.74	0.69
JJA	0.86	0.78	0.60
SON	0.82	0.78	0.69

confidence in the applicability of the technique to modified climatic conditions. The ratio of inter annual variability between reconstructed series and observed is usually between 40 and 90% (43, 72 and 56% for T_{max} , T_{min} and rain days in Beauvais). It is worth noting that, although the analogue technique reproduced the observed warming trend for T_{max} better than for T_{min} , the inter-annual variability in the reconstructed time series resembles the observed one more closely for T_{min} than for T_{max} (Fig. 15 illustrates this for a particular station and season).

The SM skill in representing interannual variability and observed trends shows the sensitivity of the technique to changing climatic conditions. It remains to be determined whether future realizations of climate can be drawn from daily situations observed for present conditions. A simple test is to compare the Euclidean distance used to define the matching analogue (Barnett and Preisendorfer 1978). This is a measure of the closeness of the analogue with the matching situation for both the control and transient simulations and shown without units since it is a normalized value (Table 12). The same predictors over the same domains were used for all applications of the SM to GCM outputs, therefore all distances are comparable. The distance increases between control and transient runs for all three GCMs. These differences are small: from 2% for the BMRC

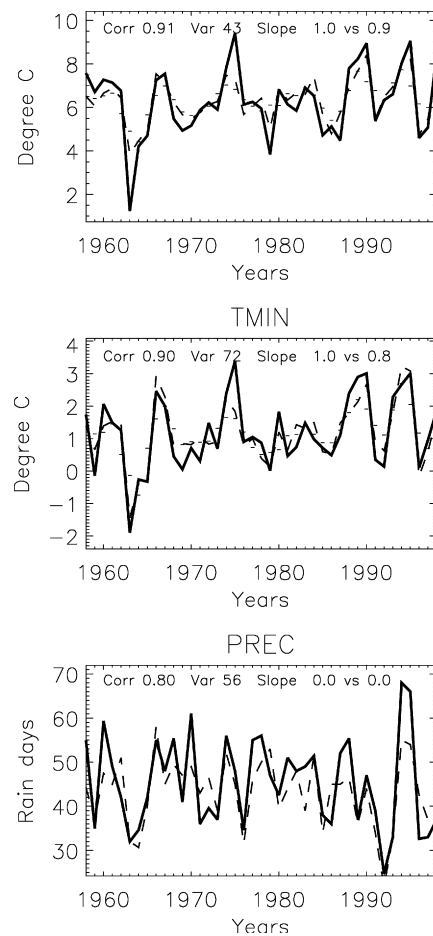
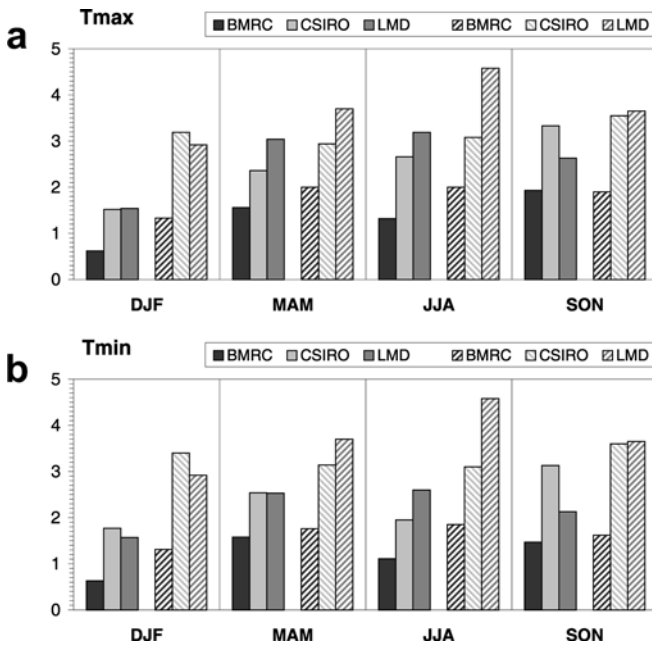


Fig. 15 Interannual variability observed (solid line) in Beauvais, in winter, for T_{max} (top), T_{min} (middle) and rain days (bottom) compared with the reconstructed series using the analogue based method (dashed line). Correlation between the two curves (corr), ratio of the variance of the analogue series over the observed one (var) and linear trend (slope) observed versus analogue are given on each graph

Table 12 Euclidean distance between days and their analogue average over 20 year-periods for control and transient runs

	BMRC	CSIRO	LMD
1 * CO ₂	3.17	2.82	2.63
2 * CO ₂	3.24	2.96	2.81
4 * CO ₂			3.52

**Fig. 16** a Estimated warming for T_{max} and b T_{min} , for each season, using the downscaled projections (left 3 bars) and direct GCM outputs (right 3 bars)

model to 7% for the LMD model and remains below the spread amongst the models. By contrast, the same measure was applied to an LMD run stabilized at the 4 * CO₂ level, and showed an increase larger than 30%. The results seem to indicate that actual climate has enough variability to contain a range of situations wide enough to encompass future possible situations due to a doubling of CO₂ concentration. However, for more drastic climate changes such as 4 * CO₂ climate, the application of the analogue technique is questionable.

Seasonal local warmings averaged over the 17 stations are shown for both T_{max} and T_{min} , in both the downscaled and direct model projections (Fig. 16). As the LMD model does not incorporate the diurnal cycle it only provides the mean daily temperature, which has been used for both T_{max} and T_{min} . This makes the direct LMD estimate somewhat inconsistent with the other estimates.

A striking feature of Fig. 16 is a general reduction of the expected local warming with downscaled projections compared with the direct GCM output. This reduction is large (up to 2 °C in some cases), significant (e.g. in winter, downscaled projections show only half the warming predicted using direct model outputs) and consistent across all cases (apart for T_{max} in autumn with the BMRC model). Uncertainties related to the

warming obtained with the downscaling technique were assessed by using several sets of variables. It was found that the spread was small and the signal robust. The warming obtained for T_{min} is larger when *PWTR* is used, while it was found earlier that including *PWTR* had a negative effect on the reproduction of observed trend for both T_{min} and T_{max} (the returns are smaller than observed trends). This suggests that although the statistical relationship does not explain fully the observed variance and reproduces only part of the observed trend, this may not link directly with the ability of the technique to reproduce future trends and therefore does not explain the reduced warming in the downscaled projections.

Another important difference between direct and downscaled projections concerns the annual signature of the warming trend. Warming tends to be smaller in winter (in all three downscaled cases) and peaks in spring and summer with the LMD scenarios or in autumn, with the CSIRO scenario. In the BMRC model warming peaks in both transition seasons, spring and autumn. In all three models the annual cycle in the warming is quite large. The warming in autumn for T_{max} , for example, is about three times the winter value in the BMRC scenario and this tends to enlarge the control annual cycle. This annual cycle of the warming, in the direct GCM outputs, is only seen with the LMD model (amid a much larger estimated warming), and is not apparent in the two other direct GCM projections. Observed trends in the past 40 years (Table 10) shown earlier also suggested an annual cycle in the warming trend, but although summer stands out as exhibiting the largest warming, winter warming in recent decades was not the smallest in the other three seasons.

Observed trends indicate a slightly larger warming for T_{max} than for T_{min} in the past half century over western France, in the annual mean. Similar findings are apparent using the downscaled projections. However differences arise when seasons are considered. Projected warming is larger for T_{max} in summer and autumn (up to 0.7 °C) while T_{min} is predicted to rise slightly more in winter (up to 0.2 °C). There is an overall agreement between the three downscaled projections when comparing T_{min} and T_{max} warmings. Such agreement was not seen for direct GCM scenarios with the BMRC model indicating larger warming for T_{max} , whereas the opposite is true in the case of the CSIRO model.

Overall, the consistency in the detailed estimates of future warming amongst downscaled scenarios increases confidence in such estimates. Another advantage is the ability of the downscaling technique to provide more detailed scenarios from one location to another. Some interesting features appear when results are analyzed station by station. We only discuss here the features that are consistent amongst all models and therefore more reliable; direct GCM outputs are not considered since their poor horizontal resolution prevents them from providing a detailed estimate at the local scale. The large warming seen for T_{max} in spring, summer and autumn is maximum inland and tends to be reduced along the

English Channel, the Atlantic Ocean and the Mediterranean Sea (not shown). This reduction increases with the mean warming: in the BMRC scenario which shows an average warming of 1.5 °C, the warming along the coast is 50% less than inland while in the CSIRO model the difference is about a factor one and a factor of two in the LMD model where the warming inland exceeds 4 °C and remains below 2 °C along the coast. Such behaviour is not observed for T_{min} , where the strongest warming is observed in the south near the Pyrenees and decreases to a minimum in the north. The north-south gradient remains below 1 °C.

T_{max} in winter shows differences amongst the projections (Fig. 17). In the CSIRO and LMD model, the warming is very small along the north-west coast (Royan, Rennes, Cherbourg and Dunkerque) but increases rapidly when moving inland: e.g. it doubles between Royan and Bordeaux or Angoulême, only 50 km in land. In the BMRC model, the oceanic influence extends across the whole western part of France. Thus the warming is only slightly larger inland than near the coast. The BMRC climatology in winter was shown to be too zonal (Fig. 2), although this tendency was reduced in the transient experiment (Fig. 5). It is anticipated that this strong zonal influence could limit the warming trend expected over western France. This therefore limits the confidence one would place in this particular case. Another sharp differences in warming trends between nearby stations is seen between Pau and Toulouse, 100 km apart. Situated in the foothills of the Pyrenees, Pau's response to large-scale forcing is 50% higher in all three projections than in Toulouse, which is located further into the Garonne Plain.

The analogue technique itself cannot infer what might happen to record-breaking temperatures since future realizations are drawn from the current climate. However, the modification of PDFs can be analyzed. There is a general tendency for the PDFs (not shown) to shift towards warmer temperatures. It is only in the case of

stronger warming (such as in summer) that there are changes in the shape, with more frequent very warm days. The GEV technique, which was shown earlier to give realistic estimates of 10, 25, 50 and 100-years return period temperatures, shows a general upward trend when applied to downscaled projections. This trend varies from north to south with some negative values near the English Channel to large positive values near the Mediterranean Sea. Values from two extreme stations, Bordeaux and Dunkerque, are shown for both downscaled projections and direct model outputs (Table 13). Only results from one GCM (the CSIRO model) are shown for clarity; but similar behaviour was evident in the other models. The warming trend is smaller using the downscaled projections compared with direct GCM outputs which show some extremely large values (up to 10 °C for the 100 year return value in Bordeaux). The estimates provided by applying the GEV to analogue reconstructed time series are conservative but still large enough to have significant impact (e.g. the 100-year return value increases by 3 °C, a figure larger than the mean warming in Bordeaux and most of the southwest of France).

An estimate of the total rainfall variation under warmer conditions was not attempted since the technique was shown to provide reliable estimates for rain occurrence only. However, information regarding rain days are relevant to impact studies. Downscaled projections tend to show a slight reduction of total rain days all year round (Table 14) similar to direct GCM outputs (not shown), but based on a more reliable estimate of total rain days in the control run (see previous section). This reduction rarely exceeds 10%, except when using the LMD model. Summer, which is considered the most critical season (since any rainfall diminution would have greater consequences) shows contrasting results. There is a small decrease in rain days in both the BMRC and CSIRO models but a large increase with the LMD

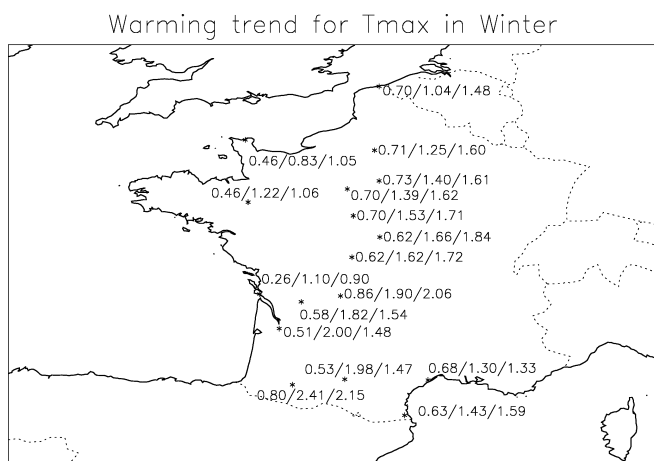


Fig. 17 Geographical repartition of the expected warming trend on T_{max} in winter. Results for each location are obtained using downscaled scenarios with the BMRC/CSIRO/LMD models

Table 13 Differences for 10, 25, 50 and 100-year return values for T_{max} in summer between control and transient simulations using the downscaling technique applied to the CSIRO GCM and using direct model outputs at two locations

	Bordeaux		Dunkerque	
	Analogue	GCM	Analogue	GCM
10	0.8	6.3	0.0	4.1
25	1.2	7.5	0.1	4.3
50	2.3	8.9	-0.2	5.2
100	3.3	10.4	0.1	4.9

Table 14 Ratio of total wet days between control and transient simulations, using the downscaling technique applied to the GCMs

	BMRC	CSIRO	LMD
DJF	0.89	0.96	0.88
MAM	0.96	1.05	0.85
JJA	0.93	0.94	1.22
SON	0.95	0.97	0.89

results. For wet and dry spells, no large or coherent differences were found between control and transient scenarios. In most seasons, there were more uncertainties between the three models available than between control and transient scenarios. In the light of the small signal obtained with rain days and without further information on rain amount no strong conclusions could be drawn on estimates of future rainfall.

5 Conclusions

A statistical downscaling model previously developed for Australia has been used to provide projections for future climate change in western France. The technique complements dynamical approaches for climate change studies performed with regional coupled models and allows finer time and spatial resolutions. It has been used to provide projections for daily temperature extremes and rain occurrence, which are particularly critical for impact studies. Data from 17 high quality stations have been used. The data quality was carefully checked and historical jumps were reduced using a homogenization procedure. The homogenization method was originally designed to apply to monthly values. By applying the statistical model (SM) to both raw and homogenized daily values, it was shown that the homogenization method has positive impacts on daily values. Predictors were chosen according to their predictive skills and their suitability as GCM outputs. The most effective domain for the predictors is rather small and varies between the most southerly and northerly coastal stations and the rest of the inland stations, reflecting local influences. It was shown that the quality of the dataset used influences the skill of the statistical model, stressing the need for further improved observations and reanalyses.

The reconstructed time series shows good agreement with observations: the mean is well reproduced and local probability functions are realistic for all stations. However, a general tendency toward smaller variance was noted for temperature. Wet and dry occurrences were reproduced with some skill but not rain amounts. The analogue technique has skill throughout the year to reproduce T_{max} and T_{min} . An important aspect of this study was to apply the SM to several coupled GCM simulations for both control and transient scenarios. The ability of coupled model control runs to reproduce the main large-scale features of present-day climate has been carefully checked for the main predictors used ($MSLP$, T_{850} and $PWTR$). Although the models show large biases, daily variability was found to resemble observed modes. No particular model appears to outperform the others. The three models produced markedly different climate change scenarios for the main atmospheric predictors. This helps to quantify uncertainties associated with both future large-scale climate changes and model sensitivities.

The SM applied to the control simulation of the models provided reliable estimate of local predictands

series: unbiased but with a reduced variance. The characteristics of the reconstructed time series based on analogues are much more realistic than those obtained directly from the nearest GCM grid point. Added value when using this downscaling approach is particularly visible when dealing with extreme events (anomalous spells or return period of record events). The technique reproduces partially recently observed trends and inter-annual variability. These two elements support the idea that this technique is robust for altered climatic conditions. Although results suggest a possible limitation due to the incomplete explained variance by the statistical technique. It was further noted that there was no significant increase in the difficulty of finding suitable analogues when the SM was applied to $2 * CO_2$ scenarios. However, the technique might fail for more drastic climatic change such as a $4 * CO_2$ increase, as the pool of observed situations used does not seem to be fully representative of the conditions encountered in this latter case.

However, one must keep in mind that the method relies on GCM large-scale projections and therefore may not be a reliable estimate of future climatic change if very large errors are present in coupled model predictions. Possible alterations of the statistical link between predictors and predictands cannot be completely ruled out neither. However, within the largely accepted framework of dynamical prediction of climate changes using coupled GCMs, the downscaling technique presented here has been shown to provide detailed local climate change projections.

Acknowledgements We are grateful to O. Mestre who provided the surface observations. Thanks to J. Hilt who performed the return value calculations and to R. Colman and D. Howard (from BMRC), G. Whetton (from CSIRO), and L. Fairhead (from LMD) for providing data from their coupled model simulations. Useful comments on earlier version of the manuscript were made by R. Colman, S. Power, N. Nicholls and two anonymous reviewers. BT was funded by the Australian Greenhouse Office.

References

- Barnett TP, Preisendorfer RW (1978) Multifield analog prediction of short-term climate fluctuations using a climate state vector. *J Atmos Sci* 35: 1771–1787
- Braconnot P, Marti O, Joussaume S (1997) Adjustment and feedbacks in a global coupled ocean–atmosphere model. *Clim Dyn* 13: 507–519
- Cess RD, Potter GL, Blanchet J-P, Boer GJ, Del Genio AD, Déqué M, Dymnikov V, Galin V, Gates WL, Ghan SJ, Kiehl JT, Lacis AA, Le Treut H, Li Z-X, Liang X-Z, McAvaney BJ, Meleshko VP, Mitchell JFB, Morcrette J-J, Randall DA, Rikus L, Roeckner E, Royer J-F, Schlese U, Scheinin DA, Slingo A, Sokolov AP, Taylor KE, Washington WM, Wetherald RT, Yagai I, Zhang M-H (1990) Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models. *J Geophys Res* 95: 16,601–16,615
- Cess RD, Zhang M-H, Ingram WJ, Potter GL, Alekseev V, Barker HW, Cohen-Solal E, Colman RA, Dazlich DA, Del Genio AD, Dix MR, Dymnikov V, Esch M, Fowler LD, Fraser JR, Galin V, Gates WL, Hack JJ, Kiehl JT, Le Treut H, Lo KK-W, McAvaney BJ, Meleshko VP, Morcrette J-J, Randall DA, Roeckner E, Royer J-F, Schlesinger ME, Sporyshev PV, Taylor

- KE, Timbal B, Volodin EM, Wang W-C, Wetherald RT (1996) Cloud feedback in atmospheric general circulation models: an update. *J Geophys Res* 101: 12,791–12,794
- Colman RA (2002) Geographical contributions to global sensitivity in a general circulation model. *Global Planet Change* 32: 211–243
- Colman RA, Fraser JR, Rotstayn L (2001) Climate feedbacks in a general circulation model incorporating prognostic clouds. *Clim Dyn* 18: 103–122
- Conoway D, Wilby RL, Jones PD (1996) Precipitation and air flow indices over the British Isles. *Clim Res* 7: 169–183
- Cubasch U, Meehl GA, Boer GJ, Stouffer R, Dix M, Noda A, Senior CA, Raper S, Yap KS (2001) Climate change 2001 the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change IPCC, University Press Cambridge, UK
- Giorgi F, Hewitson B, Christensen J, Hulme M, Von Storch H, Whetton P, Jones R, Mearns L, Fu C (2001) Climate change 2001, the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change IPCC, University Press, Cambridge, UK, ch 10
- Giorgi F, Marinucci MR, Visconti G (1990) Use of limited area model nested in a general circulation model for regional simulations over Europe. *J Geophys Res* 95: 18,413–18,431
- Gordon HB, O'Farrel SB (1997) Transient climate change in the CSIRO coupled model with dynamic sea-ice. *Mon Weather Rev* 125: 975–1007
- Guilbaud S, Obled C (1998) L'approche par analogues en prévision météorologique (in french). *La Météorologie* 24: 21–35
- Gyalistras D, von Storch H, Fischlin A, Beniston M (1994) Linking GCM-simulated climatic changes to ecosystem models: case studies of statistical downscaling in the Alps. *Clim Res* 4: 162–189
- Hewitson BC, Crane RG (1996) Climate downscaling: techniques and application. *Clim Res* 7: 85–95
- Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, Maskell K, Johnson CA (2001) Climate change 2001, the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change IPCC, University Press, Cambridge, UK
- Hulme M, Briffa KR, Jones PD, Senior CA (1996) Validation of GCM control simulations using indices of daily airflow types over the British Isle. *Clim Dyn* 9: 95–105
- Jolliffe I (1989) Rotation of ill-defined principal components. *Appl Statist* 38: 139–147
- Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Derber J, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetma A, Reynolds R, Jenne R, Joseph D (1996) The NCEP/NCAR 40-year reanalysis project. *Bull Am Meteorol Soc* 77: 437–471
- Kattenberg A, Giorgi F, Grassl H, Meehl GA, Mitchell JFB, Stouffer RJ, Tokioka T, Weaver AJ, Wigley TML (1996) Climate change 1995, the science of climate change, Contribution of Working Group I to the Second Assessment Report of the Intergovernmental Panel on Climate Change IPCC, University Press, Cambridge, UK, ch 6
- Lambert SJ, Boer GJ (2001) CMIP1 evaluation and intercomparison of coupled climate models. *Clim Dyn* 17: 83–106
- Leggett J, Pepper WJ, Swart RJ (1992) Climate change 1992, the supplementary report to the IPCC scientific assessment, the Intergovernmental Panel on Climate Change IPCC, University Press, Cambridge, UK, ch 3
- Lorenz (1969) Atmospheric predictability as revealed by naturally occurring analogues. *J Atmos Sci* 26: 636–646
- Martin E, Timbal B, Brun E (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
- McAvaney BJ, Colman RA (1993) The BMRC AGCM: AMIP configuration. Tech Rep 38, Bureau of Meteorology Research Centre, Melbourne, Australia
- McAvaney BJ, Covey C, Joussaume S, Kattsov V, Kitoh A, Ogana W, Pitman AJ, Weaver AJ, Wood RA, Zhao Z-C (2001) Climate change 2001, the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change IPCC, University Press, Cambridge, UK, ch 8
- Mearns LO, Hulme M, Carter TR, Leemans R, Lal M, Whetton P (2001) Climate change 2001, the scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change IPCC, University Press, Cambridge, UK
- Meehl GA, Boer GJ, Covey C, Latif M, Stouffer RJ (2000) The Coupled Model Intercomparison Project (CMIP), *Bull Am Meteorol Soc*, 81: 313–318
- Mestre O (1996) Ruptures multiples et points aberrants dans un modèle Gaussien, application à l'homogénéité relative des longues séries climatologique (in French). Note du SCEM 20, Météo-France
- Mestre O (2000) Méthodes statistiques pour l'homogénéisation de longues séries (in French). PhD Thesis, Université Paul Sabatier, Toulouse, France, pp 238
- Murphy J (1999) An evaluation of statistical and dynamical techniques for downscaling local climate. *J Clim* 12: 2256–2284
- Palutikof JP, Winkler JA, Goodess CM, Andresen JA (1997) The simulation of daily temperature time series from GCM output. Part I: comparison of model data with observations. *J Clim* 10: 2497–2513
- Power SB, Colman RA, McAvaney BJ, Dahni RR, Moore AM, Smith NR (1993) The BMRC Coupled atmosphere/ocean/sea-ice model. Research Rep 37, BMRC, pp 58
- Preisendorfer RW, Barnett JP (1983) Numerical model-reality intercomparison tests using small-sample statistics. *J Atmos Sci* 40: 1884–1896
- Preisendorfer RW, Mobley CD (1988) Principal component analysis in meteorology and oceanography. Elsevier, Amsterdam
- Raper SCB, Cubasch U (1996) Emulation of the results from a coupled general circulation model using a simple climate model. *Geophys Res Lett* 23: 1107–1110
- Richman MB (1986) Rotation of principal components. *Int J Climatol* 6: 293–335
- Rummukainen M (1997) Methods for statistical downscaling of GCM simulations. SMHI Meteorology Climatology Reports 80, Swedish Meteorological and Hydrological Institute
- Skelly WC, Henderson-Sellers A (1996) Grid box or grid point: what type of data do GCMs deliver to climate impact researchers? *Int J Climatol* 16: 1079–1086
- Timbal B, McAvaney BJ (2001) An analogue-based method to downscale surface air temperature: application for Australia. *Clim Dyn* 17: 947–963
- Van Del Dool HM (1989) A new look at weather forecasting through analogues. *Mon Weather Rev* 117: 2230–2247
- Van Del Dool HM (1994) Searching for analogues, how long must we wait? *Tellus* 46A: 314–324
- von Storch H (1999) On the use of “inflation” in statistical downscaling. *J Clim* 12: 3505–3506
- von Storch H, Zorita E, Cubasch U (1993) Downscaling of global climate estimates to regional scale: an application to Iberian rainfall in wintertime. *J Clim* 6: 1161–1171
- Wilby RL, Wigley TM, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS (1998) Statistical downscaling of general circulation model output: a comparison of methods. *Water Resource Res* 34: 2995–3008
- Wilks DS (1995) Statistical methods in the atmospheric sciences. Academic Press, New York
- Zorita E, Hughes JP, Lettemaier DP, von Storch H (1995) Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *J Clim* 8: 1023–1042
- Zorita E, von Storch H (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *J Clim* 12: 2474–2489
- Zwiers FW, Kharin VV (1998) Change in the extremes of the climate simulated by CCC GCM2 under CO₂ doubling. *J Clim* 11: 2200–2222