STARDEX

STAtistical and Regional dynamical Downscaling of EXtremes for European regions

EVK2-CT-2001-00115

Deliverable D13

Recommendations on the most reliable predictor variables and evaluation of inter-relationships

FOREWARD

The STARDEX project on STAtistical and Regional Dynamical downscaling of EXtremes for European regions is a research project supported by the European Commission under the Fifth Framework Programme and contributing to the implementation of the Key Action "global change, climate and biodiversity" within the Environment, Energy and Sustainable Development.

STARDEX will provide a rigorous and systematic inter-comparison and evaluation of statistical and dynamical downscaling methods for the construction of scenarios of extremes. The more robust techniques will be identified and used to produce future scenarios of extremes for European case-study regions for the end of the 21st century. These will help to address the vital question as to whether extremes will occur more frequently in the future.

For more information about STARDEX, contact the project co-ordinator Clare Goodess (c.goodess@uea.ac.uk) or visit the STARDEX web site:
http://www.cru.uea.ac.uk/projects/stardex/

STARDEX is part of a co-operative cluster of projects exploring future changes in extreme events in response to global warming. The other members of the cluster are MICE and PRUDENCE. This research is highly relevant to current climate related problems in Europe. More information about this cluster of projects is available through the MPS Portal:
http://www.cru.uea.ac.uk/projects/mps/

STARDEX is organised into five workpackages including Workpackage 3 on 'Analysis of GCM/RCM output and their ability to simulate extremes and predictor variables' which was responsible for the production of this deliverable (D13). Workpackage 3 is co-ordinated by Christoph Frei from the Swiss Federal Institute of Technology, ETH, Zürich, Switzerland.

STARDEX PROJECT MEMBERS

UEA	University of East Anglia, UK
KCL	King's College London, UK
FIC	Fundación para la Investigación del Clima, Spain
UNIBE	University of Berne, Switzerland
CNRS-INLN	Centre National de la Recherche Scientifique, France
ARPA-SMR	Servizio Meteorologico Regionale, ARPA-SMR Emilia-Romagna, Italy
ADGB	University of Bologna, Italy
DMI	Danish Meteorological Institute, Denmark
ETH	Swiss Federal Institute of Technology, Switzerland
FTS	Fachhochschule Stuttgart – Hochschule für Technik, Germany
USTUTT-IWS	Institut für Wasserbau, Germany
AUTH	University of Thessaloniki, Greece

D13 AUTHORS AND VERSION HISTORY

Lead author:
Christoph Frei, ETH

Contributing authors:
Stefano Alberghi, ADGB
Maura Amici, ARPA-SMR
Christina Anagnostopoulou, AUTH
András Bárdossy, USTUTT-IWS
Carlo Cacciamani, ARPA-SMR
Clare Goodess, UEA
Malcolm Haylock, UEA
Yeshewatesfa Hundecha, USTUTT-IWS
Panagiotis Maheras, AUTH
Stefano Marchesi, ARPA-SMR
Antonella Morgillo, ARPA-SMR
Valentina Pavan, ARPA-SMR
Guy Plaut, CNRS-INLN
Jürg Schmidli, ETH
Torben Schmith, DMI
Eric Simonnet, CNRS-INLN
Konstantia Tolika, AUTH
Rodica Tomozeiu, ARPA-SMR
Ennio Tosi, ADGB
Rob Wilby, KCL

Version 1.0:	April 2004
Version 2.0:	May 2004
Version 3.0:	January 2005

CONTENTS

See also partner contributions from UEA, CNRS-INLN, ARPA-SMR, ADGB, DMI, ETH, USTUTT-IWS and AUTH, available from the STARDEX web site http://www.cru.uea.ac.uk/projects/stardex/.

1. Introduction

One of the key factors for the reliability of regional climate change scenarios is the accuracy with which General Circulation Models (GCMs) can simulate the responses in the continental to sub-continental scale weather patterns. These are important drivers of the regional surface climate and the occurrence of extreme events. Therefore, parameters of the atmospheric circulation, its temperature and moisture, provide valuable predictors for statistical downscaling of regional climate change from GCMs.

A GCM's skill in reproducing such predictors under present-day climate is an important criterion in assessing the reliability of regional climate change scenarios. Clearly, it is not the only criterion and robustness of downscaling depends on a range of other criteria (see e.g. Wilby et al. 2004). Nevertheless, comparison of GCM skills for a palette of large-scale parameters may be a valuable consideration when choosing predictors for statistical climate change downscaling.

This report on STARDEX Deliverable 13 summarizes results from an evaluation of predictor variables as simulated by the control experiment of the standard GCM used in STARDEX. This evaluation marks one step in the procedure towards identifying robust downscaling methods. A range of different potential predictors are considered, some of which are common to several downscaling methods used across the consortium, others are specific to downscaling methods of individual partners and certain study regions. A brief description of the various statistical downscaling methods in STARDEX, including the pertinent predictors, is given in Schmith and Goodess (2004, see synthesis report of STARDEX deliverable 12).

2. Data and Methods

The GCM considered is the Hadley Centre atmospheric GCM HadAM3P, the successor version of HadAM3H (Pope et al. 2000, Jones et al. 2001, see also Johns et al. 2003). For the present evaluations a simulation is considered which is driven from observed sea-surface temperatures and sea-ice distributions for the period 1961-1990. The evaluations are based on an ensemble of three 30-year integrations. The GCM data (which originally is on a 1.25°x1.875° latitude x longitude grid) was interpolated onto a 2.5 x 2.5 degree grid prior to comparison against the NCEP reanalysis (Akima and Hiroshi 1984).

The NCEP reanalysis (for the 30 years 1961-1990) is used as the 'observation' reference (Kalnay et al. 1996). (There is one exception to this to be explained later.) It should be noted, that the NCEP reanalysis has its own attendant uncertainties (see Reid et al. 2001), and this may influence the present evaluation, at least for some of the parameters (e.g. specific humidity). An alternative reanalysis would have been available in terms of the recently completed ECMWF ERA40. However, the NCEP reanalysis was decided to form the primary reference for STARDEX because of the experience in previous downscaling projects, and because of availability at the start of the project.

The evaluation encompasses, on the one hand, a continental-scale comparison of atmospheric key parameters common to many of the downscaling methods. These are mean sea level pressure (MSLP), geopotential height (Z), temperature (T) and specific humidity (Q). The latter three parameters are considered at levels of 850, 700 and 500 hPa. Beside the long-term seasonal mean fields, we also consider the daily standard deviation of these parameters, which is relevant for downscaling using predictors at the daily time scale. Diagrams for the full set of these key parameters and the whole European area are assembled on the STARDEX web site. Currently: http://www.iac.ethz.ch/staff/freich/download/STARDEX/D13_web/ (is referred to as 'D13-web' hereafter.)

On the other hand, our evaluation comprises several specific comparisons, which were undertaken by individual partners to address the reliability of predictors and the accuracy of predictor-predictand relationships (see e.g. Wilby and Wigley, 2000) in the context of a particular study area and/or downscaling methodology. Table 1 lists these specific analyses and the sections to follow summarize results from the evaluation of both the common key predictors and the region/method specific predictors.

Table 1: List of specific analyses of predictor reliability by STARDEX partners

Partner	*Study region*	*Analysis*
UEA	UK, Europe	Principal components of seasonal MSLP over Europe (60°W-60°E, 20-80°N); explained variance and patterns)
CNRS-INLN	Alps, Alpes Provence and Cote d'Azur	Structure and transition between circulation patterns (Z700). Inter-relationships between circulation patterns and heavy precipitation.
ARPA-SMR	Emilia Romagna	Principal components of seasonal MSLP and Z500 over Europe (comparison of patterns). Blocking frequency.
ADGB	Alps, Northern Italy	Daily statistics of Z500, relative humidity at 700 hPa and geostrophic wind direction at 500 hPa for selected grid points. Pdfs and exceedences of critical thresholds.
DMI	Europe	Vorticity (based on MSLP) for Europe.
ETH	Alpine region	Daily statistics of precipitation from direct GCM output. Comparison against station observations.
USTUTT-IWS	Rhine basin, Germany	Bias and variability of eastward moisture flux at 700 hPa, divergence and vorticity at 850, 700, and 500 hPa. Frequency and persistence of objective circulation patterns based on MSLP.
AUTH	Greece	Frequency statistics and composite patterns of circulation types (based on Z500 and 1000-500 hPa thickness) for the eastern Mediterranean.

3. Main Findings

a) Continental-scale analysis of key parameters

Comparison of the HadAM3P data with the NCEP reanalysis reveals, in general, quite an accurate model representation of the continental-scale mean patterns in key predictors for statistical downscaling. As an example, Figure 1 displays a comparison of MSLP for winter (DJF) and summer (JJA).

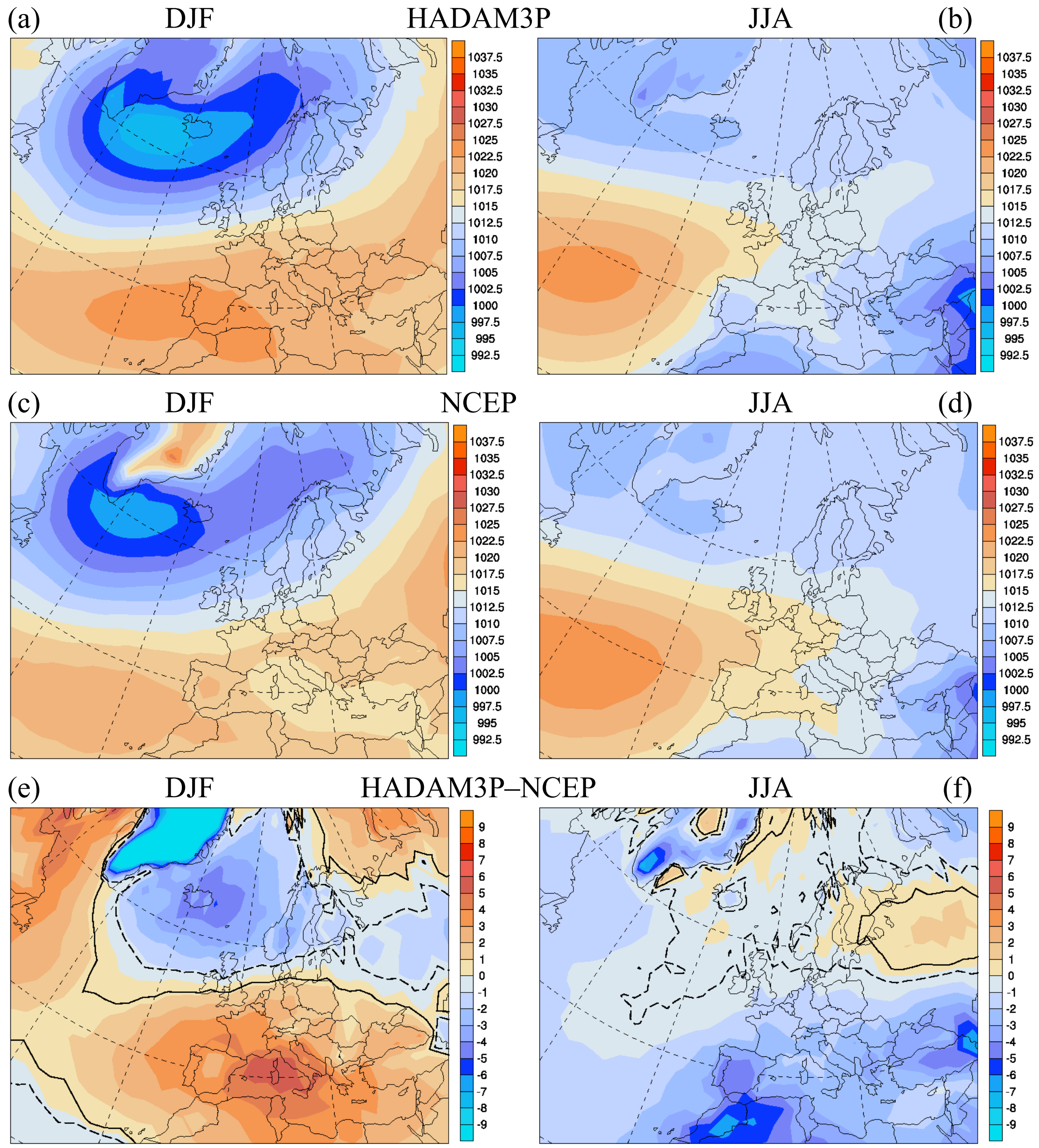

Figure 1: Mean sea level pressure (in hPa) as simulated by HadAM3P (a, b), from the NCEP reanalysis (c, d) and model biases (e, f). Seasonal mean values for 1961-1990 are shown for winter (DJF) and summer (JJA). Black lines in (e, f) depict areas with statistically significant differences (dashed: negative, full: positive) based on a two sample t-test and a significance level of 5%.

In winter, the general features of the simulated pressure fields are in good correspondence with the reanalysis. The North Atlantic mean Low (Icelandic Low) is, however, too deep and the Atlantic subtropical High (Azores High) is too high and has its centre shifted eastward towards the western Mediterranean. MSLP biases are mostly smaller than 5 hPa. (Note that the large bias over Greenland in winter is due to errors in the reanalysis (Reid et al., 2001).) The bias implies too strong low-level westerlies over northwestern Europe (especially the UK and Northern France). A similar bias pattern is found for the geopotential height fields throughout the lower troposphere (850-500 hPa, see D13 web). But a cold temperature bias in the lower troposphere over the western North Atlantic and a warm bias over the continent are also associated with an increase of the geostrophic wind bias with elevation and a slight tilting from westerlies at the surface to south-westerlies at mid-tropospheric levels over the northern parts of the continent.

As regards the daily variability of MSLP, the GCM has slightly too weak activity over most of Europe and the North Atlantic in winter. The standard deviation pattern associated with the North Atlantic storm tracks is reproduced in position and structure quite well but underestimated in magnitude. Over the European continent the underestimate is in the order of 5-10%.

In summer, again, the main continental-scale patterns of the pressure fields are reproduced (see Fig. 1). However, the Azores High is less strong than in the reanalysis and it does not reach as far into the continent and into the western Mediterranean as observed. MSLP is underestimated by 1-3 hPa over most of Europe, except over northern Russia where the bias is slightly positive. In contrast to winter, this pressure pattern is associated with too weak westerlies over northern Europe, but the bias is smaller than in winter.

Tropospheric temperatures in summer (Fig. 2) show quite substantial overestimates over the eastern European continent and the Iberian Peninsula. The bias reaches up to 4 to 5 degrees at the 850 hPa level in these areas. Over the remaining parts of the continent, the temperature bias is smaller but still positive. The temperature bias decreases with elevation. It appears that the STARDEX study regions are not overtly affected by large in-situ temperature biases, but the warm bias, even if in remote areas, may somewhat restrict our trust in low-level summer temperatures. This is even more evident from the standard deviation of daily summer temperatures (Fig. 2), which is overestimated by more than 30% over large areas of the continent up to the 700 hPa level. A similar tendency was noted for HadCM3, the parent atmosphere-ocean GCM of HadAM3P (Collins et al. 2001).

Most likely, the bias in summer temperature variance is related to a common tendency seen in many GCMs (and also regional climate models) to simulate too warm surface temperatures and associated sensible heat fluxes. The reason for this tendency is not fully understood but problems in the soil-atmosphere moisture exchange, convective parameterization and cloud physics have been mentioned as possible reasons (see e.g. Wild et al. 1997, Noguer et al. 1998, Hagemann et al. 2001).

The bias of specific humidity on the 850 hPa level is similar in winter and summer. There is a tendency for too dry conditions over the continent and too wet conditions over the UK and Scandinavia. The magnitude of the bias is about 10% of absolute values. However in summer, areas with too low specific humidity are colocated with too high temperatures and hence both of the biases contribute to a substantial underestimate of lower tropospheric relative humidity

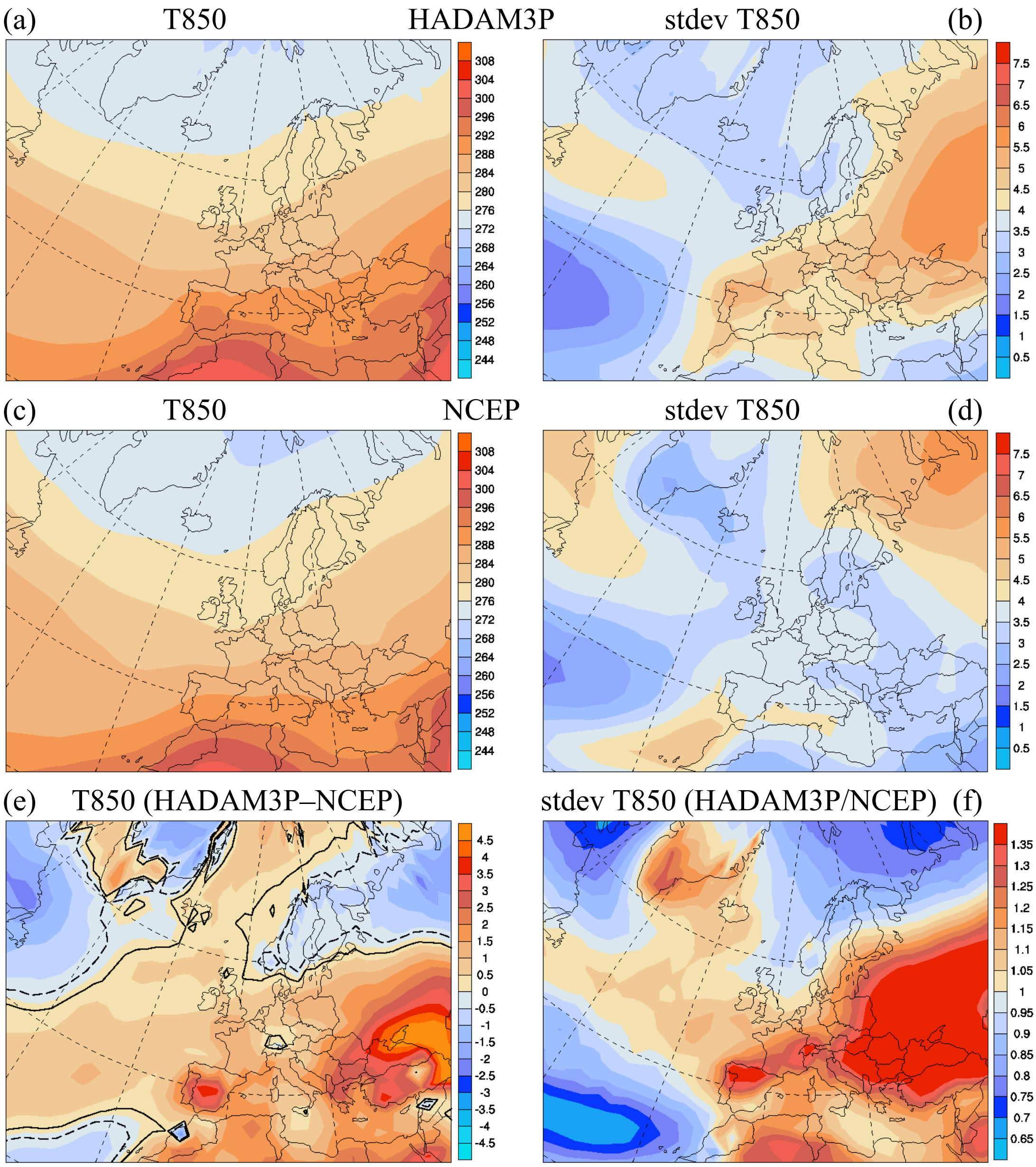

Figure 2: Summer mean T850: (a) HadAM3P, (c) NCEP, (e) bias (HadAM3P-NCEP). Daily standard deviation of summer T850: (b) HadAM3P, (d) NCEP, (f) bias (HadAM3P/NCEP). For 1961-1990. Black lines in (e) depict areas with statistically significant differences (dashed: negative, full: positive) based on a 2 sample t-test and a significance level of 5%.

especially over the Iberian Peninsula and south-eastern Europe. A similar superposition of errors is found for winter over areas of the Mediterranean. A further notable feature of humidity bias (especially at 500 hPa) is the large overestimate (attaining values of more than 40%) over northern Africa. In interpreting these errors it should be noted that the NCEP reanalysis is likely affected by larger observation and assimilation errors in the humidity fields than in the temperature and mass fields.

It is interesting to compare the magnitude and pattern of biases seen in HadAM3P to results from earlier versions of the same model. For example, Jones et al. (1995, see Figures 3 and 5) compare mean sea level pressure and its standard deviation from a mixed layer ocean GCM with the UK Met. Office operational analyses. It is evident that the patterns of these fields deviate much more substantially from observations than in HadAM3P. Again, Noguer et al. (1998) and Johns et al. (1997) show results from a previous atmosphere-only GCM and a coupled model (HadCM2) respectively. In winter, the bias patterns of MSLP were quite different from those in HadAM3P. For example, HadCM2 exhibits too weak westerlies over northern Europe as a result of a too weak Icelandic Low (cf. Figure 10 in Johns et al. 1997, see also Osborn et al. 1999). In summer, the surface pressure bias of the atmosphere-only GCM was positive all over the continent reaching more than 4 hPa in the Northeast (see Figure 3 in Noguer et al. 1998). Although a quantitative comparison with earlier model versions is difficult from the Figures alone (and not necessarily fair when comparing to a coupled GCM), it appears that some of the prominent biases in mean sea level pressure have been improved with HadAM3P.

b) Analyses specific to STARDEX study regions and downscaling methods

Some of the previously mentioned predictor biases have also been noted in the specific analyses undertaken for the STARDEX study regions and some of these analysis have also pointed to issues of predictor reliability that are not directly evident in the continental-scale mean and standard deviation fields. Here, we give a brief summary of the most important results. More detailed discussions can be found in the individual partner reports avaliable from the STARDEX web site.

The downscaling method of ADGB for heavy precipitation in Northern Italy and the Alps consists of a preselection step for rainy days and an amount prediction step. Predictors are taken at selected grid points around the Alps. Predictors based on geopotential height and geostrophic wind were found to reproduce the observed probability density function quite realistically. The effect of systematic biases was compensated for in the downscaling method by defining anomalies with respect to a latitude-time average. More significant model errors were, however, found for relative humidity, which, in accord with the continental-scale analysis, is considerably underestimated at Mediterranean grid points, and results in a bias in the selection process. A remedy to this problem was found in the selection of a more northerly grid point, where relative humidity biases happen to be smaller.

The original intention of DMI was to use daily geostrophic vorticity calculated from mean sea level pressure as a predictor. Inspection of vorticity fields in the reanalysis, however, reveals very noisy patterns especially near Greenland, Scandinavia and the Mediterranean area. Errors in pressure reduction to sea level or in the transformation from the spectral to a grid representation are possible reasons. Anticipating problems with the calibration of the downscaling model with vorticity, it was decided to use mean sea level pressure as the predictor instead.

Statistics of objective circulation patterns (CPs) relevant for the climate in Greece were evaluated by AUTH in an independent classification of HadAM3P and NCEP 500 hPa geopotential and 1000-500 hPa thickness. Both in summer and winter, the frequency of all anticyclonic CPs and all cyclonic CPs together was found to correspond very closely to the values in NCEP. However, the GCM tends to show biases in the prominence of certain cyclonic and anticyclonic patterns. For example, there is an indication that cyclones in the central and eastern Mediterranean tend to travel too far south (i.e. CP 'cyclonic south' occurs too frequently). Nevertheless, composite mean geopotential patterns for individual CPs are very similar to those in NCEP (both in winter and summer). Most composites clearly reflect the mean geopotential bias in the respective season, suggesting that the GCM circulation bias is primarily a mean phenomenon rather than triggered from specific weather situations. The bias in variability, especially the overestimation in summer thickness variance over the Mediterranean region, was noted in an excessive within-CP variance. However, this should not affect the downscaling performance.

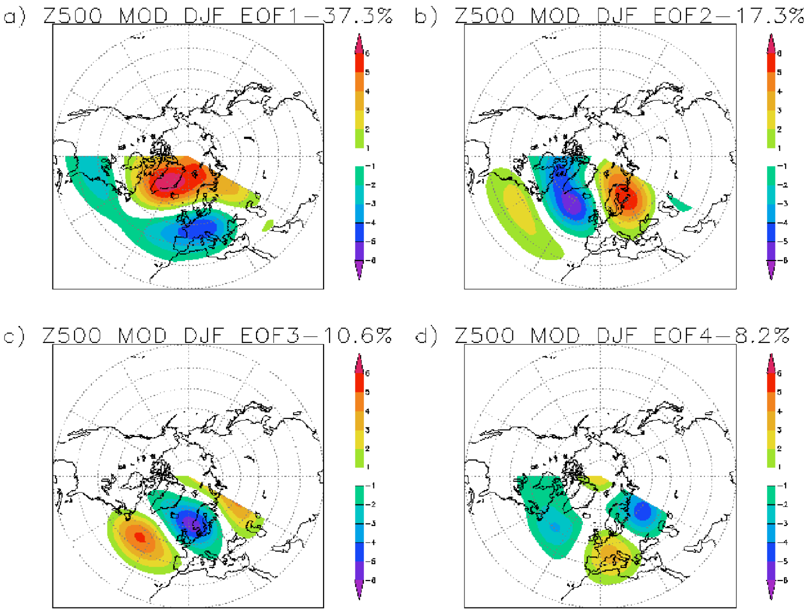

Figure 3: The leading four empirical orthogonal functions of seasonal mean 500 hPa geopotential as simulated by HadAM3P (ensemble member a, winter: DJF).

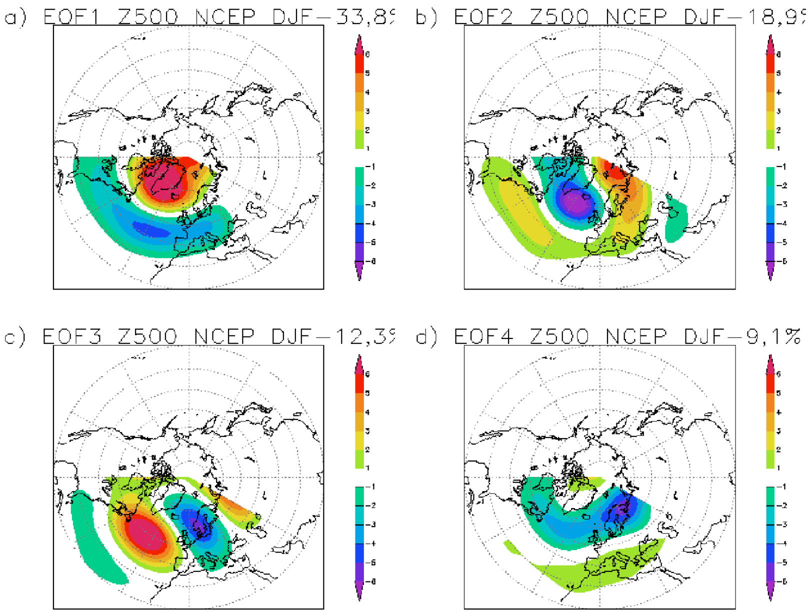

Figure 4: Same as Figure 2 but for the NCEP reanalysis.

The downscaling methods developed by UEA and ARPA-SMR are based on seasonal predictor fields and therefore their analysis has focussed on the spatial structure of internal predictor variability. Both partners have examined the representation of leading principal components by the GCM for the predictors they propose to use (MSLP by UEA and ARPA-SMR; Z500, T850 by ARPA-SMR only). Both analyses were based on large-scale European/Atlantic domains. (Slight modifications in the size of the domain did not affect the results.) While ARPA-SMR has focussed on the phase space representation by standard orthogonal EOFs (see Figs. 2 and 3), non-orthogonal rotated patterns were considered by UEA, but the results of the two approaches are very similar. HadAM3P is capable of reproducing the seasonal variations in the number of significant components and the total variance fraction explained by a few leading principal components (a smaller fraction in summer compared to winter). Also, the loading patterns of the leading modes of variability in the GCM show remarkable resemblance to modes from observations although the relative importance of modes is not always matched. Moreover, pdfs of PC scores correspond to observations within the sampling uncertainty (note that there are 30 seasonal values only in the ensemble member considered), suggesting that there is no obvious sign of variability biases in the GCM. In general, there is a tendency for lower model - observation consistency in summer compared to winter and transition seasons. This is not necessarily a GCM deficiency. It could also be related to the higher dimensionality of the phase space in summer. (An analysis conducted for a second GCM ensemble by UEA supports the latter hypothesis.)

USTUTT-IWS has compared east-west moisture flux, vorticity and divergence between HadAM3P and the NCEP reanalysis. These parameters are considered as potential predictors for statistical downscaling of heavy precipitation in the German Rhine basin. The comparison reveals an overestimation of lower-tropospheric (westerly) moisture flux in winter over North-western Europe and an underestimation over the continent (except Scandinavia) in summer. This is probably attributable to the underestimate of pressure variance (synoptic activity) in winter and the too flat pressure distribution over the Mediterranean region (see subsection a). The day-by-day variability of HadAM3P moisture transport is also underestimated by up to 20% in Southern Europe in winter. Again, for vorticity and divergence (at 850, 700, 500 hPa levels) they find reasonable representations but the accuracy is less good for the summer, where, for example, the day-by-day variability of divergence is overestimated over the Alps and parts of the Mediterranean sea. USTUTT-IWS also envisages using a circulation pattern type approach for downscaling. The fuzzy-rule based classification technique reveals that HadAM3P can reproduce the mean frequency of sea level pressure patterns associated with flooding (amounting to about 20 days per season) with an accuracy of about 2–4 days per season.

The downscaling approach of CNRS-INLN is based on an objective classification of anomaly fields in Z700 (over the North Atlantic and the European continent) using a dynamical cluster algorithm. They find that the cluster centres evaluated from HadAM3P data resemble the patterns determined from NCEP and the frequency of occurrence is similar, although the ranking in occurrence differs. A special investigation was undertaken to check the transition probabilities between the clusters. These showed certain discrepancies in the typical synoptic evolution paths between NCEP and the GCM, which however could also be due to sampling uncertainty. (Results also differ between ensemble members.) A check on the inter-relationship between the circulation types and the occurrence of heavy precipitation in the Alpes-Provence / Cote d'Azur (Southern French Alps) was undertaken by a conditional clustering of Z700 on heavy precipitation days. These results reveal a high similarity of heavy

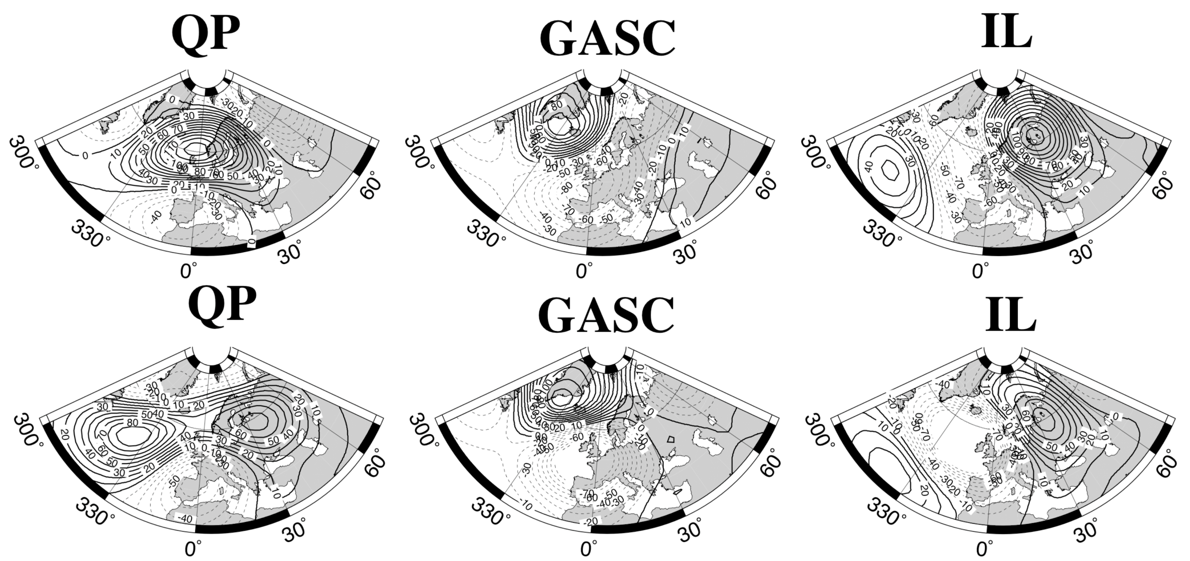

Figure 4: Ojective circulation clusters conditional to heavy precipitation in the Southern French Alps. Top: HadAM3P (ensemble member c). Bottom: NCEP with heavy precipitation events selected from observations.

precipitation related CPs between NCEP and HadAM3P, yet the relative frequency with which these patterns are found on a heavy precipitation day can vary (see Fig. 4).

An unconventional predictor is considered by ETH, who envisages a scaling of biases applied to GCM simulated precipitation. The evaluation in this case is based on a comparison with an appropriately upscaled analysis of daily rain-gauge observations in the Alpine region. It reveals that the European-scale pattern of mean seasonal precipitation by HadAM3P compares reasonably to observations and, in summer, is even closer to observations than NCEP. Similar conclusions are drawn in an evaluation for a range of daily precipitation statistics. The GCM captures the coarse pattern of the spatial distribution for precipitation intensity and frequency. There is no obvious sign that the precipitation of the GCM is severely affected by the circulation errors in the 'free' GCM (compared to the NCEP reanalysis) and by errors in the model's physical parameterisations. Thus there is no evidence so far to be concerned about using HadAM3P precipitation as a predictor. Provided the regional biases in frequency and precipitation are suitably corrected, HadAM3P precipitation is likely carrying elementary information on regional climate change.

4. Summary and Conclusions

This joint study of the representation of predictor variables for statistical downscaling by the Hadley Centre atmospheric GCM HadAM3P reveals, in general, quite promising results. HadAM3P is found to represent quite accurately the main sub-continental scale circulation, temperature and moisture patterns, including their seasonal cycle and the main modes of interannual variability. Comparison with earlier GCM versions suggests that some of the prominent biases in mean sea level pressure have been improved with HadAM3P. Nevertheless, model errors have been identified here, which may influence the representation of surface climate variables when statistical downscaling is performed using the GCM's control time slice.

For example, the too strong westerlies evident in winter over North-Western Europe and the too weak day-to-day variance of surface pressure over the entire continent are likely to influence daily surface temperature and precipitation statistics and the occurrence of extremes. However, the most significant model errors were found in summer for several of the potential predictors. Temperatures are too warm and air masses too dry in the lower troposphere over large parts of Southern Europe. Clearly, these biases are expected to influence the result of downscaling schemes using upper level relative humidity in the predictor set. Recalibration using pertinent thresholds and/or careful choices of model grid points may circumvent these problems but the magnitude of the biases also lowers the general trust in summer moisture as a predictor. Also, the warm and dry bias in summer is associated with too large temperature standard deviations, which are up to 40% larger than observed at 850 hPa. This could influence the representation of extremes in summer and care should be exercised in using this parameter as predictor.

Nevertheless the results of these analyses are not sufficient to suggest that any of the parameters considered should be rejected as a predictor from the beginning. For the purpose of assessing the robustness of downscaling methods, however, the project consortium decided to qualify some of the predictors to have a "compromised reliability" for downscaling from the criterion of GCM biases. A list of these parameters is given in Table 2, together with a brief explanation of the compromising factors. Clearly this is a somewhat subjective qualification and care should be taken in its interpretation. There is no perfect predictor and the magnitude of the bias may not be representative for a GCM's accuracy in simulating the future change of that predictor. Small biases of a GCM predictor does, in itself, not garantee the accuracy of downscaling. Moreover, the reliability into predictors can depend on the exact use in the downscaling model (e.g. whether using single grid points or large-scale patterns) and hence it can vary between downscaling models. Additional criteria will be needed to assess the role of GCM biases in downscaling performance, for example, by comparing results between downscaling from the control experiment and downscaling from NCEP, which may pinpoint to the bias sensitivity of a particular downscaling technique.

5. References

Akima, Hiroshi, 1984: On estimating partial derivatives for bivariate interpolation of scattered data. *Rocky Mountain Journal of Mathematics,* **14,** No. 1.

Collins, M., S.F.B. Tedd and C. Cooper, 2001: The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustment. *Climate Dyn.,* **17,** 61-81.

Hagemann, S., M. Botzet and B. Machenhauer, 2001: The summer drying problem over south-eastern Europe: Sensitivity of the limited area model HIRHAM4 to improvements in physical parameterization and resolutioin. *Phys. Chem. Earth (B),* **26,** 391-396.

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

Johns, T.C., J.M. Gregory, W.J. Ingram, C.E. Johnson, A. Jones, J.A. Lowe, J.F.B. Mitchell, D.L. Roberts, D.M.H. Sexton, D.S. Stevenson, S.F.B. Tett and M.J. Woodage, 2003: Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emission scenarios. *Climate Dyn.,* **20,** 583-612.

Jones, R.G., J.M. Murphy and M. Noguer, 1995: Simulation of climate change over Europe using a nested regional-climate model I: Assessment of control climate, including sensitivity to location of lateral boundaries. *Q. J. R. Meteorol. Soc.,* **121,** 1413-1449.

Jones, R., J. Murphy, D. Hassell and R. Taylor, 2001: Ensemble mean changes in a simulation of the European climate of 2071-2100, using the new Hadley Centre regional climate modelling system HadAM3H/HadRM3H. Hadley Centre Report 2001, available from prudence.dmi.dk.

Kalnay, E., M. Kanamitsu and al. Kistler, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteorol. Soc.,* **77,** 437-471.

Noguer, M., R.G. Jones and J.M. Murphy, 1998: Sources of systematic errors in the climatology of a regional climate model over Europe. *Climate Dyn.,* **14,** 691-712.

Osborn, T.J., K.R. Briffa, S.F.B. Tett, P.D. Jones and R.M. Trigo, 1999: Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model. *Climate Dyn.,* **15,** 685-702.

Pope, D. V, M. Gallani, R. Rowntree and A. Stratton, 2000: The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3. *Climate Dyn.,* **16,** 123-146.

Reid, P.A., P.D. Jones, O. Brown, C.M. Goodess and T.D. Davies, 2001: Assessments of the reliability of NCEP circulation data and relationships with surface climate by direct comparisons with station based data. *Climate Res.,* **17,** 247-261.

Schmith T. and C.M. Goodess: Downscaled extremes based on NCEP reanalysis data (1958-2000). Report of STARDEX deliverable 12. (Available from http://www.cru.uea.ac.uk/projects/stardex/.)

Wilby, R.L. and T.M.L. Wigley, 2000: Precipitation predictors for downscaling: Oberserved and General Circulation Model relationships. *Int. J. Climatol.,* **20,** 641-661.

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. (Available from the IPCC data distribution centre: http://ipcc-ddc.cru.uea.ac.uk/.)

Wild, M., A. Ohmura and U. Cubasch, 1997: GCM simulated surface energy fluxes in climate change experiments. *J. Climate,* **10,** 3093-3110.