STAtistical and Regional dynamical Downscaling of EXtremes for European regions

STARDEX

EVK2-CT-2001-00115

Detailed project report: January 2002 to July 2005

Section 6: Detailed report, related to overall project duration

Co-ordinator: Dr Clare Goodess (c.goodess@uea.ac.uk)
Climatic Research Unit, University of East Anglia,
Norwich, NR4 7TJ, UK

STARDEX home page: http://www.cru.uea.ac.uk/projects/stardex/
MICE/PRUDENCE/STARDEX portal: http://www.cru.uea.ac.uk/projects/mps/

THE STARDEX PARTNERS

1	UEA	University of East Anglia, UK
2	KCL	King's College London, UK
3	FIC	Fundación para la Investigación del Clima, Spain
4	UNIBE	University of Bern, Switzerland
5	CNRS	Centre National de la Recherche Scientifique, France
6	ARPA-SMR	Servizio Meteorologico Regional, ARPA-Emilia Romagna, Italy
7	ADGB	Atmospheric Dynamics Group, University of Bologna, Italy
8	DMI	Danish Meteorological Institute, Denmark
9	ETH	Eidgenössische Technische Hochschule, Switzerland
10	FTS	Fachhochschule Stuttgurt – Hochschule für Technik, Germany
11	USTUTT-IWS	Institut für Wasserbau, Germany
12	AUTH	University of Thessaloniki, Greece

CONTENTS

6.1 BACKGROUND

The climate of the 21[st] century is likely to be significantly different from that of the 20[th] because of anthropogenically-induced climate change. Thus climate change scenarios are needed to assess the impacts of these changes on human lives and activities and the environment. Global climate models (GCMs) provide the starting point for construction of these scenarios.

The mismatch in scales between GCM resolution and the increasingly small scales required by impact analysts can be overcome by downscaling, i.e., 'sensibly projecting the large-scale information on the regional scale' (von Storch *et al*., 1993). Two major approaches to downscaling, statistical (based on the application of relationships identified in the observed climate, between the large-scale and smaller-scale, to climate model output) and dynamical (using physically-based Regional Climate Models (RCMs)) were developed and tested 5-10 years ago by a number of different research groups, and shown to offer good potential for the construction of high-resolution scenarios (Hewitson and Crane, 1996; Wilby *et al*., 1998; Giorgi and Mearns, 1999; Mearns *et al*., 1999; Murphy, 1999; Zorita and von Storch, 1999; Murphy, 2000). In both cases, however, the focus during this period was on changes in mean climate rather than on daily extremes.

Events, such as the August 2002 floods in Central and Eastern Europe and the severe heatwaves experienced across many parts of Europe in August 2003, together with the contrasting extremes of flood and drought experienced across Europe in the summer of 2005, graphically illustrate the losses of life and high economic damages which can be caused by extreme weather events. According to estimates by Munich Re, for example, the August 2002 floods were responsible for economic losses of 21.1 billion Euro and insured losses of 3.4 billion Euro, together with over 100 fatalities (Munich Re, 2002). Events such as this also demonstrate the need for scenarios of weather extremes as well as mean climate – changes in the frequency and intensity of extreme events are likely to have more of an impact on the environment and human activities than changes in mean climate. At the same time there is a need to quantify and, where possible, reduce the uncertainties associated with climate scenarios (Karl *et al*., 1999; Beersma *et al*., 2000; Cramer *et al*., 2000; Meehl *et al*., 2000).

This need for more reliable, high-resolution scenarios of extremes was the major focus of three European Union funded projects running from 2001/2002 to 2004/2005: MICE, PRUDENCE and STARDEX. While PRUDENCE focused on the development and use of RCMs and MICE on the use of GCM and RCM output in impacts studies, STARDEX focused on the development and assessment of improved statistical downscaling methods for Europe with emphasis on the ability to construct scenarios of extremes.

6.2 SCIENTIFIC/TECHNOLOGICAL AND SOCIO-ECONOMIC OBJECTIVES

STARDEX had two general objectives:

- To **rigorously and systematically** inter-compare and evaluate statistical and dynamical downscaling methods for the reconstruction of observed extremes and the construction of scenarios of extremes for selected European regions.

- To identify the **more robust** downscaling techniques and to apply them to provide **reliable and plausible** future scenarios of temperature and precipitation-based extremes for selected European regions.

The phrases highlighted in bold were a major focus of the STARDEX work, and reflect the ambitious nature of the project. Implementing these phrases in practice required a highly co-ordinated scientific approach, together with many hours of technical debate and discussion.

Five measurable project objectives were identified:

i. Development of standard observed and climate model simulated data sets, and a diagnostic software tool for calculating a standard set of extreme event statistics, for use by all partners.

ii. Analysis of recent trends in extremes, and their causes and impacts, over a wide variety of European regions.

iii. Validation of GCM integrations, particularly for extremes.

iv. Inter-comparison of improved statistical and dynamical downscaling methods using data from the second half of the 20th century and identification of the more robust methods.

v. Development of scenarios, particularly for extremes, for the late 21st century using the more robust statistical and dynamical downscaling methods.

These measurable objectives were achieved through these 12 specific objectives:

1. To focus on an agreed, standard set of daily temperature extremes (e.g. percentiles of daily maximum/minimum temperature, frost severity and duration indices and a heatwave duration index) and daily precipitation extremes (e.g. maximum length of dry/wet spells, magnitude of the 90th percentile, percentage of rain falling on days with amounts above the 90th percentile).

2. To focus on specific regions of Europe, selected on the basis of the availability of data and the expertise of the partners, ensuring that the selected case-study regions (Iberian Peninsula, UK, Germany, Alps, Emilia Romagna, Greece) reflect the range of European climatic regimes and that the size/location of each region is appropriate for the extreme being studied.

3. To use a consistent approach (in terms of regions, observed and climate model data inputs, variables and statistics studied and time periods) for all analyses and case studies in order to allow rigorous and systematic evaluation and direct inter-comparison of the results.

4. To analyse observed data series for the second half of the 20th century from specific regions of Europe in order to identify trends in the magnitude and frequency of occurrence of extremes (and, for specific events, their losses in life and financial costs) and to investigate whether these changes are related to changes in other climatic variables (i.e. potential predictor variables derived primarily from NCEP Reanalysis data, such as large-scale and regional objective circulation indices and patterns, including the North Atlantic Oscillation, measures of atmospheric humidity and stability and sea surface temperatures).

5. To analyse output from GCMs, and RCMs driven both by these GCMs and by Reanalysis data, focusing on their ability to simulate temperature and precipitation-based extremes (including their magnitude, frequency of occurrence and trends) and potential predictor variables (including their inter-relationships and relationships with surface climate).

6. To improve existing circulation-based statistical downscaling methods (including methods based on probabilistic weather generators, canonical correlation, multiple regression, neural networks, fuzzy rules and analogue approaches) so that they are able to reproduce observed extremes. This will include the incorporation of additional predictor variables (such as humidity and stability-related parameters and sea surface temperatures) in order to address the problem of stationarity (i.e. the underlying assumption of statistical downscaling that observed large-scale/surface climate relationships remain valid under a changed climate).

7. To calibrate and validate improved 'regional' statistical downscaling methods using predictor variables derived from NCEP (1958-2000) Reanalysis data (for calibration and validation under 'perfect' conditions) and from GCM integrations (that include both anthropogenic and natural forcing) for the present day (in order to assess the effects of climate model biases).

8. To compare the results for specific European regions with output from RCMs driven by the same underlying GCMs, with RCMs driven by Reanalysis data, and with results from a two-step analogue approach to statistical downscaling applied European-wide (for a network of 400-500 stations).

9. To apply the more robust statistical and dynamical [identified in 7 and 8 on the basis of (i) present-day validation studies, (ii) inter-comparison of the scenarios obtained by different methods, and (iii) analysis of the ability of the GCMs/RCMs to reproduce the statistics and inter-relationships of the observed predictor variables] to GCMs and their associated RCMs) integrations for the end of the 21st century to provide scenarios of extreme events and the associated impacts for European regions and to assess the uncertainties associated with these scenarios.

10. To use these scenarios to identify changes in extremes, to investigate whether these changes are in accordance with recent observed changes and to consider their potential impacts in terms of losses of life and financial costs (based on the impacts of observed changes).

11. To ensure that the needs of the European climate impacts community for scenarios of extremes are taken into account, that output from the most recent climate model simulations is available for use in the project and that the work is subject to ongoing peer

review - by bringing together representatives of the stakeholder (e.g. re-insurance), climate modelling and climate impacts communities in an expert advisory panel.

12. To ensure wide dissemination of the project results to stakeholders, the scientific community and the public through the project web site and the production of reports, brochures, information sheets and scientific papers.

6.3 APPLIED METHODOLOGY, SCIENTIFIC ACHIEVEMENTS AND MAIN DELIVERABLES

6.3.1 The STARDEX methodology

The 12 specific objectives of STARDEX (see above) were achieved through five thematic workpackages (Table 1).

Table 1: The STARDEX workpackages

Work package	**Responsible for specific objectives:**	**Responsible for deliverables:**
WP1: Data set development, co-ordination and dissemination	1 to 3, 11 and 12	D1 to D8, D20
WP2: Observational analysis of changes in extremes, their causes and impacts	4	D9, D10
WP3: Analysis of GCM/RCM output and their ability to simulate extremes and predictor variables	5	D11, D13
WP4: Inter-comparison of improved downscaling methods with emphasis on extremes	6 to 8	D12, D14, D15, D16
WP5: Application of the more robust downscaling techniques to provide scenarios of extremes for European regions for the end of the 21st century	9 and 10	D17, D18, D19

In order to rigorously and systematically inter-compare and evaluate statistical and dynamical downscaling methods, a regional case-study approach was followed in all STARDEX workpackages. The case studies were selected from areas in which at least one of the partners had substantial experience. The following six regions were selected:

- Iberian Peninsula
- Greece
- Alps
- German Rhine
- UK
- Northern Italy (Emilia Romagna).

In addition, work was undertaken using a European-wide data set of observed daily maximum and minimum temperature and precipitation for 495 stations for the period 1958-2000 developed by Fundación para la Investigación del Clima (FIC). The STARDEX FIC dataset has good spatial coverage over most of Europe (Figure 1). Station data were provided by the European Climate Assessment (ECA) project (http://eca.knmi.nl) and by national meteorological services from 14 European countries. Quality control analyses of the daily temperature and precipitation time series values were undertaken by FIC and an iterative homogenisation procedure based on the approach of Moberg and Alexandersson (1997) was also applied to the mean annual maximum and minimum temperature series. Due to the relatively sparse distribution of the stations, detected inhomogeneities were not corrected, but flagged as suspect. In addition to the quality control work undertaken by FIC, considerable work on quality control and homogenisation was previously undertaken by the ECA on all temperature and precipitation series in their data set (Wijngaard *et al*., 2003), including those subsequently incorporated in the STARDEX dataset.

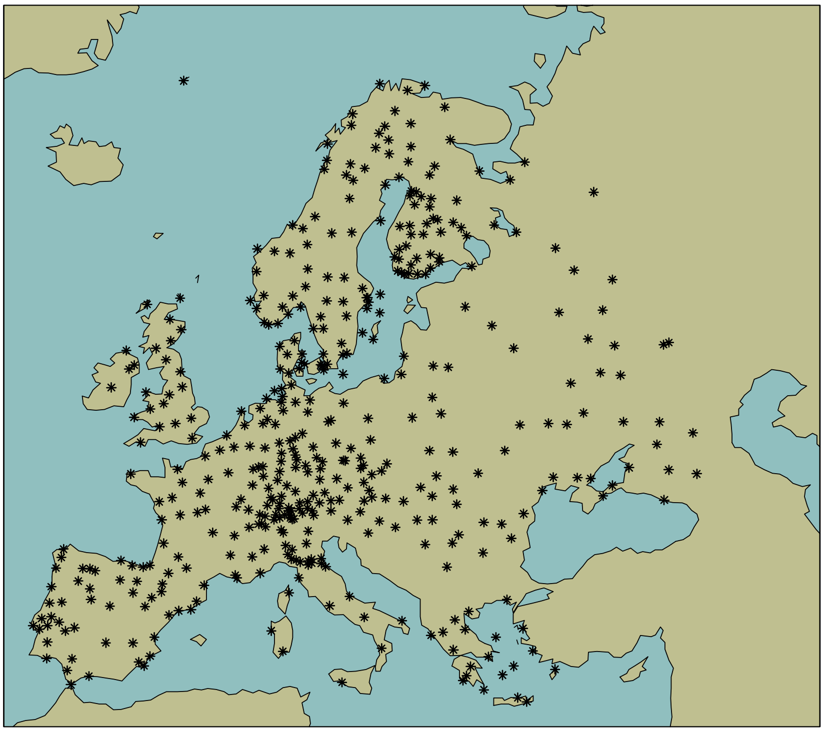

Figure 1: Location of the 495 stations in the STARDEX/FIC European-wide dataset

The European-wide data set was complemented by higher-density datasets produced by the relevant STARDEX partners for each of the six case-study regions (Table 2). For each of these regions, a subset of stations drawn from the FIC European-wide data set was also identified and used, for example, to intercompare locally- and European-wide developed downscaling methods as part of the workpackage 4 work.

Considerable effort was put into constructing standard observed and modelled datasets and making them available for use by all STARDEX partners (deliverables D3 to D7). All datasets not subject to restrictions by third parties have been made publicly available from the STARDEX web site (http://www.cru.uea.ac.uk/projects/stardex). Although the FIC European-wide dataset is subject to such restrictions, time series for 10 core indices of extremes (see below) are available from the STARDEX website (http://www.uea.ac.uk/projects/stardex).

All STARDEX statistical downscaling models were calibrated and validated using predictor variables, including sea level pressure (SLP), 500 hPa geopotential height, 1000-500 hPa thickness field and relative/specific humidity and temperature at different pressure levels, derived from the National Centers for Environmental Prediction (NCEP) reanalyses (Kalnay *et al*., 1996). This reanalysis data set was considered the most appropriate for use in STARDEX, having the major advantages of comparable spatial resolution to the current generation of GCMs and spanning the 40-year period for which suitable daily observed data are available.

In order to construct climate change scenarios, statistical predictor-predictand relationships derived between Reanalysis and observed surface data were applied to output from the Hadley Centre high-resolution atmospheric GCM HadAM3P (Pope *et al*., 2000). The Hadley Centre model was chosen as the focus of STARDEX work for consistency with the PRUDENCE and MICE projects (although most of the RCM simulations performed in the PRUDENCE project were nested within a slightly earlier version of the Hadley Centre model - HadAM3H).

Table 2: Summary of data available for the STARDEX regional case-study regions

Region	***Number of stations in the full regional dataset (and STARDEX group providing the data)***	***Number of stations in the regional subset***	***Stations in the regional subset***
Iberian Peninsula:			
Western Iberia	11 (UEA)	11	Portugal: Beja, Coimbra, Lisboa Geofisica, Santarem, Pegoes, Alvega, Mora, Penhas Douradas, Portalegre Spain: Badajoz/Talavera, Alcuescar
SE Spain	5 (UEA)	5	Albacete/Los Llano., Valencia, Alicante Ciudad Ja., Murcia/Alcantarilla, Murcia/San Javier
Greece:	22 (AUTH)	4 Western Greece	Agrinio, Ioannina, Kalamata, Kerkyra
		4 Eastern Greece	Alexandroupoli, Mytilini, Samos, Rodos
Alps:	N Alps: 27 grid points, Ticino: 15 grid points, 0.5º precipitation grid, Frei and Schär, 1998 (ETH) 21 temperature (UNIBE)	10	Austria: Innsbruck-Univ. France: Nice, Montelimar Germany: Muenchen Italy: Bologna, Lazzaro Alberoni, Bobbio Switzerland: Arosa, Locarno-Monti, Zuerich
German Rhine:	100 (USTUTT-IWS)	10	Feldberg/Schw., Karlsruhe, Mannheim, Deuselbach, Koeln-Wahn, Giessen, Wuertzburg, Saarbruecken-E., Kahler Asten, Nuernberg-Kra.
UK:			
NW UK	15 precipitation (UEA)	3	Eskdalemuir, Ringway, Shawbury
SE England	28 precipitation (UEA)	3	Cambridge, Goudhurst, Oxford
Northern Italy: (Emilia Romagna)	39 temperature (ARPA) 59 precipitation (ARPA)	8	Bobbio, Lazzaro Alberoni, Bedonia, Bologna, Alfonsine, Parma, Firenzuola, Verghereto

HadAM3P data were available for an ensemble of three 30-year members using the IPCC SRES (Nakicenov and Swart, 2000) A2 emissions scenario and one B2 scenario member. Control-period output was available for 1961-1990, while the scenario period was 2071-2100. STARDEX statistically-downscaled results were compared with ‘perfect-boundary condition’ RCM simulations (i.e., simulations forced with ERA-15 reanalyses) performed as part of the earlier EU-funded MERCURE project (Machenhauer *et al*., 1998) and HadAM3P/HadAM3H-forced RCM climate-change simulations performed during the PRUDENCE project (Christensen *et al*., 2005).

During the early stages of the STARDEX project, considerable effort was put into defining a set of indices of extremes to be the focus of work. Different indices are relevant for different applications and impacts, and a list of 54 temperature and precipitation based indices was produced by STARDEX partners. These indices are calculated by the STARDEX extremes indices software developed by KCL and UEA. This software comprises two elements: a Fortran subroutine that calculates all the indices for a single location; and a program that uses this subroutine to process station data in a standard input format. This software (deliverable D8), together with user documentation, is available from the STARDEX web site). As well as being an essential common tool for STARDEX, it is being widely used by other research groups and students in Europe and beyond.

From the 54 indices calculated by the software package, a set of 10 core indices of extremes (six for precipitation and four for temperature) was identified as the focus of STARDEX work, together with three mean indices (Table 3). Many of the indices are based on thresholds defined using percentile values rather than fixed values. This makes them transferable across the range of climatic regimes experienced across Europe. However, such 'fixed-bin' approaches do have some limitations, e.g., when exploring the contribution of extreme events to overall trends (Michaels *et al*., 2004). In order to ensure reasonable sample sizes and to avoid major difficulties in trend analysis (Frei and Schär, 2001), the focus is on 'moderate' extremes, i.e., 90th and 10th percentile values, rather than the far tails of the distributions. The core set was carefully chosen to encompass magnitude (e.g., Tmax 90th percentile), frequency (e.g., number of days with precipitation exceeding the 90th percentile) and persistence (e.g., longest dry spell length and heat wave duration) of extremes.

Table 3: The STARDEX 10 core indices of extremes and three mean indices

Precipitation related indices of extremes	
pq90	90th percentile of rainday amounts (mm/day)
px5d	Greatest 5-day total rainfall
pint	Simple daily intensity (rain per rainday)
pxcdd	Maximum number of consecutive dry days
pfl90	% of total rainfall from events > long-term 90th percentile
pnl90	Number of events > long-term 90th percentile of raindays
Temperature related indices of extremes	
txq90	Tmax 90th percentile (ºC)
tnq10	Tmin 10th percentile (ºC)
tnfd	Number of frost days Tmin < 0 °C
txhw90	Heat wave duration (days)
Mean indices	
pav	Precipitation average (mm/day)
txav	Average Tmax (ºC)
tnav	Average Tmin (ºC)

In order to ensure a rigorous and systematic approach to validation and intercomparison of the STARDEX statistical downscaling methods, a set of principles for this work was agreed. All partners used data from common predictand and predictor datasets. The core set of 10 indices of extremes was used, together with mean daily precipitation and mean maximum and minimum temperature – giving 13 indices in total (Table 3). A common verification or independent validation period was also chosen: 1979-1993, for compatibility with the ‘perfect-boundary condition’, i.e., ERA-15 forced, RCM simulations undertaken in the MERCURE project (Machenhauer *et al*., 1998). The remaining period of data, 1958-1978 and 1994-2000 was used for model calibration or training. A common set of verification statistics for the comparison of observed and downscaled annual series of seasonal indices was identified: Root Mean Square Error (RMSE) between the observed and simulated index with the bias removed; Spearman rank-correlation coefficient (CORR); and BIAS (the mean difference between the simulated and observed indices).

The STARDEX statistical downscaling methods (see Table 4) use a range of approaches suitable for different applications. Some provide multi-site information, for example. The latter methods are more applicable to denser regional station networks than European wide. Thus it was decided that it would be inappropriate to use a single case-study for verifying and intercomparing all the methods. Instead, the matrix shown in Table 5 was devised. Three groups (UEA, FIC and DMI) applied their methods European-wide, i.e., to the FIC dataset of 495 stations (Figure 1). The other nine groups undertook initial development of their method(s) in the region with which they were most familiar, e.g., ETH from Zurich, Switzerland initially developed their scaling methods for the Alps. In addition to this ‘primary’ region, eight of the nine groups also applied their method(s) in a ‘secondary’ region with a contrasting climatic regime, e.g., the UK in the case of ETH.

Table 5: The case-study regions in which STARDEX statistical downscaling methods were applied. x = method applied to the European-wide data set (Figure 1). 1 = partner’s primary region, 2 = partner’s secondary region.

Group_Method(s)	***Europe***	***Iberian Peninsula***	***Greece***	***Alps***	***German Rhine***	***NW and SE UK***	***Northern Italy***
UEA_CCA	x						
FIC_ANAL2	x						
DMI_CWG	x						
UEA_CCA and ANN		2				1	
KCL_ANN and CR		2				1	
UNIBE_CCA				1			
CNRS_PPCI		2		1			
ARPA-SMR_CCA and MLR			2				1
ETH_DYN and LOC				1		2	
USTUTT/FTS_MAR and MLR				2	1		
AUTH_ANN, CCA and MREG			1				2

This case-study approach allows a number of different intercomparisons to be undertaken by sampling the matrix shown in Table 5 either horizontally or vertically. Downscaling methods from six different groups can be directly compared in the case of Alpine precipitation, for example. While different experimental approaches could have been used, e.g., a single case-study region, this is unlikely to have altered the main findings of the STARDEX work (see Section 6.4).

Table 4: Summary of the STARDEX improved statistical downscaling methods

Method	***Predictand(s)*** ***(Unless otherwise indicated, predictands are station series)***	***Predictor(s)*** ***(See STARDEX Deliverable D10 for selection procedure)***	***Description*** ***(See STARDEX Deliverable D15 for details)***
ADGB_HYPER4	Regional DP index	GPH anomalies at 500 hPa, RH at 700 hPa, geostrophic wind at 500 hPa & precipitable water	Random sampling within the 4-dimensional hyperspace of the 4 predictors which defines conditions for high precipitation
ARPA_CCA	PIE, TIE	SLP, SH at 1000, 950, 850 and 700 hPa, and T at 850 hPa	Canonical Correlation Analysis
ARPA_MLR	PIE, TIE	Z500: First 4 PCs of 500 hPa GPH anomalies T850: First 4 PCs of 850 hPa T	Multiple Linear Regression
AUTH_ANN	DP, DT	500 hPa GPH & 1000-500 hPa thickness	Artificial Neural Network
AUTH_CCA	DP, DT	500 hPa GPH & 1000-500 hPa thickness	Canonical Correlation Analysis
AUTH_MREG	DP, DT	Circulation types for 500 hPa, 1000-500 hPa thickness	Multiple Linear Regression
CNRS_PPCI	DP	Large Scale Circulation patterns defined using 700 hPa GPH	Random selection of an analogue within a set of training days having the same 'Potential Precipitation Circulation Index' category
DMI_CWG	DP	SLP	Conditional weather generator, conditional on quantile values of a circulation index, in which precipitation occurrence and amount are modelled separately
ETH_DYN	DP – station data or mesoscale grids	Grid-box precipitation	As ETH_LOC, but with flow-dependent scaling factors
ETH_DYNI	DP – station data or mesoscale grids	Grid-box precipitation	As ETH_LOCI, but with flow-dependent scaling factors
ETH_LOC	DP – station data or mesoscale grids	Grid-box precipitation	Local scaling of GCM simulated precipitation
ETH_LOCI	DP – station data or mesoscale grids	Grid-box precipitation	Local scaling of GCM simulated precipitation with correction of precipitation frequency and intensity bias
FIC_ANAL2	DP, DT	Geostrophic fluxes at 1000 & 500 hPa,	Two-step analogue method, in which (1) the 'n' most similar

		low tropospheric humidity and thickness	days to the day being simulated are selected from a reference data set and (2) regression is performed using predictand/predictor relationships from the 'n' days data set
KCL_ANN_GA_RBF	DP	The SDSM set of predictors	Genetic algorithm used to optimise the Radial Basis Function network structure and parameters
KCL_ANN_IRBF	DP	The SDSM set of predictors	Individual Radial Basis Function artificial neural network model (i.e., applied to individual sites in each region)
KCL_ANN_MLP	DP	The SDSM set of predictors	Multi Layer Perceptron artificial neural network model
KCL_ANN_RBF	DP	The SDSM set of predictors	Radial Basis Function artificial neural network model (applied across all sites for each region)
KCL_CR	DP	The SDSM set of predictors	Conditional resampling of area average precipitation, conditional on the large-scale atmospheric forcing and a stochastic error term, and daily precipitation amounts at a 'marker site' (generated using SDSM).
UEA_ANN_GAMMA	DP	The SDSM set of predictors	Bayesian multilayer perceptron artificial neural networks, using the hybrid Bernoulli/Gamma data misfit term
UEA_ANN_GAMMAMC	DP	The SDSM set of predictors	Bayesian multilayer perceptron artificial neural networks, using the hybrid Bernoulli/Gamma data misfit term and Monte-Carlo simulation
UEA_ANN_SSE	DP	The SDSM set of predictors	Bayesian multilayer perceptron artificial neural networks, using the sum-of-squares data misfit term
UEA_CCA	PIE	CCA1: MSLP CCA4: MSLP + GPH, RH, T at 500, 700 & 850 hPa	Canonical Correlation Analysis
UNIBE_CCA	DT	SLP and GPH, T, SH & RH at 100, 850, 700, 500 and 300 hPa	Canonical Correlation Analysis
USTUTT_MAR	DP	Objective circulation patterns and: - eastward moisture flux at 700 hPa (for precipitation) - GPH at pressure level corresponding the circulation pattern	Multivariate Auto-Regressive model
USTUTT_MLR	PIE, TIE	GPH, RH, T, divergence and vorticity at several levels, eastward moisture	Multiple Linear Regression

		flux at 700 hpa level and objective circulation patterns	

DP = daily precipitation
DT = daily temperature
PIE = STARDEX core indices of precipitation extremes
TIE = STARDEX core indices of temperature extremes

GPH = Geopotential height
MSLP = Mean sea level pressure
PC = Principal Component
RH = Relative humidity
SDSM = Statistical DownScaling Model (Wilby *et al*., 2002)
SH = Specific humidity
T = Temperature

6.3.2 The STARDEX deliverables

STARDEX has produced 20 deliverables, which are summarised in Table 6.

Table 6: The STARDEX deliverables

No.	***Deliverable***	***Format***	***Outline***	***Contributing partners***	***Journal papers (published, in press, submitted)***
D1	Members' web site	Restricted web site	Provides access to restricted data sets and internal project documents	Set up by UEA – includes material from all partners.	
D2	Public web site	Public web site	http://www.cru.uea.ac.uk/projects/stardex/ - provides access to all STARDEX public deliverables (including software deliverable D8, report deliverables D9 to D13, D15, D6 and D18 to D20 and data deliverables D14 and D17).	Set up by UEA – includes material from all partners.	
D3	Standard data sets of daily temperature and precipitation time series for selected European regions and for 400-500 locations across Europe	Data sets (mainly restricted use)	10 core indices of extremes for 495 European stations available from http://www.cru.uea.ac.uk/cru/projects/stardex/	European-wide data set: FIC. Regional data sets: all partners.	Pavan *et al.,* 2005a
D4	Standard data set of NCEP Reanalysis data	Public data sets	http://www.cru.uea.ac.uk/cru/data/ncep/	UEA	Flocas *et al*., 2005; Tolika *et al*., 2005a
D5	Standard data set of objective circulation patterns and indices	Data sets (some restricted use)	See links at: http://www.cru.uea.ac.uk/cru/projects/stardex/, and STARDEX members' web site	ARPA-SMR, UEA, FTS, AUTH	Bárdossy *et al*., 2002
D6	Standard data set of GCM and RCM output	Data sets (restricted use)	Data sets from the MERCURE and PRUDENCE projects, including calculated indices of extremes, available from STARDEX members' web site	UEA, ETH	
D7	Standard data set of damages (losses of life and economic) arising	Data sets	See links at: http://www.cru.uea.ac.uk/cru/projects/stardex/	UEA, FTS	

	from specific extreme events				
D8	Diagnostic software tool for calculating a standard set of extreme event statistics	Software tool and user documentation	Calculates 54 indices of temperature and precipitation extremes. Available from http://www.cru.uea.ac.uk/cru/projects/stardex/, along with user documentation	KCL, UEA	
D9	Trends in extreme daily precipitation and temperature across Europe in the 2nd half of the 20th century, and their causes	Synthesis report and partner reports. Journal papers	Uses a dataset of nearly 500 European daily station series together with denser station networks for seven European regions, to analyse trends in extreme daily precipitation and temperature over the 2nd half of the 20th century. 10 indices of extremes are presented, calculated using the STARDEX software. The use of a common set of indices and the large number of station series used, makes this the most comprehensive and consistent study to date of European temperature and precipitation extremes.	Lead : USTUTT-IWS. ARPA-SMR, AUTH, CNRS, ETH, FTS, UEA, UNIBE	Anagnostopoulou *et al.*, 2003; 2004; Haylock and Goodess, 2004; Hundecha and Bárdossy, 2005; Maheras *et al.*, 2003; 2004; Pavan *et al.*, 2005b; Schmidli and Frei, 2005; Tomozeiu *et al.*, 2005
D10	Recommendations on methodologies for identification of the best predictor variables for extreme events	Synthesis report and partner reports	The STARDEX work described in the partner contributions indicates that the best predictor variables vary with region, season and predictand. Thus the synthesis report focuses on methodologies for identification of the best predictor variables and makes recommendations as to how the process can be carried out as objectively as possible.	Lead: UEA. ARPA-SMR, AUTH, CNRS, DMI, FIC, KCL, USTUTT-IWS	
D11	Recommendations on variables and extremes for which downscaling is required	Synthesis report and partner reports	This work differs in several respects from conventional GCM evaluations in order to provide specific insights into the need for downscaling. In particular, the NCEP reanalysis is considered as a quasi-ideal GCM. The focus is on temperature and precipitation extremes and their interannual variability. The extensive STARDEX data sets of daily stations series are used – suitably upscaled to provide a fair comparison at the scale of the GCM. Three mutually-related measures are considered: standard correlation coefficients; ratio of variance; and debiased root mean square error – displayed using Taylor diagrams. The results provide guidance on the need for downscaling and	Lead: ETH. ADGB, ARPA-SMR, AUTH, CNRS, FIC, UEA, USTUTT-IWS	

			its dependence on season, type of extremes and region.		
D12	Evaluation of downscaled extremes based on NCEP Reanalysis data	Synthesis report, regional and partner reports. Journal papers	Provides a detailed intercomparison of more than 20 improved statistical downscaling methodologies. In order to evaluate the ability of these methods to reproduce present-day indices of extremes, NCEP Reanalysis data is used to provide the predictor variables. A common set of verification principles is used to ensure that the intercomparison is fair and consistent. From the initial intercomparison exercise, it was not possible to identify the best methods. Thus more detailed regional analyses were undertaken using a larger set of stations. Several key messages and two 'headline' conclusions are drawn from the regional analyses (see Goodess *et al*., 2005).	Lead DMI/UEA. All partners	Busuioc *et* al., 2005; Goodess *et* al., 2005; Harpham and Wilby, 2004; 2005; Harpham *et* al., 2005; Haylock *et al*., 2005; Kostopoulou *et al*., 2005; Schmidli *et al*., 2005; Tolika *et al*., 2005b; Wilby *et al*., 2003
D13	Recommendations on the most reliable predictor variables and evaluation of inter-relationships	Synthesis report and partner reports. Web-based figures	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.	Lead: ETH. ADGB, ARPA-SMR, AUTH, CNRS, DMI, UEA, USTUTT-IWS	
D14	Downscaled extremes based on NCEP and GCM output for the present day	Data sets and documentation	Seasonal indices (10 core indices of extremes and 3 averages – see Table 3) for the STARDEX study regions (Table 5) and Europe as a whole, calculated using NCEP reanalyses and HadAM3P control-run data (1961-1990). For some indirect downscaling methods, daily time series are also available. All available from the central data server (accessed from the STARDEX web site) along with user documentation and metadata.	All partners. Data group led by ETH	
D15	Improved statistical downscaling methodologies	Report	Describes the 20 plus statistical downscaling methods used in the STARDEX project (multiple linear regression, canonical correlation analysis, artificial neural networks, multivariate autoregressive modelling, conditional resampling and other analogue-based methods, methods	Lead UEA. All partners	

			based on a ‘potential precipitation circulation index’ and ‘critical circulation patterns’, conditional weather generator and local and dynamical scaling) in sufficient detail to allow them to be reproduced by the reader.		
D16	Recommendations on the more robust statistical, dynamical and/or statistical-dynamical downscaling methods for the construction of scenarios of extremes	Synthesis report and partner tables	The synthesis report lists the STARDEX robustness criteria, then suggests appropriate statistical tests or analysis methods for evaluating each criterion, and finally gives examples of STARDEX applications (i.e., referring to other deliverables or papers). The format of the STARDEX application and performance criteria tables is discussed, and key recommendations and points of good practice made on a number of topics, including: planning downscaling studies; calibration and validation of statistical downscaling models; multi-model approaches; and application of statistical downscaling models.	Lead UEA. All partners	
D17	Downscaled scenarios based on GCM output for the end of the 21st century for regions of Europe and 400-500 locations across Europe	Data sets and documentation	Seasonal indices (10 core indices of extremes and 3 averages – see Table 3) for the STARDEX study regions (Table 5) and Europe as a whole, calculated using HadAM3P output for the 2071-2100 scenario period, for the IPCC SRES A2 (three ensemble members) and B2 (one ensemble member) emissions scenarios. For some indirect downscaling methods, daily time series are also available. All available from the central data server (accessed from the STARDEX web site) along with user documentation and metadata.	All partners. Data group led by ETH	
D18	Summary of changes in extremes, comparison with past changes and consideration of impacts/damages	Information sheets. Supporting figures/tables. Journal papers.	Overview information sheet on ‘What methods can be used to make projections about future changes in the occurrence of extreme weather events in Europe?’, and regional information sheets (e.g., ‘How will the occurrence of extreme rainfall events in the UK change by the end of the 21st century?’), with sections on extremes in the region, past changes in extremes, future changes in extremes, references and further reading and contacts. The web site provides more graphical information than can be included in the information sheets or journal papers (i.e.,	Lead UEA: All partners	Beniston *et al., 2005;* Frei *et al*., 2005; Haylock *et al*., 2005 Others in preparation for Greece, Germany, the Alps.

			comprehensive figures and tables showing the scenario changes in each region for the 10 core indices of extremes).		
D19	Assessment of uncertainties associated with the scenarios	Information sheet. Journal papers.	Overview information sheet and a table listing the various aspects of uncertainty (i.e., uncertaines in emissions scenarios; inter-model uncertainty; intra-model uncertainty; natural variability; and, in particular, downscaling method) with references to appropriate STARDEX, PRUDENCE and MICE reports and papers. Uncertainty issues are also discussed in the D18 information sheets and reports.	Lead UEA: All partners	Wilby and Harris, 2005 See also: Beniston *et al., 2005;* Frei *et al*., 2005; Haylock *et* al., 2005
D20	Final project report	Glossy hard-copy report (~10 pages)	Provides answers to the following 11 questions, based on the STARDEX work: What type of extremes can we make statements about?; Why is downscaling needed?; What downscaling methods are available?; Have changes in extremes occurred over the last 40 years?; Can these changes be related to changes in circulation and other aspects of the atmosphere?; Can these relationships be used to project changes in extremes?; Will extremes become more frequent and/or intense in the future?; Do these changes matter?; How confident can we be in these projections?; What is the most appropriate method of downscaling for my impacts study?; What further research is needed?	Lead UEA: All partners	

The STARDEX deliverables were produced with different principle user groups and audiences in mind (Figure 2). Thus the D3 to D8 data sets and tools, provide the key inputs to the technical reports D9 to D13, which each consist of detailed partner reports, together with a synthesis report. These reports, together with their underlying data sets, are synthesised further in D16. This is a central output of STARDEX. The D16 summary report provides guidelines for those wishing to undertake their own downscaling, building on the IPCC guidelines (Wilby *et al*., 2004; Mearns *et al*., 2003). The checklist of good practice provided here is also likely to be valuable for impacts scientists and other users who want to assess the robustness and reliability of downscaling results from other sources. The D16 partner contributions provide summary documentation concerning the robustness and reliability of the specific statistical downscaling method(s) developed and evaluated by each STARDEX group. This documentation will help users to identify the most appropriate of the STARDEX downscaling methods for their purpose, but is also directed at people wanting to use the scenario results presented in D18 – which are available from the STARDEX central data archive – for quantitative impacts modelling. For other users, the maps and figures on the web site will be more valuable, while more general needs for non-technical information are met by the D18 and D19 information sheets, and the D20 report.

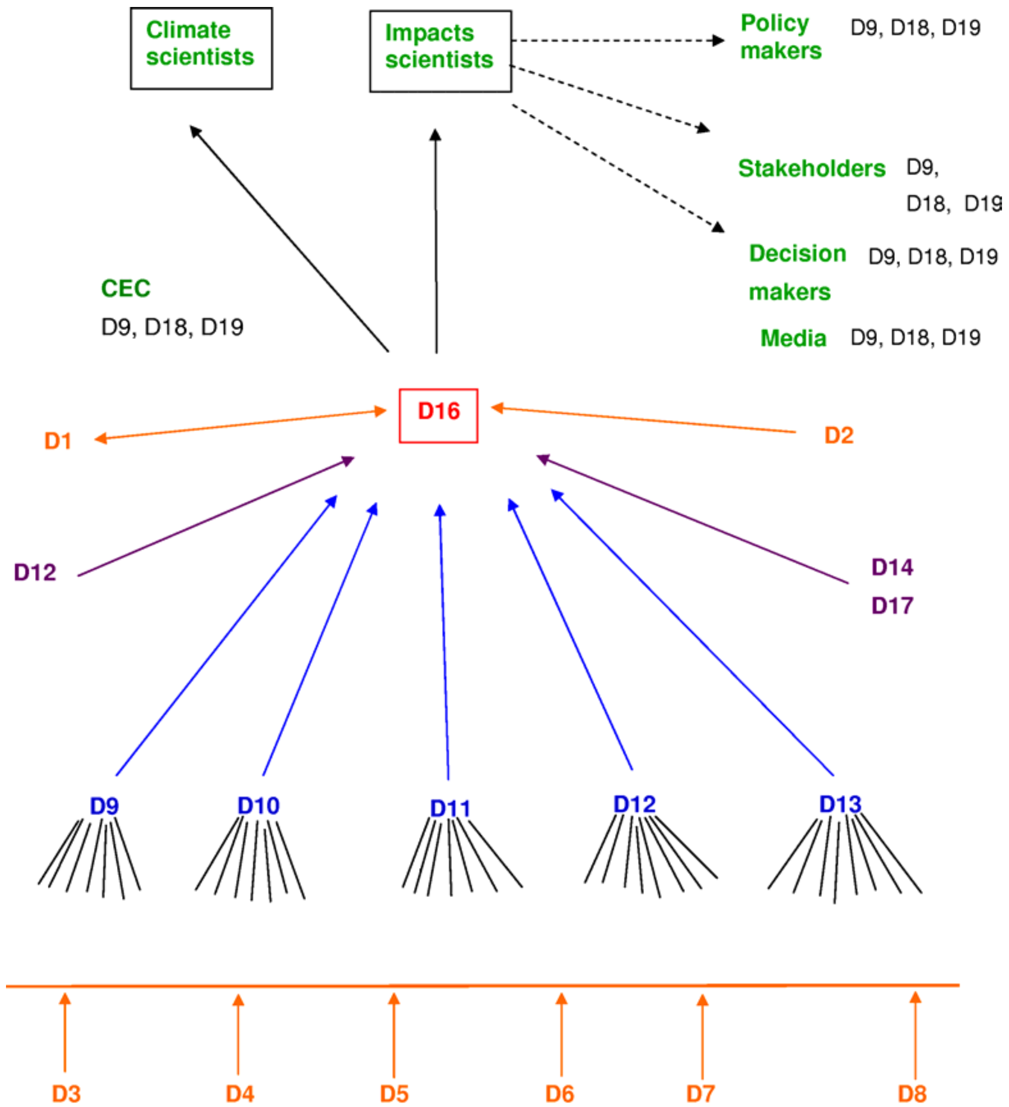

Figure 2: Schematic diagram of the STARDEX deliverables

6.3.3 STARDEX scientific achievements

Assembling the STARDEX data set deliverables D3 to D7 was time consuming, but this work provided a firm basis for the scientific work which is described in detail in the many STARDEX report deliverables (see Section 6.3.2 and Table 6). Here, selected findings from each stage of the STARDEX work are highlighted, in order to illustrate the major scientific achievements of the project. This also demonstrates how each stage of the work built upon the preceding stage(s), thus emphasising the added value of the integrated and co-ordinated approach taken by STARDEX.

Extensive analysis of observed data, supported by an understanding of the underlying physical processes, provides a sound basis for the construction of statistical downscaling methods

In statistical downscaling, relationships between larger-scale climate variables and local surface climate variables, derived from observed data, are applied to climate model output, based on two assumptions: first, that the larger-scale variables are more reliably simulated by the climate models than the local variables; and, second, that the relationships will remain valid in a changed climate (the assumption of stationarity). Since statistical downscaling is based on observed relationships, the starting point for STARDEX was the analysis of past changes in extremes and exploration of their causes, i.e., links with large-scale circulation, focusing on potential predictor variables.

Spatially coherent changes in extremes have occurred over the last 40 years (see deliverable D9 and associated journal papers)

Analysis of changes in the core indices of extremes (Table 3) over the last 40 years was carried out by STARDEX for Europe as a whole and for the case-study regions (Table 7).

Table 7: Data sets used for analysis of changes in the core indices of extremes

Region	Partner	Number of precipitation stations	Number of temperature stations
Europe	UEA	495	495
UK (SE and NW England)	UEA	40	21
German Rhine	USTUTT-IWS	611	232
Northern Italy (Emilia Romagna)	ARPA-SMR	62	44
Greece	AUTH	22	22
Switzerland	ETH	104	-
Switzerland	UNIBE	-	21
French Alps	CNRS	3 plus gridded data	3

Results for the case-study regions are summarised in Table 8. The temperature indices indicate a general shift to warmer conditions, with more extremes. This general tendency is also evident in the European-wide analysis (see for example, Figures 3 and 4 which show trends in winter frost days and summer heatwave duration respectively). The European-wide changes in temperature extremes can be summarised as follows:

Winter:
- Extreme maximum temperature increased over most of the region except the southeast
- Extreme minimum temperature increased over the entire region, apart from small decreases in parts of Greece, the Iberian Peninsula and Scandinavia
- Extreme minimum temperature increased to a larger degree than the corresponding extreme maximum temperature

Summer:
- Extreme maximum temperature increased in most areas of the region, except in northern Scandinavia, Eastern Europe and Russia
- Extreme minimum temperature increased in most areas, except that a few stations showed a decrease.

Table 8: Summary of analysis of changes in extremes for the STARDEX case-study regions
+ = positive/increasing trend, - = negative/decreasing trend
++ and - - indicate strongest, most significant trends

Region	**Winter**	**Spring**	**Summer**	**Autumn**
Extreme maximum temperature indices				
UK (SE and NW England)	+	+	+	+
German Rhine	+	+	+	
Northern Italy (Emilia Romagna)	+	+	+	+
Greece	+	+	+	+
Switzerland	++	+	+	+/-
French Alps	+	+	++	
Extreme heavy precipitation indices				
UK (SE and NW England)	++		--	
German Rhine	++	+	--	+
Northern Italy (Emilia Romagna)	-	-	+	+
Greece	-	-		
Switzerland	++	+	+	++
French Alps	Variable	Variable	Variable	+

The precipitation trends are more spatially and seasonally variable. The case-study results indicate trends towards more extreme precipitation in all seasons in Switzerland, for example, contrasting with negative trends in winter and spring and positive trends in summer and autumn in Emilia Romagna to the southeast. For the UK and the German Rhine, there is a contrast between the strong winter trend to more extreme precipitation and the summer trend towards drier conditions. These spatial complexities are revealed further in the European-wide analyses (see, for example, Figure 5 which shows changes in the number of heavy rainfall days in summer). The European-wide precipitation changes can be summarised as follows:

Winter dry-day persistence:
- Increased in the southern part of the region
- Decreased in the north
- The increase is generally greater than the decrease

Summer dry-day persistence:
- Less coherent signal

- Most stations in the central part, the UK, and southern Scandinavia showed an increase

Winter heavy precipitation extremes:
- Increased in central Europe, the UK and Scandinavia
- Decreased in Eastern Europe, Greece and western Iberian Peninsula

Summer heavy precipitation extremes:
- Increased across northern Scandinavia and Russia
- Decreased across the UK and NE Europe
- Increased across SW Europe
- Decreased across the northern Iberian Peninsula

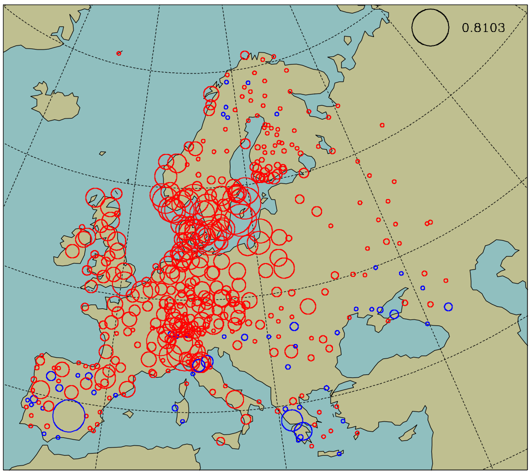

Figure 3: 1958-2000 trend in winter frost days (tnfd). Scale is days per year. Red is decreasing.

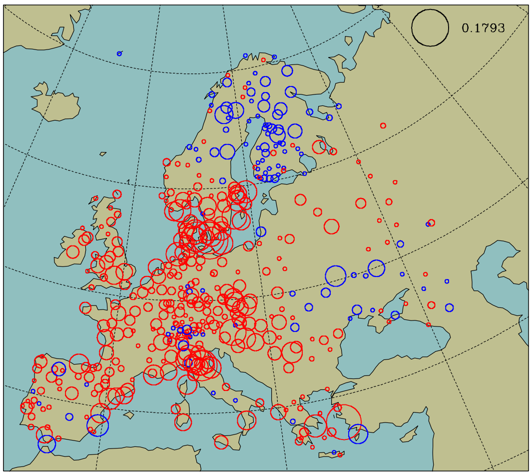

Figure 4: 1958-2000 trend in summer heatwave duration (txhw90). Scale is days per Year. Red is increasing.

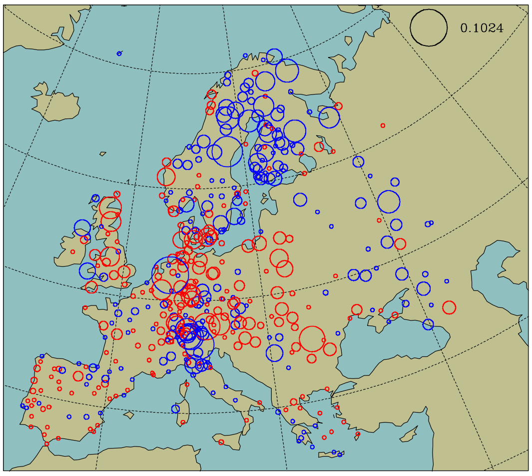

Figure 5: 1958-2000 trend in summer heavy rainfall days (pnl90). Scale is days per year. Blue is increasing.

In part, these observed changes in extremes can be explained by changes in large-scale circulation and other predictors (see deliverable D9 and supporting papers)

As well as analysing past changes in extremes, STARDEX investigated relationships between these changes and changes in large-scale circulation, focusing on potential predictor variables for statistical downscaling. A few examples of the relationships found are presented below.

Although extreme rainfall is generally not as spatially coherent as mean rainfall, Haylock and Goodess (2004), show that spatial analysis of the interannual variability of extremes can reveal interesting behaviour in the indices. In particular, they used principal component (PC) analysis and canonical correlation analysis (CCA) to explore relationships between the North Atlantic Oscillation (NAO) and European-wide trends in the number of winter heavy rain days (index pnl90). From this analysis (Figure 6), Haylock and Goodess (2004) conclude that:

> "A CCA of the indices with MSLP has revealed that the NAO is an important influence on extreme rainfall. The similarity between canonical patterns of the indices and the linear trends in the indices suggests that it is mainly changes in the NAO that have caused the observed changes in these indices".

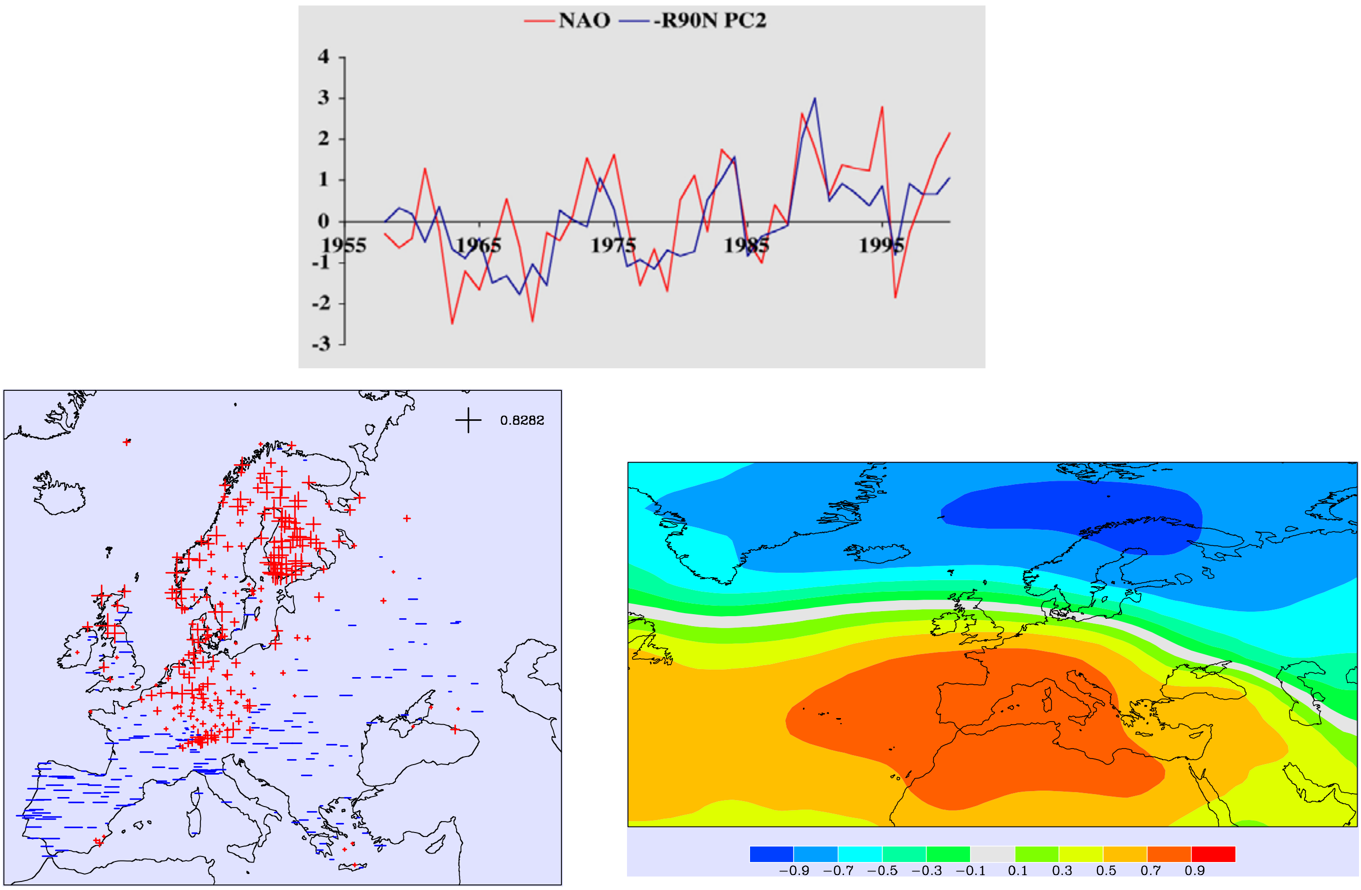

Figure 6: Correlation between the NAO and the second principal component of the number of winter heavy rainfall days (top). First canonical pattern of winter heavy rainfall days (bottom left: red + indicates a positive trend/relationship, blue - a negative trend/relationship) and the associated sea level pressure patterns (bottom right).

The second example presented here, concerns links between severe winter storms and river flooding in southwest Germany which were explored by FTS. The well-known West cyclonic (Wz) circulation pattern (CP) is very similar to the objectively defined "critical" CP11 (Figure 7). These patterns are two of the most critical identified for causing severe winter storms in Europe and river flooding in southwest Europe. The "critical" CP11 has increased significantly in frequency and persistence (Figure 8). Consistent with this, the risk of an 'extreme zonal winter' within the 'critical sector' (i.e., winters with more than 35% Wz days and a maximum persistence of more than 13 days) has increased dramatically by a factor of 23 in the period 1982/83 to 2003/04 compared to the period 1881/82 to 1981/82. Five of the winters in the 15-year period 1989/90 to 2003/04 fell within the 'critical sector' and caused economic losses due to floods and winter storms of 40 billion US$ (Table 9).

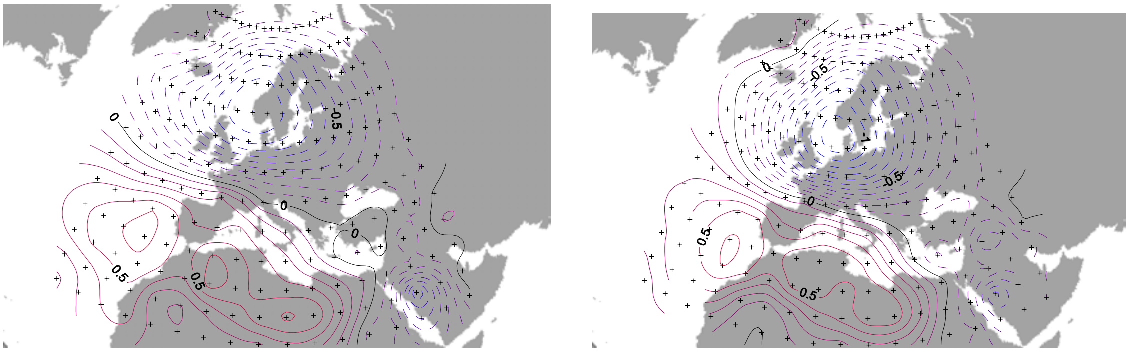

Figure 7: Sea level pressure anomalies of the subjectively classified CP Wz (left-hand side) and the objectively classified CP11 (right-hand side). CP11 is classified using runoff data for the Moselle river.

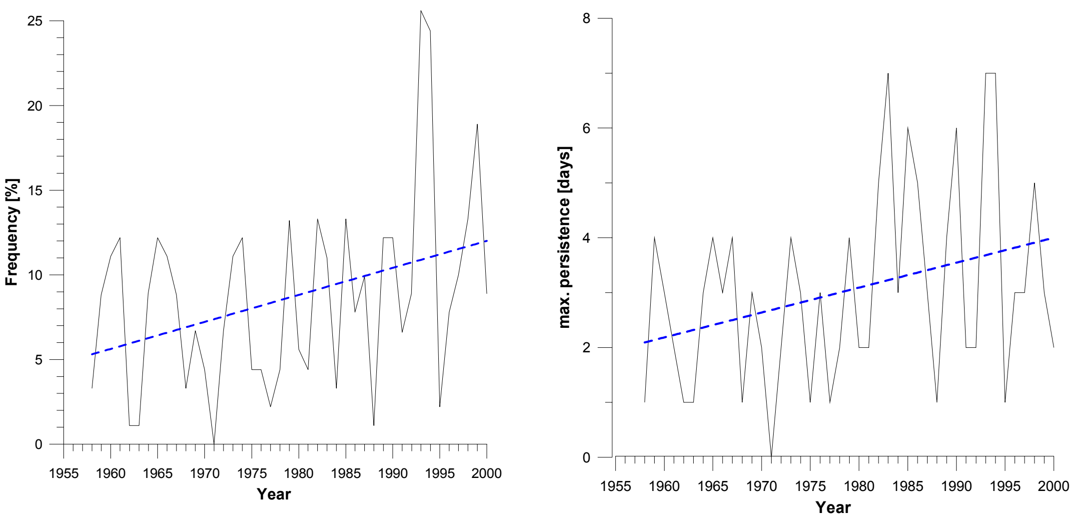

Figure 8: Critical CP11 in winter for the period 1958 to 2001 – frequency (left-hand side) and persistence (right-hand side).

Table 9: Winters in the 'critical sector' (see text), major floods and severe winter storms across western and central Europe, together with economic losses.

Winter	Event	Date	Country/Region/ River Basin	Circulation Type Wz Frequency f [%]	Circulation Type Wz max. Persistence T_{max} [days]	Economic losses [Mio. US $]	Insured losses [Mio. US $]	Fatali-ties
1989/90	Winterstorms: "Daria", "Vivian", "Wiebke"	25.01.-01.03.1990	Western Europe GB, D, F, Benelux	35,6	13	14 800	10 200	230
	River Flood	25.- 26.02.1990	Upper Danube					
1993/94	Christmasflood	20.12.-31.12.1993	Rhine, Moselle, Nahe Neckar, Enz, Kocher	43,3	17	2 000	800	14
1994/95	River Floods	19.01.-03.02.1995	Rhine, Nahe, Main, Moselle, Lahn, Benelux, Upper Danube	45,5	13	3 500	910	28
1997/98	Winterstorms	23.12.-05.01.1998	GB, IR, D, F	36,7	16	650		15
1999/2000	Winterstorms "Anatol"	03.-04.12.1999	DK, D, GB, S, PL	40,7	19	18 500	10 750	230
	"Lothar"	26/12/1999	F, D, CH, B, A					
	"Martin"	27/12/1999	F, ES, CH					
					Σ =	39 450		

While such relationships between extreme events and circulation patterns provide a sound basis for statistical downscaling, it is important to incorporate additional predictor variables, particularly humidity-based variables, into downscaling models which are going to be used for climate-change applications. Figure 9, for example, illustrates that the probability of precipitation is highly dependent on moisture flux as well as CP.

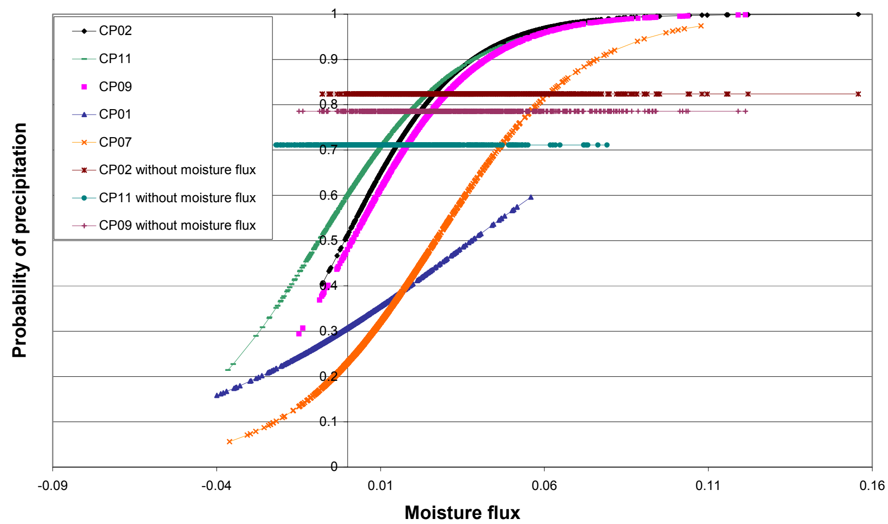

Figure 9: Probability of precipitation at a station in the German Rhine conditional on moisture flux across the region and circulation pattern.

The final example presented here, is precipitation/weather regimes identified for extreme precipitation in the French Alpes Maritimes by CNRS. Figure 10 shows the Greenland Anticyclone Sole Cyclone. This figure demonstrates that this pattern has many desirable characteristics for statistical downscaling:

- It is highly discriminating, i.e., days which most closely resemble this pattern have a very high probability of intense rainfall
- It is stable, i.e., the patterns and relationships are very similar for the two periods
- The relationships are physically realistic, i.e., consistent with the synoptic pattern
- It is associated with known extreme events, e.g., the October 12 1979 Nice airport landslide and Antibes tsunami.

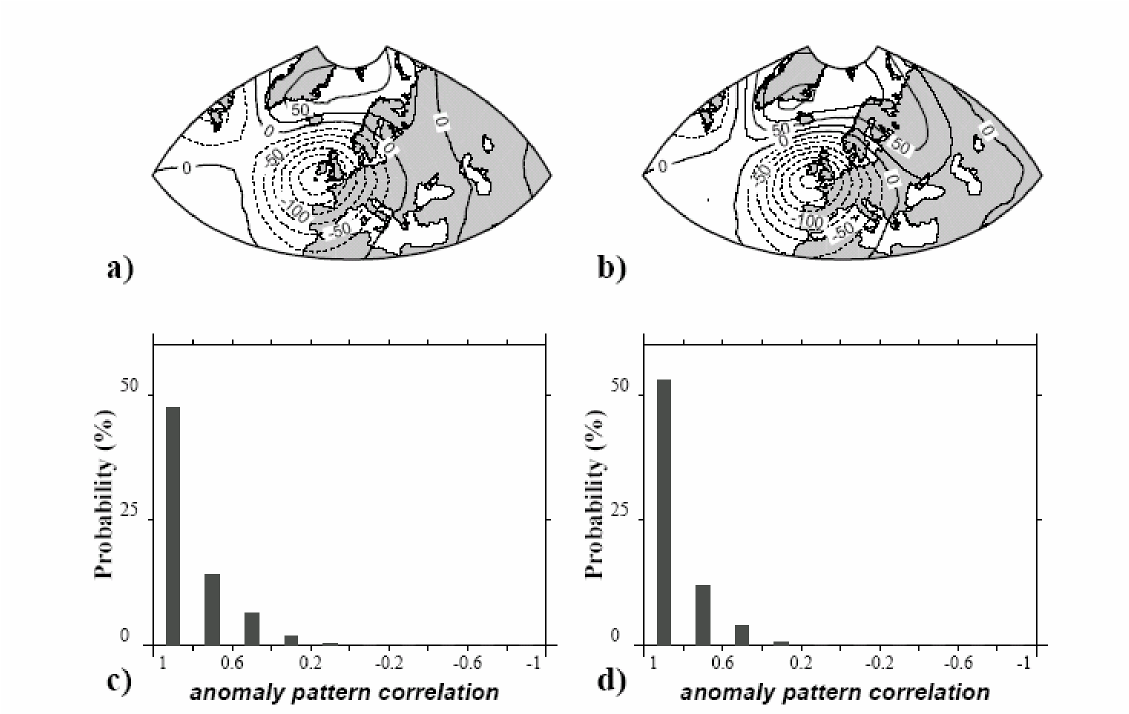

Figure 10: Circulation composites (upper panel) and relationships with extreme rainfall in the French Alpes Maritimes (lower panel) for the Greenland Anticylone Sole Cyclone for 1971-1983 (left-hand side) and 1983-1995 (right-hand side).

When identifying the best predictor variables, it is easier to make recommendations about methodologies than standard predictor sets (see deliverable D10)

Good predictor variables are defined in STARDEX Deliverable D10 as:

- having strong, robust and physically-meaningful relationships with the predictand;
- having stable and stationary relationships with the predictand;
- explaining low-frequency variability and trends;
- being at an appropriate spatial scale (in terms of both physical processes and GCM performance); and
- well reproduced by GCMs.

A number of different methods were used to select the most appropriate predictor variables for use in each STARDEX study region including: stepwise multiple regression, compositing, correlation analysis, principal components analysis and CCA. These more traditional methods proved more useful than novel methods such as a genetic algorithm approach. Automated methods for predictor selection are generally less suitable and there tends to be a need for user intervention and local expertise. Variability from season-to-season, region-to-region, and extreme-to-extreme makes it difficult to make recommendations about the specific predictors which should be used in any particular study. STARDEX has, however, been able to produce lists of potentially useful predictors (see deliverable D10) from which the most appropriate for a particular application can be selected using the recommended methodologies.

In general, the predictor variables identified for use in the STARDEX statistical downscaling models are well simulated by the GCM (HadAM3P) chosen as the basis for STARDEX work (see deliverable D13)

STARDEX deliverable D13 summarises results from an evaluation of predictor variables as simulated by the control experiment of HadAM3P (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 were 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. NCEP reanalysis was used as the observational reference. Figure 11, for example, compares the leading four empirical orthogonal functions of winter mean 500 hPa geopotential calculated by ARPA-SMR using HadAM3P and NCEP reanalysis. It can be seen that both sets of patterns are very similar.

HadAM3P was 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 were identified, 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 Northwestern 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 results of downscaling schemes using upper level relative humidity in the predictor set. 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 a predictor.

Nevertheless, the results of these analyses were not sufficient to suggest that any of the parameters considered by STARDEX should be rejected as a predictor from the beginning, although some should be used with caution. 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 not, in itself, guarantee the accuracy of downscaling. Moreover, the reliability of 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 may vary between downscaling models.

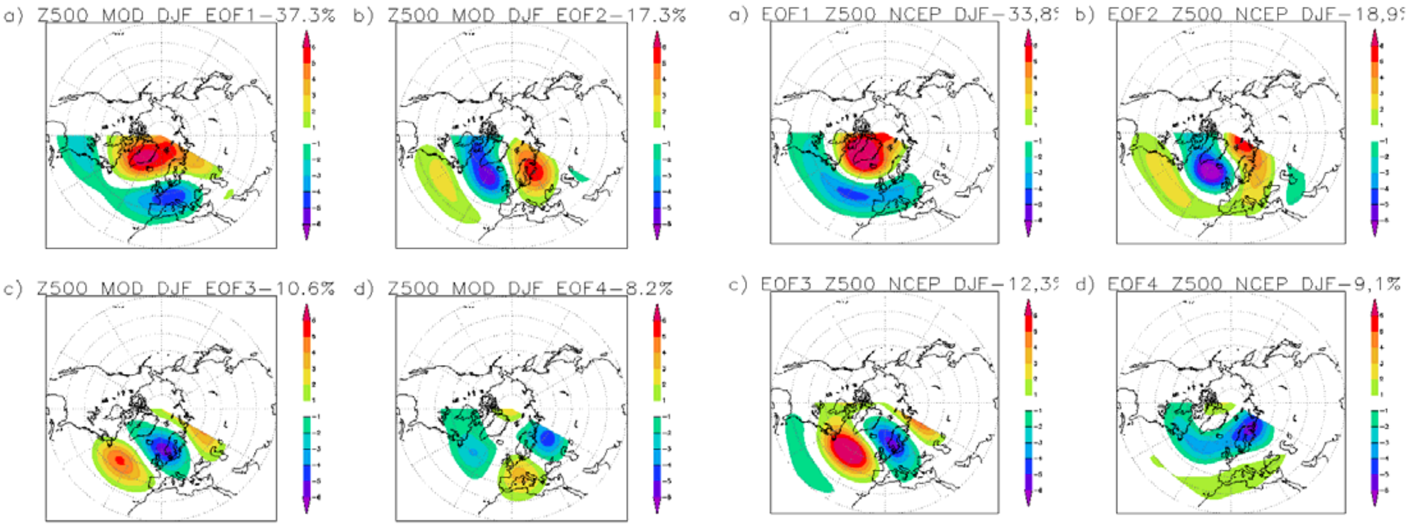

Figure 11: The leading four empirical orthogonal functions of winter mean 500 hPa geopotential as simulated by HadAM3P (left-hand side) and NCEP reanalysis (right-hand side)

STARDEX developed and evaluated a range of statistical downscaling methods of differing complexity (see deliverables D12, D15, D16 and supporting papers)

The STARDEX statistical downscaling methods (Table 4) range from standard linear regression methods (MLR, used by four STARDEX groups), through methods focusing on spatial patterns (CCA, also used by four STARDEX groups), to non-linear neural network methods (ANN) and other less-widely used approaches including some analogue-based methods. Three STARDEX groups evaluated a number of different ANN approaches. Prior to STARDEX, there had been relatively few examples of ANNs used for downscaling and even fewer applications to downscaling multi-site precipitation extremes. Thus the STARDEX study provides the most systematic evaluation of ANN methods for the latter purpose to date.

Handling many combinations of statistical downscaling methods (20+), regions (7), indices (13) and seasons (4), made intercomparison complex (see deliverable D12 and supporting papers)

Having set up a rigorous experimental approach (see Section 6.3.1), suitable approaches for handling the many combinations of different methods (22 – see Table 4), regions (seven – see Table 5), indices (13 – see Table 3) and seasons (four) had to be devised.

The experimental matrix shown in Table 5 allowed a number of specific questions to be addressed:

- Is there any systematic difference in performance of the methods between different seasons?

- Is there any systematic difference in performance of the methods between different indices?

- Is there any systematic difference in performance of the methods between different regions?

- Do direct methods in which the seasonal indices of extremes are downscaled perform better than indirect methods in which daily time series are generated and the seasonal indices then calculated from these?

- Do the regionally-developed methods perform better than the European-wide methods?

- Can a single 'best' method be identified?

These questions were addressed by undertaking a series of regional validation studies, using NCEP reanalysis as predictors, which are reported in detail in deliverable D12 and supporting papers. A preliminary inspection of the downscaled results for all regions indicated that the variation in skill from station-to-station dominates the variations in skill from index-to-index, from method-to-method and from season-to-season. This is demonstrated in Figure 12, which shows box-and-whisker plots of the Spearman correlation skill for a number of different downscaling methods applied to UK precipitation. Thus in order to address the questions above, results were averaged across neighbouring stations, as well as across the different indices, seasons and methods, as appropriate for addressing each question. While this was a pragmatic approach, designed to draw general conclusions from a large amount of downscaled data, it did not preclude more detailed analyses using different averaging methods or individual results.

A number of key messages emerged from the reanalysis-based validation studies (see deliverable D12, Goodess et al., 2005 and other supporting papers):

- Skill varies from station-to-station (in particular), season-to-season, index-to-index and method-to-method (Figure 12)

- But not systematically, which makes it hard to pick a single best method in most cases

- Methods/indices with the highest correlations are often not those with the lowest bias/root mean square error (Figure 13)

- Performance is generally better for temperature than precipitation, better for means than extremes, and best in winter and worst in summer (Figures 12 and 14)

- However, there are always exceptions to the rules, for example, in Greece, the poorest precipitation results are for autumn

- The FIC_ANAL2 European-wide method performs well for temperature (Figure 14), as well as or better than locally-developed methods

- CCA methods seem to perform better when applied locally rather than European wide

- The performance of ANN methods is generally quite good, particularly with respect to precipitation correlations (e.g., for the Iberian Peninsula) which reflect inter-annual skill

- It is particularly difficult to make statements about whether 'direct' or 'indirect' methods are consistently better for downscaling indices of extremes

- For precipitation extremes, persistence, notably the length of the longest dry spell, is better represented than magnitude or frequency characteristics (Figure 12).

Two 'headline' conclusions emerged from the reanalysis-based validation studies (see deliverable D12 and Goodess et al., 2005):

- In the majority of cases, no consistently superior statistical downscaling model can be identified, so a major recommendation is to use a range of the better statistical downscaling methods – just as the recommendation is to use a range of GCMs/RCMs

- For many regions and indices, the skill is unacceptably low for summer rainfall – thus scenarios should not be constructed for these cases

The implications of these headline conclusions are discussed in Section 6.4.

STARDEX also compared reanalysis-driven statistical and dynamical downscaling (Haylock *et al*., 2005; Schmidli *et al*., in preparation). Figure 12, for example, compares six statistical downscaling methods with two perfect-boundary condition RCMs for SE England. For this region and NW England, it was concluded that the ANN methods perform better than the other statistical downscaling methods and the RCMs. A similar comparison was undertaken for the Alps (Schmidli *et al*., in preparation), but in this case, a more direct comparison was possible, as the statistical downscaling was applied to grid-point data. It was shown that the better statistical downscaling models tend to have smaller biases than the RCMs, but that all statistical downscaling models strongly underestimate the interannual variability.

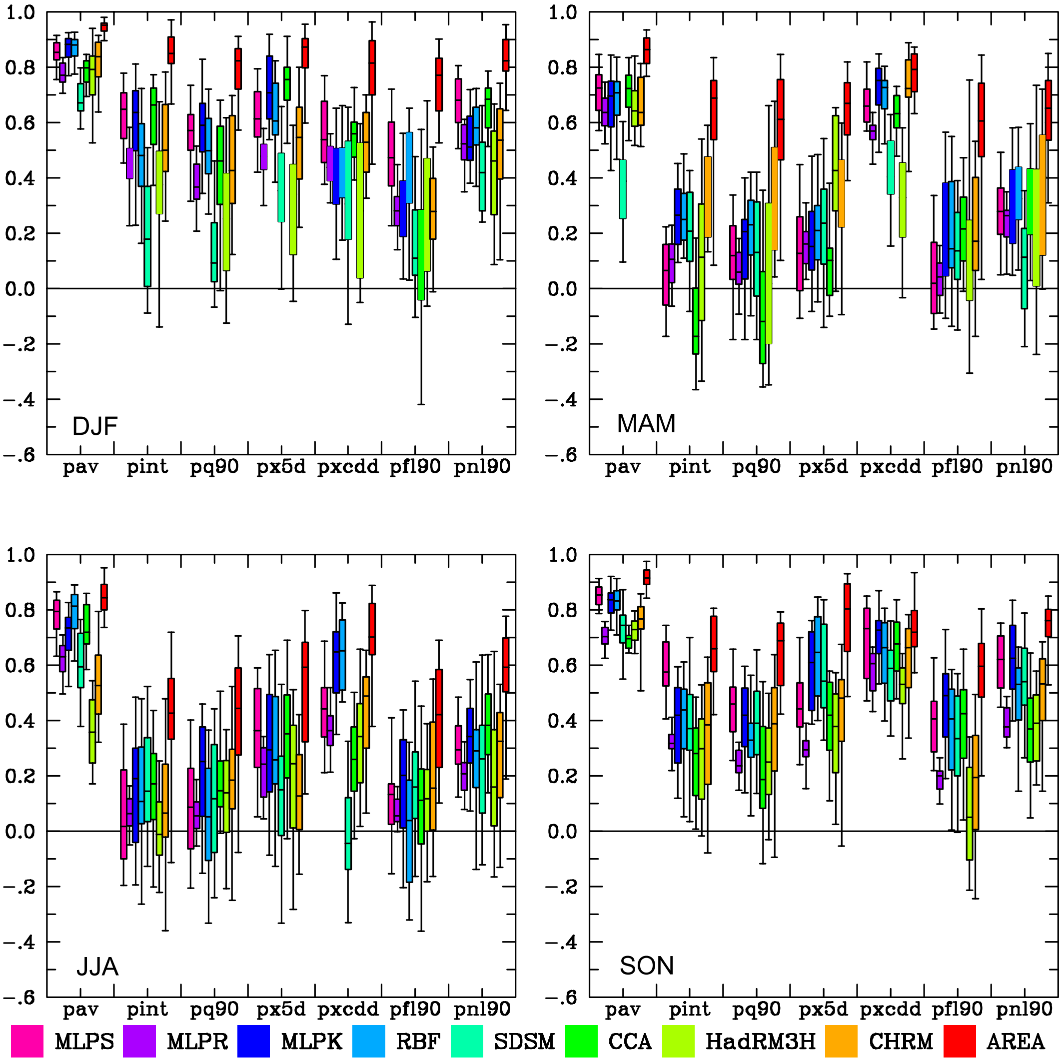

Figure 12: Correlation of modelled and observed precipitation indices for each season for SE England (28 stations). Coloured bars show inter-quartile range across the stations and median with 5th and 95th percentiles indicated by outer range. First six methods are statistical (see Table 4), HadRM3H and CHRM are RCMs. From Haylock et al., 2005.

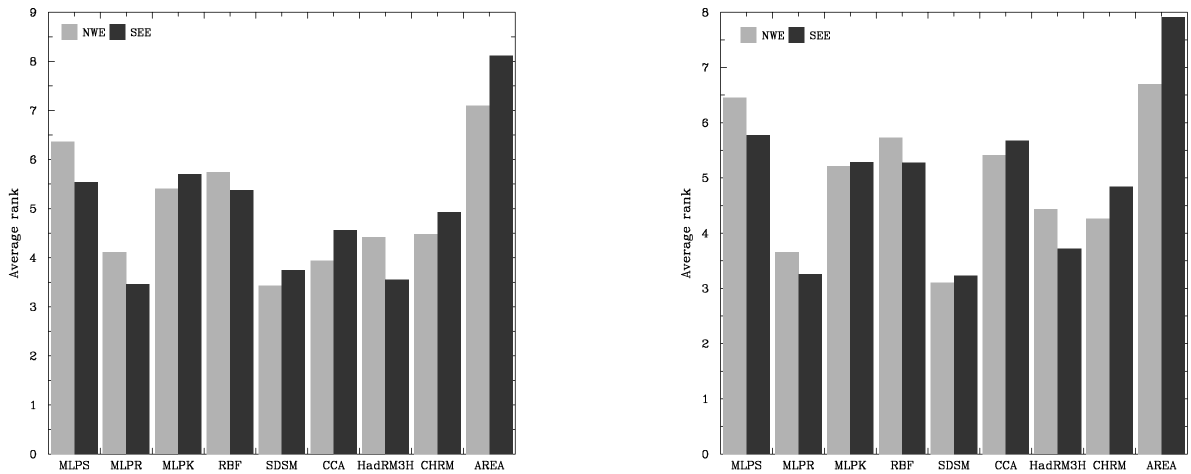

Figure 13: Average rank of correlation (left-hand side) and root mean square error (right) averaged across all indices, seasons and stations for NW and SE England. Higher ranks indicate better performance. First six methods are statistical (see Table 4), HadRM3H and CHRM are RCMs. From Haylock et al., 2005.

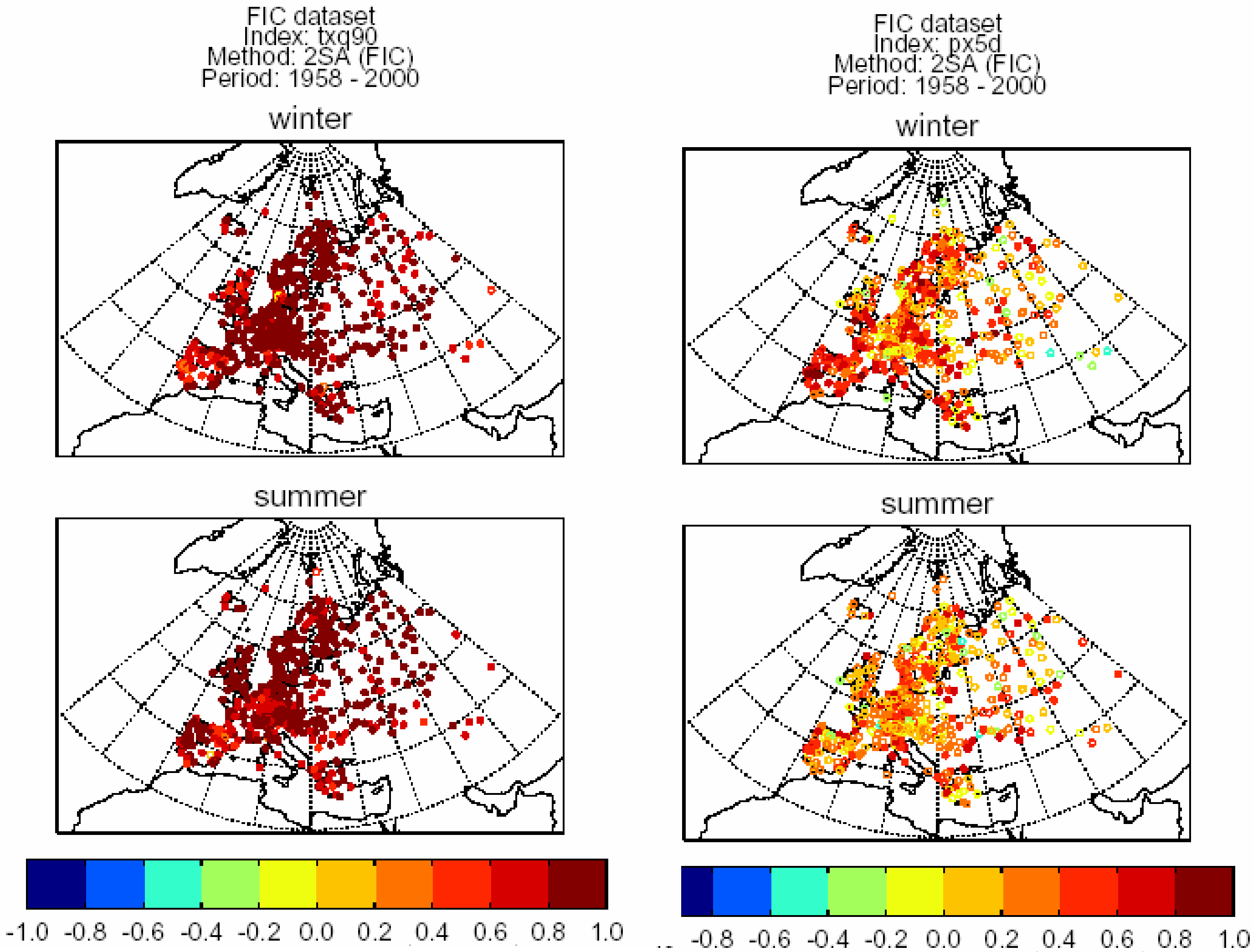

Figure 14: Spearman correlation for the 90th percentile of maximum temperature (hot-day threshold) (left-hand side) and greatest five-day rainfall total (right-hand side) downscaled using the FIC_ANAL2 method for the European-wide dataset, for winter (upper panel) and summer (lower panel).

STARDEX has developed three sets of criteria for assessing the robustness of statistical downscaling methods and the reliability and appropriateness of statistical/dynamical downscaling methods. These are referred to as robustness, performance and application criteria respectively - with the latter most focused on user needs (see deliverable D16).

The robustness criteria reflect the key assumptions of statistical downscaling (Wilby *et al.*, 2004) and comprise of four elements (Table 10). These are: strength and stability; stationarity; uniformity of performance; and, reliability of the simulation of predictors.

Uniformity of performance is self-explanatory and clearly important, as is reliability of predictor simulation. If predictors are reliably simulated, this implies that one of the assumptions of statistical downscaling – that large-scale circulation variables are better simulated than local-scale surface climate variables, is met.

Stability and stationarity are distinct issues – as indicated by the definitions below and the different assessment methods used for each (Table 10).

Stability: This concerns the stability and sensitivity of the statistical model over the calibration/validation period, i.e., over the observational period. Stability relates to both the selection of predictor variables and the relative weight or influence that particular predictors are given within the model.

Stationarity: This concerns the applicability of the statistical model to future time periods with climate change, i.e., to what extent is it legitimate to extrapolate a statistical model which has been calibrated/validated on the observational period to a future warmer period? Stationarity is considered as a separate issue to stability because a stable statistical model which performs well for the present-day is not necessarily the one that will perform best for the future.

Table 10 lists the STARDEX robustness criteria and then summarises the key questions and assessment methods used, indicating the STARDEX deliverables and papers where the latter are presented. Thus this table provides an overview of much of the work undertaken in the STARDEX project. In some cases, novel approaches are proposed for exploring the issues further in future studies.

Table 10: The STARDEX robustness criteria and summary of assessment methods. The final column indicates the STARDEX deliverables and papers where the latter are presented. For deliverables, the most relevant partner reports are indicated in brackets.

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX examples**
'Strength and stability'	Can strong predictor/predictand relationships be identified, supported by, for example, high correlation values?	D10; Haylock and Goodess, 2004
	Are these relationships physically meaningful, i.e., supported by literature review, theoretical considerations and/or local meteorological/synoptic climatology evidence and expertise?	D10; Haylock and Goodess, 2004; Maheras *et al*., 2004; Schmidli *et al*., 2005
	If different methods (e.g., correlation, stepwise multiple regression, compositing) or time periods are used for predictor selection, are similar sets of predictors obtained?	D10 (KCL); D12
	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period (e.g., relatively longer/shorter periods and/or swapping the calibration/validation periods)?	D12 (ARPA-SMR), Tolika *et al.*, 2005b
	Is the statistical model performance sensitive to other user choices, such as predictor domain, number of predictors and model parameters (e.g., choice of misfit term in neural network models)?	D10; D12 (UNIBE); Gyalistras and Schuepbach, in preparation

'Stationarity'	Minimise potential problems by the incorporation of predictors which are expected to change due to global warming (e.g., moisture flux and specific/relative humidity), based on literature review and theoretical considerations. Assess whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model (ideally using cross-validation in order to maximize the analysis period). Assess whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period. Assess whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods. Calibrate the statistical model on a 'cold' period and validate it on a 'warm' period and *vice versa*. Calibrate the statistical model in one region and apply it (without re-calibration) in a warmer region with equally simple topography.	D10 (FIC) D12 (ARPA-SIM, AUTH) D18 USTUTT-IWS, paper in preparation Proposed method not tested.
Uniformity of performance	Uniformity of statistical model performance across: - stations - regions - seasons - variables (i.e., temperature *vs* precipitation, means *vs* extremes) - indices of extremes (e.g., occurrence *vs* magnitude) Evaluated for present-day conditions using: - BIAS (mean difference between simulated and observed values) - CORR (Spearman rank-correlation coefficient) - RMSE (Root Mean Square Error) - ratio of observed : simulated standard deviations Plotted using: - Maps - Histograms - Box-whisker plots - Q-Q diagrams - Taylor diagrams	D12; D16 – see performance criteria tables; Goodess *et al*., 2005; Schmidli *et al.*, in preparation; Haylock *et al*., 2005
Reliability of simulation of predictors	Compare predictors calculated from climate model output with those calculated from Reanalysis data, taking into consideration: - raw values (e.g., sea level pressure, 500 hPa geopotential height, relative/specific humidity) - derived indices (e.g., principal components, blocking indices, synoptic circulation types) - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types.	D13

	Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	D18, e.g., Tolika *et al*., 2005b

Good practice in scenario development should include demonstration of the need for downscaling (see STARDEX deliverable D11) for each user's specific application. For some applications, this need will be more pressing, e.g., those requiring information about extremes at a high temporal and spatial resolution. The applications user needs will also determine which downscaling approach (dynamical or statistical) is most appropriate and, if statistical downscaling is adopted, which specific methods are most appropriate. The STARDEX application criteria have been developed for this purpose. They encompass spatial and temporal scale and consistency, together with resource (computing and data) requirements.

The third set of STARDEX criteria is the performance criteria. Evaluating model skill using independent data is a crucial element of any downscaling application. The STARDEX downscaling methods have been extensively evaluated for present-day conditions using Reanalysis data and focusing on extreme events (D12; Goodess *et al*., 2005). Particular emphasis was given to how well interannual variability is reproduced (measured by correlations – see Table 10), as a proxy for climate change. The reasoning is that if year-to-year changes can be modelled, you can have much more faith that you do indeed have all the relevant predictors and so your scenarios are more meaningful (than if only the biases are small). If interannual variability cannot be well modelled, this implies that relevant predictors may be missing or that noise far overshadows any model skill.

Whilst such evaluation using rigorous statistical testing is vital, the volume and detail of results are likely to be more than most users require, particularly if they are trying to inter-compare and identify a handful of most appropriate methods for their specific application. Thus the STARDEX performance criteria tables attempt to summarise the relative performance confidence and overall performance of each method, as well as indicating the optimal spatial scale and recommended impact applications.

The three sets of STARDEX criteria are described in the summary report for deliverable D16, which also includes a number of additional recommendations and points on good practice based on the STARDEX experience. The D16 partner contributions which accompany this summary report document how the STARDEX criteria have been implemented by partners – thus providing detailed evidence of their utility.

STARDEX has developed improved statistical downscaling methods which have been rigorously evaluated and inter-compared using robustness, application and performance criteria. The most robust of these methods have been used to construct scenarios of extremes for 2070-2100 for the A2 and B2 emissions scenarios for the STARDEX case-study regions and Europe as a whole (see deliverable D18 and supporting papers)

STARDEX has produced three products for each of the STARDEX case-study regions and Europe as a whole (495 stations): a set of visually attractive and uniform format scenario information sheets; a web page providing more graphical information than can be included in the information sheets or journal papers; and a journal paper. The information sheets are suitable for a non-technical audience, with more technical details in the journal papers. An overview information sheet on the STARDEX methods and approach has also been produced to accompany the regional sheets (and all are available from the STARDEX web site).

Scenario changes for the six case-study regions are outlined below, focusing more on rainfall than temperature changes:

- SE and NW England: we are confident that in winter both average and extreme rainfall are likely to increase by a factor of 1 to 1.25 in both regions, with a corresponding decrease in the longest dry period. In summer, we are less confident but the models indicate more of a decrease in average and extreme rainfall (Haylock *et al*., 2005)

- German Rhine: Significant increases in temperature extremes are expected by the end of the 21st century with more severe increases in summer accompanied by higher interannual variability. A similar increase is expected in the magnitude and frequency of occurrence of intense precipitation in winter. The cumulative 5-day rainfall total, for example, is projected to increase by up to 50% for the A2 scenario. For other seasons, nothing can be said about the possible changes of precipitation extremes due to uncertainties (i.e., the models do not agree on the direction of change).

- Emilia Romagna: Significant increases are projected in temperature extremes during winter and autumn, when the minimum temperature increases more than the maximum. In contrast, maximum temperature increases more in spring. The number of frost days is projected to decrease, and heatwave duration to increase (particularly in summer). The rainfall scenarios indicate a slight increase during summer and autumn. A significant increase in the longest dry period is projected for autumn.

- Greece: Temperature means and extremes are projected to increase. For some indices of extremes, the changes for the B2 emissions scenario are comparable to those for the higher A2 scenario. An increase in precipitation extremes is projected for winter in central continental Greece and in part of the Aegean Sea (but a decrease in autumn). A decrease is projected in other parts of the country in winter. Maximum dry spells are projected to lengthen in winter in all parts except the northwest (in the A2 scenario) and the south (B2 scenario). A significant increase in dry-spell length is indicated in summer for nearly all parts of the country.

- Iberian Peninsula: For western Iberia, the projected rainfall changes are very small (a tendency towards slightly drier conditions) in winter, contrasting with large decreases in mean rainfall and most indices of extremes in the other seasons. Changes in dry-day persistence are, however, relatively modest and there is some indication of a slight increase in the proportion of rainfall coming from heavy events. All the high temperature extremes are projected to increase, with the exception of heatwave duration. Large decreases in the number of frost days are projected. For Southeast Spain, the projected changes indicate greater uncertainty, including some counter-intuitive and contradictory changes. In spring, for example, the A2 scenario is associated with large decreases in rainfall means and extremes, whereas the B2 scenario is associated with modest increases. Heatwave duration is projected to decrease, despite increases in the other high temperature extremes. This greater uncertainty is related to the generally poorer validation results for downscaling models in this region of Spain.

- Alps: A detailed comparison of precipitation scenarios constructed using statistical (five methods) and dynamical (three RCMs) downscaling has been undertaken for the Alps (Schmidli *et al*., in preparation) and is discussed in more detail below.

The RCM (CHRM, HadRM3P and HIRHAM) simulated future change in European precipitation shows a seasonally very distinct pattern. In winter, regions north of about 45N experience an increase in mean precipitation, while in the Mediterranean region there is a tendency towards decreases (see also Frei *et al*., 2005). Results are very consistent between the three RCMs (Figure 15). All three RCMs attribute the increase in mean precipitation (MEA) about equally to an increase in wet-day frequency (FRE) and precipitation intensity (INT). In addition, the spatial patterns of relative changes are quite similar. Most of the statistical downscaling models produce an increase in mean precipitation similar to that of the RCMs, although the partition of the increase between FRE and INT varies considerably between the statistical models. Nonetheless, the generally good agreement between the downscaling models suggests that the downscaled scenarios for winter (Figure 15) are fairly reliable and robust.

In summer, the RCMs simulate a strong decrease in mean precipitation in the entire Alpine region (Figure 16). This decrease is mainly due to a substantial reduction of wet-day frequency, which also results in a large increase, 50-100%, of the maximum dry-spell length (Figure 17). In comparison to winter, the differences between the methods, particularly the statistical and dynamical models, are much larger. Even the two statistical downscaling models with very good validation results (MAR and ANA), produce almost no changes or decreases in dry-spell length (Figure 17). This suggests that the RCM simulated changes for summer are not related to circulation changes, but to physical feedback processes with, for example, the land surface. Thus Figures 16 and 17 highlight the still large uncertainties of scenario results for the summer season (a finding common to all the STARDEX case-study regions).

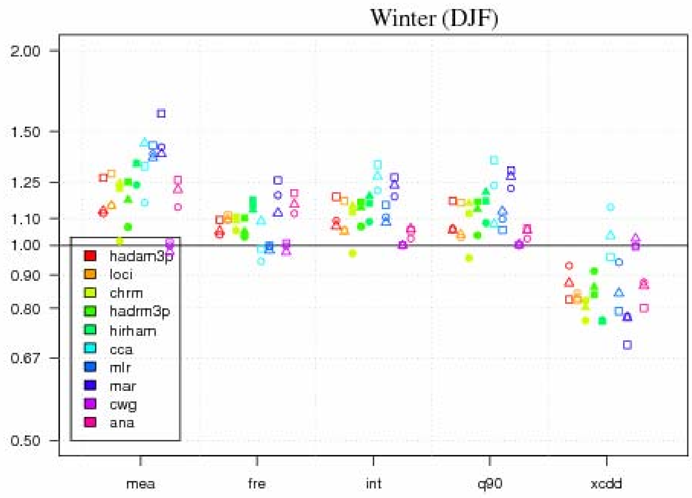

Figure 15: Simulated change (ratio of scenario:control) for the A2 scenario in regional-mean winter precipitation diagnostics for three Alpine regions (west: squares, north: circles, Ticino: triangles). Filled symbols: RCMs, open symbols: statistical downscaling methods. From Schmidli et al., in preparation.

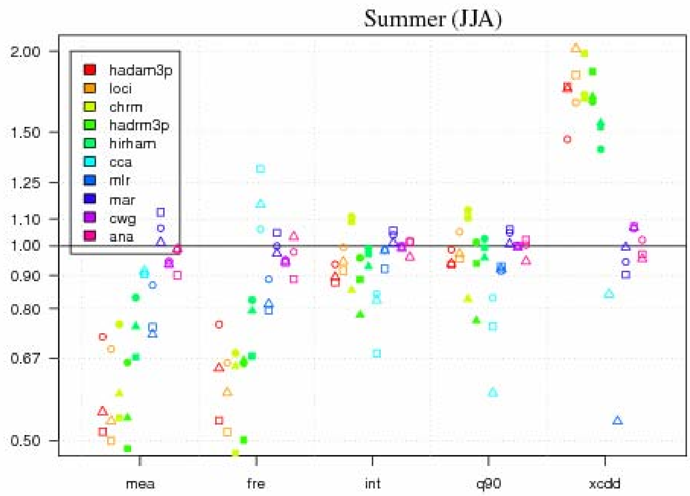

Figure 16: As Figure 15, but for summer..

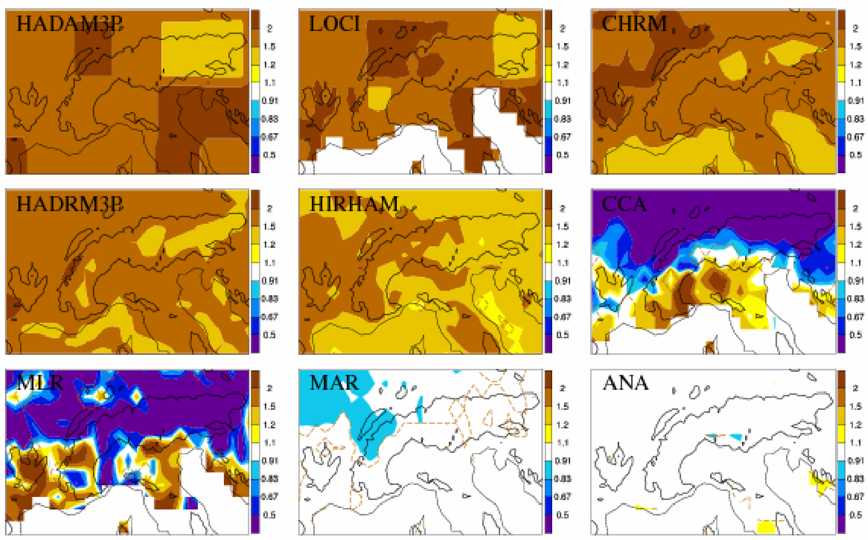

Figure 17: Ratio (scenario:control) of longest dry-spell length in summer. Results from the GCM, 3 RCMs, 5 statistical downscaling models for the A2 scenario. Note the log scale in the colour scale. The dashed line (red) indicates areas with statistically significant (5%) change, in an independent test at each model grid point. Figure from Schmidli et al., in preparation.

STARDEX has demonstrated that uncertainties due to statistical downscaling method need to be considered alongside the other sources of climate scenario uncertainty (see deliverable D19 and supporting papers)

STARDEX has explored various aspects of climate scenario uncertainty:

- Emissions scenarios: A2 and B2 scenarios were used.

- Intra-model uncertainty: downscaled scenarios based on three ensemble members of HadAM3P A2-forced simulations were compared.

- Downscaling method: scenarios were constructed using a number of different statistical downscaling methods for each region and inter-compared. For some regions statistical and dynamical downscaling methods/scenarios were inter-compared.

The main contribution to STARDEX in understanding the sources of uncertainty has been with respect to statistical downscaling method. Uncertainties arise due to choice of statistical method, predictors and their domain. Inter-model differences in the downscaled scenarios were shown to be at least as large as the differences between the emissions scenarios for a single model in some cases (e.g., Haylock *et al*., 2005). The largest uncertainties arise for summer rainfall scenarios. STARDEX work in the UK study region suggests that these difficulties are related to the lower spatial coherence (and hence predictability) of summer rainfall (Haylock *et al*., 2005).

STARDEX did not address inter-model uncertainties associated with the choice of driving GCM, since only the HadAM3P model was used. However, the PRUDENCE project (Christensen *et al*., 2005) has demonstrated that this is a major source of uncertainty with respect to dynamical downscaling, and this conclusion is also expected to be valid for statistical downscaling.

A STARDEX information sheet on scenario uncertainties has been produced which includes references to the various STARDEX deliverables and papers which incorporate relevant work. This work provides a valuable starting point for more extensive work on uncertainties and the construction of probabilistic regional climate scenarios in the ENSEMBLES project (see Section 6.4).

Many of the sources of uncertainty are considered in a STARDEX paper (Wilby and Harris, 2005) which provides a 'road-map' for exploring components of uncertainty in impacts assessments, using the example of low-flow scenarios for the River Thames, UK – an issue of great concern to UK stakeholders such as the Environment Agency. The following sources of uncertainty are considered and various weighting schemes used:

- Emissions scenario uncertainty
  - A2 = 0.5, B2 = 0.5 (i.e., equal weighting)
- Climate model uncertainty
  - Impacts Relevant-Climate Prediction Index = GCM skill, proportional to bias in summer precipitation minus potential evapotranspiration
- Downscaling uncertainty
  - SDSM model and change factor = equal weighting
- Hydrological model parameter uncertainty

– 100 parameter sets weighted by Nash-Sutcliffe score
- Hydrological model structural uncertainty
  – CATCHMOD and simpler REGMOD – weighted by correlation with observations

A Monte Carlo analysis is used to sample across these components. Analysis of the individual components suggests, in this particular case, the following order of component significance (greatest to least) with respect to uncertainties: GCM > downscaling method > hydrological model structure > hydrological model parameters > emission scenario. These sources of uncertainty can also be combined into a single distribution function either with (conditional) or without (non-conditional) weighting of the individual components (Figure 18).

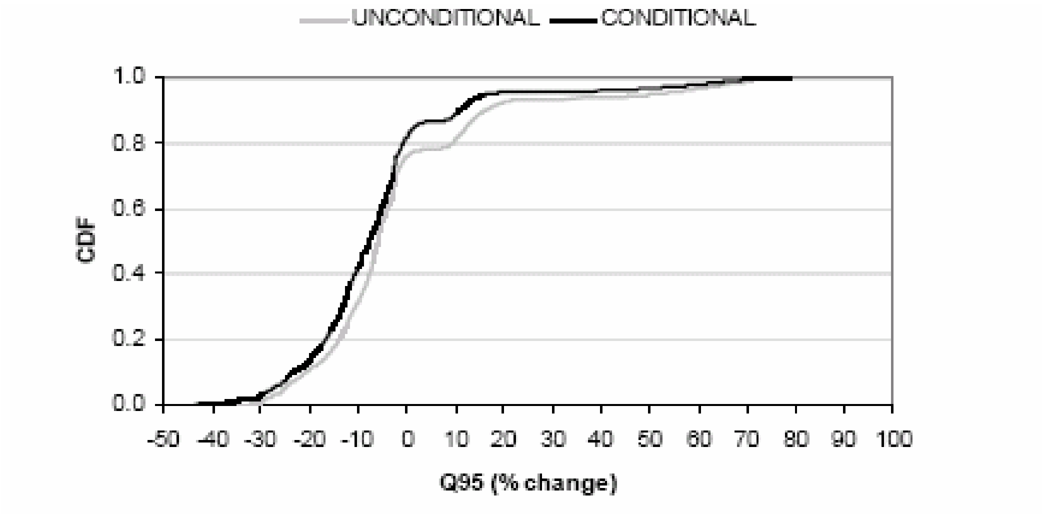

Figure 18: Cumulative distribution functions of Thames low flows (Q95) by the 2080s for unconditional and conditional experiments. Figure from Wilby and Harris, 2005.

6.4 CONCLUSIONS INCLUDING SOCIO-ECONOMIC RELEVANCE, STRATEGIC ASPECTS AND POLICY IMPLICATIONS

The scientific achievements of STARDEX can be summarised as follows from Section 6.3.3:

- Extensive analysis of observed data, supported by an understanding of the underlying physical processes, provides a sound basis for the construction of statistical downscaling methods

- Spatially coherent changes in extremes have occurred over the last 40 years

- In part, these observed changes in extremes can be explained by changes in large-scale circulation and other predictors

- When identifying the best predictor variables, it is easier to make recommendations about methodologies than standard predictor sets

- In general, the predictor variables identified for use in the STARDEX statistical downscaling models are well simulated by the GCM (HadAM3P) chosen as the basis for STARDEX work

- STARDEX developed and evaluated a range of statistical downscaling methods of differing complexity

- Handling many combinations of statistical downscaling methods (20+), regions (7), indices (13) and seasons (4), made intercomparison complex

- A number of key messages emerged from the reanalysis-based validation studies, along with two 'headline' conclusions:

    - In the majority of cases, no consistently superior statistical downscaling model can be identified, so a major recommendation is to use a range of the better statistical downscaling methods – just as the recommendation is to use a range of GCMs/RCMs
    - For many regions and indices, the skill is unacceptably low for summer rainfall – thus scenarios should not be constructed for these cases

- STARDEX has developed three sets of criteria for assessing the robustness of statistical downscaling methods and the reliability and appropriateness of statistical/dynamical downscaling methods. These are referred to as robustness, performance and application criteria respectively - with the latter most focused on user needs

- STARDEX has developed improved statistical downscaling methods which have been rigorously evaluated and inter-compared using robustness, application and performance criteria. The most robust of these methods have been used to construct scenarios of extremes for 2070-2100 for the A2 and B2 emissions scenarios for the STARDEX case-study regions and Europe as a whole

- STARDEX has demonstrated that uncertainties due to statistical downscaling method need to be considered alongside the other sources of climate scenario uncertainty

Much of the STARDEX work has most direct relevance to developers and users of climate scenarios, which provide the essential basis for all assessments of the environmental and socio-economic impacts of climate change. The guidelines and recommendations contained in deliverable 16 about robust statistical downscaling methods should, therefore, be of particular interest to these groups.

External advisors and stakeholders, e.g., from the re-insurance industry, were involved in the project, e.g., as attendees at project meetings. They helped to ensure that the final users of scenario information were not forgotten and helped to shape the format and content of project deliverables and other outputs (see Section 6.5).

While the core indices of extremes used in the project (Table 3) were defined primarily from a climate perspective for the purposes of developing and evaluating statistical downscaling methods for the construction of scenarios of extremes, they are still relevant for impacts purposes. The maximum five-day total rainfall, for example, is an important measure of extreme from the point of view of flooding in a basin like the Rhine.

During the course of the project, the STARDEX team inevitably identified some things that could have been done differently or better (for example, we could have attempted to use a single case-study region). But this is unlikely to change the main findings which are that:

- ❖ systematic changes in extremes have occurred and had an impact over the last 40 years, and major changes are projected for the future
- ❖ there are uncertainties in regional scenarios of extremes due, in part, to the downscaling method used
- ❖ there is, therefore, a need to take a multi-model approach to regional scenario construction, whether using statistical and/or dynamical downscaling methods

This implies a need for the development of new and efficient tools and techniques for combining output from multiple downscaling models. The scenario results presented in deliverable D18, for example, are based on simple averaging across methods, rather than attempting to weight each method in terms of its reliability. These are all issues for the construction of probabilistic scenarios, which are being developed in the FP6 ENSEMBLES project (http://www.ensembles-eu.org/), and for which STARDEX has provided a sound scientific starting point.

6.5 DISSEMINATION AND EXPLOITATION OF THE RESULTS

A number of different means of dissemination have been used by STARDEX.

A major focus of STARDEX dissemination effort is the public web site (http://www.cru.uea.ac.uk/projects/stardex) which provides access to all public deliverables and the central data server (see Table 6). The major report deliverables (D9 to D13, D15, D16) are presented using the same layout: Objective; Responsible author; Outline; Reports (i.e., synthesis and partner reports); and, Presentations. The latter Powerpoint files allow anyone to include STARDEX results in their conference or meeting presentations or for teaching purposes. One of the presentations is available in Spanish. Deliverables D18 and D19 are also available from the web site, but in this case, material is primarily presented in the form of information sheets suitable for a more general audience.

As well as these information sheets on regional scenarios and their associated uncertainties, three general information sheets are available from the web site:

- Information sheet 1: Camouflage, Bluff, or Real? Statistical Uncertainty of Trends in Catastrophic Extremes, produced by ETH
- Information sheet 2: The August 2002 Flood in Central and Eastern Europe and Results from the EU STARDEX Project, produced by FTS
- Information sheet 3: Record Warm Summer in Western Europe in 2003, production co-ordinated by UEA.

Scenario data sets and maps are available from the central data server. It is anticipated that these will provide a valuable resource for impacts assessment work until new regional scenarios are available from the ENSEMBLES project (e.g., output from new 25 km resolution RCM climate change simulations will be available from autumn 2007).

The STARDEX software for the calculation of indices of extremes is also freely available from the web site. It is being widely used throughout Europe and elsewhere (e.g., in Canada, Peru, China, Nigeria). Calculated indices of extremes for 495 European stations are also available.

Given the importance of the STARDEX web site in our dissemination strategy and for the above reasons, it will be maintained for a period of at least three years after the end of the project.

STARDEX partners have written many peer-reviewed papers over the course of the project (see Section 6.6) and several more are still in preparation. It is hoped that other groups will publish papers in due course, based on the scenario datasets available from the central data server.

STARDEX partners have given many presentations at national, European and international meetings and conferences over the course of the project. The cumulative publications list, for example, contains well over 50 conference-related publications. These include invited presentations made by the STARDEX co-ordinator at high-profile meetings in Canada, China and Peru. As well as conference-related items, the non-referred publications list includes newspaper and magazine articles. Some of these publications are in French, German or Italian.

STARDEX has had inputs into research programmes outside Europe. A number of Canadian participants in the Quebec-based Ouranos climate impacts consortium, for example, are using the STARDEX diagnostic software tool. Ouranos participants have also attended STARDEX project meetings. A report on 'Climate Change Scenarios for Peru 2004-2050' produced by SENAMHI, the Peruvian national hydrological and weather service, in 2004 acknowledges STARDEX and also makes use of the software tool.

A number of STARDEX members are involved in the production of the IPCC Fourth Assessment Report, as lead authors and reviewers, and it is anticipated that a number of STARDEX papers will be cited in this report.

An overview presentation on STARDEX was given at an EU sponsored side-event on the PRUDENCE project at COP10 in Buenos Aires, December 2004. A presentation is also planned for another EU sponsored side-event on the MICE, STARDEX, ENSEMBLES and ADAM projects at COP11 in Montreal, November 2005.

STARDEX partners have also fed project results into national and regional bodies. Hans Caspary, for example, has made recommendations for policy makers and hydrologic and hydraulic engineering practice in the water resources management administration of the State of Baden-Württemberg in southwest Germany. Christoph Frei was a member of PLANAT – the Swiss National Platform for Hazards. STARDEX results have also been integrated into the preparation of a 10-year regional Alpine research and educational programme for the Canton of Valais, Switzerland, and were used in an interdisciplinary workshop in June 2004 to develop an action plan for the city of Geneva, focusing on elderly people, during heatwaves.

Partners have also engaged in educational activities. UNIBE for example, disseminated STARDEX results to young scientists and stakeholders during young atmospheric scientists workshops on 'Heatwave summer 2003' (IUKB Sion and British Council Berne, December 2003) and 'Climate change and tourism' (IUKB Sion and Jungfraujoch, July 2004).

As noted in Section 6.4, the STARDEX work will be exploited and further developed in the EU FP6 ENSEMBLES project, in which a number of STARDEX partners are also involved. Thus, STARDEX has provided a capacity building opportunity for many partnerss, enhancing their ability to participate in research programmes at national, European and international levels.

6.6 MAIN LITERATURE PRODUCED

The STARDEX deliverables include a number of reports and information sheets, which are summarised in Table 6 (see also Section 6.5). These are supported by scientific papers in peer-reviewed journals. Below, are listed the 26 such papers which have been published, accepted for publication or submitted during the project period. Several more are still in preparation. The list includes joint papers by several STARDEX partners as well as joint papers by STARDEX/MICE and STARDEX/PRUDENCE participants.

Anagnostopoulou, Chr., Maheras, P., Karacostas, T. and Vafiadis, M., 2003: 'Spatial and temporal analysis of dry spells in Greece', *Theoretical and Applied Climatology,* **74**, 77-91.

Anagnostopoulou, C., Flocas, H., Maheras, P. and Patrikas, I., 2004: 'Relationship between atmospheric circulation over Greece and western – central Europe during the period 1958-1997', *International Journal of Climatology,* **24**, 1745-1758.

Bárdossy, A., Stehlik, J. and Caspary, H.J., 2002: 'Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimised fuzzy rules', *Climate Research,* **23**, 11-22.

Beniston, M., Stephenson, D.B.., Christensen, O.B., Ferro, C.A.T., Frei, C., Goyette, S., Halsnaes, K., Holt, T., Jylhä, K., Koffi, B., Palutikof, J., Schöll, R., Semmler, T. and Woth, K., 2005: 'Future extreme events in European climate: An exploration of regional climate model projections', *Climatic Change,* submitted.

Busuioc, A., Tomozeiu, R. and Cacciamani, C., 2005: 'Statistical downscaling model for winter extreme precipitation events in Emilia-Romagna region', *International Journal of Climatology,* submitted.

Flocas, H., Tolika, K., Anagnostopoulou, Chr., Patricas, I., Maheras, P. and Vafiadis, M., 2005: 'Evaluation of maximum and minimum temperature NCEP-NCAR Reanalysis data over the Greek area', *Theoretical and Applied Climatology,* **80**, 49-65.

Frei, C., Schöll, R., Fukutome, S., Schmidli, J. and Vidale, P.L., 2005: 'Future change of precipitation extremes in Europe: An intercomparison of Regional Climate Model scenarios', *Journal of Geophysical Research,* submitted.

Goodess, C.M., Anagnostopoulou, C., Bárdossy, A., Frei. C., Harpham, C., Haylock, M.R., Hundecha, Y., Maheras, P., Ribalaygua, J., Schmidli, J., Schmith, T., Tolika, K., Tomozeiu, R. and Wilby, R.L., 2005: 'An intercomparison of statistical downscaling methods for Europe and European regions – assessing their performance with respect to extreme temperature and precipitation events', *Climatic Change,* submitted.

Harpham, C. and Wilby, R.L., 2004: 'Multi-site simulation of daily precipitation amounts using Artificial Neural Networks', *Proceedings of the British Hydrological Society International Hydrology Conference, London,* in press.

Harpham, C. and Wilby, R.L., 2005: 'Multi-site downscaling of heavy daily precipitation occurrence and amounts', *Journal of Hydrology,* in press.

Harpham, C., Wilby, R.L. and Dawson, C.W., 2005: 'Downscaling daily rainfall in semi-arid regions: lessons learnt from the Iberian Peninsula', *Climate Research,* submitted.

Haylock, M.R. and Goodess, C.M., 2004: 'Interannual variability of European extreme winter rainfall and links with mean large-scale circulation', *International Journal of Climatology,* **24**, 759-776.

Haylock, M.R., Cawley, G.C., Harpham, C., Wilby, R.L. and Goodess, C.M., 2005: 'Downscaling heavy precipitation over the UK: a comparison of dynamical and statistical methods and their future scenarios', *International Journal of Climatology,* submitted.

Hundecha, Y. and Bárdossy, A., 2005: ‘Trends in daily precipitation and temperature extremes across western Germany in the second half of the 20th century’, *International Journal of Climatology,* **25**, 1189-1202.

Kostopoulou, E., Giannakopoulos, C., Anagnostopoulou, C., Tolika, K., Maheras, P., Vafiadis, M. and Founda, D., 2005: ‘Simulating maximum and minimum temperature over Greece: A comparison of three downscaling techniques’, *Climate Research*, submitted.

Maheras, P., Vafiadis, M., Kolyva-Machera, F., Anagnostopoulou, Ch. et Tolika, K., 2003: ‘Les champs des anomalies pluviométriques durant la saison humide en Grèce et leurs relations avec le géopotential à 500hPa’, *Publications de l’ Association Internationale de Climatolgie (AIC),* **15**, 312-318.

Maheras, P., Tolika, K., Anagnostopoulou, Chr., Vafiadis, M., Patrikas, I. and Flocas, H., 2004: ‘On the relationships between circulation types and changes in rainfall variability in Greece’, *International Journal of Climatology,* **24**, 1695-1712.

Pavan, V., Marchesi, S., Selvini, A., Tomozeiu, R. and Di Lorenzo, M., 2005a: ‘A data quality control scheme for a dense temperature and precipitation network in the Mediterranean region’, *International Journal of Climatology,* submitted.

Pavan, V., Collins, M., Junge, M.M., Quadrelli, R. and Stephenson, D., 2005b: ‘The role of the North Atlantic Oscillation in European climate change: CMIP2 coupled model evaluation, *Climate Dynamics,* submitted.

Schmidli, J., and Frei, C., 2005: ‘Trends of heavy precipitation and wet and dry spells in Switzerland during the 20th century’, *International Journal of Climatology,* **25**, 753-771.

Schmidli, J., Frei, C. and Vidale, P.L., 2005: ‘Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling’, *International Journal of Climatology,* in press.

Tolika, K., Maheras, P., Flocas, H. and Arseni-Papadimitriou, A., 2005a: ‘An evaluation of a general circulation model (GCM) and the NCEP/NCAR reanalysis data for the winter precipitation in Greece, *International Journal of Climatology,* in press.

Tolika, K., Maheras, P., Vafiadis, M., Flocas, H. and Arseni- Papadimitriou, A., 2005b: ‘Simulation of seasonal precipitation and raindays over Greece: A statistical downscaling technique based on Artificial Neural Networks (ANNs)’, *Climatic Change,* submitted.

Tomozeiu, R., Pavan, V., Cacciamani, C. and Amici, M., 2005: ‘Observed temperature changes in Emilia-Romagne: mean values and extremes’, *Climate Research,* in press.

Wilby, R.L., Tomlinson, O.J. and Dawson, C.W., 2003: ‘Multi-site simulation of precipitation by conditional resampling’, *Climate Research,* **23**, 183-194.

Wilby, R.L. and Harris, I., 2005: ‘A framework for assessing uncertainties in climate change impacts: low flow scenarios for the River Thames, UK’, *Water Resources Research,* submitted.

REFERENCES

For STARDEX-authored references which are published, in press, or submitted, see Section 6.6.

Beersma, J., Agnew, M.D., Viner, D. and Hulme, M., 2000: *Climate Scenarios for Water-Related and Coastal Impacts,* Proceedings of the EU Concerted Action Initiative ECLAT-2 Workshop 3, KNMI, Netherlands, May 10-12th 2000, Climatic Research Unit, Norwich, UK, 140pp.

Christensen, J.H. et al., 2005: 'Evaluating the performance and utility of regional climate models in climate change research: Reducing uncertainties in climate change projections – the PRUDENCE approach', *Climatic Change*, submitted.

Cramer, W., Doherty, R., Hulme, M. and Viner, D., 2000: *Climate Scenarios for Agricultural and Ecosystem Impacts,* Proceedings of the EU Concerted Action Initiative ECLAT-2 Workshop 2, Potsdam, Germany October 13th - 15th, 1999, Climatic Research Unit, Norwich, UK, 120pp.

Frei, C. and Schär, C., 1998: 'A precipitation climatology of the Alps from high-resolution rain-gauge observations', *International Journal of Climatology*, **18**, 873-890.

Frei, C. and Schär, C., 2001: 'Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region', *Journal of Climate,* **14,** 1564-1584.

Giorgi, F. and Mearns, L.O., 1999: 'Introduction to special section: Regional climate modeling revisited', *Journal of Geophysical Research*, **104**, 6335-6352.

Gyalistras, D. and Schuepbach, E., 2005: Statistical downscaling of seasonal temperature statistics in the European Alps: Sensitivity of model skill and climate change scenarios to choice of large-scale predictors, *Climate Dynamics*, in preparation.

Hewitson, B.C. and Crane, R.G., 1996: 'Climate downscaling: techniques and application', *Climate Research*, **7**, 85-95.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, D., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R. and Joseph, D., 1996: 'The NCEP/NCAR 40-year Reanalysis project', *Bulletin of the American Meteorological Society*, **77**, 437-471.

Karl, T.R., Nicholls, N. and Ghazi, A. (eds), 1999: 'Weather and climate extremes: Changes, variations and a perspective from the insurance industry', *Climatic Change*, **42**, 1-349.

Machenhauer, B., Windelband, M., Botzet, M., Hesselbjerg, J., Déqué, M., Jones, G.R., Ruti, P.M. and Visconti, G., 1998: *Validation and Analysis of Regional Present-day Climate and Climate Change Simulations over Europe*, Max-Planck Institute of Meteorology Report No. 275, 87pp.

Mearns, L.O., Bogardi, I., Giorgi, F., Matyasovszky, I. and Palecki, M., 1999: 'Comparison of climate change scenarios generated from regional climate model experiments and statistical downscaling', *Journal of Geophysical Research*, **104**, 6603-6621.

Mearns, L.O., Giorgi, F., Whetton, P., Pabon, D., Hulme, M. and Lal, M., 2003: *Guidelines for Use of Climate Scenarios Developed from Regional Climate Model Experiments*, Intergovernmental Panel on Climate Change (IPCC) Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA), available from: http://ipcc-ddc.cru.uea.ac.uk/guidelines/dgm_no1_v1_10-2003.pdf.

Meehl, G.A., Zwiers, F., Evans, J., Knutson, T., Mearns, L. and Whetton, P., 2000: 'Trends in extreme weather and climate events: Issues related to modeling extremes in projections of future climate change', *Bulletin of the American Meteorological Society,* **81**, 427-436.

Michaels, P.J., Knappenberger, P.C., Frauenfeld, O.W. and Davie, R.E., 2004: 'Trends in precipitation on the wettest days of the year across the contiguous USA', *International Journal of Climatology*, **24**, 1873-1882.

Moberg, A., and Alexanderson, H., 1997: 'Homogenisation of Swedish temperature data. Part II: Homogenised gridded air temperature compared with a subset of global gridded air temperature since 1861', *International Journal of Climatology*, **17**, 35-54.

Munich Re, 2002: *Flooding in Central and Eastern Europe – August 2002*. MRNatCatposter 31, www.munichre.com/default_e.asp.

Murphy, J., 1999: 'An evaluation of statistical and dynamical techniques for downscaling local climate', *Journal of Climate,* **12**, 2256-2284.

Murphy, J.M., 2000: 'Predictions of climate change over Europe using statistical and dynamical downscaling techniques', *International Journal of Climatology*, **20**, 489-501.

Nakicenovic, N. and Swart, R. (eds.), 2000: *Special Report on Emissions Scenarios*, Cambridge University Press, Cambridge, United Kingdom, pp. 612.

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

Schmidli, J., Goodess, C.M., Frei, C., Haylock, M.R., Hundecha, Y., Ribalaygua, J. and Schmith, T., 2005: Statistical and dynamical downscaling of precipitation: Evaluation, intercomparison, and scenarios for the European Alps, *Journal of Geophysical Research*, in preparation.

von Storch, H., Zorita, E. and Cubasch, U., 1993: 'Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime', *Journal of Climate,* **6**, 1161-1171.

Wijngaard, J.B., Klein Tank, A,M.G. and Konnen, G.P., 2003: 'Homogeneity of 20th century European daily temperature and precipitation series'. *International Journal of Climatology*, **23**, 679-692.

Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J. and Wilks, D.S., 1998: 'Statistical downscaling of general circulation model output: A comparison of methods', *Water Resources Research*, **34**, 2995-3008.

Wilby, R.L., Charles, S.P., Zorita, E., Timbal, B., Whetton, P. and Mearns, L.O., 2004: *Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods*, Intergovernmental Panel on Climate Change (IPCC) Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA), available from: http://ipcc-ddc.cru.uea.ac.uk/guidelines/StatDown_Guide.pdf.

Zorita, E. and von Storch, H., 1999: 'The analog method as a simple statistical downscaling technique: Comparison with more complicated methods', *Journal of Climate*, **12**, 2474-2489.