STARDEX

STAtistical and Regional dynamical Downscaling of EXtremes for European regions

EVK2-CT-2001-00115

Deliverable D16 – Partner Contributions

Recommendations on the more robust statistical and dynamical downscaling methods for the construction of scenarios of extremes

FOREWORD

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

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

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

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

STARDEX is organised into five workpackages including Workpackage 4 on 'Inter-comparison of improved downscaling methods with emphasis on extremes' which was responsible for the production of this deliverable (D16). Workpackage 3 is co-ordinated by Torben Schmith from the Danish Meteorological Institute.

STARDEX PROJECT MEMBERS

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

D16 AUTHORS AND VERSION HISTORY

Lead author:
Clare Goodess, UEA

Contributing authors:
Stefano Alberghi, ADGB
Christina Anagnostopoulou, AUTH
András Bárdossy, USTUTT-MAR
Carlo Cacciamani, ARPA-SIM
Bo Christiansen, DMI
Dimitrios Gyalistras, UNIBE
Yeshewatesfa Hundecha, USTUTT-MAR
Panagiotis Maheras, AUTH
Antonella Morgillo, ARPA-SIM
Valentina Pavan, ARPA-SIM
Jaime Ribalaygua, FIC
Jürg Schmidli, ETH
Torben Schmith, DMI
Konstantia Tolika, AUTH
Rodica Tomozeiu, ARPA-SIM
Luis Torres, FIC
Ennio Tosi, ADGB

Version 1.0: 27 July 2005
Version 2.0: 30 September 2005

CONTENTS

The following information is provided for each method:

- Summary
- STARDEX deliverables and papers
- Overall assessment
- Robustness criteria
- Application criteria
- Performance criteria
- Information provided by

A distinction is made between direct methods in which seasonal indices of extremes are downscaled and indirect methods in which daily time series are generated and the seasonal indices then calculated from these.

The STARDEX indices of extremes.

Precipitation related indices of extremes		*User-friendly name*
pq90	90^{th} percentile of rainday amounts (mm/day)	Heavy rainfall threshold
px5d	Greatest 5-day total rainfall	Greatest 5-day rainfall (amount)
pint	Simple daily intensity (rain per rainday)	Average wet-day rainfall (amount)
pxcdd	Maximum number of consecutive dry days	Longest dry period
pfl90	% of total rainfall from events > long-term 90^{th} percentile	Heavy rainfall proportion
pnl90	Number of events > long-term 90^{th} percentile of raindays	Heavy rainfall days
Temperature related indices of extremes		
txq90	Tmax 90^{th} percentile (ºC)*	Hot-day threshold
tnq10	Tmin 10^{th} percentile (ºC)**	Cold-night threshold
tnfd	Number of frost days Tmin < 0 °C	Frost days
txhw90	Heat wave duration (days)	Longest heatwave
Mean indices		
pav	Precipitation average (mm/day)	Average daily rainfall (amount)
txav	Average Tmax (ºC)	Average daily high temperature
tnav	Average Tmin (ºC)	Average daily low temperature

* Alternative definition – 10^{th} hottest day per season/36^{th} hottest day per year
** Alternative definition – 10^{th} coldest night per season/36^{th} coldest night per year

ADGB_HYPER4 Indirect Method

Summary of the method

Indirect 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	Series of sqrt($PC1^2$+$PC2^2$) PC1 and PC2 are principal components of the gridded daily precipitation over Northern Italy	GPH anomaly and geostrophic wind direction at 500 hPa, Rh at 700 hPa precipitable water all at selected grid points.	Statistical predictor predictand link

STARDEX deliverables and papers
See ADGB contributions to D10, D12, D15 and D18.

Overall assessment of the ADGB_HYPER4 method
A relatively simple method for downscaling an areal index highly correlated with extreme precipitation events over Northern Italy.

Robustness criteria for the ADGB_HYPER4 method

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX assessments**
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes, supported by high correlation values, D10(ADGB).
	Are these relationships physically meaningful?	Yes, supported by literature review and strong links with the NAO, see D12(ADGB).
	If different methods or time periods are used for predictor selection, are similar sets of predictors obtained?	Not tested.
	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period?	Relatively insensitive.
	Is the statistical model performance sensitive to other user choices?	Relatively insensitive. User choices tested: use of ERA

		reanalysis instead of NCEP, change of the number of events used to construct the observed statistic.
'Stationarity'	Predictors incorporated which are expected to change due to global warming? Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model? Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period? Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods? Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period and *vice versa*?	Relative Humidity and precipitable water due to increase in temperature. No Yes No Yes, since the same observed statistics has been used for all seasons.
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 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	The model refers to an areal index and has been applied to a single region. The CORR is good for all seasons, while the Bias increases from good in spring to acceptable in autumn.
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - derived indices	Generally well simulated, see D13 (ADGB, Summary, ETH central analysis).

	- spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	 Not much. See: D14 results

Application criteria for the **ADGB_HYPER4** *method*

ADGB_HYPER4 method provides:	**Yes/No**	**Comments/Notes**
Station-scale information	No	
Grid-box information	No	
European-wide information	No	Only one series representative of all Northern Italy.
Daily time series	Yes	But about half of total days are not downscaled (because only days in which predictors lie in a range favorable for extremes are downscaled)
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	No	
Spatially consistent multi-site information	No	
Temporally consistent multi-site information	No	
Information at sites with no observations	/	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Relatively low	
Volume of data inputs	medium	
Availability of input data	medium	

Performance criteria for the ADGB_HYPER4 method

ADGB_HYPER4 method	Relative High	Performance Medium	Confidence Low
Temperature			
Indices			
Seasons			
Regions			
Precipitation			
Indices	pnl90		
Seasons	Winter, Spring	Summer	Autumn
Regions			

Overall performance:	
Mean temperature	N/A
Temperature extremes	N/A
Mean precipitation	N/A
Precipitation extremes	Good
Optimal spatial scale:	Northern Italy
Recommended impact applications:	Any which require information on an area.

Information on ADGB_HYPER4 method provided by
Ennio Tosi
Stefano Alberghi

ARPA-SIM_CCA Direct Method

Summary of the method

Direct 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)*
CCA	Station time series of seasonal extremes – temperature and precipitation	MSLP, Z500, T850hPa, Specific humidity at 850hPa	Canonical Correlation Analysis

STARDEX deliverables and papers
See: ARPA –SIM contributions to D10, D12, D15 and D18.

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

Overall assessment of the CCA method

CCA is a multivariate statistical technique that objectively defines the most highly related patterns of potential predictors and predictands.

Additional comments on the CCA method

CCA is sensitive to the number of EOFs/CCP (Canonical Correlation Patterns) used in the downscaling model, such that many tests have to be done in order to find the best combination.

Robustness criteria for the CCA method

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX assessments**
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes, see D10(ARPA-SIM)
	Are these relationships physically meaningful?	Yes, see D10(ARPA-SIM)
	If different methods or time periods are used for predictor selection, are similar sets of predictors obtained?	. Similar set of predictors has been obtained for some temperature indices, but generally different for precipitation. Method tested: CCA against MLR

		Periods tested: 1958-1978+1994-2000
	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period?	Relatively insensitive - test made for temperature; Non-standard periods tested: -calibration 1960-1989 -verification 1990-2000
	Is the statistical model performance sensitive to other user choices?	Relatively sensitive with the area selected for predictors
'Stationarity'	Predictors incorporated which are expected to change due to global warming?	-
	Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model?	Trends and low-frequency variability are well reproduced for mean temperature and precipitation and less well for extreme precipitation (assessed for 1960-1990 period)
	Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period?	No
	Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods?	No
	Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period and *vice versa*?	No
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)	Relatively uniform across stations, more variable across regions and seasons. See: D12
	Evaluated 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	Evaluated using: BIAS,CORR;RMSE See:D12

	Plotted using: - Maps - Histograms - Box-whisker plots - Q-Q diagrams - Taylor diagrams	Histograms
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - derived indices - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	Spatial patterns and temporal trends of predictors are well simulated. See: D13 ARPA-SIM Not much. See: D13 ARPA-SIM

Application criteria for the CCA *method*

CCA method provides:	Yes/No	Comments/Notes
Station-scale information	Yes	
Grid-box information	No	
European-wide information	No	
Daily time series	No	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	No	Could be implemented if method were applied to temperature also
Spatially consistent multi-site information	Yes	
Temporally consistent multi-site information	-	
Information at sites with no observations	-	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Medium	
Volume of data inputs	Medium	
Availability of input data	Medium	

Performance criteria for the CCA method

<table>
<tr><th rowspan="2">CCA method</th><th colspan="3">Relative Performance Confidence</th></tr>
<tr><th>High</th><th>Medium</th><th>Low</th></tr>
<tr><td>Temperature</td><td></td><td></td><td></td></tr>
<tr><td>Indices</td><td>Txav, Tnav</td><td>Txq90, Tnq10, Tnfd, Txhw90</td><td></td></tr>
<tr><td>Seasons</td><td>Winter</td><td>Spring, Summer, Autumn</td><td></td></tr>
<tr><td>Regions</td><td>Emilia-Romagna, Greece</td><td>Emilia-Romagna, Greece</td><td></td></tr>
<tr><td>Precipitation</td><td></td><td></td><td></td></tr>
<tr><td>Indices</td><td>Pav</td><td>Pxcdd</td><td>Pq90, px5d, pint, pxcdd, pfl90, pnl90</td></tr>
<tr><td>Seasons</td><td>Winter</td><td>Spring, Autumn</td><td>Summer</td></tr>
<tr><td>Regions</td><td>Emilia-Romagna, Greece</td><td>Emilia-Romagna, Greece</td><td>Emilia-Romagna, Greece</td></tr>
<tr><td colspan="2">Overall performance:
Mean temperature
Temperature extremes
Mean precipitation
Precipitation extremes</td><td colspan="2">
Good
Average
Average
Poor</td></tr>
<tr><td colspan="2">Optimal spatial scale:</td><td colspan="2">Regional averages (5-10 stations)</td></tr>
<tr><td colspan="2">Recommended impact applications:</td><td colspan="2">Any high-spatial resolution applications needing information about seasonal extremes, but which do not require spatially-correlated time series for multiple sites</td></tr>
</table>

Information on CCA method provided by:
Carlo Cacciamani
Rodica Tomozeiu
Antonella Morgillo
Valentina Pavan

ARPA-SIM_MLR Direct Method

Summary of the method

Direct 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)*
MLR	Station time series of seasonal extremes – temperature and precipitation	MSLP, Z500, T850hPa	Multiple Linear Regression (MLR)

STARDEX deliverables and papers

See ARPA –SIM contributions to D10, D12, D15 and D18.

Overall assessment of the MLR method

A relatively simple method for downscaling applied to station data. The predictors are the first four PCs of the MSLP, Z500 and T850. Performs well for temperature, but should be used with caution for precipitation.

Additional comments on the MLR method

MLR needs to test, before the construction of the model, the relationships between predictors and predictand.

Robustness criteria for the MLR method

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX assessments**
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes, see D10(ARPA-SIM)
	Are these relationships physically meaningful?	Yes, see D10(ARPA-SIM)
	If different methods or time periods are used for predictor selection, are similar sets of predictors obtained?	Similar sets of predictors have been obtained for some temperature indices but different sets of predictors for precipitation.

		Method tested: MLR against CCA Periods tested: 1958-1978+1994-2000
	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period?	Relatively insensitive - test made for temperature Non-standard periods tested: -calibration 1960-1989 -verification 1990-2000
	Is the statistical model performance sensitive to other user choices?	Relatively sensitive with the area selected for predictors
'Stationarity'	Predictors incorporated which are expected to change due to global warming? Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model? Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period? Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods? Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period and *vice versa*?	- Trends and low-frequency variability are well reproduced for mean temperature and precipitation and less for extremes (assessed for 1960-1990 period) No No No
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 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	Relatively uniform across stations, more variable across regions and seasons. See: D12(ARPA-SIM) BIAS, CORR, RMSE See: D12 (ARPA-SIM) Histograms

	- Q-Q diagrams - Taylor diagrams	
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - derived indices - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	Spatial patterns and temporal trends of predictors are well simulated. See: D13 ARPA-SIM Not much. See: D13 ARPA-SIM

Application criteria for the MLR *method*

MLR method provides:	Yes/No	Comments/Notes
Station-scale information	Yes	
Grid-box information	No	
European-wide information	No	
Daily time series	No	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	No	Could be implemented if method were applied to temperature also
Spatially consistent multi-site information	Yes	
Temporally consistent multi-site information	-	
Information at sites with no observations	-	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Medium	
Volume of data inputs	Medium	
Availability of input data	Medium	

Performance criteria for the MLR method

MLR method	Relative Performance Confidence		
	High	**Medium**	**Low**
Temperature			
Indices	Txav, Tnav	Txq90, Tnq10, Tnfd, Txhw90	
Seasons	Winter	Spring, Summer, Autumn	
Regions	Emilia-Romagna, Greece	Emilia-Romagna, Greece	
Precipitation			
Indices	Pav	Pxcdd	Pq90, px5d, pint, pxcdd, pfl90, pnl90
Seasons	Winter	Spring, Autumn	Summer
Regions	Emilia-Romagna, Greece	Emilia-Romagna, Greece	Emilia-Romagna, Greece

Overall performance:	
Mean temperature	Good
Temperature extremes	Average
Mean precipitation	Average
Precipitation extremes	Poor
Optimal spatial scale:	Regional averages (5-10 stations)
Recommended impact applications:	Any which require high-spatial resolution information about seasonal extremes, but do not require spatially-correlated time series for multiple sites

Information on MLR method provided by:
Carlo Cacciamani
Rodica Tomozeiu
Antonella Morgillo
Valentina Pavan

AUTH - ANN Direct Method

Summary of the method

Direct method	*Predictand(s)*	*Predictor(s)* *(See STARDEX Deliverable D10 for selection procedure)*	*Description* *(See STARDEX Deliverable D15 for details)*
AUTH-ANN	Seasonal indices of temperature/precipitation extremes	- 500hPa geopotential heights - 1000-500hPa thickness field	Downscaling model based on artificial neural networks

STARDEX deliverables and papers

- AUTH contributions to D10, D12, D15 and D18.
- Tolika K, Maheras P, Vafiadis M, Flocas HA, Arseni- Papadimitriou A (2005): Simulation of seasonal precipitation and raindays over Greece: a statistical downscaling technique based on artificial neural nets. Submitted for publication in *Climatic Change.*
- Kostopoulou E, Giannakopoulos C, Anagnostopoulou Chr, Tolika K, Maheras P, Vafiadis M (2005): Simulating Maximum and Minimum Temperatures over Greece: A comparison of three modeling techniques. Submitted for publication in *Climate Research.*

Overall assessment of the AUTH-ANN method

The ANN method captures some of the non-linear aspects of the circulation-local climate relationship. It proved to be skillful in representing relationships in the presence of noisy data. However, it requires too much computing time for daily data. It was also found weak in representing the observed variability of the data.

Robustness criteria for the AUTH-MLRct method

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX assessments**
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes
	Are these relationships physically meaningful?	Yes
	If different methods or time periods are used	

	for predictor selection, are similar sets of predictors obtained? Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period? Is the statistical model performance sensitive to other user choices?	Relatively different. See: Tolika et al. (2005) Relatively sensitive: Relatively sensitive/ User choices tested: different predictors.
‘Stationarity’	Predictors incorporated which are expected to change due to global warming? Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model? Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period? Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods? Calibrated the statistical model on a ‘cold’ period and validated it on a ‘warm’ period and *vice versa*?	Yes , see D13 Yes No No No
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 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 - Taylor diagrams	Relatively uniform across stations, more variable across regions and seasons. Averages and duration indices more uniform than intensity indices. See: D10-D12

Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. - Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	Generally well simulated see D13 Yes See: Tolika et al., 2005

Application criteria for the **AUTH-MLRct** *method*

AUTH-MLRct method provides:	**Yes/No**	**Comments/Notes**
Station-scale information	Yes	
Grid-box information	Yes	
European-wide information	Yes, potentially	
Daily time series	No	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	Yes	
Spatially consistent multi-site information	Yes	
Temporally consistent multi-site information	Yes	
Information at sites with no observations	No	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Low	
Volume of data inputs	Low	
Availability of input data	High	

Performance criteria for the AUTH-ANN method

AUTH-ANN method	Relative Performance Confidence		
	High	**Medium**	**Low**
Temperature			
Indices	Txav, tnav Txq90, tnq10, tnfd, txhw90		
Seasons	Winter, Spring,	Summer, Autumn	
Regions	w. Greece	e. Greece	
Precipitation			
Indices	Pav, Pxcdd, pint	Pq90, px5d	pfl90, pnl90
Seasons	Winter, Spring	Summer, Autumn	
Regions	w. Greece	e. Greece	

Overall performance:	
Mean temperature	Good
Temperature extremes	Good
Mean precipitation	Good
Precipitation extremes	Average
Optimal spatial scale:	Regional averages (22 stations)
Recommended impact applications:	

Information on AUTH-ANN method provided by
Panagiotis Maheras
Christina Anagnostopoulou
Konstantia Tolika

AUTH_CCA Direct Method

Summary of the method

Direct 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)*
AUTH_CCA	Seasonal indices of temperature/precipitation extremes	500hPa (precipitation) and thickness field 1000-500hPa (temperature)	Canonical Correlation Analysis

STARDEX deliverables and papers

- AUTH contributions to D10, D12
- Kostopoulou E, Giannakopoulos C, Anagnostopoulou Chr, Tolika K, Maheras P, Vafiadis M (2005): Simulating Maximum and Minimum Temperatures over Greece: A comparison of three modeling techniques. Submitted for publication in *Climate Research*

Overall assessment of the AUTH-CCA method

The AUTH-CCA method results were inferior to the other two AUTH methods and so it was not applied to the scenario data. (*the comments on the tables refer only to the control run data*)

Application criteria for the AUTH-CCA *method*

AUTH - CCA method provides:	Yes/No	Comments/Notes
Station-scale information	Yes	
Grid-box information	Yes	
European-wide information	No	
Daily time series	No	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	No	
Spatially consistent multi-site information	No	
Temporally consistent multi-site information	No	
Information at sites with no observations	No	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Low	
Volume of data inputs	Low	
Availability of input data	High	

Performance criteria for the AUTH-CCA method

AUTH-CCA method	**Relative Performance Confidence**		
	High	**Medium**	**Low**
Temperature			
Indices		Txq90, tnq10, Txav, tnav	tnfd, txhw90
Seasons		Winter Spring, Summer,	Autumn
Regions		W. Greece	E. Greece
Precipitation			
Indices		Pxcdd .Pav	Pq90, px5d, pint, pxcdd, pfl90, pnl90
Seasons		Winter Spring,	Summer Autumn
Regions		W. Greece	E. Greece
Overall performance:			
Mean temperature	Average		
Temperature extremes	Average		
Mean precipitation	Poor		
Precipitation extremes	Poor		
Optimal spatial scale:	Regional averages (22 stations)		
Recommended impact applications:			

Information on AUTH-CCA method provided by
Panagiotis Maheras
Christina Anagnostopoulou
Konstantia Tolika

AUTH - MLRct Direct Method

Summary of the method

Direct method	*Predictand(s)*	*Predictor(s)* *(See STARDEX Deliverable D10 for selection procedure)*	*Description* *(See STARDEX Deliverable D15 for details)*
AUTH-MLRct	Seasonal indices of temperature/precipitation extremes	**Primary Data** - 500hPa geopotential heights (precipitation) - 1000-500hPa thickness field (temperature) **Predictors** Daily calendar of 14 circulation types for the two data sets	Multiple Linear Regression based on a circulation type approach

STARDEX deliverables and papers

- AUTH contributions to D10, D12, D15 and D18.
- Maheras P, Tolika K, Anagnostopoulou Chr, Vafiadis M, Patrikas I, Flocas H (2004): On the Relationships between circulation types and changes in rainfall variability in Greece. *International Journal of Climatology* **24**: 1695-1712
- Kostopoulou E, Giannakopoulos C, Anagnostopoulou Chr, Tolika K, Maheras P, Vafiadis M (2005): Simulating Maximum and Minimum Temperatures over Greece: A comparison of three modeling techniques. Submitted for publication in *Climate Research*

Overall assessment of the AUTH-MLRct method

The MLRct method provides a strong, relative stable predictor-predictant relationship. It also provides a strong signal in predictor change and gives more physical meaning into the relationships than a purely statistical approach. However, it requires a different classification for every time period and for every study region. It was also found weak in representing the observed variability of the data.

Additional comments on the AUTH-MLRct method

The classification of the circulation types is easy applicable to any region but it proved to be very sensitive to the choice of the study period

Robustness criteria for the AUTH-MLRct method

Robustness criteria	Key questions and recommended assessment methods	STARDEX assessments
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes, high correlation coefficients for all seasons, for both temperature and precipitation indices (except autumn precipitation indices).
	Are these relationships physically meaningful?	Yes, supported by Maheras et al., 2004; Kostopoulou et al., 2005
	If different methods or time periods are used for predictor selection, are similar sets of predictors obtained?	Yes Methods tested: Cross validation Period tested: 1958-2000 (1 year step)
	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period?	Relatively sensitive
	Is the statistical model performance sensitive to other user choices?	Relatively sensitive User choices tested: cross validation (Maheras et al., 2004)
'Stationarity'	Predictors incorporated which are expected to change due to global warming?	Yes, the method is based on a circulation type approach and the circulation types are sensitive to global warming (changes in their frequency) see D13
	Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model?	Yes
	Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period?	No
	Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods?	Yes
	Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period	No

	and *vice versa*?	
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 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 - Taylor diagrams	Relatively uniform across stations, more variable across regions and seasons. Averages and duration indices more uniform than intensity indices. See: D10, D11,D12
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	Generally well simulated with the exception of the frequencies of the a small number of circulation types Yes See: D10 – D12

Application criteria for the* AUTH-MLRct *method

AUTH-MLRct method provides:	**Yes/No**	**Comments/Notes**
Station-scale information	Yes	
Grid-box information	Yes	
European-wide information	No	
Daily time series	No	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	Yes	
Spatially consistent multi-site information	Yes	
Temporally consistent multi-site information	Yes	
Information at sites with no observations	No	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Low	
Volume of data inputs	Low	
Availability of input data	High	

Performance criteria for the AUTH-MLRct method

AUTH-MLRct method	Relative Performance Confidence		
	High	**Medium**	**Low**
Temperature			
Indices	Txav, tnav Txq90, tnq10, tnfd	txhw90	
Seasons	Winter Spring, Summer, Autumn		
Regions	W. Greece, E. Greece		
Precipitation			
Indices	Pav, pxcdd, pint	Pq90, px5d,	pfl90, pnl90
Seasons	Winter Spring	Summer	Autumn
Regions	W. Greece,	E. Greece	
Overall performance:			
Mean temperature	Good		
Temperature extremes	Good		
Mean precipitation	Good		
Precipitation extremes	Average		
Optimal spatial scale:	Regional averages (22 stations)		
Recommended impact applications:			

Information on AUTH-MLRct method provided by
Panagiotis Maheras
Christina Anagnostopoulou
Konstantia Tolika

DMI_CWG Indirect Method

Conditional weather generator, conditional on quantile values of a circulation index, in which precipitation occurrence and amount are modelled separately.

Direct 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)*
DMI_CWG	Daily Precipitation	Circulation index calculated from MSLP	Conditional weather generator

STARDEX deliverables and papers
See DMI contributions to D12, D15 and D18.

Goodess, C.M., Anagnostopoulou, C., Bárdossy, A., Haylock, M.R., Hundecha, Y., Maheras, P., Ribalaygua, J., Schmidli, J., Schmith, T. and Tomozeiu, R.: An intercomparison of statistical downscaling methods for Europe and European regions – assessing their performance with respect to extreme temperature and precipitation events. *Submitted to Climatic Change.*

Schmidli, J., Haylock, M., Hundecha, Y., Schmith, T. and Ribalaygua, J.: Statistical and Dynamical Downscaling of Precipitation: Evaluation, Intercomparison, and Scenarios for the European Alps. *In preparation.*

Overall assessment of the DMI_CWG method
A simple method with one predictor suggests robustness. It turns out to be best in winter and near the Atlantic and for simple statistics.

Robustness criteria for the DMI_CWG method

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX assessments**
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes, the objectively identified MSLP-pattern is statistically significant
	Are these relationships physically meaningful?	Yes, the pattern identified for a particular station often corresponds to a lower pressure near the stations when precipitation occurs.
	If different methods or time periods are used for predictor selection, are similar sets of	Very similar

	predictors obtained? Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period? Is the statistical model performance sensitive to other user choices?	Periods tested: 1958-1980/1981-2000 Relatively insensitive Non-standard periods tested:1958-1980/1981-2000 Not tested
'Stationarity'	Predictors incorporated which are expected to change due to global warming? Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model? Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period? Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods? Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period and *vice versa*?	Yes (daily MSLP), which is expected to change due to shift in storm tracks etc. Yes: The overall shape of the scenario-control climate is conistent with an increase in the NAO, which is also obtained directly from the GCMs No No No
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 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 (x)	Best performance in Atlantic-influenced regions and during winter. Averages and occurrence indices perform better than magnitude. See: D12 European report. Mainly CORR

	- Histograms (x) - Box-whisker plots - Q-Q diagrams - Taylor diagrams	
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - derived indices - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	MSLP turned out to be more useful as a predictor than vorticity, see D13 (AMI), Summary, ETH central analysis). Not evaluated

Application criteria for the DMI_CWG *method*

DMI_CWG method provides:	**Yes/No**	**Comments/Notes**
Station-scale information	Yes	
Grid-box information	No	
European-wide information	Yes	
Daily time series	Yes	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	No	Could be implemented if method were applied to temperature also
Spatially consistent multi-site information	No	
Temporally consistent multi-site information	No	
Information at sites with no observations	No	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Low	
Volume of data inputs	Low	
Availability of input data	High	

Performance criteria for the DMI_CWG method

<table>
<tr><td rowspan="2">DMI_CWG method</td><td colspan="3">Relative Performance Confidence</td></tr>
<tr><td>High</td><td>Medium</td><td>Low</td></tr>
<tr><td>Temperature
Indices
Seasons
Regions
Precipitation
Indices

Seasons
Regions</td><td>
-
-
-

Pav

Winter
Atlantic influenced</td><td>
-
-
-

Pxcdd</td><td>
-
-
-

Pq90, px5d, pint, pxcdd,
pfl90, pnl90
Summer
Easten
Mediterranean/Europe</td></tr>
<tr><td colspan="2">Overall performance:
Mean temperature
Temperature extremes
Mean precipitation
Precipitation extremes</td><td colspan="2">
-
-
Average
Poor</td></tr>
<tr><td colspan="2">Optimal spatial scale:</td><td colspan="2">Regional averages (5-10 stations)</td></tr>
<tr><td colspan="2">Recommended impact applications:</td><td colspan="2">Any which require European-wide information about winter extremes, but do not require spatially-correlated time series for multiple sites</td></tr>
</table>

Information on DMI_CWG method provided by
Torben Schmith
Bo Christiansen

ETH_LOCI Direct Method

Summary of the method

Direct 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)***
ETH_LOCI	daily precipitation (station or grid point data)	GCM precipitation	Local rescaling of GCM simulated precipitation

STARDEX deliverables and papers
See ETH contributions to D12, D13, D15 and D18.

Schmidli, J., Frei, C., and Vidale, P.L, 2005a: Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology*, accepted.

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

Overall assessment of the ETH_LOCI method
A relatively simple method for downscaling precipitation. Performance depends on the quality of the GCM simulated precipitation with respect to temporal variations.

Robustness criteria for the ETH_LOCI method

Robustness criteria	Key questions and recommended assessment methods	STARDEX assessments
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes, Schmidli et al., 2005a.
	Are these relationships physically meaningful?	Yes, Schmidli et al., 2005a.
	If different methods or time periods are used for predictor selection, are similar sets of predictors obtained?	NA
	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period?	No
	Is the statistical model performance sensitive to other user choices?	No
'Stationarity'	Predictors incorporated which are expected to change due to global warming?	Yes, GCM precipitation.
	Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model?	NA
	Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period?	NA
	Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods?	No
	Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period and *vice versa*?	No

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 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	Variable across regions and seasons. (see Schmidli et al. 2005b)
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - derived indices - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	Quality of GCM predictors is comparable to reanalysis predictors (see D13 ETH partner report). Depends on GCM and reanalysis used.

Application criteria for the ETH_LOCI method

ETH_LOCI method provides:	**Yes/No**	**Comments/Notes**
Station-scale information	Yes	
Grid-box information	Yes	
European-wide information	Potentially	Depends on available observations
Daily time series	Yes	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	NA	Method not tested for temperature
Spatially consistent multi-site information	Yes	
Temporally consistent multi-site information	Yes	
Information at sites with no observations	No	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Low	
Volume of data inputs	Low	
Availability of input data	High	

Performance criteria for the ETH_LOCI method (applied to ERA40)

ETH_LOCI method	**Relative Performance Confidence**		
	High	**Medium**	**Low**
Temperature			
Indices	-	-	-
Seasons	-	-	-
Regions	-	-	-
Precipitation			
Indices	pfre, pav	pxcdd, pint	pq90, px5d
Seasons	Winter	Spring, Autumn	Summer
Regions	-	-	-
Overall performance:			
Mean temperature	-		
Temperature extremes	-		
Mean precipitation	Good		
Precipitation extremes	Average		
Optimal spatial scale:	Grid box and larger scales		
Recommended impact applications:	Any which require high-spatial resolution information about seasonal extremes		

Information on ETH_LOCI method provided by
Jürg Schmidli

FIC_ANAL2 Indirect Method

Summary of the method

Direct 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)*
FIC_ANAL2	Daily Precipitation and Temperature	Geostrophic fluxes at 1000 & 500 hPa, low tropospheric thickness (and sine of the day of the year and 10 previous days temperature)	Two-step analogue method, in which (1) the 'n' most similar 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

STARDEX deliverables and papers
See FIC contributions to D10, D12, D15 and D18.

Overall assessment of the FIC_ANAL2 method
A robust and physically meaningful method for downscaling to an unlimited number of stations or grid points, adapted to the whole of Europe. Performs very well for temperature, but should be used with caution for precipitation.

Robustness criteria for the FIC_ANAL2 method

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX assessments**
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes, predictors are selected under theoretical considerations, looking for predictors that are physical forcings of the predictands, see D10 (FIC)
	Are these relationships physically meaningful?	Yes, enhanced using time and spatial scales as similar as possible to physical processes, see D10 (FIC)
	If different methods or time periods are used for predictor selection, are similar sets of predictors obtained?	Selection under theoretical considerations
	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to	Relatively insensitive (tested in previous versions of the method)

	changes in calibration/validation period? Is the statistical model performance sensitive to other user choices?	 Not tested
'Stationarity'	Predictors incorporated which are expected to change due to global warming? Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model? Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period? Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods? Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period and *vice versa*?	Yes (especially thickness and 10 previous days temperature), based on theoretical considerations:, see D10 (FIC) Yes: observed trends, in previous versions of the method; low-frequency variability in D12 (FIC) Yes, in previous versions of the method, in progress within STARDEX No: it is assumed that these relationships don't change, because predictors are physical forcings of the predictands, see D10 (FIC) Yes, in previous versions of the method,
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 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	Relatively uniform across stations (spatial coherence), more variable across regions and seasons. Averages and duration indices more uniform than intensity indices. Temperature downscaled better and more uniformly than precipitation. See D12 (FIC): Yes Yes Yes Yes Yes Yes No No

	- Taylor diagrams	No
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	Generally well simulated , see D13 (Summary, ETH central analysis). Theoretical selection of predictors that are well simulated by GCMs Yes. See: D14 results (FIC)

Application criteria for the FIC_ANAL2 *method*

FIC_ANAL2 method provides:	Yes/No	Comments/Notes
Station-scale information	Yes	
Grid-box information	Yes	
European-wide information	Yes	
Daily time series	Yes	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	Yes	Tested in version used for operational forecasting: consistency due to the use of the same analogous days for both variables estimation
Spatially consistent multi-site information	Yes	
Temporally consistent multi-site information	Yes	
Information at sites with no observations	No	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Low	
Volume of data inputs	Low	
Availability of input data	Medium	

Performance criteria for the FIC_ANAL2 method

FIC_ANAL2 method	Relative Performance Confidence		
	High	**Medium**	**Low**
Temperature			
Indices	Txav, tnav ,Txq90, tnq10, tnfd	txhw90	
Seasons	Winter, Spring, Autumn	Summer	
Regions	Europe		
Precipitation			
Indices	Pav	Pxcdd, pnl90	Pq90, px5d, pint, pxcdd, pfl90
Seasons	Winter	Spring, Autumn	Summer
Regions	NW Europe	SW Europe	E Europe
Overall performance:			
Mean temperature	Good		
Temperature extremes	Good		
Mean precipitation	Average		
Precipitation extremes	Poor		
Optimal spatial scale:	Single sites		
Recommended impact applications:	Very good performance for daily series or extremes of temperature. Any which require high-spatial resolution information about daily series or seasonal extremes, although it should be used with caution for precipitation.		

Information on FIC_ANAL2 method provided by
Jaime Ribalaygua
Luis Torres

UNIBE_CCA Direct Method

Summary of the method

Direct 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)*
UNIBE_CCA	Seasonal indices of temperature extremes	Best performing predictor field(s) selected from: 6 different large-scale sectors containing the study area; seasonal means of SLP, and GPH, T, SH, RH, at 300, 500, 700 & 850 hPa.	Canonical Correlation Analysis

STARDEX deliverables and papers
See UNIBE contributions to D12, D14, D15, D17 and D18.

Gyalistras, D. and Schuepbach, E.: 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.

Overall assessment of the UNIBE_CCA method
Strengths: Relatively simple and flexible method, modest input data requirements (seasonal fields), yields generally plausible results (interpretation of canonical map pairs), is little sensitive to biases in GCM-simulated fields, provides site-specific scenarios.

Weaknesses: Does not perform well for "sophisticated" seasonal statistics such as number of frost days and heat wave duration.

Additional comments on the UNIBE_CCA method
High sensitivity of downscaled scenarios to choice of predictor(s).

Robustness criteria for the UNIBE_CCA method

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX assessments**
'Strength and stability'	Can strong predictor/predictand relationships be identified? Are these relationships physically meaningful? If different methods or time periods are used for predictor selection, are similar sets of predictors obtained?	Yes, see deliverable D12 by UNIBE Yes, see deliverable D12 by UNIBE Yes, see deliverable D12 by UNIBE

	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period? Is the statistical model performance sensitive to other user choices?	No, see deliverable D12 by UNIBE There is some sensitivity to the choice of the number of EOFs used for CCA (see, e.g., Gyalistras et al. 1994, Clim. Res. 4(3): 167-189)
'Stationarity'	Predictors incorporated which are expected to change due to global warming? Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model? Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period? Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods? Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period and *vice versa*?	Yes, in particular large-scale temperature Yes, work in progress No, but changes in GCM-simulated temperature fields for the late 21st century are generally known to lie outside the observed 20th century variability range. No No
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 using: - BIAS (mean difference between simulated and observed values) - CORR (Spearman rank-correlation coefficient) - RMSE (Root Mean Square Error) - Reduction of Variance - ratio of observed : simulated standard deviations Plotted using: - Maps - Histograms - Box-whisker plots - Q-Q diagrams	 medium small large large large Yes Yes Yes Yes No Yes Yes No No

	- Taylor diagrams	No
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - derived indices - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	No such analyses carried out by UNIBE – but see D13 No significant degradation detected

Application criteria for the **UNIBE_CCA** *method*

UNIBE_CCA method provides:	**Yes/No**	**Comments/Notes**
Station-scale information	Yes	
Grid-box information	No	
European-wide information	No	Could be implemented. Use of different predictor sectors for different regions recommended (e.g., shifting spatial window for predictor fields).
Daily time series	No	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	Yes	Could be implemented if method were applied to precipitation also. See, e.g., Gyalistras et al. 1994, Clim. Res. 4(3): 167-189
Spatially consistent multi-site information	Yes	
Temporally consistent multi-site information	Yes	Note the reduction in temporal variance due to the use of a regression approach. See, e.g., von Storch 1999, J. Clim. 12(12): 3505-3506.
Information at sites with no observations	No	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Intermediate	Model fitting requires computation of eigenvectors for a large number of predictor variables.
Volume of data inputs	Low	(as compared to "direct" methods)
Availability of input data	Low	(as compared to "direct" methods)

Performance criteria for the UNIBE_CCA method

UNIBE_CCA method	Relative Performance Confidence		
	High	**Medium**	**Low**
Temperature			
Indices	txav, tnav	txq90, tnq10	tnfd, txhwd
Seasons	Winter, Spring, Summer, Autumn		
Regions	N-Alps	(S- and E-Alps)	
Precipitation	—	—	—
Indices	—	—	—
Seasons	—	—	—
Regions	—	—	—
Overall performance:			
Mean temperature	Good		
Temperature extremes	Average to poor		
Mean precipitation	—		
Precipitation extremes	—		
Optimal spatial scale:	Not known		
Recommended impact applications:	Useful for a spatially consistent, first-order assessment of likely changes in seasonal, site-specific temperature statistics. Can be combined with stochastic weather generators to produce monthly, daily or hourly weather data, see, e.g., Gyalistras & Fischlin 1999, Petermanns geogr. Mitt. 143(4): 251-264.		

Information on UNIBE_CCA method provided by
Dimitrios Gyalistras, 10-Aug-2005

USTUTT_MAR Indirect Method

Summary of the method

Indirect 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)*
USTUTT_MAR	Daily series of precipitation and temperature	Objective circulation patterns of MSLP and eastward moisture flux for precipitation HGT and the corresponding Objective circulation patterns (CPs) of the 700 hPa for temperature	Multivariate Autoregressive model

STARDEX deliverables and papers
See USTUTT contributions to D10, D12, D15 and D18.

Overall assessment of the USTUTT_MAR method
A stochastic method for downscaling of daily series of precipitation and temperature at stations or grids by maintaining the spatial covariance structure in particular European regions. Performs relatively better for temperature indices and mean precipitation. For certain indices of extremes of precipitation, it should be used with caution in seasons other than Winter and Spring.

Robustness criteria for the USTUTT_MAR method

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX assessments**
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes, supported by high correlation values and the ability of the classified CPs to explain the variability of precipitation or temperature, see D10(USTUTT)
	Are these relationships physically meaningful?	Yes, supported by literature review.
	If different methods or time periods are used for predictor selection, are similar sets of predictors obtained?	Not tested

	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period? Is the statistical model performance sensitive to other user choices?	Relatively sensitive Non-standard periods tested: Calibrated for 'warm' period and validated for 'cold' period and vice versa Not tested
'Stationarity'	Predictors incorporated which are expected to change due to global warming? Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model? Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period? Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods? Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period and *vice versa*?	Yes (e.g. moisture flux for precipitation), based on literature review. No No No Yes. Calibration during 'cold' periods shows a better prospect of extrapolating to the warmer period for precipitation related indices. A similar tendency is noted for temperature indices in Winter and Spring.
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 using: - BIAS (mean difference between simulated and observed values)	Relatively uniform across stations and between seasons as well as indices for temperature indices. Considerable variability across stations and seasons for precipitation indices. Better performance for mean precipitation than extremes. The occurrence of the extreme is better reproduced than the magnitude of the extremes. The best performance is obtained in winter and Spring, while the worst is in summer. Model shows relatively higher bias for temperature indices. It shows a slight negative bias for precipitation indices except in autumn where it shows a positive bias in the mean

		precipitation. For temperature indices, it generally shows overestimation.
	- 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	Performance as discussed above. Performance as discussed above Generally, interannual variability of indices is underestimated by the model. Histograms
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - derived indices - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	Generally well simulated. The frequency of circulation patterns associated with wet situations is slightly overestimated in all seasons except in summer where the opposite situation is noted, while the persistence of the CPs is generally slightly overestimated in all seasons. The moisture flux is overestimated over northwestern Europe in winter and underestimated in the Southern part of Europe in summer. See D13 (USTUTT, Summary, ETH central analysis). Not much. In winter, indices of precipitation related to magnitude are estimated with higher bias when predictors are taken from GCM. Underestimation of the variability of the indices is more or less the same in both cases. Temperature indices are estimated with a relatively higher bias when GCM predictors are used. The variability of the mean and the extreme maximum temperature are overestimated when GCM predictors are used. For the other indices, the variability is underestimated in both cases by similar extent.

Application criteria for the USTUTT_MAR *method*

USTUTT_MAR method provides:	**Yes/No**	**Comments/Notes**
Station-scale information	Yes	
Grid-box information	Yes	
European-wide information	Not verified	
Daily time series	Yes	
Seasonal indices of extremes	Yes	Calculated from the daily series
Temporally consistent temperature and precipitation	Yes	
Spatially consistent multi-site information	Yes	
Temporally consistent multi-site information	Yes	
Information at sites with no observations	Yes	Computed from the spatial-temporal structure
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Relatively high	
Volume of data inputs	Relatively high	
Availability of input data	High	

Performance criteria for the USTUTT_MAR method

USTUTT_MAR method	Relative Performance Confidence		
	High	**Medium**	**Low**
Temperature			
Indices	Txav, tnav, txq90, tnq10	tnfd, txhw90	
Seasons	Winter, Spring, Summer	Autumn	
Regions	German Rhine		
Precipitation			
Indices	Pav	Pxcdd, px5d, pnl90	Pq90, pint, pfl90
Seasons	Winter, Spring	Autumn	Summer
Regions	German Rhine	Alps	

Overall performance:	
Mean temperature	Good
Temperature extremes	Good/Average
Mean precipitation	Good
Precipitation extremes	Poor
Optimal spatial scale:	Performs better when indices are calculated from aggregated precipitation at a catchment scale.
Recommended impact applications:	Any which require high-spatial resolution information of spatially-correlated time series for multiple sites.

Information on USTUTT_MAR method provided by

András Bárdossy
Yeshewatesfa Hundecha

USTUTT_MLR Direct Method

Summary of the method

Direct 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)*
USTUTT_MLR	Seasonal indices of temperature and precipitation extremes	GPH, RH, T, VORT, DIVG at 500, 700 & 850 hPa, Eastward moisture flux at 700 hPa, Frequency of CPs of MSLP	Multiple Linear Regression

STARDEX deliverables and papers
See USTUTT contributions to D10, D12, D15 and D18.

Overall assessment of the USTUTT_MLR method
A simple method for downscaling of indices of extremes at stations or grids in particular European regions. Performs well for temperature indices and mean precipitation. For certain indices of extreme of precipitation, it should be used with caution in seasons other than winter.

Additional comments on the USTUTT_MLR method
Seasonal measures of mean and extremes of predictors are used. As the measures of extremes, the seasonal 90th and 10th percentile values of the predictors are used. A stepwise regression is used to select predictors from among the potential ones.

Robustness criteria for the USTUTT_MLR method

Robustness criteria	**Key questions and recommended assessment methods**	**STARDEX assessments**
'Strength and stability'	Can strong predictor/predictand relationships be identified?	Yes, specially the humidity related predictors show high correlation with the extreme indices, see D10(USTUTT)
	Are these relationships physically meaningful?	Yes, supported by literature review.
	If different methods or time periods are used for predictor selection, are similar sets of predictors obtained?	Relatively similar Methods tested: Stepwise regression Periods tested: years arranged according to annual mean temp. and selection done for the first and last third years separately.

	Is the strength of the predictor/predictand relationships and/or the performance of the statistical downscaling model sensitive to changes in calibration/validation period? Is the statistical model performance sensitive to other user choices?	Relatively sensitive Non-standard periods tested: Calibrated for 'warm' period and validated for 'cold' period and vice versa Not tested
'Stationarity'	Predictors incorporated which are expected to change due to global warming? Assessed whether the direction and magnitude of observed trends in the predictand, together with low-frequency variability, are reproduced by the statistical model? Assessed whether the projected changes in predictor variables lie outside the range of variability observed over the calibration/validation period? Assessed whether predictor/predictand relationships calculated from GCM/RCM output change between the control and perturbed periods? Calibrated the statistical model on a 'cold' period and validated it on a 'warm' period and *vice versa*?	Yes (e.g., Moisture flux), based on literature review. No No No Yes. Except in summer, the model shows a potential for extrapolation of the indices related to the magnitude and frequency of heavy precipitation to a warmer climate. For temperature related indices, it shows a similar tendency in winter and autumn. In Summer, the model shows a better prospect for extrapolation to a colder climate.
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 using: - BIAS (mean difference between simulated and observed values) - CORR (Spearman rank-	Relatively uniform across stations and between seasons as well as indices for temperature indices. Considerable variability across stations and seasons for precipitation indices. Better performance for mean precipitation than extremes. The occurrence of the extremes is better reproduced than the magnitude of the extremes. The best performance is obtained in winter and the worst in summer. Generally, the model slightly underestimates the mean and extreme precipitation. Performance as discussed above.

	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	Performance as discussed above. Generally, interannual variability of indices is underestimated by the model. Histograms
Reliability of simulation of predictors	Predictors calculated from climate model output compared with those calculated from Reanalysis data, taking into consideration: - raw values - derived indices - spatial patterns - temporal trends - frequency and persistence, and day-to-day transitions, of circulation/weather types. Is performance of the statistical model for the control period degraded when predictors are taken from climate model output rather than Reanalysis data?	Generally well simulated except that the moisture flux is overestimated over northwestern Europe in winter and underestimated in the Southern part of Europe in summer, see D13 (USTUTT, Summary, ETH central analysis). Yes, Higher bias is obtained when predictors are taken from GCM.

Application criteria for the USTUTT_**MLR** ***method***

USTUTT_MLR method provides:	**Yes/No**	**Comments/Notes**
Station-scale information	Yes	
Grid-box information	Yes	
European-wide information	Not tested	
Daily time series	No	
Seasonal indices of extremes	Yes	
Temporally consistent temperature and precipitation	-	
Spatially consistent multi-site information	-	
Temporally consistent multi-site information	-	
Information at sites with no observations	No	
Method requirements :	**Relatively high/medium/low**	**Comments/Notes**
Computing resources	Low	
Volume of data inputs	Relatively high	
Availability of input data	High	

Performance criteria for the USTUTT_MLR method

<table>
<tr><td rowspan="2">USTUTT_MLR method</td><td colspan="3">Relative Performance Confidence</td></tr>
<tr><td>High</td><td>Medium</td><td>Low</td></tr>
<tr><td>Temperature</td><td></td><td></td><td></td></tr>
<tr><td>Indices</td><td>Txav, tnav,txq90, tnq10, tnfd</td><td>Txhw90</td><td></td></tr>
<tr><td>Seasons</td><td>Winter, Spring, Summer, Autumn</td><td></td><td></td></tr>
<tr><td>Regions</td><td>German Rhine</td><td></td><td></td></tr>
<tr><td>Precipitation</td><td></td><td></td><td></td></tr>
<tr><td>Indices</td><td>Pav</td><td>Pxcdd, px5d, pnl90</td><td>Pq90, pint, pfl90</td></tr>
<tr><td>Seasons</td><td>Winter, Spring</td><td>Autumn</td><td>Summer</td></tr>
<tr><td>Regions</td><td>German Rhine</td><td>Alps</td><td></td></tr>
<tr><td colspan="2">Overall performance:</td><td colspan="2"></td></tr>
<tr><td colspan="2">Mean temperature</td><td colspan="2">Good</td></tr>
<tr><td colspan="2">Temperature extremes</td><td colspan="2">Good</td></tr>
<tr><td colspan="2">Mean precipitation</td><td colspan="2">Good</td></tr>
<tr><td colspan="2">Precipitation extremes</td><td colspan="2">Poor</td></tr>
<tr><td colspan="2">Optimal spatial scale:</td><td colspan="2">Station scale</td></tr>
<tr><td colspan="2">Recommended impact applications:</td><td colspan="2">Any which require high-spatial resolution information about seasonal extremes, but do not require spatially-correlated time series for multiple sites</td></tr>
</table>

Information on USTUTT_MLR method provided by

Yeshewatesfa Hundecha
András Bárdossy