

An Uncertainty Analysis of Monthly Temperature and Precipitation Scenarios for Switzerland

Dimitrios Gyalistras¹

September 2002

1. INTRODUCTION

During the course of the 20th century significant climatic shifts have been observed in the European Alps and more specifically in Switzerland (e.g., BÖHM *et al.*, 2001; JUNGO & BENISTON, 2001; WIDMANN & SCHÄR, 1997; SCHMIDLI *et al.*, 2002). A number of future climate scenarios for the 21st century have also been derived for the Alpine and Swiss regions, based on a variety of approaches. Reviews of the scenarios can be found in GYALISTRAS *et al.* (1998) and GYALISTRAS (2000). These studies suggest that the Alps could experience a much larger warming than the globe on average, and that there is a large potential for dramatic changes also with regard to precipitation.

The assessment of possible climatic changes for any particular region involves many uncertainties. These relate to (1) the natural variability of climate; (2) the unknown future forcings of the climate system, in particular the man-made forcings; (3) the complexity of the climate system and the limitations of the global models that are used to assess its response to any given future forcing scenarios; and finally (4) the need to project climatic changes at a spatial scale which is either not resolved, or not correctly represented in the global climate models (problem of regionalization or "downscaling").

The role of (1) has been studied in the context of climate change projections by GIORGI & FRANSISCO (2000a, b). They investigated the changes simulated by different coupled atmosphere-ocean General Circulation Models (AO-GCMs) between two 30-year periods that reflected present-day and possible future (end of the 21st century) climatic conditions and found that for most regions the projected changes were relatively insensitive to sampling errors due to the simulated natural variability of climate.

The importance of (2) and (3) was also assessed by GIORGI & FRANSISCO (2000a, b). The choice of future forcing was found to be generally more important than natural variability, but in the end the differences in the projected regional-scale changes were dominated by the choice of GCM. Large inter-model variability of projected changes for Europe was also found in further studies by HULME & CARTER (2000) and GIORGI & MEARNES (2002).

¹ Climatology and Meteorology, Institute of Geography, University of Berne, Hallerstr. 12, CH-3012 Berne, Switzerland

GYALISTRAS *et al.* (1998) and GYALISTRAS (2000) concluded from an analysis of regional climate change scenarios for the Alps that for this region the regionalization uncertainties (4) are at least as important as the choice of GCM. However, the scenario comparisons were severely hampered by methodical problems, such as the consideration of too small time windows to assess present-day and future climates in the various regionalization studies, or the use of widely varying definitions of the climate change signal across studies.

A further incentive for re-considering Alpine scenarios is that since the IPCC Third Assessment Report (TAR) a new vintage of fully transient climate change scenarios has become available (an overview is given in CUBASCH & MEEHL, 2002). The Alps are located exactly at the border of the "Northern Europe" and "Southern Europe" (or "Mediterranean Basin") regions that have been used in the IPCC TAR and other studies (e.g., KITTEL *et al.*, 1997; GIORGI & FRANSISCO 2000a, b; GIORGI & MEARNES, 2002) such that the available analyses are of limited value with regard to the Swiss region. Moreover, to my knowledge, the new vintage of scenarios has not been analyzed in detail with regard to Switzerland until now.

The present work aims at assessing the relative importance of the choice of forcing, GCM and regionalization procedure with regard to future climate change in Switzerland. The study uses results from several GCM-simulations and derives spatially highly resolved scenarios based on a new, ~5 km gridded data set for the monthly mean Swiss temperature and precipitation. The next sections describe the data and methods used to derive and analyze the regional scenarios, followed by a presentation of the results and a brief discussion. The paper ends with some concluding remarks.

2. DATA & METHODS

The investigation of a larger number of forcings and GCMs implied the use of relatively simple regionalization methods. Two such methods were considered: The first consisted in directly evaluating the simulated changes at GCM gridpoints that were located in the vicinity of the Alps. In the following this method is called "GP". The second method was the statistical downscaling ("SD") method proposed by GYALISTRAS *et al.* (1994). This method first establishes a statistical relationship between large-scale and regional climate variables, which then is applied to GCM output in order to infer possible changes in the regional climate.

2.1 Observed Data

The fitting and validation of the statistical downscaling procedure used large-scale and regional monthly mean data from the period 1951-1999.

The large-scale data were (i) the monthly mean near-surface temperature (NST) data set of JONES *et al.* (2001), and (ii) the monthly mean sea-level pressure (SLP) data set by TRENBERTH & PAULINO (1980, updated). Both data sets were provided on a 5° x 5° longitude-latitude grid. The SLP data were interpolated to the same grid as the NST data using procedures from the PINGO software package (WASZKEWITZ *et al.*, 1996). For both fields

was considered a sector containing 17 x 9 gridpoints over the North Atlantic and Europe that extended from 37.5° E to 42.5° W, respectively from 32.5° N to 72.5° N.

The regional data were monthly mean temperature and precipitation anomalies (ΔT , respectively ΔP) on a 5 km x 5 km gridd covering the entire area of Switzerland (~41'000 km², elevation range 200 to over 4'500 m.a.s.l.). A detailed description of the method used to produce the dataset and an assessment of its accuracy are given in GYALISTRAS (2002).

The data set was derived based on data from 35 long-term climate stations for ΔT and 147 stations for ΔP which were irregularly distributed over the Swiss area. It provides point estimates for the monthly means and accounts for the influence of the relief in a sophisticated manner. The interannual variability of Swiss climate at individual gridpoints is very accurately reproduced. The spatial pattern of local long-term temporal statistics, in particular of long-term trends is less well reproduced, but averages over a larger number of gridpoints are again trustworthy. The largest systematic errors and error variances were estimated to occur at locations above ~2000 m.a.s.l (cf. Fig. 1) and in data sparse areas in S/SE Switzerland. The results obtained for these regions should be interpreted with caution (GYALISTRAS, 2002).

For the present study the gridded data set was interpolated for technical reasons to a 0.05° x 0.05° (ca. 3.9 km x 5.6 km) grid. The gridpoint values of the new grid were determined from the respective closest nine gridpoints of the original data set. This was done with the aid of a partial quartic equation that defined a surface that passed exactly through all nine neighbouring gridpoint values (ZEVENBERGEN & THORNE, 1987).

2.2 GCM Data

All needed GCM data were obtained from the IPCC Data Distribution Center (IPCC-DDC, <http://ipcc-ddc.cru.uea.ac.uk>). Data for the following variables were downloaded: NST, SLP and monthly mean total precipitation (PRECIP). Used were results from 18 simulations with 7 different GCMs that were driven with four different forcing scenarios (Table 1).

The choice of the simulations was mainly guided by data availability. An additional requirement was that per GCM at least two different forcing scenarios should be available. Data from further state-of-the-art GCMs, such as the GFDL R30 model (DELWORTH *et al.*, 2002), could in principle have been obtained directly from the various modelling centers. However, each center has its own storage and downloading conventions, and the additional effort that would have been needed for the preparation of these data sets was not possible in the context of the present study.

For some of the models shown in Table 1 results from older model versions would also have been available from the IPCC-DDC. The assessment of differences in the regional scenarios due to different versions of the individual GCMs was however again beyond the scope of the present study. The only exception was the use of two model versions of the CGCM-model (Nos. 2 and 4 in Table 1) since these results were published quite recently and in quick succession.

No.	GCM Name	Resolution	GG	GS	A2	B2	References
1	CCSR/NIES1	A: T21 (~5.6° x 5.6°) L20 O: 2.8° x 2.8° L17	x	x	-	-	EMORI <i>et al.</i> (1999)
2	CGCM1	A: T32 (~3.75° x 3.75°) L10 O: 1.8° x 1.8° L29	x	x	-	-	BOER <i>et al.</i> (2000)
3	ECHAM4/OPYC3	A: T42 (~2.8° x 2.8°) L19 O: 2.8° x 2.8° L11	x	x	-	-	ROECKNER <i>et al.</i> (1999)
4	CGCM2	A: T32 (~3.75° x 3.75°) L10 O: 1.8° x 1.8° L29	-	-	x	x	FLATO & BOER (2001)
5	DOE PCM	A: T42 (~2.8° x 2.8°) L18 O: 0.66° x 0.66° L32	-	-	x	x	WASHINGTON <i>et al.</i> (2000)
6	CSIRO	A: R21 (~3.2° x 5.6°) L9 O: R21 L21	x	x	x	x	GORDON & O'FARRELL (1997)
7	HadCM3	A: 2.5° x 3.75° L19 O: 1.25° x 1.25° L20	x	x	x	x	JOHNS <i>et al.</i> (2001)

Table 1: Overview of GCM simulations used. CCSR/NIES = Center for Climate System Research, University of Tokyo/National Institute for Environmental Studies (Japan); CGCM1/2 = Canadian Global Coupled Model 1/2; ECHAM4/OPYC3 = ECMWF Model, modified in Hamburg/Isopycnal Ocean Model (Germany); DOE PCM = Department of Energy Parallel Climate Model (Boulder/Colorado, USA); CSIRO = Commonwealth Scientific & Industrial Research Organisation (Melbourne, Australia); HadCM3 = Hadley Centre Climate Model 3 (Bracknell, UK). A = atmosphere; O = ocean; T_{nn} = spectral model, triangular truncation at wavenumber *nn*; *nm* = spectral model, rhomboidal truncation at wavenumber *nn*; L_{nn} = number (*nn*) of vertical levels. GG = simulation driven by increases in greenhouse-gases only, according to IPCC-Scenario IS92a (IPCC, 1992); GS = forcing by greenhouse-gases and sulphate aerosols according to IPCC-Scenario IS92a; A2, B2 = forcing according to the IPCC marker scenarios A2 and B2 (NAKICENOVIC & SWART, 2000); "x": simulation results from the corresponding scenario run were used in the present study.

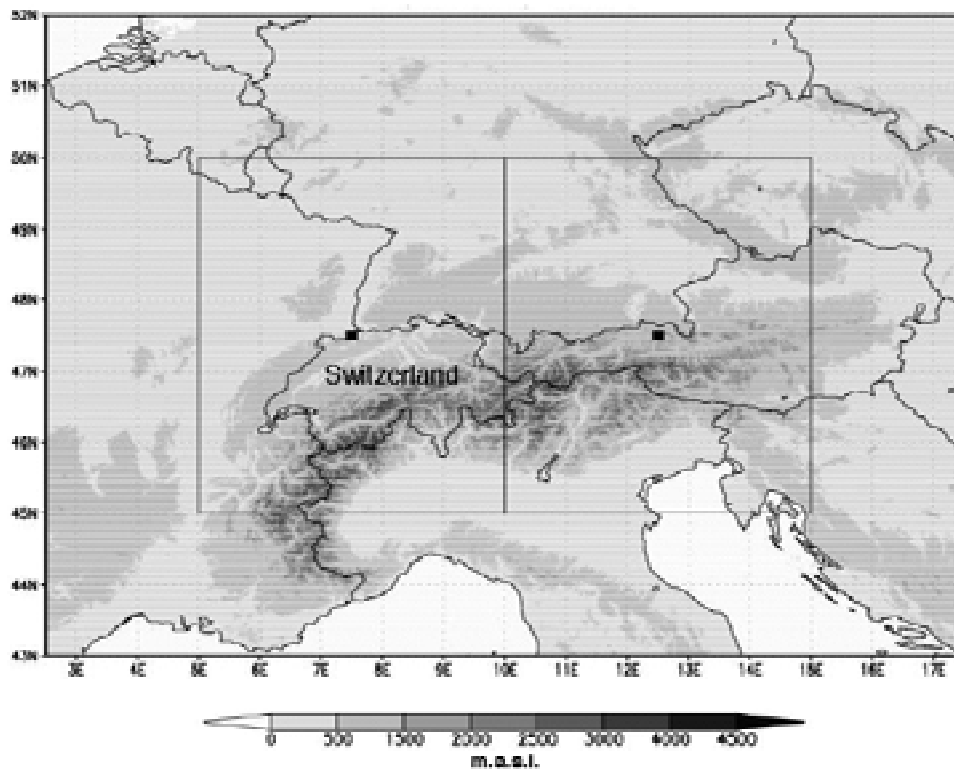


Figure 1: Location of the study region and of the GCM gridboxes used to derive the gridpoint (GP) scenarios.

The preparation of the GCM data is described in KERNEN & GYALISTRAS (2002). It involved interpolation of the original GCM fields, which had different horizontal resolutions, to the same $5^\circ \times 5^\circ$ longitude latitude grid as the observations. The 5° grid corresponded to the resolution of the coarsest available GCMs (cf. Table 1).

All interpolations were done again using the software by WASZKEWITZ *et al.* (1996). The used algorithm computed the quantity of interest at a target gridpoint as the weighted mean of the values found at surrounding gridpoints, such that the interpolation implied some spatial smoothing of the original fields. The possible effect of the interpolation on the various down-scaled results was not systematically investigated in the present study. Experimentation with gridpoint data from the vicinity of the Alps suggested however that the GP scenarios were qualitatively robust and did not depend much on the use of an interpolation procedure.

The GP scenarios for NST and PRECIP were obtained by considering averaged data from two gridpoints in the vicinity of the Alps (Figure 1). Due to the regridding of the GCM fields these gridpoints already represented model-dependent averages from several gridpoints over Europe. The use of spatially averaged information rather than of individual gridpoint values was motivated by the fact that the output of GCMs can be trusted only at a scale of several gridpoint distances (VON STORCH, 1995; WIDMANN & BREHERTON, 2000).

2.3 Statistical Downscaling

The statistical downscaling of the GCM results was based upon multivariate linear regression models that were determined by means of Canonical Correlation Analysis (CCA) in the space spanned by the first few Empirical Orthogonal Functions (EOFs) of the predictor (X) and predictand (Y) fields (VON STORCH & ZWIERS, 1999). In the present study a varying number of predictor fields was used, according to the method proposed by GYALISTRAS *et al.* (1994).

The choice of the large-scale predictors (independent variables) was strongly constrained by the availability of GCM data. The only predictors considered were SLP and NST. Usage of information from upper levels, such as the 500 or 300 hPa geopotential height fields, or of moisture-related variables would have been important, but unfortunately for most GCM simulations no corresponding data were available. PRECIP was not used as a predictor, because it is generally not so reliably simulated by GCMs (KITTEL *et al.*, 1998; GIORGI & FRANCISCO, 2000b).

The predictands (dependent variables) were given by the Swiss gridded ΔT and ΔP fields (GYALISTRAS, 2002). CCA was performed separately for each Y field and calendar month using data from the period 1951-1999 ($n = 49$).

Prior to fitting of the CCA models the data were prepared as follows: Firstly, the annual cycle was removed from all X and Y gridpoint time series by subtracting from each monthly value the 1951-1999 long-term mean of the respective month. Secondly, all data were detrended by first computing for each calendar month and gridpoint the respective linear 49-year trend and then subtracting it from the gridpoint timeseries. Finally, in order to account for the decreasing grid-cell size with increasing latitude, the detrended X variables were weighted with the square root

of their latitude cosine. The detrended Y time series were scaled to unit variance, such that all gridpoints entered the subsequent analysis with the same weight.

The choice of large-scale predictors is known to strongly affect not only the performance of statistical models under present-day conditions (e.g., HUTH, 1999; WILBY & WIGLEY, 2000), but also the statistically downscaled climatic changes (e.g., CHARLES *et al.*, 1999). Moreover, even for a given set of predictors, the results may strongly depend on the exact specification of the downscaling model, e.g. the number of EOFs used for CCA (GYALISTRAS *et al.*, 1994). In order to quantify these sources of uncertainty four different statistical downscaling model variants were investigated, as summarized in Table 2.

SD Model Identifier	X field(s)	N X-EOFs	TVX X-EOFs (%)	N ΔT -EOFs	N ΔP -EOFs	TVX Y-EOFs (%)
SD1	SLP	3–5	80	1–2	6–13	95
SD2	SLP	11–17	99	4–6	15–24	99
SD3	SLP & NST	5–9	80	1–2	6–13	95
SD4	SLP & NST	26–33	99	4–6	15–24	99

Table 2: Overview of the statistical downscaling (SD) models used. X = predictors (independent variables); Y = predictands (dependent variables); N = numbers of Empirical Orthogonal Functions (EOFs) used to fit the SD models (shown is the range of N for all months); TVX = proportion of the total variance of a data set explained by its first N EOFs; SLP = monthly mean sea-level pressure; NST = monthly mean near-surface temperature.

The first two statistical downscaling models (SD1 and SD2) used only SLP as a large-scale predictor, the other two (SD3 and SD4) considered in addition NST. The common EOF-analysis of the SLP and NST fields prior to CCA was done as follows: First all anomaly time series of each field were multiplied by a given factor. This factor was determined such, that at the end both fields had the same total variance. Then the field vectors were concatenated and a normal EOF analysis was performed (see GYALISTRAS *et al.*, 1994).

The models SD1 and SD3 used relatively small numbers (N) of X- and Y-EOFs. The N were determined by inspecting the eigenvalue spectra of the various data sets. It was found that the total variance explained by the X-EOFs levelled-off at ~80%, i.e. the inclusion of further EOFs beyond those needed to reach 80% did not bring much improvement in describing the overall variability of the X data. The corresponding threshold for the Y data sets was found to be at around 95%. For the sake of simplicity the values of 80%, respectively 95% were used to determine N for all months. Typically, larger N values were needed to reach the above thresholds for the summer as compared to the winter months (cf. ranges shown in Table 2).

The models SD2 and SD4 accounted for a very large proportion ($\geq 99\%$) of the total variance of the X and Y data sets. These models were used in order to explore an extreme alternative to the statistically "reasonable" models SD1 and SD3. The usage of a large number of EOFs can lead to overfitting of the regression equations. Nevertheless, it should be noted that the N used in

the models SD2 and SD4 were still relatively small as compared to the total number of X variables (153 per large-scale predictor field), respectively Y variables (~3'000 for ΔT or ΔP).

The performance of all downscaling models was assessed by means of a leave-one-out cross-validation procedure. The X-Y data pairs for every individual month within the 1951-1999 period were first excluded from the observational data sample and then the month's regional anomaly field was predicted from a downscaling model that was fitted using the remaining n=48 data pairs for that particular month. The prediction was based on the observed monthly SLP (or SLP plus NST) anomalies. These anomalies were derived from non-detrended fields.

The model quality was measured by the monthly areal averages of the following statistics, which were evaluated at every gridpoint (index i): the coefficient of determination (R^2_i) between the observed (Y_i) and the cross-validated (\hat{Y}_i) gridpoint time series (n=49); and the proportion of explained variance, $PEV_i = 1 - \{ \text{Variance}(E_i) / \text{Variance}(Y_i) \}$, where $E_{i(t)} = Y_{i(t)} - \hat{Y}_{i(t)}$. A PEV-value smaller than zero indicated a very poor performance of the downscaling procedure, and a value close to zero suggested that the downscaling model was not better than a trivial model that always uses the mean of Y to predict $Y_{i(t)}$. Positive PEV-values suggested a useful result in the sense that the error variance is smaller than the total variance of the observations.

In addition to the R^2 and PEV statistics the trends and climate sensitivities (see below) of the cross-validated variables were computed for comparison with their observed counterparts.

Regional climate scenarios were finally obtained by applying the CCA regression models to SLP and NST anomaly fields that were derived from the various GCM simulations. Most simulations started well before 1950 and extended to the end of the 21st century. The GCM anomaly fields were always defined as departures from the 1961-1990 simulated means, except for the PCM model. Due to lack of data the PCM anomalies were defined relative to the 1981-2000 mean for scenario A2 and relative to the 2001-2020 mean for scenario B2.

2.4 Climate Change Statistics

The temporal change in the observed and modelled regional data sets was analyzed by considering (i) the linear trends, and (ii) the climate sensitivities of the regional variables.

The trends provided a measure of change in absolute units per time interval. They were estimated by the slope of the linear regression of a given time series against the year number. The regional climate sensitivities were used in order to be able to compare regional climate responses across different forcing scenarios and GCMs. They were defined as the change of a regional variable per degree change in the globally and annually averaged near-surface temperature (T_g) and were estimated by the slope of the linear regression from T_g . A similar approach was also used at the continental scale by ANISIMOV (2001).

The trends and sensitivities from the observations were determined for the period 1951-1999. T_g was derived from the global NST data set of JONES *et al.* (2001). The scenario trends and sensitivities were determined using all available scenario data for the years ≥ 2000 .

The Tg for each scenario was computed from the NST of the the respective GCM simulation. For most simulations ca. 100 years of data were used to estimate the change statistics. The only exception was the ECHAM4-GS scenario, where only 50 simulated years were available for the 21st century.

3. RESULTS

3.1 Performance of Downscaling Models

Figure 2 shows the areal mean R^2 and PEV statistics for the various SD model variants. For ΔT the models SD2 and SD4 performed always better than the SLP-only models. The use of a larger number of EOFs yielded a substantial improvement for the model SD2 as compared to SD1, and a smaller improvement for model SD4 as compared to SD3.

The cross-validated R^2 for ΔP (Figure 2, right) were similar for all model variants and showed a clear minimum during the summer months. The areal mean PEV was mostly negative, except for model SD4. The models SD2 and SD4 performed with regard to PEV always better than their more parsimonious counterparts.

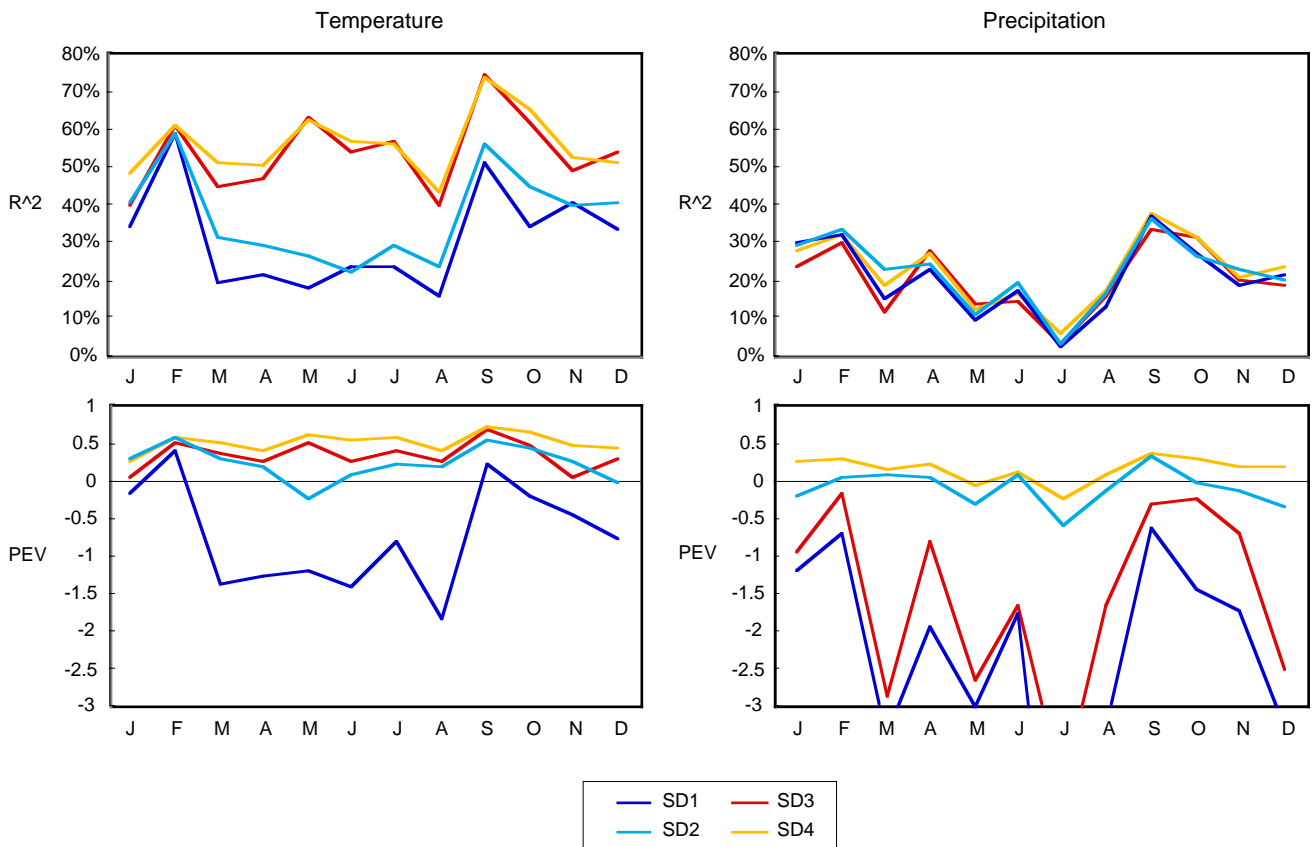


Figure 2: Cross-validation statistics of different statistical downscaling (SD) models for temperature (left) and precipitation (right) as a function of month. R^2 = coefficient of determination; PEV = proportion of explained variance.

Figure 3 compares the cross-validated, areal and seasonal mean trends and sensitivities with the corresponding statistics from the observations. The observations showed for ΔT the largest changes in the solstitial seasons. The trends were reasonably well reproduced by the models SD3 and SD4, but for the sensitivities less good results were obtained. In particular, all models strongly overestimated the winter sensitivities and underestimated the Autumn sensitivities.

The ΔP change statistics were generally not so well reproduced (Figure 3, right). For winter and spring neither of the SLP-only models got the sign of the trends correctly, whereas models SD3 and SD4 strongly underestimated the observed decrease in areal mean summer precipitation. The sensitivities showed a similar picture as the trends, except for Autumn, where none of the downscaling models captured the positive signal that was detected in the observations.

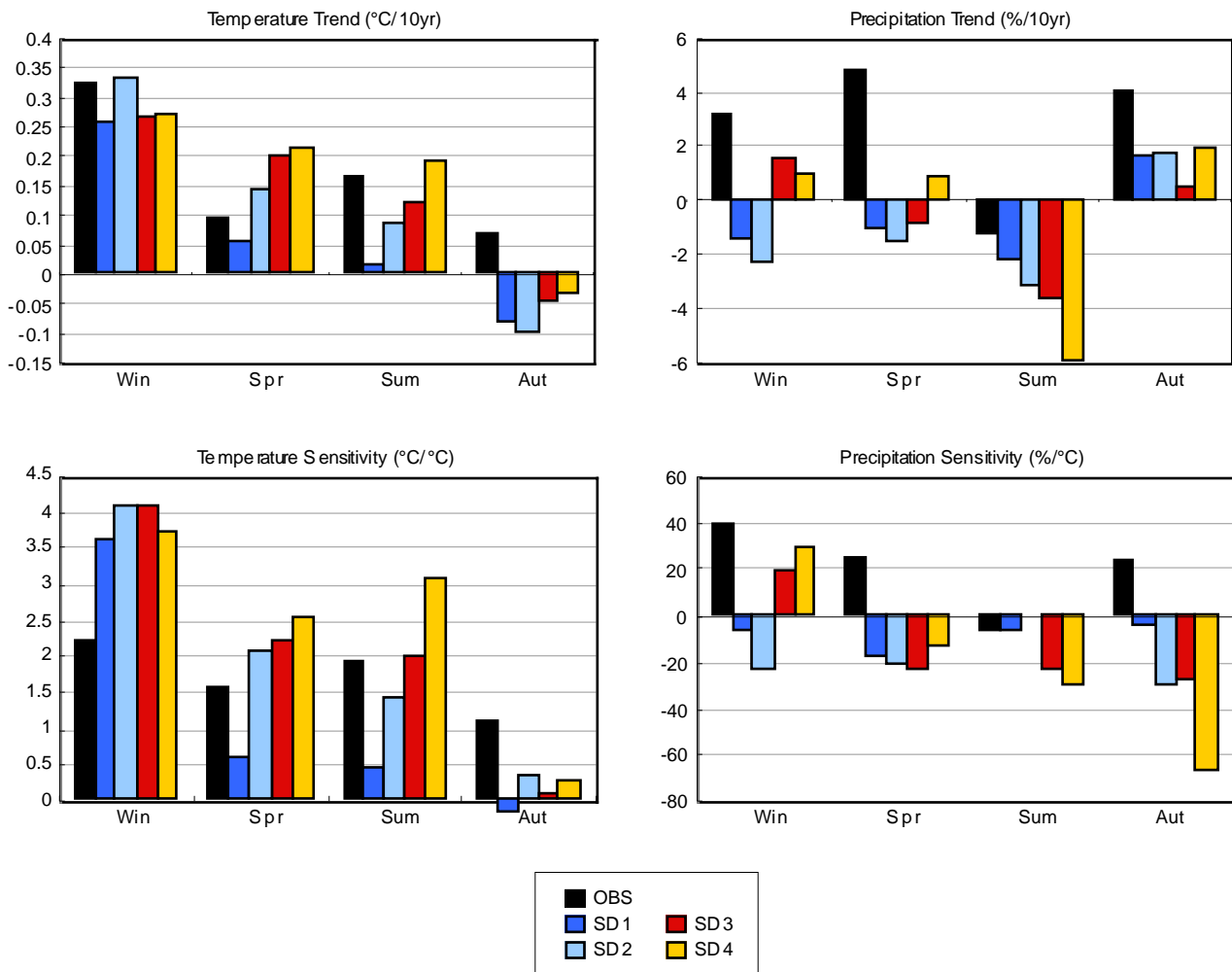


Figure 3: Comparison of observed (OBS) and by means of different statistical downscaling (SD) models reconstructed, regionally and seasonally averaged climate change statistics for temperature (left) and precipitation (right) as a function of month. All results refer to the period 1951-1999.

3.2 Global Mean Temperature Trends

The T_g trends simulated by the various GCMs for the 21st century are juxtaposed in Figure 4. Note that different sets of models were available for the GG and GS forcing scenarios as compared to the A2 and B2 scenarios. Accordingly, the responses obtained for the GG or GS forcings should not be compared with the responses from the A2 or B2 forcings by considering the entire groups of simulations, but rather only individual GCMs.

It can be seen that the GG forcing yielded generally a more rapid warming than the GS forcing, except perhaps for GCM No. 7 (HadCM3, see Table 7). All models reacted also stronger to the A2 forcing than to the B2 forcing. The pattern of inter-model variability was very similar between the GG and GS forcings. The same was also true with respect to the A2 and B2 forcings. Under the GG and GS forcings model No. 2 (CGCM1) always showed a particularly strong response, whereas under the A2 and B2 forcings the GCM No. 5 (DOE PCM) always simulated a smaller trend than the other three models.

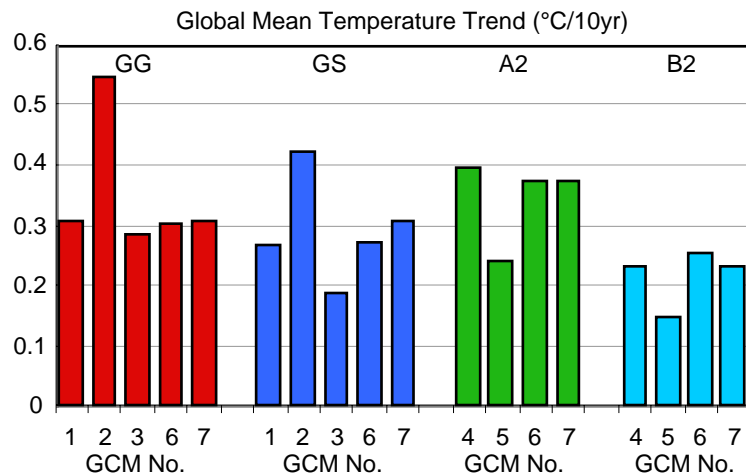


Figure 4: Comparison of 21st century trends in the globally and annually averaged near-surface temperature simulated by different GCMs (cf. Table 1). GG = simulation driven by increases in greenhouse-gases only, according to IPCC-Scenario IS92a (IPCC, 1992); GS = forcing by greenhouse-gases and sulphate aerosols according to IPCC-Scenario IS92a; A2, B2 = forcing according to the IPCC marker scenarios A2 and B2 (NAKICENOVIC & SWART, 2000).

3.3 Areal Mean Scenarios

The regionally averaged temperature sensitivities that were obtained from the two methods GP and DS3 are shown in Figure 5. The GP method yielded generally higher sensitivities than the DS3 method. For instance the annual mean sensitivity for the GG (for the A2) scenario with method GP was 1.3 °C/°C (respectively 1.1 °C/°C), whereas for method DS3 it amounted to only 1.0 °C/°C (respectively 0.85 °C/°C). Under both regionalization methods the GG and the A2 forcing yielded larger annual mean sensitivities than the GS, respectively B2 forcing.

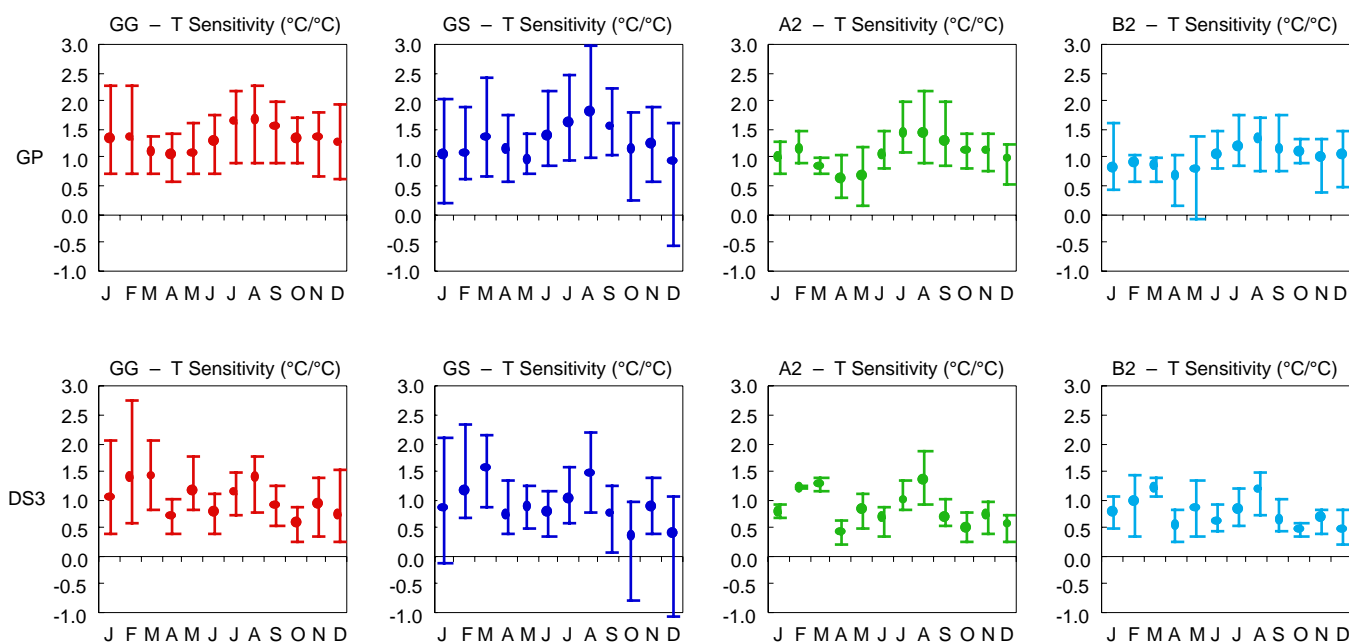


Figure 5: Comparison of temperature sensitivities obtained from the regionalization methods GP (top) and DS3 (bottom) as a function of month for different forcing scenarios. Error bars denote the range of sensitivities obtained from 5 GCMs under the forcing scenarios GG and GS, respectively 4 GCMs under the forcing scenarios A2 and B2.

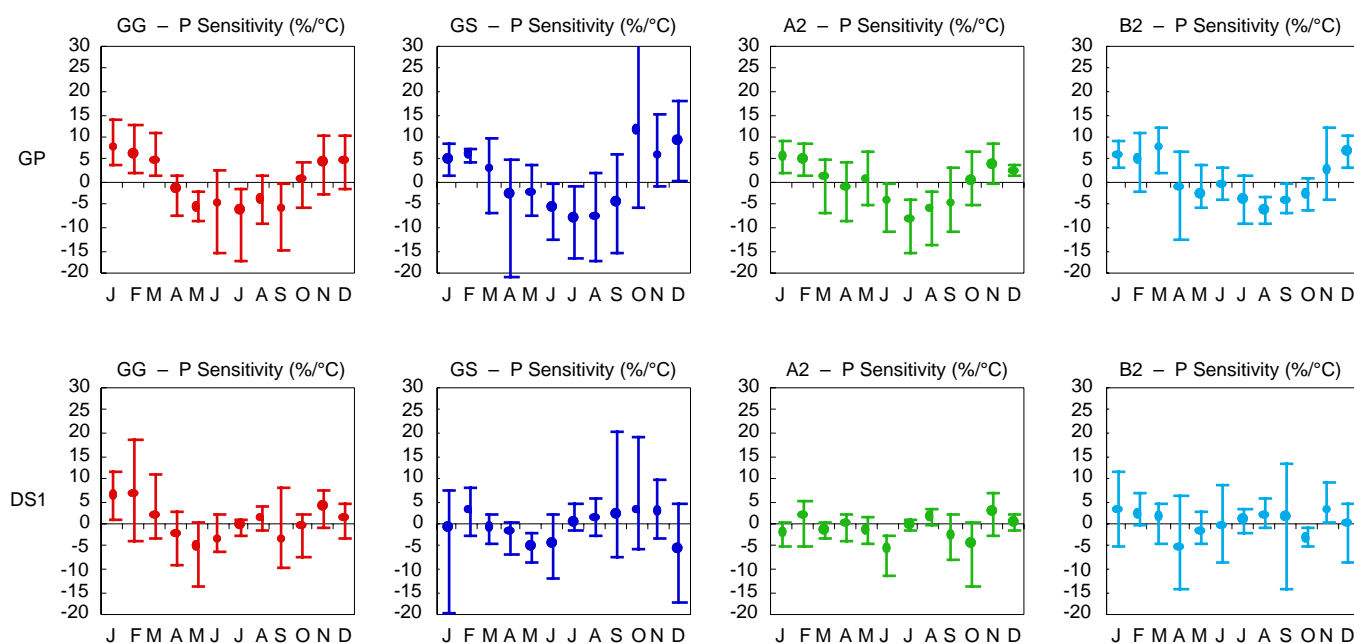


Figure 6: Comparison of the precipitation sensitivities obtained from the regionalization methods GP (top) and DS1 (bottom) as a function of month for different forcing scenarios. Error bars denote the range of sensitivities obtained from 5 GCMs under the forcing scenarios GG and GS, respectively 4 GCMs under the forcing scenarios A2 and B2.

The annual cycles of the sensitivities differed quite strongly between the two regionalization methods; the highest correlation coefficient between the mean sensitivities from GP and DS3 was found for the forcing scenario GS ($r = 0.53$, $n = 12$). Quite differently, the annual cycles of changes projected by a given regionalization method for different forcing scenarios were highly correlated, with r ranging from 0.69 to 0.97 for the GP scenarios and $r \geq 0.96$ for the DS3 scenarios.

A similar analysis for the precipitation sensitivities is shown in Figure 6. Here the results from methods GP and DS1 are compared. The two methods agreed on the sign of the annually averaged precipitation sensitivity for the GHG and B2 forcings (positive sensitivity) and for the A2 forcing (negative mean sensitivity).

However, the details of the annual cycles varied quite strongly. In particular, all GP scenarios showed negative sensitivities for the summer months which were generally not reproduced by the SD1 models. The most similar annual cycles of sensitivities from the two regionalization procedures were obtained for the GG scenario ($r = 0.85$).

The annual cycles of changes obtained for different forcings from the GP method showed correlation coefficients between 0.66 and 0.89. Quite differently, the seasonal pattern from the SD1 scenarios depended much more on the choice of forcing ($r \leq 0.67$).

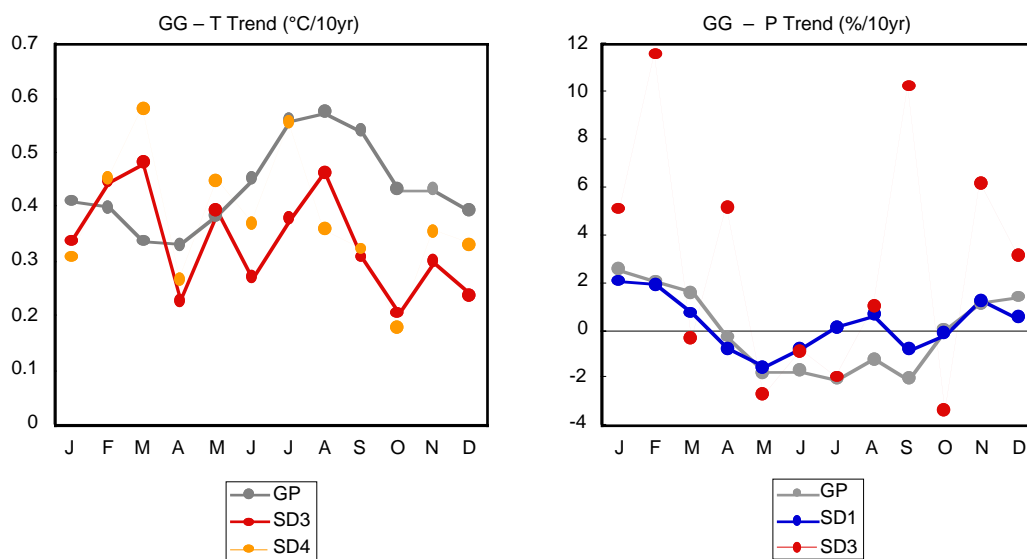


Figure 7: Comparison of trends for temperature (left) and precipitation (right) obtained from different regionalization methods as a function of month. All data refer to the GG forcing scenario and present the mean signal downscaled from 5 GCM simulations.

Figure 7 illustrates the uncertainty due to the choice of regionalization procedure. It compares the average trends that were obtained from the application of different downscaling procedures to the output of the five GCM-runs that were forced with the GG-scenario (cf. Table 2).

For temperature (Figure 7, left) the method GP yielded the strongest response. The outputs of the SD3 and SD4 models were more similar to each other than to the GP response. For precipitation (Figure 7, right) the model SD3 gave for several months rather dramatic trends. GP and SD1 agreed quite well, except during the summer months.

3.4 Spatial Patterns

Figure 8 shows the spatial pattern of the average temperature sensitivities that were downscaled from the four A2 and B2 simulations using method SD3.

The temperature sensitivities were everywhere positive and showed for winter (Figure 8, top) a much larger spatial variability than for summer (Figure 8, bottom). The B2 forcing yielded somewhat smaller sensitivities than the A2 forcing (cf. Figure 5, bottom), but the spatial pattern of the sensitivities for the two different forcing scenarios was very similar (Figure 8, left vs. right).

Generally, for all pairs of forcing scenarios the monthly sensitivities showed always very high spatial correlation coefficients, $r_s \geq 0.96$ (analyses not shown). Note, however, that this result applied to the mean downscaled sensitivities from several GCM simulations per forcing scenario. The pattern similarity for individual GCM runs was not analyzed.

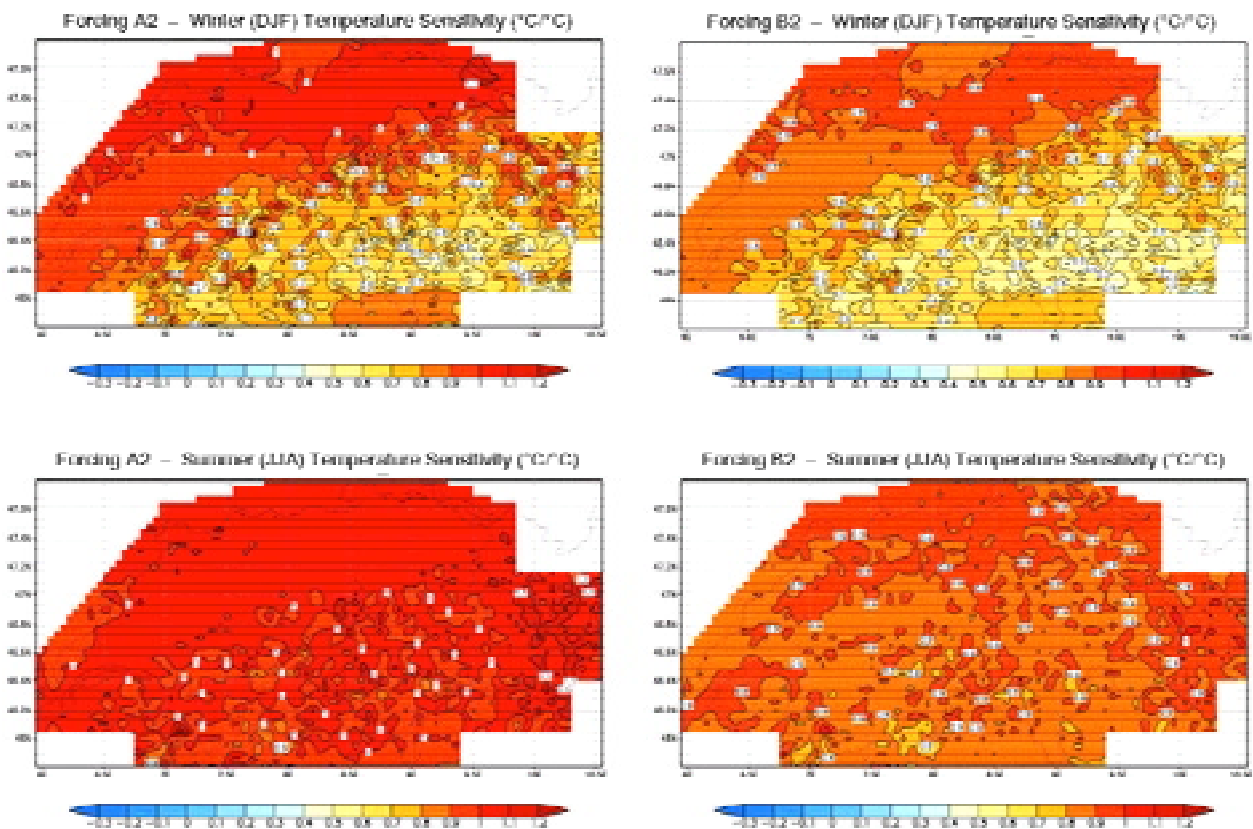


Figure 8: Comparison of winter (top) and summer (bottom) mean temperature sensitivities obtained for the forcing scenarios A2 (left) and B2 (right). Shown is the mean signal from four GCM simulations (see Table 1), as downscaled with the aid of method SD3 (see Table 2).

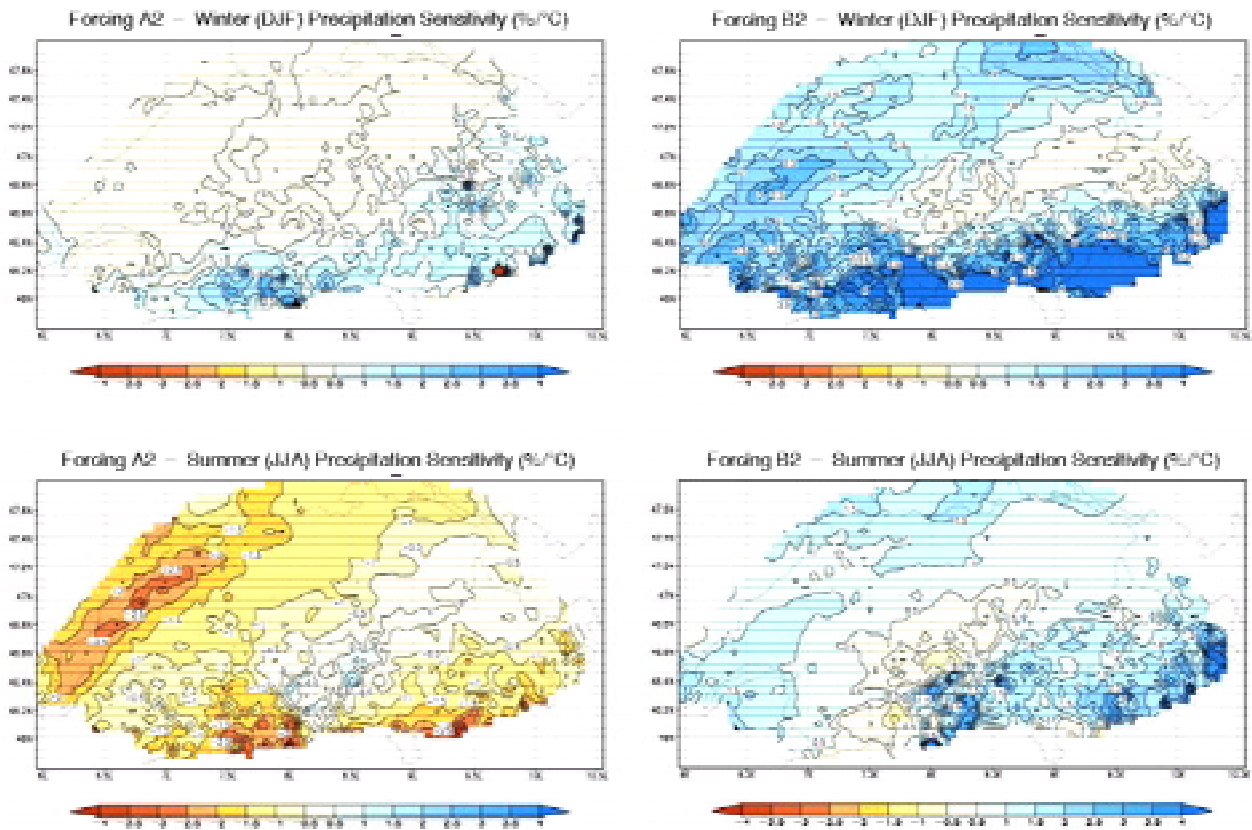


Figure 9: Comparison of winter (top) and summer (bottom) mean precipitation sensitivities obtained for the forcing scenarios A2 (left) and B2 (right). Shown is the mean signal from four GCM simulations (see Table 1), as downscaled with the aid of method SD1 (see Table 2).

Figure 9 shows the average precipitation sensitivities that were downscaled for the A2 and B2 forcings using method SD1.

Both forcings yielded over most of the Swiss area positive precipitation sensitivities for winter (Figure 9, top), but disagreed widely on the sign of the sensitivities for summer (Figure 9, bottom; cf. Figure 6, bottom). The r_s for the two winter sensitivity maps was -0.36, for the two summer maps it was close to zero.

The monthly r_s between any two forcing scenarios were distributed as follows ($n = 6 \times 12 = 72$): In 12% of all cases was $r_s < -0.7$, in 82% of all cases was $r_s > 0$, and in 60% of all cases was $r_s > 0.7$. The largest discrepancies were obtained for December, January, March and September. For all other months was always $r_s \geq 0.66$.

4. DISCUSSION

The scenarios showed marked changes, but these were often smaller than the observed trends and sensitivities for the last 50 years (Figure 3 vs. Figure 5). The stability of this result needs to be further investigated because the role of the natural variability was not quantified in the present work. Also, the use of linear regressions to estimate the trends and sensitivities was somewhat problematic with regard to the observations because the observed Tg was more or less constant from 1951 to ~1975, and showed an increase only in the second half of the analysis interval.

However, at least with regard to the scenarios, the various change statistics were probably quite robust, since they were typically estimated from ~100 years of data. Moreover, no obvious non-linear behaviour was found in the projected time-evolution of the regional climate variables or in their dependence from the simulated Tg (not shown).

The use of the NST field in addition to SLP as a large-scale predictor gave a clear improvement for the downscaling of the monthly temperatures, but less so with regard to precipitation (Figure 2). A similar result was also obtained by GYALISTRAS *et al.* (1994). Possibly, the slight improvement that was obtained for precipitation was because the temperature field provides some information on the atmospheric moisture content.

A closer inspection of the various statistical downscaling models (not shown) revealed that the models describe physically plausible relationships between the large-scale and regional climate anomalies (see also GYALISTRAS *et al.*, 1994, 1998). However, the choice of the most appropriate downscaling model is not trivial, because different validation statistics may give diverging results (Figures 2 and 3).

On average, it appears that with regard to temperature the downscaling model variant SD3 is more appropriate than variant SD4, since the latter tends to overestimate the observed sensitivities and trends (Figure 3). With regard to precipitation the model variant SD1 is probably most appropriate because it gave similar results to variant SD2, but based on a smaller number of predictor EOFs, and because it yielded more plausible change signals than variants SD3 (Figure 7) and SD4 (results not shown). The issue of optimal model selection needs however to be investigated further.

The GP and SD scenarios showed large differences. These could be on the one hand due to systematic errors in the GCMs, for instance due to problems in the simulation of the large-scale flow, or due to errors in the representation of subgrid-scale process. In particular, the consistent decrease in summertime precipitation found in the GP scenarios (Figure 6, top) might be due to a too simplistic representation of land-surface processes in the GCMs, which may cause excessive evaporation and unrealistic drying in summer (SENEVIRATNE *et al.*, 2001). The large model biases reported by GIORGI & MEARNS, (2002) for central and southern Europe render the GP signal also questionable.

On the other hand, the discrepancies could also be due to deficiencies of the downscaling procedure, for instance due to the presence of non-linear processes, the need to extrapolate beyond the range of the data used for model calibration, or due to the omission of relevant large-scale predictors.

The present work considered only selected aspects of the scenarios. The now available, extensive scenario data base enables several further analyses, for instance with regard to the altitudinal distribution of future climate change (cf. GIORGI *et al.*, 1997), or the combined effects of changes in temperature and precipitation on a variety of systems. Thanks to the use of a gridded data set, the statistically downscaled scenarios could also be easily compared with the results from regional climate model simulations.

5. CONCLUDING REMARKS

The pattern and magnitude of future temperature and precipitation changes in Switzerland must be considered very uncertain. Some scenarios show similar seasonal or spatial patterns of change, but the reality and robustness of these patterns needs to be further investigated.

The present study suggests that with regard to the assessment of the areal mean regional climate *sensitivities* the largest uncertainties are due to the regionalization step, followed by the choice of GCM, and by the choice of global forcing.

A different order may hold with regard to *absolute changes* or trends, where the magnitude of the global forcing and the climate sensitivities of the various GCMs are of major importance. This point should be investigated in future work.

The derived scenarios were based on a limited range of global forcings GCMs and regionalization procedures such that they probably underestimate the true range of uncertainty. Also, it is not possible to assign any objective probabilities to the individual scenarios. In my opinion, all scenarios taken together represent our current state of knowledge in the best possible way.

Further, systematic evaluations of the reliability of GCMs and regionalization methods could help to determine which scenarios can be trusted more and which less. However, it can be expected that substantial uncertainties will always remain.

Acknowledgements

Research by DG was supported by the Swiss Federal Research Station for Agroecology and Agriculture (FAL, project CS-MAPS), and the Swiss Agency for the Environment, Forests and Landscape (BUWAL, project No. 2001.L.03 / TREWALP).

7. REFERENCES

- ANISIMOV, O.A. (2001). Predicting Patterns of Near-Surface Air Temperature Using Empirical Data. *Clim. Change* 50(3): 297-315.
- BOER, G.J., FLATO, G.M. & RAMSDEN, D. (2000). A transient climate change simulation with historical and projected greenhouse gas and aerosol forcing: projected climate for the 21st century. *Clim. Dyn.* 16: 427-450.
- BÖHM, R., AUER, I., BRUNETTI, M., MAUGERI, M., NANNI, T. & SCHÖNER, W. (2001). Regional temperature variability in the European Alps: 1760-1998 from homogenized instrumental time series. *Int. J. Climatol.* 21(14): 1779-1801.
- CHARLES, S.P., BATES, B.C., WHETTON, P.H. & HUGHES, J.P. (1999). Validation of downscaling models for changed climate conditions: case study of southwestern Australia. *Clim. Res.* 12: 1-14.
- CUBASCH, U. & MEEHL, G.A. (2002) Projections of future climate change
- DELWORTH, T.L., STOUFFER, R.J., DIXON, K.W., SPELMAN, M.J., KNUTSON, T.R., BROCCOLI, A.J., KUSHNER, P.J. & WETHERALD R.T. (2002). Review of simulations of climate variability and change with the GFDL R30 coupled climate model. *Clim. Dyn.* 19: 555-574.
- EMORI, S., NOZAWA, T. ABE-OUCHI, A. NUMAGUTI, A. KIMOTO, M. & NAKAJIMA, T. (1999). Coupled ocean-atmosphere model experiments of future climate change with an explicit representation of sulfate aerosol scattering. *J. Meteorol. Soc. Japan* 77: 1299 - 1307.
- FLATO, G.M. & BOER, G.J. (2001). Warming Asymmetry in Climate Change Simulations. *Geophys. Res. Lett.*, 28: 195-198.
- GIORGI, F. (1995). Perspectives for regional earth system modelling. *Global Planet. Change* 10: 23-42.
- GIORGI, F., HURRELL, J.W., MARINUCCI, M.R. & BENISTON, M. (1997). Elevation dependency of the surface climate change signal: a model study. *J. Clim.* 10: 288-296.
- GIORGI, F. & FRANCISCO, R. (2000a). Uncertainties in regional climate change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM. *Clim. Dyn.* 16(2/3): 169-182.
- GIORGI, F. & FRANCISCO, R. (2000b). Evaluating uncertainties in the prediction of regional climate change. *Geophys. Res. Lett.* 27(9): 1295-1298.
- GIORGI, F. & MEARNES, L.O. (2002). Calculation of Average, Uncertainty Range, and Reliability of Regional Climate Changes from AOGCM Simulations via the "Reliability Ensemble Averaging" (REA) Method. *J. Clim.* 15(10): 1141-1158.
- GORDON, H.B. & O'FARRELL, S.P. (1997). Transient climate change in the CSIRO coupled model with dynamic sea ice. *Mon. Wea. Rev.* 125: 875-907.
- GYALISTRAS, D., VON STORCH, H., FISCHLIN, A. & BENISTON, M. (1994). Linking GCM-simulated climatic changes to ecosystem models: case studies of statistical downscaling in the Alps. *Clim. Res.* 4(3): 167-189.
- GYALISTRAS, D., SCHÄR, C., DAVIES, H.C. & WANNER, H. (1998). Future Alpine climate. In: Cebon, P., U. Dahinden, H.C. Davies, D. Imboden, and J. C. (eds.): *Views from the Alps: regional perspectives on climate change*, MIT Press, Boston, pp. 171-223.
- GYALISTRAS, D. (2000). Klimaszenarien für den Alpenraum und die Schweiz: Neuester Stand und Vergleich. In: Wanner, H., Gyalistras, D., Luterbacher, J., Rickli, R., Salvisberg, E. & Schmutz, C.: *Klimawandel im Schweizer Alpenraum*. vdf, Hochschulverlag AG an der ETH Zürich, pp 197-235.
- GYALISTRAS, D. (2002). Development and validation of a high-resolution monthly gridded temperature and precipitation data set for Switzerland (1951-2000). Manuscript submitted to *Clim. Res.*, 37pp. Document available online from <http://www.giub.unibe.ch/~gyalistr/DG/DGGreyLit.html#37>.

- HULME, M. & CARTER, T.R. (2000): The changing climate of Europe. In: *Assessment of potential effects and adaptations for climate change in Europe: The Europe ACACIA project*, Parry, M.L. (ed.). Jackson Environment Institute, University of East Anglia, Norwich, United Kingdom, pp. 47-84.
- HUTH, R. (1999). Statistical downscaling in central Europe: evaluation of methods and potential predictors. *Clim. Res.* 13: 91-101.
- IPCC, 1992: *Climate Change 1992: The Supplement Report to the IPCC Scientific Assessment*, Houghton, J.T., B.A. Callander and S.K. Varney (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 200pp.
- JOHNS, T.C., GREGORY, J.M., INGRAM, W.J., JOHNSON, C.E., JONES, A., LOWE, J.A., MITCHELL, J.F.B., ROBERTS, D.L., SEXTON, D.M.H., STEVENSON, D.S., TETT, S.F.B. & WOODAGE, M.J. (2001). Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios. Hadley Centre Technical Note 22. Met Office, Bracknell, United Kingdom, 60 pp.
- JONES, P.D., OSBORN, T.J., BRIFFA, K.R., FOLLAND, C.K., HORTON, E.B., ALEXANDER, L.V., PARKER, D.E. & RAYNER, N.A. (2001). Adjusting for sampling density in grid box land and ocean surface temperature time series. *J. Geophys. Res.* 106: 3371-3380. Data set available online from <http://www.cru.uea.ac.uk/cru/data/temperature> (data set HadCRUTv).
- JUNGO, P. & BENISTON, M. (2001). Changes in the anomalies of extreme temperature anomalies in the 20th century at Swiss climatological stations located at different latitudes and altitudes. *Theor. Appl. Climatol.* 69(1/2): 1-12.
- KERNEN, R. & GYALISTRAS, D. (2002). GCM DAT Version 1.1 – A data base of General Circulation Model results. Internal report, Climatology and Meteorology Research Group, University of Bern. Bern, Switzerland, 27 pp. Document available online from <http://www.giub.unibe.ch/~gyalistr/DG/DGGreyLit.html#34>.
- KITTEL, T.G.F., GIORGI, F. & MEEHL, G.A. (1997). Intercomparison of regional biases and doubled CO₂-sensitivity of coupled atmosphere-ocean general circulation model experiments. *Clim. Dyn.* 14(1): 1-15.
- NAKICENOVIC, N. & SWART, R. (eds.) (2000). *Special Report on Emissions Scenarios*. Cambridge University Press, Cambridge, United Kingdom, 599 pp.
- ROECKNER, E., BENGTTSSON, L., FEICHTER, J., LELIEVELD, J. & RODHE, H. (1999). Transient climate change simulations with a coupled atmosphere-ocean GCM including the tropospheric sulfur cycle. *J. Clim.* 12: 3004-3032.
- SCHMIDLI, J., SCHMUTZ, C., FREI, C., WANNER, H. & SCHÄR, C. (2002). Mesoscale precipitation variability in the region of the European Alps during the 20th Century. *Int. J. Climatol.* 22: 1049-1074.
- SENEVIRATNE, S.I., PAL, J.S., ELTAHIR, E.A.B., & SCHÄR, C. (2001). Summer dryness in the US midwest: a process study with a regional climate model. *Clim. Dyn.*, submitted.
- TRENBERTH, K. & PAOLINO, D. A. (1980). The Northern Hemisphere sea-level pressure data set: trends, errors and discontinuities. *Mon. Wea. Rev.* 108: 855-872. Data set available online from <http://dss.ucar.edu/data/sets/ds010.1>.
- VON STORCH, H. & ZWIERS, F. W. (1999). *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, United Kingdom, 484 pp.
- VON STORCH, H. (1995). Inconsistencies at the interface of climate impact studies and global climate research. *Meteorol. Z.* 4: 72-80.
- WASHINGTON, W.M., WEATHERLY, J.W., MEEHL, G.A., SEMTNER JR., A.J., BETTGE, T.W., CRAIG, A.P., STRAND JR., W.G., ARBLASTER, J., WAYLAND, V.B., JAMES, R. & ZHANG, Y. (2000). Parallel climate model (PCM) control and transient simulations. *Clim. Dyn.* 16: 755-774.
- WASZKEWITZ, J., LENZEN, P. & GILLET, N. (1996). The PINGO package – Procedural Interface to Grib formatted Objects. DKRZ Technical Report No. 11. Deutsches Klimarechenzentrum, Hamburg, Germany, 225 pp. Document available online from <http://www.dkrz.de/forschung/reports/report11-1.1/pingo-1.html>.

- WIDMANN, M. & BRETHERTON, C.S. (2000). Validation of mesoscale precipitation in the NCEP reanalysis using a new gridpoint data set for the northwestern US. *J. Clim.* 13: 1936–1950.
- WIDMANN, M. & SCHÄR, C. (1997). A principal component and long-term trend analysis of daily precipitation in Switzerland. *Int. J. Climatol.* 17(12): 1333-1356.
- WILBY, R.L. & WIGLEY, T.M.L. (2000). Precipitation predictors for downscaling: observed and general circulation model relationships. *Int. J. Climatol.* 20: 641-661.
- ZEVENBERGEN, L.W. & THORNE, C.R. (1987). Quantitative analysis of land surface topography. *Earth Surface Processes and Landforms* 12: 47-56.