Methodological details regarding the E-SD and mapping the results

R.E. Benestad

Meteorological Institute, PO Box 0313, Oslo, Norway

Empirical-Statistical Downscaling

The present results were obtained using a similar common empirical orthogonal function framework as *Benestad* [2002b] (referred to as 'B2002b'), but using a stepwise multiple regression for single sites as opposed to a canonical correlation analysis (CCA) - based model for a *set* of locations. The predictor used for calibration against the local temperature was the monthly mean large-scale T(2m) anomalies from the ERA40 re-analyses [*Simmons et al.*, 2004]. Hence, the present analysis used ERA40 for predictor calibration, as opposed to the 1873–1998 gridded analysis [*Benestad*, 2000] used to derive the B2002b results. The corresponding predictor for local precipitation was the total precipitation from the ERA40, whereas B2002b used sea level pressure (SLP) as a predictor. The E-SDS was carried out using a tool called '`clim.pact`' [*Benestad*, 2004a] (version 2.1-5), written for the `R`-environment [*R Development Core Team*, 2004]. Note that the `clim.pact` package is open source and freely available from CRAN (http://cran.r-project.org), and the method is described in a number of earlier publications [*Benestad*, 2004a; and references therein].

Whereas the B2002b empirical-statistical downscaling (E-SDS) analysis used a fixed set of predictor domains common for all stations studies, the present analysis selected the predictor domain on an individual station basis based on the criterion that the predictor region should only encompass the region where the large-scale anomaly field is positively correlated with the local variable [*Benestad*, 2004a]. Once the domain was determined for a given location, the GCM results were interpolated to the observed (here ERA40 re-analysis) grid for this domain, and its anomalies concatinated with those of the observations. Then an empirical orthogonal function (EOF) [*Lorenz*, 1956] analysis was applied to the combined data set (common EOF, [*Barnett*, 1999]), and the EOF products were used for E-SDS model calibration and prediction. This method is evaluated and described in further detail by *Benestad* [2001]. The observations were de-trended prior to model calibration and a stepwise screening using the Akaike information criterion (AIC, [*Wilks*, 1995], p.301-302) was used to exclude non-important principal components and hence avoid so-called over-fit. In `clim.pact`, this objective downscaling approach is implemented by the function '`objDS`' [*Benestad*, 2004a]. The downscaling was applied separately to single series for one given station (in B2002b the downscaling was applied to groups of stations simultaneously using CCA and cross-validation instead of stepwise regression for a single location) and a given calendar month. The annual series for each location was constructed from 12 individual downscaling exercises in order to represent all calendar months. No form for so-called 'inflation' [*von Storch*, 1999] was applied.

The calibration period in the present analysis is short relative to the prediction interval (1958–2002 for Norwegian stations updated with data from the national archive, 1958–1999 for Nordklim and NARP stations, and 1958–1990 for NACD stations), and the spatial EOF patterns may therefore be governed by GCMs rather than the observations, unless other measures are introduced. In order to ensure that the observed spatial patterns dominate the EOF products, the GCM data were scaled down (multiplication with $a = 0.25 \times n_{\mathrm{ERA40}}/n_{\mathrm{GCM}}$, where n_{ERA40} is the ERA40 record length and n_{GCM} the GCM record length) prior to the EOF analysis, and subsequently re-scaled to describe the original variance before the stepwise regression analysis. A set of tests was conducted to check whether the scaling process produced significant differences of adverse effect, but the results were similar to the unscaled analysis for selected locations and GCM (not shown). The motivation for using a scaling factor of 0.25 is to avoid splitting the ERA40 and GCM into different modes (principal components) if the spatial structure differ substantially.

Quality control

E-SDS incorporates a form of quality control by requiring similar spatial structures on anomalies in the observations and the GCM. Post-process quality control based on the E-SDS diagnostics involved comparisons between the spatial weight patterns from the stepwise regression analysis and the correlation maps between large-scale anomalies and the local variable. In this case, a spatial correlation r_{xy}(month) was estimated between these two types of maps for each calendar month, assuming that they are similar for a realistic regression and when the spatial modes in the GCMs look like those in the ERA40 data. Another criterion was that the linear trend estimates for the individual months should change smoothly over the 12 calendar months. An index of unrealistic trend 'spikes' was defined as $n_m = 1 - 1/12 \sum_{i=1}^{12} \mathcal{H}(3s - m_i)$, where m_i is the difference between the linear trend for calendar month $i \in [1, 12]$ and subsequent month, s is the standard deviation of the trend differences, and $\mathcal{H}()$ is the Heaviside function. Further quality diagnostics included flagging (f_s) if the summer temperature variability is more pronounced than winter variability ($f_s = \mathcal{H}(s_{jja} - s_{djf})$), or lower summer temperatures than in winter, spring and autumn and whether the winters are colder than the other seasons ($f_m = \mathcal{H}(\overline{x}_{jja} - \overline{x}_{djf})$). Additional quality indices can be extracted from the E-SDS regression R^2 statistics, and was here defined as the number of $n_{R^2<0.6} = \sum_{i=1}^{12} \mathcal{H}(R_i^2 - 0.6)$ for the R^2-statistic of month i. Further quality-control models included the number of cases $n_{p>0.1}$ where the regression p-value[1] exceeds 0.1, number of cases with negative temperature trend $n_{m<0}$ (i.e. cooling), and tests whether the predictor domain is very large (when longitude range exceeds 90°E or latitude range exceeds 40°N) or small (longitude range $<$ 20°E or latitude range $<$ 10°N), flag denoted as D. Unrealistically large or small domains may be an indication of problems associated with the automatic determination of the optimal domain. *Benestad* [2002a] showed that large domains encompassing both regions with positive and negative correlation, such as the North Atlantic Oscillation (NAO) seesaw, can produce spurious results. A final quality test involved checking whether the downscaled results had realistic variance by comparing the downscaled results obtained by the ERA40 to that of the GCM. The flag $f_{var} = 1/24(\sum_{i=1}^{12} \mathcal{H}[s(\text{ERA40}) - 0.5s(\text{GCM})] + 1/12 \sum_{i=1}^{12} \mathcal{H}[s(\text{GCM}) - 1.5s(\text{ERA})])$ indicates mismatch between the downscaled results from the re-analysis and from GCM, s here being the standard deviation of the downscaled results. These quality-control models are by no means objective and only give a crude indication about the GCM qualities. Thus, these should only be regarded as 'rule-of-thumb' indication about the GCM's ability to describe the temperature or precipitation over northern Europe. The quality indeces f_s, f_m and $n_{m<0}$ were only applied to temperature. The final weighting for each station i was taken as $w_i = 1/9 \sum_{j=1}^{M} (r_{xy} + n_m + f_s + f_m + n_{R^2<0.6} + n_{p>0.1} + n_{m<0} + D + f_{var})$ for temperature and $w_i = 1/6 \sum_{i=j}^{M} (r_{xy} + n_m + n_{R^2<0.6} + n_{p>0.1} + D + f_{var})$ for precipitation, summing over the multi-model ensemble size $M = 23$ for temperature and $M = 21$ for precipitation (auxiliary Table 2).

[1] http://en.wikipedia.org/wiki/P-value

Auxiliary Table 3 shows the results of quality control based on the E-SDS diagnostics for the temperature. According to these indices, the UKMO-HadCM3, CCSM-2.0, MRI, IPSL, GISS-AOM, CNRM and the INM-CM3.0 were flagged for a high rate (∼10–20%) with too high summer variance. All models had a low rate with unrealistic seasonal mean temperatures (f_m), but GISS-AOM and CCSM-2.1 were flagged with up to 37% cases with negative temperature trends ($n_{m<0}$). Only GISS-ER was flagged with high number of weak regression ($n_{R^2<0.6}$), whereas $n_{p>0.1}$ and D was reasonably low for all GCMs. Several models had problems with describing realistic variance, with MRI the worst($f_{var} \sim 50\%$). IPSL, NCAR-PCM, and GFDL-2.x had relatively few cases with calibration-prediction variance mismatch. The spatial correlation score was similar for most models (r_{xy} ∼40–46%) and only INMCM3 indicated significantly erratic seasonal evolution of trend estimates ($n_m \sim 7.5\%$). The threshold value (sensitivity) for setting these flags has been set fairly arbitrarily which means that some weights contribute more to the overall weighting scheme than others. Hence, the flags with higher number of cases, such as r_{xy}, f_s and f_{var} tend to be more important than the others for deriving the map products.

Auxiliary Figures 1–2 show time series plots of the (unweighted) downscaled multi-model results for both the 20th and 21st (SRES A1b) centuries. The shaded areas show the inter-model spread and the lines show 3rd-order polynomial trend fits [*Benestad*, 2003] to the multi-model median. The downscaled temperatures for the 20th century (auxiliary Figure 1) show a good agreement with the corresponding observed values, fulfilling the minimum requirement that the analysis must provide a reasonable description of the past trends [*Benestad*, 2003]. Furthermore, the inter-model spread is smaller than the long-term increase, suggesting a reasonably high signal-to-noise ratio. The comparison between E-SDS results for the precipitation for the past and observations suggests that the analysis does not capture as much of the variability as for temperature. Hence, the magnitude of the positive precipitation trends seen for the future are likely to be underestimated. The signal-to-noise ratio for precipitation tends to be much smaller than for temperature, as the long-term change is comparable to the inter-model spread.

Mapping

The mapping of the results was based on similar analysis as in *Benestad* [2004b]. Multi-model ensemble w_i quality-weighted mean linear trends for annual mean values over the period 2000–2099 were estimated for each station location, thus providing a Bayesian-type quality-weighted trend estimate. The technique used in present analysis differs from that of *Benestad* [2004b] by using the square-root distance from the coast ($\sqrt{d}$) as opposed to a linear relation with distance, and including two additional geographical predictors: north–south slope and east–west slope (Auxiliary Table 3). The east–west and north–south slopes were estimated through a stepwise multiple regression fit to $N_\theta = N_\phi = 35$ harmonics to the topographical cross-section profile following equation (1) and then solving for the derivatives according to equation (2). Only the locations on the European continent (marked with grey symbols) were used in calibrating the GRM, but trend estimates for islands in the Norwegian Sea, Iceland, and Greenland are given as numbers in Figures 1–2.

$$z(\theta) = z_0 + \sum_{i=1}^{N_\theta}[a_\theta(i)\cos(\omega_\theta(i)\theta) + b_\theta(i)\sin(\omega_\theta(i)\theta)],$$
$$z(\phi) = z_0 + \sum_{i=1}^{N_\phi}[a_\phi(i)\cos(\omega_\phi(i)\phi) + b_\phi(i)\sin(\omega_\phi(i)\phi)], \quad (1)$$

$$\frac{\partial\hat{z}(\theta)}{\partial\theta} = \sum_{i=1}^{N_\theta}\omega_\theta(i)[-\hat{a}_\theta(i)\sin(\omega_\theta(i)\theta) + \hat{b}_\theta(i)\cos(\omega_\theta(i)\theta)],$$
$$\frac{\partial\hat{z}(\phi)}{\partial\phi} = \sum_{i=1}^{N_\phi}\omega_\phi(i)[-\hat{a}_\phi(i)\sin(\omega_\phi(i)\phi) + \hat{b}_\phi(i)\cos(\omega_\phi(i)\phi)], \quad (2)$$

Since, spherical coordinates were used, a transformation was done to x- and y-coordinates following equation (3).

$$\frac{d\hat{p}(x)}{dx} = \frac{1}{a\cos(\phi)}\frac{d\hat{p}(\theta)}{d\theta},$$
$$\frac{d\hat{p}(y)}{dy} = \frac{1}{a}\frac{d\hat{p}(\phi)}{d\phi}, \quad (3)$$

The harmonics fit, differentiation, and transformation was done in the `R`-environment, using the `geoGrad` function in the contributed `cyclones`-package (version 1.1-4). Both the `R`-environment and `cyclones` are open source and freely available from http://cran.r-project.org (henceforth referred to as 'CRAN').

Auxiliary Table 4 shows the ANOVA of the geographical regression model (GRM) for both temperature and auxiliary Table 5 for precipitation. The use of ANOVA tables is a standard way of assessing regression models in statistics, where high R^2 scores, F-ratios and low p-values indicate strong and significant relationships [*Wilks*, 1995, p. 165–169]. The R^2 obtained for the annual mean T(2m) trend was 66%, an F-ratio of 32.2 and a p-value of 5×10^{-15}, all showing a strong relationship. Only the parameters identified as important are shown in Auxiliary Table 4–5, suggesting that the local annual mean temperature trend is not strongly influenced by e.g. the north–south slope. The corresponding statistical relationship for precipitation was weaker, however, still statistically significant to a high degree: 33% of the precipitation trends, F-ratio=10.8, and p-value of $4{\times}10^{-7}$.

The GRMs can be assessed further in split-sample tests, where part of the data was used for model calibrating (dependent) and the rest as independent data for evaluation [*Benestad*, 2004c]. Auxiliary Figures 3–4 show scatter plots between the original data and the predicted values for both the dependent (grey) and independent data (black) and provide a verification of the GRM.

A kriging analysis [*Matheron*, 1963] was applied to the residuals of the GRM in order to spatially interpolate the part of the trends that could not be related to geographical parameters (using the `geoR`-package from CRAN). Kriging is a standard method used for spatial interpolation in geo-sciences, and an evaluation of the kriging methodology is outside the scope of this paper. Whereas *Benestad* [2004b] used longitude and latitude as two independent variables representing the coordinates, the present analysis east–west and north–south displacements from the central point of the set of locations, in units of 10km. Due to the Earth's curvature, a difference of one degree at high latitudes corresponds to a smaller zonal displacement than at lower latitudes.

Digital data

The two maps presented in present paper are available digitally and stored as netCDF files in which the values represent the trends in the units °C/decade and mm/month per decade (file names: 'Europe_E-SDS_t2m-trend_map.nc' & 'Europe_E-SDS_t2m-precip_map.nc'). The weighted multi-model ensemble mean trends for the station locations are also available in ASCII format (file names: 'Europe_E-SDS_t2m-trend_tab.txt' & 'Europe_E-SDS_t2m-precip_tab.txt').

To give an idea of the scale of the analysis presented here, the grand total number downscaling exercises, i.e. the number of different combination of calendar month, location, scenario (20th century & SRES A1b), GCM and run, was $12{\times}8934$ for temperature and $12{\times}8657$ for precipitation (the analysis takes a couple of weeks on an ordinary Linux PC). The quality control and mapping analysis come on top of this downscaling analysis.

References

Barnett, T.P.(1999) Comparison of Near-Surface Air Temperature Variability in 11 Coupled Global Climate Models *J. Clim.*, *12*, 511–518

Benestad, R.E.(2004a) Empirical-Statistical Downscaling in Climate Modeling *Eos*, *85* (42), 417

Benestad, R.E.(2004b) Tentative probabilistic temperature scenarios for northern Europe *Tellus*, *56A*, 89–101

Benestad, R.E.(2004c) Are temperature trends affected by economic activity? Comment on McKitrick & Michaels *Clim. Res.*, *27*, 171–173

Benestad, R.E.(2003) What can present climate models tell us about climate change? *Climatic Change*, *59* 311–332

Benestad, R.E.(2002b) Empirically downscaled multi-model ensemble temperature and precipitation scenarios for Norway, *J. Climate*, *15*, 3008–3027

Benestad, R.E.(2002a) Empirically downscaled temperature scenarios for northern Europe based on a multi-model ensemble *Clim. Res.*, *21*, 105–125

Benestad, R.E.(2001) A comparison between two empirical downscaling strategies *Int. J. Climatology*, *21*, 1645–1668, DOI 10.1002/joc.703

Benestad, R.E.(2000) Analysis of gridded sea level pressure and 2-meter temperature for 1873-1998 based on UEA and NCEP re-analysis II *The Norwegian Meteorological Institute*, KLIMA-report *03/00*, Oslo, Norway, www.met.no

Blackmon, M.B., B. Boville, F. Bryan, R. Dickinson, P. Gent, J. Kiehl, R. Moritz, D. Randall, J. Shukla, S. Solomon, G. Bonan, S. Doney, I. Fung, J. Hack, E. Hunke, J. Hurrell and et al. (2001) The Community Climate System Model *Bul.Am.Met.Soc.*, *82*, 2357–2376

Delworth, L.D. et al. GFDL's CM2 global coupled climate models - Part 1: Formulation and simulation characteristics *J. Clim*, *submitted*, 2004

Déqué M. and Dreveton C. and Braun A. and Cariolle D. (1994) The ARPEGE-IFS atmosphere model: a contribution to the French community climate modelling *Clim. Dyn.*, *10*, 249–266

Dufresne, J.L. and P. Friedlingstein (2000) Climate-carbon feedbacks associated with CO2 anthropogenic emissions using the IPSL coupled model: an amplification effect? *Lettre IGBP-France*, *N.11*

Diansky N.A. and Volodin E.M (2002) Simulation of present-day climate with a coupled Atmosphere-ocean general circulation model *Izv. Atmos. Ocean. Phys. (Engl. Transl.)*, *38* (6), 732–747

Giorgetta, M.A. and Manzini, E. and Roeckner, E. (2002) Forcing of the quasi-biennial oscillation from a broad spectrum of atmospheric waves *GRL*, *29*, doi 10.1029/2002GL014756

Gordon, C., C. Cooper, C.A. Senior, H. Banks, J.M Gregory, T.C. Johns, J.F.B. Mitchell and R.A. Wood (2000) The Simulation of SST, Sea Ice Extents and Ocean Heat Transports in aversion of the Hadley Centre Coupled Model without Flux Adjustments *Clim.Dyn.*, *16*, 147–168

Hasumi, H. and S. Emori (2004) K-1 Coupled GCM

(MIROC) Description *K-1 Technical Report*, 1, Center for Climate System Research (CCSR), University of Tokyo; National Institute for Environmental Studies (NIES); Frontier REsearch Center for Global Change (FRCGC)

Kiehl, J. T. and P. R. Gent The Community Climate System Model, version 2 *J.Clim.*, *accepted 2004*

Kitoh, A. and A. Noda and Y. Nikaidou and T. Ose and T. Tokioka (1995) AMIP simulations of the MRI GCM *Papers in Meteorology and Geophysics*, *45*, 121–148

Lorenz, E.N. (1956) Empirical Orthogonal Functions and Statistical Weather Prediction *Sci. rep.*, *1*, Department of Meteorology, MIT, USA

Lucarini, V. and G.L. Russell (2002) Comparison of Mean Climate Trends in the Northern Hemisphere between N.C.E.P. and Two Atmosphere-Ocean Model Forced Runs *JGR*, *107*, ACL 7, doi:10.1029/2001JD001247

Matheron, G (1963) Principles of geostatistics *Economic Geology*, *58*, 1246–1266

R Development Core Team (2004) *R: A language and environment for statistical computing* ISBN 3-900051-07-0, Vienna, Austria, http://www.R-project.org

Russell, G.L. and J.R. Miller and D. Rind (1995) A Coupled Atmosphere-Ocean Model for Transient Climate Change Studies *Atmosphere-Ocean*, *33*, 683–730

Schmidt et al. Present day atmospheric simulations using GISS ModelE: Comparison to in-situ, satellite and reanalysis data *J.Clim.*, *submitted*, 2004

Simmons, A.J., P.D. Jones, V. da Costa Bechtold, A.C.M. Beljaars, P.W. Kållberg, S. Saarinen, S.M. Uppala, P. Viterbro and N. Wedi (2004) Comparison of trend and variability in CRU, ERA40 and NCEP/NCAR analyses of monthly-mean surfaceair temperatures *ERA40 Project Report Series* 18, ECMWF, www.ecmwf.int

von Storch, H. (1999) On the Use of "Inflation" in Statistical Downscaling *J.Clim.*, *12*, 3505–3506

Wilks, D.S. (1995) Statistical Methods in the Atmospheric Sciences *Academic Press*, Orlando, Florida, USA

R.E. Benestad, Meteorological Institute, PO Box 0313, Oslo, Norway (e-mail: rasmus.benestad@met.no)

(Received September 27, 2005)

Table 1. Summary of the GCMs used to simulate future climates following the SRES A1b emission scenario. The GCM results were taken from PCMDI.

Centre	Country	GCM	reference
National Center Atm. Research	USA	CCSM3	[*Blackmon et al*, 2001]
Météo-France and Centre National de Recherches Météo.	France	CNRM-CM3	[*Déqué et al.*, 1994]
Max Planck Inst. Meteorology	Germany	ECHAM5/MPI-OM	[*Giorgetta et al.*, 2002]
US Dept. Commerce, NOAA and Geophysical Fluid Dynamics Lab (GFDL)	USA	GFDL-CM2.0	[*Delworth et al*, 2004]
NOAA and GFDL	USA	GFDL-CM2.1	
NASA / Goddard Inst. for Space Studies	USA	GISS-AOM	[*Russell et al.*, 1995]
NASA / Goddard Inst. for Space Studies	USA	GISS-EH	[*Schmidt et al.*, 2004]
NASA / Goddard Inst. for Space Studies	USA	GISS-ER	[*Lucarini and Russell*, 2002]
Inst. Numerical Mathematics	Russia	INM-CM3.0	[*Diansky and Volodin*, 2002]
Institut Pierre Simon Laplace	France	IPSL-CM4	[*Dufresne and Friedlingstein*, 2000]
National Inst. Env. Studies and Frontier Res. Center Glob. Change			
Meteor. Research Institute	Japan	MRI-CGCM2.3.2	[*Kitoh et al.*, 1995]
National Center Atm. Research	USA	PCM	[*Kiehl and Gent*, 2004]
UK Met Office / Hadley Centre	UK	UKMO-HadCM3	[*Gordon et al.*, 2000]

Table 2. Number of different GCM runs used in the multi-model ensemble.

GCM	Run	m_{GCM}
Temperature		
CNRM-CM3	1	1
GFDL-CM2.0	1	1
GFDL-CM2.1	1	1
GISS-AOM	1,2	2
GISS-EH	1–3	3
GISS-ER	4	1
INM-CM3.0	1	1
IPSL-CM4	1	1
ECHAM5/MPI-OM	1–3	3
MRI-CGCM2.3.2	1–5	5
CCSM3	1,2	2
PCM	2	1
UKMO-HadCM3	1	1
		23
Precipitation		
CNRM-CM3	1	1
GFDL-CM2.0	1	1
GISS-AOM	1,2	2
GISS-EH	1–3	3
GISS-ER	4	1
INM-CM3.0	1	1
IPSL-CM4	1	1
ECHAM5/MPI-OM	1,3	2
MRI-CGCM2.3.2	1–5	5
CCSM3	1,2	2
PCM	2	1
UKMO-HadCM3	1	1
		21

Table 3. Results from quality-control for GCMs with SRES A1b results for T(2m). The values are given in %.

GCM	f_s	f_m	$n_{m<0}$	$n_{R^2<0.6}$	$n_{p<0.1}$	D	f_{var}	r_{xy}	n_m	N^*
CNRM-CM3	10.34	0.57	4.02	8.05	2.30	4.885	46.265	45.98	1.72	174
GFDL-CM2.0	4.60	0.57	0.57	7.47	2.30	4.885	20.405	43.68	0.57	174
GFDL-CM2.1	5.78	0.00	16.76	7.51	2.31	4.915	29.190	43.93	1.73	173
GISS-AOM	15.52	0.29	29.60	7.18	1.72	4.885	49.715	44.83	0.57	348
GISS-EH	3.26	0.19	6.32	6.90	2.30	4.885	34.005	41.38	0.57	522
GISS-ER	5.75	0.00	2.87	10.34	2.87	4.885	46.550	43.10	1.72	174
INM-CM3.0	17.82	0.57	0.57	7.47	2.30	4.885	41.380	44.83	7.47	174
IPSL-CM4	16.67	0.00	1.15	6.90	2.30	4.885	25.575	44.83	0.00	174
ECHAM5/MPI-OM	13.63	0.38	0.77	7.49	2.69	4.895	42.805	47.60	0.77	521
MRI-CGCM2.3.2	16.91	0.57	0.00	6.97	2.29	4.855	50.345	45.03	1.37	875
CCSM3	18.29	0.57	0.57	7.14	2.29	4.855	43.430	45.14	1.43	350
PCM	0.57	0.57	1.14	7.43	2.29	4.855	29.145	41.71	0.57	175
UKMO-HadCM3	23.43	0.57	1.14	6.86	2.29	4.855	41.715	45.71	0.00	175

Table 4. A summary of the multiple regression used to model the geographical distribution of the annual mean temperature warming rate. The values of the residuals were in the range -0.098–0.19, with a residual standard error of 0.0452. The multiple R^2 was 0.6578 (adjusted R-squared: 0.6374), the F-statistic= 32.2 on 4 and 67 degrees of freedom, and the p-value= 5×10^{-15}. The significance codes in the sixth columns are for p-values 0 ('***'), 0.001 ('**'), and 0.01 ('*').

Coefficients:

	Estimate ± Std. Error	t value	Pr(> $\|t\|$)	Signif.
(Intercept)	$(2.261 \pm 0.100)\times 10^{-1}$	22.598	< 2e-16	***
$\sqrt{\text{dist (10 km)}}$	$(1.020 \pm 0.612)\times 10^{-2}$	1.668	0.1000	
z (m)	$(1.321 \pm 0.593)\times 10^{-4}$	2.224	0.0295	*
y (10 km)	$(3.749 \pm 1.424)\times 10^{-4}$	2.634	0.0105	*
x (10 km)	$(5.997 \pm 1.326)\times 10^{-4}$	4.523	2.55e-05	***

Table 5. Same as Table 4, but for preciptation trends. The residual standard error is 0.2339 on 87 degrees of freedom, with a range of -0.5780295 – 0.8590512, the multiple R^2 was 0.3312 (Adjusted R^2: 0.3005), F-statistic= 10.77 on 4 and 87 DF, and the p-value= 4×10^{-7}.

Coefficients:

	Estimate ± Std. Error	t value	Pr(> $\|t\|$)	Signif.
(Intercept)	$(5.351 \pm 0.255)\times 10^{-1}$	20.966	< 2e-16	***
y (10 km)	$(8.620 \pm 5.595)\times 10^{-5}$	1.541	0.12702	
x (10 km)	$(1.899 \pm 0.614)\times 10^{-4}$	3.096	0.00264	**
dz/dy (m/m)	$(1.822 \pm 0.625)\times 10^{4}$	2.916	0.00451	**
dz/dx (m/m)	$(5.835 \pm 4.066)\times 10^{3}$	1.435	0.15485	

Figure 1. Plume plots showing the E-SDS results for the 20th century (grey) and 21st century (blue) together with the actual observations (black points). The darker grey/blue regions mark the multi-model inter-quartile range whereas the lighter regions show the 5%–95% range. The thick line is a 3rd order polynomial trend fit to the multi-model median. Panel (a) shows the E-SDS results for Oslo, panel (b) for Bergen and panel (c) for Tromsø.

Figure 1. Plume plots showing the E-SDS results for the 20th century (grey) and 21st century (blue) together with the actual observations (black points). The darker grey/blue regions mark the multi-model inter-quartile range whereas the lighter regions show the 5%–95% range. The thick line is a 3rd order polynomial trend fit to the multi-model median. Panel (a) shows the E-SDS results for Oslo, panel (b) for Bergen and panel (c) for Tromsø.

Figure 2. Same as Figure 1, but for precipitation.

Figure 2. Same as Figure 1, but for precipitation.

Figure 3. Split-sample evaluation of the GRM for temperature, where half of the data was used for model calibration (dependent) and the remaining as independent data for verification.

Figure 3. Split-sample evaluation of the GRM for temperature, where half of the data was used for model calibration (dependent) and the remaining as independent data for verification.

Figure 4. Same as Figure 3, but for precipitation.

Figure 4. Same as Figure 3, but for precipitation.

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

BENESTAD: MULTI-MODEL IPCC AR4

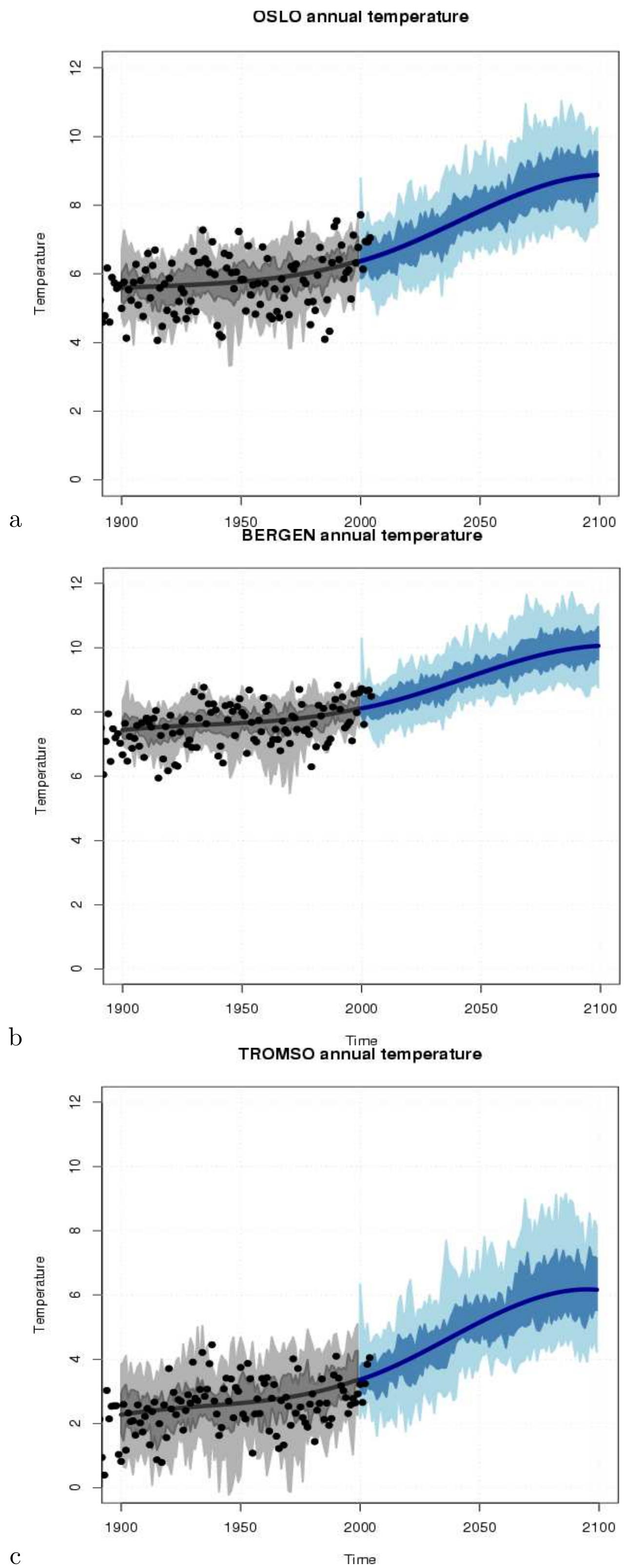
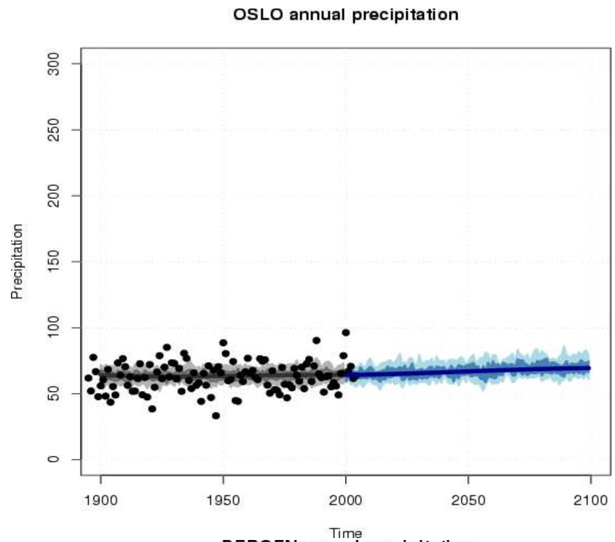
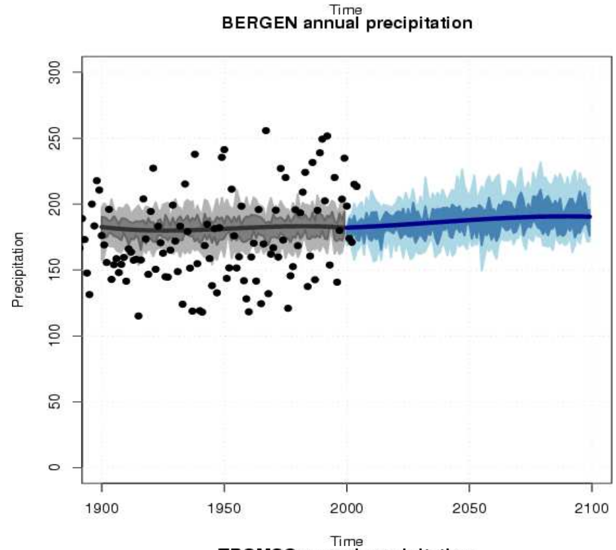

OSLO annual temperature
Temperature
a
BERGEN annual temperature
Temperature
Time
b
TROMSO annual temperature
Temperature
Time
c
0
2
4
6
8
10
12
1900
1950
2000
2050
2100

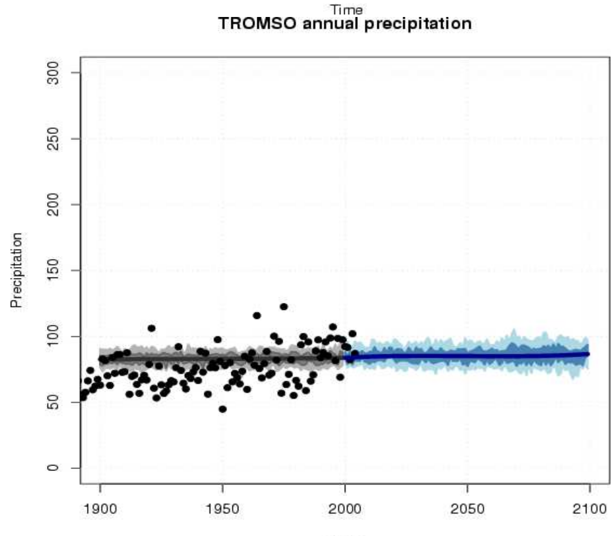
Figure 1:


OSLO annual precipitation
Precipitation
0
50
100
150
200
250
300
1900
1950
2000
2050
2100
Time

a


BERGEN annual precipitation
Precipitation
0
50
100
150
200
250
300
1900
1950
2000
2050
2100
Time

b


TROMSO annual precipitation
Precipitation
0
50
100
150
200
250
300
1900
1950
2000
2050
2100
Time

c

Figure 2: