6.2

FROM GLOBAL MEAN RESPONSES TO REGIONAL SIGNALS OF CLIMATE CHANGE: SIMPLE PATTERN SCALING, ITS LIMITATIONS (OR LACK OF) AND THE UNCERTAINTY IN ITS RESULTS

Claudia Tebaldi,* Doug Nychka and Linda O. Mearns

National Center for Atmospheric Research

1   INTRODUCTION

The simple-pattern-scaling idea (Santer (1990)) has been recently tested for average temperature and precipitation change signals, and proven to perform extremely satisfactorily (Ruosteenoja (2003), R&al.03 from now on).

In this paper we want to address two important aspects of the application of pattern-scaling to climate-change projections, that have been insufficiently studied so far. The first one is the evaluation of the scaling method for different horizontal scales of resolution, i.e. for different scales of regional aggregation. The second is the computation of uncertainty measures, to be attached to the scaled values, in order to provide a degree of confidence in the robustness of the scaled signal.

We use runs of the NCAR PCM-DOE model (Washington (2000)), performed under the two SRES scenarios A2 and B2 (Nakicenovic (2000)), and we study average surface temperature and precipitation patterns of change between the two periods of 1961-1990 and 2070-2099, separately for the four seasons.

The different horizontal resolutions at which the method is tested are defined either by recursively partitioning the 32 regions used in R&al.03, along perpendicular lines passing through the regions' centroids, or by a climatology-based definition of regions, in which observed climatological variables are used for clustering gridpoints, at different levels of aggregation. The first method provides regions encompassing contiguous areas. The second method creates a nested series of partitions based on climatological similarities, but the regions may not be connected.

The main goal of this study is to explore how the performance of the scaling method varies when going from signals aggregated over vast regions to signals at finer levels of aggregation, and how the uncertainty characterizing the results varies in the process.

---
* *Corresponding author address*: Claudia Tebaldi, National Center for Atmospheric Research, Environmental and Societal Impacts Group, Boulder, CO 80307-3000; email: tebaldi@ucar.edu

2   THE SIMPLE PATTERN SCALING METHODOLOGY

For more than a decade the idea of using fixed, AOGCM derived, geographical patterns of climate change – under the hypothesis that they can be considered constant across different emission scenarios, at least if the latter are characterized mainly by forcings from long-lived, well mixed gases – and rescale them on the basis of a global mean change signal, perhaps derived, for any scenario, from a simpler energy-balance model, has been demonstrated to be a valid one (Santer (1994), Mitchell (1999)).

The essence of the method can be summarized as follows: Define the normalized (per 1 deg C mean global warming) pattern of change for variable $\mathbf{X}$,

$$\Delta \hat{\mathbf{X}}(x)$$

estimated for each spatial location x, e.g. from the run of a fully coupled GCM under some forcing pattern. Then, under the hypothesis that the response signal patterns are largely independent of the forcing patterns and remain approximately constant in time, the signal pattern at any future time t can be computed as

$$\Delta \mathbf{X}(x,t) = \Delta T(t) \cdot [\Delta \hat{\mathbf{X}}(x)]$$

where $\Delta T(t)$ is the global mean *temperature* change. Notice that the scaling is done using the global mean change in temperature, for any variable $\mathbf{X}$.

For our exercise, both $\Delta T(t)$ and $\Delta \hat{\mathbf{X}}(x)$ are derived from the NCAR PCM-DOE runs under respectively, the A2 and B2 emission scenarios. We will be thus computing the scaled $\Delta \mathbf{X}(x,t)$ for the A2 emission scenario from normalized spatial patterns of change derived from the B2 run, rescaled by the global mean temperature change signal as simulated under the A2 run. In this idealized situation, we will be able to compare the scaled signal of change to the true one, i.e. the one derived directly from the A2 run of the PCM. In a more realistic setting, $\Delta T(t)$ would be derived by running a simpler model (an energy balance model, for

example, that just produces global mean responses) under the A2 scenario. Such simpler models can be tuned to reproduce global average responses of fully coupled AOGCMs, through setting of some free parameters.

Following R&al.03, we scale surface temperature and percent precipitation changes. "Change" is defined as the difference between 30-yr averages computed from year 1961 through 1990 and 2070 through 2099. We consider this large time interval in the spirit of the method, which dictates the use of the *signal of change* for the scaling exercise, that we take to be the difference between two "long enough" time averages taken at "distant enough" times. We do so separately for the four seasons defined as usual (DJF/MAM/JJA/SON).

3 FROM GRIDPOINTS TO REGIONS AND BACK 'DOWN' AGAIN

We start by scaling the two variables' pattern of change at the finest level, i.e. the gridpoint values. For the sake of conciseness we show images of a single season (DJF for temperature, MAM for precipitation). Figure 1 show global images of the temperature change signals from the A2 and B2 scenarios in the first two panels, and the scaled A2 derived through simple pattern scaling from the B2 run in the third panel. The last panel shows the difference between scaled signal and true signal, divided by an estimated measure of natural variability at each gridpoint. The latter has been derived by computing the standard deviation of the corresponding — gridpoint specific — time series of temperature from an all-forcings run used to create the baseline climate of the 1961–1990 period. Figure 2 shows the corresponding images for the precipitation change signal (as % of present-day average precipitation). Notice, first of all, the different spatial structure of the temperature fields compared to the precipitation fields. The former are much smoother, the errors' value being close to zero over most of the globe, with a few patches of underestimated change over the ocean and areas of overestimated change (slightly less in absolute value) over the high latitudes of the northern hemisphere and the land masses of South America, and Africa. From a comparison across seasons (not shown), winter (DJF) seems to present the largest errors in absolute value, summer (JJA) the smallest. Precipitation fields are relatively harder to interpret, because of the more irregular nature of the spatial patterns. It is still true that the scaling method performs well on average, with most of the field well within the limits of natural variability (i.e. the values being for the most part less than 1) but the error field appears interspersed of small areas where the error is as large as twice the natural variability in absolute value. The range of the errors

appears similar for temperature and precipitation.

The values represented in the second and third panel of Figures 1 and 2, the true and scaled signals of temperature and precipitation change, and the corresponding ones for the three seasons not shown, constitute the basic ingredients of the first part of our analysis. We will aggregate the scaled and true values into regional averages, for regions of different sizes, look at how the two sets of averaged values compare to each other, and how the comparison changes at different spatial scales of aggregation. As for the *error values*, i.e. the differences between the scaled and true patterns of change represented in the fourth panel of Figures 1 and 2, we will model them at the gridpoint level, and use the estimated statistical structure to attach uncertainty measures at each scale of aggregation.

3.1 BOXY REGIONS AND THEIR REPARTITIONING

We start from the regional aggregation used by R&al.03. The 32 regions, represented in Figure 3, have been derived on the basis of climatological principles, and have fairly different sizes, although they all range over subcontinental scales of around $10^6/10^7 \mathrm{km}^2$. Figure 4 for temperature change and 5 for percent precipitation change reproduce the results in R&al.03, showing how the scaled values aggregated over each region compare to the true values, identically aggregated. The four plots in each panel correspond to the four seasons. A perfect match would be represented by points on the straight $45\,\mathrm{deg}$ line plotted as a reference, and the points in each plot do not depart dramatically from this reference line. The root mean square error (RMSE) and the R^2 value for each scatterplot are also printed.

It is evident how, at this spatial scale, the correspondence of scaled to true values is extremely tight, for both temperature and percent precipitation changes. Notice though how such representation does not allow for any assessment of uncertainty in the results, which is one of our goals. Also, there are differences in the performance of the method across seasons, with the cooler (adopting a Northern hemisphere point of view) seasons, DJF and SON, being less accurately approximated, at least in terms of RMSE than the warmer seasons, JJA and MAM. Lastly, precipitation change is comparatively harder to scale than temperature change, as it is to be expected.

Next, we recursively subdivide each region into four subregions (quadrants) and recompute RMSE and R^2 values, for each of the new set of regions. Table 1 and 2 show these values for temperature and precipitation, respectively. Notice how the values of RMSE become larger and, correspondingly, the R^2 values de-

crease, although both do so in a fairly gradual fashion. There does not seem to exist a particular scale at which the validity of the scaling method starts to fall apart. Rather, the linearity assumption is increasingly eroded for finer and finer scales of resolution. Still, even at the finest scale (close to the gridpoint level), the linearity of the relation is obvious. What these measures of performance cannot address, though, is the possible presence of a systematic spatial pattern in the errors. Again, this is the object of the second part of our analysis, when we model the error fields derived from the difference between true and scaled values at the gridpoint level.

3.2 CLIMATOLOGY-BASED SUBREGIONS

On the basis of a climatology dataset available on the IPCC Data Distribution Center (DDC) website (http://ipcc-ddc.cru.uea.ac.uk, New (1999)), we build a cluster analysis of the model gridpoints based on the observed and regridded variables:

1. Temperature: mean, maximum and minimum;

2. Precipitation: occurrence and intensity.

.

Each of the variable, at each gridpoint, has been aggregated into 4 seasonal values, representing averages for the period 1961-1990. We regridded the DDC product (the grid format (5 deg resolution) available from the DDC is different from the PCM grid (2.5 deg resolution), by means of linear interpolation. Also, the matrix comprised of the $20 = 4 \times (3 + 2)$ climatology variables, whose rows correspond to gridpoints, was decomposed by a principal component analysis. Seven principal components were found to explain more than 99% of the variance of the original matrix, and the clustering was based on the Euclidean norm of these seven leading rotated components. We inverted the order of the seasons for gridpoints in the Southern hemisphere, in order to allow regions with similar climatology to be clustered together, independently of their geographical location. Figure 6 shows the tree structure of the hierarchical clustering. By 'cutting' the structure at different heights one can identify different numbers of clusters, and on average, given the homogeneous form of the cluster aggregation, the higher the cut, the larger the size of the clusters identified. Conversely, one can start from the top of the structure, corresponding to evaluating the scaling performance at the global level (perfect by construction) and recursively do so for gradually smaller clusters of gridpoints, to end with each gridpoint separately forming its own cluster.

We start by deriving 22 clusters/regions, the same number of land regions as in R&al.03 (notice that the climatology dataset on which we base the clustering is available only over land). Figure 7 shows the corresponding regions, each cluster identified by a different color. By two new sets of scatterplots, Figure 8 for temperature change and Figure 9 for percent precipitation change, we show the performance of scaling for this scale of aggregation, separately for each season, and we also write values of RMSE and R^2 in each of the panels. The performance at this scale is very similar to what was found for the 32 regions derived from R&al.03.

Again, a comparison of the performance measures for different spatial scales of aggregation, obtained by gradually 'relaxing' the cluster structure, is available in Table 3 for temperature and Table 4 for precipitation and shows only a gradual deterioration of the agreement between scaled and true values, without an obvious 'threshold effect' when spanning the spectrum of resolutions. This result is also consistent with what was found by the alternative method of regionalisation.

A few features, common to both types of disaggregation, are noteworthy:

- as expected, temperature changes are scaled more accurately than precipitation;

- the deterioration of the agreement between scaled and true quantities is faster for precipitation than temperature changes;

- temperature changes are scaled consistently well at all spatial scales and in all seasons, with a very slow degradation of the results, with respect to both measures of goodness of fit;

- error in scaling precipitation are significantly larger in SON;

- there is strong consistency in the results between the two methods of aggregation, for both measures of goodness of fit.

4 STATISTICAL MODELS FOR THE GRIDPOINT ERROR: HOW UNCERTAIN ARE THE SCALING RESULTS?

4.1 ERROR-CHARACTERIZATION FOR ΔT

In order to characterize the uncertainty in the scaling results we model the error patterns shown in the fourth panel of Figure 1, and the corresponding patterns for the three seasons not shown as gaussian random fields. After characterizing mean and covariance function of

the season-specific gaussian process, we will simulate replicates of the field and compute summary statistics of the errors, on the basis of the two methods of spatial aggregations, for various scales of resolution.

Modeling these fields requires addressing two issues, immediately apparent when inspecting the images in Figure 1 and 2. The error structure seems to be elongated longitude-wise, rather than extending homogeneously in all directions, and the structure appears different at high latitudes than at low latitudes. Other kind of spatial structure were tested and ruled out, like for example dependence on land/ocean locations and on topography.

Thus we choose to transform the set of grid coordinates according to two geometric anysotropy parameters estimated through sample variograms, computed separately for the East-West and North-South directions, for the high latitude regions. In the transformed coordinates, an isotropic gaussian model is fitted, whith Matern correlation structure, separately for high latitude (between (-90,-30) and (30,90) degrees) and low latitude (between -30 and 30 degrees) regions.

Generating sample replicates after having estimated all model parameters (mean, range, scale and smoothness) follows these steps:

1. A gaussian field, mean zero and variance one of uncorrelated random variables is generated, for the entire grid.

2. The two mean and covariance structures – estimated for low and high latitudes – are applied separately to the same field, in order to transform it into two fields having the desired spatial structure.

3. The two fields are merged by applying smoothly varying (latitude-wise) weights, close to one in the high latitudes and close to zero in the low latitudes for one of the fields, and viceversa for the second (each corresponding couple of weights summing to one).

4. The same procedure is applied as many times as we want replicates, for each seasonal error fields (we choose 100 replicates per season)

Figures 10 shows example replicates of winter (DJF) error fields. One of the panels is the true error field. The captions indicate which one, but the reader may want to guess its position. If the guess turns out to be a difficult one, that will be a first order assessment of the quality of the simulation procedure. Notice that the color palette was chosen to emphasize regions where the error is in absolute value greater than one, thus discounting errors within the range of natural variability

(between -1 and 1) and highlighting errors outside the range.

The goal of this part of our analysis is an assessment of the uncertainty in the scaling results, as a function of season, and spatial scale of aggregation. We are trying to add a distributional characterisation to the single seasonal error realization of Figure 1. Of course, one could do so in abstract, as a function of the Matern covariance parameters, but we choose to use the simulation device, as a more effective depiction of the results. For each season, we are going to aggregate every one of the 100 error fields into the regions defined in R&al.03 and our climatology-based clustering, at their various scales of resolution, and compute standard deviations of the error values within each region, across the 100 error fields. The mean being zero by construction, what matters is the variability of the error values from one realization to the next, in order to assess the uncertainty in the results of a single scaling exercise. We can then compare the distribution of these standard deviations as a function of the spatial scale of aggregation, the season and the method of regionalisation.

A series of boxplots in Figure 12, summarize these distributions. In each column the four plots correspond to the four season (DJF,MAM,JJA,SON). In each figure, plots on the left correspond to the regionalisation based on recursively partitioning the 22 original land regions from R&al.03, while plots on the right correspond to the climatology-based method of regionalisation. The six boxplots in each panel correspond to different spatial scales of aggregation. Corresponding to each boxplot is the approximate average size of the regions expressed in square miles (on a $\log_{10}$ scale). In the legend we also list the number of regions corresponding to each boxplot. For example, the first boxplot in the left panels corresponds to a discretisation into 1792 land regions, the average area of each being on the order of 10,000 square miles. The discretisation is coarser and coarser from left to right, with the last boxplot corresponding to only 18 regions, with an average area of 10 million square miles.

It is striking how similar the series of boxplots in the left column are to the corresponding ones in the right columns, in both figures. This confirms the results obtained in the previous sections, where the pattern-scaling method seemed to perform similarly for the two methods of discretisation. As for seasonal differences, they are not striking, but one could argue that spring shows the least amount of variability.

An important question that the series of boxplots in each panel should answer is if the error values, when averaged across regions, behave differently at different scales of regional discretisation. In fact, an apprecia-

ble difference can be read in the distributions, shifting from higher values in mean and wider ranges to smaller values and tighter ranges with decreasing resolution, suggesting that the variability of average error values for smaller regions from one simulated error field to another is higher than for larger regions. Evidently, this is due to the spatial structure (extent) of the error areas, that are more easily averaged out over larger regions than over smaller ones.

Notice though that all the values are well within natural variability limits, the standard deviations being below 60% even at the finest level of resolution.

4.2 ERROR-CHARACTERIZATION FOR ΔP

We proceed to a similar analysis of the seasonal error fields for % precipitation change, the spring (MAM) case of which was shown in Figure 2. The error values are expressed as percent precipitation change with respect to true percent precipitation change under the A2 scenario, divided by a measure of natural "percent standard deviation". The latter measure was computed with respect to the average precipitation field in present day climate, derived from an all-forcings run used to represent the baseline climate of the 1961 through 1990 period, similarly to what we described in the analysis of temperature.

The same spatial analysis performed for the error fields in the scaling of temperature change was applied. Because of the heterogeneous characteristics of the error fields, easily detectable by eye, we divided each image into bands of low-latitude regions (within ±50 degree latitude) and high latitude regions (outside the ±50 degree limits) and estimated two different spatial covariance models, the one for high latitude anisotropic, i.e. allowing for the covariance structure in the zonal direction to be different from the one in the meridional direction. Similarly to the analysis of the error fields for ΔT, this result was produced by first estimating a deformation of the coordinate system and then applying an isotropic model to the transformed coordinates. Figure 11 shows true and simulated error fields for the MAM case, again with the true error field identified only in the caption. Figure 13 is also addressing the question of how variable, and thus uncertain, are the error values computed at each scale of resolution. Again, the message is a significant increase of the variability for finer scales, and rather constant behavior across seasons, and methods of regionalisation. Still the highest variability at the finest resolution is within 60% of the natural variability.

5 CONCLUSIONS

We have studied the performance of simple pattern-scaling as a function of spatial aggregation, for both temperature and precipitation changes, and found that the method performs accurately (as measured by RMSE and R2 between true and scaled values), down to very fine scales of spatial resolution. The degradation of the performance is gradual, and even at its worst it remains well within the measures of natural variability estimated from a present-day climate run of the same AOGCM used for the future projections.

By modeling the error fields at their finest resolution, the gridpoint level, as gaussian random fields (after a transformation of the coordinates that, in first approximation, accounts for having stretched the globe into a flat rectangle) we obtained measures of uncertainty in the form of the expected standard deviations of the regional-scale errors. By way of simulations, we presented the distribution of the errors as a function of spatial aggregation, confirming once again that the method performs within the range of natural variability.

This is work in progress. We need to make our exercise more realistic, by substituting the "true" global mean measure under the scenario of interest, that we obtained from the same AOGCM run from which the "true" signal was extracted, by the one produced by an intermediate model. This will introduce an additional source of approximation and uncertainty, even if it will be limited to the mean compoent of the error fields, not to its spatial character. Most importantly, we need to conduct the same analysis between many different pairs of scenarios. We have found similar results when scaling B2 to A2, and when scaling a "business as usual" scenario from any of the two SRES scenarios. Additionally, and more interestingly, we need to study the validity of scaling for those scenarios that are characterised by spatially heterogeneous emissions (sulfate aerosols). Scaling among these will undoubtedly require more complex forms of statistical estimation than the simple linear method that worked so well in this instance.

The goal of this work is ultimately to optimize the use of expensive computer experiments by substituting scaled projections (and corresponding measures of uncertainty) to actual AOGCMs' runs. We think however that this will require some action "upstream", in the form of optimal design of computer experiments that effectively explores the space of alternative scenarios, and predicated on which interpolation by statistical means will perform at its best.
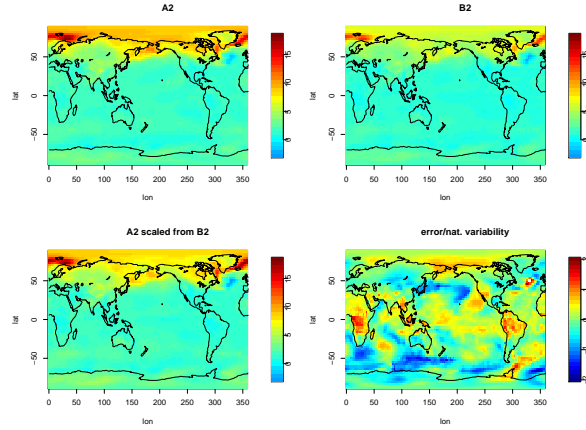
Figure 1: *Signal of temperature change, winter (DJF) season: A2 scenario, B2 scenario, A2 scaled from B2, error between scaled A2 and true A2, divided by a measure of natural variability computed at each gridpoint separately on the basis of an all-forcings, present-day climate PCM run.*
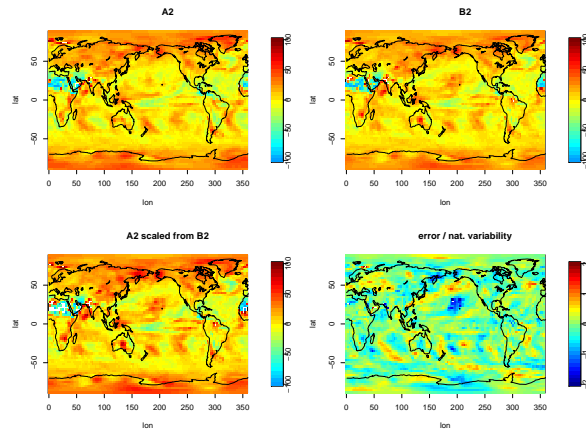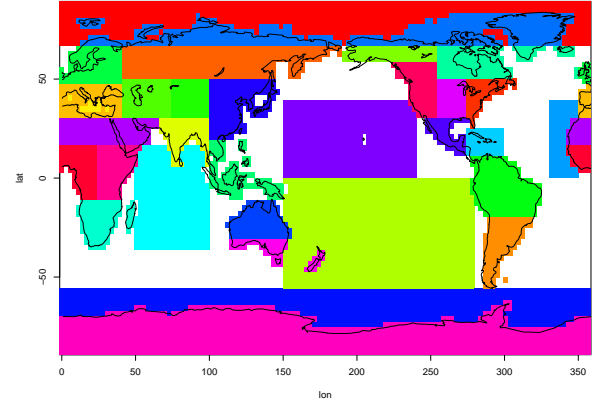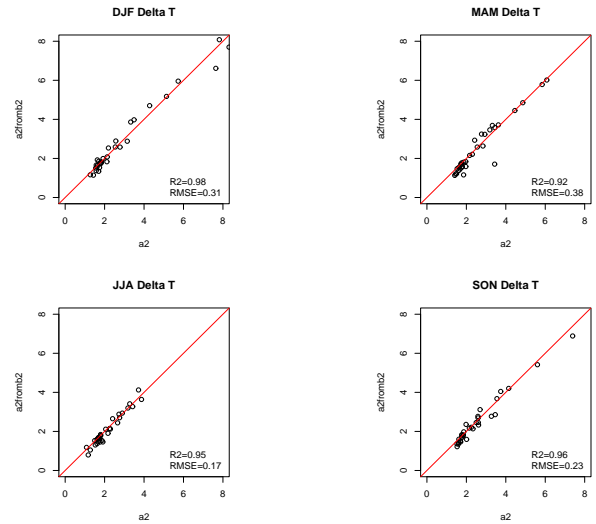

Figure 2: *Signal of % precipitation change, spring (MAM) season: A2 scenario, B2 scenario, A2 scaled from B2, error between scaled A2 and true A2, divided by a measure of natural variability computed at each gridpoint separately on the basis of an all-forcings, present-day climate PCM run.*


Figure 3: *The 32 regions from Ruosteenoia et al.*

Figure 4: *Scatterplot representation (one for each season) of scaled vs. true temperature change for the 32 regions from Ruosteenoia et al.*
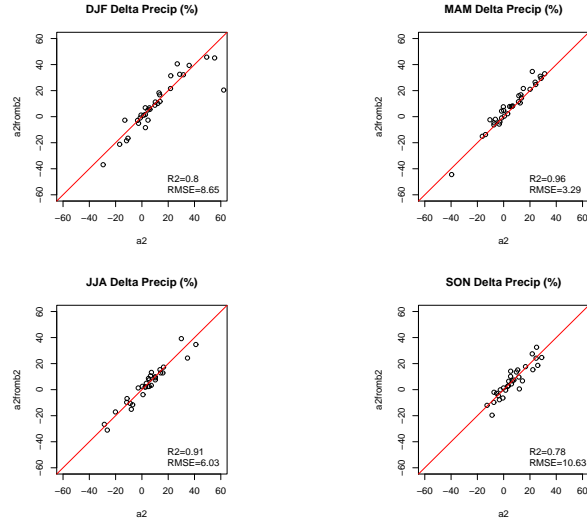
Figure 5: *Scatterplot representation (one for each season) of scaled vs. true % precipitation change for the 32 regions from Ruosteenoia et al..*

Figure 7: *22 land regions derived by climatology-based clustering.*

Figure 6: *The tree structure of the clustering based on climatology records at each gridpoint. Different size clusters are obtained by "cutting" the tree structure at different heights, the extremes being a cut at the very top, returning a single cluster containing all the gridpoints, and a cut at the very bottom, returning as many single-occupancy clusters as there are gridpoints.*
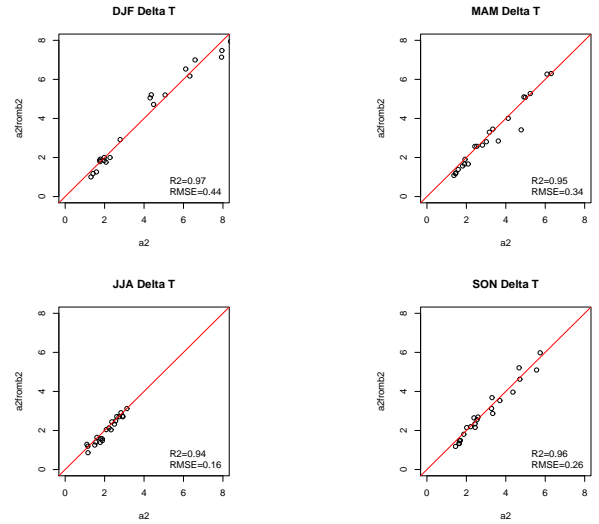
Figure 8: *Scatterplot representation (one for each season) of scaled vs. true temperature change for the 22 land regions based on clustering gridpoints according to climatology.*
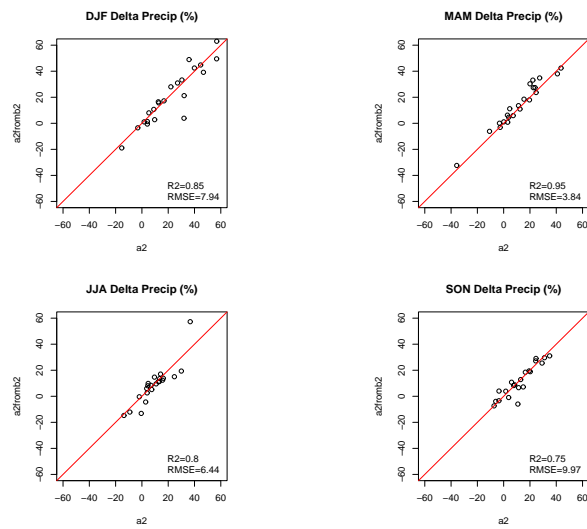
Figure 9: *Scatterplot representation (one for each season) of scaled vs. true % precipitation change for the 22 land regions based on clustering gridpoints according to climatology.*
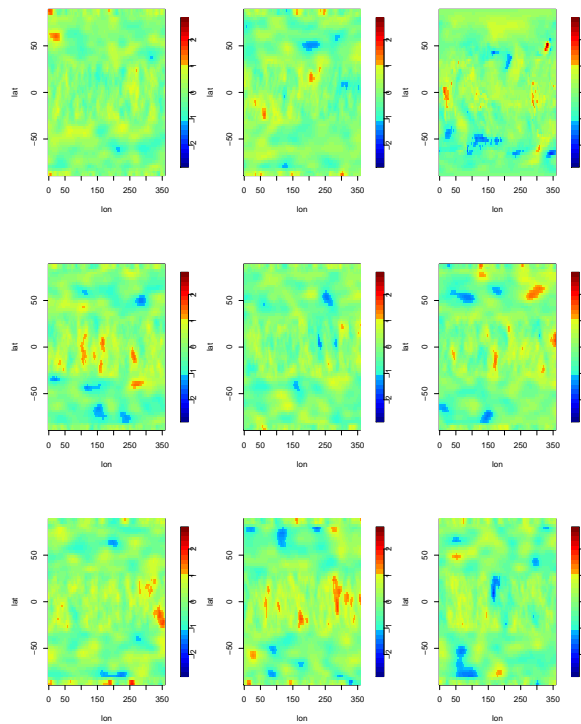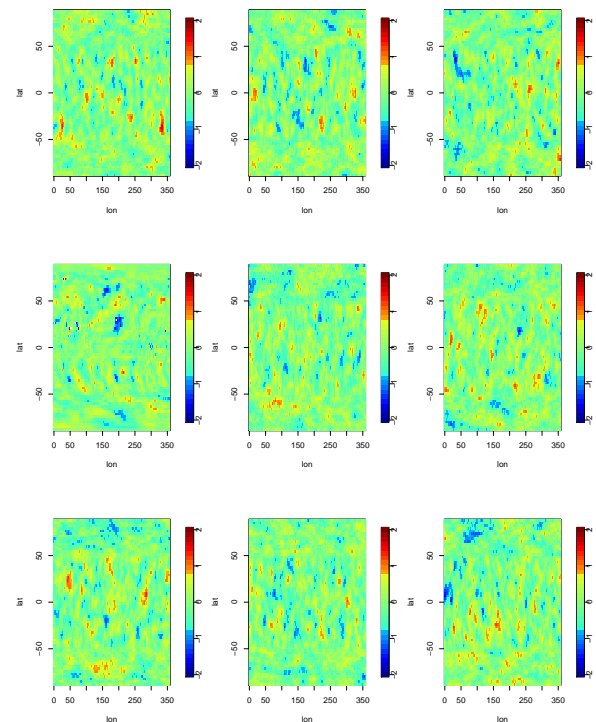

Figure 11: *True and simulated error fields for spring (MAM) scaling of % precipitation change. The true field is the first of the second row.*

Figure 10: *True and simulated error fields for winter (DJF) scaling of temperature change. The true field is the last of the first row.*
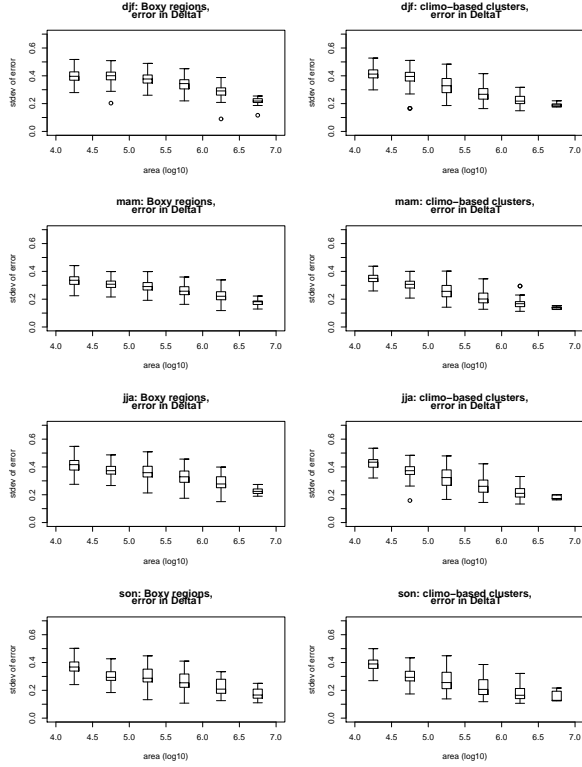
Figure 12: *Distribution (across regions) of the standard deviation of 100 simulated (and averaged over each region) error fields (ΔT). Left column plots pertain to the regionalisation of Ruosteenoia (2003), right column to the regionalisation derived by climatology-based clustering. Left to right boxplots correspond to finer to coarser spatial resolutions for the regionalisation. The number of regions in the 6 boxplots of the left column is 1792, 3155, 550, 236, 76, 18. The number of regions in the 6 boxplots of the right column is 2176, 2117, 530, 171, 58, 8.*
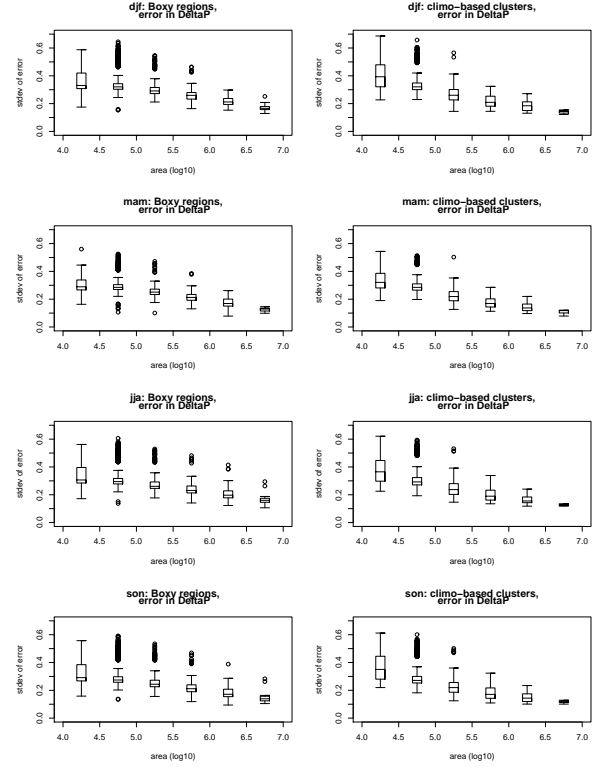
Figure 13: *Distribution (across regions) of the standard deviation of 100 simulated (and averaged over each region) error fields (% ΔP). Left column plots pertain to the regionalisation of Ruosteenoia (2003), right column to the regionalisation derived by climatology-based clustering. Left to right boxplots correspond to finer to coarser spatial resolutions for the regionalisation. The number of regions in the 6 boxplots of the left column is 1792, 3155, 550, 236, 76, 18. The number of regions in the 6 boxplots of the right column is 2176, 2117, 530, 171, 58, 8.*

season	measure	32 regions	128 regions	444 regions	1370 regions	4006 regions
DJF	R^2	0.98	0.97	0.96	0.96	0.96
DJF	RMSE	0.31	0.38	0.48	0.50	0.49
MAM	R^2	0.92	0.93	0.92	0.93	0.93
MAM	RMSE	0.38	0.40	0.42	0.41	0.41
JJA	R^2	0.95	0.94	0.92	0.91	0.90
JJA	RMSE	0.17	0.23	0.27	0.29	0.32
SON	R^2	0.96	0.94	0.92	0.92	0.93
SON	RMSE	0.23	0.32	0.37	0.39	0.38

Table 1: R^2 and RMSE for scaled ΔT with respect to true ΔT, as a function of the number of regions, obtained by recursively repartitioning the 32 regions from Ruosteenoja (2003)

season	measure	32 regions	128 regions	444 regions	1370 regions	4006 regions
DJF	R^2	0.80	0.72	0.73	0.72	0.69
DJF	RMSE	8.65	12.22	13.30	14.85	15.24
MAM	R^2	0.96	0.74	0.63	0.50	0.47
MAM	RMSE	3.29	9.86	14.23	18.93	18.94
JJA	R^2	0.91	0.82	0.76	0.67	0.66
JJA	RMSE	6.03	10.42	17.31	24.10	19.62
SON	R^2	0.78	0.75	0.74	0.61	0.57
SON	RMSE	10.63	13.15	14.99	20.90	19.48

Table 2: R^2 and RMSE for scaled $\%\Delta P$ with respect to true $\%\Delta P$, as a function of the number of regions obtained by recursively repartitioning the 32 regions from Ruosteenoja (2003)

season	measure	22 regions	60 regions	120 regions	240 regions	450 regions	1000 regions	1400 regions
DJF	R^2	0.97	0.97	0.96	0.96	0.95	0.95	0.94
DJF	RMSE	0.44	0.44	0.48	0.53	0.56	0.60	0.61
MAM	R^2	0.95	0.96	0.96	0.94	0.93	0.92	0.91
MAM	RMSE	0.34	0.33	0.35	0.42	0.45	0.50	0.53
JJA	R^2	0.94	0.93	0.94	0.93	0.92	0.92	0.91
JJA	RMSE	0.16	0.18	0.20	0.21	0.24	0.26	0.28
SON	R^2	0.96	0.95	0.94	0.92	0.91	0.91	0.90
SON	RMSE	0.26	0.30	0.35	0.39	0.43	0.44	0.46

Table 3: R^2 and RMSE for scaled ΔT with respect to true ΔT, as a function of the number of regions obtained by hierarchical clustering of the gridpoints on the basis of observed climatology.

season	measure	22 regions	60 regions	120 regions	240 regions	450 regions	1000 regions	1400 regions
DJF	R^2	0.85	0.75	0.73	0.71	0.70	0.62	0.60
DJF	RMSE	7.94	11.77	13.68	17.03	18.20	22.40	22.83
MAM	R^2	0.95	0.85	0.81	0.76	0.70	0.61	0.51
MAM	RMSE	3.84	7.47	9.80	11.06	14.10	19.91	24.22
JJA	R^2	0.80	0.84	0.82	0.75	0.57	0.63	0.58
JJA	RMSE	6.44	8.90	11.73	14.63	20.60	26.05	29.16
SON	R^2	0.75	0.75	0.77	0.68	0.66	0.56	0.5
SON	RMSE	9.97	14.06	15.81	17.56	19.97	29.53	32.02

Table 4: R^2 and RMSE for scaled $\%\Delta P$ with respect to true $\%\Delta P$, as a function of the number of regions obtained by hierarchical clustering of the gridpoints on the basis of observed climatology.

REFERENCES

Mitchell, J.F.B., T.C. Johns, M. Eagles, W.J. Ingram and R.A. Davis, 1999: Towards the construction of climate change scenarios. *Climate Change*, **41**, 547-581.

Nakicenovic, N., J. Alcamo, G. Davis, B. de Vries, J. Fenhann, S. Gaffin, K. Gregory, A. Grubler, T.Y. Jung, T. Kram, E.L. La Rovere, L. Michaelis, S. Mori, T. Morita, W. Pepper, H. Pitcher, L. Price, K. Raihi, A. Roehrl, H.-H. Rogner, A. Sankovski, M. Schlesinger, P. Shukla, S. Smith, R. Swart, S. van Rooijen, N. Victor and Z. Dadi, 2000: *Emissions scenarios. A special report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, NY, USA, 599 pp.

New, M., M. Hulme and P. Jones, 1999: Representing twentieth-century space-time climate variability. Part I: Development of a 1961-1990 mean monthly terrestrial climatology. *J. of Climate*, **12**, 829-856.

Ruosteenoja, K., T.R. Carter, K. Jylha and H. Tuomenvirta, 2002: Future climate in world regions: an intercomparison of model-based projections for the new IPCC emissions scenarios. *Working Paper*

Santer, B.D., T.M.L. Wigley, M.E. Schlesinger and J.F.B. Mitchell, 1990: Developing climate scenarios from equilibrium GCM results. Rep. No. 47, Max-Plank-Institut-fur-Meteorologie, Hamburg, 29 pp.

Santer, B.D.,W. Brueggemann, U. Cubasch, K. Hasselmann, H. Hoeck, E. Maier-Reimer and U. Mikolajewicz, 1994: Signal-to-noise analysis of time-dependent greenhouse warming experiments. *Climate Dynamics*, **9**, 267-285.

Sokolov, A.P. and P.H. Stone, 1998: A flexible climate model for use in integrated assessments. *Climate Dynamics*, **14**, 291-303.

Washington, W.M., J.W. Weatherly, G.A.Meehl, A.J.Semtner Jr., T.W. Bettge, A.P.Craig, W.G. Strand Jr., J. Arblaster, V. B. Wayland, R. James and Y. Zhang, 2000: Parallel climate model (PCM) control and transient simulations. *Climate Dynamics*, **16**, 755-774.