Tellus (2008), 60A, 398–410
Printed in Singapore. All rights reserved

TELLUS

Extended Big-Brother experiments: the role of lateral boundary data quality and size of integration domain in regional climate modelling

By MORTEN KØLTZOW[1*], TROND IVERSEN[2] AND JAN ERIK HAUGEN[1], [1]*Norwegian Meteorological Institute, P.O. Box 43, Blindern, 0313 Oslo, Norway,*
[2]*Department of Geosciences, University of Oslo, Norway*

(Manuscript received 30 April 2007; in final form 10 December 2007)

ABSTRACT

Dynamical downscaling by atmospheric Regional Climate Models (RCMs) forced with low-resolution data should produce fine scale climate details with skill. This is investigated by adopting and extending the Big-Brother approach of Denis et al. (2002). A reference climate is established from a fine resolution RCM simulation in a large domain (the Big-Brother). These Big-Brother (BB) data are degraded by removing small scales, and then used for downscaling by the RCM (the Little Brother) with the same resolution as the BB in three domains of different size. Differences between the Little- and BB are attributed to errors caused by the downscaling. We have furthermore extended the original BB method and investigated the impact of the quality of the driving data. The RCM manage to reproduce the general large scale climate features of the BB when forced with high quality data, but show deficiencies when the driving data differ both in phase and scale from the BB. Forced with data with lower quality on a sufficiently large integration domain and in regions influenced by strong local forcing, the RCM significantly improve the climate statistics for local variables (2 m air-temperature, 10 m wind speed, precipitation). We even found that the improvement increased with domain size.

1. Introduction

Global projections for anthropogenic climate change are provided by coupled global climate models (GCMs) and scenarios for constituent emissions to the atmosphere. Typical horizontal resolution in GCMs for IPCC (2001) was 250–300 km in the atmosphere and 100–150 km in the oceans (Lambert and Boer, 2001). In the 4th Assessment Report of IPCC (2007), resolution has been increased a factor 1.5–2 by some GCMs, but this is still insufficient for most studies of impacts of climate change. The low resolution also constitutes a source of errors in the regional aspects of climate simulation results. Downscaling global projection data by employing regional climate models (RCMs) with higher resolution has therefore become common procedure in the recent years (see e.g. IPCC 2001 and 2007). There still is some discussion related to the proper set-up of dynamical downscaling, and to what extent the downscaling augments the value of the global calculations for the region.

This paper revisits these challenges by employing a further developed type of the 'Big-Brother Experiments' (BBE) introduced by Denis et al. (2002). We only consider pure atmospheric downscaling which makes use of nesting (e.g. Jones et al., 1995; Christensen et al., 1997; Noguer et al., 1998). The GCM data are imposed at the open lateral boundaries of the RCM and by the sea surface temperature and sea ice cover. We do not consider the method of spectral nudging (Kida et al., 1991; von Storch et al., 2000).With this method the large scale part of the GCM data are forced in the interior of the RCM integration domain. Another method for more fine scale simulations are GCMs with variable resolution and focus over the area of interest (e.g. Fox-Rabinovitz et al., 2001).

The rational behind dynamical downscaling is that RCMs should produce realistic fine-scale details over a region when fed by information from coarser-resolution GCM data at the boundaries, and that the results are better suited for impact assessment studies. In RCMs there should be at least three sources of motion systems on scales not resolved by GCMs: (1) improved description of ground surface structures and contrasts, (2) more explicit description of nonlinear dependencies in the dynamics and (3) improved resolution and more explicit description of hydrodynamic instabilities (Denis et al., 2002). There are two quality aspects which need to be addressed in this respect: (1) the ability to reproduce observed historical and present-day climate conditions and (2) the ability of the RCM to improve, or add value,

* Corresponding author.
e-mail: morten.koltzow@met.no
DOI: 10.1111/j.1600-0870.2008.00309.x

to the climate statistics of the GCM which provides data for downscaling.

There are also a number of issues concerning the use of the nesting technique and their outcome for regional climate modelling (Giorgi and Mearns, 1991, 1999; Denis et al., 2002). One issue is to force the RCM with GCM data in a proper way. This concerns not only the mathematical formulation, but also the difference in spatial resolution between the RCM and the GCM, the frequency of boundary data updates, and errors introduced by temporal and spatial interpolation of the GCM data. A second issue is inconsistent physical parametrizations in the GCM and RCM models. This was addressed by, for example, Rummukainen et al., (2001) who found that the relative humidity in the vicinity of the lateral boundaries in a RCM is dependent on differences in the way precipitation processes is handled in the GCM and in the RCM. Thirdly, there are aspects concerning the choice of integration domain for the RCM. Jones et al., (1995 and 1997) emphasize that the RCM circulation should be constrained by the fields of the driving model for features resolved by the latter and advocated the use of small RCM integration domains in order to secure this whilst others have proposed spectral nudging (Kida et al 1991). This opinion has been supported by others (e.g. von Storch et al., 2000), but not generally (e.g. Machenhauer et al., 1994; Christensen et al., 1997). In the Norwegian research project Regional Climate Development Under Global Warming (RegClim, 1997–2006) attention was not paid to this issue (Bjørge et al., 2000; Haugen and Iversen, 2008). Less controversial is the requirement that fine-scale features developed in the RCM should not be damped or destroyed by the open lateral boundary conditions (Jones et al., 1995). A fourth aspect concerns the application of atmosphere only RCMs. Castro et al., (2005) showed that the ground surface boundary conditions are important in generating large scale variability in RCMs, and that this importance increases as the integration domain increases. Thus, large scale atmospheric variability tends to be increasingly underestimated as the domain size increases when the lower boundary conditions are not allowed to respond to the atmosphere. In atmospheric RCMs the continental ground surface variables are adjustable whilst the ocean surface is fully prescribed from the GCM. Only a few RCMs include coupling to the ocean (e.g. Döscher et al., 2002; Rinke et al., 2003).

One step in the validation of the quality of atmospheric downscaling with RCMs is the performing of Perfect Boundary Experiments (PBEs) (e.g. Christensen et al., 1997; Rinke et al., 1999). When GCM-data are downscaled with a RCM, systematic and random errors imposed at the lateral boundary and at the ocean surface are unavoidable due to the imperfections inherent in any GCM. In PBEs the RCM is driven with analysed atmospheric fields at the lateral boundaries and similar for the ocean surface. Hence, a PBE is a demanding test of the quality of the RCM with regard to downscaling, as the results can be compared directly with observed climate data over the specific time-period. However, as advocated by Denis et al., (2002), regular atmospheric climate observations often lack the spatial and temporal resolutions needed for adequate validation of the fine scale features calculated by the RCM. Fields from very-high resolution data assimilation are contaminated by inaccuracies of the assimilation method and errors in the model used for the purpose. Finally, PBEs does not differ between errors originating from the RCM itself and from the downscaling technique.

Denis et al., (2002) therefore suggested a new type of PBEs which are fully based on models. This is nick-named the BBE and enables evaluation of nesting strategies for dynamical downscaling. First a reference climate from a RCM simulation is established as a pure model product using a large integration domain and the high resolution intended for use when downscaling GCM data. This simulation is called the Big Brother (BB). In the method proposed by Denis et al., (2002) and used in later experiments (Denis et al., 2003; Antic et al., 2004; Herceg et al., 2006; Diaconescu et al., 2007) data from BB are then degraded towards the common resolution used in the atmospheric components of the global climate models (GCM) by simply removing, or filtering of, the smaller scales. The resulting filtered fields are then used as lateral boundary data to drive an RCM (called the Little Brother, LB) which is integrated using the same resolution as the BB, in a subarea of the BB domain. The climate statistics of the LB is validated by comparing with the unfiltered BB data in the LB domain. Differences between the two statistics can unambiguously be attributed to errors associated with the dynamical downscaling technique, and not to model errors or observational limitations.

Denis et al., (2002, 2003) introduced the concept of the BBE, and performed it for a one month long simulation. They found that time mean and variability of fine scale features are successfully reproduced over regions with strong small-scale surface forcing. Over ocean and away from the surface, less reproducibility is achieved. Antic et al., (2004) extended the time period to yield 4 winter months on a domain covering the west coast of North America. These results are in agreement with predictability aspects of forced versus free flows (Anthes et al., 1985; Boer, 1994).

These results were also interpreted to be in favour of the idea that the RCMs should not be able to change the driving GCM data on the scales resolved by the latter (e.g. Jones et al., 1995, 1997). We advocate that this interpretation can be a false effect of the fact that the driving data for the LB are not taken from a coarse resolution model with associated errors, but from data which by construction are perfect on the coarser scales. In a BBE experimental set-up the large scale patterns of the LB should stay close to the BB data. We therefore take the original BBE a step further by comparing the original experiments with results for which the BB data have errors typical for coarse resolution models.

In Section 2, the RCM HIRHAM used in the study is described, before a more detailed description of the experiment is given. Then the results are presented in Section 4, before a summary is given and conclusions are drawn in Section 5.

2. The regional climate model HIRHAM

Our version of the HIRHAM RCM was imported from Max Planck Institute for Meteorology (MPI) in 1997. A similar version was used at the Danish Climate Centre (Christensen et al., 1996, 1997, 1998). As described in Bjørge et al., (2000) the model has been used for dynamical downscaling of global climate scenarios on a rotated spherical grid with 0.5° (∼55 km) horizontal grid resolution and 19 levels in the vertical). The physical processes are adapted from the ECHAM4 ACGM (Roeckner et al., 1996).

A more recent version (Haugen and Haakenstad, 2006) of the HIRHAM RCM is used in this study. In this version, the dynamical core uses a two-time-level three-dimensional semi-Lagrangian semi-implicit time-integration scheme, adopted from the more recent version (no. 5) of the HIRLAM numerical weather prediction model (Undén et al., 2002.). The physical parametrization package of HIRHAM was modified to couple to this scheme. Furthermore, the lateral boundary relaxation scheme, the horizontal diffusion, and the surface albedo parametrization (Køltzow, 2007) was updated. The lateral boundary relaxation is based on Davies (1976) and is commonly in use in RCM's (e.g. used in 6 out of 8 RCM's in Rinke et al., 2006). The linear interpolation in time of the boundary data was replaced by a cubic interpolation. The changes were implemented in successive steps. In Haugen and Haakenstad (2006) the model output was validated by downscaling re-analyzed ERA40 data and compared to monthly CRU TS 2.1 gridded observations. Temperature and precipitation errors were slightly reduced for Europe in general, and the precipitation errors in particular in some mountainous areas were much reduced. In addition, the computational efficiency was improved.

3. The extension of the Big-Brother experiments

In this study the perfect prognosis approach called the poor man's Big-Brother Experiment (hereafter referred to as BB) is followed (Denis et al., 2002). Two BB simulations are done; the first uses of 0.5° × 0.5° horizontal grid and 31 levels on the big domain shown in Fig. 1. This simulation is referred to as the fine-resolution BB and is considered to represent the truth in the experiments. Small scales features in this fine-resolution BB are removed by interpolating to a 2.8° × 2.8° grid with 19 levels by a bi-linear interpolation in the horizontal and by interpolating from 31 to 19 vertical levels. As in Denis et al., (2002) this degraded data-set, hereafter referred to as *the filtered fine resolution BB data*, can be dynamically downscaled in a smaller domain, thus investigating to what extent the downscaling can regenerate the filtered small scale structures.

Compared to Denis et al., (2002) we extend the BB technique with a second BB simulation in order to more realistically mimic the actual situation where the coarse-resolution data have to be taken from a coarse-resolution model. To generate large-scale data, the model is run with coarser resolution (2.8° × 2.8° and 19 levels) directly in the same big domain, which is similar to the approach used by Diaconescu et al., (2007) independent of our present study. We refer to this simulation as the coarse-resolution BB.

The filtered fine resolution data must be regarded as representing the upper bound in the potential quality of coarse resolution data, since they contain resolution errors without phase shifts. Differences between results obtained from downscaling filtered fine resolution BB data and those from downscaling coarse resolution BB, represent a potential for improvement in dynamical downscaling.

Compared to output from GCM's, the BB data sets simulations are produced in a limited area with the same model tool as used for downscaling, thereby referred to as poor man's BB

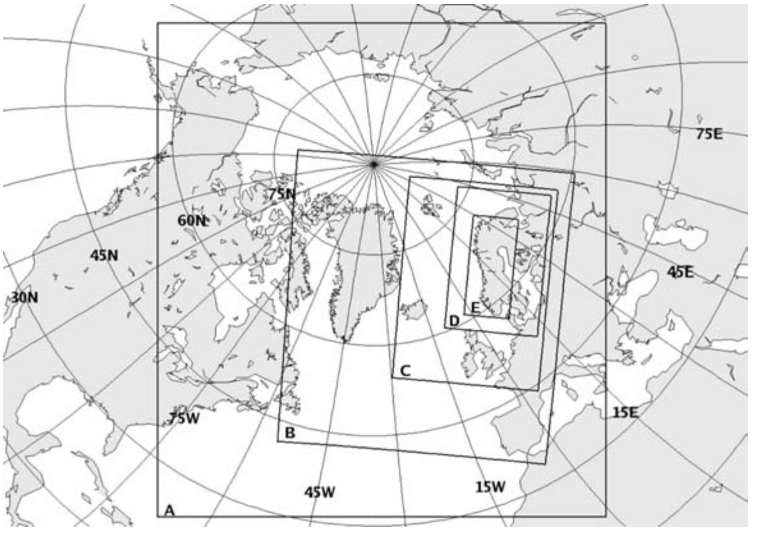

Fig. 1. The different integration domains used in the study. The large domain (marked A) is used in the fine- and coarse-resolution Big-Brother (BB) simulations. The smaller embedded areas are the large (marked B), medium (marked C) and small (marked D) Little Brother (LB) model domains. The common verification domain which is used in Figures 2–5 are the innermost area marked E.

experiments. Both BB runs are forced with the ECMWF ERA40 re-analysis data set at the lateral and surface boundary (Uppala et al., 2005). All simulations are for the period 1970–1990. The first year is considered as spin-up and not included in the analysis that thus contains 20 yr.

The dynamical downscaling of BB data uses the same model as used to generate the BB data, and is termed Little Brother simulations (hereafter referred to as LB). In this paper we apply three alternative LB runs based on three different integration domain sizes; small, medium and large. For each of the three LBs, data from both the filtered BB and the coarse-resolution BB data sets are downscaled. In the LB runs, data at the lateral boundaries are updated every 6 h, as supported by others (e.g. Dimitrijevic and Laprise, 2005; Antic et al., 2004). All the three integration domains are given in Fig. 1. Results from the different domains are intended to highlight the importance of the domain size.

It is expected that the filtered BB data set deviates from the coarse-resolution BB with respect to circulation patterns. Comparison between downscaled versions of the two data-sets is therefore expected to give valuable insight into the abilities of the downscaling to improve the climatology of the coarse-resolution BB. The original experiment by Denis et al., (2002) was not properly designed to address this problem, since the filtered data were perfect except for the finest scales. To generalize the BB experiments one step further towards practical realism, the downscaling should be run with several other models than the one used for producing the BB data. This is not attempted in this study.

4. Results

We focus our attention on the potential qualities of dynamical downscaling for Norway, even though some aspects of downscaling elsewhere are discussed. The 20-yr climatic averages of the fine-resolution BB, that is, our version of the 'truth', is compared with those of the three different LB's inside a common verification area for all model domains (see Fig. 1). Similarly, the quality of daily values of mean sea level pressure (MSLP), 2 m air temperature (T2m), precipitation (PREC), and wind speed at 10 m heights (V10) is discussed on the basis of bias error, root mean square error, and temporal correlations relative to the fine-resolution BB for all model grid points in Norway. Finally, the day-by-day variability is studied, including the reproduction of extreme values.

4.1. Reproduction of climate averages

Figure 2 show the temporal averages of MSLP for the years 1971–1990 in the fine- and coarse-resolution BB simulation, along with deviations of each LB version from the fine-resolution BB. The fine-resolution BB has low MSLP over the ocean west of Norway and higher over land with a gradient condensed along the western coastline. In the coarse-resolution BB the zone of strong MSLP gradient does not follow the coast to the same extent, and the extremes and gradients are weaker. All LB versions yield too high MSLP. When filtered fine resolution BB data are downscaled, the LB on small and medium domains give a 1–1.5 hPa bias with little geographical variation, whilst in the large domain the bias varies from 1.5 hPa in the north to 3.5 hPa in the south. With coarse-resolution BB for downscaling the bias is between 1 and 3 hPa with largest values to the west and south. The quality of the data for downscaling is shown to be particularly important, when the LB is applied to downscale filtered fine resolution BB data in the two smallest domains. For the largest domain, however, the MSLP bias is slightly worse when

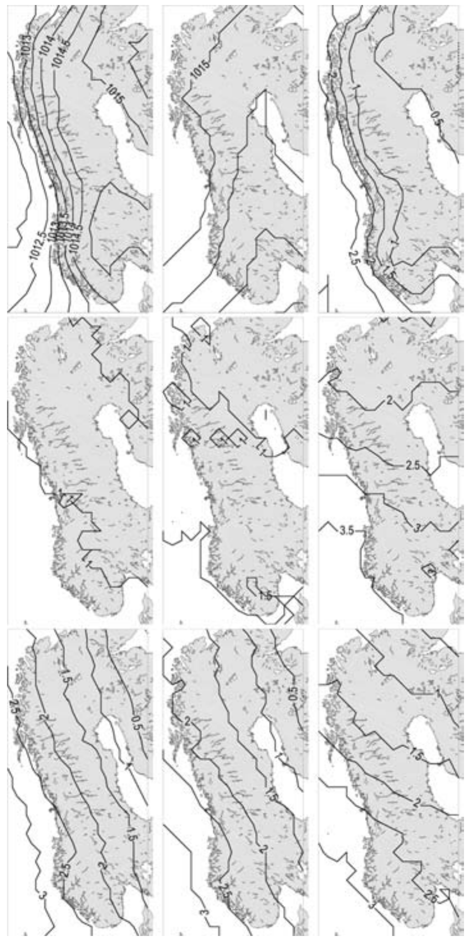

Fig. 2. Results for mean sea level pressure, with equidistance 0.5 hPa, averaged over the years 1971–1990. Upper left-hand side: fine-resolution BB; upper middle: coarse-resolution BB and upper right-hand side: the difference between the coarse- and fine-resolution BB. The middle row shows the downscaling errors (the difference between LB results and the fine-resolution BB) for the LB simulations driven by the filtered BB over the small (left-hand side), medium (middle) and large (right-hand side) LB integration domain. The bottom row shows the corresponding errors of the three LB simulations when driven by the coarse-resolution BB.

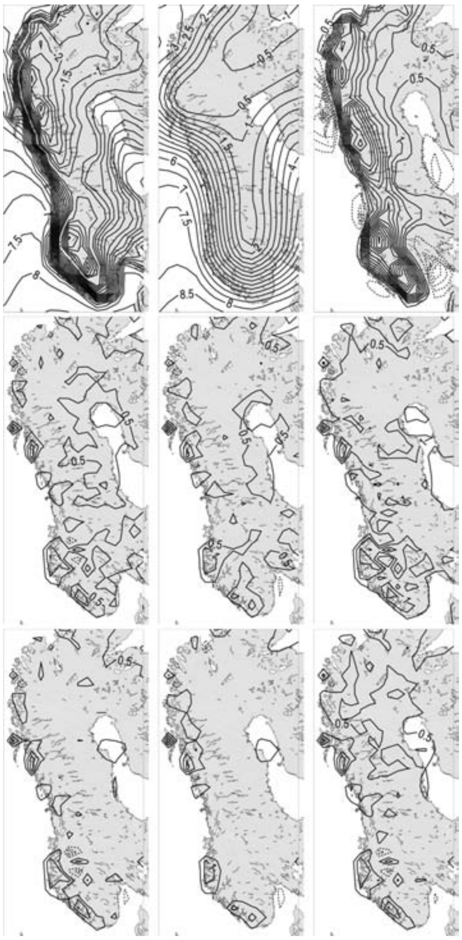

Fig. 3. Same as Fig. 2, but for air temperature 2 m above the ground surface with equidistance 0.5 °C.

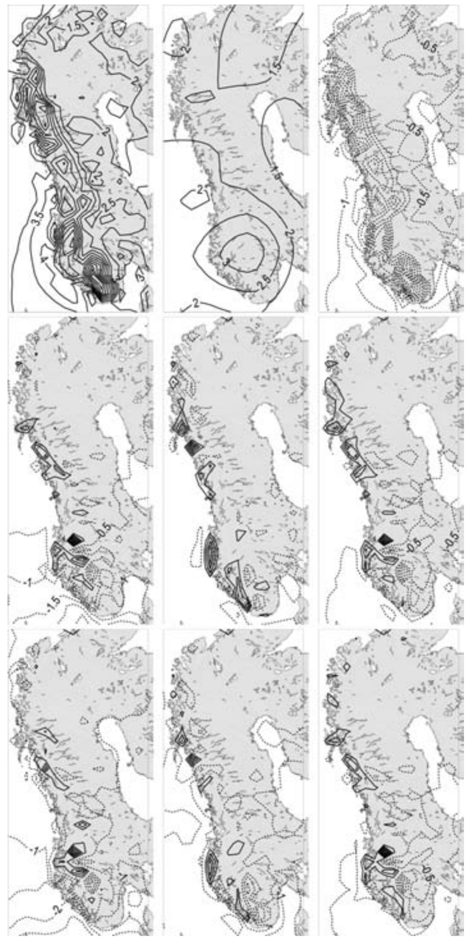

Fig. 4. Same as Fig. 2, but for 24 h accumulated precipitation with equidistance 0.5 mm $(24\ h)^{-1}$.

downscaling filtered data, indicating that LB in that case is able to develop more of its own model-specific features in the interior. Comparing with the MSLP bias of the coarse-resolution BB, however, there is otherwise little evidence of improved MSLP-patterns in the results from downscaling of the coarse-resolution BB data. One exception is seen over the ocean North West of Norway (mostly outside the area covered by Figure 2), which will be discussed below under the comparison of areas outside Norway.

Figure 3 shows 20-yr climate averages for 2-m temperature (T2M), which is strongly influenced locally by the different resolution of the topography and other ground-surface features. Hence, over the ocean the coarse and the fine scale BB are similar since prescribed sea surface temperatures in both cases are taken from the ERA40 analysis. The Scandinavian mountains are too smooth and too low in the coarse resolution BB, which therefore is too warm in these areas. (The T2M from the coarse resolution BB is not height corrected by post-processing.) However, all LB computations show little bias error. In fact, the biggest deviations are found in the large-domain LB when filtered fine resolution BB data is downscaled. This case is ~0.5° warmer than the other LB results. Some regions at the west coast of Norway have a 0.5° too warm climate in all LB simulations, originating from situations in autumn and winter together with an increased on-shore wind component compared to the fine scale BB (not shown).

Figure 4 show the 20-yr average climate statistics for precipitation. Also precipitation patterns are highly related to the topography, resulting in strong features over the westward facing coastlines in the fine resolution BB which are absent in the coarse resolution BB. All LB simulations show considerable agreement with the fine scale BB and there are almost no signals that can be related to the quality of the two data sets for downscaling. This is in agreement with, for example, Antic et al., (2004) who stated that complex topography enhances the downscaling ability of precipitation. Nevertheless, an important deviation is that precipitation is underestimated in LB with the small domain. This is a shortcoming of the closeness to the influx lateral boundary. Even though specific humidity and cloud water is included in the boundary forcing, the HIRHAM RCM needs a downwind distance time to establish dynamical balance between fields like vertical motions and moisture. For the small domain there is insufficient (lagrangian) time to establish this balance in air downstream of the lateral boundary before it hits the mountains in our area of interest. Also, differences in land-sea mask between the coarse and fine grids close to the lateral boundary of the small domain, are prone to influence the evaporation from the sea surface. However, we have not quantified this error. Furthermore, also the precipitation in the large-domain LB is underestimated in some regions, leaving the LB on medium-sized domain with smallest precipitation bias. Downscaling the filtered fine resolution data yields the smallest bias, but the difference from downscaling the coarse resolution BB is only tiny. This indicates that the medium-sized domain is closest to the optimal trade-off between the requirement of necessary moisture spin-up and a proper control of the large scale flow pattern provided by the driving data. Too much precipitation at the West Coast of Norway is produced in the same areas where the T2M in the LB's show positive bias. Again these deviations are related to an increased on-shore wind component during autumn and winter (not shown).

Figure 5 shows the average climatology for wind speed at 10 m above ground surface. Also this parameter is strongly influenced by ground surface features such as topography and coast-lines. In the fine resolution BB, the pronounced average wind maxima occur over the oceans along the western coasts of Norway, which the coarse resolution BB fails to reproduce. This bias is to some extent corrected when downscaling the coarse resolution BB with LB applied in the two smallest domains. With the large domain, the LB results are less restricted by the lateral boundaries, and the biases are to a large extent corrected. Downscaling filtered fine-resolution BB data yields small biases independent of integration domain, but the improvement above downscaling coarse-resolution BB data in the largest domain is small.

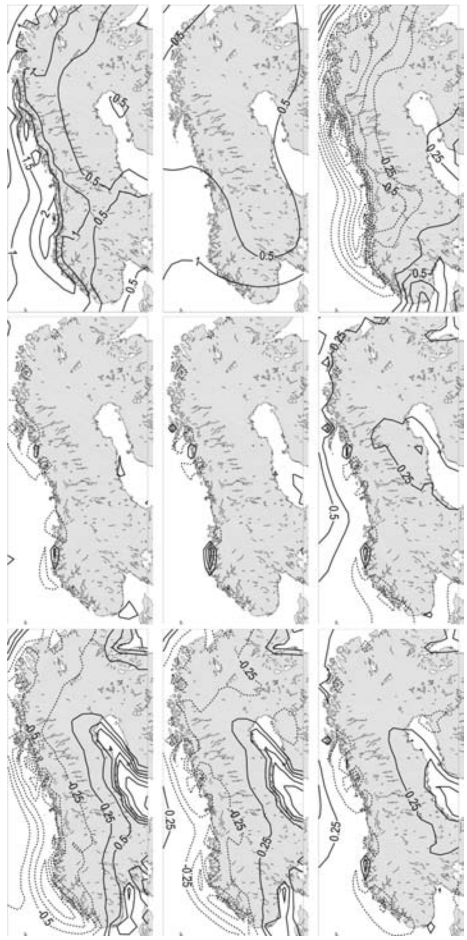

Fig. 5. Same as Fig. 2, but for wind speed at 10 m height with equidistance 0.5 m s^{-1}.

4.2. Daily statistics for Norway

In this section daily values are presented by monthly statistics in order to emphasize seasonal variations between BB and LB computations. In Fig. 6 20 yr statistics for MSLP error, that is, differences between fine resolution BB and all other model calculations, are given for all land grid points in Norway. The error measures are the bias, the error in variance of daily values as a measure of the day-to-day variability, the standard deviation (*SD*) of the error, are all shown along with the correlation with the fine resolution BB. The diagrams show that best verification results for MSLP are obtained for all statistical parameters when LB is applied with a small integration domain and filtered fine resolution BB data are downscaled. The annual cycles in statistics for errors in the coarse resolution BB are also present in the LB results obtained when coarse resolution BB data are downscaled, irrespective of domain size.

Errors are large in autumn and winter. In these seasons the reproduction of the average climatology of T2M and precipitation along the Norwegian West Coast was problematic (not shown), and the extra-tropical system of cyclones shows its highest activity. Stronger winds at the lateral boundaries increase the downwind distance that the RCM needs to develop smaller scaled features. At the same time, strong winds are associated with large gradients in pressure and temperature, and therefore an increased risk of phase errors in the simulations of synoptic and mesoscale features such as cyclones and fronts. For all measures the importance of the quality of the data used for downscaling is demonstrated. For MSLP, it is only for day to day variability, and to some extent for the correlation, obtained after downscaling with LB in a large domain, that it is possible to detect an added value of dynamical downscaling from the coarse resolution BB.

Figure 7 shows similar statistics for daily data for temperature 2 m above the ground. Downscaling decreases the temperature errors compared to those in the coarse resolution data throughout the year. This is mainly because the Scandinavian topography is considerably improved with better horizontal resolution. Annually averaged, the T2M is in good agreement with the fine scale BB, but this hides the fact there are underestimates in

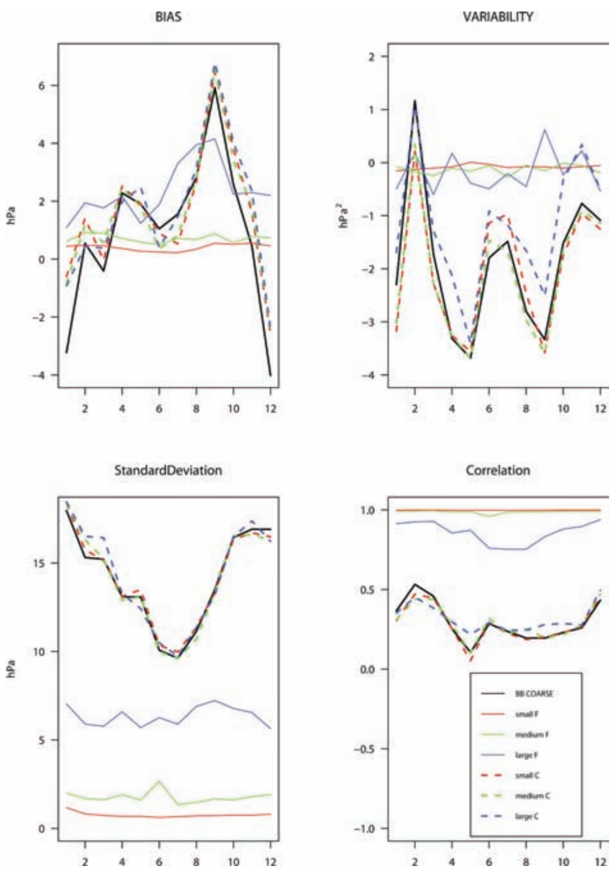

Fig. 6. Monthly mean error statistics, relative to fine resolution BB, of daily mean sea level pressure (hPa) over 20 yr and all land-based grid points in Norway. Shown are bias error (upper left-hand side), day-to-day variability (upper right-hand side), standard deviation (lower left-hand side), and correlation (lower right-hand side) for the six LB simulations and the fine resolution BB.

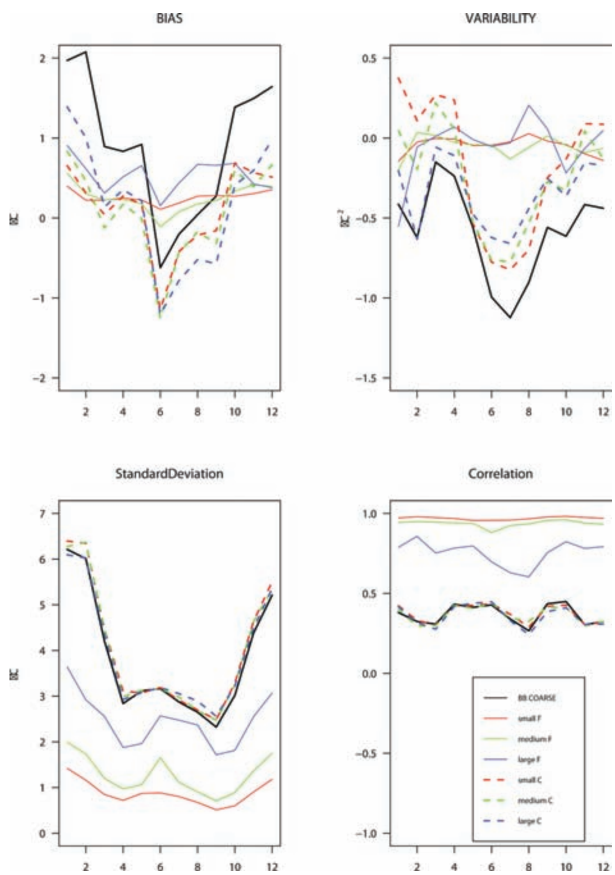

Fig. 7. Same statistics as in Fig. 6, but for 2 m air temperature T2M.

summer and autumn and overestimates during winter. The errors are considerably more evenly distributed through the year when the LB's downscale filtered fine resolution BB data. It is noticeable that during summer the smallest bias is calculated with LB in the medium-sized domain, whilst in winter the smallest bias is found with the smallest LB domain. This reflects that local forcing is relatively more important during the summer when information from the lateral boundary is slowly penetrating the integration domain. For *SD* and correlation no added value of dynamical downscaling of the coarse-resolution BB is found. For the day to day variability, however, a clear improvement in the downscaled data is apparent.

The monthly error statistics for daily precipitation is shown in Fig. 8. Again the best results are obtained when filtered fine resolution BB data are downscaled, but the smallest bias error is obtained when LB is applied with the medium sized domain. This domain is sufficiently small such that the synoptic circulation does not depart significantly from the data imposed at the boundaries (Jones et al., 1995), but is large enough to allow moisture to adjust downstream from the boundaries over the region of interest.

Even more important, however, is the added value in precipitation data obtained by downscaling the coarse resolution BB data. This is evident for bias error and the error of the day to day variability, and the best results are obtained with the largest LB domain. This contradicts the argument that smaller domains should be used to ensure that the LB stay close to the large scale flow in the coarse resolution model. For *SD* and correlation, downscaling of the coarse resolution BB worsen slightly or make no impact. The seasonal differences show less correlation and higher *SD* during summer and autumn when precipitation is strongly influenced by convection with a more stochastic nature. This confirms the findings of Dimitrijevic and Laprise (2005) based on simulations of 5 July months.

Figures 7 and 8 show that surface temperature and precipitation from downscaled scenarios are not crucially dependent on the quality of the data used for downscaling. This agrees with Noguer et al., (1998) who stated that parts of the distribution of the mentioned variables are controlled by internal processes in a RCM. However, seasonal deficiencies in the data used for

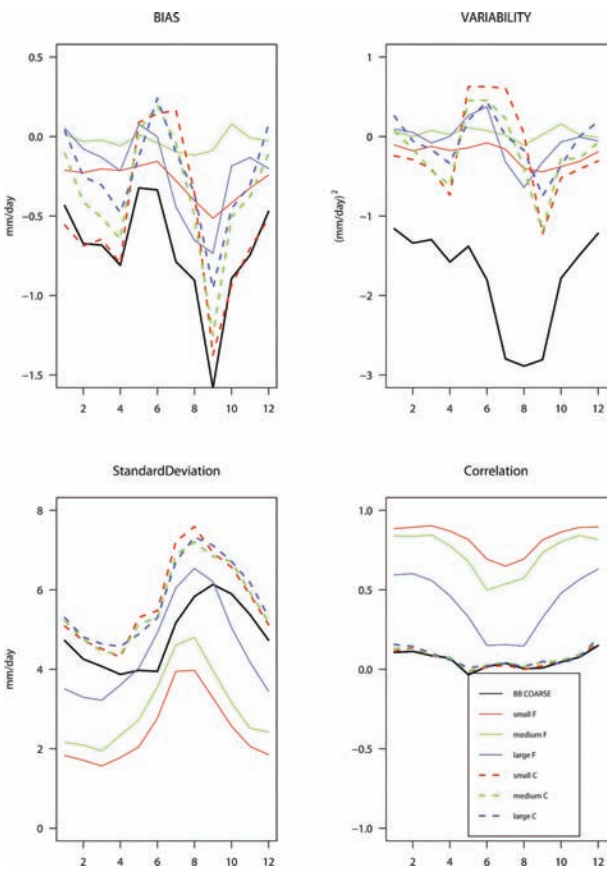

Fig. 8. Same statistics as in Fig. 6, but for valid for 24 h accumulated precipitation.

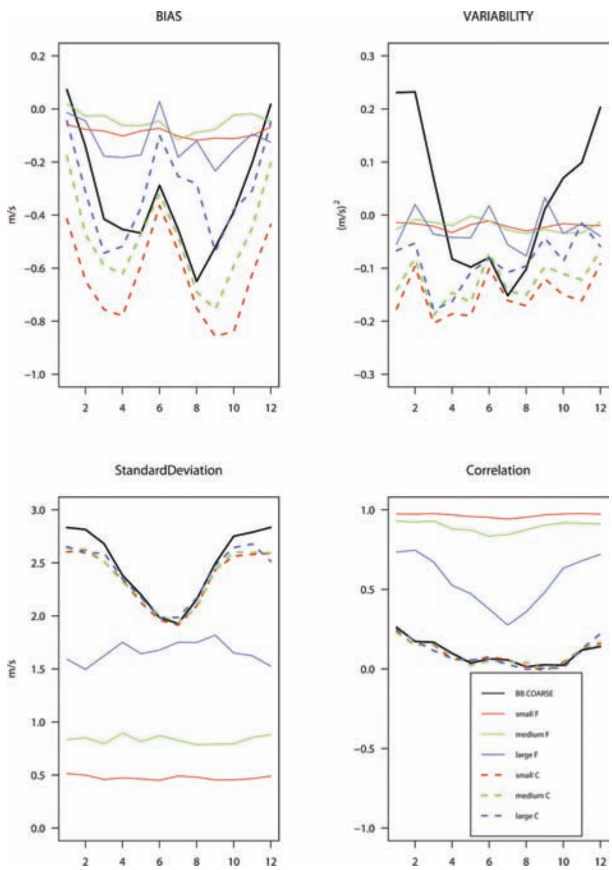

Fig. 9. Same statistics as in Fig. 6, but for 10 m wind speed.

downscaling are to some degree also seen in the downscaled data from the LB.

Figure 9 is similar to the three previous figures, but is valid for wind speed 10 m above the ground. The main features and differences in wind are seen over the ocean immediately west of Norway (see Fig. 5). The majority of these grid points are not included in the statistics of Fig. 9. Again the quality of the data used for downscaling is important. Downscaling of the coarse resolution BB adds value in summer, but only for the largest LB domain for bias error, and in winter for day to day variability for all LB domains.

From a deterministic point of view dynamical downscaling is not very efficient, since the *SD* is quite large for the different LB simulations. This was also one of the findings of de Elía et al., (2002). However, the method of dynamical downscaling shows skill in reproducing the climatology and the day to day variability. This is especially seen when the coarse resolution forcing is of high quality. However, also when the driving data deviates from the correct large scale circulation, dynamical downscaling is able to reproduce data in agreement with the fine resolution BB.

The analysis done in this section has also been performed in regions of Norway (Northern Norway, the West coast and the South East in-land). No differences in the ability of dynamical downscaling are found between the regions (not shown). This do not support earlier findings (Dimitrijevic and Laprise, 2005; Denis et al., 2003) which found that the skill in reproducing precipitation over the West Coast of North America is higher than the skill over the East Coast of North America, a region with weaker topographic forcing. This can be explained by the fact that in contrast to the wide North American continent, the southeast part of Norway is strongly influenced by improved topography in the fine resolution grid. The region with best agreement between the fine and the coarse scale BB varies. For MSLP and precipitation the southeast part of Norway is in best agreement while for wind speed, Northern Norway obtains the best agreement. For two meter temperature, only minor differences between regions are found.

4.3. Extreme values for Norway

From the 20-yr-long simulations we extract the 20 highest (MSLP, T2M, precipitation and wind) and lowest (MSLP and T2M) daily values from each model simulation. Spatial mean values and Mean Absolute Error (MAE), *SD* and Correlation of the extremes compared to the fine resolution BB are listed in

Table 1. Each cell in this table gives the average over Norway of the 20 highest and lowest extremes for MSLP, T2M, precipitation and wind speed at 10 m height in each grid square. Also the average over Norway of the mean absolute error, the standard deviation of the errors, and the Correlation with the fine resolution Big-Brother are given in each cell of the table.

	Fine resol. BB	Coarse resol. BB	Small LB forced with BBF	Medium LB forced with BBF	Large LB forced with BBF	Small LB forced with BBC	Medium LB forced with BBC	Large LB forced with BBC	
Max MSLP	1054.52	1049.78	1054.76	1055.05	1054.47	1049.30	1049.74	1052.16	Mean
		4.74	0.26	0.53	0.28	5.22	4.78	2.36	MAE
		1.98	0.19	0.27	0.35	1.40	1.28	1.04	*SD*
		0.33	0.99	0.99	0.97	0.71	0.79	0.85	Corr
Max T2m	293.54	289.81	293.52	293.33	293.71	290.23	290.58	291.08	Mean
		3.90	0.71	0.77	0.73	3.36	3.00	2.52	MAE
		3.04	1.33	1.30	1.28	2.45	2.09	1.86	*SD*
		0.73	0.96	0.96	0.96	0.84	0.91	0.94	Corr
Max Precip	33.89	19.05	33.06	34.34	34.19	37.08	35.60	35.28	Mean
		14.84	3.91	3.04	3.32	6.33	4.43	4.14	MAE
		8.36	5.56	3.99	4.50	8.57	5.32	5.89	*SD*
		0.42	0.83	0.91	0.89	0.71	0.84	0.83	Corr
Max wind	11.30	11.62	11.13	11.15	11.17	10.36	10.63	10.99	Mean
		2.34	0.64	0.64	0.65	1.30	1.08	0.78	MAE
		2.71	1.13	1.17	1.17	1.43	1.31	1.21	*SD*
		0.88	0.98	0.98	0.98	0.97	0.97	0.97	Corr
Min MSLP	970.72	970.17	971.39	971.78	972.79	973.07	973.74	972.06	Mean
		1.17	0.67	1.08	2.06	2.36	3.01	1.52	MAE
		1.32	0.25	0.59	0.87	1.04	1.22	1.19	*SD*
		0.85	0.99	0.90	0.88	0.87	0.86	0.89	Corr
Min T2m	252.00	254.92	252.37	253.56	252.22	253.60	252.02	253.26	Mean
		3.39	0.97	0.99	1.52	1.43	1.29	1.48	MAE
		4.10	1.54	1.30	1.58	1.84	1.32	1.57	*SD*
		0.94	0.99	0.99	0.99	0.99	0.99	0.99	Corr

Table 1. In each grid point in Norway, the average over the 20 most extreme values are calculated. The numbers given in Table 1 are averages of all the thus obtained grid point specific extremes in Norway. Table 1 also gives the MAE, the *SD* of the error, and the correlation with the fine-resolution BB data ('the truth'). The coarse resolution BB clearly underestimates the highest pressures, while the LB simulations stay close to their driving data. However, when the downscaling domain is sufficiently large, an improvement for events with high MSLP appears. For low surface pressures we do not see the same improvement.

The LB results which are produced by downscaling the filtered fine resolution BB data, reproduce the maximum temperatures well. The coarse resolution BB underestimates the highest temperatures considerably, and this also influences the results from the LB downscaling. However, the underestimation decreases when the LB domain increase in size. For the minimum T2M the coarse resolution BB overestimates, but for the LB downscaling the results' quality depend less critically on the data used for downscaling. Extremely low temperatures are often determined as a result of weak turbulent fluxes in the stable planetary boundary layer which is strongly influenced by long-wave radiative cooling during nights with anticyclonic subsidence and clear skies. Also extremely high temperatures are generated locally during anticyclonic subsidence and a stable boundary layer. The local component is further enhanced by the fact that evaporation from the ground is an important moderator (Schär and Jendritzky, 2004). However, whilst radiative cooling during night feeds positively back on the stabilization of the boundary layer, the strong solar radiation during heat-waves tends to destabilize the column. Extremely high temperatures therefore depend on dynamical processes which suppress destabilization, such as strong anticyclonic subsidence coupled with horizontal advection of deep tropospheric layers of potentially warm air. High temperatures are therefore more influenced by the lateral boundary conditions than low temperatures.

The maximum precipitation is considerably underestimated in the coarse resolution BB, but any downscaling with LB significantly improve this situation, even though there is a tendency of to overestimate when coarse resolution BB data are downscaled. Furthermore, relative high values of MAE and *SD* suggest that the localization is probably not the same in the downscaled data as in the fine scale BB. The LB results obtained when using the medium or large domains are considerably better than those obtained when the small domain is used.

Little difference between the experiments is found for extreme winds, although surprisingly the strongest winds are produced with the coarse resolution BB. However, compared with the fine resolution BB the MAE, *SD* and correlation show considerable improvements by downscaling with the LB. The statistical parameters indicate that the coarse resolution BB compensate artificially for higher/lower wind speeds compared to the fine resolution BB. The high correlations for the LB which downscales the coarse resolution BB data, documents an increased ability to localize wind extremes. Also in this case, the best results are obtained with the largest LB domain.

4.4. *Climate of the large LB domain*

Figure 10a show the differences in climatological average of MSLP between the coarse- and the fine resolution BB. The highest resolution gives a deeper low pressure system in the North Atlantic Ocean, the Barents Sea, and parts of the westward facing European regions. Pressure is otherwise higher over major continental areas in Eurasia, Canada and Greenland. The difference over Greenland is probably a consequence of higher Greenland topography in the fine resolution version influencing the post-processing of MSLP.

The difference in climatology between the coarse resolution BB and its downscaled results with the large domain LB, is given in Fig. 10b. The patterns are to a large extent equal, but opposite in signs to those of Fig. 10a. Thus, downscaling partly repairs the MSLP errors over Greenland. Over the North Atlantic Ocean, however, the downscaling slightly increases the error in the MSLP present the coarse resolution BB, except over the Barents Sea where downscaling reduces MSLP and thus repairs the error in the coarse resolution BB. Over continental Eurasia (Finland/Russia) the LB increases the MSLP towards the fine resolution BB. It has earlier been stated that RCMs are not able to reduce circulation errors in coarse resolution driving

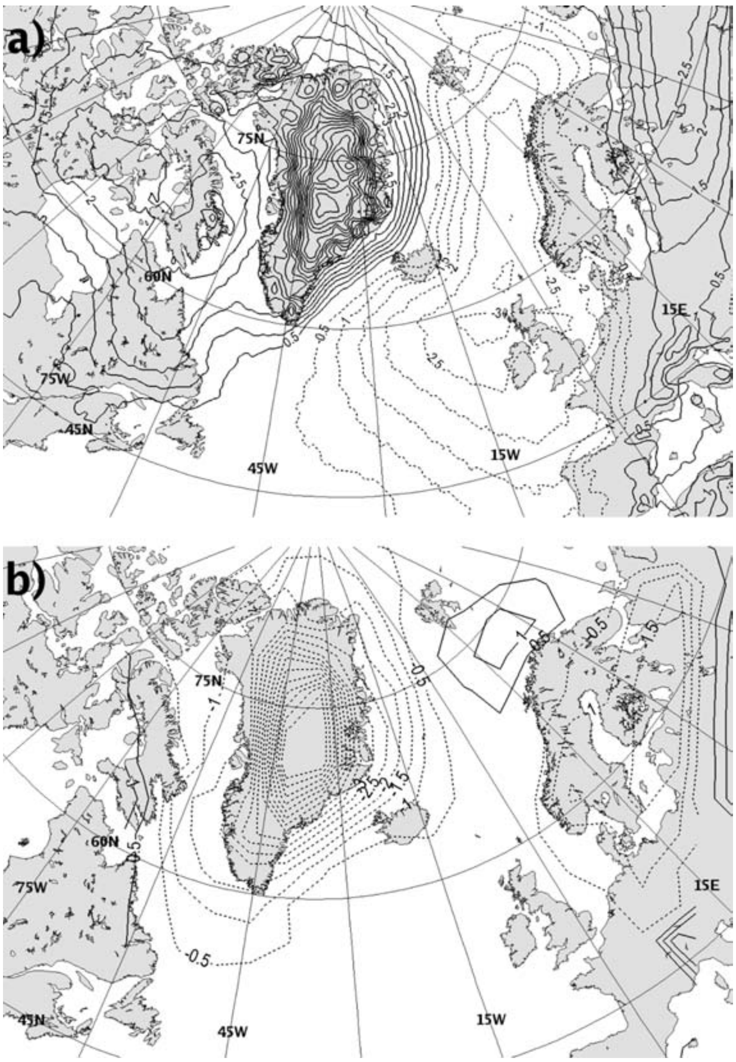

Fig. 10. (a) Fine scale Big-Brother MSLP (average 1971–1990) minus coarse resolution Big-Brother MSLP (1971–1990). Equidistance is 0.5 (b) Coarse resolution Big-Brother MSLP (1971–1990) minus MSLP (1971–1990) from downscaled coarse resolution Big-Brother with LB in the largest domain.

data (Noguer et al., 1998). However, this study indicates that at least in some regions and with a sufficiently large integration domain, it is possible to improve the climate of the large scale circulation.

Except for differences close to the lateral boundaries, the only major differences found in T2m, precipitation and wind speed are those already discussed for the Scandinavian area. Elsewhere they are remarkable similar (not shown).

5. Discussion and conclusions

In this study abilities and limitations of dynamical downscaling of global climate model data have been investigated by employing an extended version of the BB approach introduced by Denis et al., (2002). Compared to earlier studies, the simulation length (20 yr), the size of the integration domain, and the geographical scope are new. Furthermore, the investigation of the impact of the quality of the data being downscaled is an extension compared to the original method of Denis et al., (2002), and follows the same approach as Diaconescu et al., (2007).

Eight simulations were performed in total. Two 20-yr BB simulations are made, one with 0.5° horizontal resolution and 31 levels and another with 2.8° horizontal resolution and 19 levels. The first of these were established to mimic the truth, to which all the LB simulations (on a small, medium and large domain) were compared. The LBs were forced with both filtered data from the fine-resolution BB run ('the truth') and with data from the coarse BB simulation. The second constituted the data for dynamical downscaling. This is a step towards reality compared to the procedure of downscaling fine resolution BB data after removal of the fine-scale structures by filtering, which represent the highest possible quality of coarse resolution data.

When the LB is used to downscale high quality data they manage to reproduce the climatology and variability of the fine resolution BB very well. However, when the integration domain is large the LB circulation pattern may deviate from the fine resolution BB. This is especially seen in the MSLP fields. In a small domain the low-level circulations (i.e. the MSLP-pattern) is to a large extent steered by the lateral forcing and agree well with the fine resolution BB, but in this case the LB fails to regenerate some important features of precipitation because the closeness to the lateral boundary hampers upstream moisture adjustment. This illustrates that the optimal domain should be small enough to constrain the large scale circulation, but sufficiently large to allow small scale features to develop (Jones et al., 1995). On the other hand, this is not necessarily true when the quality of forcing data at the lateral boundaries is inferior. The regeneration of extreme values shows skill both in values and in localization.

The basic assumption behind dynamical downscaling is that RCMs should produce realistic fine scale features just by being fed by large scale and low resolution information at the boundaries. This study supports this assumption, although we find this positive impact in a region strongly dominated by local forcing by local topography and complex coastlines. This confirms suggestions by Anthes et al (1985) and Boer (1994), and does not contradict the results of de Elía and Laprise (2003) since they focused on unforced variability.

When the LB is used to downscale coarse resolution BB data which have lower quality than the filtered fine resolution data, it has problems in regenerating the correct large scale patterns, such as those characterized by the MSLP. This is in agreement with the findings of Diaconescu et al., (2007) with a rather small integration domain (4500 km $\times$ 4500 km) for the LB. However, our results indicate that with a sufficiently large domain, the LB is able to correct some of the deficiencies. For more local weather-related variables such as T2M, precipitation, and 10 m wind, skillful climate information is generated by the LB. We find that a large integration domain enhances the improvements, both in climate averages and for the day to day variations. Extremes in weather are potentially important when assessing consequences of climate change, and they are often related to features which are not properly resolved in GCMs. An important issue is therefore to what extent dynamical downscaling can reproduce extreme events when they are not explicitly present in the data applied for downscaling. Also in this respect we find that the LB has ability to regenerate extremes with quality which increases with increasing size of the integration domain.

In conclusion, we find that there is little doubt that downscaling may add value to coarse resolution data in areas dominated by local quasi-stationary forcing. We have also seen that, for a high-quality RCM, when the downscaling configuration allows improving the quality of the coarse resolution data, even better results may be obtained.

This conclusion should be used with some caution. For example, real dynamical downscaling of data from global climate scenarios represents even one step further in BB experimentation compared to our work. This step concerns that the coarse resolution GCM data in reality are produced with a considerably different model than the LB. Some aspects of our conclusions may depend on the simplification that the coarse resolution data are produced with the same model system as the LB except for resolution and integration domain. Deficiencies revealed in BBE may interact with other sources of error in a realistic downscaling set-up. Such aspects are recently discussed by Jacob et al., (2007), but without formally addressing the domain size.

Another caution relates to the fact that our 'truth' actually is a model product of the same type as used to produce the calculated data that are validated. Hence, the missing source of climate variability clearly present in a pure atmospheric RCM, is also missing both in our version of the 'truth' and in the coarse resolution BB data. Our experiments are therefore not designed to properly account for insufficient climate variability related to interactions with, for example, the oceans (Castro et al., 2005).

6. Acknowledgment

The work reported here is funded in part by the Norwegian Research Council, through grant no. 120656/720 to the Norwegian Meteorological Institute/RegClim. Computational costs are covered by a grant from the Research Council's Programme for Supercomputing.

References

Anthes, R. A., Kuo, Y. H., Baumhefner, D. P., Errico, R. M. and Bettge, T. W. 1985. Predictability of mesoscale motions. *Advances in Geophysics* Volume 28B, Academic Press, New York, USA 159–202.

Antic, S., Laprise, R., Denis, B. and de Elia, R. 2004. Testing the downscaling ability of a one-way nested regional climate model in regions of complex topography. *Clim. Dyn.* **23**, 473–493.

Bjørge, D., Haugen, J. E. and Nordeng, T. E. 2000. *Future climate in Norway*. DNMI Research Rep. No. 103, 41 pp. Norwegian Meteorological Institute, P. O. Box 43 Blindern, N-0313 Oslo, Norway, ISSN 0332-9879.

Boer, G. J. 1994. Predictability regimes in atmospheric flow. *Mon. Wea. Rev.* **122**, 2285–2295.

Castro, C. L., Pielke Sr, R. A. and Leoncini, G. 2005. Dynamical downscaling: assessment of value retained and added using the Regional Atmospheric Modeling System (RAMS). *J. Geophys. Res.* **110**, D05108, doi:10.1029/2004JD004721.

Christensen, J. H., Christensen, O. B., Lopez, P., van Meijgaard, E. and Botzet, M. 1996. *The HIRHAM4, Regional Atmospheric Climate Model*, Danish Meteorological Institute–Scientific Report–Copenhagen, Denmark.

Christensen, J. H., Machenhauer, B., Jones, R. G., Schar, C., Ruti, P. M., and co-authors. 1997. Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions. *Clim. Dyn.* **13**, 489–506.

Christensen, O. B., Christensen, J. H., Machenhauer, B. and Botzet, M. 1998. Very high-resolution climate simulations over Scandinavia–present climate. *J. Climate* **11**, 3204–3229.

de Elía, R. and Laprise, R. 2003. Distribution-oriented verification of limited-area models in a perfect-model framework. *Mon. Wea. Rev.* 131, 2492–2509.

de Elía, R., Laprise, R. and Denis, B. 2002. Forecasting skill limits of nested, limitied-area models: a perfect-model approach. *Mon. Wea. Rev.* **130**, 2006–2023.

Davies, H. C. 1976. A lateral boundary formulation for multi-level prediction models. *Q. J. R. Meteorol. Soc.* **102**, 405-418.

Denis, B., Laprise, R., Caya, D. and Cote, J. 2002. Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment. *Clim. Dyn.* **18**, 627–646.

Denis, B., Laprise, R. and Caya, D. 2003. Sensitivity of a regional Climate model to the spatial resolution and temporal updating frequency of lateral boundary conditions. *Clim. Dyn.* **20**, 107–126.

Dimitrijevic, M. and Laprise, R. 2005. Validation of the nesting technique in a regional climate model and sensitivity tests to the resolution of lateral boundary conditions during summer. *Clim. Dyn.* **25**, 55–580.

Diaconescu, E. P., Laprise, R. and Sushama, L. 2007. The impacts of lateral boundary data errors on the simulated climate of a nested regional climate model. *Clim. Dyn.* **28**, 333–350.

Döscher, R., Willen, U., Jones, C., Rutgersson, A., Meier, H. E. M. and co-authors. 2002. The development of the coupled regional ocean-atmosphere model RCAO. *Boreal Eviron. Res.* **7**, 1221–1234.

Fox-Rabinovitz, M. S., Takacs, L. L., Govindaraju, R. C. and Suarez, M. J. 2001. A variable-resolution stretched-grid general circulation model: regional climate simulation. *Mon. Wea. Rev.* **129**, 453–469.

Giorgi, F. and Mearns, L. O. 1991. Approaches to the simulation of regional climate change: a review. *Rev. Geophys.* **29**, 191–216.

Giorgi, F. and Mearns, L. O. 1999. Introduction to special section: regional climate modelling revisited. *J. Geophys. Res.* **104**(D6), 6335–6352.

Haugen, J. E. and Haakenstad, H. 2006. The development of HIRHAM version 2 with 50 km and 25 km resolution. RegClim Gen. Techn. Rep. No. 9. Norwegian Meteorological Institute, P. O. Box 43 Blindern, N-0313 Oslo, Norway.

Haugen, J. E. and Iversen, T. 2008. Response in extremes of daily precipitation and wind from a downscaled multi-model ensemble of anthropogenic global climate change scenarios. *Tellus* ***60A***, 10.1111/j.1600-0870.2008.00315.x.

Herceg, D., Sobel, A. H., Sun, L. and Zebiak, S. E. 2006. The Big Brother Experiment and seasonal predictability in the NCEP regional spectral model. *Clim. Dyn.* **27**, 69–82.

IPCC. 2001. *Climate Change 2001: The Scientific Basis* (eds J. T. Houghton, and co-editors), Cambridge University Press, UK. 881 pp.

IPCC. 2007. Climate change 2007: the physical science basis. *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (eds S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, and co-editors), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

Jacob, D., Barring, L., Christensen, O. B., Christensen, J. H., de Castro, M. and co-authors. 2007. An intercomparison of regional climate simulations for Europe: model performance in present-day climate. *Clim. Change* **81**, 31–52.

Jones, R. G., Murphy, J. M. and Noguer, M. 1995. Simulation of climate change over Europe using a nested regional-climate model. I: assessment of control climate, including sensitivity to location of lateral boundaries. *Quart. J. Roy. Meteorol. Soc.* **121**, 1413–1449.

Jones, R. G., Murphy, J. M., Noguer, M. and Keen, A. B. 1997. Simulation of climate change over Europe using a nested regional-climate model. II: comparison of driving and regional model responses to a doubling of carbon dioxide. *Quart. J. Roy. .Meteorol. Soc.* **123**, 265–292.

Kida, H., Koide, T., Sasaki, H. and Chiba, M. 1991. A new approach to coupling a limited area model with a GCM for regional climate simulation. *J. Met. Soc. Japan* **69**, 723–728.

Køltzow, M. 2007. The effect of a new snow and sea ice albedo scheme on regional climate model simulations. *J. Geophys. Res.,* **112**, D07110, doi:10.1029/2006JD007693.

Lambert, S. J. and Boer, G. J. 2001. CMIP1 evaluation and intercomparison of coupled climate models. *Clim. Dyn.* **17**, 83–106.

Machenhauer, B., Jacob, D. and Bozet, M. 1994. Using the MPI's nested limited area model for regional climate simulations. CAS/JSC WGNE Report No. 19. WMO/TD-No 592, pp. 758–760.

Noguer, M., Jones, R. and Murphy, J. 1998. Sources of systematic errors in the climatology of a regional climate model over Europe. *Clim. Dyn.* **14**, 691–712.

Rinke, A., Dethloff, K., Spekat, A., Enke W. and Christensen, J. H. 1999. High resolution climate simulations over the Arctic. *Polar Res.* **18**, 1–9.

Rinke, A., Gerdes, R., Dethloff, K., Kandlbinder, T., Karcher, M. and co-authors. 2003. A case study of the anomalous Arctic sea ice conditions during 1990: insights from coupled and uncoupled regional climate model simulations. *J. Geophys. Res.* **108**(D9), 4275, doi:10.1029/2002JD003146.

Rinke, A., Dethloff, K., Cassano, J. J., Christensen, J. H., Curry, J. A. and co-authors. 2006. Evaluation of an ensemble of Arctic regional climate models: spatiotemporal fields during the SHEBA year. *Clim. Dyn.* **27**, 459–472.

Roeckner, E., Arpe, K., Bengtsson, L., Christoph, M., Claussen, M. and co-authors. 1996. The atmospheric general circulation model ECHAM4: model description and simulation of present-day climate: MPI Rep No. 218.

Rummukainen, M., Räisänen, J., Bringfelt, B., Ullerstig, A., Omstedt, A. and co-authors. 2001. A regional climate model for northern Europe: model description and results from the downscaling of two GCM control simulations. *Clim. Dyn.* **17**, 339–359.

Schär, C. and Jendritzky, G. 2004. Hot news from summer 2003. *Nature* **432**, 559–560.

Undén, P., Rountu, L., Järvinen, H., Lynch, P., Calvo, J. and co-authors. 2002. *HIRLAM-5 Scientific Documentation*. HIRLAM-5 project, c/o Per Undén SMHI, S-60176 Norrkøping, Sweden.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., da Costa Bechtold, V. and co-authors. 2005. The ERA-40 re-analysis. *Quart. J. Roy. Meteorol. Soc.* **131**, 2961–3012. doi:10.1256/qj.04.176.

Von Storch, H., Langenberg, H. and Feser, F. 2000. A spectral nudging technique for dynamical downscaling purposes. *Mon. Wea. Rev.* **128**, 3664–3672.