

Statistical Downscaling Based on Dynamically Downscaled Predictors: Application to Monthly Precipitation in Sweden

Cecilia HELLSTRÖM*¹ AND Deliang CHEN^{1,2}

¹*Earth Sciences Centre, Gothenburg University, Gothenburg, Sweden*

²*National Climate Center, China Meteorological Administration, Beijing, China*

(Received 2 September 2002; revised 3 July 2003)

ABSTRACT

A prerequisite of a successful statistical downscaling is that large-scale predictors simulated by the General Circulation Model (GCM) must be realistic. It is assumed here that features smaller than the GCM resolution are important in determining the realism of the large-scale predictors. It is tested whether a three-step method can improve conventional one-step statistical downscaling. The method uses predictors that are upscaled from a dynamical downscaling instead of predictors taken directly from a GCM simulation. The method is applied to downscaling of monthly precipitation in Sweden. The statistical model used is a multiple regression model that uses indices of large-scale atmospheric circulation and 850-hPa specific humidity as predictors. Data from two GCMs (HadCM2 and ECHAM4) and two RCM experiments of the Rossby Centre model (RCA1) driven by the GCMs are used. It is found that upscaled RCA1 predictors capture the seasonal cycle better than those from the GCMs, and hence increase the reliability of the downscaled precipitation. However, there are only slight improvements in the simulation of the seasonal cycle of downscaled precipitation. Due to the cost of the method and the limited improvements in the downscaling results, the three-step method is not justified to replace the one-step method for downscaling of Swedish precipitation.

Key words: downscaling, multiple regression, atmospheric circulation indices, monthly precipitation, Sweden

1. Introduction

The awareness of global warming caused by the enhanced greenhouse effect during the industrial period has increased the demand of climate change scenarios. On a global scale, such scenarios are best produced by GCMs. At present, GCMs, due to their relatively low spatial resolution, fail to reliably simulate the climate at regional and local scales. However, climate scenarios on regional and local scale are essential for impact studies. As a consequence, various methods aiming at downscaling the GCM output have been developed. The methods can be divided into two groups: (1) dynamical downscaling, in which a high-resolution Regional Climate Model (RCM) is nested within a GCM (Giorgi, 1990); and (2) statistical downscaling (Karl et al., 1990; Busuioc et al., 2001a), in which relationships between observed large-scale and regional/local climate are identified and applied to GCM output.

Both dynamical and statistical downscalings have

their own advantages and disadvantages. Dynamical downscaling is essentially forced by a GCM through boundary conditions at the boundary of a RCM domain, which allows the RCM to develop its own climate within the domain. Alternatively, the large-scale features contained in the RCM simulation in the upper air can be forced towards the large-scale features of the driving GCM, while the small-scale ones from the RCM simulation, especially those near the surface, are left untouched (Von Storch et al., 2000). The nesting of the RCM within the GCM is based on the assumption that the large-scale climate is well simulated by the GCM. In Busuioc et al. (2001a), it was found that the SLP variability of a North Atlantic-European sector is fairly well reproduced by the HadCM2 GCM. However, in the present study it is shown that the simulation of the atmospheric circulation over Scandinavia, represented by circulation indices of geostrophic wind components and vorticity, is connected with significant deficiencies. Due to its finer resolution, the

*E-mail: cecilia@gvc.gu.se

small-scale variability, especially that near the surface, is more realistically represented in the RCM than in the driving GCM. It may be assumed that the inclusion of the small-scale information in the upscaled large-scale field should have positive effects for scales equal to and larger than the RCM resolution, including scales equal to and larger than the GCM resolution. The RCM resolution increases the ability for better representations of topography, water/land distribution, and local circulation systems. If this assumption can be approved, it will open a new possibility to do better statistical downscaling with the help of dynamical downscaling. Although the GCM-RCM coupling here is one-way, there are two-way interactions within the RCM domain. The need to perform a statistical downscaling even when a dynamical downscaling is available can be justified by the facts: (1) statistical downscaling can easily handle a variety of scales including the station scale; (2) there are climate variables (e.g., sea ice) for which it may be difficult or even impossible to achieve realistic simulations by current RCMs (Omstedt and Chen, 2001).

There are a number of factors contributing to uncertainties in regional/local scenarios. First of all, downscaling models are dependent on the skill of GCMs to simulate the present and future large-scale climate. However, even if the skills of different GCMs are about the same, they can still produce different scenarios (Räisänen, 2000; Barnett et al., 2000). Additionally, different kinds of downscaling methods with equal skill in simulating the present climate can give divergent regional climate scenarios (Mearns et al., 1999; Murphy, 2000). In statistical downscaling, the reliability of the results depends on, for example, how well the predictors are simulated by GCMs (Palutikof et al., 1997; Winkler et al., 1997) and how stable the connections between predictors and predictands are.

Very often, the statistical and dynamical downscalings are used independently, although RCM simulations have been used as predictors in statistical downscaling by Charles et al. (1999). A statistical-dynamical downscaling scheme was presented and applied to Alpine precipitation by Fuentes and Heimann (1996; 2000). They used a regional model to simulate the weather during the most typical multi-day episodes of quasi-stationary circulation classes. The precipitation climate was then calculated from the regional model results which were statistically weighted with the climatological frequencies of the circulation classes. Here, another approach to combine the strength of the two methods is presented. This study tests the ability of such an approach to reduce the uncertainty in statistical downscalings by taking advantages of dynamical

downscaling. What is tested here is the use of predictors taken from upscaled RCM simulations (three-step method) rather than directly from the GCM (one-step method). The basic steps of the method are illustrated schematically in Fig. 1. The method rests on the assumption that the reliability of the predictors at the large scale can be increased by upscaled RCM simulations compared with GCM simulations.

The method is applied to monthly precipitation at 42 Swedish stations. This is done by comparing results from statistical downscaling using predictors upscaled from RCM simulations and directly from GCM output. The comparison is based on how well the predictors from the RCM and GCMs represent the observed seasonal cycle and interannual variability for the present climate. Then, the reproductions of the station precipitation created by the statistical and dynamical downscaling models are compared. The simulation of the seasonal cycle is examined since this is a way to test the skill of the model in responding to a basic and clear signal of external forcing. Also, a correct simulation of the seasonal cycle is important in impact assessment of, for example, agriculture, hydropower, and dam constructions for which the timing of the rain is critical. The quality of the reproduction of interannual variability is essential in analyses of, for example, extreme events. Further, the effects of increased reliability of the predictors are examined in terms of the spread of the precipitation scenarios obtained by the dynamical and statistical downscaling methods.

2. Data, models, and method

The statistical downscaling model is a multiple regression model (Hellström et al., 2001), incorporating monthly anomalies of the large-scale atmospheric circulation and large-scale specific humidity at 850 hPa (q_{850}). A model for each month and station is developed by using NCEP reanalysis data (Kalnay et al., 1996) for the period 1958–1997. The large-scale circulation is represented by two geostrophic wind components (u and v) and total vorticity (ξ), calculated from monthly mean sea level pressure (MSLP) data on a 5° latitude by 10° longitude grid-point basis bounded by latitudes 52.5° – 72.5° N and longitudes 5° – 27.5° E (Chen, 2000). All the circulation indices have the units of hPa per 10° latitude at 60° N. The circulation indices have been widely applied in Sweden (Blenckner and Chen, 2003; Chen and Li, 2003; Hellström et al., 2001; Omstedt and Chen, 2001). The large-scale specific humidity is an average over the area between 55° N and 70° N and between 10° E and 25° E. The selection of the area for q_{850} is based on how much the area-

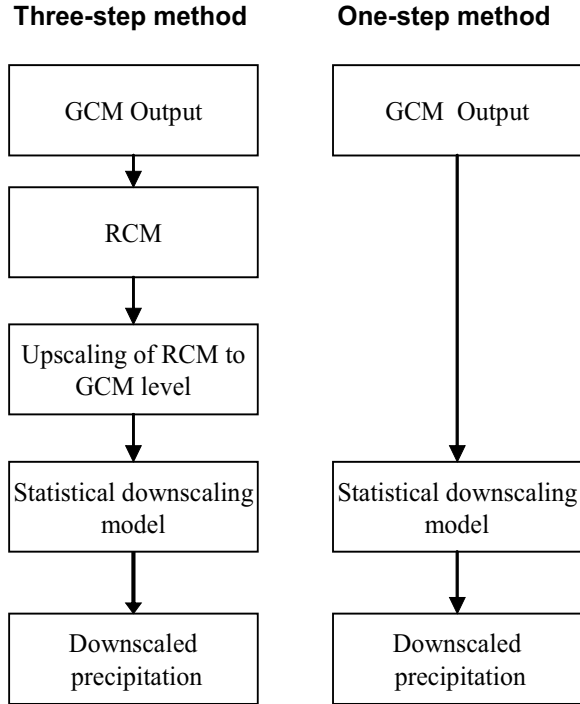


Fig. 1. Flowchart of the basic steps of the three- and one-step downscaling methods.

averaged index contributes to the regression model. The precipitation data used for fitting the statistical model and for verification of the control simulations are monthly time series of measurements at 42 Swedish stations from 1958 to 1997. The data have been homogenized/corrected through the North Atlantic Climatological Dataset (NACD) program (Frich et al., 1996). The data have recently been updated and provided by the Swedish Meteorological and Hydrological Institute (SMHI). Details of the stations are given by Hellström et al. (2001). The positions of the stations can be found in Fig. 2.

Four present day (control) and future (scenario) precipitation climates are produced by the statistical model using predictors from simulations of two GCMs and associated RCM experiments. The GCMs used are HadCM2 (Johns et al., 1997; Mitchell and Johns, 1997) and ECHAM4/OPYC3 (Oberhuber, 1993; Roeckner et al., 1996, 1999). A time slice of 10 years is used to represent the present and future climates. All the RCM experiments are made with version RCA1 of the Rossby Centre Regional Atmospheric climate model for northern and central Europe forced by HadCM2 and ECHAM4 (Rummukainen et al., 2001; Räisänen et al., 2001). The model is run at a 44-km horizontal resolution.

The MSLP and q_{850} data of the RCA1 experiments are upscaled prior to the predictor calculations by be-

ing interpolated and aggregated to a 5° latitude by 10° longitude grid by a linear scheme. The scheme works as a simple filter with equal weights for all grid points within the $5^\circ \times 10^\circ$ grid. The MSLP and q_{850} data of the GCMs are interpolated in the same manner. The seasonal cycles of the predictors from the reanalysis, the GCMs, and RCA1 of the control runs are compared. The observed seasonal cycles of precipitation at the 42 stations for the period 1957–1998 are calculated and compared with the precipitation statistically downscaled from the GCMs and RCA1 predictors as well as those of the RCA1 and GCM experiments. Precipitation amounts at all the stations from the GCM and RCA1 simulations are obtained by a linear interpolation to the coordinates of the 42 stations. In Fig. 2, the 42 stations are shown as well as the grids of HadCM2 and RCA1.

3. Results and discussion

For the sake of brevity, the RCA1 experiments driven by HadCM2 and ECHAM4 will be referred to as DDH (Dynamical Downscaling of HadCM2) and DDE (Dynamical Downscaling of ECHAM4), respectively, in the following text. The statistical model applied to the HadCM2, ECHAM4, DDH, and DDE predictors will be referred to as GSDH (GCM based Statistical Downscaling, with HadCM2 predictors), GSDE (GCM based Statistical Downscaling, with ECHAM4 predictors), RSDH (RCA1 based Statistical Downscaling driven by HadCM2), and RSDE (RCA1 based Statistical Downscaling driven by ECHAM4), respectively.

3.1 Seasonal cycles of the predictors

The seasonal cycle is one of the most important features of climate variability and has been used widely in validation of climate models (e.g., Covey et al., 2000). The seasonal cycles of the modeled predictors by the GCMs and the upscaled RCA1 results are compared with those derived from the reanalysis. In order to quantify the differences, some statistics are needed. The correlation coefficient (R), variance (γ^2 ; normalized with respect to the observed variance), and root mean squared errors (RMSE) are used to reveal the similarities. While the correlation coefficient describes how well the shape of the seasonal cycle is reproduced, the bias tells us about the systematic deviations of the annual precipitation and the variance reflects the inter-monthly variability. The RMSE can be resolved into two parts, one part that is related to differences in the total mean (E_0), or bias, and another which reflects

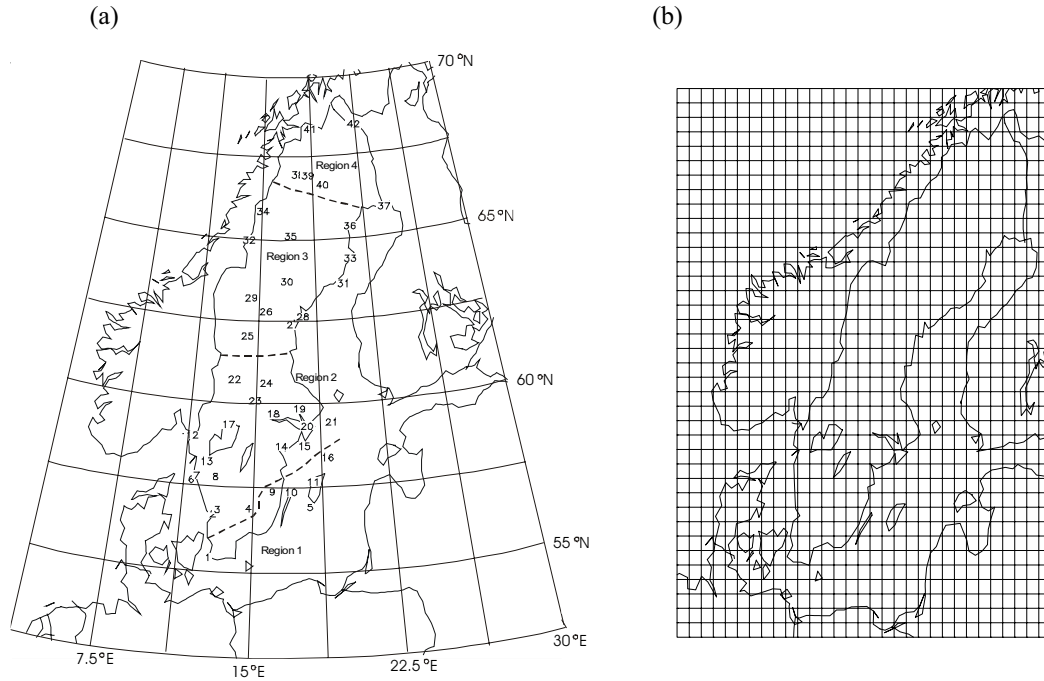


Fig. 2. (a) The grid of the HadCM2 GCM (low resolution), the 42 stations used, and the four regions. (b) The grid of the RCA1 (high resolution).

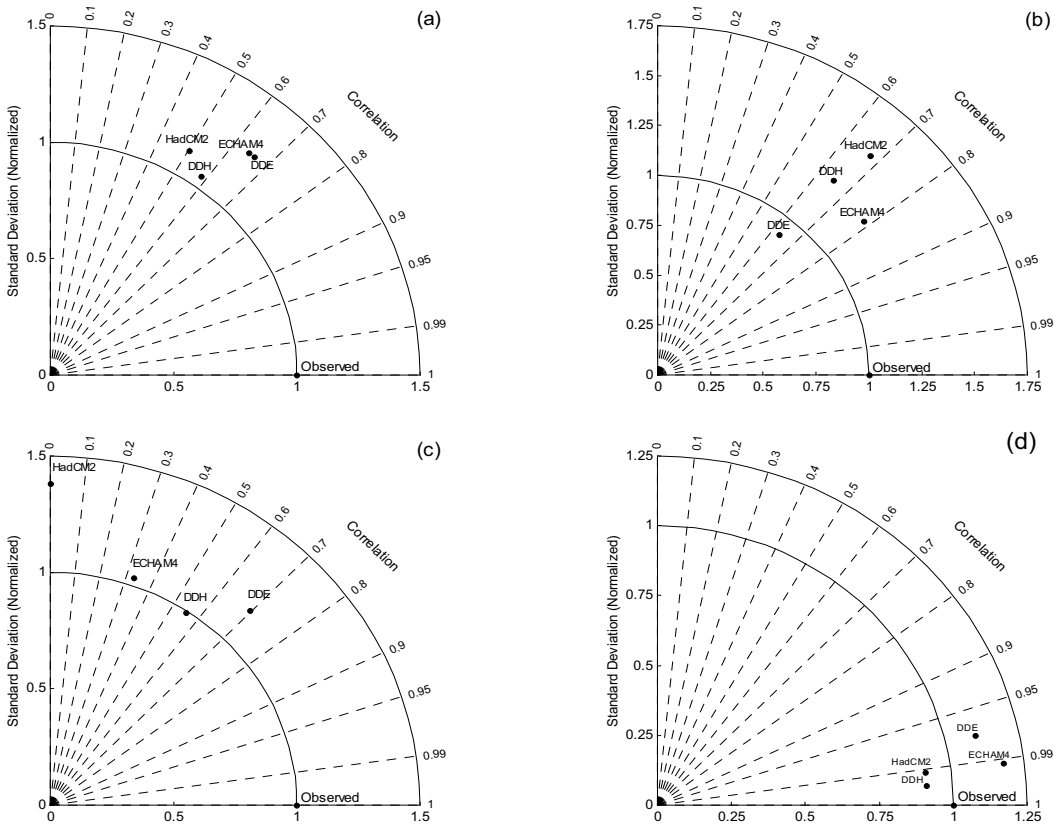


Fig. 3. Pattern statistics (Taylor diagram) of the seasonal cycle of predictors of the GCMs and the RCM runs. Standard deviations are normalized against observed data. (a) u ; (b) v ; (c) ξ ; and (d) q_{850} .

differences in the pattern of variations (E'). E' is related to R and γ^2 by the following equation (Taylor, 2001):

$$E'^2 = 1 + \gamma^2 - 2\gamma R.$$

Figure 3 shows the correlation coefficients between the reanalysis and the modeled predictors as well as the normalized variance of the modeled predictors. Differences between modeled and observed patterns of variations (E') are in the diagram proportional to the distance between the observed point and the points of the models (Taylor, 2001). In Fig. 4, the bias and the total RMSE of the monthly values are plotted.

For ξ , which is found to be most strongly related to precipitation in Sweden (Hellström et al., 2001), great improvements in the correlation are evident in the RCA1 experiments (DDH and DDE) compared to the GCM runs. Somewhat higher correlation is also found in u . It should be mentioned that u is highly correlated with NAO, thereby making it an important factor in determining the regional and local climate in Sweden (e.g., Chen, 2000; Busuioc et al., 2001b). The correlation of the seasonal cycles of v is higher for ECHAM4 than for DDE and about the same for HadCM2 and DDH. The correlation coefficients related to q_{850} are all very high and the differences between the models are therefore insignificant. It is found in the comparison of the RCA1 and GCM experiments that the skill of the models in reproducing the intermonthly variability is on average about the same for the GCMs and the RCA1. For u , the variances in the DDH experiments are closer to the observed ones than those in HadCM2. Also, DDH shows better results than HadCM2 for v . The difference between the observed and modeled variance is smaller in ECHAM4 than in DDE for ξ . The intermonthly variability is generally lower in the DDH and DDE experiments than in the corresponding GCMs. In short, the RCA1 predictors are associated with smaller E' than the GCM predictors with the exception of v and q_{850} in DDE. Furthermore, all differences in correlation and variance between the output of the GCMs and those of the corresponding RCA1 runs are statistically insignificant at the 95% level, except for ξ in HadCM2 and DDH.

The bias is smaller in HadCM2 than in DDH for all predictors except u . In general, the bias in v is relatively small, however. DDE is associated with a much lower bias than ECHAM4 for ξ and q_{850} . It is interesting to note that both of the RCA1 experiments are associated with a loss of humidity compared with the GCM experiments. This leads to a relatively large negative bias in DDH. RCA1 has a positive bias in precipitation, which may partly explain the systematic

loss of q_{850} (see below). Also, for the other predictors, RCA1 tends to reduce the mean values, except for u in DDH. With respect to the RMSE, RCA1 is superior to the GCMs for vorticity.

3.2 Seasonal cycle of downscaled precipitation

It is expected that the improved seasonal cycles of the predictors have positive effects on the statistical downscaling of the precipitation at the 42 stations. However, the different influences on the bias have a varying effect on the downscaled precipitation at different stations. To examine the combined effect of the two sets of predictors from RCA1 and the GCMs, the local precipitation at all the stations is calculated from the statistical model using the two sets of predictors, and is compared with that directly interpolated from the GCMs and RCA1.

To simplify the comparison, the statistics calculated for the predictors are also calculated for the modeled precipitation for four regions with similar precipitation variability (Fig. 2). The four regions were distinguished in an earlier work (Busuioc et al., 2001b) by using EOF and cluster analysis of annual precipitation. The four regions represent: (1) the south easternmost part of Sweden (6 stations); (2) the southern part of the country (18 stations; excluding region 1); (3) the middle and the southern part of northern Sweden (12 stations; $\sim 61^\circ$ – 66° N); and (4) northernmost Sweden (6 stations; $\sim 66^\circ$ – 69° N). In Hellström et al. (2001), it was concluded that the statistical and dynamical downscaling models are superior to the GCMs with respect to their ability to reproduce the seasonal cycle of the precipitation at the four stations/regions. It is noted also that the GCMs fail to simulate the summer maximum and the late winter-early spring minimum found in the observed data. Generally, the statistically downscaled cycles based on the GCM and RCA1 predictors follow the same patterns, but the RCA1-based cycles are smoother and closer to the observed ones.

Since the improvements of the RCA1 predictors compared to those of the GCMs vary according to predictor, the final effects of the two sets of predictors on all the station precipitation values need to be assessed individually.

The correlation coefficients between the observed and modeled averaged monthly precipitation of all stations within the specific regions as well as the normalized standard deviations (γ^2) and E' produced by the statistical and dynamical downscaling models are found in Fig. 5. The results directly interpolated from the two GCMs are also shown. It is evident that the differences between the three-step and one-step downscalings vary with region.

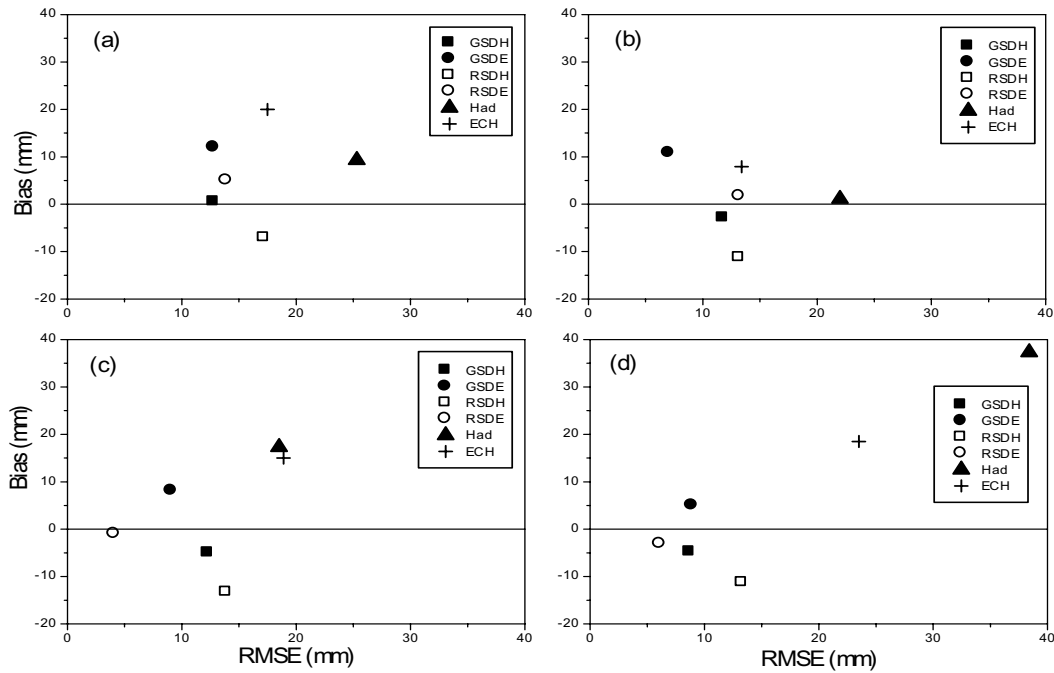


Fig. 6. Bias and total RMSE of seasonal cycle of downscaled precipitation. (a) Region 1; (b) Region 2; (c) Region 3; and (d) Region 4.

With respect to correlation, the three-step and one-step methods produce about the same results. For southern Sweden, the RSDE gives better results than GSDE, and GSDH better than RSDH. The opposite is true for the northern part of the country. The inter-monthly variability is, for all regions, closer to the observed in RSDE than in GSDE. The one-step method based on the HadCM2 predictors produces an inter-monthly variability closer to the observed than the RSDH for all regions except Region 1. Totally, the combined effect of R and γ^2 gives somewhat lower values of E' in the three-step method. All differences in correlation and variance between the precipitation derived from the one-step and three-step methods are statistically insignificant at the 95% level.

The biases are generally larger and more negative in RSDH than in GSDH (Fig. 6). The negative biases are probably connected to the negative bias in ξ and q_{850} used by DDH. RSDE has, for all the regions, a smaller bias than GSDE. It should be noted that the biases in the downscaled precipitation are still considered to be small, since the maximum is only 4.4% of the monthly averages. In terms of RMSE, the three-step method performs worse than the one-step method in all regions, except for RSDE-GSDE in Regions 3 and 4.

4. Conclusions

It is demonstrated that the seasonal cycles of large-

scale predictors can be improved in upscaled RCA1 results compared to GCM results. The increased skill in reproducing the seasonal cycle is greatest for ξ and u for both of the RCA1 experiments, driven by HadCM2 and ECHAM4, respectively.

Despite the improvement in the reproduction of the seasonal cycle of, for example, ξ , there is only a slight improvement in the simulation of the seasonal cycle of precipitation. Considering this together with the availability and cost of RCM experiments, the use of the three-step method is not recommended.

Acknowledgments. The authors thank Jouni Räisänen for providing the GCM and RCM data and Hans Alexandersson for providing the station precipitation data. This research was supported by grants from the Chinese Ministry of Science and Technology (2001BA611B-01), the Chinese Academy of Sciences, and SWECLIM which is financed by MISTRA and SMHI.

REFERENCES

- Barnett, T. P., G. Hegerl, T. Knutson, and S. Tett, 2000: Uncertainty levels in predicted patterns of anthropogenic climate change. *J. Geophys. Res.*, **D105**, 15525–15542.
- Blenckner, T., and D. Chen, 2003: Comparison of the impact of regional and north-Atlantic atmospheric circulation on an aquatic ecosystem. *Climate Research*, **23**, 131–136.
- Busuioc, A., D. Chen, and C. Hellström, 2001a: Performance of statistical downscaling models in GCM val-

- idation and regional climate change estimates: Application for Swedish precipitation. *Int. J. Climatol.*, **21**, 557–578.
- Busuioc, A., D. Chen, and C. Hellström, 2001b: Temporal and spatial variability of precipitation in Sweden and its link with the large-scale atmospheric circulation. *Tellus*, **53A**, 348–367.
- Charles, S. P., B. C. Bates, P. H. Whetton, and J. P. Hughes, 1999: Validation of downscaling models for changed climate conditions: Case study of southwestern Australia. *Clim. Res.*, **12**, 1–14.
- Chen, D., 2000: A monthly circulation climatology for Sweden and its application to winter temperature case study. *Int. J. Climatol.*, **20**, 1067–1076.
- Chen, D., and X. Li, 2003: Scale dependent relationship between maximum ice extent in the Baltic Sea and atmospheric circulation. *Global and Planetary Change* (in press).
- Covey, C., and coauthors, 2000: The seasonal cycle in coupled ocean-atmosphere general circulation models. *Climate Dyn.*, **16**, 775–787.
- Frich, F., and coauthors, 1996: North Atlantic Climatological Dataset (NACD Version 1)—Final Report. *DMI Report*, Copenhagen, Denmark, 47pp.
- Fuentes, U., and D. Heimann, 1996: Verification of statistical-dynamical downscaling in the Alpine region. *Climate Res.*, **7**, 151–168.
- Fuentes, U., and D. Heimann, 2000: An improved statistical-dynamical downscaling scheme and its application to the Alpine precipitation climatology. *Theor. Appl. Climatol.*, **65**, 119–135.
- Giorgi, F., 1990: Simulation of regional climate using a limited area model nested in a general circulation model. *J. Climate.*, **3**, 941–963.
- Hellström, C., D. Chen, C. Achberger, and J. Räisänen, 2001: A comparison of climate change scenarios for Sweden based on statistical and dynamical downscaling of monthly precipitation. *Climate Research*, **19**, 45–55.
- Johns, T. C., R. E. Carnell, J. F. Crossley, J. M. Gregory, J. F. B. Mitchell, C. A. Senior, S. F. B. Tett, and R. A. Wood, 1997: The second Hadley Centre coupled ocean-atmosphere GCM: Model description, spinup and validation. *Climate Dyn.*, **13**, 103–134.
- Kalnay, E., and coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471.
- Karl, T. R., W. C. Wang, M. E. Schlesinger, R. W. Knight, D. Portman, 1990: A method of relating general circulation model simulated climate to observed local climate. Part I: Seasonal statistics. *J. Climate*, **3**, 1053–1079.
- Mearns, L. O., I. Bogardi, F. Giorgi, I. Matyasovszky, and M. Palecki, 1999: Comparison of climate change scenarios generated from regional climate model experiments and statistical downscaling. *J. Geophys. Res.*, **104**, 6603–6621.
- Mitchell, J. F. B., and T. C. Johns, 1997: On modification of global warming by sulphate aerosols. *J. Climate*, **10**, 245–267.
- Murphy, J., 2000: Predictions of climate change over Europe using statistical and dynamical techniques. *Int. J. Climatol.*, **20**, 489–501.
- Oberhuber, J. M., 1993: Simulation of the Atlantic circulation with a coupled sea ice-mixed layer-isopycnal general circulation model. Part I: Model description. *J. Phys. Oceanogr.*, **22**, 808–829.
- Omstedt, A., and D. Chen, 2001: Influence of atmospheric circulation on the maximum ice extent in the Baltic Sea. *J. Geophys. Res.*, **106**, 4493–4500.
- Palutikof, J. P., J. A. Winkler, C. M. Goodess, and J. A. Andresen, 1997: The simulation of daily temperature time series from GCM output. Part I: Comparison of model data with observations. *J. Climate.*, **10**, 2497–2513.
- Räisänen, J., 2000: CO₂-induced climate change in northern Europe: Comparison of 12 CMIP2 experiments. Reports Meteorology and Climatology, No. 87, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden, 59pp.
- Räisänen, J., M. Rummukainen, and A. Ullerstig, 2001: Downscaling of greenhouse gas induced climate change in two GCMs with the Rossby Centre regional climate model for northern Europe. *Tellus*, **53A**, 168–192.
- Roeckner, E., and coauthors, 1996: The atmospheric general circulation model ECHAM-4: Model description and simulation of present-day climate. Max-Planck-Institut für Meteorologie, Report No.218, Hamburg, Germany, 90pp.
- Roeckner, E., L. Bengtsson, J. Feichter, J. Lelieveld, and H. Rodhe, 1999: Transient climate change simulations with a coupled atmosphere-ocean GCM including the tropospheric sulfur cycle. *J. Climate*, **12**, 3004–3032.
- Rummukainen, M., J. Räisänen, B. Bringfelt, A. Ullerstig, A. Omstedt, U. Willén, and U. Hansson, 2001: A regional climate model for northern Europe—Model description and results from the downscaling of two GCM control simulations. *Climate Dyn.*, **17**, 339–359.
- Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183–7192.
- Von Storch, H., H. Langeberg, and F. A. Feser, 2000: Spectral nudging technique for dynamical downscaling purposes. *Mon. Wea. Rev.*, **128**, 3664–3673.
- Winkler, J. A., J. P. Palutikof, J. A. Andresen, and C. M. Goodess, 1997: The simulation of daily temperature time series from GCM output. Part II: Sensitivity analysis of an empirical transfer function methodology. *J. Climate*, **10**, 2514–2532.