Evaluation and Calibration of Dynamically Downscaled Precipitation over Norwegian Mainland

Ola Haug, Elisabeth Orskaug, Ida Scheel, Arnoldo Frigessi, Douglas Maraun and Peter Guttorp

Abstract Turning future climate projections into robust and reliable information available at a local scale is imperative for the successful modeling of impacts of climate change. In this paper we investigate ERA-40 reanalysis precipitation data downscaled by the regional model HIRHAM, and compare them by statistical methods to observed precipitation on a local grid over Norwegian mainland. In general, our results indicate that the regional model has too many but too small rain events for all seasons. Although the model is to some extent skillful in describing the lower quartile of the precipitation distribution, the evaluation shows increasing discrepancies over the distributional range of the model data. Based on Doksum's shift function, we suggest a spatially smoothed model forming a full quantile calibration.

Ola Haug
Norwegian Computing Center, PO Box 114 Blindern, N-0314 Oslo, Norway, e-mail: ola.haug (a) nr.no

Elisabeth Orskaug
Norwegian Computing Center, PO Box 114 Blindern, N-0314 Oslo, Norway, e-mail: elisabeth.orskaug (a) nr.no

Ida Scheel
Department of Mathematics, University of Oslo, PO Box 1053 Blindern, N-0316 Oslo, Norway, e-mail: idasch (a) math.uio.no

Arnoldo Frigessi
Department of Biostatistics, University of Oslo, PO Box 1122 Blindern, N-0317 Oslo, Norway, e-mail: arnoldo.frigessi (a) medisin.uio.no

Douglas Maraun
Ocean Circulation and Climate Dynamics, Leibniz-Institut of Marine Sciences, Kiel University (IFM-GEOMAR), Duesternbrooker Weg 20, D-24105 Kiel, Germany, e-mail: dmaraun (a) ifm-geomar.de

Peter Guttorp
Department of Statistics, University of Washington, Box 354322, Seattle, WA 98195-4322, USA, and Norwegian Computing Center, PO Box 114 Blindern, N-0314 Oslo, Norway, e-mail: peter (a) stat.washington.edu

Ninth International Geostatistics Congress, Oslo, Norway, June 11. – 15., 2012

1 Introduction

The intensification of climate research over the past decade produces a steadily increasing number of data sets combining different global circulation models, CO_2 emissions scenarios and downscaling techniques. Turning future projections into robust and reliable information available at a local scale is imperative for the successful modeling of impacts of climate change in nature and society. The comprehensive financial and safeguarding challenges of mitigation and adaptation call for thorough validation, improvement and extensions of current downscaling techniques.

The purpose of regional climate models is to give stakeholders and decision makers a representation, possibly a reliable projection, at a useful spatio-temporal scale, of future weather events. In the insurance industry, for instance, the interest lies in precipitation projections under various possible futures to assess the changing risk of damages to buildings or flooding [7]. Typically scenario runs are done with the regional model forced by a global coupled ocean-atmosphere model. The question then becomes how reliable these regional models are, at the scale needed by the actual effect study. For example, to understand patterns of risk for the insurance of buildings, precipitation at a mesoscale level are needed, say on a 25×25 km^2 grid, at least daily.

Like weather forecasts in general, being just an incomplete representation of the physics involved, the downscaling introduces inaccuracies and errors in the resulting weather variables [1]. The scale of the errors varies geographically depending on the current state of the atmosphere best represented by what is denoted reanalysis data. Feeding the model with reanalysis data along the boundary of an integration area, emphasizes the errors of the downscaling process itself and minimizes the contribution added from propagating discrepancies inherently present in the boundary conditions. Reliant on the downscaling, the 25×25 km^2 downscaled reanalysis data are still supposed to possess properties similar to real weather locally over longer time periods.

This paper uses findings of a downscaling evaluation study [6] as a motivation for constructing a full distributional calibration model for downscaled reanalysis precipitation data.

2 Data

The data used in this study constitute 40 years of daily precipitation values for the Norwegian mainland, covering the period 1961 to 2000. The data set is twofold, where one part consists of downscaled ERA-40 reanalysis model data and the other is based on observations. A more thorough description of the data is given in [6].

ERA-40 Reanalysis Data

In vague terms, reanalysis data express the best estimate available for the current state of the atmosphere. They are formed in retrospect from feeding various sources of meteorological observations into a computerized atmospheric model that smoothes the observations and brings them into consistency. ERA-40 reanalysis data are a product of the ECMWF (European Center for Medium-Range Weather Forecasts) in the UK. Our source data outputs the state of the atmosphere through daily meteorological variables on a 125×125 km^2 grid.

Downscaled ERA-40 reanalysis data are collected from the ENSEMBLES project web-site [1]. Gridded large scale ERA-40 data along the boundary of an integration area covering most of Europe are dynamically downscaled to weather variables on a grid with a spatial resolution of 25×25 km^2, which amounts to 777 grid cells over Norwegian mainland. The downscaled ERA-40 reanalysis data [3] will be referred to as dERA40 in this paper. The downscaling is done by the Norwegian Meteorological Institute by means of their HIRHAM Regional Climate Model [4].

Observation Based Data

Precipitation is observed daily by stations irregularly distributed across Norway. Based on all observations of precipitation available at every time, high-resolution precipitation grids (1×1 km^2) are estimated applying a Delaunay triangulation [5]. The interpolated precipitation values are adjusted locally by taking the deviations between triangulated station elevations and ground heights as given by a real terrain model into account. Also, prior to the interpolation, observed precipitation is corrected for exposure dependant undercatch due to wind loss.

In order to compare the two data sets, the 1×1 km^2 observation grid is aggregated into the larger 25×25 km^2 grid of dERA40. This is obtained by collecting all 1×1 km^2 grid cells with centre points within an ERA-40 cell, and taking their mean as a representation of the measured precipitation inside that grid cell. We use the abbreviation OBS for this data.

3 Evaluation Study

Using reanalysis data as forcing, we have investigated by statistical techniques how well the Norwegian regional model HIRHAM compares to triangulated and aggregated station measurement data on a 25×25km^2 grid over Norwegian mainland. Statistical testing is shown to contribute to the purpose of a standardized assessment for dynamical downscaling models. Methods considered are the Kolmogorov-Smirnov two-sample test, a Fisher exact test for equality of quantiles, an Extreme Value Theory test, where equality of the one-year return levels are tested, and equality of wet-day frequency. All tests are performed seasonally.

The regional model is skillful in describing the lower quartile of the precipitation distribution, but underestimates higher levels of precipitation. A sample result of the quantile test for the summer season is shown in Fig. 1. The 0.05 quantile panel shows no significant difference between the OBS and dERA40 data sets for most grid cells, whereas the 0.95 quantile panel contains mostly dark colour shades indicating OBS quantiles are significantly larger as compared to dERA40. Our results also indicate that the regional model has too many but too small rain events for all seasons. In brief, the evaluation says that for the purpose of using regional models to produce realistic rainfall at the detailed level needed for regional planning and other impact studies, there is still a way to go. Further details are available from [6].

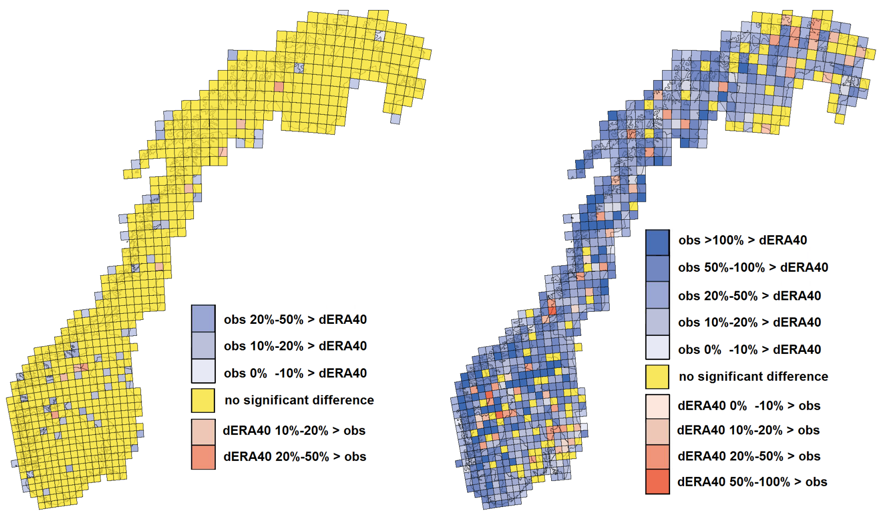

Fig. 1 Quantile test: Equality in quantiles (grid cell by grid cell) of the two data sets for the summer season at the $\alpha = 5\%$ significance level. **Left** 0.05 quantile. **Right** 0.95 quantile.

4 Calibration Model

The results of the evaluation underlines the need for enhanced climate projections at a local scale. Generally, discrepancies between the two distributions exist for the whole range of data, leaving demand for a full quantile calibration function. We address this issue through developing a spatially smoothed model that will make the model distribution closer to that of the observed precipitation data. The transfer functions between the two distributions are characterised using Doksum's shift function.

Assume Norway is divided into grid cells. For grid cell $i, i = 1, \ldots, S$, let X_i denote the precipitation modelled by dERA40 and Y_i actual precipitation. Let F_i be the cumulative distribution function of X_i and G_i be the cumulative distribution function of Y_i. We have dERA40 model output precipitation x_{it} and observed precipitation y_{it} for historic $t = 1, \dots, T$ (daily resolution for 40 years). Thinking of dERA40 reanalysis data as a substitute for global climate model data, our interest lies in distributional coherence between the downscaled model output and the actual precipitation rather than daily correspondence between x_{it} and y_{it}.

Doksum's shift function [2],

$$D_i(x_i) = G_i^{-1}(F_i(x_i)) - x_i \tag{1}$$

implies that $D_i(X_i) + X_i$ has the same distribution as Y_i. If this shift function is constant, it means that there is only a difference in location between the two distributions (and particularly constant equal zero implies no difference in distribution), if it is linear then a location-scale transformation is implied.

Doksum's shift function can be approximated by

$$\widehat{D}_i(x_i) = \widehat{G}_i^{-1}(\widehat{F}_i(x_i)) - x_i \tag{2}$$

where $\widehat{F}_i$ and $\widehat{G}_i$ are the empirical cumulative distribution functions.

Now, introduce

$$z_{it} = \widehat{D}_i(x_{it}) = \widehat{G}_i^{-1}(\widehat{F}_i(x_{it})) - x_{it} \tag{3}$$

where $\widehat{F}_i$ is estimated from x_{it}, $t = 1, \dots, T$ and $\widehat{G}_i$ from y_{it}, $t = 1, \dots, T$. z_{it}, $i = 1, \dots, S$, $t = 1, \dots, T$ will act as our "data" (and then we "forget" about the y_{it}'s).

We model z_{it} by piecewise (binned) constant term regression, using B bins separated by break points $\{q_{ib}\}_{b=0}^{B}$, and with nearest neighbour spatial smoothing on the regression coefficients. The binning is done in either of two ways. First, in what we call *quantile binning*, for each grid cell the data are distributed evenly into the bins so that they all contain the same number of observations. The break points $\{q_{ib}\}_{b=0}^{B}$ represent some common set of quantiles across all grid cells. Here, data values exist for all bins for all grid cells. However, a bin b may represent quite different quantities of precipitation even for neighbouring grid cells. In particular, this means that smoothing tail bins may imply smoothing bins with significantly different amounts of precipitation. Alternatively, data could be arranged by *absolute binning* where $\{q_{ib}\}_{b=0}^{B}$ are set at fixed, common values in all grid cells. Depending on the break points used, this approach tends to make some bins empty in certain grid cells. The argument behind this kind of smoothing is the one of borrowing strength across areas with sparse data.

Implicitly, we model the pair (Z_{it}, X_{it}) through

$$\begin{aligned}
Z_{it} &\sim N(\alpha_{ib(X_{it})}, \sigma^2_{b(X_{it})}), \;\text{ where } b(X_{it}) = \{b : q_{i(b-1)} \le X_{it} < q_{ib}\} \in \{1, 2, \dots, B\} \\
\alpha_{ib(X_{it})} &\sim \exp\big(-\phi_{b(X_{it})} \sum_{j \sim i} w_{ji} (\alpha_{ib(X_{it})} - \alpha_{jb(X_{it})})^2 \big) \qquad (4) \\
\sigma^2_{b(X_{it})} &\sim \text{Maybe Inverse-Gamma} \\
\phi_{b(X_{it})} &\sim \text{Maybe Unif}(0, U_\alpha)
\end{aligned}$$

(5)

where $j \sim i$ is the set of neighbours of grid cell i. The weights w_{ji} are calculated from the difference δ_{ij} between mean observed precipitation for grid cells j and i which is meant to be a proxy for meteorological closeness.

The data are divided into a training set and a test set in order to evaluate the posterior predictive performance. The training set is used for the posterior analysis, and then the posterior predictive distribution for the test period is compared to the observed test set.

We have started to experiment with various priors and various levels of smoothing, for both binning strategies. So far we are not able to show sufficiently robust results. For example, it is not clear what level of smoothing is chosen a posteriori, which means that we do not know yet if smoothing is justified and useful for our actual data. However, we believe that the method has potentials, beyond the current application. The poster we will present at the conference will hopefully have some clear conclusions.

References

1. Christensen, J.H., Kjellstrom, E., Giorgi, F., Lenderink, G., Rummukainen, M.: Assigning relative weights to regional climate models: Exploring the concept. Climate Research **44**, 179–194 (2010)
2. Doksum, K.: Empirical probability plots and statistical inference for nonlinear models in the two-sample case. Ann. Statist. **2**(2), 267–277 (1974)
3. Haugen, J.E., Haakenstad, H.: The development of HIRHAM version 2 with 50km and 25km resolution. In RegClim General Technical Report 9, Norwegian Meteorological Institute (2006)
4. Haugen, J.E., Iversen, T.: Response in extremes of daily precipitation and wind from a downscaled multi-model ensemble of anthropogenic global climate change scenarios. Tellus A **60**(3), 411–426 (2008)
5. Jansson, A., Tveito, O.E., Pirinen, P., Scharling, M.: NORDGRID - a preliminary investigation on the potential for creation of a joint Nordic gridded climate dataset. Tech. Rep. 03/2007, Norwegian Meteorological Institute (2007)
6. Orskaug, E., Scheel, I., Frigessi, A., Guttorp, P., Haugen, J.E., Tveito, O.E., Haug, O.: Evaluation of a dynamic downscaling of precipitation over the Norwegian mainland. Tellus A **63**(4), 746–756 (2011)
7. Scheel, I., Hinnerichsen, M.: The impact of climate change on precipitation-related insurance risk: A study of the effect of future scenarios on residential buildings in Norway. Accepted for publication in The Geneva Papers on Risk and Insurance - Issues and Practice.