Joanna Wibig, Douglas Maraun, Rasmus Benestad, Erik Kjellström,
Philip Lorenz, and Ole Bøssing Christensen

Abstract

General (global) circulation models (GCMs) are a useful tool for studying how climate may change in the future. Although GCMs have high temporal resolution, their spatial resolution is low. To simulate the future climate of the Baltic Sea region, it is necessary to downscale GCM data. This chapter describes the two conceptually different ways of downscaling: regional climate models (RCMs) nested in GCMs and using empirical and/or statistical relations between large-scale variables from GCMs and small-scale variables. There are many uncertainties in climate models, including uncertainty related to future land use and atmospheric greenhouse gas concentrations, limits on the amount of input data and their accuracy, and the chaotic nature of weather. The skill of methods for describing regional climate futures is also limited by natural climate variability. For the Baltic Sea area, the lack of an oceanic component in RCMs and poor representation of forcing by aerosols and changes in land use are major limitations.

J. Wibig (✉)
Department of Meteorology and Climatology, University of Lodz, Lodz, Poland
e-mail: zameteo@uni.lodz.pl

D. Maraun
Ocean Circulation and Climate Dynamics, Geomar Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

R. Benestad
Research Department, Norwegian Meteorological Institute, Oslo, Norway

E. Kjellström
Rossby Centre, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

P. Lorenz
Institute of Meteorology, Freie Universität, Berlin, Germany

O.B. Christensen
Danish Climate Center, Danish Meteorological Institute, Copenhagen, Denmark
e-mail: obc@dmi.dk

10.1   Introduction

The development of general circulation models (GCMs) has created a useful tool for projecting how climate may change in the future. Such models describe the climate at a set of grid points, regularly distributed in space and time and with the same density over land and ocean. Their temporal resolution is relatively high, but their spatial resolution is limited by computing power. Many important processes, such as cloud formation, convection, and precipitation, occur at spatial scales much smaller than the distance between grid points. This means that these so-called sub-grid processes are not explicitly simulated by the models, but must be approximated with simplifying algorithms referred to as parameterisations. The low spatial resolution also means that the topography, coastline, and processes at the land–air, ocean–air, and land–ocean boundaries are coarsely represented in GCMs.

The resolution of present-day GCMs, defined as a distance between two neighbouring points, is of the order of 100–300 km. However, the skilful scale (i.e. the scale at which the climate models are able to capture climate features) is larger, at about 8 grid point distances (Grotch and MacCracken 1991; von Storch et al. 1993), so about 1000–2500 km. This means that GCMs are well able to simulate the atmospheric state at scales greater than the skilful scale in spite of providing values within a grid scale (von Storch et al. 1993). However, the work of Grotch and MacCracken (1991) was based on old models with a small number of vertical levels and simple ocean–atmosphere coupling. A more comprehensive discussion about the skilful scale issue was given by Benestad et al. (2008).

To generate estimates of regional climate, that is, at a scale smaller than the skilful scale, it is necessary to downscale GCM results. Downscaling is understood as a process linking large-scale variables with small-scale variables. There are two conceptually different ways of downscaling. The first uses regional climate models (RCMs) nested in GCMs. RCMs have much higher resolution and can describe local features better, but are still able to simulate the atmospheric state in a realistic manner in their skilful scales. The second uses empirical and/or statistical relations between the large-scale results from GCMs and small-scale variables that describe regional and/or local climate conditions.

Climate projections differ significantly from weather forecasting. Forecasts cannot predict weather with high accuracy beyond a few days. Numerical weather forecasts take observations as a starting point. The number of observations is limited as is the accuracy with which they are made. Small disturbances in the data can cause a large effect on weather after some time. Lorenz (1963) referred to this as the 'butterfly effect'. Climate models are not concerned with weather on a particular day or month or even year but with the statistical features of states of the atmosphere over long periods.

There are also other differences between weather and climate. Weather is forecast for a relatively short time—a few days, generally less than two weeks. This is because changes in weather are caused mainly by changes in the atmosphere. Even changes in oceanic processes have only a very limited influence on the weather because of the longer timescales of typical processes occurring in the oceans. In the case of climate, however, other factors must be taken into account. Climate variations are also caused by changes in the environment: ocean, vegetation, ice, sun, and the composition of the atmosphere. Some of these can be predicted with high accuracy, while others cannot. Among those that cannot are land-use change and the composition of the atmosphere, especially in relation to greenhouse gases (GHGs) and aerosols. As future climate change is to a high degree related to the extent of change in these environmental variables, predicting the future climate requires reliable estimates of the future composition of the atmosphere and land use. As the concentration of GHGs and aerosols in the future atmosphere is so difficult to predict because of the many influencing factors, scenarios are developed based on projections of the future evolution of the world population and economy (see Chap. 11, Sect. 11.2) and it is these scenarios that are used as the basis for projections of future climate.

Beside the uncertainty related to the limited information on land use, and the atmospheric concentrations of GHGs and aerosols, there are also other sources of uncertainty in models. These include limited amounts of input data and their limited accuracy. Due to the chaotic nature of the climate system, a very small difference in initial conditions can generate different climate features, as each simulation generates a different set of realisations. If this were the only source of uncertainty, the differences between simulations should remain within the range of typical climate variability. However, this is not the case. Many sub-grid-scale processes must be simulated in models in a more or less complex form and are not well described by the models. For example, simulations of cloud formation, their optical and radiative features, and the creation of precipitation still carry considerable model error.

For climate models to be useful, they need to be evaluated. As future climate predictions cannot be evaluated by direct comparison with observations, models are evaluated by comparing simulations with observations of the past climate. In theory, this should make it possible to select the best model, but this is not the case in practice. One model can usually describe a particular parameter better than another model, while the second model better describes a different variable or even the same variable, but in another part of the world. There are no objective ways to choose the best model, because none are able to exactly reproduce the observed mean climate and its variability. Differences between simulations and real climate data can be estimated on the basis of a so-called reference period (in the past) for which observational data are available. The differences, usually referred to as 'biases', vary in space and typically also in daily and annual cycles.

The models describe climate at a set of grid points. Because of numerical constraints in GCMs and RCMs, model results at neighbouring grid points are more correlated than actual measurements from two observation points at the same distance (Déqué 2007). This is one reason why the distributions of simulated variables are usually smoothed in comparison with measured station data. Simulations tend to underestimate the highest values and overestimate the lowest (Déqué 2007). This means that the bias is different in different parts of the distribution.

There are a number of sources of uncertainty in climate projections, and thus, preparing scenarios for future change in climate variables is a big challenge. No single method can be used for all variables and all regions.

Natural variability is an important source of uncertainty in climate projections (Deser et al. 2012a). The term 'natural climate variability' refers to variations in climate unrelated to human influences (BACC Author Team 2008). Deser et al. (2012a) analysed how the amplitude of natural variability varies with location in North America. There has been no similar study for the Baltic Sea basin, but these results are also relevant for this region. The analysis showed that natural variability is generally smaller in summer than that in winter and at lower latitudes rather than at higher latitudes. Also, that regional averaging does not always reduce the uncertainty in climate projections (Deser et al. 2012a). Natural climate variability cannot be reduced by better models, downscaling techniques or improved GHG emission scenarios and as a result limits climate predictability. However, natural climate variability can be described and to some extent quantified in an ensemble approach (see Deser et al. 2012a).

10.2 Dynamical Downscaling

The methodology used to achieve climate simulations in high resolution for a specific region by applying RCMs is referred to as 'dynamical downscaling'. RCMs are based on atmospheric limited-area models used in numerical weather prediction. The first application of RCMs for long-term simulations goes back to the work of Dickinson et al. (1989) and Giorgi and Bates (1989). Today, RCMs are used by many institutions and have been applied for a large number of studies, and RCM climate change projections have been undertaken for regions on all continents. There are several recent reviews of RCM methodology and their application (e.g. Giorgi 2006; Foley 2010; Rummukainen 2010).

10.2.1 Methodology for Dynamical Downscaling

Owing to limitations in computational power, the spatial and temporal resolution of GCMs covering the whole globe cannot be refined arbitrarily. For long-term climate change simulations, state-of-the-art GCMs can go down to nominal horizontal resolutions of about 100 km on current supercomputing systems. As atmospheric systems can be resolved only within several grid boxes, their effective resolution is much coarser, however. Therefore, GCMs can simulate large-scale climate features (i.e. synoptic lows), but not mesoscale atmospheric features (e.g. regional winds

generated by mountains), which are necessary for a realistic simulation of regional climate.

Consequently, the principal concept of RCMs is to perform long-term climate change projections with an increased spatial resolution (down to about 50–10 km) for a specific region of interest only. RCMs are limited-area versions of three-dimensional atmospheric circulation models, which in principle use the same set of dynamic equations and physical parameterisations as GCMs. Like GCMs, they include for land grid points a model describing the thermodynamic properties of the upper soil levels. The main difference between RCMs and GCMs (apart from sometimes different parameterisation schemes) is their lateral boundary, as they do not work globally. Because the RCM does not have any information outside its modelling domain, it needs to be provided with information about the atmospheric state at its lateral boundaries, the so-called lateral boundary conditions (LBC). In contrast to GCMs, the solution of an RCM consequently transforms from an initial-value problem into a lateral boundary value problem for longer integration times. The information at the lateral boundaries is taken from the output of the 'driving model', which can be a GCM, a global (re-) analysis, or—when using a 'double nesting' technique —from RCM output simulated on a larger domain in coarser resolution. In order to provide a smooth transition and to avoid numerical problems, a careful LBC treatment is essential for RCM integrations. In the early 1970s, Davies (1976) invented the 'sponge zone', a zone of around 5–10 grid boxes at all lateral boundaries, in which the LBC and the internal solution of the RCM are merged with decreasing weight of the LBC from the boundary towards the centre of the domain. This kind of treatment of the lateral boundaries is still used in most RCM simulations. Additionally, at the lower boundary over sea areas, values for sea-surface temperature (SST) and ice coverage have to be prescribed during the integration. This information is mostly extracted from the driving model like the LBC, as most RCMs are still pure atmospheric models without a coupled ocean component.

10.2.2 Performance of RCMs in Reproducing Recent Climate

A benchmark test for RCMs is that they can reproduce the main features of the climate of the past few decades when forced with realistic boundary conditions. In this respect, it is common to evaluate simulations in which RCMs have been downscaling reanalysis data. Extensive model evaluation has been undertaken for single RCMs (e.g. Samuelsson et al. 2011) or for a large number of models (e.g. Christensen et al. 2010). However, studies on RCM performance focusing on the Baltic Sea region remain few (e.g. Lind and Kjellström 2009). This section therefore presents results from a range of

Fig. 10.1 Simulated mean
temperature bias with respect to
the daily gridded observational
data set based on European
Climate Assessment & Dataset
information (E-OBS) for 1961–
2000. The maps show the
pointwise smallest (*left*), median
(*middle*), and largest (*right*)
biases from an ensemble of nine
RCMs with lateral boundary
conditions from ERA-40
reanalysis data. *Upper row* shows
summer (JJA) biases and *lower
row* winter (DJF)
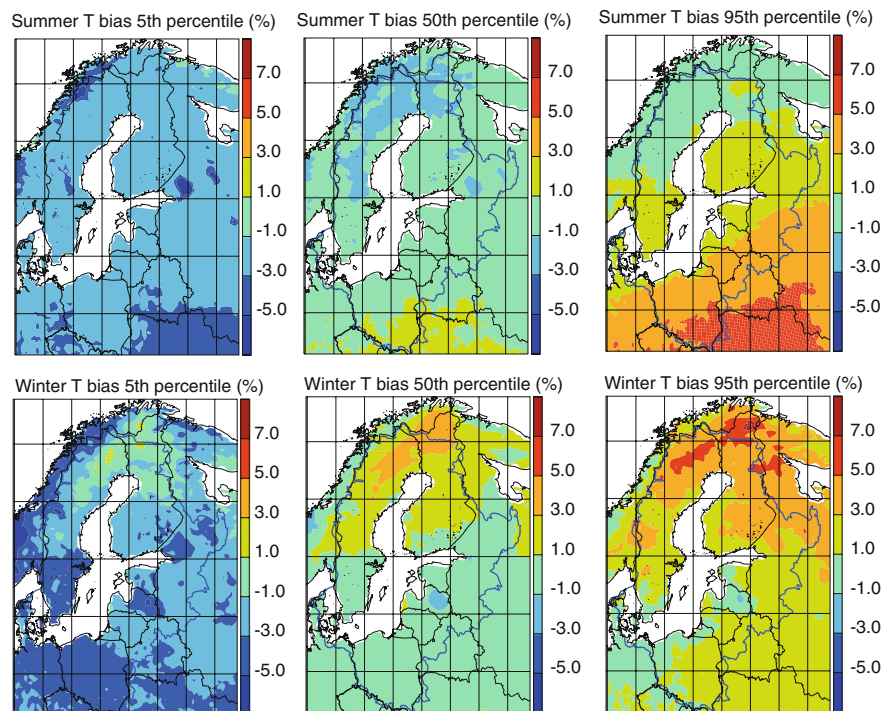
RCMs from the ENSEMBLES project (Christensen et al. 2010) to illustrate the degree to which current RCMs can reproduce the recent past climate. The nine RCMs are listed in Chap. 11 and Table 11.1, and this section presents the results for their forcing by ERA-40 reanalysis data (Uppala et al. 2005) rather than GCM output at their lateral boundaries. As example of model performance, this section shows comparisons of how the ensemble of ERA-40-driven simulations reproduces seasonal mean temperature (Fig. 10.1) and precipitation (Fig. 10.2) for the Baltic Sea region with respect to the daily gridded observational data set based on European Climate Assessment & Dataset information (E-OBS) (Haylock et al. 2008). Nine RCMs were used: C4IRCA3, KNMI-RACMO2, DMI-HIRHAM5, ETHZ-CLM, HadRM3Q0, HadRM3Q16, MPI-REMO, Had-RM3Q3, and SMHIRCA (for documentation on the individual models, see Christensen et al. 2010; data are available from http://ensemblesrt3.dmi.dk/). The maps show grid-point-wise model performance, and as an estimate of the spread, the nine sets of results for each grid point are sorted resulting in an approximate 5th percentile corresponding to the lowest value, a median, and an approximate 95th percentile corresponding to the largest value.

In summer, the temperature climate is reproduced to within ±3 °C in all models in most of the region (Fig. 10.1). An exception is the southernmost part where maximum errors are greater than 5 °C in the warmest (95th percentile) model. Another exception is the relatively large local negative biases found over the big Russian lakes: Lake Ladoga

and Lake Onega. These biases are unlikely to be real but probably reflect that the E-OBS data build on land-based observations, while some of the RCMs include lake models. Such a lake model has the effect of delaying the summertime maximum temperature by about one month, implying that the June–July–August average is lower compared to the surrounding land areas (Samuelsson et al. 2010). In the north, on the other hand, no models overestimate the temperatures indicating a systematic cold bias in most models in that area. In winter, most models tend to be too warm in parts of the northern basin indicating a too weak annual cycle, while in the south, there are both models over- and underestimating temperature. An interesting feature is the local cold bias in eastern Latvia. As Christensen et al. (2010) pointed out, there is no reason why the RCMs should have a local bias like this and it may therefore indicate that it is in fact the E-OBS data that are biased.

Precipitation seems to be overestimated in the Baltic Sea region in most RCMs, both in winter and summer. An exception is again the southern part of the basin where there are models with a dry bias in summer. The driest model is also the model with the largest positive bias in temperature, indicating a possible feedback between precipitation, soil moisture, and temperature. The biases in wintertime precipitation are apparently large, in the wettest models more than 50 % in most of the region. However, it should be noted that the wintertime observations of precipitation in this area may be biased due to undercatch related to snow and wind (e.g. Rubel and Hantel 2001). A local overestimation

Fig. 10.2 Simulated precipitation bias with respect to the daily gridded observational data set based on European Climate Assessment & Dataset information (E-OBS) for 1961–2000. The maps show the pointwise smallest (*left*), median (*middle*), and largest (*right*) biases from an ensemble of nine RCMs with lateral boundary conditions from ERA-40 reanalysis data. *Upper row* shows summer (JJA) biases and *lower row* winter (DJF)
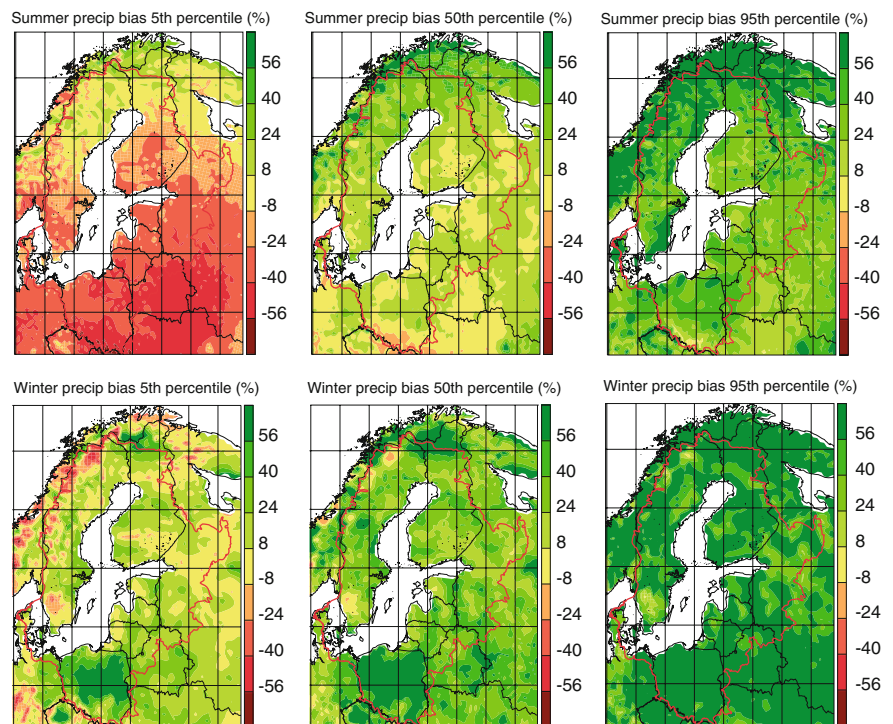
appears over large parts of Poland (Fig. 10.2). But, as discussed by Christensen et al. (2010), this could also be the result of a bias in the observations as there is no reason why the RCMs should show such a strong local deviation from the surrounding areas (see Chap. 11).

10.2.3 Developing and Extending RCMs

In global climate projections, coupled atmosphere–ocean models are state of the art (Meehl et al. 2007). However, RCM climate change projections are in general still carried out for the atmosphere only, prescribing SST data taken from the driving model (Christensen et al. 2007; Kjellström et al. 2013). Consequently, the quality of the prescribed SST/sea ice data depends on the quality of the global modelling

system. In particular, for a relatively small and semi-enclosed sea like the Baltic Sea, data quality might be limited by the coarse resolution of the global ocean component. Figure 10.3 shows the land–sea mask of the global ocean model MPI-OM in grid resolution 1.5° (GR15), as used as one of the main coupled atmosphere–ocean GCMs (AOGCMs) for driving the RCM model suite within the EU project ENSEMBLES (www.ensembles-eu.org). A better representation of the water body of such oceans can be generated by the use of high-resolution regional ocean components, which can be coupled to the atmospheric RCM (analogue to global coupled model systems).

Pioneering work in this area has been done for the Baltic Sea region. This includes establishing atmosphere–ocean–sea ice models (e.g. RCAO, Döscher et al. 2002), some including additional river routing schemes allowing the

Fig. 10.3 Land–sea mask of the global ocean model MPIOM in resolution GR15 (courtesy of M. Böttinger, DKRZ)

modelling of the complete hydrological cycle (e.g. BALTI-MOS, Lehmann et al. 2004). Recently, Meier et al. (2011) showed that a coupled RCM of this type (RCAO) has the potential to improve the results in downscaling experiments driven by GCMs considerably because SSTs and sea ice concentrations are more realistic than those taken directly from the driving GCM. This adds a major caveat to the utility of other downscaling methods relying on SSTs from GCMs in the area, for example the RCM simulations from the ENSEMBLES project.

In conventional RCM simulations, the driving model data are used only in the lateral boundary zone, while in the inner model domain, the RCM is not forced to the driving model. This causes an ill-posed boundary value problem and can lead to a different large-scale flow in the RCM simulation with respect to the driving model. In a perfect boundary setting, this 'freedom' of the RCM may lead to considerable deviations between the real and simulated local climate state (e.g. Winterfeldt and Weisse 2009). With a nudging technique, the solution of the driving model can be prescribed for the whole RCM domain. However, by a scale-independent nudging method, the desired small-scale circulation features generated by the RCM would also be suppressed. In order to circumvent this clear disadvantage, the method of 'spectral nudging' was introduced (von Storch et al. 2000; Feser et al. 2001), in which just the large-scale circulation is relaxed towards the driving model in the inner RCM domain, while the small-scale circulation remains untouched (large-scale constrained RCM simulation). This method leads to a system where empirical data (i.e. large-scale flow and surface details) are systematically combined with the theoretical understanding (i.e. the RCM). While the spectral nudging technique is now becoming more popular (e.g. Miguez-Macho et al. 2004; Castro et al. 2005), a debate on this technique is still ongoing; however, improvements through the application of spectral nudging are evident when the driving model represents a realistic large-scale flow, such as using reanalysis data as the driving model (Winterfeldt and Weisse 2009). In contrast, in the situation with a coarse GCM having an unrealistic large-scale circulation (caused by poorly represented topography due to the coarse resolution) as the driving model, even an RCM using spectral nudging could not alter the prescribed large-scale flow. In regions with complex terrain, the simulated flow in the reanalysis data and therefore the nudging constraints might themselves be biased. In such situations, spectral nudging might lead to unrealistic local-scale flow (Radu et al. 2008). The ability of a non-nudged RCM to improve the large-scale climate inside its domain can be evaluated using the 'Big Brother' approach (Denis et al. 2002). In this approach, a reference climate is established by performing a large-domain high-resolution RCM simulation

termed 'the Big Brother'. Here, the short scales are filtered out and this filtered reference is used to drive the same nested RCM (the 'Little Brother') integrated in the smaller domain but with the same resolution. Differences between climate statistics of both can be attributed to errors associated with the nesting and downscaling technique, allowing them to be distinguished from model errors. This ability is model and region dependent.

At present, most RCMs still use the hydrostatic approximation, assuming the vertical structure to be in hydrostatic equilibrium, and consequently neglecting vertical acceleration. This assumption is valid for nominal horizontal resolutions roughly above ~ 10 km. Most current RCM climate change projections still use coarser nominal horizontal resolutions, between 50 and 20 km (e.g. PRUDENCE and ENSEMBLES), but due to increasing computer power, the resolution of some RCM climate change simulations is increasing to about 10 km. The expected further increase in computational resources will presumably mean a further increase in RCM resolution, leading to the use of non-hydrostatic RCMs. Kendon et al. (2012) reported on a recent study showing results from dynamical downscaling with a non-hydrostatic RCM at 1.5 km grid spacing.

To date, climate change projections have been carried out in a one-way nesting mode, meaning that the RCM does not give information back to the driving model. The first studies of two-way nesting, allowing feedback from the RCM to the GCM (Lorenz and Jacob 2005; Inatsu and Kimoto 2009), indicate the potential for improving the driving global simulation, even in regions far from the two-way nested RCM domain (there are no examples demonstrating the two-way nesting approach for the Baltic Sea region available yet).

In addition to ocean models, lake models have also been coupled to RCMs. This is an important development for the Baltic Sea basin where a large number of lakes exist and a large fraction of the land area is covered by lakes. In a study with an RCM coupled to a lake model, Samuelsson et al. (2010) found that including lakes warmed the climate and that the largest warming occurred in autumn and winter in southern Finland and western Russia where differences of more than 1 °C were obtained.

In recent years, RCMs have begun to incorporate more processes. One example is the work of Wramneby et al. (2010) where a process-based model of vegetation dynamics and biogeochemistry has been coupled to an RCM. They showed that including dynamic vegetation that responds to climate change has an impact on the climate simulated. For the Baltic Sea region in particular, they found reduced albedo resulting from the snow-masking effect of forest expansion when dynamic vegetation is included. This leads to an enhancement of the winter warming trend.

10.3   Statistical Downscaling

Statistical downscaling is an approach that bridges the gap between model output (GCMs or RCMs) and regional or local-scale climate. Rummukainen (1997) distinguished between the model output statistics (MOS) methods and the perfect prognosis (PP) methods. The MOS methods find relationships between model simulations and observations in the historical (reference) period and then use them in future climate simulations (Wilby and Wigley 1997; Maraun et al. 2010). The PP methods identify empirical relationships linking large-scale atmospheric predictors and local/regional predictands. This relationship is assessed on the base of observations in the historical (reference) period and is then used in simulations of future climate.

10.3.1   Model Output Statistics

The biggest disadvantage of the RCM methodology is probably the occurrence of systematic biases in the present climate simulations. These systematic biases, also seen in GCMs, are because dynamic climate simulations carried out with GCMs and RCMs are bound only to changing atmospheric GHG concentrations. Due to their coarse resolution and parameterisations, GCMs and RCMs are not perfect; so even the mean climatological values produced by these models deviate from the corresponding observations. In RCM climate projections, the systematic biases are nonlinear combinations of the systematic errors of the driving GCM and the systematic errors of the RCM itself. Another limitation is that there is still the need to downscale area averages given as grid values in model output to point values necessary for impact studies (Xu et al. 2005). Given the discrepancies between observations and model results for present-day climate, a method is needed to cope with the biases. Given that good observation data sets exist, more realistic data sets of forcing fields incorporating the projected changes can be created and used for impact studies (Piani et al. 2010). This can be achieved through the methods known as MOS. These are statistical models linking simulated variables to observations. There are generally two groups of MOS methods: one is known as the bias correction method (Déqué et al. 2007; Piani et al. 2010), while the other is known as the perturbation of observed data (POD) or the delta change (DC) method (Hay et al. 2000; Lenderink et al. 2007a; van Roosmalen et al. 2011). A review of MOS methods was reported by Maraun et al. (2010).

10.3.1.1   Bias Correction Method

Validating models by comparison with observations makes it possible to quantify model biases, defined as differences in the mean as well as higher order statistical moments. An assessment of bias is the first step before using the model output to force impact models. Unfortunately, model bias is not uniform in space or time and so its identification needs long and homogenised data sets with high spatial resolution. A bias has a seasonal cycle, so its correction often means applying it to individual months or seasons separately. Because GCM and RCM outputs are given in a set of grid points, they are usually volume averages and cannot be directly compared with observations as these are point values. Volume averaging is a type of smoothing that makes high values lower and low values higher, so the range of volume averages is usually much lower than the range of point values. It means that the bias can vary also within different parts of the distribution.

Bias correction or scaling is based on the assumption that the statistical relationship between observations and RCM simulations for the present-day climate is the same as that between the future climate and RCM simulations of the future climate, which may not be true (Christensen et al. 2007; Boberg and Christensen 2012). The bias correction values are calculated by comparing observations with RCM simulations for the same period. Since two climates are compared—the real one and a simulated one—the study period should be relatively long, covering at least 30 years. The corrections can be additive or multiplicative, depending on the variable.

In some cases, the impact models need only seasonal or monthly mean values. Then, it is enough to compare long-term means of observations and RCM simulations for the present-day climate (Schmidli et al. 2006; Graham et al. 2007b). The corrections calculated for RCM simulations under the present-day climate are then applied to the RCM simulations for the future climate to generate more accurate future scenarios. In many cases, the bias correction factors are considered individually for different intensities (i.e. parts of the variable distribution). This is sometimes referred to as distribution-based scaling (DBS; Yang et al. 2010; van Roosmalen et al. 2011). Déqué (2007) and Piani et al. (2010) gave a detailed description of the method. It is generally a quantile mapping approach, where quantiles are empirical cumulative distribution functions or statistical distributions fitted to simulated and observed data.

10.3.1.2   Perturbation of Observed Data

The second MOS method is the DC or POD method (Hay et al. 2000; van Roosmalen et al. 2011). In this approach, the long-term mean additive or multiplicative change factor is calculated on the basis of an RCM projection of the future and present-day (reference) climate and applied to the observation record (Yang et al. 2010; van Roosmalen et al. 2011). These factors can differ seasonally and for different part of the frequency distribution (Olsson et al. 2009). In the DC method, there is no need to identify the bias. Instead, the

absolute or relative delta change factors (DCF) are assessed by comparing the climate model outputs representing present-day and future climate (Semadeni-Davies et al. 2008; Olsson et al. 2009). The observed variable is then rescaled and used as input for impact models.

One of the differences between MOS and other statistical downscaling methods is that MOS calibration is specific to the numerical model for which it has been developed and cannot be used with other numerical models (Maraun et al. 2010). Calibrations can be based on RCM driven by reanalysis or GCM climate simulation forced by external factors. In the first case, there is a direct correspondence between simulated and observed variables; in the second, only statistics of simulated and observed variable distributions can be compared. Because, in this second case, the simulated climate is one randomly selected from many possible choices it is always a risk that the bias determined is an artefact generated by this random choice. On the other hand, calibrating with a reanalysis-driven RCM is of little use, because no reanalysis of the future is available.

10.3.2 The 'Perfect Prognosis' Approach

The PP approaches establish the statistical relationship between large-scale predictors and regional or local-scale predictands. The local variable of interest, denoted by y, depends not only on the large-scale predictors X, but also on the local geographic parameters denoted by g. Mathematically, this can be expressed as follows:

$$y = f(X, g) + \eta,$$

where η means a residual noise term.

Figure 10.4 illustrates how the local conditions depend on both the geography and the large-scale situation. In this case, the snow only stays where the temperature is below freezing, which is only above a certain altitude. Furthermore, the large coherent extent of the snow shows that the local temperature is part of a larger pattern. Although the exact value of y may vary from location to location (small-scale noise η), it is possible to say from this photograph that the temperature in the snow-covered region shown is mainly below freezing. In this example, the large-scale condition X is the snow cover, but it is better to use a predictor with a more direct physical relationship to the predictand. X can often be the mean sea-level pressure (SLP) or the large-scale temperature pattern.

Two steps can be distinguished in the downscaling procedure: the identification of large-scale predictors and the development of a statistical model linking the local predictand with the large-scale predictors.

There are four requirements that the predictors should fulfil. Most important is the existence of a strong statistical relationship between predictors and predictand, typically manifest by high-correlation coefficients. The relationship between predictors and predictand needs to be stable over time. Suitable predictors should also be reasonably well

Fig. 10.4 The Rondane mountain range in Norway during autumn, illustrating how local conditions such as snow cover depend on both geography and large-scale weather (*photograph* R.E. Benestad)

simulated by GCMs. For climate change analysis, it is important that predictors capture the global warming signal (Wilby et al. 2000).

The predictor, being a large-scale variable, is defined at a huge number of grid points. It is therefore convenient to reduce this dimensionality, because there is usually a high correlation between values at neighbouring grid points. One way of doing this is to decompose the field variable into a smaller number of modes of variability. The large-scale variability can be described in terms of orthogonal empirical functions (EOFs) (Lorenz 1956; North et al. 1982; Benestad 2001). The spatial structures of EOFs describe a set of spatially coherent 'modes' that describe the variations in the gridded data. The leading modes describe the structures that are most pronounced and have the greatest spatial scales, and the higher order modes are associated with less variance and smaller spatial scales.

Often, only a small number of leading EOFs describe the major part of field variability (Wilks 1995). It is therefore possible to describe the main features of the gridded data in terms of a relatively small number of EOFs. Each spatial EOF pattern is associated with a vector of weights, describing how strongly this pattern is present at any time of the record. This time series is often referred to as a 'principal component' (PC). The PCs are the basis for the downscaling model calibration, for instance a multiple regression against the predictand. The benefit of using EOFs is that they are orthogonal and make the model calibration easier and more robust (no co-linearity).

The reduction in dimensionality can also be obtained by a transformation of field values into other indices. In the case of SLP fields, these can be the indices of zonal and meridional flow, vorticity, or other indices, such as the North Atlantic Oscillation index (Conway and Jones 1998; Wilby and Wigley 2000). Weather types represent another type of transformation. Here, the large-scale field, usually SLP or geopotential height, is mapped into a set of categories—weather types—by a clustering algorithm like k-means.

The transformation procedure should generate a predictor that has high predictive power, that is, explains a high percentage of the variability of the predictand. Some methods, such as canonical correlation analysis (CCA) or the singular value decomposition (SVD) method, directly seek the modes having the highest correlation or covariance with the predictand field, while others do not.

10.3.2.1 A Brand of Calibration Strategies

The brand 'PP methods' describe a class of empirical–statistical downscaling models that involve a specific strategy for model calibration (Wilks 1995). These use observations [raw and gridded data, or re-analyses (Kalnay et al. 1996; Simmons and Gibson 2000)] to calibrate against an observed

predictand. First, a predictor is taken from historical data, and then, a relation is found with the predictand (downscaling model calibration). Then, the climate model results are compared with the predictors used to calibrate the downscaling model, and steps are taken to ensure that the model results correspond with the calibration data (e.g. through a regression analysis). The PP method may involve linear and nonlinear methods.

10.3.2.2 Regression Methods

Regression models include linear and nonlinear relationships between predictors and the predictand (Benestad et al. 2008). Among them are the multiple regression (Murphy 1999), the CCA method (Busuioc et al. 1999), and the SVD method (Bretherton et al. 1992). The difference between these approaches is that the multiple regression minimises the root-mean-square errors (distance between predictions and observations), the CCA maximises the correlation, and the SVD maximises the covariance between two fields. Artificial neural networks also represent nonlinear regression models (Crane and Hewitson 1998).

10.3.2.3 Weather Classification Methods

The weather classification methods involve various strategies, such as analogues (Zorita and von Storch 1999; Timbal et al. 2008), circulation classification schemes (Bárdossy and Caspary 1990; Jones et al. 1993), cluster analysis (Corte-Real et al. 1999; Huth 2000), and neural nets. The analogue model involves searching the record of past events and taking the day that most closely matches the situation wanted to predict. Cluster analysis bases the predictions on a number of closest states (Wilks 1995), either by taking the mean of the days with close matches or by using the observed values for all days that match the predicted state, and constructing a statistical distribution (histogram). From this sample, or a fitted probability density distribution, a random value may be drawn. Neural nets involve various adaptive learning algorithms, such as 'artificial intelligence' (Wilby et al. 1998; Hewitson and Crane 2002). The analogue model, circulation classification schemes, and cluster analysis all involve a re-sampling of past measurements. These re-sampling techniques suffer from one caveat that the tails of the distributions will be distorted because the sampling cannot produce new record-breaking values (Benestad 2008). Even stationary series are expected to produce new record-breaking events, given sufficiently long intervals for observations. Theory of independent and identically distributed (iid) series shows that the expected occurrence of new record-breaking events will converge towards zero, but never actually become zero. Nevertheless, this implies that the upper and lower tails of the distribution of the results from the re-sampling methods may be distorted and that the

results may have to be re-calibrated. A re-calibration can be performed once the theoretical probability distribution function is known through local quantile mapping.

10.3.2.4 Weather Generators

Stochastic weather generators are statistical models producing high-resolution local-scale time series of a suite of elements such as temperature and precipitation among others, whose large-scale statistics follow the required criteria (Richardson 1981; Wilks and Wilby 1999; Olsson et al. 2009; Willems and Vrac 2011). Among many applications, they can serve as a computationally effective tool to produce site-specific data sets at the required time resolution (Semenov et al. 1998).

The distribution used is usually different for different climate variables. For temperature, the normal distribution is the most popular (Semenov et al. 1998). More complicated is the generation of precipitation data, and different functions are used. Among the most popular are the Markov chain, the semi-empirical, and the Neyman–Scott rectangular pulse (NSRP) weather generator. In the Markov chain generator, precipitation occurrence and totals are produced separately (Sunyer et al. 2012). Two states are possible: wet or dry days. The amount of precipitation on a rainy (wet) day is most often generated using a gamma or exponential distribution (Benestad 2007). In the semi-empirical generator, a few distributions can be defined, for instance for wet and dry spell lengths and precipitation amount. In the NSRP weather generator, Kilsby et al. (2007) proposed four different steps. A storm origin is described by the Poisson process. Separate rain cells within a storm are separated by time intervals taken from exponential distribution. The duration and intensity of each rain cell are also described by exponential distributions, and their sum gives a rainfall total.

Weather generators can be used when the observation records are relatively short. They can also supply many weather 'realisations' having the same overall statistics. A wide suite of statistics can be used to fit the model: mean, variance, skewness, autocorrelation, and many others. Weather generators can also serve to produce data for locations where there is information about the statistical distribution and time structure. For places with only short records of high-temporal-resolution data but longer series with data of low resolution, it is possible to use information from the longer records to make inferences about the distributions, and it is in principle possible to produce projections for temporal scales higher than those usually produced by RCMs (6 h).

10.3.2.5 Randomisation

Models generally underestimate the local-scale variance. To resolve this, Karl et al. (1990) proposed the use of a scaling factor to ensure that the variance of the projected surface values will match the observed variance. But this could increase the error of the estimates, a phenomenon called 'inflation'. Von Storch (1999) argued that this was not a good method because of the need to relate the variance of the predictor to the variance of the predictand. Instead, this author proposed a randomisation method that relied on adding a noise (not necessarily a white one, a random signal with a constant power spectral density was adequate). Another method of resolving the issue of underestimating local-scale variance was developed by Bürger (1996) and called the 'variance-optimised' version of expanded downscaling. Bürger and Chen (2005) compared all these methods. They found that inflation for multi-site downscaling did not describe spatial correlation. Randomisation has a problem with simulating variance in a future climate. The Bürger (1996) method is very sensitive to the quality of normalisation.

10.4 Ensembles, How to Use Them and How to Assess an Error of Projection

All techniques developed to derive regional-scale climate information are associated with uncertainties. This is true both for the direct use of global climate model output and for information emanating from dynamic or statistical downscaling techniques. Uncertainties related to forcing, climate sensitivity, and natural variability can, at least to some degree, be treated by utilising climate change information from ensembles including a large number of climate change experiments (Benestad 2011).

10.4.1 Different Types of Ensembles

Ensembles of climate change simulations can be constructed such that they sample different GCMs with different climate sensitivity under different GHG emission scenarios starting from different initial conditions. Such climate change experiments could be performed by the use of multi-model ensembles (e.g. van der Linden and Mitchell 2009). Under a given forcing scenario, the spread between the different model results can then be taken as an indicator of uncertainty related to structural differences between models, differences in parameterisations, and different initial conditions. In total, there are around 20–30 different coupled AOGCMs worldwide that can constitute such a multi-model ensemble (status as of 2012).

A problem in the context of uncertainty is that different climate models are not totally independent of each other but rather share parts of the code. This means that any multi-model ensemble will contain members that are related to each other. Furthermore, the degree of freedom in a GCM is very large, implying that even if all different GCMs are used,

the full range of model uncertainty will not be sampled by a multi-model ensemble. As an alternative, perturbed physics ensembles with a much larger number of ensemble members have been developed (e.g. Murphy et al. 2007). In these ensembles, one model is used as a reference. In addition to the reference simulation, a large number of simulations with the same model are performed where one or more of the model parameters have been altered within their uncertainty bounds. In this way, the parametric uncertainty can be addressed along with the uncertainty related to initial conditions.

Even if the number of simulations is much larger in a perturbed physics ensemble compared to that in multi-model ensembles, this type of experiment will not sample the structural differences between different GCMs, and therefore, the full model uncertainty is not sampled by perturbed physics ensembles, either. Recently, comparisons have been performed between perturbed physics ensembles based on different GCMs (Yokohata et al. 2010) and between perturbed physics ensembles based on one GCM and multi-model ensembles (Collins et al. 2011). In the ENSEMBLES project (van der Linden and Mitchell 2009), uncertainties due to structural effects as determined from the multi-model CMIP3 GCM ensemble were added to the parametric uncertainties from the HadCM3 perturbed physics ensemble to yield a total uncertainty that could be used in the production of probabilistic climate change projections. In both multi-model ensembles and perturbed physics ensembles, it is not possible to distinguish between uncertainty related to model formulation and that related to initial conditions unless several ensemble members sampling also initial conditions are performed for each multi-model or perturbed physics ensemble member.

10.4.2 Are Ensemble Projections Better Than Those Based on Single Climate Projections?

The multi-model ensemble means have been shown to outperform the single model simulations. This has been shown to result from the fact that models are overconfident, that is, they have a too small spread in the ensemble, centred at the wrong value (Weigel et al. 2008). The good performance of the multi-model ensemble means holds true in a general sense, although for individual variables, seasons, and regions, it is possible to find single models that are better than the ensemble mean. This has been shown in a number of studies at the European scale based on RCMs downscaling reanalysis data in the ENSEMBLES project (e.g. Kjellström et al. 2010; Lenderink 2010; Lorenz and Jacob 2010). This is also illustrated for the Baltic Sea region in Figs. 10.1 and 10.2. There is no reason why the ensemble

mean (or median in this case) should systematically show the smallest biases. For instance, the warmest model is better at reproducing the temperature in the far north in summer and the coldest model is better in the north in winter. Similarly, the driest model appears to outperform the ensemble average in summertime precipitation in the far north. A practical problem here is that different models perform best for different aspects; no one model performs best for everything (e.g. Christensen et al. 2010). This makes it difficult to know which model to choose and favours the use of the multi-model ensemble mean over the results of any single model.

10.4.3 Performance-Based Weighting of Ensembles

Climate models differ in their agreement with observations. The idea of performance-based weighting of ensembles is to utilise these differences to derive weights that can be applied when results from different models are to be combined in a common climate change signal. The rationale would be to give models with a better agreement to observations greater weight than those with less good agreement. However, there are a number of issues. For example, a model can have a good agreement for one variable but not for others, for one season but not for others, and the agreement can be due to compensating errors, etc. Furthermore, any performance-based weights will need to be calculated based on agreement in past decades and so are not necessarily applicable to future climate conditions. Also, regardless of how objective the methods used to derive weights are, there is a high degree of subjectivity as to which metrics to use and what observational data should be used in the analysis (e.g. Christensen et al. 2010).

In the ENSEMBLES project, a weighting system was designed and tested. It consists of a combination of a series of weights derived from evaluating different aspects of RCM performance. These aspects include reproduction of large-scale atmospheric circulation patterns, mesoscale patterns, daily temperature and precipitation distributions and extremes, trends, and the annual cycle (Christensen et al. 2010). Christensen and co-workers found no compelling evidence of an improved description of mean climate states when the weights were used. Furthermore, they concluded that using model weights added another level of uncertainty to the generation of ensemble-based projections. A particular problem related to RCM ensembles was that the underlying GCM simulation largely governed the results. Application of weights that are determined for RCMs in reanalysis-driven simulations (Christensen et al. 2010) on GCM-driven simulations with the same RCMs may therefore not lead to an improvement in the overall ensemble skill (Déqué and Somot 2010).

10.4.4 Design and Use of GCM-RCM Ensemble Regional Climate Projections

Traditionally, climate change ensembles are 'ensembles of opportunity', that is, they are the result of a compilation of more or less coordinated climate change experiments. This means that there have not been any deliberate attempts to design the ensemble so as to sample uncertainty in any specific way. Recently, however, there have been some attempts to design GCM-RCM ensembles in order to sample various kinds of uncertainty in a more systematic way. The PRUDENCE project mainly addressed uncertainty related to RCM formulation with 11 RCMs downscaling one and the same GCM under the same GHG emission scenario, but there were also other GCMs and emission scenarios included in that project (Christensen and Christensen 2007). Based on these results, Déqué et al. (2007) concluded that uncertainty in future European climate change is generally more associated with the choice of GCM than with which RCM is used, particularly for temperature. Consequently, in the ENSEMBLES project, there was an emphasis on having a larger ensemble with more GCMs involved (van der Linden and Mitchell 2009). In a recent study, Déqué et al. (2012) investigated sources of uncertainty in the ENSEMBLES GCM-RCM ensemble. This new study confirmed the results of Déqué et al. (2007) in that the choice of GCM is the dominant source of uncertainty. But there are exceptions, such as for summertime precipitation, when it is RCM formulation that may be the dominant source of uncertainty. Other examples of GCM-RCM ensembles involve ensembles with the Norwegian RCM sampling several GCMs (Haugen and Iversen 2008) or the Swedish RCM sampling a range of different GCMs under different GHG emission scenarios and in some cases with different initial conditions (Kjellström et al. 2011). Based on the results from the ENSEMBLES simulations and the Swedish model, Kendon et al. (2010) also concluded that sampling GCM uncertainty is most important, but RCM uncertainty also needs to be sampled, at least for some regions and seasons.

10.5 Validation Techniques

Any downscaled simulation of present-day climate or a future climate scenario is a more or less simplified representation of reality. A validation against observational data is therefore crucial to assess the quality of the simulation, in particular for a further use in impact studies. To this end, a set of indices is usually derived to describe the properties of interest from the reference data set and the model simulation to be validated. Agreement between the reference and the model is quantified by suitably chosen measures. As discussed in Sect. 10.4, the errors and uncertainties of downscaled climate simulations arise from an imperfect model formulation, uncertain future concentrations of GHGs, and internally generated climate variability. In a downscaling context, the uncertainty due to imperfect model formulation originates from three parts: errors of the driving GCM, errors inherent in the downscaling approach, and errors in observations themselves. The first two types of error are of interest in the validation.

When validating a downscaling system with boundaries from a GCM against observational data, the combined GCM/downscaling error can be evaluated. The influence of the driving GCM on the downscaled simulation can be assessed by combining a single downscaling method with different GCMs and then comparing the different results (e.g. Nikulin et al. 2011). In such a control run setting, care must be taken not to mix the model error and internal climate variability on long timescales. In particular, the estimation of extreme properties requires long time series and the typical 30-year period might not be long enough to gain robust estimates (Kendon et al. 2008). The downscaling error can be separated from the GCM error by driving the downscaling method with 'perfect boundary conditions' (Frei et al. 2003), that is, observational data or—as a proxy—reanalysis data. In a perfect boundary setting, the simulated and reference weather sequences are more or less synchronised, allowing for relatively short validation periods (although care should be taken not to be dominated by individual events). The nesting procedure for RCMs into large-scale low-resolution data at the lateral boundaries of the RCM domain is often supported by a spectral nudging technique that poses additional large-scale constraints onto the largest waves in the interior of the RCM domain (von Storch et al. 2000; Feser et al. 2001). To isolate the error due to nesting in dynamical downscaling in both control run and perfect boundary setting, an approach to separate different error sources in an RCM pseudo-reality, the Big Brother approach (see Sect. 10.2.3) can be used (Denis et al. 2002).

Before using a regional climate projection for follow-up studies, the assessment of not only the downscaling error but also the GCM error is essential, as misrepresentation of large-scale patterns (e.g. the position of the storm tracks) or temporal structure (e.g. blocking frequency and duration) is important practical limitations.

10.5.1 Validation Data

Ultimately, the reliability of any validation depends on the observational data used, either as a reference data set or to provide the forcing in a perfect boundary setting. The typical problems with reference data are inhomogeneities, outliers, and biases (e.g. Jones 1995). Inhomogeneities are systematic

changes in the observational data such as slow creeping trends or jumps in the time series (its mean or other moments) due to changes in the measurement system or the surrounding environment; they might increase uncertainties and induce spurious trends (e.g. Yang et al. 2006; BACC Author Team 2008, Annex 5). Outliers are erroneously high (or low) values, such as caused by multiple-day counts of precipitation measurements; they are particularly detrimental for the estimation of extreme properties but may also affect the validation of other quantities. Biases are caused by systematic peculiarities that lead to a misrepresentation of the local climate by the measurement device, such as wind shadows due to buildings or wind-induced precipitation undercatch. Depending on the property of interest, addressing these issues might be essential for a reliable validation. Another common issue is the availability of long reference data sets, which are needed for robust estimates of the indices of interest, especially for extremes and long-term variability. In particular processes with strong small-scale variability such as precipitation, station data cannot directly be compared with regional climate data, which are considered to represent areal averages instead of point measurements (Chen and Knutson 2008). To overcome this spatial mismatch, gridded data sets have been derived by interpolation and averaging from dense station networks. Prominent examples are the UK Met Office gridded daily precipitation data set (Perry et al. 2009) and the E-OBS daily data set of temperature and precipitation (Haylock et al. 2008) derived from the European Climate Assessment & Dataset database (http://eca.knmi.nl; Klok and Klein Tank 2009) as part of the ENSEMBLES project. Crucial for the usefulness of gridded precipitation data sets is the density of the underlying rain gauge network. For instance, it has been shown that the first version of the E-OBS data set has incorporated too few rain gauges to represent extreme precipitation in some mountain regions (Maraun et al. 2011).

To validate large-scale features, reanalysis data are often taken as reference such as the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) (Kalnay et al. 1996) or the ERA-40 (Uppala et al. 2005) and ERA-Interim (Dee et al. 2011) reanalysis. These are numerical model hindcasts into which observational data have been assimilated. As the output from numerical models, these data are globally complete at the given resolution and provide a sequence of climate states (usually provided every 6 h) consistent within the numerical model. However, due to model biases, the reanalysis data can substantially deviate from reality. Also, reanalysis data do not resolve small-scale features and are therefore not suitable for validation on scales typically relevant for impact studies. Furthermore, it is necessary to be aware whether the observations representing the variable of interest have been assimilated into the model. For instance, precipitation is generally not assimilated into the reanalysis model but fully generated by the model parameterisations; such data are obviously not suitable as reference for validation. Recent projects such as the North American Regional Reanalysis (Mesinger et al. 2006) therefore assimilate further variables such as precipitation. Their completeness and consistency make reanalysis data an ideal candidate to provide boundary conditions for a perfect boundary validation.

10.5.2 Validation Indices

To validate climate simulations, several indices have been proposed, depending on the application of the downscaled product. Comprehensive lists of indices are available from the 'Expert Team on Climate Change Detection and Indices' (Peterson et al. 2001), the STARDEX project (Goodess et al. 2005), and the ENSEMBLES project (van der Linden and Mitchell 2009). Typical validated indices characterise statistics of the variable of interest such as mean, variance, or even the spatial and temporal structure.

The indices to validate the distribution of the variable of interest are statistics such as mean and variance or specific quantiles. For instance, a widely used index for strong but not yet extreme events is the 90th percentile. More generally, the indices can be the parameters of a parameterised formulation of the distribution such as the shape parameter describing the tail behaviour. To obtain results as robust as possible, the representation of extreme events should, if possible, be based on parametric distributions motivated by the extreme value theory, that is, the generalised extreme value (GEV) distribution to validate maxima of long blocks and the generalised Pareto distribution (GPD) to validate excesses of high thresholds (Coles 2001). The spatial indices are, for example, spatial correlations, cluster sizes, and indices describing spatial patterns. The temporal indices are autocorrelation functions, the annual cycle, variability on interannual to decadal timescales, and trends. Other temporal indices describe the length of events such as droughts or wet spells, and the transition probability between different states (e.g. from dry to wet). The corresponding extremal indices (which do not necessarily follow the extreme value theory) would be the maximum length of an event in a defined period, such as a season. To increase confidence in future projections, it is also important to assess the representation of relevant physical processes (e.g. Schär et al. 1999; Lenderink and van Meijgaard 2008; Kendon et al. 2010; Maraun et al. 2011). Of course for every validation procedure, particularly if hypothesis tests and statistical models are involved, the assumptions to be made should be clearly laid out.

An ongoing debate concerns whether the validation should use the data directly with grid box resolution, or whether the data should be smoothed in advance. On the one hand, it is argued that regional climate simulations are not

meant to be interpreted on a grid box level and so the former choice would be too rigid. While on the other, it is a matter of fact that RCM simulations are often used on the grid box level, and a validation should not influence the corresponding performance. Furthermore, in impact studies, the simulated unsmoothed fields are often required even when they are not interpreted on a grid box level. Smoothing might then hide important spatial properties such as the spatial correlation structure.

The validation indices need to be carefully selected. In particular, they need to be independent of calibration or tuning. That is, for PP statistical downscaling and MOS, calibration and validation need to be carried out as a cross-validation on different data sets (e.g. different time periods). Even in cross-validation, the significance of apparently good performance needs to be critically assessed. If the indices are the predictands explicitly modelled in the PP approach or corrected using MOS, they will probably closely resemble the reference indices even in the validation period. Here, good agreement does not necessarily imply a high skill to represent future climate. A similar argument holds for RCMs, as these are in general tuned to properly simulate the observed climate of a specific region. This is discussed in more detail in Sect. 10.5.6.

10.5.3   Validation Measures

To quantify the discrepancy between the modelled and reference validation indices, a range of validation measures has been defined. In some validation studies, the discrepancies have not been quantified at all, but have only been visually inspected. On the other end of the range are statistical tests which explicitly address the significance of the discrepancies. In all cases, deviations should be interpreted carefully. Whereas a visual inspection might overlook important misspecifications, a significance test might as well be misleading (BACC Author Team 2008, Annex 8). Apart from false-positive results, the power of a test might simply be too low to detect model errors due to a lack of data or, in contrast, a significant deviation might simply be completely irrelevant.

The validation in a control run setting is fundamentally different from that in a perfect boundary setting. In a control run setting, the weather sequences between the model simulation and the validation data are independent. The validation can therefore only be based on long-term (climatological) statistics or, more general, distributions. In a perfect boundary setting, the modelled and observed weather sequences are more or less synchronous, given spectral nudging or a small domain and strong lateral forcing. Therefore, in addition to a distribution-wise validation, measures developed for the validation of weather forecasts can be applied for an eventwise validation.

10.5.4   Measures for Distribution-Wise Validation

Simple measures that can be applied to either spatial fields or time series are absolute and relative biases, for example, in mean and standard deviation. Spatial fields can furthermore be validated by their pattern correlation and (root)-mean-squared error relative to the reference pattern, which can be visualised in Taylor diagrams (Taylor 2001). It should be noted, however, that Taylor diagrams do not address the overall biases and provide no confidence intervals. They have been introduced in the AMIP project to synthesise the results of a large number of models in a single diagram. Further insight can be gained by calculating corresponding measures for quantiles or parameters of distributions. For the comparison of the overall distribution, the chi-square test or the Kolmogorov–Smirnov test might be applied (e.g. Semenov et al. 1998; Bachner et al. 2008). The graphical tools for the comparison of distributions are probability (PP) plots and, in particular for extremes, quantile (QQ) plots (e.g. Déqué 2007; Coles 2001). For a list of measures to validate distributions, see Ferro et al. (2005).

10.5.5   Measures for Eventwise Validation

In a perfect boundary setting, a broad range of additional validation measures can be applied. If the modelled and observed time series are synchronous and their phases are expected to match, measures can be applied that have been applied to validate weather forecasts. The same measures that are only applicable to spatial fields in a distribution-wise validation can in this context be applied to validate individual time series. These are, for example, cross-correlations and (centred) root-mean-squared errors, which then, for variables close to normally distributed, can also be visualised by Taylor diagrams. The measures to validate the occurrence of events are the hit rate and the false alarm rate, which are summarised in contingency tables (e.g. Wilks 1995). From these, it is possible to derive frequency biases and odds ratios. Also, continuous variables can be compared using these measures by defining suitable thresholds. Several downscaling approaches predict local-scale probability density distributions rather than specific values; their performance can be validated by probability scores. The classic measure to validate the occurrence of events is the Brier score (Brier 1950). Continuous events (i.e. intensities) can be validated by the continuous ranked probability score (e.g. Jolliffe and Stephenson 2003) and the quantile verification score (e.g. Friederichs and Hense 2007). Absolute score values are often difficult to interpret; therefore, they are usually compared with a reference forecast such as the

climatology or the best-performing method. Such relative measures are skill scores, which can be derived from the aforementioned scores. A comprehensive list of further scores is available from, for example, Wilks (1995) and Jolliffe and Stephenson (2003). As an alternative to simple cross-correlations, one can assess the performance on different timescales using the squared coherence (Brockwell and Davis 1991); for an example, see Maraun et al. (2011). In essence, in this setting, a more rigorous validation is possible, as the capability of a model to simulate the occurrence and magnitude of individual events can be assessed. Of course, this setting does not in general allow for the assessment of GCM errors.

10.5.6 Validation in a Climate Change Context

A high skill of a downscaling method in the current climate does not necessarily imply a high skill in a future climate (e.g. Christensen and Christensen 2007). In PP statistical downscaling, the predictor–predictand relationships might be non-stationary in time, for example, because not all relevant factors controlling the local-scale variable have been included in the model. Also, it is not a priori clear whether the parameterisations of RCMs might capture the changing climate conditions. Finally, biases are not stationary under climate change (e.g. Christensen et al. 2008; Maraun 2012).

To at least partly address these shortcomings, it has been suggested to choose time periods climatically, as different as possible, to calibrate and validate statistical downscaling models (Maraun et al. 2010). This approach is of course limited by the availability of long time series of high quality. For dynamical downscaling, a similar approach is to check whether a RCM performs well in different present-day climates (Christensen et al. 2007). Consensus between different simulations is often seen as a measure of skill. Similarly, a comparison of statistical and dynamical downscaling might provide some insight into the reliability of future simulations. For instance, relationships within statistical downscaling models have been used to validate dynamical climate models (e.g. Busuioc et al. 2001; Maraun et al. 2011). Closely related is the use of RCMs as pseudo-realities to assess the stationarity of predictor–predictand relationships and model biases (e.g. Frias et al. 2006; Vrac et al. 2007; Maraun 2012). The value of model consensus and related concepts is, however, limited as deficiencies might be common to all models. Therefore, understanding the relevant underlying processes and the quality of their representation by the models used is essential to assess the reliability of future climate simulations (Maraun et al. 2010).

10.6 Skill of Downscaling Methods

This section gives a brief overview of the advantages and disadvantages of different downscaling methods. A more detailed discussion can be found in Benestad et al. (2008) and Maraun et al. (2010).

The quality of a downscaling product stands and falls with the ability of the forcing GCM to provide meaningful large-scale boundary conditions. As downscaling aims to correct local-scale misrepresentations due mainly to topographic and small-scale circulation effects, it cannot correct the misrepresentation of the large-scale atmospheric flow. For northern Europe and the Baltic Sea, the most obvious shortcoming of many GCMs is the position, strength, and variability of the main westerly flow. The circulation in many GCMs is too zonal (van Ulden et al. 2007). The large-scale circulation plays a dominant role in the European winter climate (Hurrel and van Loon 1997; Wibig 1999), but also strongly influences summer precipitation in northern Europe (Wibig 1999; Boé et al. 2009). A major shortcoming of the current generation of GCMs is the representation of blocking events (e.g. Palmer et al. 2008; Hinton et al. 2009). Consequently, a large part of the uncertainty in northern and central European temperature and precipitation projections stems from the driving GCM (Déqué et al. 2007).

The main rationale for using dynamical downscaling is that RCMs are based on physical laws. As a consequence, RCMs are in general expected to adequately describe climate change on regional scales. Although the related stationarity issues are more severe for statistical downscaling, it should be noted that parameterisations are developed and tuned for specific climates and might be at least slightly misspecified under future climate conditions. As RCMs calculate the state of the atmosphere regularly in three-dimensional space and in time, output can be generated for a large number of variables at or close to the surface as well as for levels above at temporal frequencies down to the internal computational time step of the respective RCM on a regular grid.

A practical advantage of dynamical downscaling approaches is that they are in principle applicable to any region of the world, whereas statistical downscaling approaches rely on high-quality data for the calibration. As parameterisations must be tuned for different climatic regions, however, RCM simulations for regions without proper validation data should not be taken face value. Although station coverage in the Baltic Sea region is generally very dense (e.g. van Engelen et al. 2008), this problem is not negligible here. Lind and Kjellström (2009) have shown that the observational estimates of precipitation differ to such a high degree that RCM evaluation was affected.

RCMs have been shown to adequately simulate European daily temperature and precipitation intensities, although considerable biases must be expected (e.g. Fig. 10.5; Jacob

Fig. 10.5 A schematic overview of seasonal bias in the PRUDENCE regional models. In each panel, *rows* are the analysis areas, and *columns* correspond to models. *Rows* of panels signify the four seasons, the *left column* of panels is temperature biases (*left colour bar*, °C), while the *right column* of panels signifies precipitation (*right colour bar*, relative change). Areas not covered by a particular model are indicated by *black squares* (after Jacob et al. 2007)
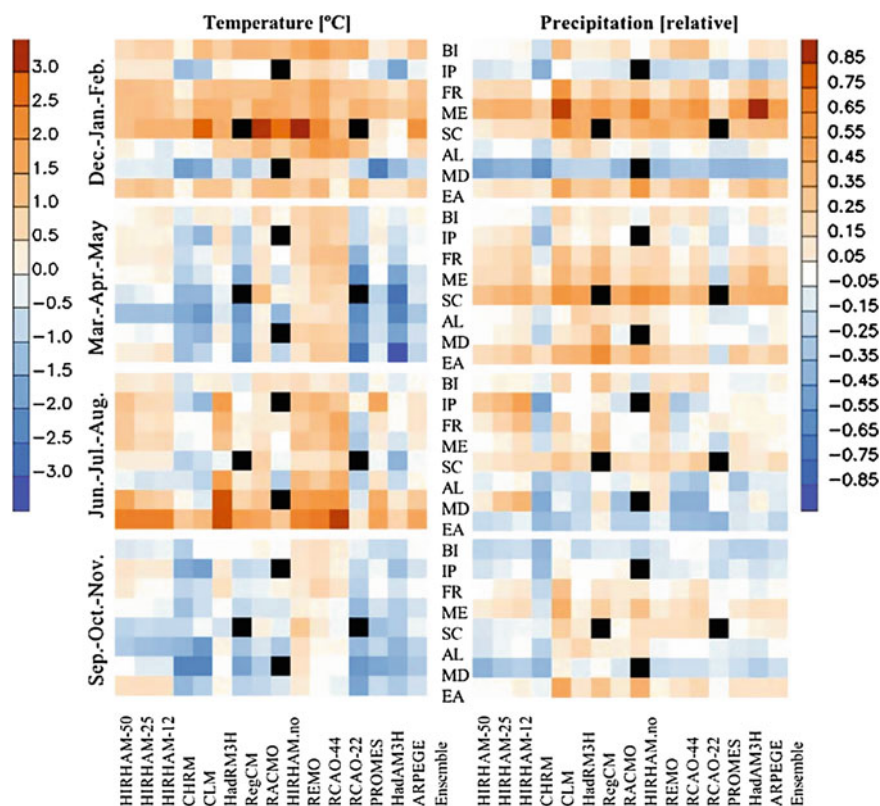
et al. 2007). For instance, in winter, model results tend to be too wet in northern Europe, too warm in summer and winter, and too cold in spring and autumn (Jacob et al. 2007). On the one hand, RCMs generally overestimate the number of wet days; this 'drizzle effect' is partly because RCMs simulate area averages rather than point values. While on the other, RCMs underestimate heavy precipitation events (e.g. Fowler et al. 2007b). Generally, bias is different in different part of the distribution (e.g. Jeong et al. 2011; Fig. 10.6).

A major advantage of RCMs is the simulation of spatially coherent fields. In general, RCMs with a typical resolution of 25 km overestimate the spatial coherence of precipitation events, in particular for convective precipitation. It should be noted that RCMs provide meaningful information only on the scale of a few grid cells (e.g. Fowler and Ekström 2009). In particular, local precipitation is dominated by internal climate variability (Maraun 2012).

As RCMs integrate the equations governing the atmospheric circulation, they in principle provide a coherent picture. However, biases in one variable may propagate into strong biases in dependant variables (e.g. Fig. 10.7); for example, Yang et al. (2010) have shown for Sweden that small temperature biases may, via the nonlinear interaction with precipitation around the melting point, lead to large biases in spring river run-off. Inconsistencies arise in particular where parameterisations come into play. For example, Graham et al. (2007b) have shown for the drainage areas

to the total Baltic Sea basin and Bothnian Sea basin that the partition of precipitation into run-off and evapotranspiration is in general biased towards the latter.

Only a few RCM validation studies consider sub-daily scales, which are particularly relevant for heavy precipitation events. Jeong et al. (2011) have shown that the spatial pattern of the diurnal precipitation cycle in Sweden is reasonably captured by the RCA3 model (Uppala et al. 2005), but the afternoon peak occurs too early and is spatially too uniform. The RACMO2 model accurately simulates the intensity scaling of heavy hourly precipitation with temperature for very intense precipitation, but fails to represent the temperature influence on moderate precipitation intensities beyond 20 °C (Lenderink and van Meijgaard 2008).

In general, increasing model resolution improves model simulations, in particular for precipitation in complex terrain (e.g. Salathé 2003; Hohenegger et al. 2009).

MOS (see Sect. 10.3.1) aims to improve misspecification of dynamical downscaling, although recent work has demonstrated the ability to directly apply MOS to GCMs (Eden et al. 2012). An underlying assumption of MOS is stationarity of the bias. Yet Christensen et al. (2008) inferred a dependence of biases on temperature, indicating potential non-stationarities. In a pseudo-reality, Maraun (2012) found biases in seasonal temperature and precipitation to be relatively stable across Europe, but identified non-stationarities for some regions and seasons; for the Baltic Sea region,

Fig. 10.6 Probability density of precipitation intensity on the eastern coast of Sweden in the warmest months of the year (April–September). *Inner box* highlights the probability of precipitation intensity from 1–3 mm h^{-1}. The 90th, 95th, and 99th percentiles are marked for the observations (*solid vertical lines*) and the model simulation (*dashed vertical lines*) (Jeong et al. 2011)
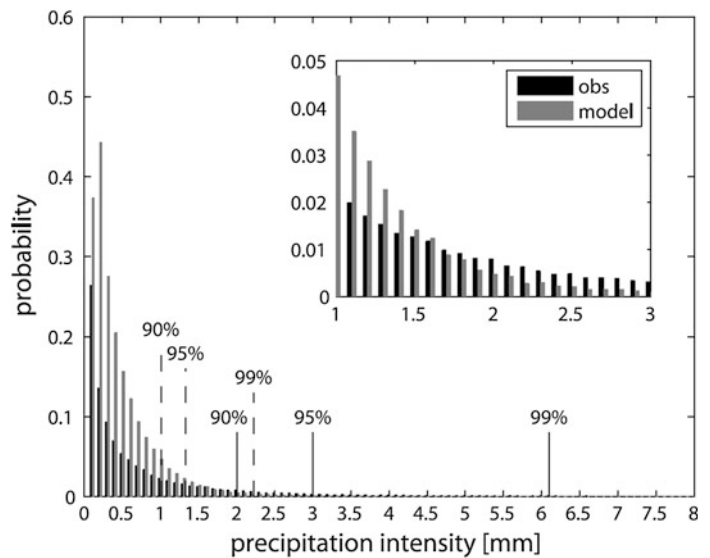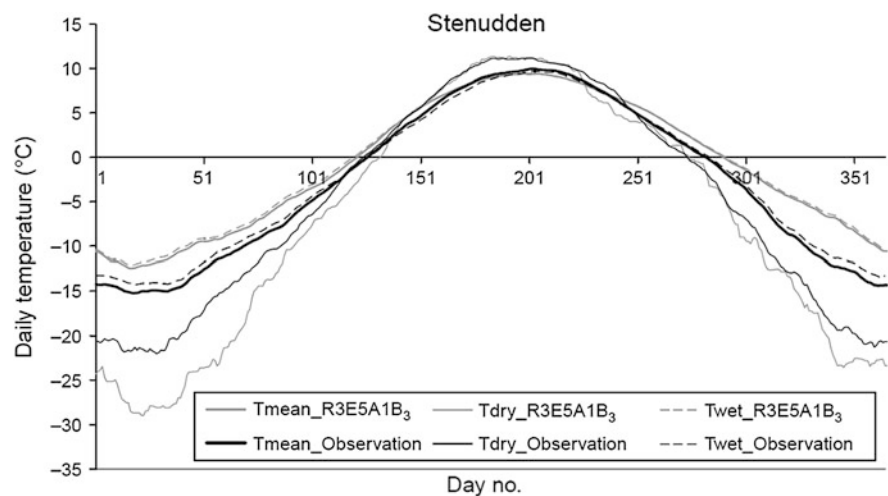

Fig. 10.7 Impact of precipitation on temperature bias for Stenudden, northern Sweden. Annual cycle of daily temperature from observations (1961–1990) and a model (R3E5A1B3) simulation for the same period. *Tmean* mean daily temperature; *Twet* mean daily temperature for days with precipitation; *Tdry* mean daily temperature for dry days (Yang et al. 2010)

temperature biases appear to be non-stationary because of uncertainties in sea ice parameterisations.

MOS has been shown to successfully correct temperature biases as well as biases in precipitation intensities and the number of wet days (e.g. Hay and Clark 2003; Lenderink et al. 2007b; Piani et al. 2010). MOS is particularly suitable to correct orographic effects on precipitation intensity in regions where the topography is misrepresented by the coarse model grid. Furthermore, Widmann et al. (2003) developed a non-local MOS that corrects systematic spatial displacements of precipitation. Yang et al. (2010) applied MOS to improve the correlation between simulated temperature and precipitation.

The good example of the skill of the DBS methodology was presented by Piani et al. (2010). Bias corrections were assessed for the 10-year period 1961–1970 and then applied to the simulated data for the period 1991–2000 and compared with observations from this period. The periods were

chosen to maximise the time lag between them and test whether the bias correction estimated in one period can be applied in the other period with different climatic conditions. The results were surprisingly good (Fig. 10.8). Not only did the mean and higher moments of the scenario data fit well with the observed data, but also indices depending on autocorrelation spectra, such as drought and heavy precipitation, were well projected.

MOS is not capable of correcting the misrepresentation of the temporal structure of a simulated variable; for example, MOS cannot correct errors in the length of dry, hot, or cold spells inherited from GCMs. In particular, the POD method should be considered carefully. Because it only scales observed time series, owing to its construction, it ignores any changes in the atmospheric dynamics which might change the temporal structure of future weather. Nevertheless, applying MOS separately to seasons, individual months or even shorter parts of the year might improve the

Fig. 10.8 Validation of methodology: seasonal mean daily precipitation. Application of bias correction, derived from simulated and observed data for 1961–1970, to model data for 1991–2000. **a** Mean observed daily precipitation for winter (DJF) 1991–2000, **b** as '**a**' but for corrected simulated data, **c** as '**a**' but for uncorrected simulated data, and **d–f** as '**a–c**' but for summer (JJA) (Piani et al. 2010)
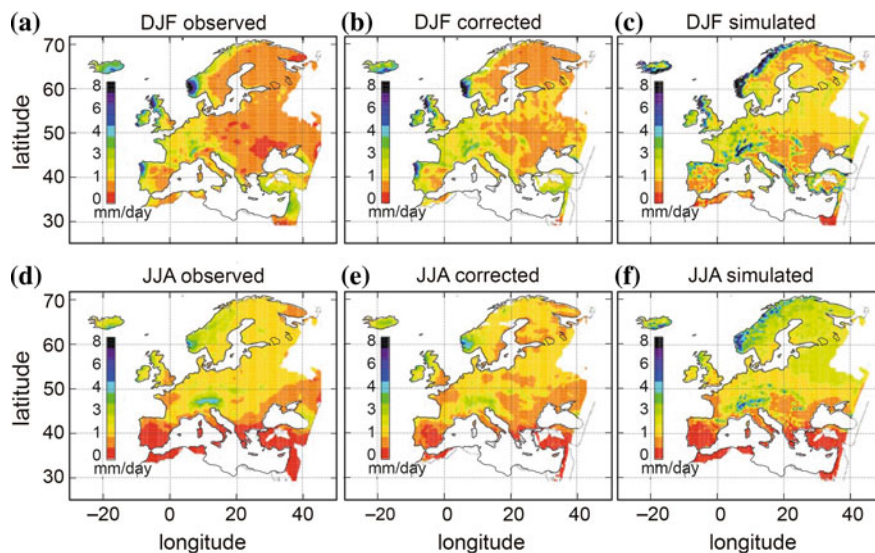
representation of the annual cycle (e.g. Boé et al. 2007; Leander and Buishand 2007).

PP statistical downscaling (see Sect. 10.3.2) is often a computationally cheap alternative to dynamical downscaling. Its main rationale is the explicit use of empirical knowledge by including observational data in statistical models. By construction, the properties which are directly modelled as predictands should be simulated bias free over the calibration period, a characteristic often required by impact modellers. In particular, over complex terrain, PP intrinsically accounts for local effects which might not be captured by the coarse topography and imperfect parameterisations of RCMs. The predictor selection is crucial for the performance of PP approaches; non-stationarities may arise if the predictors do not capture the climate change signal. Predictors should therefore be physically motivated and as close to the underlying processes as possible, see Benestad et al. (2008) for details.

In general, PP methods perform better during winter than summer; Wetterhall et al. (2007) demonstrated this tendency for Sweden.

A shortcoming of many traditional PP methods is the underrepresentation of temporal variability. Most PP methods can be interpreted as some kind of linear or nonlinear, continuous, or categorical regression models. Such models are in general intended to predict the mean of a distribution, disregarding the variability around the mean. As previously discussed, inflation (Karl et al. 1990) or similar approaches do not resolve this problem, but rather create time series with an incorrect temporal structure (von Storch et al. 2000). Instead, many randomisation procedures have been suggested, ranging from generalised linear models (e.g. Chandler 2005) via mixture models (e.g. Vrac and Naveau 2007) to full weather generators. These models provide a realistic

temporal structure, which might be explicitly modelled by Markov processes on short temporal scales, and imposed on longer timescales by large-scale predictors (Wilks and Wilby 1999). Weather generators have also been constructed to simulate sub-daily precipitation (e.g. Cowpertwait et al. 1996; Jones et al. 2009).

A main disadvantage of PP approaches is the handling of spatial coherence. Many PP approaches are used for single locations. Here, the downscaling to local scales is a major advantage over RCMs, which operate on scales of tens of kilometres. The large-scale predictors might impose a coherent spatial structure, which however is often too smooth and can be improved by the addition of a stochastic factor—a weather generator. Randomisation leads to improved local temporal variability, but at the same time might destroy spatial coherence. Therefore, spatial dependence needs to be modelled explicitly by complex multi-site models (e.g. Yang et al. 2005), which provide output at discrete points in space. The development of downscaling methods to full spatial fields for climate change studies is still in its early stages; an example is the Gaussian process-based disaggregation of areal rainfall by Onibon et al. (2004).

A practical advantage of the PP approach is its computational cost; for single sites, it can easily be applied to large ensembles of GCMs, and conditional on one GCM, numerous realisations can be carried out by randomisation.

In general, whether dynamical downscaling or PP is preferable depends on the problem addressed. In many situations, both methods are complementary and should be used in combination. With the availability of large RCM ensembles from the ENSEMBLES project (van der Linden and Mitchell 2009), MOS corrections have become increasingly attractive, attempting to combine the best of both worlds.

10.7 Added Value of Dynamical Downscaling

It is assumed that GCMs are able to provide a reliable description of large-scale weather phenomena and their dynamics. RCMs can resolve mesoscale atmospheric features explicitly and they add small-scale structures to the large-scale circulation provided by the driving model (Feser 2006). Local climate is influenced by large-scale dynamics, regional physiographic features such as local orography, land-sea contrasts land use, and soil type, as well as by small-scale atmospheric features such as frontal systems or convective cells (Lenderink et al. 2007b; Feser et al. 2011). This is particularly the case for the simulation of precipitation. Consequently, the simulated mean precipitation patterns as well as the extreme values are enhanced, especially for complex terrain (e.g. Christensen and Christensen 2001; Feldmann et al. 2008; Suklitsch et al. 2008). For many variables, the explicit treatment of small-scale atmospheric features leads to added value (AV) with respect to the driving model. Assessment of the AV of large-scale constrained versus unconstrained simulations was discussed by Castro et al. (2005) and Rockel et al. (2008). For different realisations of an RCM simulation, generated by small changes in the set-up of the RCM (e.g. domain size/location, initialisation time), substantial variability between the individual realisations is well known (e.g. Ji and Vernekar 1997; Rinke and Dethloff 2000; Weisse et al. 2000), demonstrating the need for ensemble RCM simulations with a large number of realisations. Many publications demonstrate that RCMs are able to realistically simulate climate in comparison with raw or gridded observations or reanalysis. Most of these state the superiority of RCM simulations compared to those from GCMs, but without giving proof.
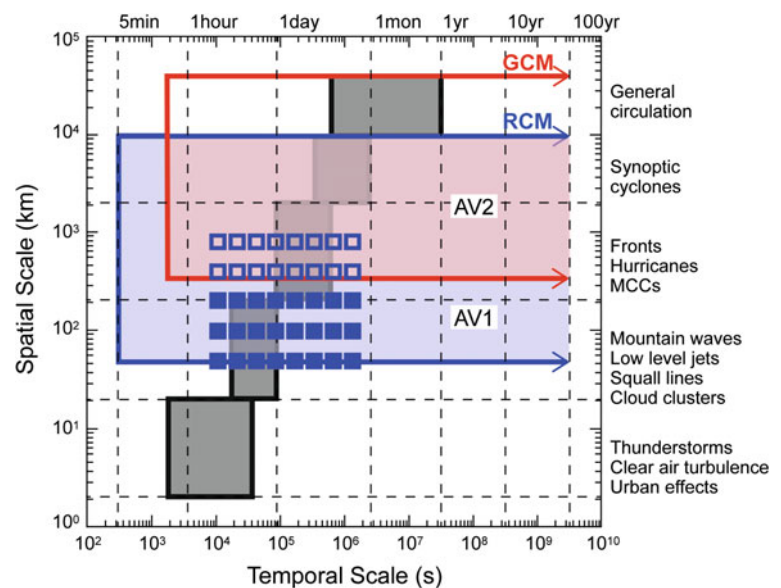
One of the most important purposes of regional climate modelling is increasing knowledge of the real world (Laprise 2005). This additional knowledge is commonly termed 'added value' (Feser 2006). Identification of AV is not an easy task. Small-scale atmospheric fields are usually less energetic than large-scale fields (Laprise 2005), so scale decomposition is sometimes necessary to separate the finer scales.

Di Luca et al. (2012) used a diagram, adapted from Orlanski (1975) and von Storch (2005), to illustrate the concept of AV for the range of scales represented by global and regional models, relative to the characteristic temporal and spatial scales of atmospheric processes (Fig. 10.9). Regional climate modelling is mainly expected to add value at regional dimensions below 300 km and temporal scales less than 30 min, which are absent in GCMs.

Evaluation of a hypothesis of AV implies a comparison of the performance of the RCM with that of the driving GCM (Feser et al. 2011). To date, the number of studies in which the AV of RCMs is directly analysed is limited and only a few concerns the Baltic Sea region.

RCMs could provide AV by adding variability at scales not well resolved by GCMs (at Fig. 10.9 referred to as AV1). However, RCMs can also improve climate simulation at scales resolved by both RCMs and GCMs. This component of AV is referred as AV2 in Fig. 10.9. Because separation of scales is usually made while assessing AV, it is convenient to analyse both types of AV separately (Di Luca et al. 2012). RCMs operate in a limited domain, and two-way interactions between the regional domain and the rest of the globe do not usually occur. In many simulations, the spectral nudging technique ensures that the large scales are not altered too much by the regional model. For all these reasons, AV2 has not been clearly identified and existing analyses are usually limited to AV1.

Fig. 10.9 Characteristic temporal and horizontal spatial scales of atmospheric processes (in *black*) and the range of scales represented in RCMs (*blue line*) and GCMs (*red line*). *Red* and *blue* shaded regions represent the added value of type 1 (AV1) and 2 (AV2), respectively (redrawn from Di Luca et al. 2012)

As RCMs can resolve mesoscale atmospheric features explicitly, they do add small-scale structures to the large-scale circulation provided by the driving model (Feser 2006). This explicit treatment of small-scale atmospheric features leads, for many variables, to an AV with respect to the driving model. This is particularly the case for the simulation of precipitation, which also depends strongly on topography and land–sea contrast, which are better represented at the increased resolution of the RCM. Consequently, the simulated mean precipitation patterns as well as the extreme values are enhanced, especially for complex terrain (e.g. Christensen and Christensen 2001; Feldmann et al. 2008; Suklitsch et al. 2008). For the Baltic Sea region, Walther et al. (2013) demonstrated the improved simulation of the daily precipitation cycle for spring and summer with increasing RCM resolution; an example for a station in central southern Sweden is displayed in Fig. 10.10.

Winterfeldt et al. (2011) analysed AV in dynamically downscaled wind speed fields. They used the Brier skill score (BSS) to detect the AV of the regionally modelled (with spectrally nudged-REMO) wind in comparison with the global reanalysis (NCEP). As seen in Fig. 10.11, the RCM provides AV along the coasts and in narrow bays and straits, in places with complex coastlines or topography. Over open seas and oceans, as well as the interior of Baltic Sea, the BSS is negative, indicating that in these regions, dynamical downscaling does not add value.

Feser (2006) analysed AV in the case of SLP and 2 m air temperature provided by the REMO RCM in comparison with NCEP reanalysis data. Spatial filters were used to separate the data into two domains: that represented best at the large-scale and that represented well by the REMO model (Fig. 10.12). The effect of spectral nudging was also analysed. For SLP, no AV is provided by RCM simulation without nudging. The small improvement was obtained when spectral nudging was applied. For 2 m air temperature,

Fig. 10.10 Estimated diurnal cycle of precipitation amount from observation and the RCA3 regional climate model (RCM developed by the Rossby Centre of SMHI) simulations for four different resolutions for the station 'Malexander' in central southern Sweden (Walther et al. 2013)

significant AV was obtained at both scales when spectral nudging was applied. Without nudging, only the improvement in the scale represented well by RCM was provided (Feser 2006). AV is small in the case of SLP and only for the scale well resolved by RCM, because it is the driving fields (from the GCM) that are the most relevant for SLP. For 2 m air temperature, regional and local factors exert a strong impact on its spatial distribution and the AV of RCMs can be significant (Feser et al. 2011).

Zahn et al. (2008) have shown that RCMs can provide AV in describing mesoscale phenomena such as polar lows. Figure 10.13 presents the SLP and 10 m wind speed fields filtered with a digital bandpass filter that allows better

Fig. 10.11 Brier skill score using QuikSCAT level 2B12 as the source of ground-truth data, global reanalysis (NCEP reanalysis) as the reference forecast, and a regional model (spectrally nudged-REMO) as the forecast, after Winterfeldt et al. (2011)
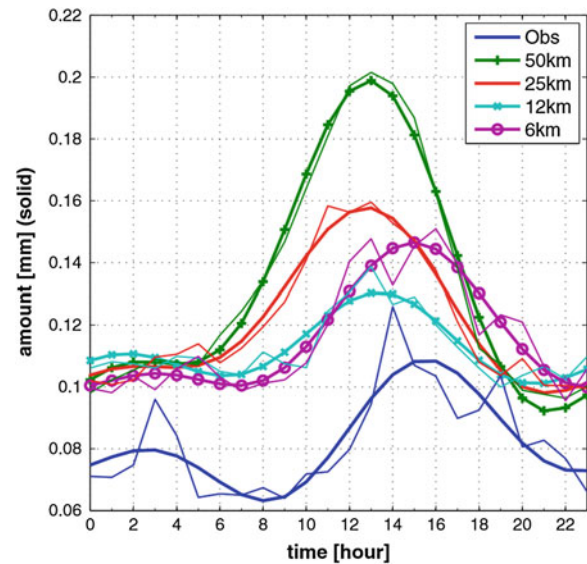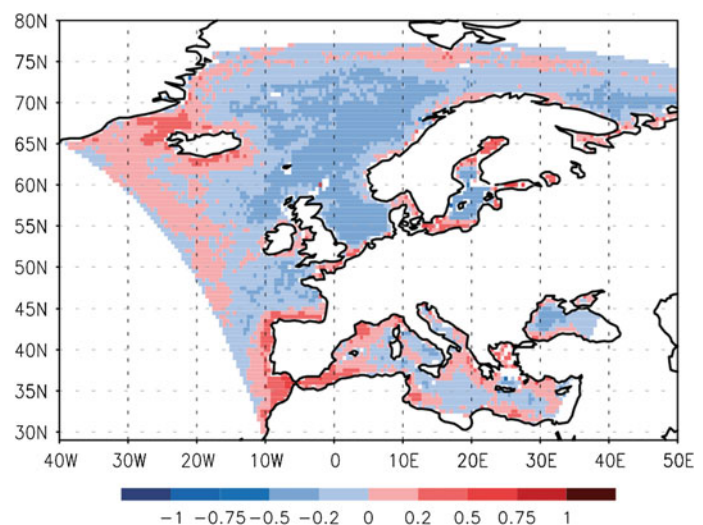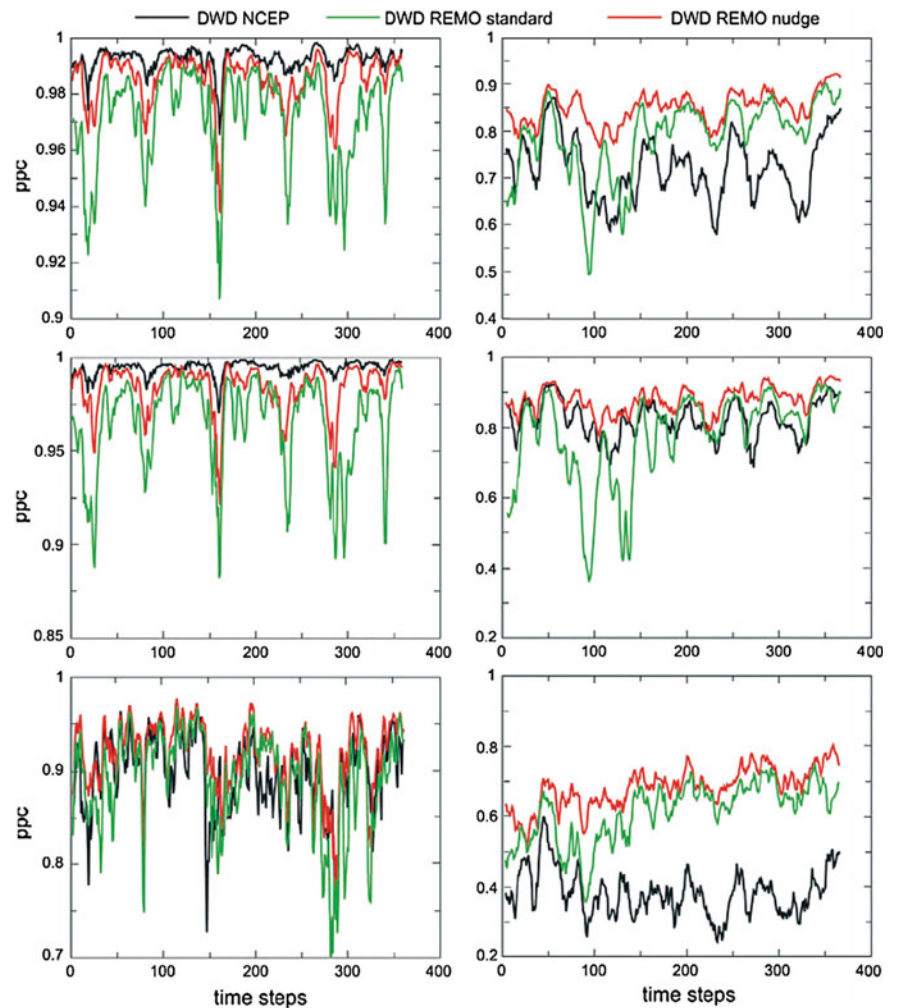
presentation of phenomena at the 200–600 km spatial scale. The community land model (CLM) was able to identify the polar low along the Norwegian coast, although it was still too coarse to describe fine detail.

Di Luca et al. (2012) also introduced the concept of potential AV, as a type of necessary condition for AV. It is a fine spatial scale variability that would be present in regional climate statistics but absent on a coarser grid. The presence of *potential* AV in RCM simulations indicates the possible existence of AV but does not prove it. Di Luca et al. (2012) investigated the existence of potential AV at the temporal scale for different regions and seasons. They showed that for precipitation, potential AV increases for short temporal scales, the summer season and in regions with complex
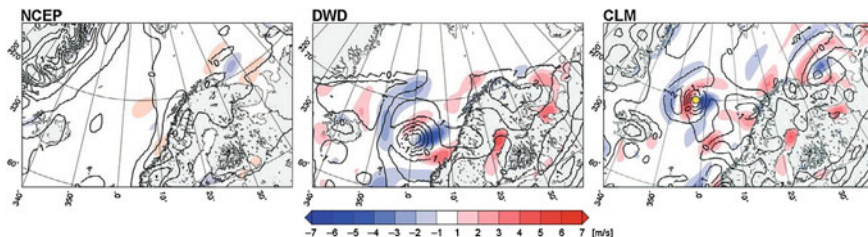

Fig. 10.13 Bandpass-filtered mean sea-level pressure (*isolines* hPa) and 10 m wind speed anomalies (*shaded*) at 0600 UTC 15 October 1993: NCEP analysis, DWD (Deutscher Wetterdienst—German weather service) analysis data, and a simulation by the regional climate model CLM (after Zahn et al. 2008). The position of the pressure minimum of the polar low in the CLM simulation is indicated (*yellow dot*), after Feser et al. (2011)

orography, and decreases when statistics are averaged over long time periods or over a wide spatial domain.

10.8 Downscaling in the Context of Climate Change Impact Studies

The GCMs were not designed for direct application in impact models. Prudhomme et al. (2002) stated that the quality of their output did not allow for direct use in hydrological impact studies, because the spatial and temporal scales were too coarse. Wilby et al. (1999) recommended the use of downscaling techniques before the GCM output data could be used in impact studies. There are many possibilities for downscaling GCM output: direct use of RCM output (Wood et al. 2004), use of bias-corrected RCM output (Wood et al. 2004; Fowler et al. 2007a), statistical downscaling (Wilby et al. 2000; Müller-Wohlfeil et al. 2000), stochastic weather generators (Evans and Schreider 2002) or weather typologies, and/or indices (Pilling and Jones 2002). The skills of different downscaling methods differ considerably between variables and regions.

Hydrologic simulation was found to be sensitive to biases in the spatial distribution of temperature and precipitation at the monthly level, especially where the seasonal snow pack transfers run-off from one season to the next (Fowler et al. 2007c).

Using the example of the Lule River in northern Sweden and two GCMs used to force the same RCM, Graham et al. (2007a) have shown that the choice of driving GCM has a greater impact on results than the choice of GHG emission scenario. The strong impact of the choice of GCM was also emphasised by Widmann et al. (2003), Jasper et al. (2004), Salathé (2005), and Wilby et al. (2006).

Fowler et al. (2007a) stated that at least two variables—temperature and precipitation—had to be downscaled for impact studies in hydrology. In impact models, the physical consistency between variables is very important. To obey this requirement for physical consistency, multi-variate methods should be applied which yield simultaneous correction of relevant variables. This is possible when RCMs are used (Fowler and Kilsby 2007; Fowler et al. 2007a; Graham et al. 2007a, b), but is generally not in statistical downscaling. A multi-site approach should be used when spatial consistency is needed.

10.9 Conclusion

GCMs are a useful tool for studying how climate may change in the future. Such models describe the climate on a set of grid points, regularly distributed in space and time using the same density over land and ocean. Their temporal resolution is relatively high; however, their spatial resolution is low. To simulate regional climate, that is, at a scale smaller than the skilful scale, it is necessary to downscale the GCM results. Downscaling is understood as a process that links large-scale variables with small-scale variables. There are two conceptually different ways of downscaling. One uses RCMs nested in GCMs; RCMs have much higher resolution and can describe local features better and are still able to simulate the atmospheric state in a realistic manner in their skilful scales. The other group of downscaling methods uses empirical and/or statistical relations between the large-scale variables simulated by GCMs and small-scale variables describing regional and/or local climate conditions.

There are many sources of uncertainty in climate model results. These include uncertainty related to limited information on future land use and atmospheric GHG concentrations, limits on the amount of input data and their accuracy, and the chaotic nature of weather. Many sub-grid processes must be represented in models in a simplified form and are not well described by the models. For example, the modelling of cloud formation, the optical and radiative features of clouds, and the creation of atmospheric precipitation still carry considerable model error. The skill of methods for describing regional climate futures is also limited by natural climate variability.

The quality of a downscaling product rests with the ability of the forcing GCM to provide meaningful large-scale boundary conditions, because a large part of the uncertainty in northern and central European temperature and precipitation stems from the driving GCM (Déqué et al. 2007). The main shortcoming of GCMs in Europe is that in many, circulation is too zonal in winter (van Ulden et al. 2007).

RCMs are able to simulate spatially coherent fields, but the parameterisations are developed and tuned for specific climates and might be slightly misspecified under future climate conditions. RCMs have been shown to adequately simulate European daily temperature and precipitation intensities, although considerable biases must be expected (e.g. Jacob et al. 2007). The biases in one variable may propagate into strong biases in dependant variables (e.g. Yang et al. 2010).

Using an ensemble of RCMs is one way of filtering a random error and assessing uncertainty. However, the models within an ensemble are not fully independent because of using shared codes. There is still debate about ensemble design. There have been some recent attempts to design GCM-RCM ensembles in order to sample various kinds of uncertainty in a more systematic way. The uncertainty in future European climate change is generally more associated with the choice of GCM than RCM (Déqué et al. 2007), although for summer precipitation, the RCM formulation may be the dominant source of uncertainty.

Only a few RCM validation studies consider sub-daily scales. Jeong et al. (2011) showed that the diurnal precipitation cycle in Sweden is reasonably well captured by the RCM at SMHI, but that the afternoon peak in precipitation occurs too early and is spatially too uniform. Increasing model resolution will in general improve model simulations, particularly for precipitation in complex terrain (Salathé 2003).

Using coupled AOGCMs is state of the art for global climate projections (Meehl et al. 2007). RCM climate change projections are in general still carried out for the atmosphere only, prescribing SST data from the driving GCM (Christensen et al. 2007). The quality of the prescribed SST/sea ice data depends on the quality of the global modelling system. For a relatively small and semi-enclosed water body like the Baltic Sea, data quality might be limited by the coarse resolution of the global ocean component.

Non-GHG forcings, such as aerosols and land-use change, are not fully represented in RCMs. This can be a source of major uncertainty in projections of future climate as a large part of the simulated multi-decadal variance in North Atlantic SSTs depends on the atmospheric levels of aerosols.

Natural climate variability limits the skill of future climate predictability in many regions (Deser et al. 2012a). In locations where the amplitude of natural variability is high, predictability is low. Conversely, in locations with low natural variability, predictability is higher. The uncertainty of future climate projections is largely a consequence of the chaotic nature of large-scale atmospheric circulation patterns, and improving models or GHG scenarios cannot eliminate this uncertainty (Deser et al. 2012b).

Open Access This chapter is distributed under the terms of the Creative Commons Attribution Noncommercial License, which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

BACC Author Team (2008) Assessment of Climate Change for the Baltic Sea Basin. Regional Climate Series, Springer Verlag, Berlin Heidelberg

Bachner S, Kapala A, Simmer C (2008) Evaluation of daily precipitation characteristics in the CLM and their sensitivity to parameterizations. Meteorol Z 17:407-419

Bárdossy A, Caspary HJ (1990) Detection of climate change in Europe by analyzing European circulation patterns from 1881 to 1889. Theor Appl Climatol 42:155-167

Benestad RE (2001) A comparison between two empirical downscaling strategies. Int J Climatol 21:1645-1668

Benestad RE (2007) Novel methods for inferring future changes in extreme rainfall over northern Europe. Clim Res 34:195-210

Benestad RE (2008) A simple test for changes in statistical distributions. Eos 89:389-390

Benestad RE (2011) A new global set of downscaled temperature scenarios. J Clim 24:2080-2098

Benestad RE, Hanssen-Bauer I, Chen D (2008) Empirical-Statistical Downscaling. World Scientific Publishing Co

Boberg F, Christensen JH (2012) Overestimation of Mediterranean summer temperature projections due to model deficiencies. Nat Clim Change 2:433-436

Boé J, Terray L, Habets F, Martin E (2007) Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. Int J Climatol 27:1643-1655

Boé J, Hall A, Qu X (2009) Current GCMs' unrealistic feedback in the Arctic. J Clim 22:4682-4695

Bretherton CS, Smith C, Wallace JM (1992) An intercomparison of methods for finding coupled patterns in climate data. J Clim 5:541-560

Brier G (1950) The statistical theory of turbulence and the problem of diffusion in the atmosphere. J Atmos Sci 7:283-290

Brockwell PJ, Davis RA (1991) Time series: Theory and methods. Springer-Verlag

Bürger G (1996) Expanded downscaling for generating local weather scenarios. Clim Res 7:111-128

Bürger G, Chen Y (2005) Regression-based downscaling of spatial variability for hydrologic applications. J Hydrol 311:299-317

Busuioc A, von Storch H, Schnur R (1999) Verification of GCM-generated regional seasonal precipitation for current climate and of statistical downscaling estimates under changing climate conditions. J Clim 12:258-272

Busuioc A, Chen D, Hellström C (2001) Performance of statistical downscaling models in GCM validation and regional climate change estimates: application for Swedish precipitation. Int J Climatol 21:557-578

Castro CL, Pielke Sr. RA, Leoncini G (2005) Dynamical downscaling: Assessment of value retained and added using the Regional Atmospheric Modeling System (RAMS). J Geophys Res D 110:1-21

Chandler RE (2005) On the use of generalized linear models for interpreting climate variability. Environmetrics 16:699-715

Chen CT, Knutson T (2008) On the verification and comparison of extreme rainfall indices from climate models. J Clim 21:1605-1621

Christensen JH, Christensen OB (2001) Regional climate scenarios - a study on precipitation. In: Jørgensen AMK, Fenger J, Halsnæs K (eds), Climate Change Research, Danish Contributions. Gads Forlag, Copenhagen, Denmark p 151-166

Christensen JH, Christensen OB (2007) A summary of the PRUDENCE model projections of changes in European climate by the end of the century. Climatic Change 81(Suppl 1):7-30

Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon W-T, Laprise R, Magaña Rueda V, Mearns L, Menéndez CG, Räisänen J, Rinke A, Sarr A, Whetton P (2007) Regional climate projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom

Christensen JH, Boberg F, Christensen OB, Lucas-Picher P (2008) On the need for bias correction of regional climate change projections of temperature and precipitation. Geophys Res Let 35:L.20709. doi:10.1029/2008GL035694

Christensen JH, Kjellström E, Giorgi F, Lenderink G, Rummukainen M (2010) Weight assignment in regional climate models. Clim Res 44:179-194

Coles S (2001) An Introduction to Statistical Modeling of Extreme Values. Springer

Collins M, Booth BB, Bhaskaran B, Harris GR, Murphy,JM, Sexton DMH, Webb MJ (2011) Climate model errors, feedbacks and

forcings: a comparison of perturbed physics and multi-model ensembles. Clim Dynam 36:1737-1766

Conway D, Jones PD (1998) The use of weather types and air flow indices for GCM downscaling. J Hydrol 212-213:348-361

Corte-Real J, Qian B, Xu H (1999) Circulation patterns, daily precipitation in Portugal and implications for climate change simulated by the second Hadley Centre GCM. Clim Dynam 15:921-935

Cowpertwait P, O'Connell P, Metcalfe A, Mawdsley J (1996) Stochastic point process modelling of rainfall. I. Single-site fitting and validation. J Hydrol 175:17-46

Crane RG, Hewitson BC (1998) Doubled CO_2 precipitation changes for the Susquehanna basin: down-scaling from GENESIS general circulation model. Int J Climatol 18:65-76

Davies HC (1976) A lateral boundary formulation for multi-level prediction models. Q J Roy Meteorol Soc 102:405-418

Dee DP, Uppala SM, Simmons AJ et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J Roy Meteorol Soc 137:553-597

Denis B, Laprise R, Caya D, Côté J (2002) Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment. Clim Dynam 18:627-646

Déqué M (2007) Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values. Global Planet Change 57:16-26

Déqué M, Somot S (2010) Weighted frequency distributions express modelling uncertainties in the ENSEMBLES regional climate experiments. Clim Res 44:195-209

Déqué M, Rowell DP, Lüthi D, Giorgi F, Christensen JH, Rockel B, Jacob D, Kjelström E, De Castro M, Van Den Hurk B (2007) An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. Climatic Change 81 (Suppl 1):53-70

Déqué M, Somot S, Sanchez-Gomez E, Goodess CM, Jacob D, Lenderink G, Christensen OB (2012) The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability. Clim Dynam 38:951-964

Deser C, Knutti R, Solomon S, Phillips AS (2012a) Communication of the role of natural variability in future North American climate. Nat Clim Change 2:775-779

Deser C, Phillips A, Bourdette V, Teng H (2012b) Uncertainty in climate change projections: the role of internal variability. Clim Dynam 38:527-547.

Di Luca A, de Elia R, Laprise R (2012) Potential for added value in precipitation simulated by high-resolution nested Regional Climate Models and observations. Clim Dynam 38:1229-1247

Dickinson RE, Errico RM, Giorgi F, Bates GT (1989) A regional climate model for western United States. Climatic Change 15:383-422

Döscher R, Willen U, Jones C, Rutgersson A, Meier HEM, Hansson U (2002) The development of the coupled regional ocean-atmosphere model RCAO. Boreal Env Res 7:183-192

Eden JM, Widmann M, Grawe D, Rast S (2012) Skill and correction of GCM-simulated precipitation. J Clim 25:3970-3984

Evans J, Schreider S (2002) Hydrological impacts of climate change on inflows to Perth, Australia. Clim Change 55:361-393

Feldmann H, Früh B, Schädler G, Panitz H-J, Keuler K, Jacob D, Lorenz P (2008). Evaluation of the precipitation for south-western Germany from high resolution simulations with regional climate models. Meteorol Z 17:455-465

Ferro CAT, Hannachi A, Stephenson DB (2005) Simple nonparametric techniques for exploring changing probability distributions of weather. J Clim 18:4344-4354

Feser F (2006) Enhanced detectability of added value in limited-area model results separated into different spatial scales. Mon Weather Rev 134:2180-2190

Feser F, Weisse R, von Storch H (2001) Multi-decadal atmospheric modeling for Europe yields multi-purpose data. EOS Trans AGU 82:305-310

Feser F, Rockel B, von Storch H, Winterfeldt J, Zahn M (2011). Regional climate models add value to global model data: A review and selected examples. Bull Am Meteorol Soc 92:1181-1192

Foley AM (2010) Uncertainty in regional climate modelling: A review. Prog Phys Geogr 34:647-670

Fowler HJ, Ekström M (2009) Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes. Int J Climatol 29:385-416

Fowler HJ, Kilsby CG (2007) Using regional climate model data to simulate historical and future river flows in northwest England. Climatic Change 80:337-367

Fowler HJ, Blenkinsop S, Tebaldi C (2007a) Review. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. Int J Climatol 27:1547-1578

Fowler HJ, Ekström M, Blenkinsop S, Smith AP (2007b) Estimating change in extreme European precipitation using a multi-model ensemble. J Geophys Res D 112:D18104. doi: 10.1029/2007JD008619

Fowler HJ, Kilsby CG, Stunell J (2007c) Modelling the impacts of projected future climate change on water resources in northwest England. Hydrol Earth Syst Sci 11:1115-1126

Frei C, Christensen JH, Déqué M, Jacob D, Jones RG, Vidale PL (2003) Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. J Geophys Res D 108:4124. doi:10.1029/2002JD002287

Frias M, Zorita E, Fernández J, Rodriguez-Puebla C (2006) Testing statistical downscaling methods in simulated climates. Geophys Res Lett 33:L19807. doi: 10.1029/2006GL027453

Friederichs P, Hense A (2007) Statistical downscaling of extreme precipitation events using censored quantile regression. Mon Weather Rev 135:2365-2378

Giorgi F (2006) Regional climate modeling: Status and perspectives. J Phys IV, 139:101-118

Giorgi F, Bates GT (1989) The climatological skill of a regional model over complex terrain. Mon Weather Rev 117:2325-2347

Goodess CM, Anagnostopoulou C, Bárdossy A, Frei C, Harpham C, Haylock ML, Hundecha Y, Maheras P, Ribalaygua J, Schmidli J, Schmith T, Tolika K, Tomozieu F, Wilby RL (2005) An intercomparison of statistical downscaling methods for Europe and European regions assessing their performance with respect to extreme weather events and the implications for climate change applications. CRU Research Report 11

Graham LP, Andréasson J, Carlsson B (2007a) Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods - A case study on the Lule River basin. Climatic Change 81(Suppl 1):293-307

Graham LP, Hagemann S, Jaun S, Beniston M (2007b) On interpreting hydrological change from regional climate models. Climatic Change 81(Suppl 1):97-122

Grotch S, MacCracken M (1991) The use of general circulation models to predict regional climate change. J Clim 4:286-303

Haugen JE, Iversen T (2008) Response in extremes of daily precipitation and wind from a downscaled multimodel ensemble of anthropogenic global climate change scenarios. Tellus A 60:411-426

Hay LE, Clark MP (2003) Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. J Hydrol 282:56-75

Hay LE, Wilby RL, Leavesley GH (2000) A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States. J Am Water Resour Assoc 36:387-397

Haylock M, Hofstra N, Klein Tank AMG, Klok E, Jones PD, New M (2008) An European daily high-resolution gridded dataset of surface temperature and precipitation for 1950-2006. J Geophys Res 113: D20119. doi:10.1029/2008JD010201

Hewitson BC, Crane RG (2002) Self-organizing maps: applications to synoptic climatology. Clim Res 22:13-26

Hinton T, Hoskins B, Martin G (2009) The influence of tropical sea surface temperatures and precipitation on north Pacific atmospheric blocking. Clim Dynam 33:549-563

Hohenegger C, Brockhaus P, Bretherton CS, Schär C (2009) The soil moisture-precipitation feedback in simulations with explicit and parameterized convection. J Clim 22:5003-5020

Hurrel JW, van Loon H (1997) Decadal variation in climate associated with the North Atlantic Oscillation. Climatic Change 36:301-326

Huth R (2000) A circulation classification scheme applicable in GCM studies. Theor Appl Clim 67:1-18

Inatsu M, Kimoto M (2009) A scale interaction study on East Asian cyclogenesis using a General Circulation Model coupled with an interactively nested regional model. Mon Weather Rev 137:2851-2868

Jacob D, Bärring L, Christensen OB, Christensen JH, de Castro M, Déqué M, Giorgi F, Hagemann S, Hirschi M, Jones R, Kjellström E, Lenderink G, Rockel B, Sánchez E, Schär C, Seneviratne SI, Somot S, van Ulden A, van den Hurk B (2007) An intercomparison of regional climate models for Europe: model performance in present-day climate. Climatic Change 81:31-52

Jasper K, Calanca P, Gyalistras D, Fuhrer J (2004) Differential impacts of climate change on the hydrology of two alpine river basins. Clim Res 26:113-129

Jeong J-H, Walther A, Nikulin G, Chen D, Jones C (2011) Diurnal cycle of precipitation amount and frequency in Sweden: Observation versus model simulation. Tellus A 64:664-674

Ji Y, Vernekar AD (1997) Simulation of the Asian summer monsoons of 1987 and 1988 with a regional model nested in a global GCM. J Clim 10:1965-1977

Jolliffe IT, Stephenson DB (eds) (2003) Forecast Verification, A Practitioner's Guide in Atmospheric Science. Wiley

Jones PD (1995) Land surface temperatures - is the network good enough? Climatic Change 31:535-548

Jones PD, Hulme M, Briffa KR (1993) A comparison of Lamb circulation types with an objective classification scheme. Int J Climatol 13:655-663

Jones PD, Harpham C, Kilsby C, Glenis V, Burton A (2009) Projections of future daily climate for the UK from the Weather Generator, Tech. Rep. 1, UK Met Office

Kalnay E, Kanamitsu M, Kistler R et al (1996) The NCEP/NCAR 40-Year Reanalysis Project. Bull Am Met Soc 77:437-471

Karl TR, Wang WC, Schlesinger ME, Knight RW, Portman D (1990) A method of relating general circulation model simulated climate to observed local climate. Part I: seasonal statistics. J Clim 3:1053-1079

Kendon EJ, Rowell DP, Jones RG, Buonomo E (2008) Robustness of future changes in local precipitation extremes. J Clim 21:4280-4297

Kendon EJ, Jones R, Kjellström E, Murphy J (2010) Using and designing GCM-RCM ensemble regional climate projections. J Clim 23:6485-6503

Kendon EJ, Roberts NM, Senior CA, Roberts MJ (2012) Realism of rainfall in a very high resolution regional climate model. J Clim 25:5791-5806

Kilsby CG, Jones PD, Burton A, Ford AC, Fowler HJ, Harpham C, James P, Smith A, Wilby RL (2007) A daily weather generator for use in climate change studies. Environ Model Software 22:1705-1719

Kjellström E, Boberg F, Castro M, Christensen JH, Nikulin G, Sanchez E (2010) On the use of daily and monthly temperature and precipitation statistics as a performance indicator for regional climate models. Clim Res 44:135-150

Kjellström E, Nikulin G, Hansson U, Strandberg G, Ullerstig A (2011) 21st century changes in the European climate: uncertainties derived from an ensemble of regional climate model simulations. Tellus A 63:24-40

Kjellström E, Thejll P, Rummukainen M, Christensen JH, Boberg F, Christensen OB, Maule CF (2013) Emerging regional climate change signals for Europe under varying large-scale circulation conditions. Clim Res 56:103-109

Klok EJ, Klein Tank AMG (2009) Updated and extended European dataset of daily climate observations. Int J Climatol 29:1182-1191

Laprise R (2005) High-resolution climate modelling: Assessment, added value and applications, A Foreword. In: Bärring L, Laprise R (eds), High-resolution climate modelling: assessment, added value and applications. Lund University electronic reports in physical geography No. 5, Lund, Sweden

Leander R, Buishand TA (2007) Resampling of regional climate model output for the simulation of extreme river flows. J Hydrol 332:487-496

Lehmann A, Jacob D, Lorenz P (2004) Modelling the exceptional Baltic Sea inflow events in 2002–2003. Geophy Res Lett 31: L21308. doi:10.1029/2004GL020830

Lenderink G (2010) Exploring metrics of extreme daily precipitation in a large ensemble of regional climate model simulations. Clim Res 44:151-166

Lenderink G, van Meijgaard E (2008) Increase in hourly precipitation extremes beyond expectations from temperature changes. Nat Geosci 1:511-514

Lenderink G, Buishand A, van Deursen W (2007a) Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. Hydrol Earth Syst Sci 11:1145-1159

Lenderink G, van Ulden A, van den Hurk B, Keller F (2007b) A study on combining global and regional climate model results for generating climate scenarios of temperature and precipitation for the Netherlands. Clim Dynam 29:157-176

Lind P, Kjellström E (2009) Water budget in the Baltic Sea drainage basin: Evaluation of simulated fluxes in a regional climate model. Boreal Env Res 14:56-67

Lorenz EN (1956) Empirical orthogonal functions and statistical weather prediction department of meteorology. Massachusetts Institute of Technology, Department of Meteorology, Issue 1 of Science Report

Lorenz E (1963) Deterministic nonperiodic flow. J Atm Sci 20:130-141

Lorenz P, Jacob D (2005) Influence of regional scale information on the global circulation: A two-way nesting climate simulation. Geophys Res Lett 32:L18706. doi: 10.1029/2005GL023351

Lorenz P, Jacob D (2010) Validation of temperature trends in the ENSEMBLES regional climate model runs driven by ERA40. Clim Res 44:167-177

Maraun D (2012) Nonstationarities of regional climate model biases in European seasonal mean temperature and precipitation sums. Geophys Res Lett 39:L06706. doi: 10.1029/2012GL051210

Maraun D, Wetterhall F, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themeβl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich I (2010) Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. Rev Geophys 48:RG3003. doi:10.1029/2009RG000314

Maraun D, Osborn TJ, Rust HW (2011) The influence of synoptic airflow on UK daily precipitation extremes. Part I: observed spatio-temporal relations. Clim Dynam 36:261-275

Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gaye AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z-C (2007) Global climate projections. In: Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M Miller HL (eds), Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK

Meier HEM, Höglund A, Döscher R, Andersson H, Löptien U, Kjellström E (2011) Quality assessment of atmospheric surface fields over the Baltic Sea from an ensemble of regional climate model simulations with respect to ocean dynamics. Oceanologica 53:193-227

Mesinger F, DiMego G, Kalnay E, Mitchell K, Shafran PC, Ebisuzaki W, Jovic D, Woollen J, Rogers E, Berbery EH, Ek MB, Fan Y, Grumbine R, Higgins W, Li H, Lin Y, Manikin G, Parrish D, Shi W (2006) North American regional reanalysis. Bull Am Met Soc 87:343-360

Miguez-Macho G, Stenchikov GL, Robock A (2004) Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations. J Geophys Res 109:D13104. doi: 10.1029/2003JD004495

Müller-Wohlfeil DI, Bürger G, Lahmer W (2000) Response of a river catchment to climate change: application of expanded downscaling to northern Germany. Climatic Change 47:61-89

Murphy JM (1999) An evaluation of statistical and dynamical techniques for downscaling local climate. J Clim 12:2256-2284

Murphy JM, Booth BBB, Collins M, Harris GR, Sexton DMH, Webb MJ (2007) A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles. Phil Trans Roy Soc Lond A 365:1993-2028

Nikulin G, Kjellström E, Hansson U, Strandberg G, Ullerstig A (2011) Evaluation and future projections of temperature, precipitation and wind extremes over Europe in an ensemble of regional climate simulations. Tellus A 63:41-55

North GR, Bell TL, Cahalan RF (1982) Sampling errors in the estimation of empirical orthogonal functions. Mon Weather Rev 110:699-706

Olsson J, Berggren K, Olofsson M, Viklander M (2009) Applying climate model precipitation scenarios for urban hydrological assessment: A case study in Kalmar City, Sweden. Atmos Res 92:364-375

Onibon H, Lebel T, Afouda A, Guillot G (2004) Gibbs sampling for conditional spatial disaggregation of rain fields. Water Resour Res 40:W08401

Orlanski I (1975) A rational subdivision of scales for atmospheric processes. Bull Am Met Soc 56:527-530

Palmer TN, Doblas-Reyes FJ, Weisheimer A, Rodwell MJ (2008) Toward seamless prediction. Calibration of climate change projections using seasonal forecasts. Bull Am Met Soc 89:459-470

Perry M, Hollis D, Elms M (2009) The generation of daily gridded datasets of temperature and rainfall for the UK. Climate Memorandum No.24. National Climate Information Centre, Met Office, Exeter, UK

Peterson TC, Folland C, Gruza G, Hogg W, Mokssit A, Plummer N (2001) Report on the Activities of the Working Group on Climate Change Detection and Related Rapporteurs 1998-2001, WMO-TD 1071. World Meteorological Organisation Rep. WCDMP-47, Geneva, Switzerland

Piani C, Haerter JO, Coppola E (2010) Statistical bias correction for daily precipitation in regional climate models over Europe. Theor Appl Climatol 99:187-192

Pilling CG, Jones JAA (2002) The impact of future climate change on seasonal discharge, hydrological processes and extreme flows in the Upper Wye experimental catchment, mid-Wales. Hydrolog Process 16:1201-1213

Prudhomme C, Reynard N, Crooks S (2002) Downscaling of global climate models for flood frequency analysis: Where are we now? Hydrolog Process 16:1137-1150

Radu R, Déqué M, Somot S (2008) Spectral nudging in a spectral regional climate model. Tellus A 60:2461-2481

Richardson CW (1981) Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour Res 17:182-190

Rinke A, Dethloff K (2000) On the sensitivity of a regional Arctic climate model to initial and boundary conditions. Clim Res 14:101-113

Rockel B, Castro CL, Pielke Sr. RA, von Storch H, Leoncini G (2008) Dynamical downscaling: Assessment of model system dependent retained and added variability for two different regional climate models. J Geophys Res A 113:D21107. doi:10.1029/2007JD009461

Rubel F, Hantel M (2001) BALTEX 1/6-degree daily precipitation climatology. Meteorol Atmos Phys 77:155-166

Rummukainen M (1997) Methods of statistical downscaling of GCM simulations. Reports Meteorology and Climatology 80, Swedish Meteorological and Hydrological Institute, SE-601 76 Norrköpping, Sweden

Rummukainen M (2010) State-of-the-art with regional climate models. WIREs Clim Change 1:82-96

Salathé EP (2003) Comparison of various precipitation downscaling methods for the simulation of streamflow in a rain shadow river basin. Int J Climatol 23:887-901

Salathé EP (2005) Downscaling simulations of future global climate with application to hydrologic modelling. Int J Climatol 25:419-436

Samuelsson P, Kourzeneva E, Mironov D (2010) The impact of lakes on the European climate as simulated by a regional climate model. Boreal Env Res 15:113-129

Samuelsson P, Jones CG, Willén U, Ullerstig A, Gollvik S, Hansson U, Jansson C, Kjellström E, Nikulin G, Weser K (2011) The Rossby Centre Regional Climate Model RCA3: Model description and performance. Tellus A 63:4-23

Schär C, Luthi D, Beyerle U, Heise E (1999) The soil-precipitation feedback: A process study with a regional climate model. J Clim 12:722-741

Schmidli J, Frei C, Vidale PL (2006) Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods. Int J Climatol 26:679-689

Semadeni-Davies A, Hernebring C, Svensson G, Gustafsson L-G (2008) The impact of climate change and urbanization on drainage in Helsingborg, Sweden: Combined sewer system. J Hydrol 350:100-113

Semenov MA, Brooks RJ, Barrow EM, Richardson CW (1998) Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. Clim Res 10:95-107

Simmons AJ, Gibson JK (2000) The ERA-40 Project Plan. ECMWF, ERA-40 Project Report Series, #1, www.ecmwf.int

Suklitsch M, Gobiet A, Leuprecht A, Frei C (2008) High resolution sensitivity studies with the regional climate model CCLM in the Alpine region. Meteorol Z 17:467-476

Sunyer MA, Madsen H, Ang PH (2012) A comparison of different regional climate models and statistical downscaling methods for extreme rainfall estimation under climate change. Atmos Res 103:119-128

Taylor KE (2001) Summarizing multiple aspects of model performance in a single diagram. J Geophys Res D 106:7183-7192

Timbal B, Hope P, Charles S (2008) Evaluating the consistency between statistically downscaled and global dynamical model climate change projections. J Clim 21:6052-6059

Uppala SM, Källberg PW, Simmons AJ et al (2005) The ERA-40 reanalysis. Q J Roy Met Soc 131:2961-3012

van der Linden P, Mitchell JFB (eds) (2009) ENSEMBLES: Climate Change and Its Impacts: Summary of Research and Results From the ENSEMBLES Project. Met Office Hadley Centre, UK

van Engelen A, Klein Tank AMG, van de Schrier G, Klok L (2008) Towards an operational system for assessing observed changes in climate extremes. ECA&D Report 2008

van Roosmalen L, Sonnenborg TO, Jensen KH, Christensen JH (2011) Comparison of hydrological simulations of climate change using perturbation of observation and distribution-based scaling. Vadose Zone J 10:136-150

van Ulden A, Lenderink G, van den Hurk B, van Mejigaard E (2007) Circulation statistics and climate change in Central Europe: PRUDENCE simulations and observations. Climatic Change 81:179-192

von Storch H (1999) On the Use of "Inflation" in statistical downscaling. J Clim 12:3505-3506

von Storch H (2005) Models of global and regional climate. In: Anderson MG (ed), Encyclopaedia of Hydrological Sciences, Part 3. Meteorology and Climatology. J. Wiley, Chichester, England, p 478-490

von Storch H, Zorita E, Cubasch U (1993) Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. J Clim 6:1161-1171

von Storch H, Langenberg H, Feser F (2000) A spectral nudging technique for dynamical downscaling purposes. Mon Weather Rev 128:3664-3673

Vrac M, Naveau P (2007) Stochastic downscaling of precipitation: From dry events to heavy rainfalls. Water Resour Res 43:W07402. doi:10.1029/2006WR005308

Vrac M, Marbaix P, Paillard D, Naveau P (2007) Nonlinear statistical downscaling of present and LGM precipitation and temperatures over Europe. Clim Past 3:669-682

Walther A, Jeong J-H, Nikulin G, Jones C, Chen D (2013) Evaluation of the warm season diurnal cycle of precipitation over Sweden simulated by the Rossby Centre regional climate model RCA3. Atmos Res 119:131-139

Weigel AP, Liniger MA, Appenzeller C (2008) Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts? Q J Roy Met Soc 134:241-260

Weisse R, Heyen H, von Storch H (2000) Sensitivity of a regional atmospheric model to a sea state-dependent roughness and the need for ensemble calculations. Mon Weather Rev 128:3631-3642

Wetterhall F, Halldin S, Xu C (2007) Seasonality properties of four statistical-downscaling methods in central Sweden. Theor Appl Climatol 87:123-137

Wibig J (1999) Precipitation in Europe in relation to circulation patterns at the 500 hPa level. Int J Climatol 19:253-269

Widmann M, Bretherton CS, Salathé EP (2003) Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. J Clim 16:799-816

Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. Prog Phys Geog 21:530-548

Wilby RL, Wigley TML (2000) Precipitation predictors for downscaling observed and General Circulation Model relationships. Int J Climatol 20:641-661

Wilby RL, Wigley TML, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS (1998) Statistical downscaling of general circulation model output: A comparison of methods. Water Resour Res 34:2995-3008

Wilby RL, Hay LE, Leavesley GH (1999) A comparison of downscaled at raw GCM output: implication for climate change scenarios in the San Juan River Basin, Colorado. J Hydrol 225:67-91

Wilby RL, Hay LE, Gutowski WJJ, Arritt RW, Takle ES, Pan Z, Leavesley GH, Clark MP (2000) Hydrological responses to dynamically and statistically downscaled climate model output. Geophys Res Lett 27:1199-1202

Wilby RL, Whitehead PG, Wade AJ, Butterfield D, Davis RJ, Watts G (2006) Integrated modelling of climate change impacts on water resources and quality in a lowland catchment: River Kennet, UK. J Hydrol 330:204-220

Wilks DS (1995) Statistical Methods in the Atmospheric Sciences. Academic Press, Florida, USA

Wilks DS, Wilby RL (1999) The weather generation game: a review of stochastic weather models. Prog Phys Geogr 23:329-357

Willems P, Vrac M (2011) Statistical precipitation downscaling for small scale hydrological impact investigations of climate change. J Hydrol 402:193-205

Winterfeldt J, Weisse R (2009) Assessment of value added for surface marine wind speed obtained from two Regional Climate Models (RCM). Mon Weather Rev 137:2955-2965

Winterfeldt J, Geyer B, Weisse R (2011) Using QuikSCAT in the added value assessment of dynamically downscaled wind speed. Int J Climatol 31:1028-1039

Wood AW, Leung LR, Sridhar V, Lettenmaier DP (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change 62:189-216

Wramneby A, Smith B, Samuelsson P (2010) Hot spots of vegetation-climate feedbacks under future greenhouse forcing in Europe. J Geophys Res 115:D21119. doi:10.1029/2010JD014307

Xu C, Widen E, Halldin S (2005) Modelling hydrological consequences of climate change – Progress and challenges. Adv Atmos Sci 22:789-797

Yang C, Chandler RE, Isham VS, Wheater HS (2005) Spatial-temporal rainfall simulation using generalized linear models. Water Resour Res 41:W11415. doi:10.1029/2004WR003739

Yang C, Chandler RE, Isham VS, Wheater HS (2006) Quality control for daily observational rainfall series in the UK. Water Environ J 20:185-193

Yang W, Andréasson J, Graham LP, Olsson J, Rosberg J, Wetterhall F (2010) Distribution based scaling to improve usability of RCM regional climate model projections for hydrological climate change impacts studies. Hydrol Res 41:211-229

Yokohata T, Webb MJ, Collins M, Williams KD, Yoshimori M, Hargreaves JC, Annan JD (2010) Structural similarities and differences in climate responses to CO_2 increase between two perturbed physics ensembles. J Clim 23:1392-1410

Zahn M, von Storch H, Bakan S (2008) Climate mode simulation of North Atlantic polar lows in a limited area model. Tellus A 60:620-631

Zorita E, von Storch H (1999) The analog method as a simple statistical downscaling technique: comparison with more complicated methods. J Clim 12:2474-2489