Journal of the Meteorological Society of Japan, Vol. 94A, pp. 31−45, 2016
DOI:10.2151/jmsj.2015-042

Reconsidering the Quality and Utility of Downscaling

Izuru TAKAYABU

Meteorological Research Institute, Tsukuba, Japan

Hideki KANAMARU

Food and Agriculture Organization of the United Nations, Rome, Italy

Koji DAIRAKU

National Research Institute for Earth Science and Disaster Prevention, Tsukuba, Japan

Rasmus BENESTAD

The Norwegian Meteorological Institute, Oslo, Norway

Hans von STORCH

Institute of Coastal Research, Helmholtz Zentrum Geesthacht für Material- und Küstenforschung, Geesthacht, Germany

and

Jens Hesselbjerg CHRISTENSEN

Danish Meteorological Institute, Copenhagen, Denmark

(Manuscript received 30 October 2014, in final form 10 July 2015)

Abstract

Dynamical downscaling (DDS) is performed using regional climate models (RCMs) with global atmospheric states as the input, but there is no consensus among researchers on how to define and estimate the resolvable scale of the various climatic variables obtained by DDS. Sources of RCM uncertainties, including both internal model and intermodel variability, have been assessed by performing ensemble simulations and model intercomparisons, sometimes under the controversial assumption that model bias is independent of the climatic state. Compared with low-resolution global climate simulations, DDS can add value in several ways. For example, because they consider high-resolution topographic data, RCMs can often capture mesoscale phenomena and can better represent climate dynamics. Another downscaling method, empirical statistical downscaling (ESD), is complementary to DDS because it is based on a different philosophy (i.e., sources of information) and on a mostly different set of assumptions. More collaboration and communication should be encouraged among those who develop models, those who use models and perform downscaling, those who use downscaling data, and those who make decisions

Corresponding author: Izuru Takayabu, Atmospheric Environment and Applied Meteorology Research Department, Meteorological Research Institute, Nagamine 1-1, Tsukuba, Ibaraki 305-0052, Japan
E-mail: takayabu@mri-jma.go.jp

based on the scientific results provided by models. In addition, ensemble experiments should be devised that can more effectively benefit impact studies. Using DDS and ESD, separately or in combination, users can maximize the utility of local climate information.

Keywords dynamical downscaling; empirical statistical downscaling; ensemble experiments; skillful scale; uncertainty

1. Introduction

Products of atmosphere-ocean general circulation models generally need to be downscaled to represent climate changes at regional and local scales to increase their usefulness to studies of the impacts of and adaptation to climate change. Indeed, demands for both more accurate and higher resolution climate projection data are steadily increasing. The practice of dynamical downscaling (DDS) global climate simulations using a regional climate model (RCM) was pioneered by Dickinson et al. (1989) and Giorgi et al. (1990).

With regard to RCM performance, Giorgi and Mearns (1991) reported that (1) DDS increases the accuracy with which orographic precipitation can be simulated and (2) RCM accuracy is strongly dependent on features of the driving global model used for the lateral boundary conditions. The Third Assessment Report from Working Group 1 of the Intergovernmental Panel on Climate Change (IPCC 2001) reported the importance of validating model results with high-resolution observation data to determine the performance of an RCM (Rummukainen 2010). In addition, Wang et al. (2004) reviewed the application of DDS in Asian monsoon research and stressed the necessity of refining the physical processes used by the RCM to improve model performance.

In recent reviews, Hong and Kanamitsu (2014) and Xue et al. (2014) discuss the role of the initial conditions in RCM performance in applications of DDS to weather forecasting, including intraseasonal and seasonal forecasting. However, because initial conditions do not greatly affect climate change projections, we do not discuss them further in this paper.

In this work, we try to re-consider downscaling methods, not only DDS but also empirical statistical downscaling (ESD), in climate applications. We identify many of the challenges that are faced by climate downscalers. In particular, we discuss (1) improving the overall performance of RCMs, (2) extending the utility of dynamically downscaled data by coupling DDS with ESD, and (3) enhancing RCM accuracy by adopting ensemble techniques. We believe that framing downscaling in this way will benefit not only climate downscalers but also the users of the downscaled data who need to be aware of both the potential and the limitations of the downscaled regional climate data that they use in their applications.

The remainder of the paper is organized as follows. In section two, we summarize technical issues of DDS. In section three, we discuss issues with the use of downscaled data, whether obtained by DDS or ESD, in studies of climate impacts. In section four, we discuss ensemble DDS experiments. The final section is a summary.

2. Technical issues of dynamical downscaling

2.1 *Origin of dynamical downscaling*

The technique of DDS was originally developed with the aim of applying the methodology of limited-area modeling, developed for weather forecasting, to climate problems over longer time periods. Because this is an ill-posed problem, a technical trick, called the sponge zone, was developed for dealing with the problem of noise generation by inconsistent phase speeds (Davies 1976).

In weather forecasting, the additional skill achieved by DDS over global forecasting runs comes from both the finer resolution of the downscaled model, which provides more detail to the relatively coarse structures at the lateral boundaries of the RCM domain, and the detailed features of the initial field. The daily routine of numerical weather forecasting justifies this approach and has demonstrated the skill of the downscaled models.

However, when limited-area models are run for longer periods of time, for example, years or decades, the conventional DDS approach can no longer utilize the small-scale features of the initial field. Instead, the added value must be generated exclusively by processing the coarse-grid boundary values with additional small-scale forcing (including processes allowed because of the higher resolution) such as the land–sea contrast in the more detailed topography.

In the early years of the development of DDS (i.e., in the early 1990s), it was hoped that this conventional approach would not only add details on smaller scales but also overcome systematic errors present in the coarse-grid model driving the downscaled model; however, this hope did not materialize. It became apparent that only those components that are influenced by the dynamics within the downscaled domain and are strong enough to overcome the coarse-resolution value constraints inflowing across the lateral boundaries can be improved; however, few components meet these criteria.

An example of the successful application of this conventional approach to DDS is found in the field of urban climate research. Kusaka et al. (2012) dynamically downscaled local climate around three major metropolitan areas in Japan and succeeded in separating the heat island effect from the general climate warming around these areas. In this case, temperature within the urban canopy was controlled mainly at the scale of the urban area, and the local model could successfully represent dynamic effects arising at that scale.

2.2 *Scale-selective nudging*

RCMs derive the dynamical state of a region from the lateral boundary conditions, but these do not always determine a unique solution in the interior of the downscaled domain. Rather, several different interior states are consistent with a given set of lateral boundary conditions. It might be possible to overcome this mathematically ill-posed boundary problem by framing the regional modeling as a state–space problem. Based on this concept, a dynamical model can augment existing knowledge about the regional state using the large-scale state of the atmosphere above a certain vertical level, where the influence of regional physiographic details is small (Kida et al. 1991; Sasaki et al. 1995; Waldron et al. 1996; von Storch et al. 2000). The spectral nudging technique, in which the regional model is exposed not only to lateral boundary values but also to scale-selective constraints in the interior that are considered well described by the global analysis or the global simulation, is based on this concept. This general approach was first acknowledged in the Third Assessment Report of the IPCC (Giorgi et al. 2001).

2.3 *Dynamical downscaling in climate studies*

Originally, most regional modeling was aimed at the detailed analysis of mesoscale weather events as well as at weather forecasting; therefore, only short time periods were simulated. Later, multi-year simulations were conducted with present-day conditions, mostly for determining the extent to which the models could reproduce the climatic distributions of, for instance, monthly temperature and precipitation amounts. These simulations were performed mainly for quality-control purposes. Then, a very popular line of application was the determination of possible future developments in so-called “scenario”-simulations, when regional detail was adopted to global climate change simulations.

In recent years, an additional application of DDS has emerged, namely, a type of regional re-analysis that processes not local observations but the (hopefully) homogeneous information about a coarse-scale state, with the objective of arriving at a homogeneous description of mesoscale phenomena and their statistical characteristics. For example, impact models, which project the potential impacts of, say, storm surges or ocean waves on a particular system, are sometimes added to limited-area models (Weisse et al. 2009). Such simulations are valuable because they help to characterize changes in complex quantities, for example, the frequency of mesoscale storms such as polar lows, intra-seasonal percentiles of wind speed, or other complex impact-relevant parameters. These simulations are also used to derive realistic estimates of the risk to, for instance, offshore activities.

2.4 *Remaining issues with dynamical downscaling*

A significant open problem of DDS is the question of how many grid points are needed to describe a phenomenon realistically. In other words, what should the resolution of the regional model be (as opposed to the grid resolution derived from the Nyquist frequency)? The appropriate number may depend on the time-averaging interval and the variables involved. So far, only ad hoc recommendations have been made (e.g., Pielke 1991); systematic studies are lacking.

A clear strategy for defining a “skillful scale” has not yet been developed. Both global climate models (GCMs) and RCMs have a minimum skillful scale that is not the same as the spatial resolution (e.g., it may be four grid boxes). Moreover, the minimum skillful scale varies, depending on the climate element and location. A straightforward definition of skillful scale has been elusive up to now (Benestad et al. 2008), and its definition may also depend on the model design. The skillful scale may derive partly from the fact that both GCMs and DDS use numerical algorithms to solve continuous functions in terms of discrete numbers, parameterization of surface and sub-grid processes, and approximations of primitive equations.

Climate downscaling itself may be another source of uncertainty because none of the numerical algorithms used in DDS provide a perfect description of the relationship between the large and small scales. Hence, it is important to address the question of whether downscaling adds value to climate projections compared with the value offered by the GCM. It is usually claimed that the added value of DDS consists of "more spatial detail." This is a claim that should be, but hardly ever has been, examined in light of the extent to which such detail can be generated by much simpler methods such as geostatistical interpolation (using topography as the explanatory factor). Livezey (1995) described some standard procedures for determining the skill of weather forecast methods. A number of studies have identified enhanced variability at medium scales and more realistic spatial detail, not only of physiographic effects (e.g., coasts) but also of dynamical processes, such as polar lows, typhoons, or medicanes (Mediterranean tropical-like cyclones), as values added by downscaling (Zahn et al. 2008; Chen et al. 2012; Cavicchia and von Storch 2012; Feser et al. 2011; Di Luca et al. 2011).

Taken together, enhanced skill in reproducing the present-day climate by downscaling methods is a necessary but not sufficient condition for credibility of projections in changes in statistics on many regional-scale phenomena. It is not obvious whether DDS can be used to skillfully represent changes in statistical descriptions of weather phenomena and other aspects of the climate system. Changes in, for example, the frequencies of synoptic disturbances, typhoons, and heat waves, reflect both internal variability and nonlinear feedback processes, and these might be different in a warmer climate. We still have no clear answer on designing experiments that more clearly demonstrate potential skill under perturbed climate conditions.

3. The application of dynamical downscaling and empirical statistical downscaling in climate impact studies

3.1 *Dynamical downscaling and empirical statistical downscaling*

Studies investigating the local impacts of climate change need reliable information about local climate characteristics. Moreover, several scientific issues link climate downscaling, whether statistical or dynamical, with studies of climate impacts on various sectors, though they are distinct fields of research.

As is well known, although GCMs do not describe local climatic details accurately, a downscaling process can potentially provide a description of such details. Downscaling is based on the fact that local climate is affected by both local geographical conditions and the surrounding large-scale climate (Benestad et al. 2008; von Storch et al. 1993).

Both DDS and ESD are based on two different philosophies: DDS solves equations for local wind, temperature, and moisture through direct formulations of all known relevant processes (dynamics and thermodynamics), whereas ESD makes use of information from empirical data that may also embed unknown processes. In other words, DDS relies on information that is based on our understanding of physical processes, whereas ESD relies on information obtained by the statistical analysis of observed past climate. Unlike DDS, ESD makes model results (calculations performed with either GCMs or RCMs) comparable with point observations through the introduction of statistical transfer functions, and the skillful scale concept does not apply to ESD as it does to RCMs. While the results obtained from DDS and GCMs may not be directly comparable with observations because they describe a mean value over a volume rather than a point measurement, ESD results may be directly comparable with the observed values used to calibrate the statistical models used in ESD. For most practical purposes, users of climate information such as impact modelers often need information on a local scale (about specific points or about a limited area such as a watershed, valley, or region) that is obtained by either DDS or ESD. Hence, the results obtained by these two downscaling strategies should be compared in terms of the suitability of the information provided to the stakeholders.

Furthermore, different GCMs tend to exhibit different biases in their descriptions of regional and local climate features. For example, mean temperatures may be either too high or too low, too many days with drizzle may be estimated, or the spatial shape of natural phenomena such as El Niño Southern Oscillation (ENSO) may be distorted. In general, downscaling does not correct biases in GCMs, and incorrect wind patterns in a GCM will be passed on to the downscaled results, especially those obtained by DDS. In other words, downscaling is not a "magic" operation that fixes GCMs; nevertheless, biases in the mean state may be avoided using ESD, which tends to focus on the variability rather than on the mean state or the annual cycle of a variable (because these aspects are trivial and can be obtained directly from the empirical data). DDS results, however, often have to be adjusted before they can be used in impact studies (Themeßl et al. 2011; Piani et al. 2010; Schmidli et al. 2006). This topic is discussed further in Section 3.4.

ESD is well suited for large GCM ensembles because of its low computational cost (Benestad 2011); moreover, ESD can provide diagnostics about the projections: the degree of skill with which the GCMs reproduce the spatial characteristics of the predictor; the strength of the relationship between the predictors and the predictands; and the patterns of co-variability between large and small scales. ESD can also be used in some situations to predict the shape of probability distribution functions such as that for the 24-h precipitation, hence providing a description of conditional probabilities for a given set of predictors from a GCM (Benestad 2007; Pryor et al. 2005a, b). In addition, ESD can be used to derive changes in statistics of state such as percentiles (von Storch and Reichardt 1997), full probability distributions, or even parameters in impact study models (Busuioc et al. 1999). Pattern-scaling and bootstrapping methods can also provide information about uncertainty in a climate analog (Ishizaki et al. 2012).

It is also possible to combine ESD and DDS. In this case, ESD uses reanalysis-driven (such as ERA40) DDS results as predictors and observations as predictands, and the calibrated statistical model in ESD is subsequently fed by DDS results describing the future to derive downscaled future climate. This approach is similar to the "model output statistics" (MOS) technique, and it avoids the predictor similarity problem (Wilks 1995). The statistical model used in ESD is able to detect the connection between large-scale conditions and local effects, even if the DDS depiction of the large-scale conditions has systematic errors. Hence, problems due to model biases in the GCMs may be bypassed through this MOS-like approach. However, regional biases in the GCMs driving the DDS for the future projections also produce biased DDS results, and these biases need to be considered. For example, (i) in the GCMS, storm tracks in the North Atlantic may be confined to a narrower range of latitudes than observed storm tracks, which also implies a narrower latitude range in the DDS results; (ii) errors in sea surface temperature may be caused by an ocean model's inability to represent ocean heat transport in sufficient detail; (iii) the South Asia monsoon system may be misrepresented with consequences for the downscaled results; and (iv) ENSO characteristics may be incorrect and result in incorrect downscaled rainfall. Reanalyses, on the other hand, use observations to force the results to look more like the real world than they would otherwise (using assimilation techniques with observations ranging from satellite data to surface measurements). Biases inherited from GCMs may be even more severe than those produced by the RCM itself, and such biases cannot be reduced by sequentially applying the ESD-technique MOS because it assumes perfect boundary conditions (reanalyses) that are different from those provided by the GCM. In some cases, the MOS technique can be applied to DDS driven by an atmosphere-only GCM, if observed sea surface temperatures are used for the boundary condition in the GCM. However, this exception is not relevant to climate change studies, because to fully capture a future state, it is necessary to consider changes in the coupled ocean–atmosphere system. Because in the coupled ocean–atmosphere GCM results, unforced natural variations (such as ENSO) are not synchronized with those in the real world, using the MOS approach is impossible. Moreover, the approach requires consistency between large-scale fields of the DDS and reanalysis data; such consistency can be ensured through, for example, spectral nudging (von Storch et al. 2000). Another technique is to use "common empirical orthogonal functions" (common EOFs) as a hybrid between the MOS approach and the traditional "perfect prog" approach used in ESD when the calibration of the statistical model is based on observed (historical) local and large-scale gridded data (Benestad 2001). Common EOFs can provide a reference frame for the statistical models in ESD, and when they are used, the results from both GCMs and reanalyses are combined into a single data matrix before calibrating the statistical model used in ESD. Some data pre-processing is required because anomalies from both the reanalysis and GCM results must be combined on the same spatial grid. Then, EOFs (Lorenz 1956) are derived from this data matrix. The common EOF approach avoids biases in the mean climatology but requires a good match between the spatio-temporal patterns of the reanalysis and the GCM results. Inspection of the common EOFs can also reveal systematic differences in terms of the predominant modes of variabilities, their common states, and their variance. Although such information can be used to weigh the model results based on their similarity with the reanalysis, it is also possible to bias-correct the common EOFs before the ESD stage.

3.2 Uncertainties

The concept of uncertainty in the context of climate services refers to information about what we know and what we do not know. Climate modeling and downscaling are currently active research arenas, in which the tools (models) are not yet settled, and the potential for predictability often is still not known. Uncertainty is a given in leading-edge research, in contrast to established science, which represents what we know. Therefore, it is necessary to change the focus from uncertainty to certainty. One of the fundamental uncertainties of climate modeling is future emissions. Individual GCMs also have shortcomings, and different GCMs tend to have different biases. Hence, a crude approach for exploring the range of uncertainty is to use an ensemble (see Section 4.2) of GCMs with a downscaling model (whether with DDS or ESD). The divergence of ensemble model results provides an indication of the range of uncertainty if it is not known which of the models forming the ensemble are more reliable; however, it does not properly describe probabilities of the possible outcome because the ensemble of models does not constitute a valid statistical sample.

Downscaling may make the model uncertainties more visible because downscaling using different DDS strategies often provides divergent results. Moreover, the process of downscaling can cause additional uncertainty, even though it produces a more realistic description of the local climate and reduces the differences between the description provided by the GCMs and observations. The role of downscaling in terms of uncertainties depends, of course, on the aim of the modeling exercise, but predicted changes in local climate are expected to be sensitive to the different parameterization schemes ("physics") being used in the RCM and the driving GCMs. There are a number of different parameterization schemes, but it may be difficult to say which scheme is more appropriate for a given location. The sensitivity to the choice of parameterization scheme will result in physical inconsistencies and a different local climatology in the downscaled product, and is one factor connected to the uncertainties associated with regional climate modeling.

The relationship between unresolved physical processes and their large-scale effect described by parameterization schemes may not be stationary. Likewise, ESD may potentially involve nonstationary links between the large and small scales that are also expected to produce biases. Part of these biases is invisible in GCMs, which only describe large-scale phenomena. Potential non-stationarities in the link between large and small scales are part of the overall uncertainty associated with the ESD. There is also the question of whether GCMs represent added value in terms of their ability to predict realistic trends at larger scales, which are translated to the local climate through downscaling. For example, the predictions of different GCMs differ as to how ENSO may change in the future (Collins et al. 2010).

Observations are key to modeling local climate and clarifying the uncertainties with both DDS and ESD. Obviously, ESD requires observations for model development, but observations are also needed to evaluate both ESD and DDS models. Without real data, there is no information about the skill with which DDS can reproduce a given climate element. This fact also implies that the expected skill of regional models to provide better descriptions of certain aspects of local climate cannot be proven without real data. Instead, the existence of such skill becomes a mere claim, one that may or may not be true. Of course, this situation always prevails when models are used to describe features and dynamics that cannot be documented with observations. Moreover, biases in one quantity, such as precipitation or snow, may not be directly related to biases in another (e.g., temperature) (Maraun 2013).

We encourage climate impact researchers to learn about these different sources of uncertainties and their relative roles at different time scales. By considering uncertainties in downscaled climate information, as well as those introduced by impact modeling, researchers will be better able to design impact studies and interpret the results appropriately. Uncertainties do not need to be an obstacle; rather, they are themselves useful information about what we know and what we do not know. One way to deal with uncertainties is to explore the importance of different factors, or to carry out sensitivity tests (Brown and Wilby 2012). Furthermore, a range of statistical methods can be applied to study the quality of model results and the dependency of these results on the choices made in the model set-up. It is also important to apply proper validation to the results (see COST action VALUE; http://www.value-cost.eu/). To date, statisticians, who are experts at making sense out of information, have not been extensively involved in climate science, and climate scientists need to distil the most relevant and reliable information from a range of different sources: observations, model results, and statistical framework. An improved dialog among impact researchers, statisticians, and climate researchers would possibly lead to progress in this area; however, because scholars from different communities speak different "dialects," communication issues sometimes hamper progress. For this reason, several initiatives have been instituted to set up controlled vocabularies, common meta-data structures, and data structure syntax (i.e., CORDEX-ESDM, COST-VALUE, and NPPC; https://www.earthsystemcog.org/projects/downscalingmetadata/).

3.3 Evaluation of downscaling assumptions and outputs

In ESD, four assumptions are made (Benestad et al. 2008; Maraun et al. 2010): (1) the present relationship between small and large scales will also hold in the future (criterion of stationarity); (2) GCMs can skillfully reproduce the predictor used as input for ESD; (3) there is a strong relationship between predictor and predictand; and (4) the predictor contains the climate change "signal." The validity of these assumptions can be assessed to some extent. It is also important to keep in mind that nonstationarity is also an issue in parameterization in both GCMs and DDS.

Evaluations of DDS have suggested that it tends to cause overestimation of downward short-wave radiation throughout the year, and there are large inter-model differences in the relationship between the snow-depth bias and temperature biases. Snow parameterizations in some dynamically downscaled models can be very coarse, and they vary considerably from RCM to RCM, but linking the RCM with a sophisticated snow cover model may improve understanding of projected future changes, based on global warming scenarios, in the characteristics of snow extremes (Fujihara et al. 2008). Evaluation of such climatic variable interrelationships would also improve our understanding of sources of climate downscaling uncertainty'; for agricultural examples, see Iizumi et al. (2008).

Models should be evaluated to gain a better understanding of physical processes in both DDS and ESD. ESD should be viewed as an advanced analysis technique that provides a set of diagnostics in addition to projections (Benestad 2004). ESD can also be used to analyze DDS results with regard to variable interrelationships between large and small scales as well as teleconnections. DDS can be used in numerical experiments to test various hypotheses and examine specific processes. Recently, the VALUE research network (http://www.value-cost.eu/) was established to use a statistical framework to validate and develop downscaling methods for climate change research. Murphy et al. (2010) called for a comprehensive approach for validating downscaled results. In particular, a physics-based validation framework is also needed to check the physical consistency between the GCM and the RCM. For example, the different parameterizations and rain climatologies between a GCM and a RCM affect how vertical energy flow is modeled. This raises the question of whether the output of the RCM represents true added value or simply reflects this discrepancy in vertical energy flows.

3.4 Communication between different communities

Through most of the history of downscaling research, climate scientists have downscaled basic meteorological variables (e.g., surface temperature and precipitation) without adequately considering the utility of downscaling for impact model researchers. The traditional approach, as adopted by the European project FP7-ClipC (http://www.ceda.ac.uk/projects/clipc/; a "one-stop shop" for climate data), has been to archive the downscaled data on a web site (e.g., Copernicus; http://www.ecmwf.int/en/about/what-we-do/copernicus); then, an impact researcher visits the site and downloads the data of choice to drive his impact study model. This one-way approach to data dissemination works well provided that the downscaled data have very high accuracy and the user does not require any guidance in their use, but the reality is very different. Existing models have a limited ability to downscale climate signals, and the data currently being provided by the downscaling community do not necessarily meet user requirements with regard to, for example, choice of variables, accuracy, or temporal and spatial scales. Recently, efforts have been made toward establishing a provider–user dialog by asking users about what they need and in what form they would like to have the data (e.g., FP7-ClipC and FP7-EUPORIAS [http://www.euporias.eu]).

There are several good recent examples of collaboration between dynamical downscalers and impact model researchers, in which they work together from the beginning of a research project to design experiments that consider the needs and perspectives of both communities. The European ENSEMBLES project (http://www.ensembles-eu.org/), which started in 2004, is an early example of downscaling researchers working with climate impact researchers. Currently, one focus of the Coordinated Regional

Climate Downscaling Experiment (CORDEX; http://wcrp-cordex.ipsl.jussieu.fr/) is to enhance interaction between downscalers and users working on climate change adaptation and impact assessment in various regions of the world. The FP7-EUPORIAS project works closely with a number of stakeholders who are interested in regional climate information and services on seasonal-to-decadal timescales. One research theme of JPI Climate (http://www.jpi-climate.eu/), a collaboration among European countries to coordinate climate research, is climate service development and deployment. At the global level, it was decided at World Climate Conference-3 to establish the Global Framework for Climate Services (http://www.gfcs-climate.org/), a United Nations-led initiative to guide the development and application of climate information and services for better decision-making in climate-sensitive sectors. The Climate Services Partnership (http://www.climate-services.org/), in contrast, is an open, informal coalition of "climate information users, providers, donors, and researchers" to improve the development and provision of climate services worldwide.

Collaborative frameworks such as these benefit downscaling research as well as impact studies, and they enable the downscaling community to provide downscaled data with much higher utility and applicability. Despite these recent developments, shortcomings associated with downscaling, whether by DDS or ESD, are still not always communicated to users interested in climate impacts (Maslin and Austin 2012), because the impact research community, which includes climate change adaptation practitioners in a broad sense, is rapidly expanding.

Climate change adaptation policy has often been formulated by local governments in an ad hoc fashion, without coordination among stakeholders. In Denmark, for example, the government has issued a climate change adaptation strategy (Government of Denmark 2008) that has been heavily criticized by practitioners (e.g., by the Danish Society of Engineers, IDA) for not offering any advice on how to coordinate efforts (IDA 2012). Often the information on which such policy-making is based is derived from only one scenario. Furthermore, biases in RCM results are often strong in the case of quantities such as precipitation (Orskaug et al. 2011), and such results must be adjusted before they can be used for hydrological impact assessment. These biases can be corrected using techniques such as daily scaling, quantile–quantile mapping, regression, analogs, discrete wavelet transform, or a combination of these (Themeßl et al. 2011; Piani et al. 2010; Schmidli et al. 2006). However, such adjustment may do no more than sweep these problems under the carpet if the question of whether biases in temperature and precipitation affect the predicted changes is not answered. In some cases, an adjustment (e.g., altitude correction) may be needed because the model is not expected to reproduce the observations with a sufficient degree of precision for the users (e.g., the model resolution may be too coarse) when the correction merely addresses a limitation that does not necessarily affect projected trends (e.g., a general warming trend or a change in storm tracks). However, the possibility that bias correction might mask a more fundamental problem with the representation of important processes (e.g., thermodynamics of air and surface, cloud representation) is concerning. Hence, before such data are used, they must be rigorously evaluated by assessing the model's ability to predict corresponding changes in the past. One difficulty in performing such rigorous evaluations, however, is that local temperature and precipitation are subject to both forced and unforced variability, both of which may be equally pronounced (Deser et al. 2012). The effect of internal unforced variations, however, may be reduced by using statistical techniques such as regression (Van Oldenborgh et al. 2009; Lean and Rind 1998) or by examining linear trends over long time intervals (Benestad 2001). The time interval needs to be long enough to make the error estimates small compared to the slope coefficient.

The strengths and weaknesses of ESD and DDS are both different and independent; hence, these two downscaling techniques should be regarded as complementary. Thus, a convergence of results from the two approaches increases confidence in the results because of both their different technical caveats and their different philosophical approaches (sources of information). The use of both ESD and DDS, combined with the examination of past climate, trends, and relevant climatic phenomena, maximizes information about local climate, which may help in dealing with uncertainties. It is also important to evaluate the GCMs on which the downscaling is based. Communication barriers between the separate ESD and DDS downscaling communities, however, have often led to a failure to take advantage of the different strengths of the two downscaling methods.

It is legitimate to ask whether the traditional top-down approach is the most fruitful for stakeholders when vulnerability to climate change is a function of potential impacts (exposure and sensitivity to exposure) and adaptive capacity. In some cases, a bottom-up approach might be more valuable. In this approach, the impact research community examines the sensitivity of their subject matter to various climatological and nonclimatological parameters, and then assesses what can be said about these parameters in terms of downscaled scenarios (Pielke et al. 2012). This type of assessment is first based on historical observations before including information derived through downscaling.

Both the traditional top-down approach and the bottom-up approach require good communication among climate modelers, impact modelers, and policy-makers for climate change adaptation. In the top-down approach, the results obtained by climate modeling are fed into impact models. In the bottom-up approach, factors affecting societal vulnerability are identified first, and then climate adaptation is regarded as a form of risk management where climate is just one of several potential risk factors. The assessment of risk involves considering the probability of several risk factors occurring together (i.e., "contextual vulnerability"), not all of which may be affected by climate either directly or indirectly.

Sometimes there are unrealistic expectations about which variables models can provide data for, for example, in the case of impact research focused on very short time scales or very local conditions. It is therefore important to ask what types of data have been used before (what were previous decisions based on?) and to investigate whether downscaling can add value to such data by introducing further relevant information. A sensitivity analysis may clarify the relative importance of different weather or climate variables. Furthermore, providers of downscaled climate products need to realize that their job does not end when they publish the data. Rather, there must be a sustained dialog between decision-makers and scientists. In fact, climate projections, downscaling, and impact assessments should be part of a sustained and evolutionary process, where new information and improved models and analysis provide updated projections and assessments. In the work of downscaling for climate impact assessment, climatologists should devise methodologies together with modelers characterizing vulnerability and policy-makers managing adaptation to climate change. Regarding adaptation, some important considerations include the following: Are we adapted to the present climate? To what extent and when will we adapt to the future climate? And to which parameters? For climate adaptation and risk management, a quantitative risk index that shows how the range of risk varies between high- and low-risk areas can be easier to use than a qualitative risk index.

3.5 Remaining issues with empirical statistical downscaling

Various statistical downscaling techniques need to be assessed in a systematic fashion with regard to accuracy scores, distributional similarity scores, and robustness under climate change conditions. The time frequency analysis method is useful for correcting biases in long-term temporal variations. One of the most important issues in ESD is whether the relationship between predictor and predictand will remain stationary in the future. To some extent, past and future stationarity can be tested using GCM/RCM results (Busuioc and von Storch 2003). Furthermore, statistical models in ESD can be calibrated with a subset of data that excludes the highest or lowest values, and then they can be evaluated against these independent data.

4. Ensemble dynamical downscaling experiments

4.1 Origin of dynamical downscaling ensemble experiments

The DDS modeling community was late in recognizing the chaotic nature of regional climate dynamics (Ji and Vernekar 1997; Weisse et al. 2000). Already in the 1970s, global modelers understood (Chervin and Schneider 1976) that running the same model with slightly different conditions (such as slightly different initial states) would generate a different simulation trajectory. This insight has important consequences for the validation strategy of global climate models and for determining the response of such models to prescribed experimental modifications. In contrast, many regional modelers, even to this day, consider one short-duration simulation sufficient to determine both the quality of the model in reproducing "reality" and the model's sensitivity to changes in components. Given the constraints imposed by boundary conditions, it has been more-or-less assumed that internal model variability is insignificant in comparison to that of the large-scale forcing. The first ensemble studies based on analyses with multiple RCMs were published in the 1990s (Ji and Vernekar 1997; Takle et al. 1999), and convincing examples demonstrating the importance of internal variability have been published by Weisse et al. (2000) and Rinke and Dethloff (2000). More systematically, Christensen et al. (2001) compared the level of internal variability generated in RCMs to that of a GCM and showed that, although the internal variability of the RCMs was by comparison substantially reduced, it was not negligible.

4.2 Multi-model ensemble projects

The earlier experiments, which mainly served to explore the existing possibilities for applying regional models, have been superseded by more systematic multi-model approaches and large-scale international modeling efforts. In Europe, the first major attempt to apply a systematic DDS approach with multiple RCMs was pioneered by the PRUDENCE (Prediction of Regional scenarios and Uncertainties for Defining EuropeaN Climate change risks and Effects) project, which was initiated in 2001 (Christensen et al. 2002, 2007). PRUDENCE adopted a coordinated framework and identified the importance of carrying out long-term simulations (30 years) so that intramodel variabilities could be addressed. PRUDENCE was set up and designed to assess the role of intra- and inter-model variability associated with DDS in future projections for Europe, and to a lesser degree to assess the role of the driving GCM. Christensen et al. (2007) summarized the main results of PRUDENCE, and Déqué et al. (2007) more specifically addressed the sources of chaotic behavior in the projections.

After the PRUDENCE project, the idea of portraying uncertainty in regional projections by systematically addressing the known sources of uncertainty in modeling using DDS has been further advanced by studying different regions and by more comprehensively addressing the sources of uncertainty imposed by the driving GCM. The North American Regional Climate Change Assessment Program (NARCCAP) was initiated in 2004, about the time that PRUDENCE became operational (Mearns et al. 2012). NARCCAP focused more extensively on the role of the GCM as a driver of uncertainty in regional climate change projections by designing a DDS set-up that used a wider combination of GCMs and RCMs. While PRUDENCE used only a 2 × 2 matrix combination, NARCCAP was set up to implement a balanced fractional factorial design in which half of a 4 × 6 matrix was sampled in a statistically meaningful way to maximize the amount of information that could be obtained from the experiment; moreover, NARCCAP applied additional statistical techniques to fill in empty combinations in the matrix (Mearns et al. 2012).

In Asia, the Regional Climate Model Intercomparison Project (RMIP; Fu et al. 2005) was established to build on the developments in Europe and North America with similar goals, but with a different geographical focus than PRUDENCE and NARCCAP and with quite different climatological drivers. These included, for example, the Asian monsoon and the subcontinent (i.e., the Tibetan Plateau) downwind of the large-scale flows crossing the Eurasian continent.

The European ENSEMBLES project (see Section 3.4) specified an even larger GCM/RCM matrix, but with a relatively sparse population. This makes the interpretation of the results somewhat more challenging, because the statistical approach for sampling the empty combinations was more difficult (Déqué et al. 2012). One of the major scientific objectives of ENSEMBLES was to establish a means of optimizing the information obtainable from a multi-model DDS ensemble. A major effort was concentrated on defining model evaluation metrics that would highlight performance indices that could be associated as uniquely as possible with RCM performance rather than with the driving boundary conditions. A series of metrics was developed, but the overall conclusion was that establishing a method to weigh the better models higher than the poorer models based on these metrics did not produce a different result or reduce the spread of the projections (Christensen et al. 2010; Déqué et al. 2012).

Other more specialized ensembles have been set up as well to highlight special national needs (see Iizumi et al. 2012; UKCIP 09 [http://ukclimateprojections.defra.gov.uk]; Kjellström et al., 2011). The World Climate Research Programme has adopted a DDS agenda based on ensemble techniques in support of CORDEX (Giorgi et al. 2009; Jones et al. 2011). This project has initiated a multi-faceted exploration of the DDS technique across many geographical regions, although at present it has no clear overall scientific objective. Instead, CORDEX is mainly building on the expressed need for DDS information at local scales across the populated parts of the world. In fact, the aim of CORDEX is to offer this information to potential users in a coordinated and streamlined fashion. Thus, CORDEX is defining a framework under which further research can take place.

In the future, the scientific objectives underpinning the use of multi-model ensembles of downscaled models needs to be clear by more explicitly defining scientific goals. At present, the use of ensembles in DDS climate scenarios is still in its infancy.

4.3 Implementation of an ensemble approach

Ordinarily, many computer resources are needed to cover all GCM and DDS combinations of an ensemble. Therefore, for ensemble projects, an efficient experimental design has been developed with the help of statistical mathematicians. The North American NARCCAP project (e.g., Mearns et al. 2009) used four atmosphere-ocean GCMs (AOGCMs) downscaled by six RCMs. However, producing the full suite of 24 simulations (4 AOGCMs × 6 RCMs) was not possible owing first and foremost to funding limitations. Instead, a statistical design framework was used, in which the full matrix was sampled in a balanced manner with each AOGCM providing boundary conditions for three RCMs and each RCM using boundary conditions from two different AOGCMs. Another possible way to save resources is to implement the ensemble increment method proposed by Yoshimura and Kanamitsu (2013).

When a multi-model ensemble method is adopted, it is assumed that the bias of each model is of about the same magnitude and has the same degree of randomness. Because these assumptions may not hold, a perturbed physics ensemble can be adopted instead by changing the values of the main parameters controlling the behavior of the physical models incorporated in the RCM. Such "detuning experiments" are now being carried out in GCM research (Murphy et al. 2004; Shiogama et al. 2012). It might be possible to extend this method to DDS (Bellprat et al. 2012). Apart from providing a framework for assessing the realism of climate and climate change simulations using DDS, the use of model ensembles also allows data analyses to focus on more general model behavior than what is possible when a single model is used. In a study based on the ENSEMBLES RCMs, Boberg and Christensen (2012) found a common systematic bias in many downscaling models that would affect the downscaling of future projections. In essence, model errors appeared to vary depending on the overall temperature change (Christensen et al. 2008), hence violating the invariant bias assumption inherent in all climate change projections relying on direct model output, as well as in most statistically based bias correction methods that are used. If biases in DDS are not invariant, the so-called delta-change approach (e.g., Hay et al. 2000) is not valid. Understanding the role of systematic model errors in future projections is a scientific goal that can be better pursued using ensembles of models. An essential feature of this study is that all models in the ensemble were used to conduct similar experiments. The simulations representing both the present-day climate and climate change demonstrate a substantial degree of overall warming bias. The former set of experiments allowed the authors to identify a systematic model bias that varies across the ensembles of models, while the latter set of experiments allowed them to validate the role of the bias identified. Firm conclusions could not have been deduced from the study if an ensemble of models had not been used. Moreover, Bellprat et al. (2013) cautioned against over-interpretation of the results because the underlying cause of the differential bias could be related to feedback mechanisms that would not continue to be a scalable effect for all plausible futures. Christensen and Boberg (2012) subsequently expanded their analysis to the Coupled Model Intercomparison Project 5 models and overall found quite similar behaviors in global models. Hence, it is possible to use the knowledge gained from analyzing ensembles of RCMs to identify needs for more general analysis targeted aspects also relevant when using GCM ensemble-based information.

5. Summary

At present, one of the main topics in downscaling research is how to make climate information more useful for the many different kinds of application studies, because downscaled data are increasingly being used as input data for impact models.

It should also be stressed that we need to clarify the "skillful scale" used in downscaling research. Would the skillful scale differ if histograms of hourly data were being examined instead of the climatic value of seasonal averages? Unfortunately, currently we have no measure by which the skillful scale of a model can be judged. If the target region of DDS is sufficiently small, DDS can be regarded as a response function, and we can assume that the local climate is actually controlled at the selected scale. However, the accuracy of the GCM used to set the lateral boundary conditions for DDS becomes an issue when downscaling is considered as a technique for making future projections.

There are many different sources of uncertainties in downscaled products, including future emission scenarios, internal RCM variability, and inter-RCM variability. To advance downscaling science and to aid the users of downscaled data, the characteristics of these uncertainties need to be systematically assessed. A number of ensemble simulations and model intercomparison projects have been carried out to address this issue.

ESD is a complementary approach to DDS; however, the DDS and ESD communities have been mainly working independently, and there is much room for better collaboration and communication between them. Users of downscaling products and impact study researchers would benefit by gaining access to both DDS and ESD products, which have different strengths and weaknesses.

Another important consideration is "added value." What is the added value provided by a GCM in the first place? Does downscaling provide added value comparable to that obtained by simpler methods such as spatial interpolation or by different strategies such as a bottom-up approach and sensitivity analysis? Does DDS or ESD produce added value in addition to that obtained by using a GCM? The answer to this question may differ depending on the intended application of the downscaled product.

Acknowledgments

This commentary was written following a discussion at the 3rd International Workshop on Downscaling, held in Tsukuba, Japan, on 17–19 October 2011, with the support of the S-5-3 Project of the Ministry of the Environment, Japan. We acknowledge the co-chairpersons of the workshop, Dr. Hideo Shiogama, Prof. Tomonori Sato, and Prof. Kei Yoshimura. Some content was also inspired by discussions that took place during the World Climate Research Programme's Open Science Conference, held at Denver, Colorado, USA, on 24–28 October 2011. We acknowledge all of the participants in both the workshop and conference. We received financial support for this paper from the SOUSEI Research Project, funded by the Ministry of Education, Culture, Sports, Science and Technology, Japan, and JHC was supported by the Danish Council for Strategic Research through the project Centre for Regional Change in the Earth System (CRES; http://www.cres-centre.net/) under contract no: DSF-EnMi09-066868.

References

Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär, 2012: Exploring perturbed physics ensembles in a regional climate model. *J. Climate*, **25**, 4582–4599.

Bellprat, O., S. Kotlarski, D. Lüthi, and C. Schär, 2013: Physical constraints for temperature biases in climate models. *Geophys. Res. Lett.*, **40**, 4042–4047.

Benestad, R. E., 2001: A comparison between two empirical downscaling strategies. *Int. J. Climatol.* **21**, 1645–1668.

Benestad, R. E., 2004: Empirical-statistical downscaling in climate modeling. *Eos*, **85**, 417–422.

Benestad, R. E., 2007: Novel methods for inferring future changes in extreme rainfall over Northern Europe. *Climate Res.*, **34**, 195–210.

Benestad, R. E., 2011: A new global set of downscaled temperature scenarios. *J. Climate*, **24**, 2080–2098.

Benestad, R. E., I. Hanssen-Bauer, and D. Chen, 2008: *Empirical-Statistical Downscaling*. World Scientific Publishing Company, Singapore, 228 pp.

Boberg, F., and J. H. Christensen, 2012: Overestimation of Mediterranean summer temperature projections due to model deficiencies. *Nat. Climate Change*, **2**, 433–436.

Brown, C., and R. L. Wilby, 2012: An alternate approach to assessing climate risks. *Eos,* **93**, 401–402.

Busuioc, A., and H. von Storch, 2003: Conditional stochastic model for generating daily precipitation time series. *Climate Res.*, **24**, 181–195.

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

Cavicchia, L., and H. von Storch, 2012: The simulation of medicanes in a high resolution regional climate model. *Climate Dyn.*, **39**, 2273–2290.

Chen, F., B. Geyer, M. Zahn, and H. von Storch, 2012: Toward a multi-decadal climatology of North Pacific Polar Lows employing dynamical downscaling. *Terr., Atmos. Ocean. Sci.*, **23**, 291–301.

Chervin, R. M., and S. H. Schneider, 1976: On determining the statistical significance of climate experiments with general circulation models. *J. Atmos. Sci.*, **33**, 405–412.

Christensen, J. H., and F. Boberg, 2012: Temperature dependent climate projection deficiencies in CMIP5 models. *Geophys. Res. Lett.*, **39**, L24705. doi: 10.1029/2012GL053650.

Christensen, J. H., T. R. Carter, and F. Giorgi, 2002: PRUDENCE employs new methods to assess European climate change. *Eos,* **83**, 147–147.

Christensen, J. H., T. R. Carter, M. Rummukainen, and G. Amanatidis, 2007: Evaluating the performance and utility of regional climate models: the PRUDENCE project. *Climatic Change*, **81**, 1–6.

Christensen, J. H., F. Boberg, O. B. Christensen, and P. Lucas-Picher, 2008: On the need for bias correction of regional climate change projections of temperature and precipitation. *Geophys. Res. Lett.*, **35**, L20709, doi:10.1029/2008GL035694.

Christensen, J. H., E. Kjellström, F. Giorgi, G. Lenderink, and M. Rummukainen, 2010: Weight assignment in regional climate models. *Climate Res.*, **44**, 179–194.

Christensen, O. B., M. A. Gaertner, J. A. Prego, and J. Polcher, 2001: Internal variability of regional climate models. *Climate Dyn.*, **17**, 875–887.

Collins, M., S.-I. An, W. Cai, A. Ganachaud, E. Guilyardi, F.-F. Jin, M. Jochum, M. Lengaigne, S. Power, A. Timmermann, G. Vecchi, and A. Wittenberg, 2010: The impact of global warming on the tropical Pacific

Ocean and El Niño. *Nat. Geosci.*, **3**, 391–397.

Davies, H. C., 1976: A lateral boundary formulation for multilevel prediction models. *Quart. J. Roy. Meteor. Soc.*, **102**, 405–418.

Déqué, M., D. P. Rowell, D. Lüthi, F. Giorgi, J. H. Christensen, B. Rockel, D. Jacob, E. Kjellström, M. de Castro, and B. van den Hurk, 2007: An intercomparison of regional climate simulations for Europe: Assessing uncertainties in model projections. *Climatic Change*, **81**, 53–70.

Déqué, M., S. Somot, E. Sanchez-Gomez, C. M. Goodess, D. Jacob, G. Lenderink, and O. B. Christensen, 2012: The spread amongst ENSEMBLES regional scenarios: Regional climate models, driving general circulation models and interannual variability. *Climate. Dyn.*, **38**, 951–964.

Deser, C., R. Knutti, S. Solomon, and A. S. Phillips, 2012: Communication of the role of natural variability in future North American climate. *Nat. Climate Change*, **2**, 775–779.

Dickinson, R. E., R. M. Errico, F. Giorgi, and G. T. Bates, 1989: A regional climate model for the western United States. *Climatic Change*, **15**, 383–422.

Di Luca, A., R. de Elía, and R. Laprise, 2011: Potential for added value in precipitation simulated by high-resolution nested regional climate models and observations. *Climate Dyn.*, **38**, 1229–1247.

Feser, F., B. Rockel, H. von Storch, J. Winterfeldt, and M. Zahn, 2011: Regional climate models add value to global model data: A review and selected examples. *Bull. Amer. Meteor. Soc.*, **92**, 1181–1192.

Fu, C., S. Wang, Z. Xiong, W. J. Gutowski, D.-K. Lee, J. L. McGregor, Y. Sato, H. Kato, J.-W. Kim, and M.-S. Suh, 2005: Regional climate model intercomparison project for Asia. *Bull. Amer. Meteor. Soc.*, **86**, 257–266.

Fujihara, Y., K. Tanaka, T. Watanabe, T. Nagano, and T. Kojiri, 2008: Assessing the impacts of climate change on the water resources of the Seyhan River Basin in Turkey: Use of dynamically downscaled data for hydrologic simulations. *J. Hydrol.*, **353**, 33–48.

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

Giorgi, F., M. R. Marinucci, and G. Visconti, 1990: Use of a limited-area model nested in a general circulation model for regional climate simulation over Europe. *J. Geophys. Res.*, **95**, 18413–18431.

Giorgi, F., B. Hewitson, J. H. Christensen, M. Hulme, H. von Storch, P. Whetton, R. Jones, L. Mearns, and C. Fu, 2001: Regional climate information - evaluation and projections. *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson (eds.), Cambridge University Press, Cambridge, United Kingdom, and New York, USA, 583–638.

Giorgi, F., C. Jones, and G. G. Asrar, 2009: Addressing climate information needs at the regional level: The CORDEX framework. *WMO Bull.*, **58**, 175–183.

Government of Denmark, 2008: *Strategi for tilpasning til klimaændringer i Danmark*. Regeringen, 52 pp. [Available at http://nyhedsarkiv.ens.dk/graphics/Publikationer/Klima/Klimatilpasning/klimatilpasningsstrategi_03032008.pdf] (in Danish).

Hay, L. E., R. L. Wilby, and G. H. Leavesley, 2000: A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States. *J. Amer. Water Resour. As.*, **36**, 387–397.

Hong, S.-Y., and M. Kanamitsu, 2014: Dynamical downscaling: Fundamental issues from an NWP point of view and recommendations. *Asia-Pacific J. Atmos. Sci.*, **50**, 83–104.

IDA, 2012: Klimatilpasning Af Danmark-IDAs Klimatilpasningsstrategi. [Available at http://ipaper.ipapercms.dk/IDA/Politik/Klimatilpasning/.]

Iizumi, T., M. Nishimori, and M. Yokozawa, 2008: Combined equations for estimating global solar radiation: Projection of radiation field over Japan under global warming conditions by statistical downscaling. *J. Agric. Meteor.*, **64**, 9–23.

Iizumi, T., F. Uno, and M. Nishimori, 2012: Climate downscaling as a source of uncertainty in projecting local climate change impacts. *J. Meteor. Soc. Japan*, **90B**, 83–90.

IPCC, 2001: *Climate Change 2001: The Scientific Basis*. Contribution of working group I to the third assessment report of the intergovernmental panel on climate change. Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson (eds.), Cambridge University Press, Cambridge, UK, and New York, USA, 881 pp.

Ishizaki, N. N., H. Shiogama, K. Takahashi, S. Emori, K. Dairaku, H. Kusaka, T. Nakaegawa, and I. Takayabu, 2012: An attempt to estimate of probabilistic regional climate analogue in a warmer Japan. *J. Meteor. Soc. Japan*, **90B**, 65–74.

Ji, Y., and A. D. Vernekar, 1997: Simulation of the Asian summer monsoon of 1987 and 1988 with a regional model nested in a global GCM. *J. Climate*, **10**, 1965–1979.

Jones, C., F. Giorgi, and G. Asrar, 2011: The Coordinated Regional Downscaling Experiment: CORDEX, An international downscaling link to CMIP5. *Clivar Exchanges*, **16**, 34–40.

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

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

Kusaka, H., M. Hara, and Y. Takane, 2012: Urban climate projection by the WRF model at 3-km horizontal grid increment: Dynamical downscaling and predicting heat stress in the 2070's August for Tokyo, Osaka, and Nagoya Metropolises. *J. Meteor. Soc. Japan*, **90B**, 47–63.

Lean, J., and D. Rind, 1998: Climate forcing by changing solar radiation. *J. Climate*, **11**, 3069–3094.

Livezey, R. E., 1995: The evaluation of forecasts. *Analysis of Climate Variability: Applications of Statistical Techniques*. von Storch, H., and A. Navarra (eds), Springer Verlag, 177–196.

Lorenz, E. N., 1956: *Empirical orthogonal functions and statistical weather prediction*. Sci. Rep. No. 1, Statistical forecasting project, Massachusetts Institute of Technology, Cambridge, MA, 48 pp.

Maraun, D., 2013: Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *J. Climate*, **26**, 2137–2143.

Maraun, D., F. Wetterhall, A. M. Ireson, R. E. Chandler, E. J. Kendon, M. Widmann, S. Brienen, H. W. Rust, T. Sauter, M. Themeßl, V. K. C. Venema, K. P. Chun, C. M. Goodess, R. G. Jones, C. Onof, M. Vrac, and I. Thiele-Eich, 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.

Maslin, M., and P. Austin, 2012: Uncertainty: Climate models at their limit? *Nat.*, **486**, 183–184.

Mearns, L. O., W. J. Gutowski, R. Jones, R. Leung, S. McGinnis, A. Nunes, and Y. Qian, 2009: A regional climate change assessment program for North America. *Eos*, **90**, 311–312.

Mearns, L. O., R. Arritt, S. Biner, M. S. Bukovsky, S. McGinnis, S. Sain, D. Caya, J. Correlia, Jr., D. Flory, W. Gutowski, E. S. Takle, R. Jones, R. Leung, W. Moufouma-Okia, L. McDaniel, A. M. B. Nunes, Y. Qian, J. Roads, L. Sloan, and M. Snyder, 2012: The North American regional climate change assessment program: Overview of phase I results. *Bull. Amer. Meteor. Soc.*, **93**, 1337–1362.

Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth, 2004: Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nat.*, **430**, 768–772.

Murphy, J., V. Kattsov, N. Keenlyside, M. Kimoto, G. Meehl, V. Mehta, H. Pohlmann, A. Scaife, and D. Smith, 2010: Towards prediction of decadal climate variability and change. *Procedia Environ. Sci.*, **1**, 287–304.

Orskaug, E., I. Scheel, A. Frigessi, P. Guttorp, J. E. Haugen, O. E. Tveito, and O. Haug, 2011: Evaluation of a dynamic downscaling of precipitation over the Norwegian mainland. *Tellus A*, **63**, 746–756.

Piani, C., J. O. Haerter, and E. Coppola, 2010: Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor. Appl. Climatol.*, **99**, 187–192.

Pielke, Sr., R. A., 1991: A recommended specific definition of "resolution". *Agric. For. Meteor.*, **72**, 1914–1914.

Pielke, Sr., R. A., R. Wilby, D. Niyogi, F. Hossain, K. Dairaku, J. Adegoke, G. Kallos, T. Seastedt, and K. Suding, 2012: Dealing with complexity and extreme events using a bottom-up, resource-based vulnerability perspective. *Extreme Events and Natural Hazards: The Complexity Perspective*, 345–359.

Pryor, S. C., J. T. Schoof, and R. J. Barthelmie, 2005a: Climate change impacts on wind speeds and wind energy density in northern Europe: Empirical downscaling of multiple AOGCMs. *Climate Res.*, **29**, 183–198.

Pryor, S. C., J. T. Schoof, and R. J. Barthelmie, 2005b: Empirical downscaling of wind speed probability distributions. *J. Geophys. Res.*, **110**, D19110, doi:10.1029/2005JD005899.

Rinke, A., and K. Dethloff, 2000: On the sensitivity of a regional Arctic climate model to initial and boundary conditions. *Climate Res.*, **14**, 101–113.

Rummukainen, M., 2010: State-of-the-art with regional climate models. *WIREs Climate Change*, **1**, 82–96.

Sasaki, H., H. Kida, T. Koide, and M. Chiba, 1995: The performance of long term integrations of a limited area model with the spectral boundary coupling method. *J. Meteor. Soc. Japan*, **73**, 165–181.

Schmidli, J., C. Frei, and P. L. Vidale, 2006: Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *Int. J. Climatol.*, **26**, 679–689.

Shiogama, H., M. Watanabe, M. Yoshimori, T. Yokohata, T. Ogura, J. D. Annan, J. C. Hargreaves, M. Abe, Y. Kamae, R. O'ishi, R. Nobui, S. Emori, T. Nozawa, A. Abe-Ouchi, and M. Kimoto, 2012: Perturbed physics ensemble using the MIROC5 coupled atmosphere-ocean GCM without flux corrections: Experimental design and results. *Climate Dyn.*, **39**, 3041–3056.

Takle, E. S., W. J. Gutowski, Jr., R. W. Arritt, Z. Pan, C. J. Anderson, R. R. da Silvia, D. Caya, S.-C. Chen, F. Giorgi, J. H. Christensen, S.-Y. Hong, H.-M. H. Juang, J. Katzfey, W. M. Lapenta, R. Laprise, G. E. Liston, P. Lopez, J. McGregor, R. A. Pielke, Sr., and J. O. Roads, 1999: Project to intercompare regional climate simulations (PIRCS): Description and initial results. *J. Geophys. Res.*, **104**, 19443–19461.

Themeßl, M. J, A. Gobiet, and A. Leuprecht, 2011: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *Int.*

J. Climatol., **31**, 1530–1544.

van Oldenborgh, G. J., S. Drijfhout, A. van Ulden, R. Haarsma, A. Sterl, C. Severijns, W. Hazeleger, and H. Dijkstra., 2009: Western Europe is warming much faster than expected. *Climate. Past*, **5**, 1–12.

von Storch, H., and H. Reichardt, 1997: A scenario of storm surge statistics for the German Bight at the expected time of doubled atmospheric carbon dioxide concentration. *J. Climate*, **10**, 2653–2662.

von Storch, H., E. Zorita, and U. Cubasch, 1993: Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *J. Climate*, **6**, 1161–1171.

von Storch, H., H. Langenberg, and F. Feser, 2000: A spectral nudging technique for dynamical downscaling purposes. *Mon. Wea. Rev.*, **128**, 3664–3673.

Waldron, K. M., J. Peagle, and J. D. Horel, 1996: Sensitivity of a spectrally filtered and nudged limited area model to outer model options. *Mon. Wea. Rev.*, **124**, 529–547.

Wang, Y., L. R. Leung, J. L. McGregor, D.-K. Lee, W.-C. Wang, Y. Ding, and F. Kimura, 2004: Regional climate modeling: Progress, challenges, and prospects. *J. Meteor. Soc. Japan*, **82**, 1599–1628.

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

Weisse, R., H. von Storch, U. Callies, A. Chrastansky, F. Feser, I. Grabemann, H. Guenther, A. Pluess, T. Stoye, J. Tellkamp, J. Winterfeldt, and K. Woth, 2009: Regional meteorological-marine reanalyses and climate change projections: Results for Northern Europe and potentials for coastal and offshore applications. *Bull. Amer. Meteor. Soc.*, **90**, 849–860.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences, 2nd edition.* International Geophysics Series, **59**, Academic Press, Orlando, Florida, USA, 464 pp.

Xue, Y.-K., Z. Janjic, J. Dudhia, R. Vasic, and F. D. Sales, 2014: A review on regional dynamical downscaling in intraseasonal to seasonal simulation/prediction and major factors that affect downscaling ability. *Atmos. Res.*, 68–85.

Yoshimura, K., and M. Kanamitsu, 2013: Incremental correction for the dynamical downscaling of ensemble mean atmospheric fields. *Mon. Wea. Rev.*, **141**, 3087–3101.

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