

Review

Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling

H. J. Fowler,^{a*} S. Blenkinsop^a and C. Tebaldi^b

^a *Water Resource Systems Research Laboratory, School of Civil Engineering and Geosciences, Newcastle University, UK*

^b *Institute for the Study of Society and Environment, National Center for Atmospheric Research, Boulder, CO, USA*

Abstract:

There is now a large published literature on the strengths and weaknesses of downscaling methods for different climatic variables, in different regions and seasons. However, little attention is given to the choice of downscaling method when examining the impacts of climate change on hydrological systems. This review paper assesses the current downscaling literature, examining new developments in the downscaling field specifically for hydrological impacts. Sections focus on the downscaling concept; new methods; comparative methodological studies; the modelling of extremes; and the application to hydrological impacts.

Consideration is then given to new developments in climate scenario construction which may offer the most potential for advancement within the ‘downscaling for hydrological impacts’ community, such as probabilistic modelling, pattern scaling and downscaling of multiple variables and suggests ways that they can be merged with downscaling techniques in a probabilistic climate change scenario framework to assess the uncertainties associated with future projections. Within hydrological impact studies there is still little consideration given to applied research; how the results can be best used to enable stakeholders and managers to make informed, robust decisions on adaptation and mitigation strategies in the face of many uncertainties about the future. It is suggested that there is a need for a move away from comparison studies into the provision of decision-making tools for planning and management that are robust to future uncertainties; with examination and understanding of uncertainties within the modelling system. Copyright © 2007 Royal Meteorological Society

KEY WORDS downscaling; climate change; hydrological impacts; comparative studies; extremes; uncertainty

Received 6 October 2006; Revised 27 March 2007; Accepted 6 April 2007

INTRODUCTION

General circulation models (GCMs) are an important tool in the assessment of climate change. These numerical coupled models represent various earth systems including the atmosphere, oceans, land surface and sea-ice and offer considerable potential for the study of climate change and variability. However, they remain relatively coarse in resolution and are unable to resolve significant subgrid scale features (Grotch and MacCracken, 1991) such as topography, clouds and land use. For example, the Hadley Centre’s HadCM3 model is resolved at a spatial resolution of 2.5° latitude by 3.75° longitude whereas a spatial resolution of 0.125° latitude and longitude is required by hydrologic simulations of monthly flow in mountainous catchments (Salathé, 2003). Bridging the gap between the resolution of climate models and regional and local

scale processes represents a considerable problem for the impact assessment of climate change, including the application of climate change scenarios to hydrological models. Thus, considerable effort in the climate community has focussed on the development of techniques to bridge the gap, known as ‘downscaling’.

A number of papers have previously reviewed downscaling concepts, including Hewitson and Crane (1996); Wilby and Wigley (1997); Zorita and von Storch (1997); Xu (1999); Wilby *et al.* (2004); and regionally for Scandinavia in Hanssen-Bauer *et al.* (2005). This paper differs from previous reviews as it focuses on recent developments in downscaling methods for hydrological impact studies, updating and extending the methodological study of Xu (1999). In the next two sections the concept of downscaling, the development of new methods, comparative methodological studies and the modelling of extremes are discussed. The application of downscaling to the field of climate change impacts on hydrological modelling is reviewed in ‘Downscaling for hydrological impact studies’. ‘Incorporating new developments’

*Correspondence to: H. J. Fowler, Water Resource Systems Research Laboratory, School of Civil Engineering and Geosciences, Cassie Building, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK. E-mail: h.j.fowler@ncl.ac.uk

reviews new developments in climate scenario construction, such as probabilistic modelling, pattern scaling and downscaling of multiple variables and suggests ways that they can be merged with downscaling techniques in a probabilistic climate change scenario framework to assess the uncertainties associated with future projections. The last section ‘‘Summary and next steps’’ draws these themes together to make some recommendations on future work in the field, providing an example of how probabilistic climate scenarios can be linked with downscaling methods for hydrological, and other, impact studies.

In particular, this review will try to answer five questions that we believe must be addressed for the successful use of downscaling methods in hydrological impact assessment, in both the downscaling research community and for practitioners:

1. What more (if anything) can be learnt from downscaling method comparison studies?
2. Can dynamical downscaling contribute advantages that can not be conferred by statistical downscaling?
3. Can realistic climate change scenarios be produced from dynamically downscaled output for periods outside the time period of simulation using methods such as pattern scaling?
4. What new methods can be used together with downscaling to assess uncertainties in hydrological response?
5. How can downscaling methods be better utilized within the hydrological impacts community?

Whilst this review aims to discuss recent developments in the application of climate change scenarios, through downscaling methods, to assess hydrological impacts, it will not provide a comprehensive review of all published studies. Instead it aims to concentrate on those studies

that address new concepts and real advances in downscaling for hydrological impact assessment, particularly those that address the quantification of uncertainty in the estimation of climate change impacts. Therefore, the review will concentrate on studies that compare different downscaling approaches, the outputs from multiple climate models or ensembles, and multiple emissions scenarios.

OVERVIEW OF DOWNSCALING METHODS

Two fundamental approaches exist for the downscaling of large-scale GCM output to a finer spatial resolution. The first of these is a dynamical approach where a higher-resolution climate model is embedded within a GCM. The second approach is to use statistical methods to establish empirical relationships between GCM-resolution climate variables and local climate. These two approaches are described below and the main advantages and limitations of each are summarized in Table I.

Dynamical downscaling

Dynamical downscaling refers to the use of regional climate models (RCMs), or limited-area models (LAMs). These use large-scale and lateral boundary conditions from GCMs to produce higher resolution outputs. These are typically resolved at the $\sim 0.5^\circ$ latitude and longitude scale and parameterize physical atmospheric processes. Thus, they are able to realistically simulate regional climate features such as orographic precipitation (e.g. Frei *et al.*, 2003), extreme climate events (e.g. Fowler *et al.*, 2005a; Frei *et al.*, 2006) and regional scale climate anomalies, or non-linear effects, such as those associated with the El Niño Southern Oscillation (e.g. Leung *et al.*, 2003a). However, model skill depends strongly on biases inherited from the driving GCM and the presence and strength of regional scale forcings such as orography, land-sea contrast and vegetation cover. Studies

Table I. Comparative summary of the relative merits of statistical and dynamical downscaling techniques (adapted from Wilby and Wigley, 1997).

	Statistical downscaling	Dynamical downscaling
Advantages	<ul style="list-style-type: none"> • Comparatively cheap and computationally efficient • Can provide point-scale climatic variables from GCM-scale output • Can be used to derive variables not available from RCMs • Easily transferable to other regions • Based on standard and accepted statistical procedures • Able to directly incorporate observations into method 	<ul style="list-style-type: none"> • Produces responses based on physically consistent processes • Produces finer resolution information from GCM-scale output that can resolve atmospheric processes on a smaller scale
Disadvantages	<ul style="list-style-type: none"> • Require long and reliable observed historical data series for calibration • Dependent upon choice of predictors • Non-stationarity in the predictor-predictand relationship • Climate system feedbacks not included • Dependent on GCM boundary forcing; affected by biases in underlying GCM • Domain size, climatic region and season affects downscaling skill 	<ul style="list-style-type: none"> • Computationally intensive • Limited number of scenario ensembles available • Strongly dependent on GCM boundary forcing

within the western U.S., Europe, and New Zealand, where topographic effects on temperature and precipitation are prominent, often report more skilful dynamical downscaling than in regions such as the U.S., Great Plains and China where regional forcings are weaker (Wang *et al.*, 2004).

Variability in internal parameterizations also provides considerable uncertainty. Therefore, use of model ensembles is to be recommended for a realistic assessment of climate change. Hagemann *et al.* (2004) examined the relative performance of four RCMs (HIRHAM, CHRM, REMO and HadRM3H) and a variable resolution GCM (ARPEGE) over the Danube and Baltic Sea catchments. Boundary conditions from gridded reanalysis data were used to remove the effect of errors in GCM boundary conditions, and therefore identify simulation errors resulting from internal RCM parameterizations. Over the Baltic, all models overestimated precipitation, except in summer; probably due to inaccurate parameterizations of large-scale condensation and convection schemes. For the more continental Danube, all models except ARPEGE simulated a dry summer bias. In CHRM this was related to poor soil parameterization but in the other models it was due to lack of moisture advection into the region. Additional errors were noted to result from poor simulation of snow-albedo feedback.

As dynamical downscaling is computationally expensive, model integrations have, until recently, been restricted to 'time slices'; normally ~ 30 years for a control or 'baseline' climate from 1961–1990 and for a changed climate from 2070–2100. This makes climate change impacts for other periods difficult to assess. Producing scenarios for other periods has been addressed using 'pattern scaling' (further discussed in the section "Pattern scaling") – where changes are scaled according to the temperature signal modelled for the intervening period, assuming a linear pattern of change (e.g. Prudhomme *et al.*, 2002). However, some transient RCM simulations, from 1950 to 2100, are now becoming available (Erik Kjellström, personal communication) and so this scaling issue may soon be overcome.

Using a RCM provides additional uncertainty to that inherent to GCM output. Uncertainty in RCM formulation has a small, but non-negligible, impact on future projections of UK and Ireland mean-climate (Rowell, 2006; Fowler and Blenkinsop, in press). For temperature projections, the uncertainty introduced by the RCM is less than that from the emissions scenario, but for precipitation projections the opposite is true. However, the largest source of uncertainty derives from the structure and physics of the formulation of the driving GCM. The uncertainty introduced by ten RCMs for eight European regions was evaluated by Déqué *et al.* (2005) using RCM ensemble runs and applying the same emissions scenario. The contribution of different sources of uncertainty was found to vary according to spatial domain, region and season, but the largest uncertainty was introduced by the boundary forcing i.e. choice of driving GCM, particularly for temperature. Exceptionally, for summer precipitation

the uncertainty attributable to the choice of RCM was of the same magnitude.

There has now been much assessment of the ability of RCMs to simulate climate variables, particularly those relevant to hydrological impact studies. Several studies (e.g. Leung *et al.*, 2004) have illustrated how dynamical downscaling provides 'added value' for the study of climate change and its potential impacts, as regional climate change signals can be significantly different from those projected by GCMs because of orographic forcing and rain-shadowing effects. Dynamical downscaling can also provide improved simulation of meso-scale precipitation processes and thus higher moment climate statistics (Schmidli *et al.*, 2006); producing more plausible climate change scenarios for extreme events and climate variability at the regional scale. To this end, longer duration, higher spatial resolution (e.g. Christensen *et al.*, 1998), and ensemble RCM simulations (e.g. Leung *et al.*, 2004), particularly the European FP5 Prediction of Regional scenarios and Uncertainties for defining European Climate change risks and Effects (PRUDENCE) project (Christensen *et al.*, 2007) and the North American Regional Climate Change Assessment Program (NARCCAP) project (Mearns *et al.*, 2006), are becoming more common. These improve the realism of control simulations; more accurate variability and extreme event statistics are simulated by higher spatial and temporal resolution models (e.g. Frei *et al.*, 2006). Applications to geographically diverse regions and model inter-comparison studies have allowed the strengths and weaknesses of dynamical downscaling to be better understood (Wang *et al.*, 2004). This has recently proliferated their use in impact studies (e.g. Bergstrom *et al.*, 2001; Wood *et al.*, 2004; Zhu *et al.*, 2004; Graham *et al.*, 2007 a,b); and is further discussed in "Downscaling for hydrological impact studies".

Statistical downscaling

Many statistical downscaling techniques have been developed to translate large-scale GCM output onto a finer resolution.

The simplest method is to apply GCM-scale projections in the form of change factors (CFs) – the 'perturbation method' (Prudhomme *et al.*, 2002) or 'delta-change' approach. Differences between the control and future GCM simulations are applied to baseline observations by simply adding or scaling the mean climatic CF to each day. Therefore, it can be rapidly applied to several GCMs to produce a range of climate scenarios but has a number of caveats. Firstly, the method assumes that GCMs more accurately simulate relative change than absolute values, i.e. assuming a constant bias through time. Secondly, CFs only scale the mean, maxima and minima of climatic variables, ignoring change in variability and assuming the spatial pattern of climate will remain constant (Diaz-Nieto and Wilby, 2005). Furthermore, for precipitation the temporal sequence of wet days is unchanged, when change in wet and dry spells may be an important component of climate change. Several modifications have been

proposed. Prudhomme *et al.* (2002) suggest any increase in precipitation is distributed evenly among existing rain days, added to make each third dry day wet, or distributed on only the three wettest days to simulate an increase in extremes. The effectiveness of these arbitrary parameters is however, not assessed. Harrold and Jones (2003) rank GCM daily rainfall for current and future climates and use these to scale ranked historical precipitation series, a variation which is sensitive to change in extreme rainfall and wet day frequencies.

More sophisticated statistical downscaling methods are generally classified into three groups:

- Regression models
- Weather typing schemes
- Weather generators (WGs)

Each group covers a range of methods, all relying on the fundamental concept that regional climates are largely a function of the large-scale atmospheric state. This relationship may be expressed as a stochastic and/or deterministic function between large-scale atmospheric variables (predictors) and local or regional climate variables (predictands). Predictor variables useful for downscaling typically represent the large-scale circulation, e.g. sea-level pressure and geopotential heights, but can also include measures of humidity and simulated surface climate variables such as GCM precipitation and temperature (e.g. Widmann and Bretherton, 2000; Salathé, 2005). Essentially, in these methods, the regional climate is considered to be conditioned by the large-scale climate state in the form $R = F(X)$, where R represents the local climate variable that is being downscaled, X is the set of large-scale climate variables and F is a function which relates the two and is typically established by training and validating the models using point observations or gridded reanalysis data. Performance of these methods in reproducing observed or reanalysis statistics is normally measured using correlation coefficients, distance measures such as root mean squared error (RMSE), or explained variance, although Busuioc *et al.* (2001) suggest that for climate change applications the optimum downscaling model may well be that which best reproduces low frequency variability (e.g. Wilby *et al.*, 2002a).

Several key assumptions are inherent within these statistical downscaling techniques. Firstly, predictor variables should be physically meaningful, reproduced well by the GCM and able to reflect the processes responsible for climatic variability on a range of timescales. For example, downscaling which uses only circulation-based predictors may fail to reflect change in atmospheric humidity in a warmer climate. Secondly, the predictor–predictand relationship is assumed to be stationary in time, remaining the same in a changed future climate. This assumption has shown to be questionable in the observed record (e.g. Huth, 1997; Slonosky *et al.*, 2001; Fowler and Kilsby, 2002) and is best tested using

long records or model validation on a period with different climate characteristics (Charles *et al.*, 2004). Non-stationarity may be attributed to an incomplete set of predictor variables that exclude low-frequency climate behaviour, inadequate sampling or calibration period, or temporal change in climate system structures (Wilby, 1998). The degree of non-stationarity in projected climate change has recently been assessed by Hewitson and Crane (2006) who found that this is relatively small and that circulation dynamics in particular may be more robust to non-stationarities.

Choice of predictor variables should also be given high consideration. A predictor may not appear significant when developing a downscaling model under present climate, but future changes in that predictor may be critical in determining climate change (Wilby, 1998). For example, local temperature change under a $2 \times \text{CO}_2$ scenario is dominated by change in atmospheric radiative properties rather than circulation changes (Schubert, 1998) but for local precipitation change the inclusion of low-frequency atmospheric predictors can produce enhanced simulations (Wilby, 1998). There is little consensus on the most appropriate choice of predictor variables. Circulation-related predictors, such as sea-level pressure, are attractive as relatively long observations are available and GCMs simulate these with some skill (Cavazos and Hewitson, 2005). However, it is increasingly acknowledged that circulation predictors alone are unlikely to be sufficient, as they fail to capture key precipitation mechanisms based on thermodynamics and vapour content. Thus humidity has increasingly been used to downscale precipitation (e.g. Karl *et al.*, 1990; Wilby and Wigley, 1997; Murphy, 2000; Beckmann and Buishand, 2002); particularly as it may be an important predictor under a changed climate. Indeed, the inclusion of moisture variables as predictors can lead to convergence in the results of statistical and dynamical approaches (Charles *et al.*, 1999), with the inclusion of GCM precipitation as a predictor also improving downscaling skill (Salathé, 2003; Widmann *et al.*, 2003). Cavazos and Hewitson (2005) have performed the most comprehensive assessment of predictor variables to date, assessing 29 NCEP reanalysis variables using an artificial neural network (ANN) downscaling method in 15 locations. Predictors representing mid-tropospheric circulation (geopotential heights) and specific humidity were found to be useful in all locations and seasons. Tropospheric thickness and surface meridional and mid-tropospheric wind components were also important predictors but more regionally and seasonally dependent.

The results of statistical downscaling are also dependent upon the choice of predictor domain (Wilby and Wigley, 2000). However, this is generally ignored (Benestad, 2001). A study by Brinkmann (2002) suggests that, for studies where only one grid point is used, the optimum grid point location for downscaling may be a function of the timescale under consideration and is not necessarily related solely to location. Additionally, large-scale circulation patterns over the predictor domain

may not capture small-scale processes; these may result from variability in neighbouring locations. Similar results were obtained by Wilby and Wigley (2000) who found that in many cases, maximum correlations between precipitation and mean sea level pressure (MSLP) occurred away from the grid box, suggesting that the choice of predictor domain, in terms of location and spatial extent, is a critical factor affecting the realism and stability of downscaled precipitation scenarios.

Statistical methods are more straightforward than dynamical downscaling but tend to underestimate variance and poorly represent extreme events. Regression methods and some weather-typing approaches underpredict climate variability to varying degrees, since only part of the regional and local climate variability is related to large-scale climate variations. Three approaches to add variability to downscaled climate variables are frequently employed: variable inflation, expanded downscaling and randomization. Variable inflation (Karl *et al.*, 1990) increases variability by multiplying by a suitable factor. However, von Storch (1999) suggests that variable inflation is not meaningful as it assumes that all climate variability is related to the large-scale predictor fields, recommending the alternative approach of 'randomization' where additional variability is added in the form of white noise (e.g. Kilsby *et al.*, 1998). This was found to give good results in the reproduction of 20–50 year return values of central European surface temperature (Kysely, 2002). The more sophisticated 'expanded downscaling' approach, a variant of canonical correlation analysis (CCA), was developed by Bürger (1996) and has been used by Huth (1999); Dehn *et al.* (2000) and Müller-Wohlfeil *et al.* (2000). A comparison of the three methods notes that each presents different problems (Bürger and Chen, 2005). Variable inflation poorly represents spatial correlations, whilst randomization performs well for control climate simulations but is unable to reproduce changes in variability which may represent a significant disadvantage given the expectations of future change. In contrast, expanded downscaling is sensitive to the choice of statistical process used during its application.

Regression models. The term 'transfer function' (Giorgi and Hewitson, 2001) is used to describe methods that directly quantify a relationship between the predictand and a set of predictor variables. In the simplest form, multiple regression models are built using grid cell values of atmospheric variables as predictors for surface temperature and precipitation (e.g. Hanssen-Bauer and Førland, 1998; Hellström *et al.*, 2001). Other more complex techniques include using the principal components of pressure fields or geopotential heights (e.g. Cubasch *et al.*, 1996; Kidson and Thompson, 1998; Hanssen-Bauer *et al.*, 2003) and more sophisticated methods such as ANNs (e.g. Zorita and von Storch, 1999), CCA (Karl *et al.*, 1990; Wigley *et al.*, 1990; von Storch *et al.*, 1993; Busuioc *et al.*, 2001) and singular value decomposition (SVD) (Huth, 1999; von Storch and Zwiers, 1999).

There have been a number of recent innovations in this type of downscaling. For example, Abaurrea and Asín (2005) applied a logistic regression model to daily precipitation probability and a generalized linear model for wet day amounts for the Ebro Valley, Spain. The approach simulated seasonal characteristics and some aspects of daily behaviour such as wet and dry runs well, but had low skill in reproducing extreme events. Bergant and Kajfez-Bogataj (2005) used multi-way partial least squares regression to downscale temperature and precipitation in Slovenia, more appropriate with predictor variables that are strongly correlated. The method performed better than a more conventional method based on regression of principal components but was tested only on the cold season.

Weather typing schemes. Weather typing or classification schemes relate the occurrence of particular 'weather classes' to local climate. Weather classes may be defined synoptically, typically using empirical orthogonal functions EOFs from pressure data (Goodess and Palutikof, 1998), by indices from SLP data (e.g. Conway *et al.*, 1996), or by applying cluster analysis (Fowler *et al.*, 2000, 2005b) or fuzzy rules (Bárdossy *et al.*, 2002, 2005) to atmospheric pressure fields. Local surface variables, typically precipitation, are conditioned on daily weather patterns by deriving conditional probability distributions for observed statistics, e.g. $p(\text{wet-wet})$ or mean wet day amount, associated with a given atmospheric circulation pattern (e.g. Bellone *et al.*, 2000). Climate change is estimated by evaluating the change in the frequency of the weather classes simulated by the GCM. The method assumes that the characteristics of the weather classes will not change and many classification procedures also have the inherent problem of within-class variability of climate parameters (Brinkmann, 2000). Enke *et al.* (2005a) recently described a scheme to partially address this, limiting within-type variability by deriving a classification scheme for circulation patterns that optimally distinguishes between different values of regional weather elements. The scheme is based on a stepwise multiple regression where predictor fields are sequentially selected to minimize the RMSE between forecasts and observations. This is applied in Enke *et al.* (2005b) with the aim of modelling daily extremes of not yet observed magnitudes. This uses a two-stage process; first applying the circulation pattern frequencies from the model and then using regression analysis to make alterations resulting from changes in the intensity of atmospheric processes, for example increasing geopotential thicknesses.

Weather generators. At their simplest these are stochastic models, based on daily precipitation with a two-state first-order Markov chain dependent on transition probabilities for simulating precipitation occurrence, and a gamma distribution for precipitation amounts (e.g. WGEN, Wilks, 1992), although second-order (e.g. Mason, 2004) and third-order (e.g. MetandRoll; Dubrovsky *et al.*, 2004) Markov chain models have now

been developed that are better able to reproduce precipitation occurrence or persistence. Rather than being conditioned by weather patterns, variables are conditioned on specific climatic events, e.g. precipitation occurrence, with daily climate governed by the outcome on previous days. An example is the climate research unit (CRU) daily WG developed by Jones and Salmon (1995) and modified by Watts *et al.* (2004). It generates precipitation using a first-order Markov chain model from which other variables – minimum and maximum temperatures, vapour pressure, wind speed and sunshine duration – are generated. A recent development of this approach has been the linkage of a stochastic precipitation generator – the Neyman Scott rectangular pulses (NSRP) model – to the weather component developed by Watts *et al.* (2004). This is described in Kilsby *et al.* (2007b) and has been shown to improve upon the Markov chain model approach; better describing both variability and extremes within climatic time series.

WGs may also be driven by a weather typing scheme. Corte-Real *et al.* (1999) used daily weather patterns identified from principal components of MSLP to condition a WG. Improvements in the modelling of the autocorrelation structure of wet and dry days were observed when the probability of rain is conditioned on the current circulation pattern and the weather regime of the previous day. The generator simulated fundamental characteristics of precipitation such as the distribution of wet and dry spell lengths and even extreme precipitation. Other studies have also found that conditioning on circulation and low frequency variability, including sea-surface temperatures (SSTs), improves simulation results (e.g. Wilby *et al.*, 2002a).

The relative performance of different WGs was assessed by Semenov *et al.* (1998) who indicated that LARS-WG (Racsko *et al.*, 1991) is better than WGEN at reproducing monthly temperature and precipitation means across the USA, Europe and Asia due to a greater number of parameters and the use of more complex distributions. However, both were poor at modelling inter-annual variability in monthly means and reproducing frost and hot spells due to simplistic treatment of persistence. Qian *et al.* (2005) evaluated the LARS-WG and AAFC-WG (Hayhoe, 2000) WGs and highlighted differences in performance; most notably that the AAFC-WG model was better at reproducing distributions of wet and dry spells than the LARS-WG.

The major disadvantage of WGs is that they are conditioned using local climate relationships and so may not be automatically applicable in other climates, though the extent to which this limits their usefulness has not been fully tested. They also tend to underestimate inter-annual variability. Approaches have been developed to improve the simulation of variability. For example, the incorporation of a stochastic rainfall model into a WG (e.g. Kilsby *et al.*, 2007b) improves the simulation of both variability and extremes when compared to the use of a simple Markov method. Additionally, Wilby *et al.* (2002b) developed the Statistical DownScaling Model (SDSM),

a hybrid of stochastic WG and regression methods. It uses circulation patterns and moisture variables to condition local weather parameters, and stochastic methods to inflate the variance of the downscaled climate series.

COMPARISON OF DOWNSCALING METHODS

The section entitled “Overview of downscaling methods” has indicated that there are a wide range of downscaling methods that may be employed. However, the use of different spatial domains, predictor variables and predictands, and assessment criteria makes direct comparison of the relative performance of different methods difficult to achieve. This represents a problem for their application in climate change impact assessments as the differential performance of each method creates a further level of uncertainty that is difficult to quantify.

Many recent studies have compared the performance of different downscaling methods. In this section studies that have compared different statistical downscaling techniques and those that have examined the relative performance of statistical and dynamical or statistical–dynamical methods are reviewed. The Statistical and Regional dynamical Downscaling of Extremes for European regions (STARDEX) project is the first attempt to rigorously and systematically compare statistical, dynamical and statistical–dynamical downscaling methods, focussing on the downscaling of extremes, and this is discussed further in “Downscaling of extremes”.

Inter-comparison of statistical downscaling methods

Studies comparing different statistical downscaling methods are now relatively common. Most, however, investigate the downscaling of either precipitation or temperature, with few investigating the simultaneous downscaling of multiple variables.

The earliest study of precipitation downscaling methods (Wilby and Wigley, 1997), for six North American study regions compared the performance of two ANN models, two WGs – the WGEN WG based on a two-state Markov process of rainfall occurrence, and another based on spell length – and two semi-stochastic classification schemes based on daily vorticity values. Fourteen diagnostic statistics were compared, including mean precipitation, wet day probabilities, extreme precipitation amounts and wet and dry spell lengths. Model performance varied considerably for the statistics being tested. The WGs captured wet-day occurrence and amount but were less skilful for inter-annual variability; with the opposite found for ANNs. Overall, WGs proved more skilful, with ANNs the worst because of the overestimation of wet days; use of the direct output from the HadCM2 GCM was superior to one ANN model on 12 out of 14 occasions.

Indeed, ANNs have been shown repeatedly to perform poorly in the simulation of daily precipitation, particularly for wet-day occurrence (Wilby and Wigley, 1997; Wilby *et al.*, 1998; Zorita and von Storch, 1999;

Khan *et al.*, 2006) due to a simplistic treatment of days with zero amounts, although they perform adequately for monthly precipitation (Schoof and Pryor, 2001). Harpham and Wilby (2005) addressed this by using variants of ANNs which, analogous to weather generation methods, treat the occurrence and amount of precipitation separately. These ANNs were more skilful than the SDSM for individual sites but over-estimated inter-site correlations of daily amounts due to the deterministic forcing of amounts, whereas the stochastic nature of the SDSM led to increased heterogeneity and lower inter-site relations.

A more modest comparative study by Zorita and von Storch (1999) concluded that simple analogue methods perform as well as more complex methods. Analogues reproduced the mean monthly and daily statistics of winter Iberian rainfall well, producing results comparable to a CCA method, and outperformed CCA and ANNs in simulating variability. Similarly, Widmann *et al.* (2003) found that using GCM precipitation as a predictor in downscaling improved results considerably, without the use of complex statistical methods. This approach was extended by Schmidli *et al.* (2006) to the daily temporal resolution.

Comparative studies of methods for downscaling mean temperature have also been undertaken. Huth (1999) compared several linear methods for the downscaling of daily mean winter temperature in central Europe: CCA, SVD and three multiple regression models – stepwise regression of principal components (PCs), and regression of PCs with and without stepwise screening of gridded values. PCs without screening, including all PCs within the analysis, provided the greatest skill, with the other two regression methods and CCA performing with comparable skill providing a large number of predictor PCs were used. This suggests that the stepwise screening of predictor variables may be an unnecessary step. However as GCMs simulate predictors with differing accuracy, the rating of methods may have been different had downscaling been applied to a GCM control run rather than NCEP reanalyses. Note also that as winter temperatures are relatively straightforward to downscale then it is questionable as to how much variation in skill is linked to the downscaling methodology.

Benestad (2001) compared methods for downscaling monthly mean temperature: EOFs and a conventional CCA method. Using EOFs was found to be more robust with respect to predictor domain as it reduced the number of subjective choices for model set up; therefore more appropriate for downscaling global climate scenarios. To downscale maximum temperature series, Schoof and Pryor (2001) compared ANNs with regression methods for a station in Indianapolis, U.S. The ANNs proved more skilful if lagged predictors were not included.

Downscaling method comparisons for more than one variable are rare. Dibike and Coulibaly (2005) compared SDSM, and a stochastic WG (LARS-WG) for a catchment in northern Quebec. Mean daily precipitation was

simulated well by both methods; the WG better reproduced wet and dry-spell lengths, with SDSM underestimating wet-spell lengths. For maximum and minimum temperatures both models performed well, with SDSM showing a consistent cold bias and LARS-WG showing positive and negative biases in different months. For future scenarios however, the models displayed different results. SDSM simulated a generally increasing trend in mean daily precipitation amount and variability not reproduced by the WG. Similar results were obtained by Khan *et al.* (2006) in a comparison of SDSM, LARS-WG and an ANN method. SDSM was found to perform the best, with the ANN method producing the poorest results. SDSM has also been compared with the perturbation method in the Thames Valley, UK (Diaz-Nieto and Wilby, 2005). SDSM modelled monthly totals and wet day occurrence reasonably well but underestimated dry spell length. Despite the inclusion of a lagged predictor, the persistence of the precipitation process was not captured well.

Relative performance of dynamical and statistical methods

There have been few studies of the relative performance of dynamical and statistical methods in climate change impact assessment. Kidson and Thompson (1998) compared the performance of the regional atmospheric modelling system (RAMS), RCM and a regression-based method for New Zealand. Their technique used five EOFs of geopotential height and other derived variables as predictors for daily temperature and precipitation. They noted little difference in skill for daily or monthly timescales. The dynamical RAMS model has greater skill in simulating convective precipitation but overall the relative computational efficiency favoured the statistical model. Murphy (1999) drew similar conclusions for Europe using the UK meteorological office unified model RCM and a regression method. However, applying the same methods to a future scenario (Murphy, 2000) produced divergent results for the dynamical and statistical methods. Calibrating the regression equations using GCM-simulated variables rather than observations highlighted differences in the strength of the predictor-predictand relationship in the model. Mearns *et al.* (1999) also observed large differences in projections for climate change scenarios for east Nebraska despite similar skill for current climate conditions when using the RegCM2 RCM and a statistical technique based on stochastic generation conditioned upon weather types. Climate change beyond the range of the data used to condition the model was hypothesized as a possible reason for this difference.

Hellström *et al.* (2001) compared dynamical outputs from the Rossby Centre RCM (RCA1) driven by two different GCMs (HadCM2 and ECHAM4/OPYC3) with regression models based on large-scale circulation indices and including a humidity measure. All downscaling methods improved the simulation of the seasonal cycle and statistical and dynamical methods driven by ECHAM4 showed higher simulation skill. When applied to a future

scenario, differences between the two statistical methods were larger than differences between dynamical and statistical methods.

Wilby *et al.* (2000) examined the performance of statistical and dynamical methods in a mountainous catchment using the Animas basin, Colorado. Temperature, precipitation occurrence and amount were downscaled using a multiple regression method. Overall, statistical downscaling had greater skill in simulating maximum and minimum temperatures than precipitation. Uncorrected RCM monthly temperatures showed a cold bias in maximum temperature. RCM results could however be improved by providing an elevational bias correction on the raw RCM output. Similar conclusions were derived by comparing dynamically and statistically downscaled precipitation and temperature time series for three mountainous basins in Washington, Colorado and Nevada, USA (Hay and Clark, 2003).

Haylock *et al.* (2006) compared six statistical and two dynamical downscaling methods with regard to their ability to downscale seven indices of heavy precipitation for two station networks in northwest and southeast England. Generally, winter showed the highest downscaling skill and summer the lowest; skill increases as the spatial coherence of rainfall increases. Indices indicative of rainfall occurrence processes were also found to be better modelled than those indicative of intensity. Methods based on non-linear ANNs were found to be the best at modelling the inter-annual variability but these had a strong negative bias in the estimation of extremes; circumvented by the development of a novel re-sampling method. Similar results to Murphy (2000) and Hellström *et al.* (2001) were obtained when applying six of the methods to the HadAM3P model forced by two different emissions scenarios. The inter-method differences in the future change estimates for precipitation indices were at least as large as the differences between the emissions scenarios for a single method.

There have also been comparisons of statistical and dynamical downscaling methods within the seasonal forecasting field. Díez *et al.* (2005) compared their performance in downscaling seasonal precipitation forecasts over Spain from two DEMETER models: ECMWF and UKMO. As with similar studies in the climate change literature, they conclude that different methods produce better results dependent upon season and on the study region.

The performance of the direct statistical downscaling of GCM output has been compared with the use of an intermediate dynamical downscaling step before statistical downscaling. Hellström and Chen (2003) used a three-step method to downscale Swedish precipitation. RCM predictors (using RCA1) were up-scaled to GCM level (HadCM2 and ECHAM4) using a linear interpolation scheme based on the assumption that the inclusion of small-scale information in the large-scale field should have positive effects. These were then downscaled using a multiple regression model. The intermediate dynamical downscaling step improved the seasonal cycle of the

predictors but only provided a slight improvement in the seasonal cycle of precipitation; not a reasonable return on the cost of running the regional models. Wood *et al.* (2004) came to similar conclusions in a comparison of six downscaling approaches to produce precipitation and other variables for hydrological simulation. Three relatively simple statistical downscaling methods – linear interpolation, spatial disaggregation, and bias-correction and spatial disaggregation – were each applied to GCM output directly and after dynamical downscaling with a RCM. The most important aspect of use of GCM or RCM outputs was found to be the bias-correction step (as noted by Wilby *et al.* (2000) and Hay and Clark (2003)). The dynamical downscaling step did not lead to large improvements in simulation relative to using GCM output alone. This is in contrast to comparisons in the seasonal forecasting field (Díez *et al.*, 2005) where using dynamical and statistical downscaling methods in combination was found to offer an improvement over their use alone.

Downscaling performance for different climates

It is also difficult to assess the relative performance of different downscaling methods in different climatic regimes, although it would be expected that different methods would have greater skill in different climates since seasonal differences in simulation are noted in the comparison studies reviewed in the earlier sections. Indeed, in some studies statistical downscaling is conducted for each season separately because of the strong seasonal links between large-scale circulation and local climate in the mid-latitudes (e.g. Matulla, 2005). Table II lists recent statistical downscaling publications and their study regions. Commonly, choice of downscaling method is not based on any objective criteria related to either the variable to be downscaled or climatic region, with global applicability assumed (Wetterhall *et al.*, 2006), although many researchers have commented that the accuracy of downscaling methods has a geographical and seasonal component (e.g. Huth, 1999). However, an important component of any impact study should be an assessment of the conditions under which different downscaling methods can be successfully applied.

Few studies have examined whether choice of downscaling method should be based on climatic region, downscaled climatic variable or season. Indeed, the use of homogeneous climatic regions in downscaling studies is a relatively new concept, with the prospect that statistical connections between large-scale circulation and local climate should be more stable. This was recently shown to be of benefit in a region of the Tropics by Penlap *et al.* (2004). This spatial differentiation was also used in an Austrian study by Matulla (2005) where region- and season-specific combinations of predictor variables were used in downscaling. Here, the use of homogenous regions also enhanced the performance of the downscaling models, although the improvement was only slight. Regional or seasonal dependence in downscaling has been observed by other researchers. For

Table II. Examples of recent statistical downscaling studies. This list is not intended to be exhaustive but illustrative of the geographical distribution of downscaling studies and the methods employed. Abbreviations used in this table are: REG – linear regression methods, ANN – artificial neural networks, CCA – canonical correlation analysis, OTH – other, SCA – scaling methods, SVD – singular value decomposition, WG – weather generators, WT – weather typing. For predictands, T – temperature, P – precipitation, PE – potential evapotranspiration, H – humidity variables.

Authors	Statistical Downscaling Technique	Location	Predictand	Authors	Statistical Downscaling Technique	Location	Predictand
Europe							
Abaurrea and Asín (2005)	REG	Ebro valley, Spain	P	Kyselý (2002)	REG	Central Europe	T
Bárdossy <i>et al.</i> (2002)	WT	Germany, Greece	T, P	Matulla (2005)	CCA	Austria	T, P
Beckmann and Buishand (2002)	REG	Netherlands and Germany	P	Murphy (2000)	REG	Europe	T, P
Bergant and Kajfez-Bogataj (2005)	REG	Slovenia	T, P	Tatli <i>et al.</i> (2004)	REG	Turkey	P
Cawley <i>et al.</i> (2003)	ANN	North-west UK	P	Widmann and Schär (1997)	WT	Switzerland	P
North America							
Diaz-Nieto and Wilby (2005)	WG	Thames Valley	P	Qian <i>et al.</i> (2005)	WG	Canada	T, P
Enke <i>et al.</i> (2005a,b)	WT	Germany	T, P	Schoof and Pryor (2001)	ANN, REG	Indianapolis, USA	T, P
Goodess and Jones (2002)	WT	Iberian Peninsula	P	Salathé (2003)	SCA, OTH	Washington and Oregon, UK	P
Goodess and Palutikof (1998)	Automated LWT – Markov process	Guadalentin Basin, SE Spain	P	Wood <i>et al.</i> (2004)	OTH	Pacific North West USA	T, P
Hellström <i>et al.</i> (2001); Hellström and Chen (2003)	REG	Sweden	P	Widmann <i>et al.</i> (2003)	SCA, SVD	NW USA	P
Australasia							
Huth (1999, 2002)	CCA, SVD, REG	Central Europe	T	Kidson and Thompson (1998)	REG	New Zealand	T, P
Huth <i>et al.</i> (2003)	REG, WG	Central Europe	T				
Africa							
Huth (2005)	REG	Czech Republic	H	Penlap <i>et al.</i> (2004)	CCA	Cameroon	P
Jasper <i>et al.</i> (2004)	REG	Switzerland	T, P				
Kettle and Thompson (2004)	REG	Europe (high elevation)	T				

example, the downscaling skill for mean daily temperature in central Europe was found to be dependent upon station elevation, with statistical downscaling methods showing more skill at higher elevation stations, and with increased skill at continental compared to maritime sites (Huth, 1999). In contrast, Kleinn *et al.* (2005) found that for dynamically downscaled and then bias-corrected precipitation and temperature outputs there was less skill at high elevations. Similarly, seasonal and geographical differences in downscaling skill were noted by Cavazos and Hewitson (2005) who examined the

performance of NCEP reanalysis variables in the statistical downscaling of daily precipitation using ANNs and principal components analysis (PCA). Downscaling methods had the greatest skill in mid-latitude locations for the cool/dry season and showed the largest errors in the wet season. The poorest geographical performance was in the Tropics. However, this may be due to the dominance of convective processes which are harder to predict or due to poor quality of the reanalysis data in this region rather than deficiencies in downscaling method.

Goodess *et al.* (2007), in a comparison of 22 statistical downscaling methods, as part of the FP5 STARDEX project, found that downscaling performance was better in winter than summer and generally better in wetter than drier regions, probably due to the difficulty in resolving small scale processes in the driest regions. However, locally-developed methods were not found to outperform methods developed using the whole European domain, apart from for CCA. A study in China by Wetterhall *et al.* (2006) evaluated four statistical precipitation downscaling methods on three catchments located in different climatic zones, reaching similar conclusions to Goodess *et al.* (2007). They used the STARDEX extreme indices and an inter-annual variability measure to evaluate model performance using probability scores commonly used in forecasting. The study clearly showed that model performance was better in wetter geographical locations and that winter precipitation is better simulated than summer, due to its link to large-scale circulation patterns. Inter-annual variability was improved with the addition of humidity as a predictor variable.

The difficulties in GCM parameterization in some regions present additional problems for downscaling. For example, the different treatment of sea-ice was considered an important factor in interpreting downscaled temperature from three GCMs for Svalbard in the Arctic Ocean (Benestad *et al.*, 2002), leading them to conclude that statistical downscaling models may be invalidated in future climates affected by nearly ice-free Arctic Oceans.

Downscaling of extremes

Most climate change research has been framed within the context of studies of mean global climate (e.g. Jones, 1988). However, following the Intergovernmental Panel on Climate Change's (IPCC) second assessment report (Houghton *et al.*, 1996) which asked the question 'Has the climate become more variable or extreme?', increased attention has focused on climate variability and possible changes in short-term extremes of not just temperature but also precipitation (Nicholls *et al.*, 1996). Modelling evidence (e.g. Frei *et al.*, 1998) suggests that warming may lead to an intensification of the hydrological cycle and increases in mean and heavy precipitation.

Change in variability and extremes may have the largest impact on hydrological systems but extreme climate events are not easily defined. Many studies of extremes have focussed on what Klein Tank and Können (2003) refer to as 'soft' extremes, typically 90th or 95th percentile events, principally because the detection probability of trends decreases for even moderately rare events (Frei and Schär, 2001). Other studies have examined more rare events, for example, events with return periods of 50 years (e.g. Ekström *et al.*, 2005; Fowler *et al.*, 2005a). The performance of downscaling methods in representing extremes is therefore difficult to quantify as many different extreme thresholds have been assessed and, indeed, extreme climatic events in one catchment may not be of the same magnitude as those in another.

As the reliability of GCM output decreases with increases in temporal and spatial resolution, the representation of extremes is poor. Huth *et al.* (2003) found that GCMs differ in their ability to reproduce higher order moments for central European temperatures. ECHAM4 was able to reproduce skewness and kurtosis but CCCM2 failed. GCM skill for very low and high temperatures is also limited (Kyselý, 2002). Only ECHAM4 could even partially reproduce extreme summer temperatures across the European domain. CCCM2 also has problems reproducing winter temperatures due to poor soil moisture parameterization. As subgrid scale processes are more important for extreme precipitation than temperature, the modelling of precipitation extremes is also poor. This has led to the use of both dynamical and statistical methods in the downscaling of extremes.

The ability of dynamical downscaling methods to reproduce extreme climate statistics has been assessed in a few studies concentrating on precipitation. For example, the HadRM3H regional climate model was found to represent extreme precipitation events with return periods of up to 50 years well for most of the UK (Fowler *et al.*, 2005a). The main deficiencies in the model's representation of extremes are related to the treatment of orographic rainfall processes; consequently extremes in north Scotland are over estimated with the converse in eastern rain-shadowed regions. However, other RCMs were found to perform poorly in the simulation of extreme precipitation, particularly at the 1 day level (Fowler *et al.*, in press). Such assessments of model performance are crucial if they are to be applied with confidence to the prediction of future extremes under enhanced greenhouse conditions (e.g. Ekström *et al.*, 2005). However, as most uncertainty in future climate is derived from the choice of climate model and emissions scenario (Déqué *et al.*, 2007), a better understanding of the range of possible future change may be derived by comparing a number of climate models under different emission scenarios.

Even models sharing similar parameterization schemes may produce considerably different daily precipitation statistics (Frei *et al.*, 2003). Consequently, Frei *et al.* (2006) evaluated the performance of dynamical downscaling of daily precipitation extremes for the European Alps using six RCMs driven by HadAM3H boundary conditions to derive a range of estimates for future change in extremes. The models showed some skill in reproducing the 5-year return value. Furthermore, model biases for the tails of the extreme distribution were similar to those for wet-day intensities, suggesting errors in the extremes are related to the intensity rather than the occurrence process.

The ability of statistical downscaling methods to reproduce extreme climate statistics has also been assessed. The most comprehensive study was performed by Goodess *et al.* (2007) who compared 22 statistical downscaling methods using ten extremes indices, taking into account magnitude, frequency and persistence, for six European regions and Europe as a whole. For many

indices they used percentile thresholds rather than fixed values, focussing on moderate extremes. This may be of limited applicability to hydrological studies where rare events are of interest. The comparison used common data sets, calibration and validation periods, and test statistics, aiming to compare systematic differences in model performance between different seasons, indices and regions. Comparisons were also made of differences between direct methods, in which seasonal indices of extremes are downscaled, and indirect methods, in which daily time series are generated and seasonal indices calculated.

The performance of downscaling methods varied across seasons, stations and indices, making it difficult to identify a best method. However, traditional methods, such as stepwise regression, compositing, correlation analysis, PCA and CCA were more useful than novel methods such as ANNs (e.g. Harpham and Wilby, 2005); although Haylock *et al.* (2006) developed a novel approach within one of the ANN methods to output the rainfall probability and gamma distribution scale and shape parameters for each day which allowed re-sampling methods to be used to improve the estimation of extremes. Station to station differences were found to dominate index to index, method to method and season to season differences, although methods generally performed better in winter than summer, particularly for precipitation. In general, methods were found to perform better for mean than extreme indices of temperature and precipitation. Watts *et al.* (2004) presented similar conclusions for the CRU WG for extremes based on 90th and 95th percentile values. Downscaling methods were also shown to capture some processes better than others (Haylock *et al.*, 2006; Goodess *et al.*, 2007); indices representing the frequency of extremes were better reproduced than those for the magnitude of events. Occurrence processes are better captured by the downscaling process as predictors based on large-scale circulation capture the patterns prohibiting precipitation occurrence whereas event magnitude depends on smaller local-scale processes. Haylock *et al.* (2006) applied six of the methods to the Hadley Centre global climate model HadAM3P forced by emissions according to two SRES scenarios. This revealed that the inter-method differences between the future changes in the downscaled precipitation indices were at least as large as the differences between the emission scenarios for a single model. Goodess *et al.* (2007) concluded that there was no consistently superior downscaling method and recommended using a range of statistical downscaling methods, alongside a range of RCMs and GCMs, for scenarios of extremes, but that statistical downscaling methods may be more appropriate where point values of extremes are needed for impact studies.

Other researchers have also investigated statistical downscaling methods. Coulibaly (2004) used a genetic programming (GP) method to downscale local extreme (minimum and maximum) temperatures in northern Canada. The GP method was found to outperform the

SDSM approach in the simulation of temperature minima, but both methods performed similarly for temperature maxima. Similarly, Schubert and Henderson-Sellers (1997) used a PCA of MSLP fields to derive relationships between large-scale atmospheric flow patterns and local scale temperature extremes. Huth *et al.* (2003) indicated that skewness of summer maximum and winter minimum temperature distributions for the same region were well reproduced by a stochastic WG (MetandRoll). However, downscaling with a regression method for winter minima produced the right sign skewness but not magnitude whereas summer maxima had unrealistic negative skewness.

It is likely that all downscaling methods produce extremes that are too moderate compared with observed series; possibly linked to the assumption of linearity in most downscaling methods (Kysely, 2002). A comparison of a stepwise regression method with outputs of two GCMs (ECHAM4 and CCCM2) for central European 20- and 50-year return values of annual maxima by Kysely (2002) produced different results dependent on GCM. ECHAM4 GCM output was found to be better than downscaled results, whereas downscaling from CCCM improved the representation of extremes. Downscaling results were also found to be worse for annual minima compared with annual maxima as cold extremes are more strongly influenced by radiation balance and local station setting than large-scale circulation factors.

Discussion of comparative studies

Consideration of these analyses suggests that, at least for present day climates, dynamical downscaling methods provide little advantage over statistical techniques. Murphy (2000) suggests that increased confidence in RCM estimates of change will only be established by convergence between dynamical and statistical predictions, or by the emergence of clear evidence supporting the use of a single preferred method. This, however, presupposes that such a method will emerge. Comprehensive comparison studies such as STARDEX suggest that no one method is better and thus it may be more appropriate to pursue an approach based on the construction of probabilistic scenarios as discussed in "Probabilistic projections of climate change".

DOWNSCALING FOR HYDROLOGICAL IMPACT STUDIES

GCMs were not designed for the application of hydrological responses to climate change. GCM-predicted runoff is over-simplified and lacks a lateral transfer of water within the land phase between grid cells (Xu, 1999), although results vary considerably between models (Wilby *et al.*, 1999). Additionally, off-line hydrological models driven directly by GCM output have been found to perform poorly; the quality of GCM outputs precluding their direct use for hydrological impact studies (Prudhomme *et al.*, 2002). A clear mismatch exists

between climate and hydrologic modelling in terms of the spatial and temporal scales, and between GCM accuracy and the hydrological importance of the variables. In particular, the reproduction of observed spatial patterns of precipitation (Salathé, 2003) and daily precipitation variability (Bürger and Chen, 2005) is not sufficient. However improved results can be obtained by the application of even simple downscaling methods (Wilby *et al.*, 1999)

The downscaling of climate model outputs to study the hydrological consequences of climate change is now common. However, the minimum standard for any useful downscaling procedure for hydrological applications is that 'the historic (observed) conditions must be reproducible' (Wood *et al.*, 2004). The simplest methods examine hypothetical climate change scenarios by modifying time series of meteorological variables by CFs in accordance with GCM scenario results, often on a monthly basis (e.g. Arnell and Reynard, 1996; Boorman and Sefton, 1997). However, these do not allow for change in temporal variability (Kilsby *et al.*, 1998; as discussed in "Relative performance of dynamical and statistical methods") and so recent studies have used more sophisticated methods. Hydrological climate change impact studies have used dynamically downscaled output directly (Wood *et al.*, 2004; Graham *et al.*, 2007a,b), bias-corrected dynamically downscaled output (Wood *et al.*, 2004; Fowler and Kilsby, 2007; Fowler *et al.*, 2007), simple statistical approaches such as multiple regression relationships (e.g. Wilby *et al.*, 2000; Jasper *et al.*, 2004), expanded downscaling (Müller-Wohlfeil *et al.*, 2000), stochastic WGs (e.g. Evans and Schreider, 2002), statistical links to weather typing and circulation indices (e.g. Pilling and Jones, 2002), and some studies have compared more than one approach.

Comparison of downscaling approaches

The relative performance of different downscaling methods for hydrological impact assessment has been assessed by a few studies. Wilby *et al.* (2000) examined the performance of statistical and dynamical methods in mountainous areas using data from the Animas basin in Colorado. Temperature, precipitation occurrence and amount were downscaled using a multiple regression method. Overall, statistical downscaling returned better results than RCM output for daily run-off due to well timed snow pack melt. This was found to be regulated by temperature and estimates of gross snow pack accumulation rather than the sequence of individual precipitation events. The results could still be improved by using an elevation bias correction on the raw RCM output however. Similar conclusions were derived by Hay and Clark (2003) for three mountainous basins in Washington, Colorado and Nevada, U.S. However, Kleinn *et al.* (2005) found that by using bias-corrected RCM output at 56- and 14-km grid resolution they were able to reproduce monthly streamflow variability in the Rhine basin, although the model performed more poorly at high elevations. Model resolution was found to have a limited impact upon streamflow

simulation in large catchments, but may have a significant impact in small catchments.

However, in catchments where run-off is not snow-melt driven, other climatic features important for hydrological impacts may be poorly captured by statistical downscaling methods. For example, Diaz-Nieto and Wilby (2005) demonstrated in the River Thames, UK, that the SDSM underestimates mean dry-spell length. This reflects its inability to capture the true persistence of the precipitation occurrence process. Comparisons of statistical downscaling methods have also been made, with the use of different downscaling methods resulting in significantly different hydrological impacts for the same catchment (Coulibaly *et al.*, 2005; Dibike and Coulibaly, 2005). The implication of such studies for future downscaling for hydrological impacts assessment is that the means by which downscaling skill is assessed must be tuned to the particular catchment and application in question rather than simply applying standard assessment criteria.

Simple methods have also been used for downscaling and found to be effective in simulating hydrological systems. A comparative assessment of precipitation downscaling methods was undertaken by Salathé (2003) for the Yakima River, Washington. In this area many events are associated with large-scale storm systems; downscaling must reflect the physical processes accounting for precipitation. They assessed the performance of applying a local scaling factor compared to a 'dynamical scaling' method; a modification of the local scaling method that takes account of atmospheric circulation being dependent on monthly mean 1000 hPa heights. This 'dynamical scaling' produced significant improvement in simulation in the lee of the Cascade Mountains, which are not resolved in a GCM, but little difference in other areas. Both methods were able to simulate inter-annual flow variability and capture wet and dry sequences. In the simulation of monthly flows, local scaling was even able to differentiate between years with double-peaked flow and years with a single, melt-driven peak; a feature not reproduced when the hydrological model was driven with larger scale NCEP data. Thus, simple downscaling methods can provide accurate flow simulation. However, should there be significant change in future circulation, local scaling may not capture the effects.

A large U.S. modelling study 'The Effects of Climate Change on Water Resources in the West' (Barnett *et al.*, 2004) further compared different downscaling methods, including dynamical-statistical methods (Mason, 2004; Wood *et al.*, 2004), for various basins in the western U.S. (Christensen *et al.*, 2004; Dettinger *et al.*, 2004; Payne *et al.*, 2004; van Rheezen *et al.*, 2004). The most important findings are from Wood *et al.* (2004) who investigated the performance of an intermediate dynamical downscaling step before undertaking three different statistical downscaling techniques for the Columbia River Basin. They found that a dynamical downscaling step does not lead to large improvements in hydrologic simulation relative to using GCM output alone and of the

three statistical methods assessed – linear interpolation, spatial disaggregation and bias-corrected spatial disaggregation – linear interpolation is notably poor. Hydrologic simulation was found to be sensitive to biases in the spatial distribution of temperature and precipitation at the monthly level, especially where seasonal snow pack transfers run-off from one season to the next. Bias-corrected data was found to reproduce the observed hydrology reasonably well, but the monthly scale used for bias-correction failed to rectify subtle differences between modelled and observed climate. In particular, interdependencies between precipitation and temperature, e.g. the frequency of wet-warm and wet-cold winters, are not addressed by this method.

Use of multiple model outputs and emissions scenarios

There has been little research on the effect of using multiple climate model outputs and different emissions scenarios on the hydrological response to climate change, due to the lack of available comparable climate model outputs. However, the Climate Model Intercomparison Project, (CMIP2) (Meehl *et al.*, 2000) provided an opportunity to compare the impact of using different GCMs on statistically downscaled outputs, and the FP5 PRUDENCE project (Christensen *et al.*, 2007) has now provided a set of RCM experiments using ensemble runs, different RCMs, different driving GCMs and different emissions scenarios for the European region. Similar regional modelling studies are now also underway in the U.S. (NARCCAP; Mearns *et al.*, 2006) and Canada. The investigation of the use of multiple RCM outputs in hydrological impact studies has therefore so far been restricted to Europe, although the use of multiple GCMs and emissions scenarios has been investigated elsewhere.

The uncertainty introduced by using outputs from different RCMs on the hydrological response to climate change was recently investigated by Graham *et al.* (2007a) using PRUDENCE ensemble output for the Lule River in northern Sweden. Using boundary conditions from the HadAM3H and ECHAM4/OPYC3 GCMs to drive the RCMs elicited a different hydrological response; ECHAM4 boundary conditions produced higher river flow throughout the year. They concluded that the choice of GCM in providing boundary conditions for RCMs plays a larger role in hydrological change than the choice of emissions scenario. Graham *et al.* (2007b) used the same models to obtain estimates of hydrological change for river basins in northern and central Europe. Here, RCMs were able to reproduce the seasonal cycle of precipitation but most overestimated precipitation amounts. Temperature was better modelled but most models overestimated cold season evapotranspiration. In the Baltic Basin, RCMs varied in their ability to reproduce the apportioning between run-off and evapotranspiration. Better results were obtained for the Rhine, where models appeared to show a higher apportionment of precipitation to evapotranspiration than in Central Europe. Using different RCMs with the same GCM forcing and

emissions scenario again indicated that GCM forcing is a significant determinant of results.

The uncertainty introduced by choice of driving GCM was assessed by Chen *et al.* (2006) who used 17 GCMs from CMIP2 and a statistical downscaling method to produce regional precipitation scenarios over Sweden. The uncertainty in estimated precipitation change from different GCMs was larger than that for different regions. However, there was seasonal dependence, with estimates for winter showing the highest confidence and estimates for summer the lowest. Similar results were obtained by Wilby *et al.* (2006) who used three GCMs and two emission scenarios to explore uncertainties in an integrated approach to climate impact assessment; linking established models of regional climate, water resources and water quality within a single framework. The magnitude of estimated change differed depending on choice of GCM, for precipitation in particular, and differences were largest in summer months. The uncertainties introduced by downscaling from different GCMs make policy decisions difficult; the A2 run of HadCM3 produced lower river flows and greater water scarcity in summer by the 2080s leading to reduced deployable outputs and nutrient flushing episodes following prolonged droughts, however CGCM2 and CSIRO yielded wetter summers by the 2080s with increased deployable yield.

Jasper *et al.* (2004) estimated the uncertainty introduced by choice of driving GCM and emissions scenario, producing 17 future scenarios from seven GCMs and four emissions scenarios for two Alpine river basins in Switzerland. Multivariate regression models were used to relate inter-annual variations in sea-level pressure and near-surface temperature to monthly mean temperature and monthly total precipitation. The magnitude of predicted change varied between scenarios, particularly for precipitation, indicating that large uncertainties are introduced by the choice of GCM – although the choice of emissions scenario also had a discernable impact. They also demonstrated that the accurate reproduction of observed patterns of precipitation may not be as important in high elevation basins where run-off dynamics are strongly controlled by change in snow conditions and, therefore, temperature change.

Uncertainties introduced by choice of GCM have also been assessed by Salathé (2005) for a catchment in the north-western U.S. A local scaling method, derived by the ratio between observed and modelled precipitation at each local grid point (Widmann *et al.*, 2003), was applied to data from three GCMs (ECHAM4, HadCM3 and NCAR-PCM) for use in a streamflow model. Variations in model performance were noted, with down-scaled output from ECHAM4 able to capture the timing and distribution of precipitation events. However, down-scaled output from HadCM3 produced unrealistic winter precipitation variability, less high altitude snow and more winter streamflow; resulting in an earlier summer peak.

Wilby and Harris (2006) have gone one step further and assessed the 'uncertainty cascade' resulting from different sources of uncertainty in climate scenario construction. They presented a simple probabilistic framework for combining information from four GCMs, two emissions scenarios, two statistical downscaling techniques, two hydrological model structures and two sets of hydrological model parameters to assess uncertainties in climate change impacts on low flows in the River Thames, UK. Weights were assigned to GCMs according to an index of reliability for downscaled summer effective rainfall in the Thames Basin. Hydrological models and parameter sets were weighted by their performance in simulating annual low flow series. Emissions scenarios and downscaling methods were unweighted. A Monte Carlo approach was then used to explore the uncertainties in climate change projections for the River Thames, producing cumulative distribution functions (cdf) of low flows. The cdfs were found to be most sensitive to uncertainties resulting from the choice of GCM; mainly the different behaviour of atmospheric moisture amongst the chosen GCMs. This is one of the first examples of the use of simple probabilistic methods in climate change impacts studies, which will be further discussed in the next section. However, a move from the 'norm' of scenario planning to probabilistic planning may require a step-change in thinking (Dessai and Hulme, 2004).

INCORPORATING NEW DEVELOPMENTS

Probabilistic projections of climate change

Incorporating uncertainties into climate predictions is deemed necessary due to the inherent uncertainties within the climate modelling process, such as grid resolution and process parameterization, those introduced by the choice of model and emissions scenario, and the fact that different models often disagree even on the sign of changes expected in particular regions (Giorgi and Francisco, 2000).

The first probabilistic treatments of projected temperature change at the global average scale were produced by Wigley and Raper (2001), by utilizing a simple energy balance model in a Monte Carlo approach, and by Allen *et al.* (2000) by evaluating the uncertainties in the projections of a single global climate model on the basis of its ability to reproduce current climate, through an adaptation of the optimal fingerprinting approach to climate change detection and attribution. Other studies (e.g. Andronova and Schlesinger, 2001; Forest *et al.*, 2002; Knutti *et al.*, 2002; Stott and Kettleborough, 2002) considered results from only one climate model and assessed uncertainties in emissions scenario (radiative forcing), 'climate sensitivity' (the equilibrium global-mean warming for a doubling of the CO₂ level) and the rate of oceanic mixing or heat uptake.

More recently, thanks to advances in computing power and the availability of concerted experiments among the international climate modelling community, results from

multiple global climate models (e.g. Knutti *et al.*, 2003; Benestad, 2004) or multiple parameterizations of the same climate model (e.g. Murphy *et al.*, 2004; Stainforth *et al.*, 2002, 2005) have been used to generate probability density function pdfs of global warming – with most studies concentrating on defining 'climate sensitivity' (e.g. Frame *et al.*, 2005) for the prevention of 'dangerous anthropogenic interference with the climate system' (Mastrandrea and Schneider, 2004; Schneider and Mastrandrea, 2005) by providing probabilistic projections for different CO₂ stabilization levels (e.g. Knutti *et al.*, 2005). These kinds of approaches are aimed at generating climate change forecasts based on 'super-ensembles' of climate model simulations (Dessai *et al.*, 2005); an example of which is being developed in the ClimatePrediction.Net project (e.g. Stainforth *et al.*, 2005).

Most studies have concentrated on the global scale, producing pdfs of global-mean warming, but recently there has been some interest in the production of pdfs for regional-scale change (e.g. Stott, 2003; Tebaldi *et al.*, 2004b, 2005; Dessai *et al.*, 2005; Greene *et al.*, 2006; Räisänen and Ruokolainen, 2006; Stott *et al.*, 2006), with some studies aimed directly at the production of regional change pdfs for climate change impact assessment of hydrological systems (e.g. Ekström *et al.*, 2007; Hingray *et al.*, 2007a). Furrer *et al.* (2007) have gone one step further and produced pdfs at the grid point level, although this can be adapted to produce pdfs of regionally aggregated values. However, there has still been little probabilistic analysis of variables other than temperature, with few exceptions. A study by Palmer and Räisänen (2002) used 19 climate models in a probabilistic analysis to quantify the increases in probability of extreme precipitation for different regions of the world under global warming, and a study by Tebaldi *et al.* (2004b) focused on the production of pdfs for change in precipitation for land regions of subcontinental size. Additionally, studies by Räisänen and Ruokolainen (2006); Ekström *et al.* (2007) and Hingray *et al.* (2007a) produced regional pdfs of change in both temperature and precipitation using both GCM and RCM information, with Ekström *et al.* (2007) comparing the two different approaches taken by Jones (2000a,b) and Hingray *et al.* (2007a). Both methods estimate the probability distribution for change in the regional variables by combining a pdf for global temperature change with a pdf of the scaling variables (change in regional temperature or precipitation per degree of global temperature change).

The 'optimal fingerprinting' approach (Allen *et al.*, 2000, 2003; Stott and Kettleborough, 2002) derives probabilistic projections on the basis of a single model's detection and attribution studies, assuming that robust climate predictions should be model-independent and based only on objective information such as the reproduction of observed climate and recent climate change. However, the alternative multi-model methods are predicated on the fundamental belief that no model is the true model, and there is value in synthesizing projections from an ensemble, even when the individual models

seem to disagree with one another. The two published statistical methods (Tebaldi *et al.*, 2004b, 2005; Greene *et al.*, 2006) use a Bayesian approach to estimate the probability distribution of future climate, using information from past observed climate and the corresponding GCM simulated climatologies. The Greene *et al.* method combines the ensemble of models by calibrating their past trends at regional scales to the observed trends, and using the calibration coefficients (and their estimated uncertainty ranges) to derive probabilistic projections of future trends. The Tebaldi *et al.* method uses the two criteria of bias – performance with respect to current climate – and convergence – a measure of each model's agreement with the majority – from the reliability ensemble averaging (REA) method (Giorgi and Mearns, 2002, 2003). The basic tenet of REA is that the ability of a GCM to reproduce current climate (the bias) provides a measure of its reliability. However, a small bias in the simulation of current climate does not necessarily suggest that the model will reproduce future climate accurately – therefore the introduction of the convergence criterion where if the model future climate is near to the model ensemble mean climate then the model prediction is considered more reliable.

In this way both approaches differ from earlier approaches (Räisänen and Palmer, 2001; Palmer and Räisänen, 2002) where equal weighting is applied to the results from different models in the production of pdfs of global-mean warming. Indeed, these more recent approaches take the principle that non-uniform weighting may be more appropriate as models may have unequal skill or simulation capability. Lopez *et al.* (2006) compared the method of Tebaldi *et al.* and the optimal fingerprinting approach in the production of global mean temperature change pdfs. The Bayesian approach is shown to produce a narrower pdf than the optimal fingerprinting approach and shows bimodality, as two groups of climate models converge to different points. The narrower pdf from the Bayesian approach results from the model weighting criteria. The convergence criteria gives a large weight to models that cluster around one another, and increasing the number of models within the Bayesian procedure also gives rise to a reduction of uncertainty in the pdf that is not seen in the optimal fingerprinting approach (Lopez *et al.*, 2006). Furthermore, the Bayesian approach is very sensitive to the choice of models, especially in the tails of the produced pdfs, and the prior that controls the model convergence. Therefore the use of this method to produce pdfs of regional climate change is potentially problematic since the predicted pdfs are highly conditional on the formulation of the statistical model, the priors for the model variables, and the GCM sample (Lopez *et al.*, 2006). However, the use of the optimal fingerprinting method is also dependent on assumptions used to analyse the data. In particular, the assumption of linear scaling, although well justified for global temperature change, may well not be applicable to

regional scale temperature change or change in other climatic variables such as precipitation (Allen and Ingram, 2002).

All recent approaches for the production of pdfs of change contain a model evaluation stage – although different methods are used for evaluation and the model simulation of different climatic variables may be evaluated. Approaches have ranged from very complex; Murphy *et al.* (2004) used a climate prediction index with 32 components, to simpler, but have been mainly based on the assessment of descriptive statistics of climate models, such as global means, regional averages or geographical distributions. Indeed, simpler methods examining only the simulation of surface air temperature and precipitation (e.g. Wilby and Harris, 2006), but defining model skill rather than using simple means have been found to be useful in assigning regional climate change probabilities (Dessai *et al.*, 2005). However, the ability of climate models to simulate present climate is only part of the issue and the robustness of climate change response across models must also be given consideration (e.g. Giorgi and Mearns, 2002, 2003; Tebaldi *et al.*, 2004b, 2005). It remains to be seen whether probabilistic scenarios constructed using assessment criteria based on climate model ability to reproduce variability and extreme statistics, in addition to the commonly used mean, would produce better uncertainty constraints for hydrological impact studies. Indeed it is possible that for hydrological impact studies, analysis should be made of hydrological parameter sensitivity to climatological parameters prior to the selection of climate statistics for model evaluation and weighting. This would allow models that were best able to simulate those climatic variables that most affect the hydrological parameters under investigation to be selected preferentially over other models, thus potentially improving the accuracy of estimated probability distributions of change.

Therefore, which method is the most appropriate for climate change impact studies? Probabilistic methods have not been widely used in impact studies, although they have been used for short- and long-term climate, weather and hydrological forecasting for a number of years (e.g. Räisänen and Palmer, 2001; Palmer and Räisänen, 2002; Grantz *et al.*, 2005; Palmer *et al.*, 2005). These studies examine the changing probability of the occurrence of certain types of weather events under climate change, but not their impacts. Very few applications of probabilistic distributions in climate change impact assessment can be found in the literature, due to the inadequate quantification of uncertainties in regional climate change projections until very recently with the production of regional change pdfs. However, exceptions are found in the risk analysis of agricultural production (e.g. Luo *et al.*, 2005; Trinka *et al.*, 2005), and emerging studies in the use of probabilistic scenarios in climate change impact assessment on water related activities. For example, Hingray *et al.* (2007b) applied probabilistic climate

change scenarios for regional temperature and precipitation (pdfs) developed in Hingray *et al.* (2007a) to examine the impacts on the water resources of a regulated lake system in Switzerland. They concluded that the uncertainty resulting from the climate model (i.e. from the climate scenario pdf) is larger than uncertainty introduced by the hydrological model and, indeed, contributes the largest part to the total impact prediction uncertainty. Other examples include Cunderlik and Simonovic (2005) who examined the potential impacts of climate change on hydrological extremes in a Canadian river basin, Wilby and Harris (2006) who examined the impact of climate change on low flows in the river Thames, UK, Jones (2000a) and Prudhomme *et al.* (2003) who used Monte Carlo simulation to produce probability estimates of climate change impacts, and Jeong *et al.* (2005) and Georgakakos *et al.* (2005) who used multiple-member ensembles to provide probabilistic forecasts.

The EU FP6 project ENSEMBLES aims to inter-compare different methods for representing model uncertainty, some of which have been discussed here, and to link these to downscaling methods to improve the robustness of climate change impact assessments. Multi-model ensembles of GCMs and RCMs are now available and can be used to provide pdfs of change in climatic variables such as precipitation and temperature for impact assessment. Thus probabilistic methods should be considered when assessing climate change impacts on hydrological systems (Hunt, 2005). The limited use of such methods has so far proved valuable in providing increasingly robust forecasts of climate change impacts on agriculture and hydrology, through the inclusion of uncertainty estimates which can be used in the planning of adaptation measures (Dessai and Hulme, 2004). What is unclear so far is which approaches are most suitable for the construction of pdfs for regional impact studies. Should we use outputs from multiple GCMs or RCMs, or both? How do we combine the outputs to provide a pdf? Should we 'weight' model outputs? If we do then should we weight on model ability to simulate observed climate or model ability in prediction? How should the pdfs be applied to examine impacts? Can we use percentiles from pdfs or should we use a full Monte Carlo type approach? Finally, how can the pdfs be linked to downscaling methods to examine the additional uncertainties introduced by the downscaling of climatic variables to the local, impact, scale?

What is clear, however, is that probabilistic approaches may offer an improvement over currently accepted methods such as: (a) comparing the impacts resulting from using downscaled climatic variables from different GCMs and RCMs and different emissions scenarios to provide estimates of uncertainty (e.g. PRUDENCE FP5 project; Christensen *et al.*, 2007); and (b) comparing different downscaling methods to assess uncertainties introduced by the downscaling method (as discussed in "Comparison of downscaling methods"). It is hoped that some of the remaining questions can be addressed by the European ENSEMBLES and U.S. NARCCAP modelling projects

and the UK Climate Impacts Programme (UKCIP08) scenarios that are presently underway.

Pattern scaling

Pattern scaling was originally developed by Santer *et al.* (1990) and has been used in climate change scenario construction for hydrological impact studies (e.g. Christensen *et al.*, 2001; Salathé, 2005). It is particularly useful in impact studies where climate model simulations are not available for time periods of interest. The use of pattern scaling in downscaling and hydrological impact studies is widespread. For example, most RCMs are run only for a time-slice from 2070 to 2100, whereas water resource planners may be interested in climate change impacts in the 2020s. Pattern scaling works by standardising the climate change pattern (change per degree Celsius of mean global warming) for a variable X derived from a GCM over a geographic region, x , i.e. $\Delta\hat{X}(x)$, commonly by taking a regression over the full model period. Then, under the hypothesis that the response signal patterns are largely independent of the forcing patterns and are approximately stationary in time, the change pattern at any future time t can be derived as:

$$\Delta X(x, t) = \Delta T(t) \cdot [\Delta \hat{X}(x)]$$

where $\Delta T(t)$ is the global mean temperature change. This method has been shown to be valid in scaling changes in different climatic variables for different geographic regions and time periods (e.g. Santer *et al.*, 1994; Mitchell *et al.*, 1999; Tebaldi *et al.*, 2004a).

The most extensive study of pattern-scaling to date is that of Mitchell (2000, 2003). He investigated changes in mean climate rather than extremes, finding instances where pattern scaling is not appropriate due to non-linear responses to radiative forcing. He developed a method of scaling where, if the variable in question is assumed to have a Gaussian distribution, changes in the entire probability distribution may be obtained by individually scaling the mean and standard deviation. If, as in the case of precipitation, the variable follows a non-Gaussian distribution it may be possible to estimate change by re-scaling the distribution and relating it back to the mean and standard deviation of the Gaussian distribution. This improves upon the previous assumption that change in all climate variables can be scaled linearly with temperature change, which makes little sense for change in extremes and changes in variables other than temperature.

Recent transient RCM simulations by the Rossby Centre covering the full time period 1961–2100 (Erik Kjellström, personal communication) show that even the simplest pattern-scaling in time works well for many variables for large parts of the year. However, in some parts of northern Europe in the transition seasons of autumn and spring, the pattern scaling method fails as the regional changes do not follow the increase in global mean temperature. This suggests that the assumption that regional change occurs at the same rate as global change

is not true in all cases and pattern scaling should be used with caution.

Despite this, pattern scaling has been demonstrated to be of use in probabilistic studies where regional climate change scenarios can be constructed for emissions scenarios where climate model simulations are not available by combining GCM-simulated patterns with simple climate model results (e.g. Dessai *et al.*, 2005).

Downscaling multiple climatic variables

For hydrological impact studies of climate change, at a minimum the climatic variables of temperature and precipitation must be downscaled to provide inputs of precipitation and potential evapotranspiration to hydrological models. It is important however, to preserve the correlation between the different downscaled variables. The direct use of RCM data in impact studies, or use of bias-corrected RCM data (e.g. Wood *et al.*, 2004; Fowler and Kilsby, 2007; Fowler *et al.*, 2007; Graham *et al.*, 2007a,b etc.), preserves the physical correlation between precipitation and temperature. However, most statistical downscaling studies have considered the downscaling of only one of these variables, simulating change in the other variable in an ad-hoc fashion but normally by the simple application of climate CFs to observed climatic data time series (e.g. Kilsby *et al.*, 2007a). It is important for more accurate assessment of potential hydrological impacts that more complex approaches, where a number of climatic variables are downscaled simultaneously from climate model information, are developed. This section discusses recent developments in such approaches using statistical downscaling techniques, usually based on the 'WG' type approach (Wilks and Wilby, 1999) discussed in "Weather generators".

The SDSM approach (Wilby *et al.*, 2002b) has been widely used for the generation of multiple climatic variables in hydrological impact assessment. However, the basic approach simulates climatic variables only at single sites, like many similar approaches (e.g. Easterling, 1999; Kyselý and Dubrovsky, 2005). For hydrological impact assessment, particularly in large basins, the downscaling of multiple climatic variables at multiple sites is needed. A multi-site approach has been further developed by Wilby *et al.* (2003) by extending the SDSM approach using conditional re-sampling of area-averages to produce daily time series at multiple sites, but for precipitation only. Other re-sampling approaches for the simulation of precipitation and temperature series at multiple sites have been developed (e.g. Palutikof *et al.*, 2002; Leander *et al.*, 2005), sometimes conditional on atmospheric circulation patterns (e.g. Beersma and Buishand, 2003). However these have not, as yet, been applied for climate change impact assessment.

Similarly, WGs such as the CRU WG developed by Jones and Salmon (1995), and modified by Watts *et al.* (2004) are also presently restricted to single site applications, as described in "Weather generators". Kilsby *et al.* (2007b) describe a recent development of this approach,

which improves the simulation of daily extreme values and variability due to the implementation of the third-order moment – or coefficient of skewness – within the NSRP model. The model can be used to simulate perturbed climates by applying CFs obtained from climate models to the precipitation statistics of mean, variance, proportion dry days, skewness, and lag-1 autocorrelation within the NSRP model, and the mean and standard deviation of temperature within the WG. Although the model has been calibrated at a 5-km resolution throughout the UK, it is only applicable for the simulation of climatic variables for single-site or small catchments (Kilsby *et al.*, 2007b). The model is presently undergoing extensions to simulate at an hourly resolution and at multiple sites. However, it is still unclear how to improve WG approaches to enable them to capture low frequency variability, such as multi-season droughts, which may be crucial for hydrological impact assessment.

Techniques have been developed for the generation of multiple linked climatic variables at multiple sites. However, these are not generally applied at the daily time scale needed for most hydrological impact assessments. For example, Huth and Kyselý (2000) used a multiple linear regression method to estimate monthly mean temperature and precipitation for two catchments in the Czech Republic from large-scale circulation height and thickness fields. Techniques have also been developed on reanalysis data, but not applied to climate model outputs to assess future climate change impacts. Gangopadhyay *et al.* (2005), for example, developed the K-nearest neighbour approach of Lall and Sharma (1996) to downscale the NCEP 1998 medium-range forecast output. The method uses an analogue-type approach to identify days similar to a given feature vector and, using EOFs, identifies a subset of days with similar features. These are then randomly sampled to generate ensembles. This approach has been used to generate hypothetical climate scenarios (Yates *et al.*, 2003). However, it is unclear how this type of analogue approach could be used together with outputs from GCMs or RCMs to develop climate change scenarios for hydrological impact studies unless the assumption is made that appropriate analogues exist in the training set for future climate states.

SUMMARY AND NEXT STEPS

Example application linking probabilistic scenarios to a downscaling method

This section proposes a method that links probabilistic climate change scenarios to a WG downscaling method for hydrological impact studies. This is an illustration of one possible way that downscaling may be embedded within a probabilistic framework and does not address statistical versus dynamical downscaling uncertainty. A flow diagram of the main steps in the methodology is shown in Figure 1.

We use precipitation and temperature data from six RCMs driven by boundary conditions from two different Global Climate Models from the FP5 PRUDENCE

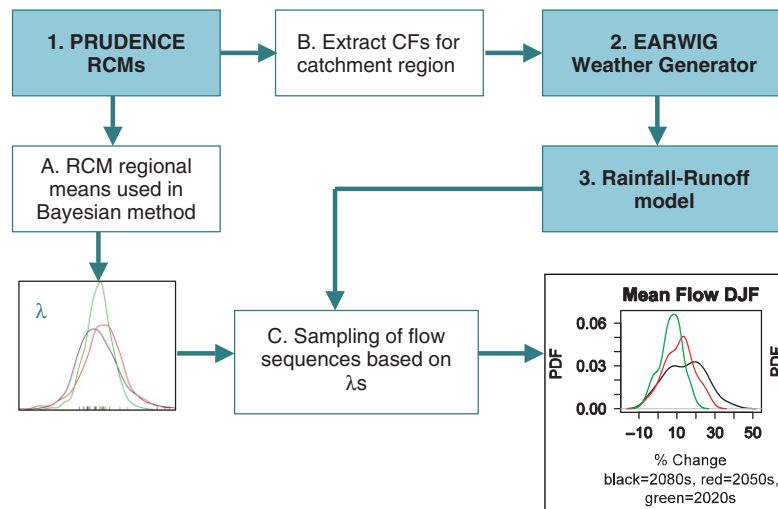


Figure 1. Flow diagram showing steps in the methodology used to produce probabilistic scenarios of climate change impacts for the River Eden. This figure is available in colour online at www.interscience.wiley.com/ijoc

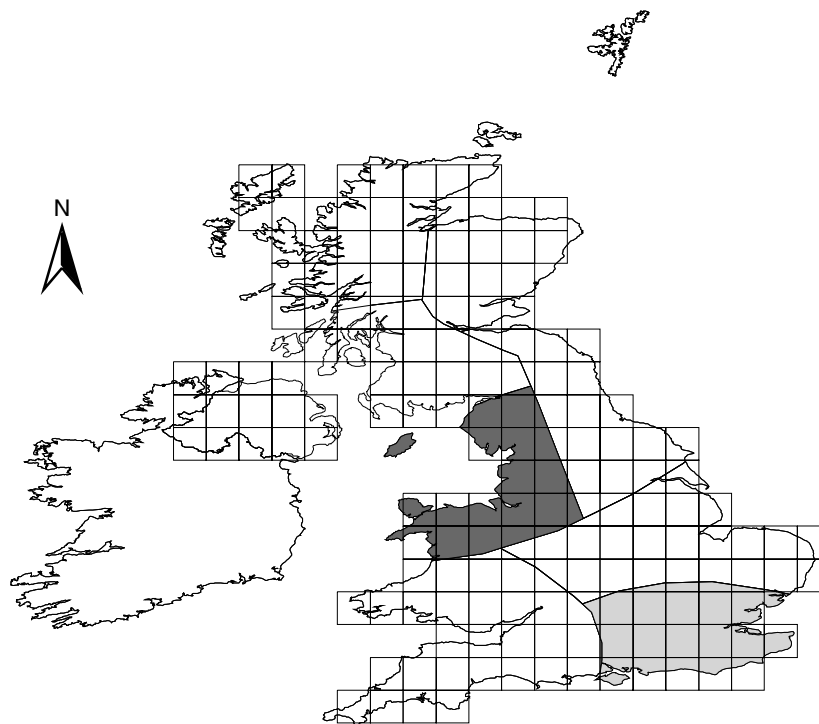


Figure 2. Map showing re-gridded RCM cells on the CRU 0.5° by 0.5° grid for the UK. The two regions used in this study are shown in grey: North West England (dark) and South East England (light).

project to provide a control (1961–1990) and future (2071–2100) scenario. Each model was re-gridded onto a common 0.5° by 0.5° grid to allow comparisons, with the future scenario based on the SRES A2 emissions scenario (Figure 2). The models used are detailed in Table III.

A Bayesian scheme (as in Tebaldi *et al.*, 2004b, 2005), described in “Probabilistic projections of climate change”, is used to fit a pdf of change in temperature and precipitation for regions in the north-west and south-east of the UK using area-averages of the model output at grid cells contained in that region. The Bayesian method assumes uninformative (i.e. diffuse) prior probability

distributions for all the unknown quantities of interest, among which the most relevant for this analysis are current and future regional average temperature and precipitation (for each season and under a specific scenario) and climate models’ reliabilities. The data from the regional models and from observation is incorporated through Bayes’ theorem, in order to derive posterior pdfs for all the unknown quantities. From these posterior pdfs, the distribution of temperature and precipitation change can be straightforwardly derived. These are shown in Figures 3 and 4.

Although the estimation of pdfs of temperature and precipitation change based on the six RCMs generates

Table III. Selection of PRUDENCE Regional Climate Models used in this study. The AquaTerra acronyms are adopted here to provide an easier understanding of the format of each model run. The first part of each acronym refers to the RCM and the second to the GCM data used to provide the boundary conditions. Scenario simulations have the further suffix A2.

	RCM	Driving Data	PRUDENCE Acronym	AquaTerra Acronym
Danish Meteorological Institute (DMI)	HIRHAM	HadAM3H A2	HC1	HIRHAM_H
		ECHAM4/OPYC A2	HS1	HIRHAM_H_A2
			ecctrl	HIRHAM_E
Swedish Meteorological and Hydrological Institute (SMHI)	RCAO	HadAM3H A2	ecscA2	HIRHAM_E_A2
			HCCTL	RCAO_H
		ECHAM4/OPYC A2	HCA2	RCAO_H_A2
			MPICTL	RCAO_E
Hadley Centre – UK Met Office	HadRM3P	HadAM3P A2	MPIA2	RCAO_E_A2
			adeha	HAD_P
Météo-France, France	Arpège	HadCM3 A2	adhfa	HAD_P_A2
			DA9	ARPEGE_C
			DE6	ARPEGE_C_A2

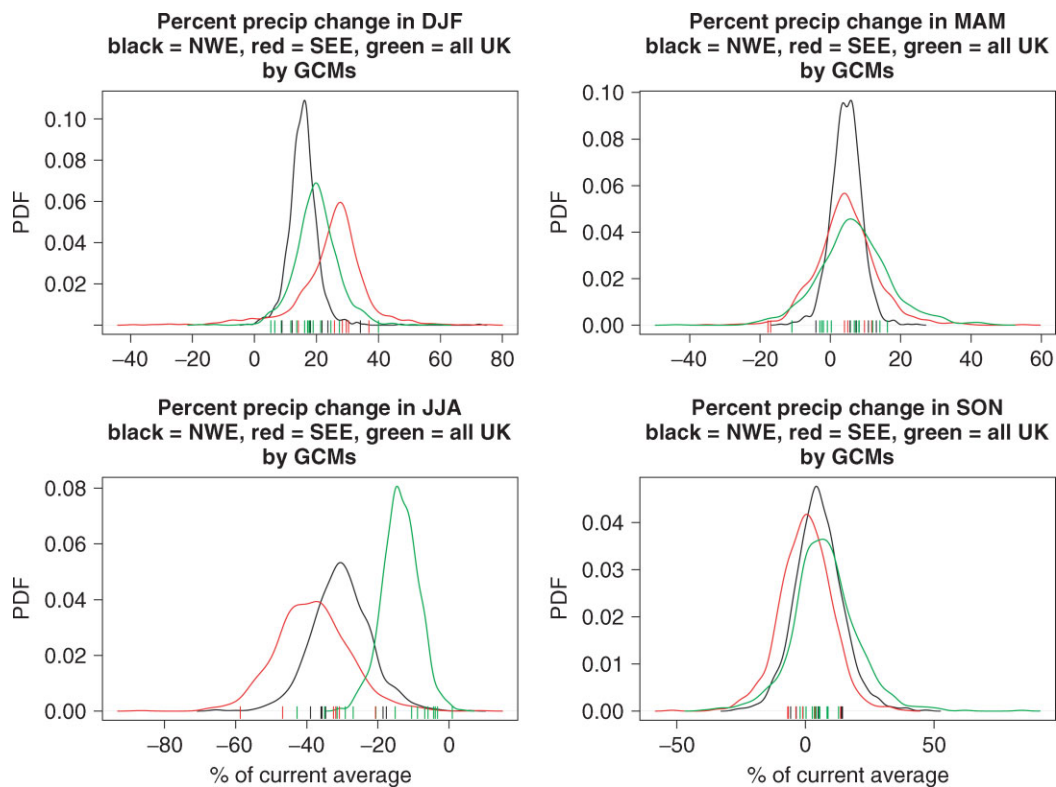


Figure 3. Percentage seasonal change in precipitation projected under the 2070–2100 SRES A2 emissions scenario from the control period, 1961–1990, for two UK regions: North West England (black) and South East England (red). Notations for seasons are winter (DJF, December to February), spring (MAM, March to May), summer (JJA, June to August) and autumn (SON, September to November). The projection from the IPCC AR4 suite of 14 GCMs for all UK is shown in green for comparison.

sensible results, being designed to handle an arbitrary number of data points, one caveat of the Bayesian approach is that independence is assumed between ensemble members. This may be an oversimplification, as many climate models share internal numerical and parameterization schemes. Here in particular, as some RCMs are driven by lateral boundary schemes from the same GCM, this assumption may be even less defensible

and generate overly confident estimates of the uncertainty in the changes. However, results for two UK regions from the set of RCMs are comparable to the results for the entire UK from the set of IPCC-AR4 GCMs (Figures 3 and 4), suggesting that these RCMs are in fact representative of the IPCC-AR4 results. There is noticeable agreement in the location and even the spread of the pdfs for all seasons but summer, where the GCM

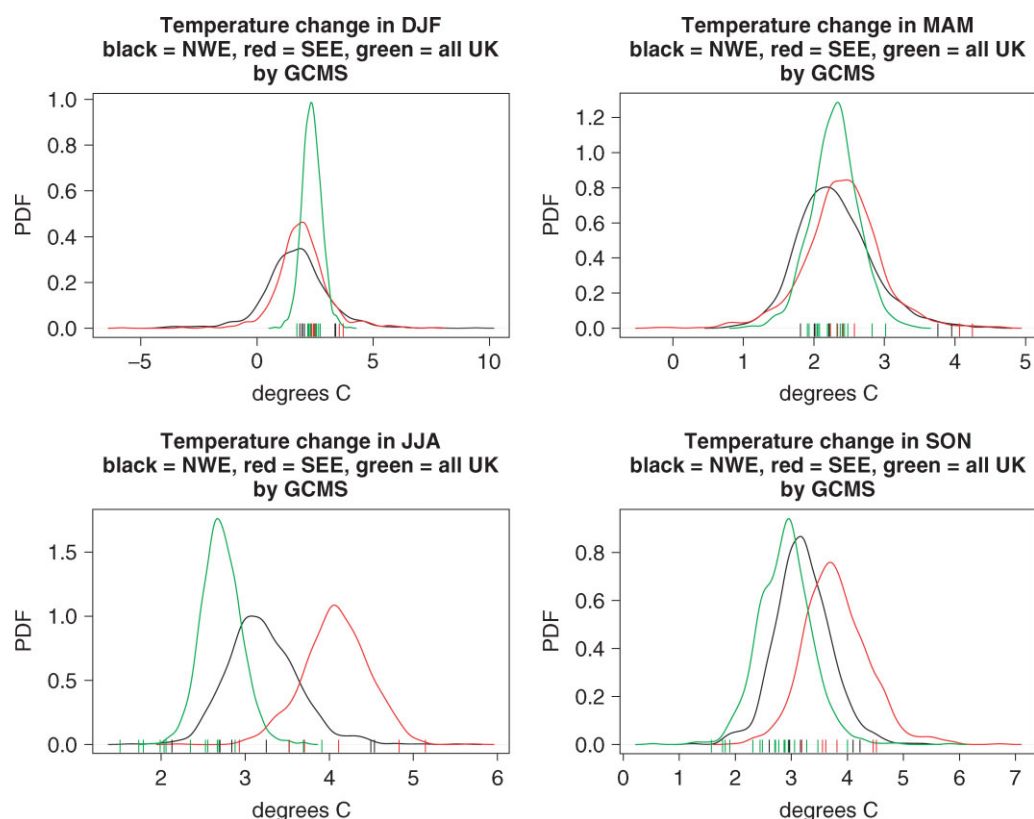


Figure 4. Percentage seasonal change in temperature projected under the 2070–2100 SRES A2 emissions scenario from the control period, 1961–1990, for two UK regions: North West England (black) and South East England (red). Notations for seasons are winter (DJF, December to February), spring (MAM, March–May), summer (JJA, June–August) and autumn (SON, September–November). The projection from the IPCC AR4 suite of 13 GCMs for all UK is shown in green for comparison.

Table IV. Seasonal mean λ values for the North West England region of the UK for the six selected regional climate models for (a) precipitation, and (b) temperature. Note that λ values for all six models have been standardized to add to 1.0 for each season.

	ARP_C	HAD_P	HIRH_E	HIRH_H	RCAO_E	RCAO_H
DJF	0.07	0.19	0.25	0.26	0.08	0.15
MAM	0.08	0.05	0.11	0.23	0.26	0.27
JJA	0.15	0.06	0.16	0.23	0.18	0.22
SON	0.11	0.11	0.21	0.20	0.14	0.23

(b)

	ARP_C	HAD_P	HIRH_E	HIRH_H	RCAO_E	RCAO_H
DJF	0.23	0.22	0.12	0.19	0.11	0.13
MAM	0.17	0.22	0.15	0.26	0.09	0.10
JJA	0.08	0.18	0.09	0.25	0.16	0.25
SON	0.12	0.23	0.16	0.24	0.13	0.12

results are significantly less warm and less dry than the RCM results.

The model-specific reliabilities parameters, λ s in Table IV, are also estimated jointly with the other relevant quantities, as a function of each model's performance in reproducing current climate (1961–1990) and each model's agreement with the ensemble consensus for future projections. Their interpretation as model weights is thus straightforward, and, once standardized,

they can be applied as such in the downscaling step. Since the Bayesian analysis produces joint posterior pdfs of the lambda parameters, we use the posterior mean of the distribution as our best estimate of model weights (Table IV).

A WG approach was then used to downscale these regional distributions of change in temperature and precipitation to the catchment scale. This advanced WG, the Environment Agency Rainfall and Weather Impacts

Generator (EARWIG), incorporates a stochastic rainfall model based on the NSRP model (Cowpertwait, 1991) and a WG model based on regression relationships between daily weather variables and daily rainfall (and their autocorrelative properties, see Kilsby *et al.*, 2007b for details). The model is able to generate synthetic daily climate data for any 5 km grid cell or small catchment in the UK, for current climate (1961–1990) and future climate scenarios.

In the Kilsby *et al.* (2007b) approach, daily outputs from the HadRM3H RCM are used to derive factors of change (CFs) from the current climate state based on the UKCIP02 climate change scenarios (Hulme *et al.*, 2002). Within the NSRP part of the WG, five multiplicative CFs are used to change future rainfall statistics of mean daily rainfall, proportion dry days, variance of daily rainfall, skewness of daily rainfall and lag-1 autocorrelation. Within the WG part, additive/multiplicative factors are derived for change in temperature mean/variance; other weather variables are dependent on rainfall and temperature and these relationships are assumed to remain constant under climate change. The CFs are based on the outputs from HadRM3H and vary at a spatial resolution of 50 km across the UK. More information on the methodology used to derive CFs can be found in Kilsby *et al.* (2007b).

Here, we use the same methodology as Kilsby *et al.* (2007b) to derive a suite of CFs for each of the six PRUDENCE RCMs for a catchment in northwest

England, the Eden. These CFs are then applied within EARWIG to produce 100 30-year synthetic climate sequences for each RCM for the SRES A2 2080s scenario and for a baseline scenario, 1961–1990. The 30-year sequences were chosen to reflect the length of the RCM control and future integrations. For the baseline, the simulation is based on observed climate statistics.

A simplified version of the Arno hydrologic model (Todini, 1996) was calibrated for the Eden catchment down to Temple Sowerby; with an area of 616 km² and elevation range from 950 to 92 m (Figure 5). Rainfall in the catchment is predominantly from the westerly quadrant, producing an average annual rainfall over 1961–1990 of 1146 mm and a mean flow of 14.44 m³ s⁻¹. Model calibration was achieved in a split-sample approach using a method of genetic algorithm optimization, the shuffled complex evolution method for global optimization (SCE-UA), developed by Duan *et al.* (1992), with the Nash and Sutcliffe efficiency measure (CE) (Nash and Sutcliffe, 1970) and water balance (WB) used as optimization criteria. Historic rainfall, PE and flow data from the period 1976–1990 was used in calibration (CE = 0.73, WB = 1.08) and for validation from 1991–1998 (CE = 0.78, WB = 1.08). More detail on the model is given in Fowler and Kilsby (2007).

Using the calibrated hydrologic model, 100 synthetic 30-year daily flow sequences were then produced for the baseline and each of the six possible futures (one from

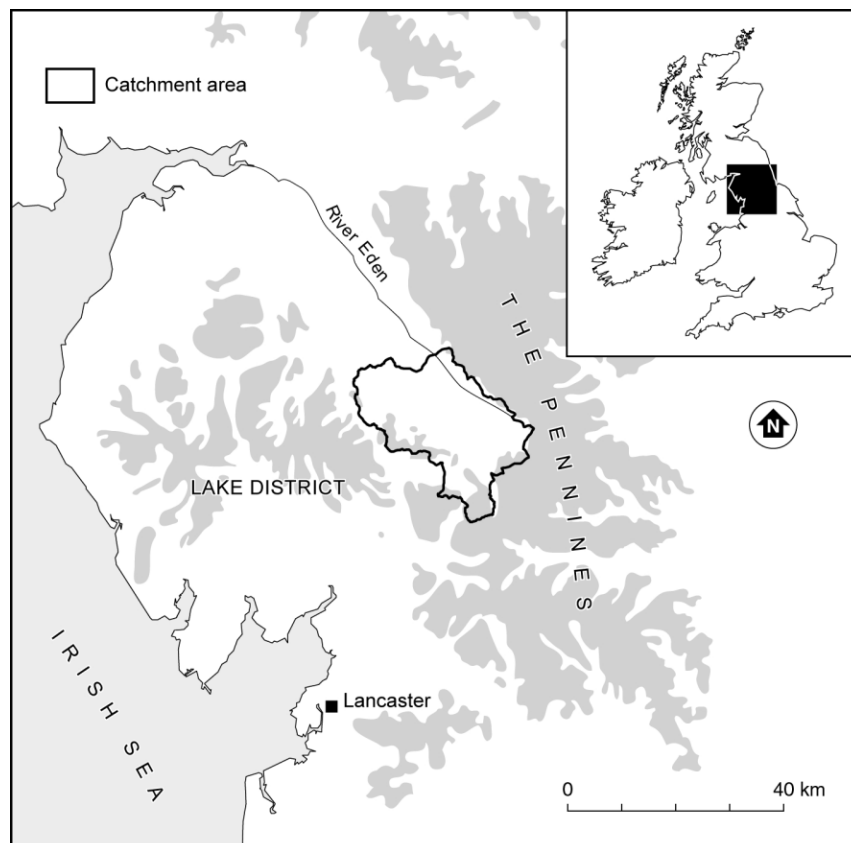


Figure 5. The Eden catchment in northwest England (to Temple Sowerby).

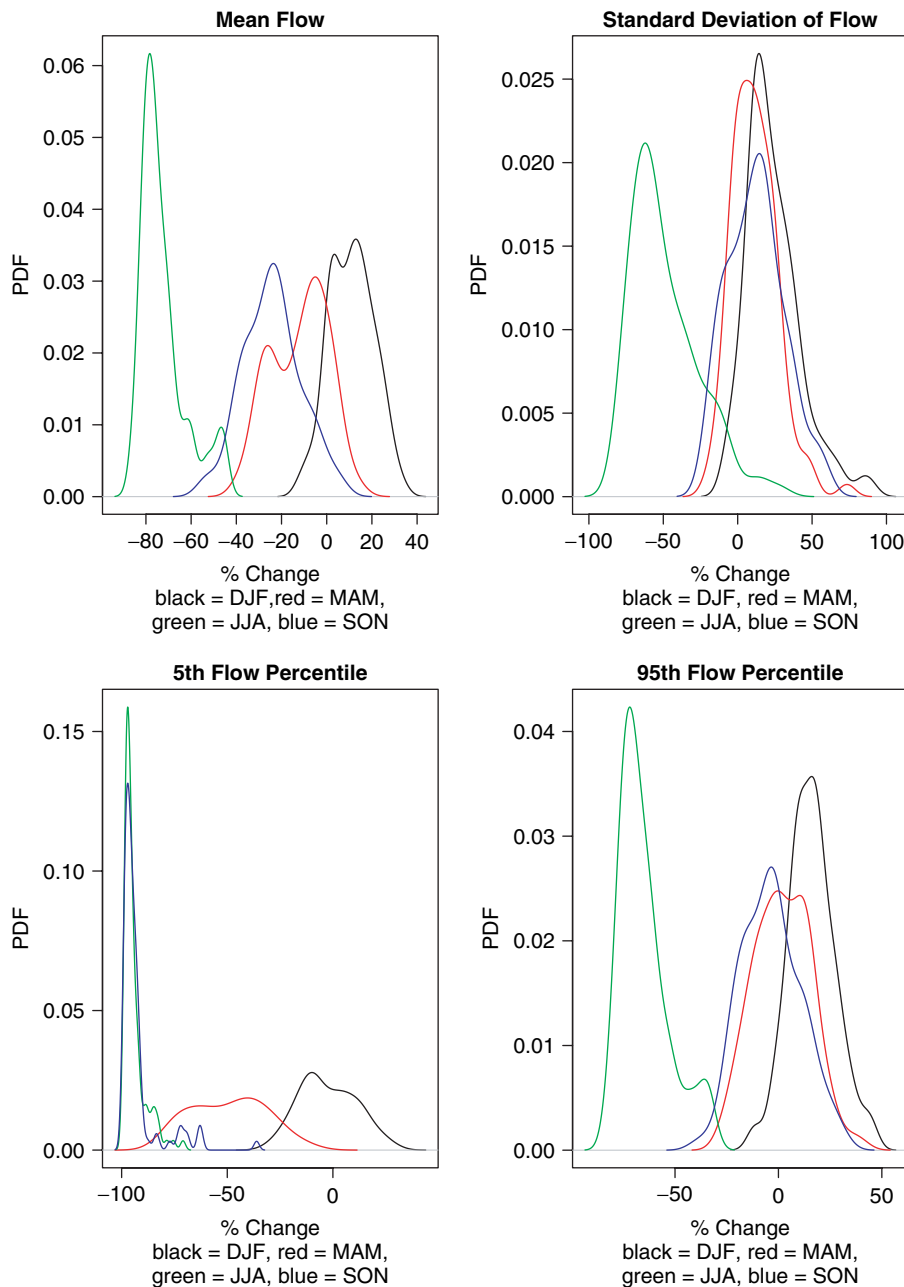


Figure 6. Projected percentage changes in flow statistics for the 2070–2100 period (change from 1961–1990): (a) mean flow, (b) standard deviation of flow, (c) 5th percentile of flow, (d) 95th percentile of flow. Seasonal change probabilities are given for winter (DJF, black), spring (MAM, red), summer (JJA, green) and autumn (SON, blue).

each RCM) using the simulated climate data from EAR-WIG as inputs. Seasonal statistics were then calculated for each 30-year sequence; providing mean, standard deviation, 5th and 95th flow percentiles.

To produce probabilistic estimates of future change in flows in the Eden catchment, the λ values produced by the Bayesian scheme (Table IV) for precipitation were then used to provide a model-specific weighting for each season, where the standardized λ values for all models within one season add to unity. The λ value for each RCM provided a weight for each model; directly proportional to the number of 30-year flow sequences from each model used in the analysis (where 0.01 = ten 30-year sections). For each baseline and future scenario

flow sequence, seasonal statistics of absolute and relative change for the mean, standard deviation, 5th and 95th flow percentiles were then calculated. When all sequences are added together this gave 1000 'change statistics' for each flow statistic (mean, standard deviation and 5th and 95th percentiles). Finally, a kernel density function was fitted to the 1000 'change statistics' to give a smooth pdf of change for each flow statistic.

This method provides a probabilistic estimate of future climate change impacts on flow statistics in the Eden catchment based on a weighted multi-model approach. Figure 6 shows the seasonal pdfs of change in different flow statistics for the Eden catchment in northwest England for 2070–2100. However, these probabilistic

estimates of future change are based on the results from multiple RCMs for only one emissions scenario and one hydrological model parameterization. Improvements to the method may include an assessment of the further uncertainty introduced by emissions scenario, and the impact of the structure and parameterization of the hydrological model upon the estimates. It may also be noted that the outcome may be different if the lambda parameters were obtained for the model-specific reliability of the resultant Q5, Q50 and Q95 statistics rather than the regional average inputs.

Discussion and recommendations

This paper has provided an assessment of recent developments in downscaling methodologies through a review of the downscaling literature. In recent years there have been a plethora of new studies in downscaling. Figure 7 shows the large growth in this literature since the mid-1990s; with a total of 64 publications listed on the ISI Web of Knowledge for 'downscaling and climate' in 2005. There has, however, been more restricted growth in publications where downscaling methods are used to examine impacts; only ~30% of all downscaling studies. Indeed, in the hydrological sciences there has been little additional use of these methods to examine impacts since their inception; a search on the ISI Web of Knowledge reveals only ten publications in 2005 when the search criteria 'downscaling and hydrol*' and impact' are used. This is despite the large increase in publications on downscaling methods themselves.

In this review paper we have tried to address five questions. In this discussion section we will try to answer these questions in turn, as well as cross-matching our research imperatives against those of Leung *et al.* (2003b) (Table V) to see if there has been any progress in the interim:

1. What more (if anything) can be learnt from downscaling method comparison studies?

Systematic efforts within the research community since the mid-1990s have provided a large amount of information on the advantages and disadvantages of different downscaling methods. This allows us to make statements about the downscalable skill of different climatic variables with some certainty. In general, temperature can be downscaled with more skill than precipitation, winter climate can be downscaled with more skill than summer due to stronger relationships with large-scale circulation, and wetter climates can be downscaled with more skill than drier climates. However, direct comparison of the skill of different methods remains difficult due to the range of climate statistics that have been assessed in the literature, the large range of predictors used, and the different ways of assessing model performance. Although we now know the theoretical strengths and weaknesses of each downscaling method, where systematic inter-comparisons have been made, e.g. STARDEX, no single best downscaling method is identifiable.

Simple statistical downscaling methods seem to perform as well as more sophisticated methods in reproducing mean characteristics; if the downscaling of mean climate is the main objective then the effort required to use more sophisticated techniques is probably not warranted by the additional downscaling skill provided. However, the importance of reproducing mean climate characteristics is dependent on the impact to be investigated, and for many applications reproducing climate statistics such as extremes may be just as important. In such instances, a selection of downscaling methods should be used as, even when methods are able to downscale to the observed 'control' climate with a degree of consistency, application to future climate model output may produce very different future projections. However, the choice of driving-GCM generally provides the largest source of uncertainty in downscaled scenarios.

Perhaps just as important as the choice of downscaling method is the choice of predictor variables

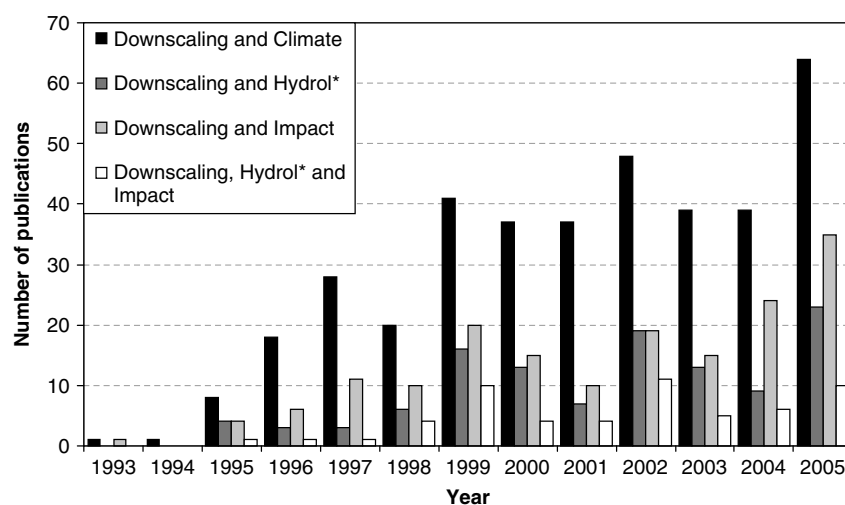


Figure 7. Analysis of downscaling literature using the ISI Web of Knowledge: 'downscaling' included in title or abstract.

Table V. Summary table of research imperatives in regional climate change research (from Leung *et al.* (2003b)).

 Recommendations

Downscaling

- We need to develop physics parameterizations for higher spatial resolution global or regional climate models. Such parameterizations may be scaleable for applications at different spatial resolutions. Regional climate models can be used as test beds for such development.
- Coordinated intercomparison and diagnostics of models (GCMs, RCMs, and SD) are needed. This will require an infrastructure for experimental protocols and community participation.
- We need to quantify predictability at the regional scale. Climate variability increases with spatial resolution, but perhaps only up to a point. Regional predictability may increase or at least be similar to larger-scale predictability at a certain regional resolution.
- Different ways of generating ensemble simulations with RCMs need to be explored to improve signal to noise or estimating uncertainty.

Evaluation and Diagnostics

- Need to further develop regional observational datasets. Downscaling may be useful to fill data gaps. Communications between the modelling and data communities needs to be improved.
- Inter-variable relationships, higher-order statistics (e.g., frequency of extreme and variability), and teleconnections (relationships and integral constraints between large-scale and regional scales) need to be more widely used to measure downscaling skill.
- Evaluating circulation is an alternative way to evaluate surface variables. Downscaled climate in process models (e.g. hydrologic models) and secondary variables should be used more often in evaluating downscaling techniques and value.
- Archiving of higher temporal frequency outputs is needed for more detailed evaluation.

Applications

- Regional climate information needs to be easy to obtain, use, and validate.
- To produce more realistic future climate scenarios for impact assessment, use of realistic driving forces and more complete representation of climate components are needed. Climate system models are moving toward incorporating biogeochemistry and lakes. Regional prediction of complex physical and socioeconomic systems is also needed for integrated assessment.
- We need to involve stakeholders in determining the resolution of regional climate information required in different impact assessments.
- Other applications of regional climate information, such as storm surges or air quality, should be tested.

Overall

- Impact assessments in the past required patching together isolated modelling, diagnostics, analyses, and assessment studies with disparate goals. We need coordinated end-to-end prediction systems to test the whole approach of impact assessment.
 - Seasonal prediction is a useful framework for assessing the added value of downscaling because results can be evaluated with observations. Projects utilizing various approaches for seasonal prediction can show whether downscaling can improve accuracy in addition to providing greater spatial detail. Such experiments could lead to seasonal forecasts for various applications. Improvements in seasonal predictions using the super-ensemble approach demonstrate the value of utilizing as many models as available.
 - Funding agencies tend to support model applications or projects that produce predictions more than development or evaluation of models that are used in making predictions. This problem needs to be addressed.
 - All downscaling techniques have been shown to be valid and produce useful results. Research on the various downscaling methods should proceed along parallel paths to the limit of funding availability.
-

in statistical schemes, particularly for the downscaling of precipitation. Failure to incorporate predictors that account for physical interactions in the climate system will produce poorly downscaled climate whichever method is employed. Additionally, a fundamental lack of predictability makes some climates difficult to downscale. This is not necessarily due to the applicability of statistical methods but rather due to a lack of good quality observational data or poor climate model parameterization, possibly due to the dominance of convective processes, in these areas. Considerable gaps consequently remain in the global distribution of downscaling studies with most studies concentrating on the mid-latitudes. Tropical and

high-latitude locations have been relatively overlooked to date.

This review suggests that we now understand the strengths and weaknesses of different standard downscaling techniques. Coordinated inter-comparisons of downscaling methods and their ability to reproduce higher-order statistics, two of the recommendations of Leung *et al.* (2003b), have been undertaken. We would recommend therefore, that additional comparison studies are not needed. Instead, the researcher should concentrate on defining the climatic variables that it is necessary to accurately downscale for each different impact application. The needs of hydrological impact studies differ considerably and in standard downscaling

studies little consideration has been given to the most appropriate downscaling method to use for a particular application. Different climates, different seasons and different climatic variables may be more accurately downscaled by using more appropriate downscaling methods. We would therefore recommend the incorporation of a 'sensitivity' step as standard in all downscaling procedures; where the climatic variables that have the largest impact on the hydrological system are first identified, and the most appropriate downscaling method then determined.

2. Can dynamical downscaling contribute advantages that can not be conferred by statistical downscaling?

Applications of dynamical downscaling in the literature have consistently shown that outputs from RCMs cannot be used in impact studies without first applying a 'bias-correction' on observations. Despite this need for additional correction before their use in impact studies however, several studies have now illustrated how dynamical downscaling provides 'added value' for the study of climate change and its potential impacts. Regional climate change signals can be significantly different from those projected by GCMs, particularly in regions with complex orography which remain unresolved by GCM simulations. RCMs are able to better capture the effects of orographic forcing and rain-shadowing and provide improved simulation of higher moment climate statistics; hence providing more plausible climate change scenarios for extreme events and climate variability at the regional scale. Despite this there is still a need, as stated by Leung *et al.* (2003b), for more research examining the statistical structures of climate signals at different spatial scales to establish whether predictability of the climate system is improved by regional modelling.

Leung *et al.* (2003b)'s review also suggested that 'to develop credible high-resolution climate simulations for impact assessment, a logical approach is to use multiple GCMs with multiple ensembles and force multiple RCMs'. The production of ensemble RCM simulations driven by multiple GCMs is now possible and has been achieved within large research projects such as PRUDENCE, ENSEMBLES and NARCCAP. This has enabled improvements to the realism of control simulations and the study of future change in extreme events. The use of ensemble simulations is very important to establish the statistical significance of changes associated with events that have low probability of occurrence. However, it is still uncertain how large an ensemble of integrations is needed to quantify these uncertainties. Alternative approaches suggested by Leung *et al.* (2003b) include 'factor analysis', where ensembles of results are produced by perturbing, for example, initial and lateral boundary conditions, perturbing model parameters, using different models, bootstrapping and re-sampling techniques. Investigations are now routinely made in the UK, by both the Hadley Centre and climateprediction.net, into the uncertainties associated with changes to initial conditions and model parameterization,

and simple statistical emulators have been developed to provide representations of outputs from complex GCMs; thus assessing further parameter combinations. This has allowed the assessment of uncertainties resulting from various sources.

Applications to geographically diverse regions and model inter-comparison studies have allowed the strengths and weaknesses of dynamical downscaling to be better understood. This has recently proliferated their use in impact studies. However, despite this recent increase, hydrological impact studies using regional climate model information are still relatively rare. Leung *et al.* (2003b) suggested that there is a 'need to involve stakeholders in determining the resolution of regional climate information required in different impact assessments'. The lack of this in the past may be one of the reasons for the poor uptake of downscaling methods by the hydrological impacts community. However, many initiatives are now rising to this challenge. For example, the next set of climate change scenarios for the UK, UKCIP08, are being developed under a steering group comprised of stakeholders among others.

3. Can realistic climate change scenarios be produced from dynamically downscaled output for periods outside the time period of simulation using methods such as pattern scaling?

Many studies have shown that realistic climate change scenarios can be produced from dynamically downscaled output for periods outside the time period of simulation. For regional temperature response over the next century, pattern scaling provides a reasonable estimate. Recent transient RCM simulations for 1961–2100 have shown that simple pattern-scaling works well for many variables for large parts of the year. However, due to the non-linear responses to radiative forcing by other climatic variables, such as precipitation, pattern scaling on linear temperature change makes little sense. Therefore, given that only temperature can be reliably pattern-scaled, the approach offers limited scope for studies of transient climate change, except perhaps where snowmelt is the dominant driver of hydrological change.

The production of many transient RCM simulations in the ENSEMBLES project will allow further investigation of this issue, and possibly even obviate the need for pattern scaling in future studies. Nonetheless, pattern scaling is the best currently available method for the production of climate change scenarios for periods outside the time period of simulation when using dynamically downscaled data. Although an alternative to pattern scaling between RCM runs exists; to downscale statistically from transient GCM outputs.

4. What new methods can be used together with downscaling to assess uncertainties in hydrological response?

Probabilistic methods seem to offer a more robust way of assessing climate change impacts with much consideration given to uncertainties in the modelling framework, although few studies have so far examined impacts on hydrological systems. These allow the inclusion of uncertainty estimates using a multi-model approach which can potentially be used in the planning of adaptation measures. Although the implementation of such methods in examining climate change impacts on hydrological systems has now begun, there are many avenues of research still to be addressed: (1) How should the outputs from multiple RCMs and GCMs be used – weighting methods? (2) Which approaches are most suitable for the construction of pdfs for regional impact studies? (3) How can the pdfs be applied to examine impacts? (4) How may pdfs be linked to downscaling methods?

Model credibility has thus far been based on the ability to reproduce the observed conditions in the ‘control integration’ as we have no way of assessing what will be the ‘true’ future response. Many different methods have been used to assess model outputs against observations. Most commonly, the regional average of the long-term mean or seasonal mean for different climatic variables is compared for observations and climate models outputs. Model weighting is then determined according to ability in reproducing the observed mean climate. However, there has been little assessment of whether ability to reproduce average observed conditions predicates ability in future prediction. More research is needed in this area on questions such as: do we need to assess model ability to reproduce change in climate conditions to assess model ability in future predictions? A measure of prediction ability rather than downscaling ability may be more appropriate for model weighting, if indeed, model weighting is appropriate. An alternative approach may be to give equal weighting to all models and thus all possible outcomes.

Emphasis has been put on the reproduction of regional mean climate in model weighting, but for many impact studies the reproduction of daily and inter-annual variability may be more important. Busuioac *et al.* (2001) for example suggest that the optimum statistical downscaling model for climate change applications will be one which has the highest skill in reproduction of low frequency variability, rather than having the most skill in terms of explained variance. Alternatively, Wetterhall *et al.* (2006) suggest that the ability to model inter-annual variability is an important measure of the sensitivity to the climate signal. Therefore, a model that is used for climate change studies should be able to differentiate between dry and wet years; again differentiating models based on variability rather than mean statistics.

It is not yet clear how model ability in future prediction may be assessed; and, such, there has been no real advance in this research area since Leung *et al.* (2003b)’s review. However, it is clear that in most cases it is impossible to designate a ‘best model’ as their simulation skill for key statistics varies between, and even within, climatic variables both temporally and spatially. Important

questions must be asked as to how such models should be assessed, as those models which well reproduce the observed mean statistics of the climate at a regional scale may not well reproduce the spatial variability over the same region. Given that within-region differences are likely to be important when considering specific impacts of climate change it may be more appropriate to assess how well models reflect the spatial characteristics of climate, large-scale circulation patterns or paired climate regions. Evaluating circulation patterns is still an alternative that is frequently overlooked. Currently, we would recommend that models are chosen that best reproduce the climatic statistics important for an impacts study; an impacts specific climate model selection. It is hoped, however, that large ongoing projects such as ENSEMBLES will be able to provide solutions to many of these questions.

The consistent downscaling of multiple climatic variables has been successful using bias-corrected RCM outputs. However, to incorporate probabilistic methods into impact assessment, use must be made of either distributed computing methods such as Climate Prediction. Net or statistical downscaling methods that allow the simulation of multiple, long climatic time series. Progress has been made in the simulation of multiple climatic variables using WGs but these still function at only the single-site level. For hydrological impact studies it is important that these types of approaches are extended to simulate at multiple sites across a region, that they can be easily re-parameterized for future climate change, and that there is assessment of their ability to produce low frequency variability.

5. How can downscaling methods be better utilized within the hydrological impacts community?

Despite a thorough exploration of the weaknesses and strengths of different downscaling methods, predictor variables and downscaling domains within the literature, there has still been little application of these methods for hydrological impact assessment. Where these have been achieved there is still little thought given to how the results can be best used for robust management decisions leading to adaptation.

Current assessments of climate change effects on hydrological systems are made mainly using offline modelling approaches. Outputs from climate model integrations are used to drive hydrological models and estimate climate change impacts. However, because feedback effects are important, coupled regional modelling systems may provide more useful predictions of climate impacts, particularly now that RCMs are available at higher spatial resolutions more appropriate for hydrological modelling. These ‘whole system models’ may be more appropriate for the assessment of adaptation and mitigation strategies. However, there has been no examination of whether on- or off-line approaches provide better estimates of future climate change impacts and what temporal and spatial scales are appropriate for coupling. There is a need for

these types of comparison studies. We still, as recommended by Leung *et al.* (2003b), need to develop 'coordinated end-to-end prediction systems to test the whole approach of impact assessment'.

Additionally, within the hydrological community, models are normally designed for use in stationary conditions. These are then used in changing or changed conditions in impact studies without testing of whether these models are predictive. In the same way that climate model ability to reproduce observations may not predicate its ability to predict future climate, traditional split-sample testing may not give a robust estimate of a hydrological model's ability to predict the effects of climate change. The main assumption, that hydrological parameters will remain the same, may simply not hold true under climate change. Although there has been some research into the problem of equifinality – that different model parameterizations produce similarly good flow simulations – more research is needed into which hydrological models are predictive and why.

CONCLUSIONS

In conclusion, since the 1990s there has been a thorough exploration of the strengths and weakness of different downscaling methods within the literature and there is no need for additional comparison studies. Although there has been a huge expansion of the downscaling literature only about one third of all downscaling studies consider impacts, and only half of these consider hydrological impacts. Within studies considering hydrological impacts there is still little consideration given to applied research; how the results can be best used to enable stakeholders and managers to make informed, robust decisions on adaptation and mitigation strategies in the face of many uncertainties about the future.

As many of the impacts of climate change will not be detectable in the near future (e.g. Wilby, 2006), there is a need for decision-making tools for planning and management that are robust to future uncertainties. In hydrological impacts research there is a need for a move away from comparison studies into the provision of such tools based on the selection of robust, possibly impact-specific, downscaling methods. This is essential, together with the examination and understanding of uncertainties within the downscaling and modelling system, as for example, attempted by Wilby and Harris (2006). Probabilistic methods seem to offer a more robust way of assessing climate change impacts, but much research is still needed on the best way to apply such methods for different impacts and in different locations. It is still unclear who exactly should be responsible for generating probabilistic hydrologic scenarios. If this is to be the decision-maker or practitioner rather than the scientist, then this implies a need for more tools where the 'hard' science and data is embedded and hidden, such as EARWIG (Kilsby *et al.*, 2007b). Nevertheless, as these allow the inclusion of uncertainty estimates using a multi-model

approach which can be used in the planning of adaptation measures, they seem to offer the most potential for advancement within both the 'downscaling for hydrological impacts' science community and for practitioners.

ACKNOWLEDGEMENTS

This work was supported by the European Union FP6 Integrated Project AquaTerra (Project no. 505428) under the thematic priority sustainable development, global change and ecosystems. We would like to thank all who participated in the H13 session 'Linking climate change modelling to impacts studies: downscaling techniques for hydrological impact studies' at the European Geosciences Union meeting in Vienna, April, 2006, for stimulating presentations, posters and discussion that helped to shape some of the ideas in this paper. Thanks also to Rob Wilby for insightful comments provided on a draft of this paper.

REFERENCES

- Abaurrea J, Asín J. 2005. Forecasting local daily precipitation patterns in a climate change scenario. *Climate Research* **28**: 183–197.
- Allen MR, Ingram WJ. 2002. Constraints on future changes in climate and the hydrologic cycle. *Nature* **419**: 224–232.
- Allen MR, Kettleborough JA, Stainforth DA. 2003. Model error in weather and climate forecast. *Proceedings of the 2002 ECMWF Predictability Seminar*. ECMWF: Reading.
- Allen MR, Stott PA, Mitchell JFB, Schnur R, Delworth TL. 2000. Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature* **417**: 617–620.
- Andronova NG, Schlesinger ME. 2001. Objective estimation of the probability density function for climate sensitivity. *Journal of Geophysical Research-Atmospheres* **106**(D19): 22605–22611.
- Arnell NW, Reynard NS. 1996. The effects of climate change due to global warming on river flows in Great Britain. *Journal of Hydrology* **183**: 397–424.
- Bárdossy A, Stehlík J, Caspary HJ. 2002. Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules. *Climate Research* **23**: 11–22.
- Bárdossy A, Bogardi I, Matyasovszky I. 2005. Fuzzy rule-based downscaling of precipitation. *Theoretical and Applied Climatology* **82**: 119–129.
- Barnett T, Malone T, Pennell W, Stammer D, Semtner B, Washington W. 2004. The effects of climate change on water resources in the west: introduction and overview. *Climatic Change* **62**: 1–11.
- Beckmann BR, Buishand TA. 2002. Statistical downscaling relationships for precipitation in the Netherlands and north Germany. *International Journal of Climatology* **22**: 15–32.
- Beersma JJ, Buishand TA. 2003. Multi-site simulation of daily precipitation and temperature conditional on the atmospheric circulation. *Climate Research* **25**: 121–133.
- Bellone E, Hughes JP, Guttorp P. 2000. A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts. *Climate Research* **15**: 1–12.
- Benestad RE. 2001. A comparison between two empirical downscaling strategies. *International Journal of Climatology* **21**: 1645–1668.
- Benestad RE. 2004. Tentative probabilistic temperature scenarios for northern Europe. *Tellus Series A – Dynamic Meteorology and Oceanography* **56**: 89–101.
- Benestad RE, Førland EJ, Hanssen-Bauer I. 2002. Empirically downscaled temperature scenarios for Svalbard. *Atmospheric Science Letters* **3**(2): 71–93, DOI: 10.1006/Asle.2002.0051.
- Bergant K, Kajfez-Bogataj L. 2005. N-PLS regression as empirical downscaling tool in climate change studies. *Theoretical and Applied Climatology* **81**: 11–23.
- Bergstrom S, Carlsson B, Gardelin M, Lindstrom G, Pettersson A, Rummukainen M. 2001. Climate change impacts on run-off in Sweden – assessments by global climate models, dynamical

- downscaling and hydrological modeling. *Climate Research* **16**: 101–112.
- Blenkinsop S, Fowler HJ. 2007. Changes in drought frequency and severity over the British Isles projected by the PRUDENCE regional climate models. *Journal of Hydrology* **342**: 50–71.
- Boorman BD, Sefton CEM. 1997. Recognising the uncertainty in the quantification of the effects of climate change on hydrological response. *Climatic Change* **35**: 415–434.
- Brinkmann WAR. 2000. Modification of a correlation-based circulation pattern classification to reduce within-type variability of temperature and precipitation. *International Journal of Climatology* **20**: 839–852.
- Brinkmann WAR. 2002. Local versus remote grid points in climate downscaling. *Climate Research* **21**: 27–42.
- Busuioc A, Chen D, Hellström C. 2001. Performance of statistical downscaling models in GCM validation and regional climate change estimates: application for Swedish precipitation. *International Journal of Climatology* **21**: 557–578.
- Bürger G. 1996. Expanded downscaling for generating local weather scenarios. *Climate Research* **7**: 111–128.
- Bürger G, Chen Y. 2005. Regression-based downscaling of spatial variability for hydrologic applications. *Journal of Hydrology* **311**: 299–317.
- Cavazos T, Hewitson BC. 2005. Performance of NCEP-NCAR reanalysis variables in statistical downscaling of daily precipitation. *Climate Research* **28**: 95–107.
- Cawley GC, Haylock M, Dorling SR, Goodess C, Jones PD. 2003. Statistical downscaling with artificial neural networks. In *Proceedings – European Symposium on Artificial Neural Networks*, Bruges, Belgium, 23–25th April 2003.
- Charles SP, Bates BC, Whetton PH, Hughes JP. 1999. Validation of a downscaling model for changed climate conditions in southwestern Australia. *Climate Research* **12**: 1–14.
- Charles SP, Bates BC, Smith IN, Hughes JP. 2004. Statistical downscaling of daily precipitation from observed and modelled atmospheric fields. *Hydrological Processes* **18**: 1373–1394.
- Chen D, Achberger C, Räisänen J, Hellstrom C. 2006. Using statistical downscaling to quantify the GCM-related uncertainty in regional climate change scenarios: a case study of Swedish precipitation. *Advances in Atmospheric Sciences* **23**: 54–60.
- Christensen JH, Carter TR, Rummukainen M, Amanatidis G. 2007. Evaluating the performance and utility of regional climate models: the PRUDENCE project. *Climatic Change* **81**(Supplement): 1–6.
- Christensen OB, Christensen JH, Machenhauer B, Botzet M. 1998. Very high-resolution regional climate simulations over Scandinavia – Present climate. *Journal of Climate* **11**: 3204–3229.
- Christensen NS, Wood AW, Voisin N, Lettenmaier DP, Palmer RN. 2004. Effects of climate change on the hydrology and water resources of the Colorado river basin. *Climatic Change* **62**: 337–363.
- Christensen JH, Räisänen J, Iversen T, Bjørge D, Christensen OB, Rummukainen M. 2001. A synthesis of regional climate change simulations – A Scandinavian perspective. *Geophysical Research Letters* **28**: 1003–1006.
- Conway D, Wilby RL, Jones PD. 1996. Precipitation and air flow indices over the British Isles. *Climate Research* **7**: 169–183.
- Corte-Real J, Xu H, Qian B. 1999. A weather generator for obtaining daily precipitation scenarios based on circulation patterns. *Climate Research* **13**: 61–75.
- Coulibaly P. 2004. Downscaling daily extreme temperatures with genetic programming. *Geophysical Research Letters* **31**: L16203, DOI: 10.1029/2004GL020075.
- Coulibaly P, Dibike YB, Ancil F. 2005. Downscaling precipitation and temperature with temporal neural networks. *Journal of Hydrometeorology* **6**: 483–496.
- Cowpertwait PSP. 1991. Further developments of the Neyman-Scott clustered point process for modelling rainfall. *Water Resources Research* **27**: 1431–1438.
- Cubasch U, von Storch H, Waszkewitz J, Zorita E. 1996. Estimates of climate change in southern Europe using different downscaling techniques. *Climate Research* **7**: 129–149.
- Cunderlik JM, Simonovic SP. 2005. Hydrological extremes in a southwestern Ontario river basin under future climate conditions. *Hydrological Sciences Journal* **50**: 631–654.
- Dehn M, Bürger G, Buma J, Gasparetto P. 2000. Impact of climate change on slope stability using expanded downscaling. *Engineering Geology* **55**: 193–204.
- Déqué M, Rowell DP, Lüthi D, Giorgi F, Christensen JH, Rockel B, Jacob D, Kjellström 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**(Supplement 1): 53–70.
- Déqué M, Jones RG, Wild M, Giorgi F, Christensen JH, Hassell DC, Vidale PL, Rockel B, Jacob D, Kjellström E, de Castro M, Kucharski F, van den Hurk B. 2005. Global high resolution versus Limited Area Model climate change projections over Europe: quantifying confidence level from PRUDENCE results. *Climate Dynamics* **25**: 653–670.
- Dessai S, Hulme M. 2004. Does climate policy need probabilities? *Climate Policy* **4**: 107–128.
- Dessai S, Lu X, Hulme M. 2005. Limited sensitivity analysis of regional climate change probabilities for the 21st century. *Journal of Geophysical Research* **110**: D19108, DOI: 10.1029/2005JD005919.
- Dettinger MD, Cayan DR, Meyer MK, Jeton AE. 2004. Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River Basins, Sierra Nevada, California, 1900–2099. *Climatic Change* **62**: 283–317.
- Díaz-Nieto J, Wilby RL. 2005. A comparison of statistical downscaling and climate change factor methods: Impacts on low flows in the River Thames, United Kingdom. *Climatic Change* **69**: 245–268.
- Díez E, Primo C, García-Moya JA, Gutiérrez JM, Orfila B. 2005. Statistical and dynamical downscaling of precipitation over Spain from DEMETER seasonal forecasts. *Tellus Series A-Dynamic Meteorology and Oceanography* **57**(3): 409–423.
- Dibike YB, Coulibaly P. 2005. Hydrologic impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. *Journal of Hydrology* **307**: 145–163.
- Duan Q, Sorooshian S, Gupta VK. 1992. Effective and efficient global optimization for conceptual rainfall-run-off. *Water Resources Research* **28**: 1015–1031.
- Dubrovsky M, Buchtele J, Zalud Z. 2004. High-frequency and low-frequency variability in stochastic daily weather generator and its effect on agricultural and hydrologic modelling. *Climatic Change* **63**: 145–179.
- Easterling DR. 1999. Development of regional climate scenarios using a downscaling approach. *Climatic Change* **41**: 615–634.
- Ekström M, Fowler HJ, Kilsby CG, Jones PD. 2005. New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Future estimates and use in impact studies. *Journal of Hydrology* **300**: 234–251.
- Ekström M, Jones PD, Fowler HJ, Lenderink G, Buishand A, Conway D. 2007. Regional climate model data used within the SWURVE project 1: Projected changes in seasonal patterns and estimation of PET. *Hydrology and Earth System Sciences* **11**(3): 1069–1083.
- Enke W, Schneider F, Deuschländer T. 2005a. A novel scheme to derive optimized circulation pattern classifications for downscaling and forecast purposes. *Theoretical and Applied Climatology* **82**: 51–63.
- Enke W, Deuschländer T, Schneider F, Küchler W. 2005b. Results of five regional climate studies applying a weather pattern based downscaling method to ECHAM4 climate simulations. *Meteorologische Zeitschrift* **14**: 247–257.
- Evans J, Schreider S. 2002. Hydrological impacts of climate change on inflows to Perth, Australia. *Climatic Change* **55**: 361–393.
- Forest CE, Stone PH, Sokolov AP, Allen MR, Webster MD. 2002. Quantifying uncertainties in climate system properties with the use of recent climate observations. *Science* **295**: 113–117.
- Fowler HJ, Kilsby CG, O'Connell PE. 2000. A stochastic rainfall model for the assessment of regional water resource systems under changed climatic conditions. *Hydrology and Earth System Sciences* **4**: 261–280.
- Fowler HJ, Kilsby CG. 2002. Precipitation and the North Atlantic Oscillation: a study of climatic variability in Northern England. *International Journal of Climatology* **22**: 843–866.
- Fowler HJ, Ekström M, Kilsby CG, Jones PD. 2005a. New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate. *Journal of Hydrology* **300**: 212–233.
- Fowler HJ, Ekström M, Blenkinsop S, Smith AP. Estimating change in extreme European precipitation using a multi-model ensemble. *Geophysical Research Atmospheres* (in press).
- Fowler HJ, Kilsby CG, O'Connell PE, Burton A. 2005b. A weather-type conditioned multi-site stochastic rainfall model for generation of scenarios of climatic variability and change. *Journal of Hydrology* **308**(1–4): 50–66.
- Fowler HJ, Kilsby CG. 2007. Using regional climate model data to simulate historical and future river flows in northwest England. *Climatic Change* **80**(3–4): 337–367.

- Fowler HJ, Kilsby CG, Stunell J. 2007. Modelling the impacts of projected future climate change on water resources in northwest England. *Hydrology and Earth System Sciences* **11**(3): 1115–1126.
- Frame DJ, Booth BBB, Kettleborough JA, Stainforth DA, Gregory JM, Collins M, Allen MR. 2005. Constraining climate forecasting: the role of prior assumptions. *Geophysical Research Letters* **32**(9): L09702, DOI: 10.1029/2004GL022241.
- Frei C, Schär C. 2001. Detection probability of trends in rare events: theory and application to heavy precipitation in the Alpine region. *Journal of Climate* **14**: 1568–1584.
- Frei C, Schär C, Lüthi D, Davies HC. 1998. Heavy precipitation processes in a warmer climate. *Geophysical Research Letters* **25**(9): 1431–1434.
- Frei C, Schöll R, Fukutome S, Schmidli J, Vidale PL. 2006. Future change of precipitation extremes in Europe: an intercomparison of scenarios from regional climate models. *Journal of Geophysical Research-Atmospheres* **111**: D06105, DOI: 10.1029/2005JD005965.
- 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. *Journal of Geophysical Research* **108**(D3): 4124, DOI: 10.1029/2002JD002287.
- Furrer R, Sain SR, Nychka D, Meehl GA. 2007. Multivariate bayesian analysis of atmosphere-ocean general circulation models. *Environmental and Ecological Statistics* (in press).
- Gangopadhyay S, Clark M, Rajagopalan B. 2005. Statistical downscaling using K-nearest neighbours. *Water Resources Research* **41**: W02024, DOI: 10.1029/2004WR003444.
- Georgakakos KP, Bae DH, Jeong CS. 2005. Utility of ten-day climate model ensemble simulations for water resources applications in Korean watersheds. *Water Resources Management* **19**: 849–872.
- Giorgi F, Francisco R. 2000. Evaluating uncertainties in the prediction of regional climate change. *Geophysical Research Letters* **27**: 1295–1298.
- Giorgi F, Hewitson BC. 2001. Regional climate information – evaluation and projections. In *Climate Change 2001: The Scientific Basis*. C, Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dia X, Maskell K, Johnson CA (eds). Cambridge University Press: Cambridge.
- Giorgi F, Mearns LO. 2002. Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the “reliability ensemble averaging” (REA) method. *Journal of Climate* **15**: 1141–1158.
- Giorgi F, Mearns LO. 2003. Probability of regional climate change based on the Reliability Ensemble Averaging (REA) method. *Geophysical Research Letters* **30**(12): 1629, DOI: 10.1029/2003GL017130.
- Goodess CM, Palutikof J. 1998. Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling. *International Journal of Climatology* **18**: 1051–1083.
- Goodess CM, Jones PD. 2002. Links between circulation and changes in the characteristics of Iberian rainfall. *International Journal of Climatology* **22**: 1593–1615.
- Goodess CM, Anagnostopoulou C, Bárdossy A, Frei C, Harpham C, Haylock MR, Hundecha Y, Maheras P, Ribalaygua J, Schmidli J, Schmith T, Tolika K, Tomozeiu R, Wilby RL. 2007. An intercomparison of statistical downscaling methods for Europe and European regions – assessing their performance with respect to extreme temperature and precipitation events. *Climatic Change* (in press).
- 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**(Supplement): 293–307.
- Graham LP, Hagemann S, Jaun S, Beniston M. 2007b. On interpreting hydrological change from regional climate models. *Climatic Change* **81**(Supplement): 97–122.
- Grantz K, Rajagopalan B, Clark M, Zagana E. 2005. A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts. *Water Resources Research* **41**: W10410, DOI: 10.1029/2004WR003467.
- Greene AM, Goddard L, Lall U. 2006. Probabilistic multi-model regional temperature change projections. *Journal of Climate* **19**: 4326–4343.
- Grotch SL, MacCracken MC. 1991. The use of general circulation models to predict regional climatic change. *Journal of Climate* **4**: 286–303.
- Hagemann S, Machenhauer B, Jones R, Christensen OB, Déqué M, Jacob D, Vidale PL. 2004. Evaluation of water and energy budgets in regional climate models applied over Europe. *Climate Dynamics* **23**: 547–567.
- Hanssen-Bauer I, Førland EJ. 1998. Long-term trends in precipitation and temperature in the Norwegian Arctic: can they be explained by changes in atmospheric circulation patterns? *Climate Research* **10**: 143–153.
- Hanssen-Bauer I, Førland EJ, Haugen JE, Tveito OE. 2003. Temperature and precipitation scenarios for Norway: comparison of results from dynamical and empirical downscaling. *Climate Research* **25**: 15–27.
- Hanssen-Bauer I, Achberger C, Benestad RE, Chen D, Forland EJ. 2005. Statistical downscaling of climate scenarios over Scandinavia. *Climate Research* **29**: 255–268.
- Harpham C, Wilby RL. 2005. Multi-site downscaling of heavy daily precipitation occurrence and amounts. *Journal of Hydrology* **312**: 1–21.
- Harrold TI, Jones RN. 2003. Downscaling GCM rainfall: a refinement of the perturbation method. In *MODSIM 2003 International Congress on Modelling and Simulation*, Townsville, Australia, 14–17 July 2003.
- 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. *Journal of Hydrology* **282**: 56–75.
- Hayhoe HN. 2000. Improvements of stochastic weather data generators for diverse climates. *Climate Research* **14**: 75–87.
- Haylock MR, Cawley GC, Harpham C, Wilby RL, Goodess CM. 2006. Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios. *International Journal of Climatology* **26**: 1397–1415.
- Hellström C, Chen D. 2003. Statistical downscaling based on dynamically downscaled predictors: application to monthly precipitation in Sweden. *Advances in Atmospheric Sciences* **20**: 951–958.
- Hellström C, Chen D, Achberger C, Räisänen J. 2001. Comparison of climate change scenarios for Sweden based on statistical and dynamical downscaling of monthly precipitation. *Climate Research* **19**: 45–55.
- Hewitson BC, Crane RG. 1996. Climate downscaling: techniques and application. *Climate Research* **7**: 85–95.
- Hewitson BC, Crane RG. 2006. Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *International Journal of Climatology* **26**: 1315–1337.
- Hingray B, Mezghani A, Buishand TA. 2007a. Development of probability distributions for regional climate change from uncertain global-mean warming and an uncertain scaling relationship. *Hydrology and Earth System Sciences* **11**(3): 1097–1114.
- Hingray B, Mouhous N, Mezghani A, Bogner K, Schaeffli B, Musy A. 2007b. Accounting for global-mean warming and scaling uncertainties in climate change impact studies: application to a regulated lake system. *Hydrology and Earth System Sciences* **11**(3): 1207–1226.
- Houghton JT, Meira Filho LG, Callander BA, Harris N, Kattenberg A, Maskell K. 1996. *Climate Change 1995: The Science of Climate Change. The Second Assessment Report of the IPCC: Contribution of Working Group I*. Cambridge University Press: Cambridge.
- Hulme M, Jenkins GJ, Lu X, Turnpenny JR, Mitchell TD, Jones RG, Lowe J, Murphy JM, Hassell D, Boorman P, McDonald R, Hill S. 2002. *Climate Change Scenarios for the United Kingdom: The UKCIP02 Scientific Report*. Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia: Norwich; 120.
- Hunt JCR. 2005. Inland and coastal flooding: developments in prediction and prevention. *Philosophical Transactions of the Royal Society – Series A* **363**: 1475–1491.
- Huth R. 1997. Potential of continental-scale circulation for the determination of local daily surface variables. *Theoretical and Applied Climatology* **56**: 165–186.
- Huth R. 1999. Statistical downscaling in central Europe: evaluation of methods and potential predictors. *Climate Research* **13**: 91–101.
- Huth R. 2002. Statistical downscaling of daily temperature in Central Europe. *Journal of Climate* **15**: 1731–1742.
- Huth R. 2005. Downscaling of humidity variables: a search for suitable predictors and predictands. *International Journal of Climatology* **25**: 243–250.
- Huth R, Kyselý J. 2000. Constructing site-specific climate change scenarios on a monthly scale using statistical downscaling. *Theoretical and Applied Climatology* **66**: 13–27.
- Huth R, Kyselý J, Dubrovský M. 2003. Simulation of surface air temperature by GCMs, statistical downscaling and weather

- generator: higher order statistical moments. *Studia Geophysica et Geodaetica* **47**: 203–216.
- Jasper K, Calanca P, Gyalistras D, Fuhrer J. 2004. Differential impacts of climate change on the hydrology of two alpine river basins. *Climate Research* **26**: 113–129.
- Jeong CS, Heo JH, Bae DH, Georgakakos KP. 2005. Utility of high-resolution climate model simulations for water resources prediction over the Korean Peninsula: a sensitivity study. *Hydrological Sciences Journal* **50**: 139–153.
- Jones PD. 1988. Hemispheric surface air temperature variations: recent trends and an update to 1987. *Journal of Climate* **1**: 654–659.
- Jones PD, Salmon M. 1995. *Development and Integration of a Stochastic Weather Generator into a Crop Growth Model for European Agriculture*. MARS Project, Final Report to Institute of Remote Sensing Applications, Agricultural Information Systems (ISPRA): UK; Contract No. 5631-93-12, ED ISP GB.
- Jones RN. 2000a. Analysing the risk of climate change using an irrigation demand model. *Climate Research* **14**: 89–100.
- Jones RN. 2000b. Managing uncertainty in climate change projections – issues for impact assessments. *Climatic Change* **45**: 403–419.
- 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. *Journal of Climate* **3**: 1053–1079.
- Kettle H, Thompson R. 2004. Statistical downscaling in European mountains: verification of reconstructed air temperature. *Climate Research* **26**: 97–112.
- Khan MS, Coulibaly P, Dibike Y. 2006. Uncertainty analysis of statistical downscaling methods. *Journal of Hydrology* **319**: 357–382.
- Kidson JW, Thompson CS. 1998. A comparison of statistical and model-based downscaling techniques for estimating local climate variations. *Journal of Climate* **11**: 735–753.
- Kilsby CG, Cowpertwait PSP, O'Connell PE, Jones PD. 1998. Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *International Journal of Climatology* **18**: 523–539.
- Kilsby CG, Tellier SS, Fowler HJ, Howels TR. 2007. Hydrological impacts of climate change on the Tejo and Guadiana rivers. *Hydrology and Earth System Sciences* **11**(3): 1175–1189.
- 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. *Environmental Modelling and Software* **22**: 1705–1719.
- Kleinn J, Frei C, Gurtz J, Lüthi D, Vidale PL, Schär C. 2005. Hydrologic simulations in the Rhine basin driven by a regional climate model. *Journal of Geophysical Research* **110**: D04102, Doi:10.1029/2004JD005143.
- Klein Tank AMG, Können GP. 2003. Trends in indices of daily temperature and precipitation extremes in Europe, 1946–99. *Journal of Climate* **16**: 3665–3680.
- Knutti R, Stocker TF, Joos F, Plattner GK. 2002. Constraints on radiative forcing and future climate change from observations and climate models ensembles. *Nature* **416**: 719–723.
- Knutti R, Stocker TF, Joos F, Plattner GK. 2003. Probabilistic climate change projections using neural networks. *Climate Dynamics* **21**: 257–272.
- Knutti R, Joos F, Müller SA, Plattner GK, Stocker TF. 2005. Probabilistic climate change projections for CO₂ stabilization profiles. *Geophysical Research Letters* **32**(20): L20707. DOI: 10.1029/2005GL023294.
- Kyselý J. 2002. Comparison of extremes in GCM-simulated, downscaled and observed central-European temperature series. *Climate Research* **20**: 211–222.
- Kyselý J, Dubrovsky M. 2005. Simulation of extreme temperature events by a stochastic weather generator: Effects of interdiurnal and interannual variability reproduction. *International Journal of Climatology* **25**: 251–269.
- Lall U, Sharma A. 1996. A nearest neighbor bootstrap for time series resampling. *Water Resources Research* **32**: 679–693.
- Leander R, Buishand A, Aalders P, de Wit M. 2005. Estimation of extreme floods of the River Meuse using a stochastic weather generator and a rainfall-runoff model. *Hydrological Sciences Journal* **50**: 1089–1103.
- Leung RL, Mearns LO, Giorgi F, Wilby RL. 2003a. Regional climate research: Needs and opportunities. *Bulletin of the American Meteorological Society* **84**: 89–95.
- Leung LR, Qian Y, Bian X, Hunt A. 2003b. Hydroclimate of the western United States based on observations and regional climate simulation of 1981–2000. Part II: mesoscale ENSO anomalies. *Journal of Climate* **16**: 1912–1928.
- Leung LR, Qian Y, Bian X, Washington WM, Han J, Roads JO. 2004. Mid-century ensemble regional climate change scenarios for the Western United States. *Climatic Change* **62**: 75–113.
- Lopez A, Tebaldi C, New M, Stainforth D, Allen M, Kettleborough J. 2006. Two approaches to quantifying uncertainty in global temperature changes. *Journal of Climate* **19**(19): 4785–4796.
- Luo QY, Jones RN, Williams M, Bryan B, Bellotti W. 2005. Probabilistic distributions of regional climate change and their application in risk analysis of wheat production. *Climate Research* **29**: 41–52.
- Mason SJ. 2004. Simulating Climate over Western North America Using Stochastic Weather Generators. *Climatic Change* **62**: 155–187.
- Mastrandrea MD, Schneider SH. 2004. Integrated assessment of “dangerous” climate change. *Science* **304**: 571–575.
- Matulla C. 2005. Regional, seasonal and predictor-optimized downscaling to provide groups of local scale scenarios in the complex structured terrain of Austria. *Meteorologische Zeitschrift* **14**: 31–47.
- Mearns LO, the NARCCAP Team. 2006. Overview of the North American Regional Climate Change Assessment Program. In *NOAA RISA-NCAR Meeting*, Tucson, AZ, March 2006.
- Mearns LO, Bogardi I, Giorgi F, Matyasovsky I, Palecki M. 1999. Comparison of climate change scenarios generated daily temperature and precipitation from regional climate model experiments and statistical downscaling. *Journal of Geophysical Research* **104**: 6603–6621.
- Meehl GA, Boer GJ, Covey C, Latif M, Stouffer RJ. 2000. The coupled model intercomparison project (CMIP). *Bulletin of the American Meteorological Society* **81**: 313–318.
- Mitchell TD. 2000. An investigation of the pattern scaling technique for describing future climates. PhD thesis, University of East Anglia, Norwich.
- Mitchell TD. 2003. Pattern scaling: an examination of the accuracy of the technique for describing future climates. *Climatic Change* **60**: 217–242.
- Mitchell JFB, Johns TC, Eagles M, Ingram WJ, David RA. 1999. Towards the construction of climate change scenarios. *Climatic Change* **41**: 547–581.
- 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 J. 1999. An evaluation of statistical and dynamical techniques for downscaling local climate. *Journal of Climate* **12**: 2256–2284.
- Murphy J. 2000. Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *International Journal of Climatology* **20**: 489–501.
- Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA. 2004. Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* **430**: 768–772.
- Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models part I - A discussion of principles. *Journal of Hydrology* **10**: 282–290.
- Nicholls N, Gruza G, Jouzel J, Karl TR, Ogallo LA, Parker DE. 1996. Observed climate change and variability. In *Climate Change 1995: The Science of Climate Change. The Second Assessment Report of the IPCC: Contribution of Working Group I*, Houghton JT, Meira Filho LG, Callender BA, Harris N, Kattenberg A, Maskell K (eds). Cambridge University Press: Cambridge; 133–192.
- Palmer TN, Räisänen J. 2002. Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature* **415**: 512.
- Palmer TN, Shutts GJ, Hagedorn R, Doblas-Reyes FJ, Jung T, Leutbecher M. 2005. Representing model uncertainty in weather and climate prediction. *Annual Review of Earth and Planetary Sciences* **33**: 193–193.
- Palutikof JP, Goodess CM, Watkins SJ, Holt T. 2002. Generating rainfall and temperature scenarios at multiple sites: examples from the mediterranean. *Journal of Climate* **15**: 3529–3548.
- Payne JT, Wood AW, Hamlet AF, Palmer RN, Lettenmaier DP. 2004. Mitigating the effects of climate change on the water resources of the Columbia river basin. *Climatic Change* **62**: 233–256.
- Penlap EK, Matulla C, von Storch H, Kamga FM. 2004. Downscaling of GCM scenarios to assess precipitation changes in the little rainy season (March–June) in Cameroon. *Climate Research* **26**: 85–96.
- 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. *Hydrological Processes* **16**: 1201–1213.

- Prudhomme C, Reynard N, Crooks S. 2002. Downscaling of global climate models for flood frequency analysis: Where are we now?. *Hydrological Processes* **16**: 1137–1150.
- Prudhomme C, Jakob D, Svensson C. 2003. Uncertainty and climate change impact on the flood regime of small UK catchments. *Journal of Hydrology* **277**: 1–23.
- Qian B, Hayhoe H, Gameda S. 2005. Evaluation of the stochastic weather generators LARS-WG and AAFC-WG for climate change impact studies. *Climate Research* **29**: 3–21.
- Racsko P, Szeidl L, Semenov M. 1991. A serial approach to local stochastic weather models. *Ecological Modelling* **57**: 27–41.
- Räisänen J, Palmer TN. 2001. A probability and decision-model analysis of a multimodel ensemble of climate change simulations. *Journal of Climate* **14**: 3212–3226.
- Räisänen J, Ruokolainen L. 2006. Probabilistic forecasts of near-term climate change based on a resampling ensemble technique. *Tellus Series A-Dynamic Meteorology and Oceanography* **58**: 461–472.
- Rowell DP. 2006. A demonstration of the uncertainty in projections of UK climate change resulting from regional model formulation. *Climatic Change* **79**(3–4): 243–257.
- Salathé EP. 2003. Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. *International Journal of Climatology* **23**: 887–901.
- Salathé EP. 2005. Downscaling simulations of future global climate with application to hydrologic modelling. *International Journal of Climatology* **25**: 419–436.
- Santer BD, Wigley TML, Schlesinger ME, Mitchell JFB. 1990. Developing climate scenarios from equilibrium GCM results. Rpt. No. 47. Max-Planck-Institut-fur-Meteorologie: Hamburg; 29.
- Santer BD, Brueggemann W, Cubasch U, Hasselmann K, Hocke H, Maier-Reimer E, Mikolajewicz U. 1994. Signal-to-noise analysis of time-dependent greenhouse gas warming experiments. *Climate Dynamics* **9**: 267–285.
- Schmidli J, Frei C, Vidale PL. 2006. Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology* **26**: 679–689.
- Schneider SH, Mastrandrea MD. 2005. Probabilistic assessment “dangerous” climate change and emissions pathways. *Proceedings of the National Academy of Science* **102**: 15728–15735.
- Schoof JT, Pryor SC. 2001. Downscaling temperature and precipitation: a comparison of regression-based methods and artificial neural networks. *International Journal of Climatology* **21**: 773–790.
- Schubert S. 1998. Downscaling local extreme temperature changes in south-eastern Australia from the CSIRO Mark2 GCM. *International Journal of Climatology* **18**: 1419–1438.
- Schubert S, Henderson-Sellers A. 1997. A statistical model to downscale local daily temperature extremes from synoptic-scale atmospheric circulation patterns in the Australian region. *Climate Dynamics* **13**: 223–234.
- Semenov MA, Brooks RJ, Barrow EM, Richardson CW. 1998. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Climate Research* **10**: 95–107.
- Slonosky VC, Jones PD, Davies TD. 2001. Atmospheric circulation and surface temperature in Europe from the 18th century to 1995. *International Journal of Climatology* **21**: 63–75.
- Stainforth D, Kettleborough J, Allen M, Collins M, Heaps A, Murphy J. 2002. Distributed computing for public-interest climate modeling research. *Computing in Science & Engineering* **4**: 82–89.
- Stainforth DA, Aina T, Christensen C, Collins M, Faull N, Frame DJ, Kettleborough JA, Knight S, Martin A, Murphy JM, Piani C, Sexton D, Smith LA, Spicer RA, Thorpe AJ, Allen MR. 2005. Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature* **433**: 403–406.
- Stott PA. 2003. Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophysical Research Letters* **30**(14): 1728, DOI: 10.1029/2003GL017324.
- Stott PA, Kettleborough JA. 2002. Origins and estimates of uncertainty in predictions of twenty-first century temperature rise. *Nature* **416**: 723–726.
- Stott PA, Kettleborough JA, Allen MR. 2006. Uncertainty in continental-scale temperature predictions. *Geophysical Research Letters* **33**: L02708, DOI: 10.1029/2005GL024423.
- Tatli H, Dalfes HN, Mentes SS. 2004. A statistical downscaling method for monthly total precipitation over Turkey. *International Journal of Climatology* **24**: 161–180.
- Tebaldi C, Nychka D, Mearns LO. 2004a. From global mean responses to regional signals of climate change: simple pattern scaling, its limitations (or lack of) and the uncertainty in its results. In *Proceedings of the 18th Conference on Probability and Statistics in the Atmospheric Sciences, AMS Annual Meeting*, Seattle, WA.
- Tebaldi C, Mearns LO, Nychka D, Smith RL. 2004b. Regional probabilities of precipitation change: a Bayesian analysis of multimodel simulations. *Geophysical Research Letters* **31**: L24213, DOI: 10.1029/2004GL021276.
- Tebaldi C, Smith RL, Nychka D, Mearns LO. 2005. Quantifying uncertainty in projections of regional climate change: a Bayesian approach to the analysis of multi-model ensembles. *Journal of Climate* **18**: 1524–1540.
- Todini E. 1996. The ARNO rainfall-runoff model. *Journal of Hydrology* **175**: 339–382.
- Trinka M, Dubrovsky M, Semeradova D, Zalud Z. 2005. Projections of uncertainties in climate change scenarios into projected winter wheat yields. *Theoretical and Applied Climatology* **77**: 229–249.
- van Rheeën NT, Wood AW, Palmer RN, Lettenmaier DP. 2004. Potential implications of PCM climate change scenarios for Sacramento–San Joaquin river basin hydrology and water resources. *Climatic Change* **62**: 257–281.
- von Storch H. 1999. On the Use of “Inflation” in Statistical Downscaling. *Journal of Climate* **12**: 3505–3506.
- von Storch H, Zwiers F. 1999. *Statistical Analysis in Climate Research*. Cambridge University Press: Cambridge.
- von Storch H, Zorita E, Cubasch U. 1993. Downscaling of global climate change estimates to regional scales: an application to Iberian Rainfall in wintertime. *Journal of Climate* **6**: 1161–1171.
- Wang Y, Leung LR, McGregor JL, Lee DK, Wang WC, Ding Y, Kimura F. 2004. Regional climate modeling: progress, challenges, and prospects. *Journal of the Meteorological Society of Japan* **82**: 1599–1628.
- Watts M, Goodess CM, Jones PD. 2004. *The CRU Daily Weather Generator*, BETWIXT Technical Briefing Note 1, Version 2, February 2004.
- Wetterhall F, Bárdossy A, Chen D, Halldin S, Xu C. 2006. Daily precipitation-downscaling techniques in three Chinese regions. *Water Resources Research* **42**: W11423, DOI: 10.1029/2005WR004573.
- Widmann M, Schär C. 1997. A principal component and long-term trend analysis of daily precipitation in Switzerland. *International Journal of Climatology* **17**: 1333–1356.
- Widmann M, Bretherton CS. 2000. Validation of mesoscale precipitation in the NCEP Reanalysis using a new gridcell dataset for the Northwestern United States. *Journal of Climate* **13**: 1936–1950.
- Widmann M, Bretherton CS, Salathé EP. 2003. Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. *Journal of Climate* **16**: 799–816.
- Wigley TML, Raper SCB. 2001. Interpretation of high projections for global-mean warming. *Science* **293**: 451–454.
- Wigley TML, Jones PD, Briffa KR, Smith G. 1990. Obtaining subgrid scale information from coarse-resolution general circulation model output. *Journal of Geophysical Research* **95**: 1943–1953.
- Wilby RL. 1998. Statistical downscaling of daily precipitation using daily airflow and seasonal teleconnection indices. *Climate Research* **10**: 163–178.
- Wilby RL. 2006. When and where might climate change be detectable in UK river flows? *Geophysical Research Letters* **33**: L19407, DOI: 10.1029/2006GL027552.
- Wilby RL, Wigley TML. 1997. Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography* **21**: 530–548.
- Wilby RL, Wigley TML. 2000. Precipitation predictors for downscaling: observed and general circulation model relationships. *International Journal of Climatology* **20**: 641–661.
- Wilby RL, Harris I. 2006. A framework for assessing uncertainties in climate change impacts: low flow scenarios for the River Thames, UK. *Water Resources Research* **42**: W02419, DOI: 10.1029/2005WR004065.
- Wilby RL, Hay LE, Leavesley GH. 1999. A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado. *Journal of Hydrology* **225**: 67–91.
- Wilby RL, Conway D, Jones PD. 2002a. Prospects for downscaling seasonal precipitation variability using conditioned weather generator parameters. *Hydrological Processes* **16**: 1215–1234.
- Wilby RL, Dawson CW, Barrow EM. 2002b. SDSM – a decision support tool for the assessment of regional climate change impacts. *Environmental Modelling & Software* **17**(2): 145–157.

- Wilby RL, Tomlinson OJ, Dawson CW. 2003. Multi-site simulation of precipitation by conditional resampling. *Climate Research* **23**: 183–194.
- Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO. 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods, *Supporting material of the Intergovernmental Panel on Climate Change*, available from the DDC of IPCC TG CIA, 27.
- 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. *Journal of Hydrology* **330**: 204–220.
- 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 Resources Research* **34**: 2995–3008.
- 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. *Geophysical Research Letters* **27**: 1199–1202.
- Wilks DS. 1992. Adapting stochastic weather generation algorithms for climate change studies. *Climatic Change* **22**: 67–84.
- Wilks DS, Wilby RL. 1999. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography* **23**: 329–357.
- 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.
- Xu CY. 1999. From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. *Progress in Physical Geography* **23**(2): 229–249.
- Yates D, Gangopadhyay S, Rajagopalan B, Strzepek K. 2003. A technique for generating regional climate scenarios using a nearest-neighbour algorithm. *Water Resources Research* **39**: 1199, DOI: 10.1029/2002WR001769.
- Zhu C, Pierce DW, Barnett TP, Wood AW, Lettenmaier DP. 2004. Evaluation of hydrologically relevant PCM climate variables and large-scale variability over the continental U.S. *Climatic Change* **62**: 45–74.
- Zorita E, von Storch H. 1997. A survey of statistical downscaling techniques. GKSS report 97/E/20, GKSS Research Center: Geesthacht.
- Zorita E, von Storch H. 1999. The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate* **12**: 2474–2489.