Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods

RL Wilby[1,2], SP Charles[3], E Zorita[4], B Timbal[5], P Whetton[6], LO Mearns[7]

[1]Environment Agency of England and Wales, UK
[2]King's College London, UK
[3]CSIRO Land and Water, Australia
[4]GKSS, Germany
[5]Bureau of Meteorology, Australia
[6]CSIRO Atmospheric Research, Australia
[7]National Center for Atmospheric Research, USA

August 2004

Document history

The *Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods* constitutes “Supporting material” of the Intergovernmental Panel on Climate Change (as defined in the *Procedures for the Preparation, Review, Acceptance, Adoption, Approval, and Publication of IPCC Reports*). The *Guidelines* were prepared for consideration by the IPCC at the request of its Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA). This supporting material has not been subject to the formal intergovernmental IPCC review processes.

The first draft of the *Guidelines* was produced by Rob Wilby (Environment Agency of England and Wales) in March 2003. Subsequently, Steven Charles, Penny Whetton, and Eduardo Zorito provided additional materials and feedback that were incorporated by June 2003. Revisions were made in the light of comments received from Elaine Barrow and John Mitchell in November 2003 – most notably the inclusion of a worked case study. Further comments were received from Penny Whetton and Steven Charles in February 2004. Bruce Hewitson, Jose Marengo, Linda Mearns, and Tim Carter reviewed the *Guidelines* on behalf of the Task Group on Data and Scenario Support for Impacts and Climate Analysis (TGICA), and the final version was published in August 2004.

Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods

1. INTRODUCTION

The climate change information required for many impact studies is of a spatial scale much finer than that provided by global or regional climate models. The ensuing problems for impact assessment have been recognised for a long time (Kim et al., 1984; Gates, 1985; Robinson and Finkelstein, 1989; Lamb, 1987; Smith and Tirpak, 1989; Cohen, 1990). Global climate models (GCMs) have resolutions of hundreds of kilometres whilst regional climate models (RCMs) may be as fine as tens of kilometres. However, many impact applications require the equivalent of point climate observations and are highly sensitive to fine-scale climate variations that are parameterised in coarse-scale models. This is especially true for regions of complex topography, coastal or island locations, and in regions of highly heterogeneous land-cover.

The most straightforward means of obtaining higher spatial resolution scenarios is to apply coarse-scale climate change projections to a high resolution observed climate baseline – the change factor method. This method is often used when RCM output are unavailable, for sensitivity studies, or whenever rapid assessments of multiple climate change scenarios (and/or GCM experiments) are required. Fine resolution climate change information for use in impact studies can also be obtained via more sophisticated statistical downscaling (SD) methods but such studies have, to date, largely restricted themselves to the use of a single driving GCM. The purpose of this report is to provide background information and guidance on the application of SD methods for climate scenario development. Guidance material on the use of regional modelling for climate scenario development is provided in a companion document (Mearns et al., 2003).

Statistical downscaling is based on the view that the regional climate is conditioned by two factors: the large scale climatic state, and regional/local physiographic features (e.g. topography, land-sea distribution and land use; von Storch, 1995, 1999). From this perspective, regional or local climate information is derived by first determining a statistical model which relates large-scale climate variables (or "predictors") to regional and local variables (or "predictands"). Then the large-scale output of a GCM simulation is fed into this statistical model to estimate the corresponding local and regional climate characteristics. One of the primary advantages of these techniques is that they are computationally inexpensive, and thus can be easily applied to output from different GCM experiments. Another advantage is that they can be used to provide site-specific information, which can be critical for many climate change impact studies. The major theoretical weakness of SD methods is that their basic assumption is not verifiable, i.e., that the statistical relationships developed for the present day climate also hold under the different forcing conditions of possible future climates – a limitation that also applies to the physical parameterizations of dynamical models.

To date, most of the SD approaches described in this document are practiced by climatologists rather than by impact analysts undertaking fully fledged, policy orientated impact assessments. This is because the scenarios have largely been regarded as unreliable, too difficult to interpret, or do not embrace the range of uncertainties in GCM projections in the same way that simpler interpolation methods do. This means that downscaled scenarios based on single GCMs or emission scenarios, when translated into an impact study, can give the misleading impression of increased resolution equating to increased confidence in the projections. However, it is increasingly recognized that comprehensive impact studies must be founded on multiple GCM outputs.

Part 2 of this guidance note provides researchers of climate impacts with background material and descriptions of the main SD techniques. We also outline some of the key assumptions and limitations applying to their usage. Part 3 continues by advising researchers to consider some broader questions about downscaling approaches, whether statistical downscaling or regional modelling. For example, is high-resolution information really needed for the application? Is the effort involved in producing high-resolution climate information appropriate in the context of all of the uncertainties associated with the project? Many of these wider issues are addressed by the companion document on regional climate modelling (Mearns et al., 2003). They are not repeated here so we recommend that both guides be read in tandem. Rather, the focus of Part 3 will be on the practical aspects of statistical downscaling implementation, beginning with the study objectives themselves. Finally, a worked case study is provided in Part 4, and a summary of the key recommendations attached to the proper implementation of SD methods is given in Part 5.

2. REVIEW OF METHODS OF STATISTICAL DOWNSCALING

Statistical downscaling involves developing quantitative relationships between large–scale atmospheric variables (predictors) and local surface variables (predictands). The most common form has the predictand as a function of the predictor(s), but other types of relationships have been used. For example, between predictors and the statistical distribution parameters of the predictand (Pfizenmayer and von Storch, 2001) or the frequencies of extremes of the predictand (Katz *et al.*, 2002). Most SD work has focussed on single-site (i.e., point scale) daily precipitation as the predictand because it is the most important input variable for many natural systems models and cannot be obtained directly from climate model output. Predictor sets are typically derived from sea level pressure, geopotential height, wind fields, absolute or relative humidity, and temperature variables. These variables are archived at the grid resolution of operational and re-analysis climate models, with the horizontal resolution typically 300–500 km. However, the grid spacing of the observed climate fields and GCM climate change projection output do not always correspond. Therefore, driving a statistical downscaling model with GCM output often requires interpolation of the GCM fields to the grid resolution of the atmospheric predictor sets used in fitting.

The following sections outline the main SD techniques under the broad headings ‘weather classification’, ‘regression models’, and ‘weather generators’. This categorization is similar to that used by IPCC TAR WG1 section 10.6 (Giorgi *et al.*, 2001). **Table 1** provides a summary of their relative strengths and weakness.

2.1 Weather classification schemes

Weather classification methods group days into a finite number of discrete weather types or “states” according to their synoptic similarity. Typically, weather states are defined by applying cluster analysis to atmospheric fields (Corte-Real *et al.*, 1999; Huth, 2000; Kidson, 2000; Hewitson and Crane, 2002) or using subjective circulation classification schemes (Bardossy and Caspary, 1990; Jones *et al.*, 1993). In both cases, weather patterns are grouped according to their similarity with ‘nearest neighbours’ or a reference set. The predictand is then assigned to the prevailing weather state, and replicated under changed climate conditions by resampling or regression functions (Hay *et al.*, 1991; Corte-Real *et al.*, 1999). Classification-based methods can have limited success in reproducing the persistence characteristics of at–site wet and dry spells (e.g., Wilby, 1994). Recent approaches include extensions to multi–site and multi–variate series (e.g., precipitation and temperature as in Bardossy and van Mierlo (2000) or Palutikof *et al.* (2002)).

Analogue approaches are examples of a weather classification method in which predictands are chosen by matching previous (i.e., analogous situations) to the current weather-state. The method was originally designed by Lorenz (1969) for weather forecasting applications but was abandoned due to its limited success. It has resurfaced for climate applications (Zorita *et al.*, 1995; Martin *et al.*, 1997) since longer series of predictors have emerged following the completion of reanalysis projects (e.g., Kalnay *et al.*, 1996). Even so, the analogue method still suffers whenever the pool of training observations is limited (Timbal *et al.*, 2003) and/or the number of classifying predictors is large (Van den Dool, 1989). However, it compares favourably with more complex regression methods (Zorita and von Storch, 1999) and is suitable for providing multi-site and multi-variate series (Timbal and McAvaney, 2001).

Another approach is to classify spatial rainfall occurrence patterns using hidden Markov models, then infer corresponding synoptic weather patterns (Hughes and Guttorp, 1994; Hughes *et al.*, 1999). A hidden Markov model represents a doubly stochastic process, involving an underlying (hidden) stochastic process that is translated into another stochastic process that yields the sequence of observations (Rabiner and Juang, 1986). The observed process (e.g., precipitation occurrence at a network of sites) is conditional on the hidden process (the weather states). Weather states evolve according to a first order Markov chain, in which transitions from one state to the next have fixed probabilities and depend only on the current state. Alternatively, non-homogeneous hidden Markov models have transition probabilities that are conditioned by atmospheric predictors and thus vary in time. These models reproduce key characteristics of precipitation such as interannual variability, occurrence and persistence of wet and dry spells at individual sites, and correlations between precipitation series for pairs of sites (Hughes and Guttorp, 1994; Charles *et al.*, 1999a).

2.2 Regression models

Regression models are a conceptually simple means of representing linear or nonlinear relationships between predictands and the large scale atmospheric forcing. Commonly applied methods include multiple regression (Murphy, 1999), canonical correlation analysis (CCA) (von Storch *et al.*, 1993), and artificial neural networks which are akin to nonlinear regression (Crane and Hewitson, 1998). von Storch (1999) and Bürger (1996) discuss the important issue of under prediction of variance often associated with regression approaches. The problem is particularly evident for daily precipitation downscaling because of the relatively low predictability of local amounts by large-scale forcing alone. Bürger (2002) uses an approach termed ‘expanded downscaling’ to increase the variance of simulated predictor series. Some multi-site regression-based methods are also becoming available in which the unexplained variance is represented by stochastic processes (e.g., Charles et al., 1999a; Wilby *et al.*, 2003).

2.3 Weather generators

Weather generators (WGs) are models that replicate the statistical attributes of a local climate variable (such as the mean and variance) but not observed sequences of events (Wilks and Wilby, 1999). These models are based on representations of precipitation occurrence via Markov processes for wet-/ dry-day or spell transitions. Secondary variables such as wet-day amounts, temperatures and solar radiation are often modelled conditional on precipitation occurrence (e.g., dry-days in summer may have on average more sunshine than wet-days). WGs are adapted for statistical downscaling by conditioning their parameters on large-scale atmospheric predictors, weather states or rainfall properties (Katz, 1996; Semenov and Barrow, 1997; Wilks, 1999). However, parameter modification for future climate scenarios can lead to unanticipated outcomes (Wilks, 1992). For example, changes to parameters governing wet-/dry-spell lengths can affect simulated temperatures and solar radiation even before modifications are applied to the parameters governing these variables. Moreover, WGs based on first-order Markov chains (i.e., one-state-to-the-next transitions) often underestimate temporal variability and persistence of precipitation (Gregory et al., 1993; Mearns et al., 1996; Katz and Parlange, 1998). However, conditioned WG methods are useful for temporal downscaling, for instance disaggregating monthly precipitation totals and rain days into daily amounts, or daily totals into sub-daily components (Kilsby et al., 1998; Fowler et al., 2000).

Table 1 A summary of the strengths and weaknesses of the main SD methods.

Method	Strengths	Weaknesses
Weather typing (e.g. analogue method, hybrid approaches, fuzzy classification, self organizing maps, Monte Carlo methods).	• Yields physically interpretable linkages to surface climate • Versatile (e.g., can be applied to surface climate, air quality, flooding, erosion, etc.) • Compositing for analysis of extreme events	• Requires additional task of weather classification • Circulation-based schemes can be insensitive to future climate forcing • May not capture intra-type variations in surface climate
Weather generators (e.g. Markov chains, stochastic models, spell length methods, storm arrival times, mixture modelling).	• Production of large ensembles for uncertainty analysis or long simulations for extremes • Spatial interpolation of model parameters using landscape • Can generate sub-daily information	• Arbitrary adjustment of parameters for future climate • Unanticipated effects to secondary variables of changing precipitation parameters
Regression methods (e.g. linear regression, neural networks, canonical correlation analysis, kriging).	• Relatively straightforward to apply • Employs full range of available predictor variables • 'Off-the-shelf' solutions and software available	• Poor representation of observed variance • May assume linearity and/or normality of data • Poor representation of extreme events

2.4 Key Assumptions

It is important to be aware of several key assumptions when downscaling climate model output for current and projected climates (Hewitson and Crane, 1996; Giorgi *et al.*, 2001):

- Predictors relevant to the local predictand should be adequately reproduced by the host climate model at the spatial scales used to condition the downscaled response(s). Prior knowledge of climate model limitations can be advantageous when screening potential predictors. Therefore, predictors have to be chosen on the balance of their relevance to the target predictand(s) and their accurate representation by climate models (see: Wilby and Wigley, 2000). This necessarily places some onus on the downscaling community to undertake GCM verification, at least for the predictors of interest.
- The relationship between the predictors and predictand remains valid for periods outside the fitting period (time invariance). This needs careful assessment for future climate projection as it is obviously impossible to check with observational data. A way around this problem is to validate the statistical downscaling model with observational data stemming from periods well separated from the fitting period, i.e. representing a "different" climate regime (e.g., Charles *et al.*, 2004). Since projected climate change may lie (partly) outside the bounds of the long-term variability of the observational record, this is not a completely satisfactory solution (see below). Thus, there is value in checking equivalent statistical relationships between predictor and predictand in climate model simulations for future climates (Gonzalez-Rouco *et al.*, 2000). Here the caveat is the limited ability of climate models to simulate the local variable, the origin of the whole downscaling "problem".
- The predictor set sufficiently incorporates the future climate change 'signal'. Some approaches, for example stepwise regression, may exclude predictors based on current climate performance that could be important in future changed climates. In order to test predictor stability, Charles *et al.* (1999b) compared CSIRO RCM 2×CO_2 grid–scale daily precipitation occurrence probabilities to those obtained by driving a downscaling model

fitted to 1×CO_2 RCM grid–scale precipitation with 2×CO_2 RCM atmospheric predictors. The downscaling model driven by 2×CO_2 RCM atmospheric predictors reproduced the 2×CO_2 RCM grid–scale precipitation occurrence probabilities only when the predictor set included information on lower atmosphere moisture saturation. Although not validation in the traditional sense, this approach increases confidence in the choice of predictors and showed that relationships derived during fitting remained legitimate for the changed climate. Busuioc *et al.* (1999) applied a similar method to a CCA of monthly precipitation data for sites in Romania.

- The predictors used for determining future local climate should not lie outside the range of the climatology used to calibrate the SD model. If this is the case, then technically the SD model is invalid. Preliminary explorations of the output of ECHAM4 suggest that the assumption of 'stationarity' is robust for geopotential heights across most of the Southern Hemisphere, but less so for atmospheric moisture in the tropics where up to 90% of the data lie outside the climatology of the control simulation (**Figure 1**). Further research is needed to establish whether these patterns are representative of the wider suite of predictors routinely used for statistical downscaling, as well as of other GCMs.

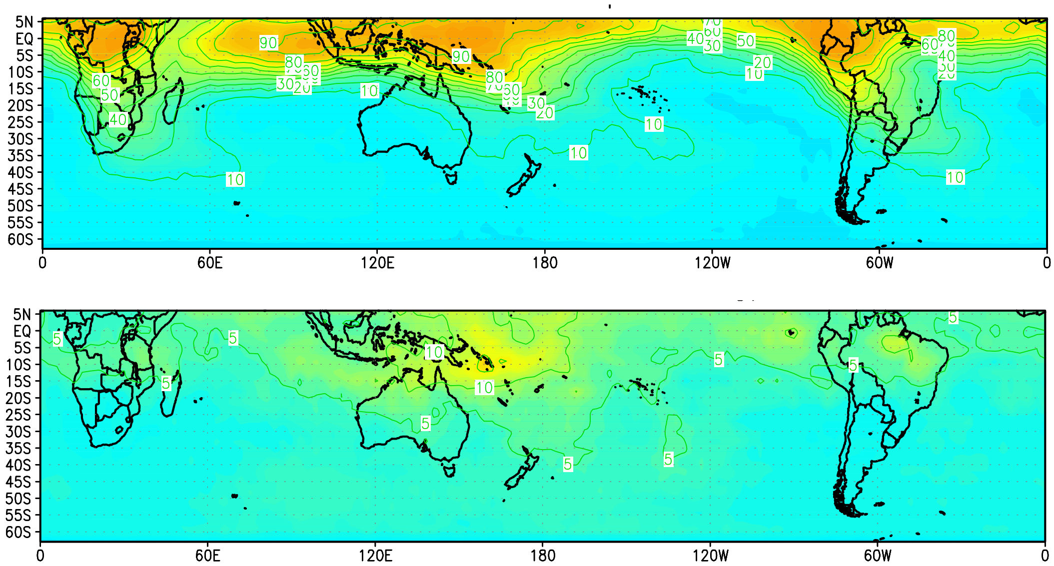

Figure 1 Percentage of days that specific humidity (upper panel) and 850 geopotential heights (lower panel) for the climate of the 2080s under the A2 emissions scenario lie outside the range of the 1961-1990 climatology of the ECHAM4 GCM. Source: Hewitson (2004).

2.5 Inter-comparison studies

A growing number of studies have compared several SD methods or compared statistical downscaling with dynamical (i.e., RCM–based) downscaling. For example, Wilby and Wigley (1997) and Wilby *et al.* (1998) compared six SD approaches (two neural nets, two weather generators, and two vorticity–based regression methods) for multiple sites across the USA using observed and GCM data. The vorticity–based regression methods were found to perform best. Performance criteria were root mean squared errors of the following diagnostics: wet day amount mean, median, standard deviation and 95th percentile; dry–dry and wet–wet day occurrence probabilities; wet day probabilities; wet and dry spell duration mean, standard deviation, and 90th percentile; and standard deviation of monthly precipitation totals. Although the GCM yielded large changes in precipitation it projected only small changes in the circulation–based predictors used — within the limits of modelled interannual variability. The various SD approaches gave significantly different future precipitation scenarios despite using common sets of GCM predictors. Given this ambiguity, Wilby *et al.*

(1998) suggested the need for additional atmospheric predictors such as moisture–based predictors to capture long-term changes in atmospheric saturation.

Kidson and Thompson (1998) compared results from a RCM with a screening regression downscaling technique based on indices of local and regional airflow. Using data from 1980 to 1994 for a network of 78 sites across New Zealand, the regression approach better explained the daily variance in precipitation anomalies. The relatively poor performance of the RCM was attributed to its inability to resolve orography, a result of the 50 km grid spacing used. It was concluded that the (linear) regression relationships would remain valid provided that predictors extend only slightly beyond the range of the observed data used in calibration. However, it was preferable to use RCMs whenever significant changes in factors such as atmospheric vapour content could influence storm intensity.

Mearns *et al.* (1999) compared a circulation classification approach (*k*–means clustering of the principal components of 700 hPa geopotential height fields) to the NCAR RegCM2 RCM nested within the CSIRO Mk 2 GCM of Watterson *et al.* (1997) for 5 years of $1\times CO_2$ and $2\times CO_2$ runs. The RCM reproduced monthly or seasonal precipitation for 12 sites in the eastern Nebraska study area quite well, partly due to compensating errors in the frequency (overestimated by a factor of 2 to 5) and intensity of precipitation events (underestimated by a factor of 2 to 14). The classification-based approach reproduced observed precipitation characteristics for the same sites, but this is expected as the model was calibrated and then validated against observed station data. The climate change projections, however, exemplify the problem of different results arising from different approaches. The two approaches did not produce mean precipitation changes of the same direction for 40% of months and locations investigated. The statistical downscaling yielded mainly increases in mean precipitation, whereas RegCM2 produced both increases and decreases for coherent subregions.

Zorita and von Storch (1999) compared an analogue method (Zorita *et al.*, 1995) to (i) a linear regression method based on CCA applied to monthly site precipitation totals and sea level pressure (SLP) fields; (ii) a method based on classification and regression trees applied to daily precipitation occurrence and SLP fields; and (iii) a neural network, as an example of a nonlinear method, applied to daily precipitation amounts and SLP anomalies for the Iberian Peninsula (southwest Europe). In general, the analogue method performed comparably, or was best at reproducing daily precipitation amounts and frequency characteristics as well as being technically simpler to implement.

Widmann *et al.* (2003) have compared a Singular Value Decomposition downscaling method with a simpler method of local rescaling of simulated rainfall (Widmann and Bretherton, 2000). The simpler method performed surprising well, especially when the resampling is improved by dynamical corrections. They were also able to demonstrate that using model rainfall as a predictor improved the downscaling techniques (see Salathé, 2003).

Finally, there is a growing number of statistical downscaling versus dynamical downscaling studies that use hydrological indices for assessing relative skill under present boundary forcing supplied by re-analysis data. This research has helped identify systematic biases in precipitation and temperature at the basin-scale, as well as the role of snowpack and soil moisture stores in integrating and/or canceling these errors at seasonal time-scales (e.g., Hay and Clark, 2003; Wilby et al., 2000; Wood et al., 2004) (**Figure 2**). Other work has focused on the reproduction of extreme events such as persistent wet- or dry-spells (e.g., Goodess et al., 2003) for flood frequency estimation (e.g., Reynard et al., 2004). Collectively, such studies indicate that SD and RCM based methods have comparable skill over both daily and monthly time-scales *at least* for present climate conditions.

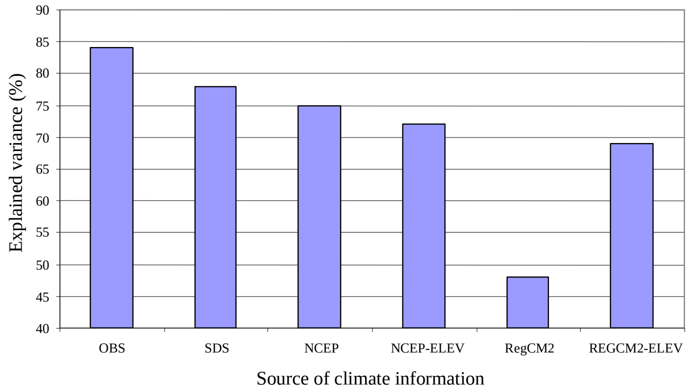

Figure 2 Percentage of explained variance in observed daily flows modelled using station (OBS), statistically downscaled (SDS), reanalyis (NCEP), elevation corrected reanalyis (NCEP-ELEV), RCM (RegCM2) and elevation corrected RCM (RegCM2-ELEV) daily precipitation and temperature series for the Animas River basin, Colorado. Source: Wilby et al. (2000).

2.6 Issues for statistical downscaling

Choice of statistical method. The choice of statistical method is to some extent determined by the nature of the local predictand. A local variable that is reasonably normally distributed, such as monthly mean temperature, will require nothing more complicated than (multiple) regression, since large scale climate predictors tend to be normally distributed – assuming linearity of the relationship. A local variable that is highly heterogeneous and discontinuous in space and time, such as daily precipitation, will probably require a more complicated non-linear approach or transformation of the raw data. Fitting such complex models often requires large amounts of observational data.

Choice of predictors. Sometimes the best predictors identified in the statistical analysis of observations are not completely adequate for climate change applications. For example, daily rainfall may be determined by geopotential heights in the extratropical areas. But changes in geopotential heights caused by global warming will contain a non-dynamical signal, which will spuriously affect the estimation of rainfall changes. This non-dynamical component should be corrected, either by subtracting the average changes of the geopotential height in a sufficiently large area, or by using geopotential thickness, instead of geopotential heights, as predictors (Burkhardt, 1999). Conversely, exclusion of key predictors for future change, perhaps due to a high degree of covariance with another variable under current climates, may result in a critical loss of information about future regional response to changes in large scale forcing.

Extremes. Statistical downscaling models are often calibrated in ways that are not particularly designed to handle extreme events for which fewer realisations are available. In many cases, the implementation of a SD technique is most successful at reproducing the mean of the signal. This is important to keep in mind for impact studies. [The 2003-2005 EU project STARDEX "Statistical and Regional dynamical Downscaling of Extremes for European regions" is investigating such issues, see: http://www.cru.uea.ac.uk/projects/stardex/].

The tropical regions. These regions may present more complex behaviour than the mid-latitudes. In the tropics, the strong ocean-atmosphere coupling makes the consideration of the role of the ocean unavoidable, thus enlarging the set of potential large-scale predictors. Also the relationships between large-scale predictors and local variables may vary strongly within the annual cycle. In the case of precipitation, statistical models especially designed for a particular month (start or end of rainy season) may be required (Jimoh and Webster, 1999).

Feedbacks
Under weak synoptic forcing, other climate subsystems, such as vegetation, may come into play in a changed climate. The role of these subsystems can be critical in feedback processes governing, for example, the onset date of seasonal rains, the development of extreme convective systems, or the reinforcement of persistent states such as dry spells. The natural response of these feedback processes to global change, or anthropogenic forcing (e.g., land use practice) can potentially seriously compromise statistical models fitted to the observational record (Zheng *et al.*, 2002).

3. GUIDELINES

With the above considerations in mind, the following sections should assist readers in deciding whether the time and resource implications of statistical downscaling are justified in terms of added value to the project. Before this, however, it is wise to consider whether or not SD techniques are appropriate for the task and needed to achieve project aims or, in fact, whether simpler procedures can deliver comparable results. Having made the decision to proceed with a SD approach, it is helpful to identify the most important technical considerations *en route* to the production of downscaled climate scenarios.

3.1 Issues to consider in deciding whether SD is required for a project

As noted at the outset, the most commonly cited rationale for downscaling is that GCMs provide only a "broad–brush" view of how climate variables, such as global temperature and rainfall patterns, might change in the future in response to rising concentrations of anthropogenic greenhouse gases. GCMs cannot resolve important processes relating to sub-grid scale cloud and topographic effects that are of significance to many impact studies. For example, assessments of future river flows may require (sub-) daily precipitation scenarios at catchment, or even station scales. Therefore, downscaling is often justified on the grounds that climate information is needed at higher temporal and/or spatial resolutions than currently delivered by GCM output. However, it is important to recognise that increased precision of downscaling does not necessarily translate to increased confidence in regional scenarios. The following sections outline when it may or may not be appropriate to employ SD methods.

3.1.1 Situations when it may be appropriate to use SD methods

SD methods are particularly useful in heterogeneous environments with complex physiography or steep environmental gradients (as in island, mountainous or land/sea contexts) where there are strong relationships to synoptic scale forcing. Indeed, SD may be the only practicable means of generating climate scenarios for point-scale processes such as soil erosion (e.g. Favis-Mortlock and Boardman, 1995). A further justification for SD is the need for better sub- GCM grid-scale information on extreme events such as heat-waves (e.g., Schubert and Henderson-Sellers, 1997), heavy precipitation (e.g., Olsson et al., 2001) or localised flooding (e.g., Pilling and Jones, 2002). A further very real pragmatic reason is when there are severe limitations on computational resources, especially in developing nations where arguably the greatest need exists.

Empirical relationships between atmospheric predictors and sub-daily climatic statistics also provide a framework for interpreting observed trends in extreme events, and for downscaling future events. For example, Bárdossy (1997) showed, using 30 years of 5 minute precipitation data for sites in the Ruhr valley, that the probability of a wet-hour and number of wet spells in a day are conditional on the season and prevailing circulation pattern. Precipitation scenarios at such fine temporal and spatial resolution are needed in order to improve the design and evaluate the future performance of urban drainage systems (see Bronstert et al., 2002).

Table 2 Selected examples of the use of SD for exotic predictands. Adapted from Giorgi et al. (2001).

Flowering times of *Galanthus nivalis* L.
Snow cover in the French Alps, snow duration in Austria
Spawning time of North Sea fish stocks
Slope stability and landsliding in SE France
Storm surge statistics for Germany
Zooplankton populations in the Netherlands
Wave heights and salinity off the Polish coast
Palaeo proxy records for Norwegian glaciers
Lake stratification and water temperature profiles in Japan
Air pollution episodes in London

A major advantage of SD methods over RCMs is their low computational demand. This can allow generation of large ensembles of climate realizations and the exploration of some aspects of climate uncertainty (due to SD model parameters and/or natural climate variability). SD schemes are also very flexible in the sense that for any local variable with predictability, a trans-scale relationship can usually be found. For example, **Figure 3** shows the weak but statistically significant associations that exist between an index of the intensity of London's urban heat island and two regional-scale predictor variables (standardised mean sea level pressure and wind speed over Eastern England). **Table 2** provides further examples of exotic predictands that are not directly supplied by either GCMs or RCMs but have been produced using SD methods.

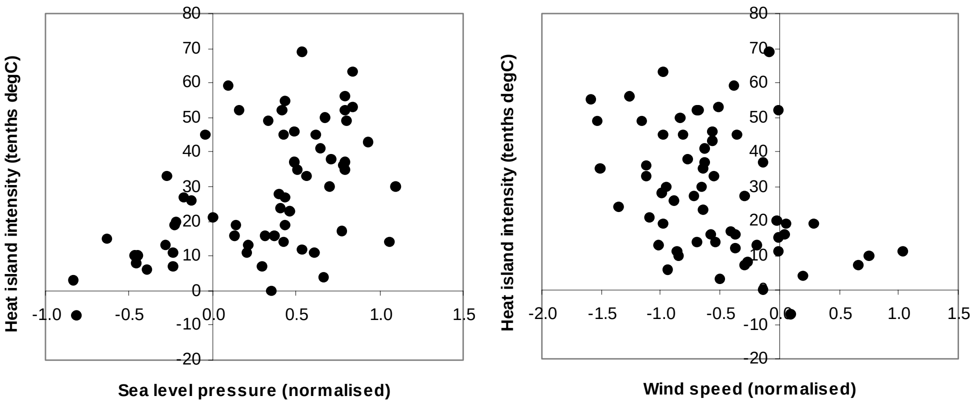

Figure 3 Association between London's nocturnal urban heat-island intensity and a) mean sea level pressure, and b) wind speed, July to August 1995. Source: LCCP (2002).

3.1.2 Situations when it may NOT be appropriate to use SD methods

To date, the majority of downscaling studies have been undertaken for temperate, mid–latitude regions of the Northern Hemisphere; relatively few have examined semi–arid or tropical locations. There has also been an elevational bias towards low altitude (often coastal) sites arising from concerns about data homogeneity and network density. In other words, application of SD schemes is restricted to data-rich areas; unlike RCMs, most SD methods can not be applied unless station data are available for model calibration. Possible exceptions include the distributing of weather generator parameters across landscapes using topographic information from digital elevation models (e.g., Kittel et al., 1995).

In contrast to RCMs, the majority of SD schemes are unable to incorporate land-surface forcing – the regional climate response is driven entirely by the free atmosphere predictor variables supplied by the GCM. This means that the climate change scenarios produced by conventional SD methods will be insensitive to changes in land-surface feedbacks. However, it is increasingly recognised that local land use practices influence regional climate, vegetation and runoff regimes in adjacent natural areas (Chase et al., 2001; Kalnay and Cai, 2003; Stohlgren et al., 1998; Reynard et al., 2001). Human modified rural landscapes tend to have a lower albedo, and higher surface roughness and soil moisture than natural surfaces. The net effect of these physical changes is to partition a larger proportion of solar energy into latent heat (associated with evapotranspiration) and less into sensible heating of the overlying atmosphere. As in the case of RCMs, land-surface related changes in cloudiness due to elevated moisture fluxes and atmospheric instability will not be captured by SD methods without explicit conditioning of model parameters in line with the evolving land-use.

Following from the above, it is a widely recognised that SD techniques should not be applied whenever statistical transfer functions are deemed to be temporally unstable. This may be a real consideration when faced with marked changes in atmospheric circulation regimes or abrupt climate changes such as a collapse of the Atlantic thermohaline circulation (Vellinga and Wood, 2002). Reliance upon circulation-based predictors alone will capture only this component of the climate change signal and, even for observed data, can fail to capture all aspects of multi-decadal variability (Wilby, 1997). For example, Widmann and Schär (1997) showed that observed trends in daily precipitation across Switzerland were not due to changes in the frequency of circulation patterns, but rather to changes *within* rain–producing weather types. Elsewhere, detailed analyses of circulation and humidity predictors have assisted explanations of abrupt climate changes. For example, the marked decline in winter precipitation over south-west Australia in the mid-1970s was linked to changes in the frequency of high pressure systems centred to the east of the region and to changes in the moisture content of the lower troposphere (CSIRO, 2001). This again highlights the need to represent all aspects of the regional forcing (not just circulation patterns), so that the SD model remains valid under changing climatic conditions.

3.1.3 Alternatives to SD for scenario generation and impact assessment

With the above factors in mind, it is acknowledged that there are many alternative techniques for generating high-resolution climate change scenarios other than the application of RCM and SD schemes. These approaches include the spatial interpolation of grid-point data to the required local-scale (sometimes called “unintelligent” downscaling); climate sensitivity analysis of impact models (also known as the “bottom-up” approach); construction of spatial/temporal analogues using historic climate data; and the use of simple change factors for strategic assessments. The latter two are discussed in more detail below.

Climate change analogues are developed from climate records that may typify the future climate of a given region. The analogue can originate from either past climate data (temporal analogue) or from another region (spatial analogue). A major advantage of the analogue approach is that the future climate scenario and associated impacts may be described at far greater temporal and spatial resolutions than might otherwise be possible. For example, in the UK the hot/dry summers of 1976 and 1995, and the mild/wet winters of 1990/91 and 1994/95 are thought to provide useful temporal analogues for future climate change (Subak et al., 2000; Palutikof et al., 2004). By the 2080s, under a high emissions scenario, approximately one summer in three is expected to be hotter and drier than the extraordinary summer of 1995 (Hulme et al., 2002).

One of the more straightforward and popular procedures for rapid impact assessment involves the use of “change factors” (e.g., Arnell, 2003a;b; Arnell and Reynard, 1996; Diaz-Nieto and Wilby, 2004; Eckhardt and Ulbrich, 2003; Pilling and Jones, 1999; Prudhomme et al., 2002). Firstly, the reference climatology is established for the site or region of interest. Depending on the application this might be a representative long-term average such as 1961-1990, or an actual meteorological record such as daily maximum temperatures. Secondly, changes in the equivalent temperature variable for the GCM grid–box closest to the target site are calculated. For example, a difference of 2.5˚C might occur by subtracting the mean GCM temperatures for 1961-1990 from the mean of the 2050s. Thirdly, the temperature change suggested by the GCM (in this case, +2.5˚C) is then simply added to each day in the reference climatology.

Although the resultant scenario incorporates the detail of the station records as well as the areal average climate change of the specified GCM grid–box, there are problems with this method. The scaled and the base–line scenarios only differ in terms of their respective means, maxima and minima; all other properties of the data, such as the range and variability remain unchanged. The procedure also assumes that the spatial pattern of the present climate remains unchanged in the future. Furthermore, the method does not easily apply to precipitation records because the addition (or multiplication) of observed precipitation by GCM precipitation changes can affect the number of rain days, the size of extreme events, and even result in negative precipitation amounts! When direct scaling is applied to the baseline precipitation series, the temporal sequencing is unchanged, so the method may not be helpful in circumstances where changes in wet-/dry-spell lengths are important to the impact assessment, such as in semi-arid and arid hydrology where runoff response to rainfall amount and timing are highly non-linear. Most critically, this approach fails to recognize that values of single grid cells are not representative of the GCM skilful scale. Consequently, this is arguably a problematic approach.

3.2 Technical guide on the appropriate application of SD schemes

The use of SD for the development of regional climate change scenarios generally requires less physical reasoning than a nested RCM simulation (but in both cases 'hidden' parameters can significantly affect model outcomes). However, the value and confidence placed in SD scenarios is enhanced by adherence to 'good-practice' guidelines. The following sections identify some of the most important step-by-step considerations (**Figure 4**).

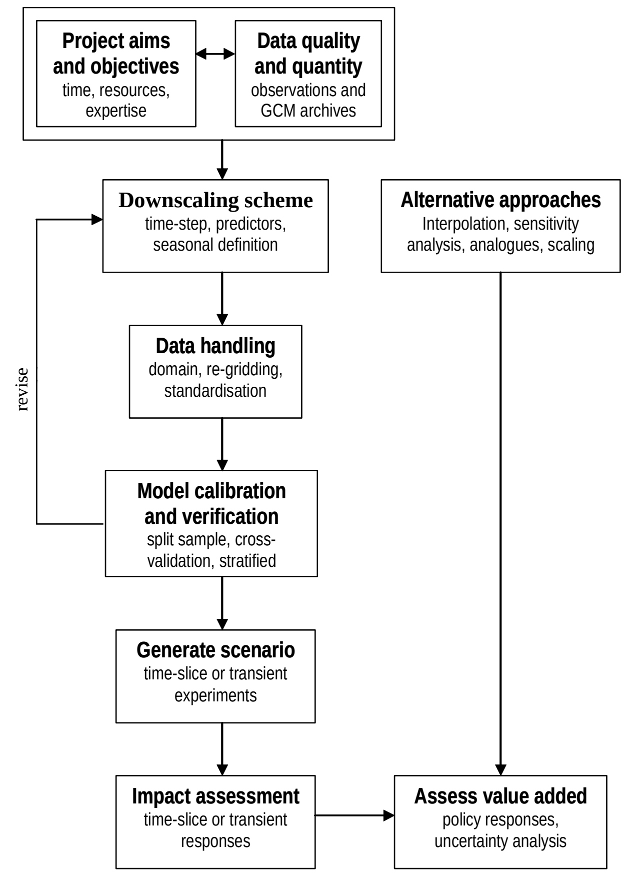

Figure 4 Key stages in statistical downscaling.

3.2.1 Specify the study objectives

Many technical aspects of SD hinge upon data quality and quantity, time and resources available, and on whether single- or multi-site information is required. A useful starting point for any SD investigation is a "bottom-up" assessment of the key climate sensitivities of the system(s) of interest (Beersma et al., 2000). For example, the reliable yield of a reservoir may depend upon the length of dry-spells at critical times of the year. Persistent dry-spells may, in turn, be linked to the occurrence of specific weather patterns over the region. This leads us to question the ability of the host GCM to replicate such weather patterns for the present climate, as well as changes in the future. Next, we must decide on the most appropriate transfer scheme and seasonal stratification for downscaling (dry-spells), and on how best to evaluate model skill. Finally, knowledge of system sensitivities or threshold states, allows us to judge whether projected changes in the downscaled scenario (dry-spell length) are of *practical* significance when compared to the cascade of model uncertainties.

3.2.2 Assess availability and quality of archived data

The viability of all SD techniques rests critically upon access to high-quality predictands and predictor variables at the space and time scales of intended use. It should also be recognised at the outset that few meteorological stations have data sets that are 100% complete and/or fully accurate. Therefore, handling of missing and imperfect meteorological records is necessary for most practical situations.

Fortunately, the downscaling community is now widely consulted by the leading climate modelling centres concerning 'wish-lists' of archived predictor variables. For example, following a survey of user needs the UK Hadley Centre archived over 20 daily variables from their HadCM3SRES A2 and B2 experiments (including temperature, humidity, energy and dynamical variables at several levels in the atmosphere). Some groups, such as the Canadian Climate Impacts Scenarios (CCIS) Project have even begun supplying gridded predictor variables on-line (see: http://www.cics.uvic.ca/scenarios/sdsm/select.cgi), with the specific needs of the downscaling community in mind. In addition, a large suite of secondary variables such as atmospheric stability, vorticity, divergence, zonal and meridional airflows may be derived from standard daily variables such as mean sea level pressure or geopotential height (Conway et al., 1996). Nonetheless, there is a general need for better meta-data from climate centres to ensure appropriate and consistent usage of their products (Beersma et al., 2000).

A major caveat underpinning both RCM and SD methodologies is that the GCM output(s) used for downscaling are realistically modelled by the GCM. The same assumption applies to the re-analysis products that are now widely used to calibrate SD models for the present climate. Verification of climate model output at the space and time scales of use is, therefore, a necessary precursor to all downscaling exercises because SD techniques propagate the uncertainty in the driving fields of the GCM, and do not improve on the base skill of the GCM (Hewitson and Crane, 2003). For example, evaluations of mean sea level pressures in the NCEP re–analysis revealed a positive bias from 1941–1967 across parts of the North Atlantic (Reid et al., 2001). Quantities such as temperatures or geopotential heights are better represented by GCMs at the regional scale than derived variables such as precipitation, but there is still an onus upon the downscaling community to verify the realism of inter-variable relationships used by multivariate SD methods.

3.2.3 Specify model type and structure

The range of downscaling techniques and applications has increased significantly since the IPCC TAR (see above). The question naturally arises as to which family of SD methods should be employed? The relative merits of each downscaling approach have already been discussed in detail elsewhere; **Table 1** summarised the key attributes of the main SD methods. Comparative studies indicate that the skill of SD techniques depends on the chosen application and region of interest. This reflects the considerable versatility of SD approaches, regarded by some as a distinct strength of empirical downscaling (Giorgi et al., 2001). In practice, the choice of model type often reflects the availability of data, ease of access to existing models (including supporting documentation), and the nature of the problem at hand (whether uni- or multi-variate, single or multi-site, etc.)

Recent studies also vary with respect to the seasonal stratification of data prior to model calibration. Primary considerations include the time-step of the SD model (hourly, daily, monthly averages, etc.) and whether SD models should be developed for individual months, seasons or years (such as wet and dry episodes). In some instances, conventional climatological seasons (i.e., December-February, March-May, and so on) may not reflect 'natural' seasons contained in data so alternative delimitations may be required (Winkler et al., 1997). Furthermore, rigid classifications of data using seasonal definitions based on present climate behaviour may not be valid under altered climate conditions. Under these circumstances model parameters should be allowed to vary at sub-seasonal time-scales. However, in some situations (e.g., downscaling precipitation amounts in semi-arid regions) it will be necessary to group data into seasons simply to ensure sufficient (wet-day) cases for model calibration.

3.2.4 Select appropriate predictor variables

The choice of predictor variable(s) is one of the most influential steps in the development of a SD scheme because the decision largely determines the character of the downscaled scenario. The selection process is complicated by the fact that the explanatory power of individual predictor variables may be low (especially for daily precipitation), or the power varies both spatially and temporally (**Figure 5**). Some of the earliest downscaling studies used GCM grid-box values of the predictand (e.g., area-average temperatures) to derive station-scale values of the same variable (e.g., local surface temperature), or employed monthly means as the predictor of daily quantiles (e.g., Wigley et al., 1990; Wilks, 1989).

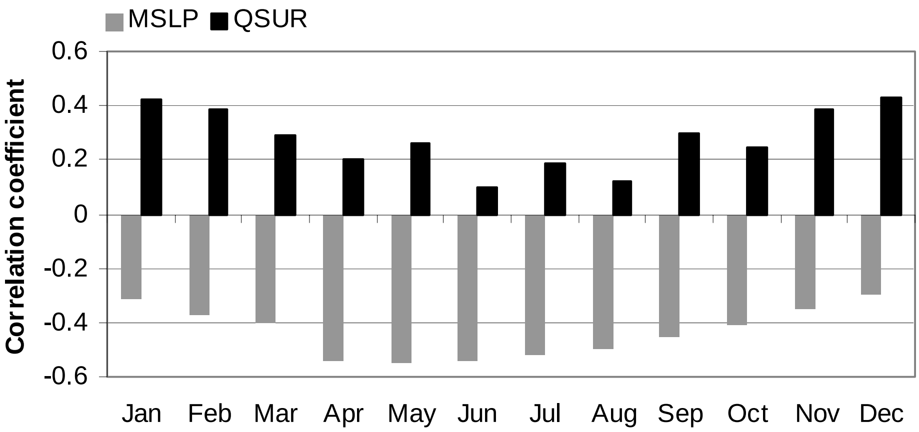

Figure 5 Monthly variations in the strength of the correlation between daily wet–day amounts at Eskdalemuir (55º 19' N, 3º 12' W) and mean sea level pressure (MSLP), and near surface specific humidity (QSUR) over the Scottish Borders region, 1961–1990. Source: Wilby et al. (2003).

The availability of re-analysis data sets has significantly increased the number and variety of candidate predictors. Unfortunately, there have been relatively few systematic assessments of different predictors (e.g., Charles et al., 1999b; Huth, 1999; Wilby and Wigley, 2000; Winkler et al., 1997). Furthermore, potentially useful predictors may be overlooked because of limited explanatory power for the present climate. For example, it is suspected that future changes in surface temperature may be dominated by changes in the radiative properties of the atmosphere rather than by circulation changes (Schubert, 1998), a response accommodated by employing both temperature and circulation fields in the downscaling process (Huth, 1999). For precipitation downscaling it is known that inclusion of humidity variables can have a significant bearing on the outcome, not only changing the magnitude of future changes but also the sign of the changes (Hewitson, 1999; Charles et al., 1999b).

Low frequency climate variability can be under-predicted by transfer function and weather typing SD methods. Some have addressed this deficiency by conditioning model parameters on slowly varying predictors such as large-scale pressure indices (Katz and Parlange, 1996) or sea surface temperatures (Wilby et al., 2002). Alternatively, the missing variability can be added by perturbing model parameters using a low-frequency stochastic model (Hansen and Mavromatis, 2001). In either event, it must be assumed that the conditioning indices are realistically produced by GCMs, and continue to be relevant to future regional climate forcing (see Osborn et al., 1999).

Ultimately, the choice of downscaling predictors is constrained by the data archived from GCM experiments because the range of re-analysis products generally exceeds that retrievable for individual GCM runs. With this in mind, simple procedures such as partial correlation analysis, step-wise regression, or information criterion may be used to screen the most promising predictor variables from a candidate suite (Charles et al., 1999b; Wilby et al., 2003). Expert judgement or local knowledge bases are also invaluable sources of information when choosing sensible combinations of predictors from the available data. The ideal SD predictor variable is strongly correlated with the target variable; makes physical sense; is realistically represented by the GCM; and captures multiyear variability. Most critically, the predictors should collectively reflect the climate change signal – meeting all other criteria yet missing this will lead to a seriously erroneous climate change scenario.

3.2.5 Specify the downscaling domain

It is necessary to specify the location and dimensions of the large-scale predictor field(s) for downscaling local weather variables (analogous to the choice of lateral meteorological boundary conditions used to drive high-resolution RCM simulations). The smaller the predictor domain, the more direct the influence of the host GCM on the downscaled scenario. The *location* of the downscaling domain is important because the skill of GCMs at reproducing observed climatology varies between models and is not uniform across space or time (see Lambert and Boer, 2001). The positioning of the domain should also reflect the dominant processes affecting the region in question (such as the trajectory of mid-latitude cyclones, influence of water bodies, orography, etc.). If the positioning of the downscaling domain is inadequate, subtle variations in present and future storm track position, for example, may not be captured.

The optimum *area* of the domain depends to a certain extent on the theoretical level of aggregation at which the host GCM is skilful – generally held to be at least several grid-points (Widmann and Bretherton, 2000). Domain averages or principal components of predictor fields have the advantage of capturing spatial patterns of predictor variable behaviour at larger spatial scales but downplay the significance of local forcing. Continental-scale predictor domains are applicable when using slowly varying properties such as sea surface temperatures (Wilby et al., 2002) or atmospheric circulation indices (Katz and Parlange, 1993) to condition multi-year variability in downscaled scenarios.

Finally, it is noted that the optimal location and dimensions of the predictor domain may vary by region. For example, as **Figure 6** shows, daily wet-day amounts across the Sierra Nevada and Oklahoma are correlated with near *in situ* humidity variables but with mean sea level pressure fields located to the west of the target station(s) (Brinkmann, 2002; Wilby and

Wigley, 2000). Use of correlation surfaces like those in **Figure 6** can be helpful in identifying the optimal geographic location and extent of predictor domains (Stidd, 1954), but it must also be recognised that the spatial pattern of predictor-predictand associations may change under altered climate conditions.

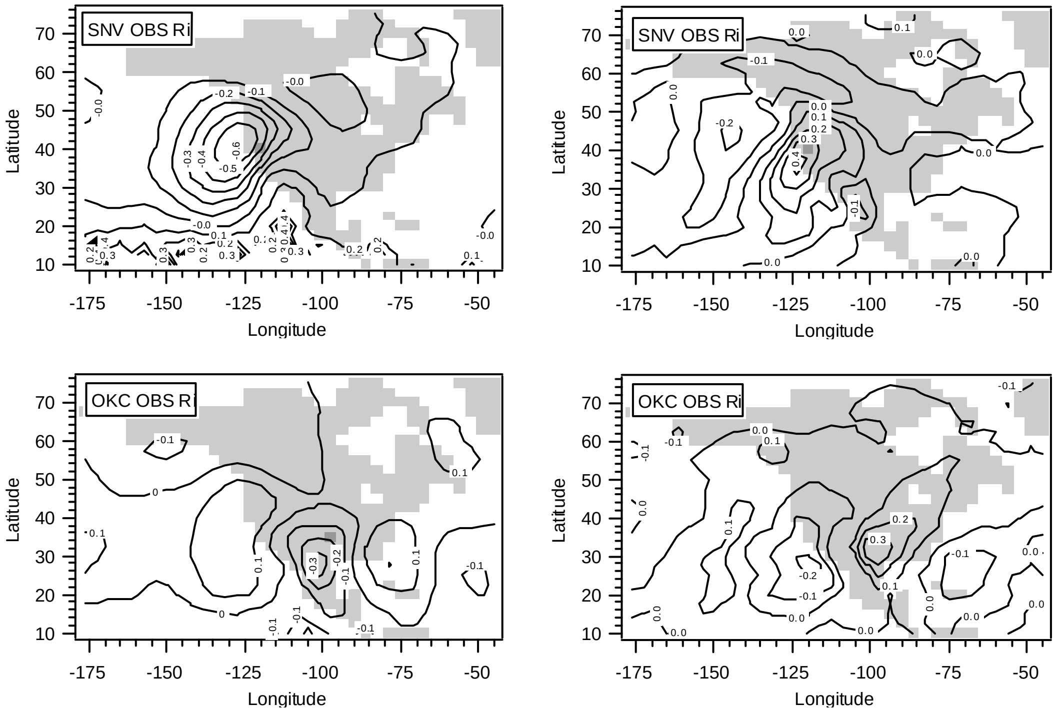

Figure 6 The correlation between observed daily wet-day amounts in winter in Sierra Nevada (SNV, top row) and Oklahoma (OKC, bottom row) and re-analysis a) mean sea level pressure (left column); b) surface specific humidity (right column). Source: Wilby and Wigley (2000).

3.2.6 Re-gridding and standardisation of data

The assembly of a candidate predictor suite can be an involved process entailing data extraction, re-gridding and standardisation techniques. Re-gridding is often needed because the grid-spacing and/or co-ordinate systems of observed and re-analysis data sets (used for SD model calibration) do not always correspond to the grid-spacing and co-ordinate systems of the GCM output (used to develop regional climate change scenarios). For example, the NCEP/NCAR re-analysis (Kalnay et al., 1996) has a grid-spacing of 2.5º latitude by 2.5º longitude whereas the HadCM3 model has a slightly coarser resolution of 2.5º latitude by 3.75º longitude. A popular re-gridding technique is to compute the weighted average of neighbouring grid-points, where the weighting decreases with separation distance following a gaussian curve up to a specified maximum separation distance.

Standardisation is widely used prior to SD to reduce systematic biases in the mean and variance of GCM predictors relative to observations (or re-analysis data). The procedure typically involves subtraction of the mean and division by the standard deviation of the predictor for a predefined baseline period. The main issues here relate to the choice of the baseline and averaging window (whether monthly, seasonal or annual). The period 1961-1990 is widely used as a baseline because it is of sufficient duration to establish a reliable climatology, yet not too long, nor too contemporary to include a strong global change signal.

3.2.7 Evaluate model skill using independent data

There is a very real danger that SD methods will be applied uncritically as “black boxes”, particularly when employing regression-based modelling techniques (and an equal danger exists with the popularisation of RCMs running on PCs). Ideally, the downscaling should be based upon physically sensible linkages between large-scale forcing and local meteorological response(s). Therefore, best practice demands rigorous evaluation of SD schemes (and equally of RCMs) using independent data. At the same time, we should be mindful of the fact that observational data used for model evaluation – such as gridded precipitation amounts – introduce further sources of uncertainty due to changes in the spatial and temporal coverage of networks, or due to the non-homogeneity of individual site records.

A standard approach to model validation involves the use of split-records: one portion for model calibration and the remainder for testing. This method is appropriate when long (>30-year) observation records are available. However, cross-validation techniques may be a more effective use of shorter records, or subsets of data (such as odd/even years) when trends are suspected. Alternatively, the model may be developed using data drawn from dry years and tested against data from wet-years, or *vice versa* (Wilks, 1999). A further strategy of ‘multiple working hypotheses’ involves the intercomparison of different statistical transfer schemes (e.g., Zorita and von Storch, 1999), or ‘benchmarking’ of SD model skill relative to RCMs (e.g., Murphy, 1999).

Success of the SD under present climate conditions does not imply that the model is valid under future climate conditions. This is because transfer functions may become invalid or the weights attached to different predictors may change. For example, atmospheric moisture content does exert some control over present-day precipitation occurrence and amounts, but is expected to assume even greater significance in the future (Hewitson, 1999). Tests of the stationarity of statistical transfer schemes using comparable relationships in RCMs, suggest that the assumption of stationarity may be robust provided that the choice of predictors is judicious (Charles et al., 1999b).

3.2.8 Generate downscaled scenarios

Having calibrated and verified the SD model performance, it is then necessary to generate ensembles of synthetic daily weather series given standardised atmospheric predictor variables supplied by the GCM (representing either the present or future climate). The predictor variables may originate from either time-slice or transient model experiments. In the former case, downscaling provides higher resolution climate scenarios for discrete periods (e.g., 2020s, 2050s, 2080s, etc.) or for equilibrium response experiments that attach no explicit times to changes in greenhouse gas forcing (only differences in atmospheric concentrations such as $1xCO_2$ or $2xCO_2$ scenarios). Alternatively, continuous downscaling of regional predictands may be undertaken using the GCM output of transient experiments. This enables downscaling of the time-varying properties of the regional climate as they evolve throughout the course of the 21st century (**Figure 7**).

Some SD techniques are capable of simultaneously delivering multiple outputs such as precipitation, maximum and minimum temperatures, solar radiation, relative humidity and wind speed (e.g., Parlange and Katz, 2000). Indeed, this is a common requirement of many impact studies. However, where predictors are downscaled independently it is necessary to verify that inter-variable relationships are realistically preserved (for instance, that the maximum daily temperature is always greater than the minimum). This is particularly important whenever analysing joint probabilities of events such as the dependence between sea surge, river flow and precipitation (Svensson and Jones, 2002).

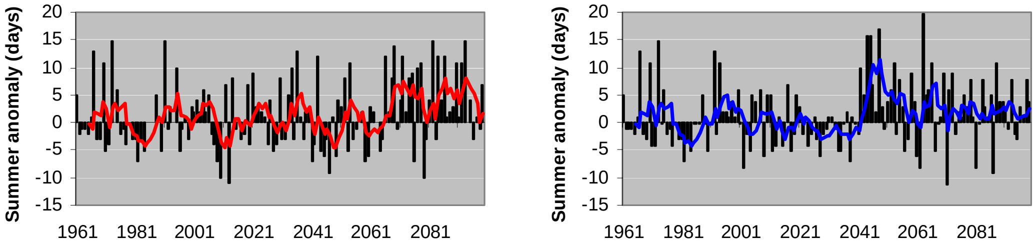

Figure 7 Changes in the frequency of summer weather patterns favouring pollution episodes over eastern England downscaled from HadCM3 output from the Medium-High Emissions (left column) and Medium-Low Emissions (right column) SRES scenarios. All anomalies are expressed with respect to the 1961 to 1990 average. Source: LCCP (2002).

Ideally, downscaling should be performed using output from a wide range of climate model experiments in order to represent the uncertainties attached to different emission scenarios, model structures, parameterization schemes, and climate sensitivities (Mearns et al., 2001). Provided that the predictor variables of different climate models have been standardised in the same way and that they are representing identical atmospheric phenomena (another good reason for high quality meta-data), repeat downscaling experiments may be undertaken using the same model structure/ transfer functions but different sources of driving variables. This is practicable for most SD methods but is seldom undertaken. For RCMs, this may be prohibitively costly unless achieved via pattern-scaling (as in Hulme et al., 2002) or risk assessment approaches (e.g., New and Hulme, 2000).

3.2.9.Evaluate 'value-added' of downscaling

Having developed a set of scenarios it is important to evaluate the extent to which the downscaling has contributed value to the impact assessment above and beyond the use of raw GCM output or simpler scaling approaches (see above). The most straightforward test is to assess the realism of GCM and downscaled variables relative to observed climatology under present climate conditions, at the temporal and spatial scale of the intended impact (Hay et al., 2000). A further step involves the comparison of derived variables generated by, for example, water balance (Wilby et al., 2000), flood frequency (Reynard et al., 2004) or agricultural (Mearns et al., 1999) impact models driven by downscaled or raw GCM output. In this case, it is important to recognise that the impact models themselves introduce additional layers of uncertainty to that of the downscaling algorithm. Furthermore, if the impact model output depends upon the temporal or spatial integration of climate information (as in the case of changes in soil moisture, groundwater or snow pack volume), dependent model processes (such as runoff) can imply higher skill than the finer resolution inputs. For example, realistic time-series of daily precipitation amounts in winter may not be produced by the downscaling but seasonal accumulations of precipitation in the snow pack, and hence spring-melt, may well be. Finally, it has been suggested that seasonal forecasting may be a useful framework for testing downscaling because the forecasts are for hitherto unseen events, yet the models may be verified 'on-line' as new data become available (Leung et al., 2003).

4. CASE STUDY

The following case study describes the use of a SD procedure for the rapid assessment of potential changes in soil moisture deficits and groundwater recharge in two headwaters of the River Thames, UK (**Figure 8**). The national UKCIP02 scenarios (Hulme et al., 2002) suggest a decrease in average soil moisture both annually and in summer that is most marked in southeast England. In the Thames region, 55% of the effective rainfall that falls annually is abstracted, amounting to about 5000 Ml/d, of which 86% is used for public water supply. Even without climate change, the present balance of supply and demand is in deficit by some

180 Ml/d (Environment Agency, 2001). The example in **Table 3** is used to illustrate the key steps described in the technical guidance (above).

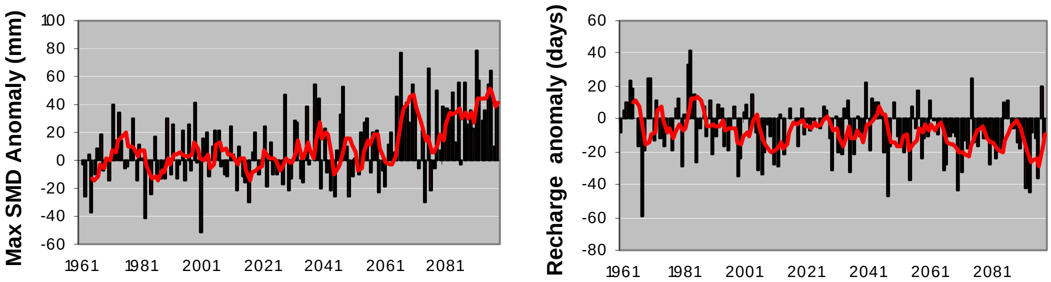

Figure 8 Changes in maximum soil moisture deficits (SMDs) and length of recharge season (days) in the River Kennet under A2 emissions, compared with the 1961-1990 average. Source: LCCP (2002).

Table 3 Procedure for the rapid assessment of future changes in soil moisture and groundwater recharge across two tributary catchments in the River Thames basin.

1. Specify the study objectives	The objective of the study was determine possible changes in soil moisture deficits and the duration of the groundwater recharge season under the A2 and B2 SRES scenarios by the 2020s , 2050s and 2080s for two tributaries of the River Thames (the Kennet and Loddon). Only two working days were available for the climate change impact assessment, suggesting the need for 'off-the-shelf' modelling solutions.
2. Assess availability and quality of archived data	Daily precipitation and runoff series, as well as monthly potential evaporation (PE) were available for the period 1961-2000. A fully calibrated rainfall-runoff model (CATCHMOD) was obtained for each river catchment. Daily predictor variables originated from the NCEP re-analysis for the period 1961-1990 and from the coupled ocean-atmosphere GCM HadCM3 for the A2 and B2 SRES scenarios for the period 1961-2099. All data were subject to rigorous quality checks.
3. Specify model type and structure	Drier soils at the end of the water year mean that more precipitation is required for re-wetting to saturated conditions under which groundwater recharge or surface runoff are assumed to occur. As a consequence, the length of the recharge season depends critically upon the timing and duration of dry-spells. This dictated that the SD method should operate on a daily time-step, produce realistic sequences of wet- and dry-spells, be applicable to PE as well as precipitation, and capture seasonality. A hybrid stochastic weather generator/ regression SD model was chosen.
4. Select appropriate predictor variables	Predictor variable selection was guided by stepwise regression and by 'expert' knowledge of large-scale atmospheric controls of precipitation and PE in the Thames basin. Several transfer functions were applied to the predictands (precipitation and PE) and predictors (NCEP variables). Predictor-predictand relationships were evaluated using multiple performance criteria against data for 1961-1990. Separate monthly models were built to capture the observed seasonality in local precipitation and PE.
5. Specify the downscaling domain	Previous analyses had indicated that the optimal predictor domains for downscaling precipitation across the region were the two HadCM3 grid-boxes located over southeast and eastern England (Goodess et al., 2003).
6. Re-gridding and standardisation of data	The NCEP re-analysis predictors were re-gridded to conform to the grid-spacing of HadCM3 using the weighted average of neighbouring grid-points. All NCEP and HadCM3 predictors were standardised by their respective means and standard deviations of the baseline period 1961-1990.
7. Evaluate model skill using independent data	The SD model was extensively tested using a split-sample approach. The most robust precipitation model had five predictor variables (mean sea level pressure, surface vorticity, meridional flow, and specific humidity, and 850hPa geopotential heights). In line with previous work, it was found that the SD model captured precipitation occurrence better than wet-day amounts.

8. Generate downscaled scenarios	The calibrated SD model was driven using HadCM3 predictors for the full transient run 1961-2099 under A2 and B2 emission scenarios. The resulting daily precipitation and PE series were used as inputs to CATCHMOD to compute water balance changes. The length of the recharge season reduces by 25% in the Kennet by the 2080s under the A2 emissions scenario (**Figure 8**).
9. Evaluate 'value-added' of downscaling	The assessment of changes in recharge would not have been plausible using raw GCM output because of the under representation of long dry-spells by the model. Experiments are ongoing to assess the relative merits of RCM and SD scenarios for water resource estimation at the catchment scale, as well as the uncertainty due to the choice of GCM used for downscaling.

5. SUMMARY RECOMMENDATIONS

The following recommendations capture the essence of the guidance on the use of statistical downscaling for climate scenarios development:

- **Carefully consider the objectives of the climate change impact study** and the potential added value of higher resolution climate scenarios. Are the time and resources involved in the production of the statistically downscaled scenarios really justified, or can comparable outcomes be achieved using more straightforward procedures such as raw GCM output, change factors, scaling methods or interpolation to finer scales?
- **Be aware of the generic strengths and weaknesses of statistical downscaling**. On the one hand SD methods may be helpful for impact studies in heterogeneous environments (e.g., islands, mountains, land/sea contrasts); whenever point-scale information is needed (e.g., localised flooding, soil erosion, urban drainage, etc.); for modelling exotic predictands (e.g., heat island indices, flowering times, wave heights, etc.); or for generating large ensembles and/or transient scenarios. On the other hand, SD methods can be data-demanding, are typically applied "off-line" (i.e., do not incorporate important land-surface feedbacks) and assume stationary predictor-predictand relationships over multi-decadal time-scales.
- Assuming that statistical downscaling is applicable, **be aware of the specific strengths and weaknesses attached to the various families of SD methods**. Before specifying the SD model type and structure evaluate the data requirements in terms of local station data (predictands) and large-scale, archived climate model information (predictors), including available meta-data. Recognise that verification (and possibly re-gridding) of GCM predictors is a time-consuming but necessary part of SD model development.
- When selecting the optimum set of large-scale predictors try to incorporate as much local knowledge of the relevant processes. **Recognise that downscaling predictability and skill varies** seasonally, as a consequence of the size and positioning of the predictor field, and between different periods of record. Remember that the ideal SD predictor is strongly correlated with the target variable, is physically sensible, well represented in the GCM control run, and captures multi-year variability.
- **Test the SD model using independent data**. Confidence in projections for future climate change scenarios will be increased if it can be demonstrated that SD model parameters are temporally stable. Then apply the model to a wide range of climate models to evaluate the uncertainties associated with different emission scenarios, GCM structures, parameterizations and climate sensitivities. Otherwise, any resulting impact studies should only be regarded as sensitivity studies of the few scenarios. **Where possible apply RCM scenarios in parallel to explore the uncertainty due to choice of downscaling method(s) (recognizing that a given RCM solution may be no nearer the truth)**.
- Finally, apply the SD scenarios to the climate change impact assessment to inform strategic decision-making and policy. **Constantly re-evaluate the added value or new insight(s) that have been gained through the use of higher resolution scenarios** above and beyond that arising from raw GCM output. Share the outcomes and experience with the broader climate modelling and impacts community.

6. ACKNOWLEDGEMENTS

The authors are extremely grateful to Elaine Barrow, Tim Carter, Bruce Hewitson and John Mitchell for their constructive remarks on earlier versions of the document.

7. REFERENCES

Arnell, N.W. 2003a. *Effects of climate change on river flows and groundwater recharge using the UKCIP02 scenarios*. Report to UK Water Industry Research Limited. University of Southampton.

Arnell, N. W. 2003b. Relative effects of multi-decadal climatic variability and changes in the mean and variability of climate due to global warming: future streamflows in Britain. *Journal of Hydrology*, **270**,19-213.

Arnell, N. W. and Reynard, N. S. 1996. The effects of climate change due to global warming on river flows in Great Britain. *Journal of Hydrology*, **183**, 397-424.

Bárdossy, A and Caspary H.J. 1990. Detection of climate change in Europe by analyzing European atmospheric circulation patterns from 1881 to 1989. *Theoretical and Applied Climatology* **42**, 155-167.

Bárdossy, A and van Mierlo J.M.C. 2000. Regional precipitation and temperature scenarios for climate change. *Hydrological Sciences Journal*, **45**, 559-575.

Bárdossy, A. 1997. Disaggregation of daily precipitation. In Casale, R., Borga, M., Baltas, E. and Samuels, P. (Eds.) Proceedings of the Workshop on Ribamod – River Basin Modelling, Management and Flood Mitigation Concerted Action, Padua, Italy, 25-26 September. 1997. EU Environment and Climate Programme, pp40-55.

Beersma, J., Agnew M.D., Viner D., and Hulme M. 2000. Climate Scenarios for Water-Related and Coastal Impacts. Proceedings of the EU Concerted Action Initiative ECLAT-2 Workshop 3. KNMI, Netherlands, May 10-12th 2000, Climatic Research Unit, Norwich 140pp.

Brinkmann, W.A.R. 2002. Local versus remote grid points in climate downscaling. *Climate Research*, **21**, 27-42.

Bronstert, A., Niehoff, D. and Burger, G. 2002. Effects of climate and land-use change on storm runoff generation: present knowledge and modelling capabilities. *Hydrological Processes*, **16**, 509-529.

Burger, G. 1996. Expanded downscaling for generating local weather scenarios. *Climate Research*, **7**, 111-128.

Bürger, G. 2002. Selected precipitation scenarios across Europe. *Journal of Hydrology*, **262**, 99-110.

Burkhardt, U. 1999. Alpine precipitation in a tripled CO2-climate. *Tellus, A* **51**, 289-303.

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

Charles, S.P., Bates B.C. and Hughes J.P. 1999a. A spatio-temporal model for downscaling precipitation occurrence and amounts. *Journal of Geophysical Research*, **104**, 31657-31669.

Charles, S.P., Bates B.C., Whetton P.H. and Hughes J.P. 1999b. Validation of downscaling models for changed climate conditions: case study of southwestern Australia. *Climate Research*, **12**, 1-14.

Charles, S.P., Bates B.C. and Hughes J.P. 2004. Statistical downscaling of daily precipitation from observed and modelled atmospheric fields. *Hydrological Processes* **in press**

Chase, T.N., Pielke, R.A., Kittel, T.G.F., Zhao, M., Pitman, A.J., Running, S.W. and Nemani, R.R. 2001. Relative climate effects of landcover change and elevated carbon dioxide combined with aerosols: a comparison of model results and observations. *Journal of Geophysical Research-Atmospheres*, **106**, 31685-31691.

Cohen, S.J. 1990. Bringing the global warming issue closer to home: the challenge of regional impact studies. *Bulletin of the American Meteorological Society,* **71**, 520-526.

Conway, D., Wilby, R.L. and Jones, P.D. 1996. Precipitation and air flow indices over the British Isles. *Climate Research*, **7**, 169–183.

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

Crane, R.G and Hewitson, B.C. 1998. Doubled CO_2 precipitation changes for the Susquehanna basin: down-scaling from the GENESIS general circulation model. *International Journal of Climatology*, **18**, 65-76.

CSIRO 2001. Towards an understanding of climate variability in south western Australia. Second Research Report of the Indian Ocean Climate Initiative, Bureau of Meteorology Research Centre.

Diaz-Nieto, J. and Wilby, R.L. 2004. The impact of climate change on low flows in the River Thames UK: A comparison of statistical downscaling and change factors. *Climatic Change*, in press.
Eckhardt, K. and Ulbricht, U. 2003. Potential impacts of climate change on groundwater recharge and streamflow in a central European low mountain range. *Journal of Hydrology*, **284**, 244-252.
Environment Agency 2001. *Water resources for the future: a strategy for the Thames Region*. Environment Agency: Reading.
Favis-Mortlock, D.T. and Boardman, J. 1995. Nonlinear responses of soil erosion to climate change: a modelling study of the UK South Downs. *Catena*, **25**, 365-387.
Fowler, H.J., Kilsby, C.G. and O'Connell, P.E. 2000. A stochastic rainfall model for the assessment of regional water resource systems under changed climatic conditions. *Hydrological and Earth Systems Science*, **4**, 263-282.
Gates, W.L. 1985. The use of general circulation models in the analysis of the ecosystem impacts of climatic change. *Climate Change*, **7**, 267-284.
Giorgi, F., Hewitson, B., Christensen, J., Hulme, M., Von Storch, H., Whetton, P., Jones, R., Mearns, L., Fu, C. 2001. Regional climate information - evaluation and projections, in *Climate Change 2001: The Scientific Basis,* Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Xiaosu D, Eds., Cambridge University Press, Cambridge, 583-638
Gonzalez-Rouco, J,F, Heyen H., Zorita E. and Valero F. 2000. Agreement between observed rainfall trends and climate change simulations in the southwest of Europe. *Journal of Climate,* **13**, 3057-3065.
Goodess, C., Osborn, T. and Hulme, M. 2003. *The identification and evaluation of suitable scenario development methods for the estimation of future probabilities of extreme weather events*. Tyndall Centre for Climate Change Research, Technical Report 4. http://www.tyndall.ac.uk/research/theme3/final_reports/it1_16.pdf
Gregory, J.M., Wigley, T.M.L. and Jones, P.D. 1993. Application of Markov models to area-average daily precipitation series and interannual variability in seasonal totals. *Climate Dynamics*, **8**, 299-310.
Hansen, J.W. and Mavromatis, T. 2001. Correcting low-frequency variability bias in stochastic weather generators. *Agricultural and Forest Meteorology*, **109**, 297-310.
Hay, L.E, McCabe G.J., Wolock D.M. and Ayers, M.A. 1991. Simulation of precipitation by weather type analysis. *Water Resources Research,* **27**, 493-501.
Hay, L.E. and Clark, M.P. 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.
Hay, L.E., Wilby, R.L. and Leavesley, G.H. 2000. A comparison of delta change and downscaled GCM scenarios for three mountainous basins in the United States. *Journal of the American Water Resources Association*, **36**, 387–397.
Hewitson, B.C. 2004. Empirical downscaling: assumptions, caveats and stationarity. *Paper presented at the Ninth International Meeting on Statistical Climatology*, Cape Town, South Africa, 24-28 May 2004.
Hewitson, B.C. 1999. Deriving regional precipitation scenarios from General Circulation Models. Water Research Commission Report 751/1/99, Pretoria, South Africa, 40pp.
Hewitson, B.C. and Crane R.G. 1996. Climate downscaling: techniques and application. *Climate Research*, **7**, 85-95.
Hewitson, B.C. and Crane, R.G. 2002. Self-organizing maps: applications to synoptic climatology. *Climate Research*, **22**, 13-26.
Hewitson, B.C. and Crane, R.G. 2003. Global climate change: Questions of downscaled climate scenarios for impact assessment. *Bulletin of the American Meteorological Society,* submitted.
Hughes, J.P and Guttorp, P. 1994. A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena. *Water Resources Research*, **30**, 1535-1546.
Hughes, J.P., Guttorp, P. and Charles S.P.. 1999. A non-homogeneous hidden Markov model for precipitation occurrence. *Applied Statistics*, **48**, 15-30.
Hulme, M., Jenkins, G.J., Lu, X., Turnpenny, J.R., Mitchell, T.D., Jones, R.G., Lowe, J., Murphy, J.M., Hassell, D., Boorman, P., McDonald, R. and Hill, S. 2002. Climate Change Scenarios for the UK: The UKCIP02 Scientific Report, Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK. 120pp.
Huth, R. 1999. Statistical downscaling in central Europe: Evaluation of methods and potential predictors. *Climate Research*, **13**, 91-101.
Huth, R. 2000. A circulation classification scheme applicable in GCM studies. *Theoretical and Applied Climatology*, **67**, 1-18.

Jimoh, O.D. and Webster, P. 1999. Stochastic modelling of daily rainfall in Nigeria: intra-annual variation of model parameters. *Journal of Hydrology*, **222**, 1-17.

Jones, P.D., Hulme, M. and Briffa, K.R. 1993. A comparison of Lamb circulation types with an objective classification scheme. *International Journal of Climatology*, **13**, 655-663.

Kalnay, E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Leetmaa A, Reynolds R, Jenne R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, **77**, 437-471.

Kalnay, E. and Cai, M. 2003. Impact of urbanization and land-use change on climate. *Nature*, **423**, 528-531.

Katz, R.W. 1996. Use of conditional stochastic models to generate climate change scenarios. *Climatic Change*, **32**, 237-255.

Katz, R.W. and Parlange, M.B. 1993. Effects of an index of atmospheric circulation on stochastic properties of precipitation. *Water Resources Research*, **29**, 2335-2344.

Katz, R.W. and Parlange, M.B. 1996. Mixtures of stochastic processes: applications to stochastic downscaling. *Climate Research*, **7**, 185-193.

Katz, R.W. and Parlange, M.B. 1998. Overdispersion phenomenon in stochastic modeling of precipitation. *Journal of Climate*, **11**, 591-601.

Katz, R.W., Parlange, M.B. and Naveau, P. 2002. Statistics of extremes in hydrology. *Advances in Water Resources*, **25**, 1287-1304.

Kidson, J.W. 2000. An analysis of New Zealand synoptic types and their use in defining weather regimes. *International Journal of Climatology*, **20**, 299-316.

Kidson, J.W. and Thompson, C.S. 1998. A comparison of statistical and model-based downscaling techniques for estimating local climate variations. *Journal of Climate*, **11**, 735-753.

Kilsby, C.G., Cowpertwait, P.S.P., O'Connell, P.E. and Jones, P.D. 1998. Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *International Journal of Climatology*, **18**, 523-539.

Kim, J.W., Chang, J.T., Baker, N.L., Wilks, D.S. and Gates, W.L., 1984. The statistical problem of climate inversion: Determination of the relationship between local and large-scale climate. *Monthly Weather Review*, **112**, 2069–2077.

Kittel, T.G.F., Rosenbloom, N.A., Painter, T.H., Schimel, D.S. and VEMAP Modeling Participants, 1995. The VEMAP integrated database for modeling United States ecosystem/vegetation sensitivity to climate change. *Journal of Biogeography*, **22**, 857-862.

Lamb, P. 1987. On the development of regional climatic scenarios for policy oriented climatic impact assessment. *Bulletin of the American Meteorological Soc*iety, **68**, 1116-1123.

Lambert, S.J. and Boer, G.J. 2001. CMIP1 evaluation and intercomparison of coupled climate models. *Climate Dynamics*, **17**, 83-106.

Leung, L.R., Mearns, L.O., Giorgi, F. and Wilby, R.L. 2003. Regional climate research: needs and opportunities. *Bulletin of the American Meteorological Society*, **84**, 89-95.

London Climate Change Partnership (LCCP), 2002. A climate change impacts in London evaluation study. Final Technical Report, Entec UK Ltd.

Lorenz, E.N. 1969. Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences*, **26**, 636-646.

Martin, E, Timbal, B and Brun, E. 1997. Downscaling of general circulation model outputs: simulation of the snow climatology of the French Alps and sensitivity to climate change. *Climate Dynamics*, **13**, 45-56.

Mearns, L.O., Rosenzweig, C. and Goldberg, R. 1996. The effect of changes in daily and interannual climatic variability on CERES-Wheat: a sensitivity study. *Climatic Change*, **32**, 257-292.

Mearns, L.O., Bogardi, I., Giorgi, F., Matyasovszky, I. and Palecki, M. 1999. Comparison of climate change scenarios generated from regional climate model experiments and statistical downscaling. *Journal of Geophysical Research*, **104**, 6603–6621.

Mearns, L.O., Giorgi, F., Whetton, P., Pabon, D., Hulme, M. and Lal, M. 2003. Guidelines for use of climate scenarios developed from Regional Climate Model experiments. Data Distribution Centre of the Intergovernmental Panel on Climate Change.

Mearns, L.O., Hulme, M., Carter, T.R., Lal, M., Leemans, R. and Whetton, P.H. 2001. Climate scenario development. In: Climate Change 2001: The Scientific Basis, Chapter 13. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, 881pp.

Murphy, J.M. 1999. An evaluation of statistical and dynamical techniques for downscaling local climate. *Journal of Climate*, **12**, 2256-2284.

New, M. and Hulme, M. 2000. Representing uncertainties in climate change scenarios: a Monte Carlo approach. *Integrated Assessment*, **1**, 203-213.

Olsson, J., Uvo, C.B. and Jinno, K. 2001. Statistical atmospheric downscaling of short-term extreme rainfall by neural networks. *Physics and Chemistry of the Earth Part B-Hydrology Oceans and Atmosphere*, **26**, 695-700.

Osborn, T.J., Briffa, K.R., Tett, S.F.B., Jones, P.D. and Trigo, R.M. 1999 Evaluation of the North Atlantic Oscillation as simulated by a coupled climate model. *Climate Dynamics*, **15**, 685-702.

Palutikof, J.P., Agnew, M.D. and Hoar, M.R. 2004. Public perceptions of unusually warm weather in the UK: impacts, responses and adaptations. *Climate Research*, **26**, 43-59.

Palutikof, J.P., Goodess, C.M., Watkins, S.J. and Holt, T. 2002. Generating rainfall and temperature scenarios at multiple sites: examples from the Mediterranean. *Journal of Climate*, **15**, 3529-3548.

Parlange, M.B. and Katz, R.W. 2000. An extended version of the Richardson model for simulating daily weather variables. *Journal of Applied Meteorology*, **39**, 610-622.

Pfizenmayer, A and von Storch H. 2001. Anthropogenic climate change shown by local wave conditions in the North Sea. *Climate Research*, **19**, 15-23.

Pilling, C.G. and Jones, J.A.A. 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.

Pilling, C. and Jones, J.A.A. 1999. High resolution climate change scenarios: implications for British runoff. *Hydrological Processes*, **13**, 2877-2895.Rabiner, L.R. and Juang, B.H. 1986. An introduction to hidden Markov models. *IEEE ASSP Magazine* 4-16.

Prudhomme, C., Reynard, N. and Crooks, S. 2002. Downscaling of global climate models for flood frequency analysis: where are we now? *Hydrological Processes*, **16**, 1137-1150.

Reid, P.A., Jones, P.D., Brown, O., Goodess, C.M. and Davies, T.D., 2001. Assessments of the reliability of NCEP circulation data and relationships with surface climate by direct comparisons with station based data. *Climate Research*, **17**, 247-261.

Reynard, N.S., Crooks, S.M. and Kay, A.L. 2004. *Impact of climate change on flood flows in river catchments*. Defra/Environment Agency Flood and Coastal Defence R&D Programme, Report W5B-01-050, Swindon, UK.

Reynard, N.S., Prudhomme, C. and Crooks, S.M. 2001. The flood characteristics of large UK rivers: potential effects of changing climate and land use. *Climatic Change*, **48**, 343-359.

Robinson, P.J. and Finkelstein, P.L. 1989. *Strategies for Development of Climate Scenarios.* Final Report to the U.S. Environmental Protection Agency. Atmosphere Research and Exposure Assessment Laboratory, Office of Research and Development, USEPA, Research Triangle Park, NC, 73 pp.

Salathé, E.P. 2003. Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. *International Journal of Climatology*, **23**, 887-901.

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. and 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, M.A. and Barrow, E.M. 1997. Use of a stochastic weather generator in the development of climate change scenarios. *Climatic Change*, **35**, 397-414.

Smith, J.B., and Tirpak, D.A. (eds.) 1989. *The Potential Effects of Global Climate Change on the United States.* Report to Congress, United States Environmental Protection Agency, EPA-230-05-89-050, Washington, DC, 409 pp.

Stidd, C.K. 1954. The use of correlation fields in relating precipitation to circulation. *Journal of Meteorology*, **11**, 202-213.

Stohlgren, T.J., Chase, T.N., Pielke, R.A., Kittel, T.G.F. and Baron, J.S. 1998. Evidence that land use practices influence regional climate, vegetation, and streamflow patterns in adjacent natural areas. *Global Change Biology*, **4**, 495-504.

Subak, S., Palutikof, J.P., Agnew, M.D., Watson, S.J., Bentham, C.G., Cannell, M.G.R., Hulme, M., McNally, S., Thornes, J.E., Waughray, D. and Woods, J.C. 2000. The impact of the anomalous weather of 1995 on the U.K. economy. *Climatic Change*, **44**, 1-26.

Svensson, C. and Jones, D.A. 2002. Dependence between extreme sea surge, river flow and precipitation in Eastern England. *International Journal of Climatology*, **22**, 1149-1168.

Timbal, B. and McAvaney, B.J. 2001. An analogue-based method to downscale surface air temperature: application for Australia. *Climate Dynamics*, **17**, 947-963.

Timbal, B., Dufour, A. and McAvaney, B. 2003. An estimate of future climate change for western France using a statistical downscaling technique. *Climate Dynamics*, **20**, 807 - 823.
Van den Dool, H.M. 1989. A new look at weather forecasting through analogues. *Monthly Weather Review*, **117**, 2230-2247.
Vellinga, M. and Wood, R.A. 2002. Global climate impacts of a collapse of the Atlantic thermohaline circulation. *Climatic Change*, **54**, 251-267
Von Storch, H. 1999. On the use of "inflation" in statistical downscaling. *Journal of Climate*, **12**, 3505-3506.
Von Storch, H., 1995. Inconsistencies at the interface of climate impact studies and global climate research. *Meteorol. Zeitschrift* **4**, 72-80.
Von Storch, H., Zorita E. and 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.
Watterson, I.G., O'Farrell, S.P. and Dix, M.R. 1997. Energy and water transport in climates simulated by a general circulation model that includes dynamic sea ice. *Journal of Geophysical Research*, **102**, 11027-11037.
Widmann, M. and Bretherton, C.S. 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. and 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, C.S. and Salathe Jr., E.P. 2003. Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. *Journal of Climate*, **16**, 799-816.
Wigley, T.M.L., Jones, P.D., Briffa, K.R. and Smith, G. 1990. Obtaining sub–grid scale information from coarse resolution general circulation model output. *Journal of Geophysical Research*, **95**, 1943–1953.
Wilby, R.L. 1994. Stochastic weather type simulation for regional climate change impact assessment. *Water Resources Research*, **30**, 3395-3403.
Wilby, R.L. 1997. Nonstationarity in daily precipitation series: implications for GCM downscaling using atmospheric circulation indices. *International Journal of Climatology*, **17**, 439–454.
Wilby, R.L. and Wigley, T.M.L. 1997. Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography*, **21**, 530-548.
Wilby, R.L. and Wigley, T.M.L. 2000. Precipitation predictors for downscaling: observed and General Circulation Model relationships. *International Journal of Climatology*, **20**, 641–661.
Wilby, R.L., Conway, D. and Jones, P.D. 2002. Prospects for downscaling seasonal precipitation variability using conditioned weather generator parameters. *Hydrological Processes*, **16**, 1215-1234.
Wilby, R.L., Hay, L.E., Gutowski, W.J., Arritt, R.W., Tackle, E.S., Leavesley, G.H. and Clark, M. 2000. Hydrological responses to dynamically and statistically downscaled General Circulation Model output. *Geophysical Research Letters*, **27**, 1199–1202.
Wilby, R.L., Tomlinson, O.J. and Dawson, C.W. 2003. Multi-site simulation of precipitation by conditional resampling. *Climate Research*, **23**, 183-194.
Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J. and Wilks, D.S. 1998. Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research*, **34**, 2995-3008.
Wilks, D.S. 1989. Conditioning stochastic daily precipitation models on total monthly precipitation. *Water Resource Research*, **25**, 1429–1439.
Wilks, D.S. 1992. Adapting stochastic weather generation algorithms for climate change studies. *Climatic Change*, **22**, 67-84.
Wilks, D.S. 1999. Multisite downscaling of daily precipitation with a stochastic weather generator. *Climate Research*, **11**, 125-136.
Wilks, D.S. and Wilby, R.L. 1999. The weather generation game: A review of stochastic weather models. *Progress in Physical Geography*, **23**, 329-357.
Winkler, J.A., Palutikof, J.P., Andresen, J.A. and Goodess, C.M. 1997. The simulation of daily temperature time series from GCM output: Part II: Sensitivity analysis of an empirical transfer function methodology. *Journal of Climate*, **10**, 2514-2535.
Wood, A.W., Leung, L.R., Sridhar, V. and Lettenmaier, D.P. 2004. Hydrological implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **62**, 189-216.
Zheng, N., Hales, K. and Neelin, J.D. 2002. Nonlinear dynamics in a coupled vegetation-atmosphere system and implications for desert-forest gradient. *Journal of Climate*, **15**, 3474-3487.

Zorita, E. and von Storch, H. 1999. The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate*, **12**, 2474-2489.

Zorita, E., Hughes, J.P., Lettenmaier, D.P. and von Storch, H. 1995. Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *Journal of Climate*, **8**, 1023-1042.