The statistical downscaling model-decision centric (SDSM-DC): conceptual basis and applications

Article

Accepted Version

The Statistical DownScaling Model – Decision Centric (SDSM-DC): Conceptual basis and applications

Wilby[1], R.L., Dawson[2], C.W., Murphy[3], C., O'Connor[3], P. and Hawkins[4], E.

[1] Department of Geography, Loughborough University, LE11 3TU, UK

[2] Department of Computer Science, Loughborough University, LE11 3TU, UK

[3] Department of Geography, National University of Ireland Maynooth, Maynooth, Ireland

[4] Department of Meteorology, University of Reading, RG6 6BB, UK

Main body word count: 7057

14 July 2014

Resubmitted to: *Climate Research*

Corresponding author: Robert Wilby (email: r.l.wilby@lboro.ac.uk)

Abstract

Regional climate downscaling has arrived at an important juncture. Some in the research community favour continued refinement and evaluation of downscaling techniques within a broader framework of uncertainty characterisation and reduction. Others are calling for smarter use of downscaling tools, accepting that conventional, scenario-led strategies for adaptation planning have limited utility in practice. This paper sets out the rationale and new functionality of the Decision Centric (DC) version of the Statistical DownScaling Model (SDSM-DC). This tool enables synthesis of plausible daily weather series, exotic variables (such as tidal surge), and climate change scenarios guided, not determined, by climate model output. Two worked examples are presented. The first shows how SDSM-DC can be used to reconstruct and in-fill missing records based on calibrated predictor-predictand relationships. Daily temperature and precipitation series from sites in Africa, Asia and North America are deliberately degraded to show that SDSM-DC can reconstitute lost data. The second demonstrates the application of the new scenario generator for stress testing a specific adaptation decision. SDSM-DC is used to generate daily precipitation scenarios to simulate winter flooding in the Boyne catchment, Ireland. This sensitivity analysis reveals the conditions under which existing precautionary allowances for climate change might be insufficient. We conclude by discussing the wider implications of the proposed approach and research opportunities presented by the new tool.

Key words

Downscaling; Climate scenario; Weather generator; Stress test; Data reconstruction; Adaptation

1. Introduction

Attitudes are changing about the production and utility of regional climate change scenarios. The notion that climate model output can be used in a deterministic sense to direct adaptation decisions is increasingly hard to defend in the face of recognised uncertainties in global and regional climate modelling – both statistical and dynamical (Pielke Sr & Wilby 2012, Stakhiv 2011). There are a few cases where downscaled products have been applied, such as establishment of precautionary allowances for flood risk in Australia, Denmark, Germany and the UK (Wilby & Keenan 2012). However, some believe that climate models are still not yet "ready for prime time" (Kundzewicz and Stakhiv, 2010). Others advocate an *assess-risk-of policy* over *predict-then-act* framework (Lempert et al. 2004, Weaver et al. 2013).

Conventional uses of downscaling include production of scenarios, data inputs for impacts modelling, evaluation of the consequences relative to present climate, and discussion of appropriate adaptation responses. Typically, large uncertainties attached to climate model scenarios cascade into even larger uncertainties in downscaled regional climate change scenarios and impacts (**Figure 1**). The decision-maker is then left with a bewildering range of possibilities, and often defaults to "low regret" decisions (World Bank 2012). A few studies use regional downscaling to explore the relative significance of uncertainty components, for example in future snowmelt (Dobler et al. 2012), high (Smith et al. 2014), low (Wilby & Harris 2006), or mean river flows (Bastola et al. 2011).

The Statistical DownScaling Model (SDSM) was originally conceived as a regional climate change scenario generator to support climate risk assessment and adaptation planning. A meta-analysis of the first decade of published work using SDSM showed that over half the 200+ studies to date refer to water and flood impacts, often with regards to the production of climate scenarios, benchmarking with other scenario tools, or refinement of downscaling techniques (Wilby & Dawson 2013). A modest but growing number of studies apply the tool in adaptation planning or climate risk management[1].

Some assert that downscaling should be used to appraise adaptation options through vulnerability-led rather than scenario-led methodologies (Wilby & Dessai, 2010). In this 'bottom-up' framework, the scenario is used to evaluate the performance (some say "stress test") adaptation measures. As such, the scenario does not need to be explicitly tied to a given

[1] For a bibliography of SDSM studies see: http://co-public.lboro.ac.uk/cocwd/SDSM/Bibliography.pdf

climate model or ensemble; plausible futures can be described by representative climates or generated from weather sequences using simple narratives of the future (such as "warmer", "drier", "more variable") (Whetton et al. 2012). Scenarios are then used to test the sensitivity of the system or decision set, ideally to reveal non-linear behaviours or break-points under prescribed climate-forcing (e.g., Prudhomme et al. 2010, Stakhiv 2011, Brown & Wilby, 2012, Lempert et al. 2012, Nazemi et al. 2013, Steinschneider & Brown, 2013; Turner et al., 2014).

Accordingly, this paper describes a suite of tools for producing daily weather series and climate scenarios **without explicit use of climate model output**. Our Decision-Centric (DC) version of SDSM is built on the premise that downscaled scenarios should be informed by but not determined by climate models. This increases the range of plausible scenarios that can be evaluated in an adaptation context. The new *Weather Generator* in SDSM-DC also provides tools for in-filling missing data and interrogating local climate information based on re-analysis predictor variables. These functions enable application in data sparse regions and leads to deeper understanding of regional climate systems.

The purpose of this paper is to introduce the new functions of SDSM-DC and to demonstrate their usage with two case studies. The following section describes the technical basis of SDSM-DC as applied to single and multiple sites. We then illustrate how SDSM-DC can be used for data reconstruction in contrasting climate regimes. These analyses address the often asked question about how much data is needed to calibrate the model to achieve a given level of skill. The second worked example shows how SDSM-DC can be used in a 'stress testing' situation. In this case, we refer to the definition of safety margins for flood risk under a changed climate in Ireland. Finally, we identify some of the research opportunities emerging from a 'bottom-up', vulnerability-based paradigm for downscaling.

2. SDSM-DC

Earlier versions of SDSM have been described elsewhere (Wilby et al. 2002, 2003, Wilby & Dawson 2013) but for completeness are brought together here. The tool enables the production of climate change time series at sites for which there are *daily* observations (the predictand) and re-analysis products describing large-scale atmospheric properties (the predictors) for model calibration. In the vintage version of SDSM, archived General

Circulation Model (GCM) output may then be used to generate scenarios for future decades. The SDSM-DC User is guided through each stage of the downscaling process by a set of screens (**Figure 2**). These address key functions such as basic quality control and transformations (as required) of input data; predictor variable selection; model set-up and calibration; weather and scenario generation; diagnostics for interrogating model output (summary statistics, frequency and time-series analysis, graphing). The following section reprises the key features of the single- and multi-site versions of SDSM then introduces the new functions of SDSM-DC.

2.1 Downscaling single sites

SDSM is best described as a conditional weather generator because atmospheric circulation indices and regional moisture variables are used to estimate time-varying parameters describing daily weather at individual sites (e.g., precipitation occurrence or daily mean temperatures). The downscaled process is either unconditional (as with wet-day occurrence or air temperature), or is conditional on an event (as with rainfall amounts).

For wet-day occurrence W_i there is a direct linear dependency on n predictor variables X_{ij} on day i:

$$W_i = \alpha_0 + \sum_{j=1}^{n} \alpha_j X_{ij}$$

under the constraint $0 \leq W_i \leq 1$. Precipitation occurs when the uniform random number [0,1] $r \leq W_i$. The threshold (mm) for a wet-day varies between locations, depending on the definition of trace rainfalls or precision of measurement. Here we define a wet-day as any day with non-zero precipitation total.

When a wet-day is returned, the precipitation total P_i is downscaled using:

$$P_i^k = \beta_0 + \sum_{j=1}^{n} \beta_j X_{ij} + e_i$$

where k is used to transform daily wet-day amounts to better match the normal distribution. Here we apply the fourth root transformation (i.e., $k = 0.25$) to P_i. Note that the same predictor set is used to downscale W_i and P_i and that all predictors v_{ij} are standardised with respect to the 1961-1990 mean $\bar{V}_j$ and standard deviation σ_j:

$$X_{ij} = \frac{v_{ij} - \bar{V}_j}{\sigma_j}$$

For unconditional processes, such as temperature, there is a direct linear relationship between the predictand U_i and the chosen predictors X_{ij}:

$$U_i = \gamma_0 + \sum_{j=1}^{n} \gamma_j X_{ij} + e_i$$

The model error e_i is assumed to follow a Gaussian distribution and is stochastically generated from normally distributed random numbers and added on a daily basis to the deterministic component. This white noise enables closer fit of the variance of the observed and downscaled distributions, but is known to degrade skill at replicating serial autocorrelation implicit to daily predictor variables. The stochastic process also enables the generation of ensembles of time-series to reflect model uncertainty.

All downscaling parameters (α_j, β_j, and γ_j) are obtained via least squares calibration of the local predictand(s) against regional predictor variables derived from the National Center for Environmental Prediction (NCEP) re-analysis (Kalnay et al. 1996) using data for any period within 1961-2000. Users are advised to calibrate SDSM using data drawn from this period because it is assumed that these decades have relatively high data quality/availability with modest risk of nonstationarity in predictor-predictand relationships due to anthropogenic forcings. Predictands are downscaled separately so any covariance must be conveyed by common predictor variables and/or correlation between predictors. Model testing suggests that this is a reasonable assumption (Wilby et al. 1998).

In common with all downscaling methods, SDSM predictor-predictand relationships are assumed to be unaffected by anthropogenic influences during the calibration period, and are applicable to conditions outside the training set. In practice, the parameters of all empirical and dynamical downscaling models are observed to vary over decadal-time scales, not least because of natural variability. Furthermore, the climate effects of land-surface changes cannot be captured by conventional statistical downscaling models (Pielke Sr. & Wilby 2011). For instance, previous work in the western US suggests that winter snow/ice cover feedbacks can lead to lower temperatures than expected by downscaling models (Wilby & Dettinger 2000). All these caveats undermine the case for applying downscaling in predict-then-act modes.

2.2 SDSM-DC functionality

Perhaps the most contentious aspect of SDSM-DC is that climate scenarios are not determined explicitly by climate model output. Rather, the range of the adjustments may be informed by palaeoclimatic evidence, expert judgement, or climate model experiments. Alternatively, the range may be designed to bracket conditions that would stress the target system(s) to failure (Steinschneider & Brown 2013). These methods represent a marked departure from main-stream downscaling ideology which is wholly contingent upon the realism of future driving variables supplied by climate models. Nonetheless, there is acceptance that even massive climate model ensembles may understate the true uncertainty in regional climate change (Stainforth et al. 2007, Deser et al. 2012). Therefore, tools are needed to generate scenarios that can test adaptation decisions and system vulnerabilities over a much wider (yet still plausible) range of climate variability and change (Steinschneider & Brown 2013, Brown & Wilby, 2012, Nazemi et al. 2013).

SDSM-DC enables the User to apply such *Treatments* to daily predictands. These are User-defined factors and functions that manipulate the unconditional occurrence process, mean, variance and trend of the original series. Input series may originate from observations[2] or from output produced by a weather generator (as in **Figure 3a**) if multiple realisations are required. Four main types of single and multiple treatments are described below.

2.2.1 Occurrence

In the following explanation we refer to precipitation as an example manipulation of event occurrence. However, this treatment might apply to any other phenomena with zero and non-zero values (such as sunshine hours). For precipitation the event threshold might be any non-zero total. In this case, the percentage change entered represents the amount by which event frequency should change. For example, a value of 10% applied to rainfall series would

[2] For sample input data, predictor variables and parameter file see: http://co-public.lboro.ac.uk/cocwd/SDSM/sdsmmain.html

increase the number of rain days by 10%; a value of -20% would reduce the number of wet-days by a fifth (**Figure 3b**).

When increasing event frequencies, new wet-days are not generated randomly across the entire range of the series but are weighted according to the baseline occurrence profile. This ensures that (for precipitation occurrence) wet months remain generally wetter than dry months and vice versa. This process involves four stages. First, input series are analysed to determine the frequency of events in each month (e.g., January 16%; February 20%, etc.). Second, a random month is selected based on the overall likelihood of occurrence (in this case, February would have a slightly higher chance of being selected than January). Third, a random non-event (dry) day in this month is selected from the concatenated series. Fourth, in order to convert this dry day into a wet day an appropriate event magnitude (wet-day amount) must be determined. This is achieved by sampling a non-zero event from the month. Steps two to four are then repeated until the required percentage change in rain days has been achieved.

Removal of events from the series operates in a similar way to the process outlined above. As before, the series is first analysed to determine the monthly occurrence profile. This likelihood is used to weight the chance of removing an event: those months with the greatest frequency of zero days are most likely to lose a non-zero event. A non-zero day is randomly selected and then removed from that month (anywhere within the entire series) by replacing it with the event threshold value. This process is repeated until the required percentage of events has been achieved.

The above processes are conditionally stochastic since addition or removal of events is weighted by monthly event frequencies, but individual days are randomly changed within months. This effectively amplifies the initial seasonality of event occurrence. Alternatively, the User can prescribe the change in occurrence for each month by setting the target likelihood profile. In this case, SDSM-DC then calculates whether to randomly add or remove events from each month in turn (across the entire series). In cases where a month has no events, magnitudes are sampled from adjacent months.

Stochastically adding or removing events from a series can affect the mean of the series. If the user wishes to preserve the initial mean despite adjusting the occurrence process, SDSM-DC scales the final series such that the overall total is the same as pre-treatment. SDSM-DC stores the event total for the series before the occurrence process is manipulated. The model then calculates how much the final series needs to be adjusted in order to preserve this original total. For example, under this set-up, reducing the frequency of events by 10% would necessitate scaling the remaining non-zero events by 10% to preserve the pre-treatment mean.

2.2.2 Mean

The mean treatment enables adjustments to individual daily values by the chosen amount. For a conditional process this treatment is only applied to values above the event threshold (for example, non-zero rainfall amounts). The treatment may be applied either as a factor (such as for precipitation) or by addition (such as for temperature). Note that this also affects other properties of the series including the maximum, quantile distribution, and variance.

2.2.3 Variance

In order to change the variance and preserve the coefficient of variation (mean divided by standard deviation) only the mean need be scaled (see above). Otherwise, for an unconditional process, the mean is first removed from each value then each data point is multiplied by the square root of the required percentage change in variance. The mean is then added back to the result thereby increasing the variance by the desired amount overall and leaving the mean unchanged. This treatment is summarised as:

$$U_m = [(U_i - \bar{U}) * (\sqrt{1+r})] + \bar{U}$$

where U_m is the transformed value, U_i is the original value, $\bar{U}$ is the mean of the series, and r is the change entered by the user ($0 \leq r \leq 1$). This simple procedure cannot be applied to highly skewed distributions (such as wet-day amounts) because the treatment would yield negative values. In this case, the variance treatment is applied after a Box-Cox transformation (Hinkley 1977, Sakia, 1992):

$$U_m = (U_i^{\lambda} - 1)/\lambda \qquad \text{where } \lambda \neq 0;$$

$$U_m = \ln(U_i) \qquad \text{where } \lambda = 0;$$

where λ lies in the range [-5, +5] and is set to minimise the skewness of the distribution of U_m. SDSM-DC determines λ via iteration until skewness is minimised. In order to evaluate the effectiveness of the transformation for each λ Hinkley's (1977) nonparametric measure of symmetry is applied, d_{IQR}. This does not depend on knowledge of the underlying distribution and may be computed using either the standard deviation or inter-quartile range as the denominator:

$$d_{IQR} = \frac{(mean - median)}{inter\_quatile\ range}$$

The inter-quartile range is used in preference to the standard deviation in SDSM-DC because the latter tends to drive values of *d* towards zero for larger values of λ. As the algorithm employed by SDSM-DC is iterative, the standard deviation may well result in large (positive or negative) values of λ being selected which by no means minimise the skewness of the data. Conversely, d_{IQR} provides similar λ value as d_{SD} but does not suffer from convergence as values increase and decrease.

Having transformed the series it is now possible to apply the factor to achieve the required variance inflation as with normally distributed data. This is not straightforward as there is no direct relationship between the required variance transformation and the Box-Cox transformed data. Therefore, SDSM-DC applies an iterative approach to determine an appropriate value of *r*. For increased variance *r* ranges from 0 to a maximum of value of 0.3; for decreases *r* ranges from 0 to a minimum value of -0.5. Through iteration, SDSM-DC derives an appropriate value of *r* to achieve the intended variance treatment, such as +50% (**Figure 3c**).

2.2.4 Trend

SDSM-DC allows three types of trend to be applied to a series: linear, exponential or logistic. A linear trend simply adds (or subtracts) the value entered at each annual increment, scaled within years by Julian day number. For example, 10 would add values from 0 to 10 in the first year, 10 to 20 in the second year, 20 to 30 the following year, etc. For a calendar year each day has added 10/365.25 multiplied by the Julian day number.

For a conditional process, event values are adjusted multiplicatively. For example, if the factor is 5, events in the first year are increased by 0 to 5% linearly (for days 1 to 365); then by 5% to 10% in the second year; and so forth. In this case, the first day would be approximately unchanged; a value in the middle of the year would be increased by ~2.5%; and a value at the end of the year by 5%.

Exponential and logistic trends are applied across the entire range of the series, rather than annually as in the linear treatment. An exponential trend adds (or subtracts) an exponential function across the entire range of the data. For example, entering +5 would add between 0 (for the first data point) to +5 (for the final data point) with intervening values scaled exponentially between these end-points (**Figure 3d**). For a conditional process the treatment is multiplicative rather than additive. For example, +10 would result in exponential scaling by 1 to 1.10 between the first and last non-zero value in the series.

The logistic trend applies an S-shaped function by addition of the chosen value between the first and last points of the unconditional series. For a conditional process the change is multiplicative rather than additive. For example, 5 results in events being scale by 1 to 1.05 across the full length of the series following the logistic curve. The logistic function is useful for introducing step changes into generated series.

2.2.5 Multiple treatments

Treatments can be implemented in isolation or combination to create more complex transformations of the series. If the latter, treatments are applied by SDSM-DC in fixed order (*Occurrence*, *Mean*, *Variance* and *Trend*). For instance, it is possible to adjust the occurrence, by say -20%, whilst preserving the mean annual precipitation total (**Figure 3e**). In this case, the generated series would have fewer wet-days but with greater mean intensity. More elaborate scenarios can be produced by simultaneously changing the occurrence, variance and trend (**Figure 3f**). These complex treatments might be applied to mimic a specific scenario, or to explore known system vulnerabilities. However, the task of interpreting associated impacts becomes much more demanding. Hence, most cases where synthetic series have been used for stress testing are uni- or two-dimensional (e.g., Prudhomme et al. 2010; Nazemi et al., 2013, Steinschneider & Brown, 2013).

2.3 Extension to multiple sites

Although the public domain version of SDSM-DC is for single sites, the basic model can be modified for multi-site applications (following Wilby et al., 2003). This involves two steps. First, a 'marker' series based on daily area averages from several sites (or a single key site) is generated using predictors X_{ij}. Second, the area-average is disaggregated to observed daily series recorded at the constituent sites. This is achieved by resampling multi-site values on the date with observed area-average closest to the downscaled area-average. For example, **Figure 4** shows that SDSM-DC reproduces the observed range of inter-site correlations for both rainfall and temperature in the Upper Colorado River Basin. Across 76 stations in this catchment, the spatial autocorrelation in daily temperature (mean r_{obs} = 0.98; r_{SDSM} = 0.98) is found to be more homogeneous than that of precipitation (mean r_{obs} = 0.72; r_{SDSM} = 0.69).

Since actual patterns of values are re-sampled by SDSM-DC, both the area average of the marker series and the spatial covariance of the multi-site array are preserved (Wilby et al. 2003, Harpham & Wilby 2005). Area averages are favoured over single site marker series because there is less risk of employing a non-homogeneous or non-representative record, and predictability is generally increased (because of larger signal-to-noise ratio). As with other resampling methods, the maximum daily value generated cannot exceed the maximum daily amount in the observations without invoking the treatments described above.

3. Worked example 1: Data reconstruction

Many of the regions that are most vulnerable to climate variability and change are also the most data sparse. For example, major data gaps exist in the Congo basin, Sahel, central Asia, and Amazon basin. One solution is to support intensive field campaigns (such as the EU African Monsoon Multidisciplinary Analysis [AMMA]) to collect data on poorly understood processes or climate regimes, especially in the Tropics. An alternative strategy is to locate, rescue, digitize, archive and share historic climate data that may be held only as paper or physical copies (as is the mission of the International Environmental Data Rescue Organization [IEDRO]). A third way is to synthesize or infill missing data using a stochastic weather generator. In the following application SDSM-DC is used to reconstruct daily temperature and precipitation series and to demonstrate the trade-off between model skill and information content of available data.

3.1 Strategies for weather simulation

There are broadly three main approaches to stochastic weather generator calibration. The most conventional way involves tuning model parameters against available series for precipitation occurrence, then dependent variables such as rainfall amount, temperature, sunshine duration and so forth (Wilks & Wilby 1999). The resulting model replicates important properties of the data (such as wet-day frequencies and amounts, wet- and dry-spell durations, and covariance amongst variables) or can be used to synthesize much longer series for analysis of extreme events. More sophisticated mixture-model variants can be tuned to simulate low-frequency behaviour of annual to multi-decadal time-scales. Such tools have found important applications in hydrologic design and crop-modelling, but are not suited for data reconstruction because of their stochastic outputs.

Others apply weather generators based on parameters (e.g., rainfall occurrence or the alpha and beta parameters of the gamma distribution) that have been prepared from gridded data (e.g., Semenov et al., 2010, 2013) or interpolated from sites where such data exist to locations where they do not (e.g., Camberlin et al. 2014, Semenov & Brooks 1999). In some cases, landscape properties such as local slope aspect, distance from coast and altitude are extracted from digital elevation models (e.g., the 1 km resolution Shuttle Radar Topography Mission of the US Geological Survey) to explicitly account for topographic controls via weighted local regressions (e.g., Wilby & Yu 2013). Such techniques are particularly helpful for estimating weather generator parameters in regions of complex topography but are not so well suited to repairing or infilling partial series.

This is where SDSM-DC potentially offers hope: observed (NCEP) predictor-predictand relationships constructed for each calendar month, season, or series as a whole can be used to estimate values on days for which there are no data, or for independently testing suspect values. If it can be assumed that other (non-climatic) forcings are constant, the main practical questions become how much data are needed for reconstruction, and what are the expected uncertainty bounds for reconstructed series? Both aspects are explored below using experiments in which daily series have been deliberately degraded in order to emulate SDSM-DC capabilities under realistic ‘field conditions’.

3.2 Minimum data requirements

The effect of reducing daily data availability is demonstrated using contrasting sites: Charlottetown on Prince Edward Island, Canada and Tunis in Tunisia (for temperature); Addis Ababa, Ethiopia and Chang wu, China (for precipitation). In each case, the length of observations presented for model calibration was varied between 10% and 100% of the available record (equating to about 4 to 40 years of data). Individual days or blocks of years were randomly removed to represent situations in which data records might be patchy or where longer sequences of data are missing. SDSM-DC skill at reproducing the artificially removed days was assessed using the Root Mean Squared Error (RMSE) for temperature; the proportion correct wet-day occurrence (PCW); and the non-parametric Kolmogorov-Smirnov (KS) *D*-statistic to test similarity of wet-day amount distributions.

Distributing "lost" data via missing year blocks yielded marginally larger RMSEs in temperature reconstructions than random data gaps, but only for records less than 10 years (**Figure 5**). This is because the random data reduction might still sample information content for extreme periods or on trends within the series that are otherwise missed when whole year blocks are removed. Both sets of results suggest that beyond 20 years of calibration data there is little reduction in RMSEs for temperature. A similar pattern emerges for precipitation occurrence with the most dramatic reduction in PCW for calibration sets less than 10 years (**Figure 6**). However, unlike temperature, there appears to be little difference between data degraded by random or block omission. In both cases, the presence or absence of a wet-day (non-zero precipitation) is simulated correctly on average ~75% of the time.

Ability to reproduce wet-day amount distributions was assessed by comparison of cumulative distributions (**Figure 7**) and the *D*-statistic (**Figure 8**). These reveal that the assumed fourth root distribution provides a fair approximation of observed wet-day amounts at both sites, particularly for occurrence of days >30 mm. The distribution of downscaled wet-day amounts appears to be robust to data reduction until very low levels (10%) of information are available for model calibration whether random days or years are removed. The type of data reduction is less important for Addis Ababa (**Figures 7a** and **7b**) than for Chang wu (**Figures 7c** and **7d**) because even the initial data set for the former site is partially fragmented.

D-statistics show little change in ensemble median but variance in the metric grows with increasing levels of data reduction, most notably at Addis Ababa (**Figure 8**). For this site, model skill at reproducing wet-day amounts is resistant to 10% random data loss. At Chang wu, where initial data quality is superior, the *D*-statistic is largely unchanged even after 80% reduction (by random day removal). The instability of the *D*-statistic for large data reduction at Addis Ababa is due to the diminished number of wet days available for downscaling parameter estimation within individual months. For example, with 90% data reduction there are fewer than 10 wet-days for model calibration in December. Large *D* can then arise when the stochasticity of the downscaling algorithm generates unexpectedly large wet-day amounts (as in **Figure 7d**). Likewise, small *D* may occur in a large ensemble when the small number of generated wet-days closely matches observations by chance.

With diminished samples of observed wet-day amounts there is larger uncertainty in parameter estimates and proportionately greater influence of any extreme event(s) captured in the sub-set. **Figure 8a** suggests that ~30 events are needed to obtain stable wet-day parameters for a given month. Moreover, choice of distribution (whether exponential, long-normal, fourth root, gamma, etc.) may be as important as the amount of data available for model calibration. The ramifications for minimum record lengths are most significant for semi-arid and hyper-arid regions where there may be very few wet-days even when there are many years of record, or when data are stratified by season rather than by calendar month. Conversely, as **Figure 6** shows, wet-day occurrence estimates are relatively robust to variations in record length and data gaps.

3.3 Reconstructed time-series

SDSM-DC was used to reconstruct daily temperature and precipitation series at the same sites as above. Models were fitted to all available data but assessed against metrics that were not applied in calibration, including extreme temperatures and annual precipitation totals. An ensemble of 20 daily series was produced in each case using NCEP predictors for the period 1961-2000. **Figures 9a** and**9b** show that SDSM-DC provides a close approximation of observed annual mean (r=0.87) and maxima (r=0.91) temperatures at Prince Edward Island and Tunis respectively. In both cases, the observations lie within the ensemble range of the downscaled series for the majority of years. The correlation between observations and downscaled series was also high for the annual frequencies of cold (r=0.76) and hot (r=0.91) days (**Figures 9c** and **9d**). Again, the majority of the hindcast values lie within the ensemble range. Results for Tunis demonstrate that even when there are strong trends in observations the NCEP predictors and downscaling are able to replicate most of the inter-annual and inter-decadal variability despite model calibration against daily performance metrics.

SDSM-DC was less skilful at replicating inter-annual variability in wet-day frequencies and totals at Addis Ababa and Chang wu (**Figure 10**). Although the majority of observed annual totals lie within the ensemble range, the correlation with the ensemble median is weak at Addis Ababa (r=0.36) compared with Chang wu (r=0.63). Correlations for the annual wet-day frequencies are marginally stronger: Addis Ababa (r=0.41) and Chang wu (r=0.71). Differences in skill between the two sites may reflect the quality and length of data available for calibration: 27 and 40 years respectively. The long-term mean at Addis Ababa is reproduced to within 3%, but 36% of observed annuals totals fall outside the ensemble range. Conway et al (2004) note that there is some ambiguity about the location of the site and that the possibility of changes in instrumentation cannot be discounted. Hence, evaluation of the downscaled series remains problematic for this site.

4. Worked example 2: Stress testing

In this application SDSM-DC is used to *stress-test* adaptation decisions for local flood risk management (O'Connor, 2013). By focusing on a specific question rather than the traditional "predict-then-act" approach the application can be categorised as a "bottom-up" approach to adaptation (Brown & Wilby, 2012). First, the option is described. Second, an impact model is calibrated for the system in question. Third, the scenario generator tool in SDSM-DC is used to construct the inputs for the impact model, and then construct a response surface showing the sensitivity of the system under a wide range of conditions. Finally, results obtained from a given climate model ensemble (such as CMIP3 or CMIP5) may be mapped onto the sensitivity surface to indicate likelihoods based on current knowledge.

4.1 Identifying the adaptation question or concern

In adapting to assumed increases in flood risk in Ireland, the Office of Public Works (OPW), the agency responsible for flood risk management, advocate precautionary allowances in design of flood defences (OPW 2009). Under this guidance an allowance of 20 % on design peak flows is recommended under a mid-range future scenario, with a 30 % allowance under a high-end future scenario. Note that OPW chose not to tie these allowances explicitly to any emissions or climate model scenario.

The value chosen for the precautionary allowance has far-reaching consequences. If too low, there is a danger of maladaptation and failure to protect lives, livelihoods and critical infrastructure; if too high, the cost of flood defences may be prohibitive or outweigh the intended benefits. Authorities have to weigh up these costs and benefits in the context of uncertainty about climate change impacts. Using an example catchment in east Ireland, SDSM-DC was used to explore the sensitivity of a 1-in-100 year design flood, to changes in key precipitation parameters.

4.2 Developing an impact model for the chosen system

The Boyne at Slane Castle in east Ireland has a catchment area of 2460 km^2, average annual precipitation 897 mm (1952-2009), Base Flow Index (BFIsoils) 0.69, and an undulating landscape dominated by pasture. The conceptual rainfall-runoff model HYSIM (Manley 2006) was used to simulate streamflow within the catchment. The model has modest data requirements and has been applied previously in Ireland (e.g., Harrigan et al. 2014, Murphy et al. 2006, Bastola et al. 2012). Daily precipitation for three rainfall stations and potential evapotranspiration for the period 1952-2009 were obtained from Met Eireann, while daily streamflow for a gauge at Slane Castle was obtained from the OPW for the same period.

We recognise that HYSIM adds uncertainty due to non-uniqueness of model parameters (Murphy et al. 2006), but apply a single behavioural parameter set for illustrative purposes. Emphasis is placed on characterising uncertainties from GCMs and emission scenarios, given their large contribution to overall uncertainty in local impacts (e.g. Dobler et al. 2012, Wilby & Harris 2006). HYSIM was trained on daily flows for the period 1981-1995 and verified for the period 1996-2007. Nash-Sutcliffe (NS) (Nash and Sutcliffe 1970) scores of 0.87 and 0.88 were derived for the full training and verification periods respectively, while NS scores of 0.80 and 0.90 for winter (DJF) flows were obtained for training and verification periods respectively, indicating good model performance (**Figure 11**). To examine changes in flood events the Generalised Logistic (GL) distribution was fitted to annual winter maximum flood series simulated using original and perturbed precipitation series (Hosking and Wallis 1997).

4.3 Generating the impact model inputs

SDSM-DC was used to derive a response surface representing the sensitivity of changes in the design (1-in-100 year) flood to prescribed changes in precipitation. The scenario generator function in SDSM-DC was used to perturb observed catchment area-average rainfall to produce daily rainfall series without explicit use of climate model inputs. Changes in rainfall are expected to influence flooding through changes in seasonal wet-day occurrence and amounts. Wide ranges of change for these precipitation attributes were employed to construct bounds within which to perturb observed precipitation. Only winter (DJF) changes are reported here for illustrative purposes.

The sensitivity domain for precipitation parameters was informed by the projections of the Coupled Model Intercomparison Project CMIP3 for the nearest grid box, together with previous impacts assessments for Irish catchments (e.g. Bastola et al. 2012; Murphy & Charlton 2006). Changes in mean winter rainfall total ranging between -30 and +30 % and changes in the occurrence of winter wet days (amounts > 0.1 mm) between -20 and +20 % were sampled at 5% increments and applied to the observed rainfall series (1952-2009). Changes in the likelihood of wet-day occurrence and amounts were applied simultaneously so, for example, -20 % likelihood of rainfall with +10 % winter total yields an increase in mean wet-day amounts. Preserving winter totals while adjusting occurrence allows sensitivity to changes in intensity to be explored. Note that these treatments are specific to evaluation of flood risk; sensitivity analysis of other characteristics such as drought would imply alternative treatments to precipitation and potentially evapotranspiration.

4.4 Constructing the response surface and mapping climate projections

Perturbed rainfall series were input to HYSIM model to explore the sensitivity of the design flood to changes in rainfall properties with results visualised in the form of a response surface (**Figure 12**). PE was held constant at observed values given low losses during winter months. The 1-in-100 year flood was found to be sensitive to changes in both mean rainfall amounts and changes in the number of wet days. For the ranges of precipitation parameters considered, changes in the magnitude of the 1-in-100 year flood span -40 to +120 %.

Even very modest changes in mean rainfall amounts (when combined with reduced wet day occurrence) result in large changes in modelled flood magnitude, delivering rainfall in greater daily amounts and resulting in elevated flood peaks. Even reductions of winter mean rainfall by 10 %, when coupled with reductions in the number of wet days by 15 %, result in changes in flood magnitude approaching the medium range scenario design allowance of an additional 20 %. With no change in wet day occurrence increases in winter mean rainfall of above 5 % result in changes in flood magnitude approaching 20 %. The results highlight the sensitivity of flooding within this catchment – not just to changes in rainfall amounts, but to how changes in rainfall amounts are distributed through time. Such sensitivities are moderated by physical catchment properties defining the rainfall-runoff response and will vary on a catchment by catchment basis.

Climate change scenarios were then mapped onto the sensitivity response surface to examine risk of exceedence of the precautionary allowances (**Figure 13**). The exemplar climate change scenarios are regionalised outputs from 17 GCMs forced with three (A1B, A2 and B1) SRES emissions scenarios from the Coupled Model Intercomparison Project CMIP3 (Bastola et al. 2012). A change factor method based on monthly output from GCMs was used to infer changes in the parameters of a weather generator related to both the magnitude and occurrence of precipitation and was employed to derive regional scenarios for synoptic rainfall stations in Ireland (Bastola et al. 2011). Here 50 realisations of precipitation (based on sampled change factors from GCMs) under each emissions scenario were used to represent uncertainty in future scenarios. For each realisation percent changes in mean winter precipitation amounts and occurrence were derived relative to control simulations for the period 1961-1990. These are then plotted onto the sensitivity response surface, represented as a contour plot, for three future time periods (**Figure 13**).

Based on the above sensitivity analysis it is concluded that flood defences with a short design life (i.e. to the 2020s) with medium-range allowance of 20 % are likely to be adequate for the Boyne catchment, but some scenarios under the A1B and B1 emissions fall close to the limit of this allowance. However, given that most hard engineering defences have a design life in excess of 50 years, particularly when designed for extremes with a low recurrence interval (such as 1-in-100 year flood) this is unlikely to be the case for the 2050s and beyond. By the 2050s (2040-69) and especially by the 2080s (2070-99) a higher proportion of scenarios exceed the medium range allowance of 20 %, under all emissions scenarios. By the 2080s a number of projections under the A1B and A2 emissions scenario exceed even the high range allowance of 30 %.

In summary, this case study reveals potential limitations in the medium range allowance to rainfall driven changes in the design flood. By the 2080s there is greater residual risk, indicated by the proportion of scenarios exceeding the 20 % precautionary allowance. Such an 'assess risk of policy' approach allows decision makers to more readily appreciate the sensitivity of the system without explicit reliance on climate models, while the latter can be readily integrated to visualise risk as represented by a large ensemble of climate change scenarios. The approach adopted also facilitates rapid appraisal of such threshold based adaptation decisions and can be extended to national assessments (e.g., Prudhomme et al. 2010) or updated as new climate change projections become available.

5. Conclusions

This paper introduced the latest version of the Statistical DownScaling Model (SDSM) which was engineered with the specific needs of adaptation options appraisal in mind – hence the Decision Centric (-DC) extension. Consistent with other innovations in the downscaling community we are moving away from complete dependence on GCM output for producing regional climate change scenarios. Tools based entirely on weather generator techniques enable synthesis of input variables for impacts modelling and adaptation planning (e.g., Nazemi et al. 2013; Steinschneider & Brown 2013) but they are not always well-suited to reconstructing and/or infilling historic series. Most weather generators are also unable to synthesize exotic variables (e.g., air quality and urban heat island metrics, wave and tidal surge heights). SDSM-DC addresses these gaps by offering functionality to support data reconstruction and basic weather generation, as well as direct simulation of decision-relevant climate indices (**Table 1**). Moreover, tests reveal that SDSM performs as well as conventional weather generators such as LARS-WG (see: Hashmi et al., 2011; Hassan et al., 2014). Hence, with these capabilities, it is hoped that SDSM-DC will support decision-making in some of the most vulnerable and data sparse regions of the world.

Two worked examples were presented to demonstrate some of these capabilities. The first showed that with 10 years of data it is possible to achieve approximately the same level of skill at simulating rainfall occurrence, amounts and temperatures as with 40 years at the chosen sites. The analysis also confirmed that the downscaling is more robust to randomly degraded data throughout a longer record than to lost year blocks. Hence, recovery and digitization of even fragmentary observations may be beneficial and sufficient to allow infilling. Moreover, the stochastic features of SDSM-DC enable confidence limits to be attached to hindcast series so, even where the estimate may be uncertain, the model can at least provide an upper and lower bound.

The second example study showed how SDSM-DC can be used to stress test an adaptation decision – in this case a climate change safety allowance for flood defence schemes. The tool enables arbitrary treatments to be applied to the synthetic series needed for systems modelling. Treatments in the occurrence, mean, variance, and trend of events can be used to elucidate thresholds in the pressure-response. The range of scenarios that are explored may be guided by GCM output but importantly the tool enables exploration of consequences beyond even a multi-model ensemble. Likelihoods can still be attached by overlaying the cloud of model results on the response surface (as in Prudhomme et al. 2010). Moreover, by shifting emphasis from the GCM, the decision-maker is free to consider more holistic narratives that may be pertinent to the decision-making process (including perhaps changes in land cover, fire risk, forest die back and so forth in the case of water resources).

To conclude, the rationale behind SDSM-DC is as much about what the specific tool can do, as how downscaling in general can be used in smarter ways to support adaptation planning. Planned technical enhancements include the ability to manipulate low frequency variability in order to assess multi-season phenomena such as droughts or wet-spells persisting over more than one year. New diagnostics are needed to evaluate expected levels of skill at series reconstruction, perhaps based on more exhaustive cross-validation against whatever data are available. Further exploration of direct downscaling potential is needed, such as for river flows (as in Tisseuil et al., 2010) or other quantities that are typically derived by feeding downscaled climate variables into impact models. Hindcasting performance needs to be tested more thoroughly in a wider range of climate regimes, building on the knowledge base that has been accumulated over the last decade of application. There is also a community-wide need for practical guidance on setting bounds to weather generation for stress testing. Again, this should look beyond the scenario-led framework that would conventionally turn to the latest climate model ensembles but, instead, be guided by knowledge of the vulnerabilities of the system of interest.

Acknowledgements

The authors thank Dr Tom Matthews for assistance in producing some of the graphics.

References

Bastola S, Murphy C, Fealy R (2012) Generating probabilistic estimates of hydrological response for Irish catchments using a weather generator and probabilistic climate change scenarios. *Hydrological Processes*, **26**, 2307-2321.

Bastola S, Murphy C, Sweeney J (2011) The role of hydrological modelling uncertainties in climate change impact assessments of Irish river catchments. *Advances in Water Resources*, **34**, 562-576.

Brown C, Werick W, Leger W, Fay D (2011) A decision-analytic approach to managing climate risks: application to the Upper Great Lakes. *Journal of the American Water Resources Association* **47**, 524-534.

Brown C, Wilby RL (2012) An alternate approach to assessing climate risks. *Eos*, **92**, 401-403.

Camberlin P, Gitau W, Oettli P, Ogallo L, Bois B (2014) Spatial interpolation of daily stochastic generation parameters over East Africa. *Climate Research*, **59**, 39-60.

Chu JT, Xia J, Xu CY, Singh VP (2010) Statistical downscaling of daily mean temperature, pan evaporation and precipitation for climate change scenarios in Haihe, China. *Theoretical and Applied Climatology*, **99**, 149-161.

Conway D, Mould C, Bewket W (2004) Over one century of rainfall and temperature observations in Addis Ababa, Ethiopia. *International Journal of Climatology*, **24**, 77-91.

Cueto ROG, Martinez AT, Ostos EJ (2010) Heat waves and heat days in an arid city in the northwest of Mexico: current trends and in climate change scenarios. *International Journal of Biometeorology*, **54**, 335-345.

Deser C, Phillips A, Bourdette V, Teng H (2012) Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, **38**, 527-546.

Dobler C, Hagemann S, Wilby RL, Stötter J (2012) Quantifying different sources of uncertainty in hydrological projections at the catchment scale. *Hydrology and Earth Systems Science*, **16**, 4343-4360.

Donovan B (2003) *An investigation into the relationship between large scale atmospheric variables, wave climate and weather related sea level variations*. Unpublished MSc Thesis, Kings College London, 57pp.

Hackney CR (2013) *Modelling the effects of climate change and sea level rise on the evolution of incised coastal gullies*. Unpublished PhD thesis, University of Southampton, 203pp.

Harpham C, Wilby RL (2005) Multi-site downscaling of heavy daily precipitation occurrence and amounts. *Journal of Hydrology*, **312**, 235-255.

Harrigan S, Murphy C, Hall J, Wilby RL, Sweeney J (2014) Attribution of detected changes in streamflow using multiple working hypotheses. *Hydrology and Earth System Sciences*, **18**, 1935-1952.

Hashmi MZ, Shamseldin AY, Melville BW (2011) Comparison of SDSM and LARS-WG for simulation and downscaling of extreme precipitation events in a watershed. Stochastic *Environmental Research and Risk Assessment*, **25**, 475-484.

Hassan Z, Shamsudin S, Harun S (2014) Application of SDSM and LARS-WG for simulating and downscaling of rainfall and temperature. *Theoretical and Applied Climatology*, **116**, 243-257.

Hinkley D (1977) On quick choice of power transformation. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **26**, 67-69.

Holloway T, Spak SN, Barker D, Bretl M, Moberg C, Hayhoe K, Van Dorn J, Wuebbles D (2008) Change in ozone air pollution over Chicago associated with global climate change. *Journal of Geophysical Research-Atmospheres*, **113**, D22306.

Hosking JRM, Wallis JR (1997) *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press, 224pp.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Wollen 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.

Kundzewicz ZW, Stakhiv EZ (2010) Are climate models "ready for prime time" in water resources management applications, or is more research needed? *Hydrological Sciences Journal*, **55**, 1085-1089.

Lempert R, Nakicenovic N, Sarewitz D, Schlesinger M (2004) Characterizing climate change uncertainties for decision-makers. *Climatic Change*, **65**, 1-9.

Lempert R, Sriver RL , Keller K (2012) *Characterizing Uncertain Sea Level Rise Projections to Support Investment Decisions.* California Energy Commission. Publication Number: CEC-500-2012-056. RAND, Santa Monica, California.

Li Z, Zheng F-L, Liu W-Z (2011) Spatiotemporal characteristics of reference evapotranspiration during 1961-2009 and its projected changes during 2011-2099 on the Loess Plateau of China. *Agriculture and Forest Meteorology*, **154-155**, 147-155.

Manley RE (2006) *A guide to using HYSIM, R. E. Manley and Water Resource Associates*, Ltd.

Murphy C, Fealy R, Charlton R, Sweeney J (2006) The reliability of an 'off-the-shelf' conceptual rainfall runoff model for use in climate impact assessment: uncertainty quantification using Latin hypercube sampling. *Area* **38**, 65-78.

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models, Part 1: A discussion of principles. *Journal of Hydrology*, **10**, 282-290.

Nazemi A, Wheater HS, Chun KP, Elshorbagy A (2013) A stochastic reconstruction framework for analysis of water resource system vulnerability to climate-induced changes in river flow regime. *Water Resources Research*, **49**, 291-305.

O'Connor P (2013) *Assessment of a decision-centric approach to climate change adaptation*. Unpublished Masters thesis. National University of Ireland, Maynooth, 80pp.

Office of Public Works (OPW) (2009) *Assessment of Potential Future Scenarios*. Flood Risk Management Draft Guidance.

Pielke RA Sr., Wilby RL (2012) Regional climate downscaling – what's the point? *Eos*, **93**, 52-53.

Prudhomme C, Wilby RL, Crooks S, Kay AL, Reynard NS (2010) Scenario-neutral approach to climate change impact studies: application to flood risk. *Journal of Hydrology*, **390**, 198-209.

Sakia RM (1992) The Box-Cox transformation technique: a review. *The Statistician*, **41**, 169-178.

Semenov MA, Brooks RJ (1999) Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain. *Climate Research*, **11**, 137-148.

Semenov MA, Donatelli M, Stratonovitch P, Chatzidaki E, Brauth B (2010) ELPIS: a dataset of local-scale daily climate scenarios for Europe. *Climate Research*, **44**, 3-15.

Semenov, MA, Pilkington-Bennett S, Calanca P (2013) Validation of ELPIS 1980-2010 baseline scenarios using the observed European Climate Assessment data set. *Climate Research*, **57**, 1-9.

Smith A, Freer J, Bates P, Sampson C (2014) Comparing ensemble projections of flooding against flood estimations by continuous simulation. *Journal of Hydrology*, **511**, 205-219.

Stainforth DA, Downing TE, Lopez RWA, New M (2007). Issues in the interpretation of climate model ensembles to inform decisions. *Philosophical Transactions of the Royal Society A*, **365**, 2163-177.

Stakhiv EZ (2011) Pragmatic approaches for water management under climate change uncertainty. *Journal of the American Water Resources Association*, **47**, 1183-1196.

Steinschneider S, Brown C (2013) A semiparametric multivariate, multi-site weather generator with low-frequency variability for use in climate risk assessments. *Water Resources Research*, **49**, 7205-7220.

Tisseuil C, Vrac M, Lek S, Wade, AJ (2010) Statistical downscaling of river flows. *Journal of Hydrology*, **385**, 279-291.

Turner, S.W.D., Marlow, D., Ekström, M., Rhodes, B.G., Kularathna, U. and Jeffrey, P.J. 2014. Linking climate projections to performance: A yield-based decision scaling assessment of a large urban water resources system. *Water Resources Research*, **50**, 3553-3567.

Wang W, Xing W, Shao Q, Yu Z, Peng S, Yang T, Yong B, Taylor J, Singh VP (2013) Changes in reference evapotranspiration across the Tibetan Plateau: Observations and future projections based on statistical downscaling. *Journal of Geophysical Research: Atmospheres*, **118**, 4049-4068.

Weaver CP, Lempert RJ, Brown C, Hall JA, Revell D, Sarewitz D (2013) Improving the contribution of climate model information to decision making: the value and demands of robust decision frameworks. *WIREs Climate Change,* **4**, 39-60.

Whetton P, Hennessy K, Clarke J, McInnes K, Kent D (2012) Use of Representative Climate Futures in impact and adaptation assessment. *Climatic Change*, **115**, 433-442.

Wilby RL (2008a) Constructing climate change scenarios of urban heat island intensity and air quality. *Environment and Planning B: Planning and Design*, **35**, 902-919.

Wilby RL (2008b) Downscaling future skew surge statistics at Sheerness, Kent. Phase 3 studies – synthesis report. Thames Estuary 2100, *Environment Agency*, 27pp.

Wilby RL, Dawson CW (2013) The Statistical DownScaling Model (SDSM): Insights from one decade of application. *International Journal of Climatology*, **33**, 1707-1719.

Wilby RL, Dawson CW, Barrow EM (2002) SDSM - a decision support tool for the assessment of regional climate change impacts. *Environmental and Modelling Software*, **17**, 145-157.

Wilby RL, Dessai S (2010) Robust adaptation to climate change. *Weather*, **65**, 180-185.

Wilby RL, Harris I (2006) A framework for assessing uncertainties in climate change impacts: low flow scenarios for the River Thames, UK. *Water Resources Research*, **42**, W02419.

Wilby RL, Keenan R (2012) Adapting to flood risk under climate change. *Progress in Physical Geography*, **36**, 349-379.

Wilby RL, Miller KA, Yates D, Kaatz L (2013) Use of narrative scenarios for evaluating drought management responses in the Upper Colorado River Basin. *American Geophysical Union, Fall Meeting 2013*, abstract #H34C-02,

Wilby RL, Tomlinson OJ, Dawson CW (2003) Multi-site simulation of precipitation by conditional resampling. *Climate Research*, **23**, 183-194.

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

Wilby RL, Yu D (2013) Rainfall and temperature estimation for a data sparse region. *Hydrology and Earth System Sciences*, **17**, 3937-3955.

Wilby, RL, Hassan H, Hanaki K (1998) Statistical downscaling of hydrometeorological variables using General Circulation Model output. *Journal of Hydrology*, **205**, 1-19.

Wilks DS, Wilby RL (1999) The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, **23**, 329-357.

Wise K (2009) Climate-based sensitivity of air quality to climate change scenarios for the southwestern United States. *International Journal of Climatology*, **29**, 87-97.

World Bank Independent Evaluation Group (2012) *Adapting to Climate Change: Assessing World Bank Group Experience*. World Bank Group, Washington DC, 193pp.

Yang T, Li H, Wang W, Xu C-Y, Yu Z (2012) Statistical downscaling of extreme daily precipitation, evaporation, and temperature and construction of future scenarios. *Hydrological Processes*, **26**, 3510-3523.

Table 1 Examples of direct downscaling of exotic variables using SDSM

Variable	Location	Source
Evaporation	Haihe, China	Chu et al. (2010)
	Loess plateau, China	Li et al. (2012)
	Tibetan plateau, Tibet	Wang et al. (2013)
	River Kennet, UK	Wilby et al. (2006)
	River Dongjiang, China	Yang et al. (2012)
Ground-level ozone and/or particulates	Chicago, US	Holloway et al. (2008)
	London, UK	Wilby (2008a)
	Tucson, US	Wise (2009)
Heat wave indices	Mexicali, Mexico	Cueto et al. (2010)
	London, UK	Wilby (2007)
Waves and tidal surge	North Sea, UK	Donovan (2003)
	Isle of Wight, UK	Hackney (2013)
	Thames Estuary, UK	Wilby (2008b)

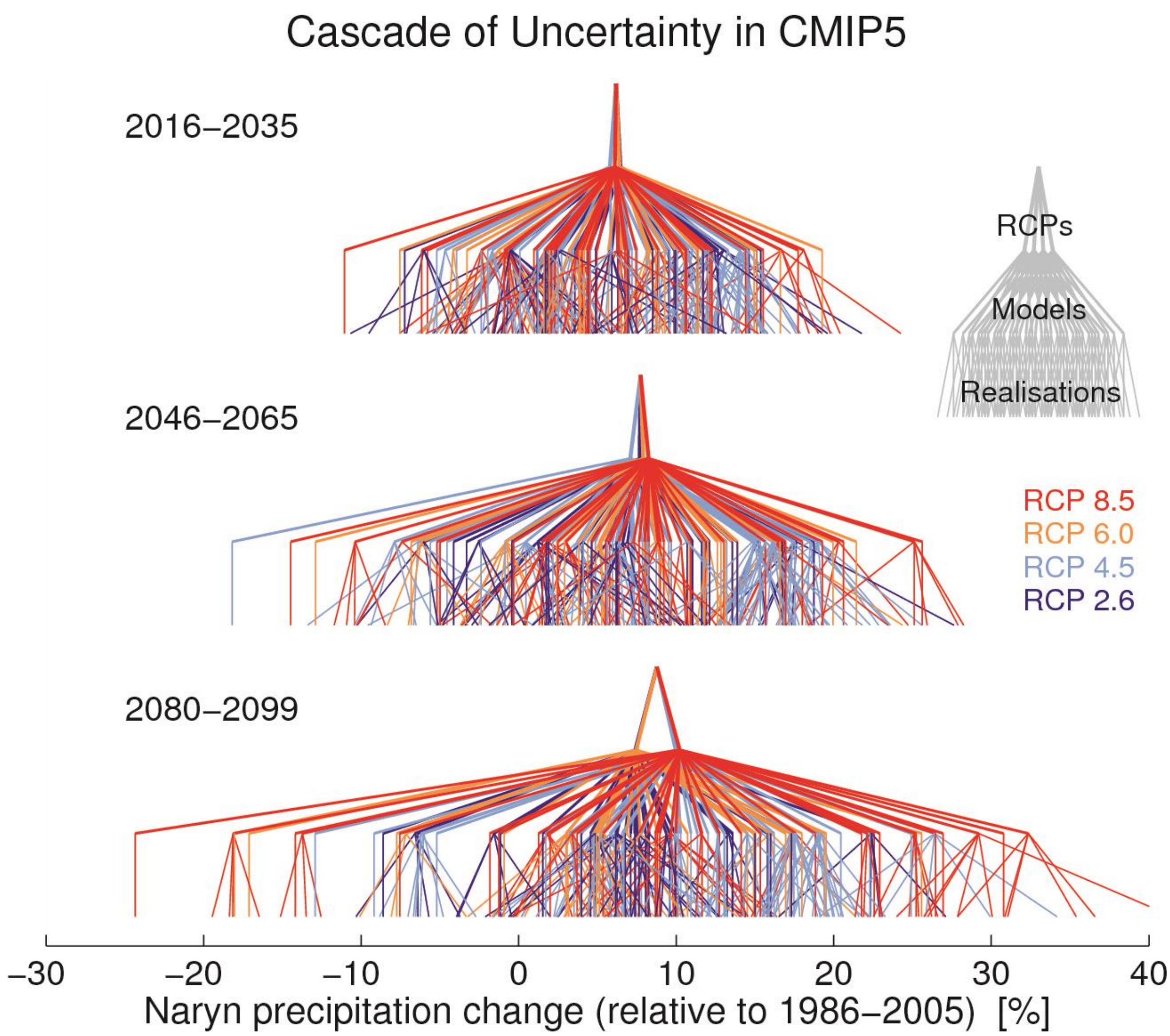


Figure 1 A 'cascade of uncertainty' in precipitation changes projected by the CMIP5 ensemble for the River Naryn basin, Central Asia (70-80°E, 40-45°N). The three levels of each pyramid illustrate uncertainty due to the choice of Representative Concentration Pathway (RCP), GCM and realisation of climate variability. Not all simulations have multiple realisations, resulting in a vertical line in the lowest layer. The intersection on the top row for each time period is the multi-scenario, multi-model, multi-realisation mean.

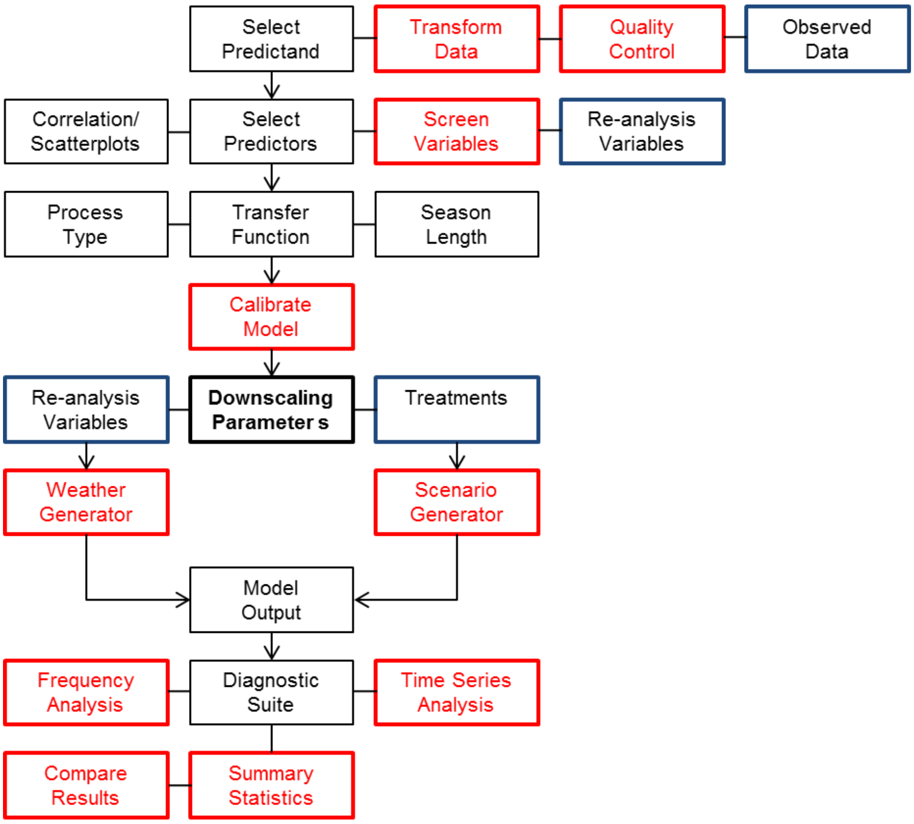


Figure 2 SDSM-DC architecture showing inputs (blue boxes) and screens (red boxes).

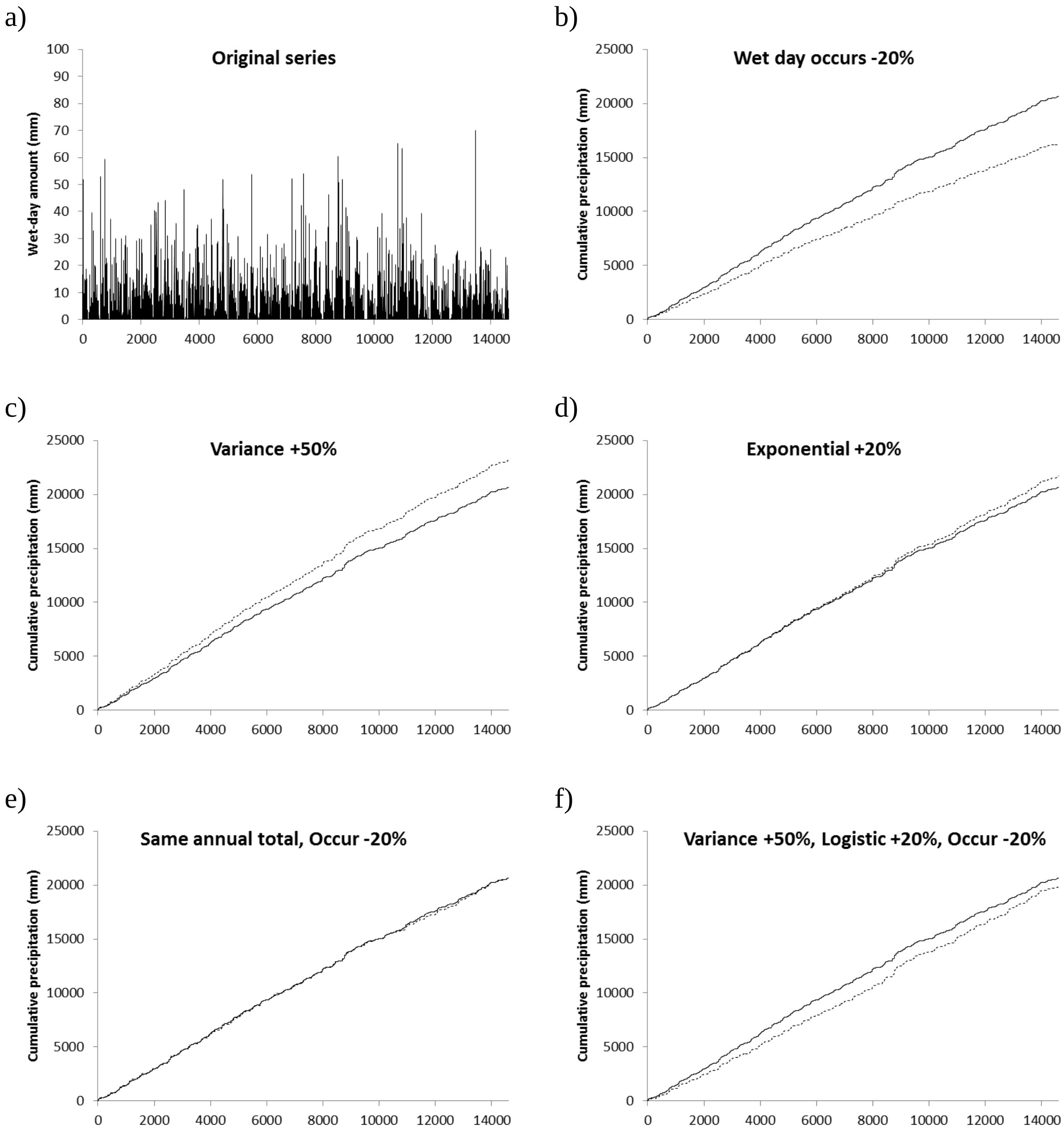


Figure 3 Example SDSM-DC treatments applied to a 40-year daily precipitation series. The dark line shows the original data and the grey line the treated series, both expressed as cumulative totals for ease of comparison.

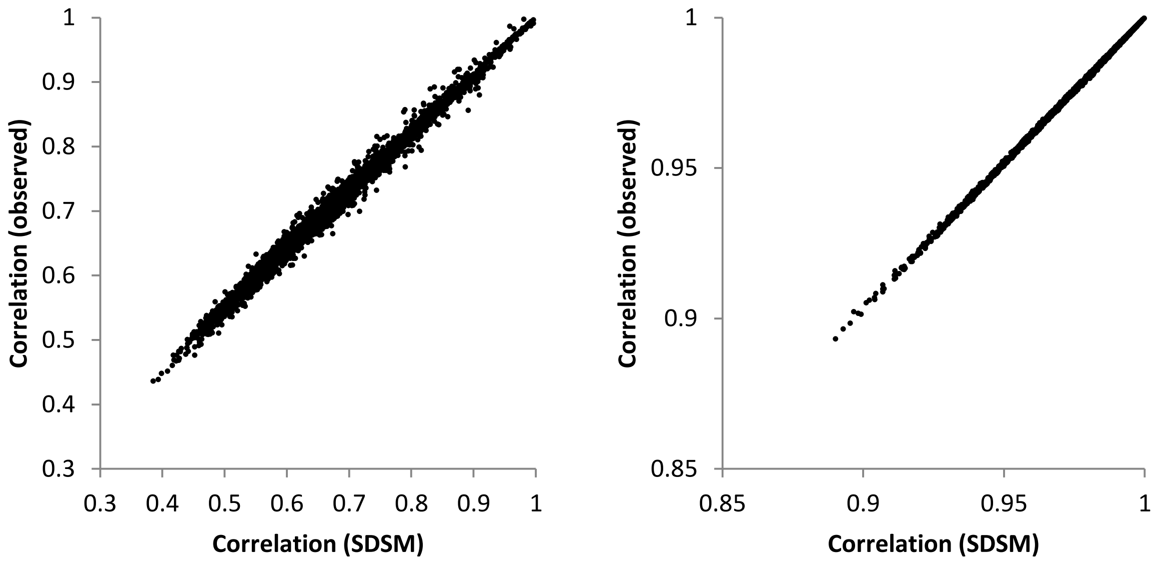
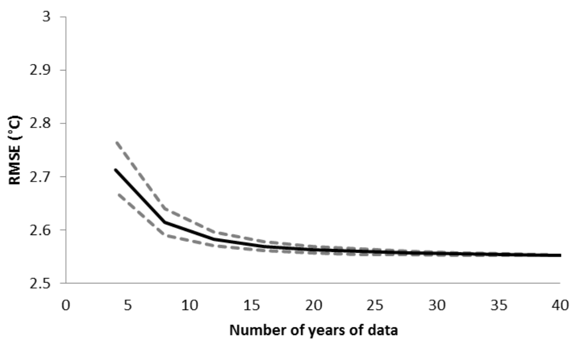
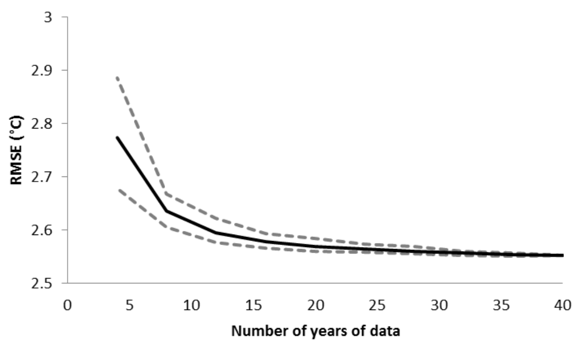


Figure 4 Pairwise correlation of observed and downscaled daily precipitation (left) and mean temperature (right) in the Upper Colorado River Basin. Source: Wilby et al. (2013).

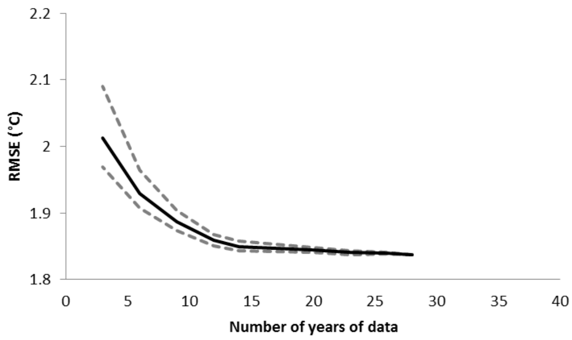
a) Prince Edward Island (day)



b) Prince Edward Island (year)



c) Tunis (day)



d) Tunis (year)

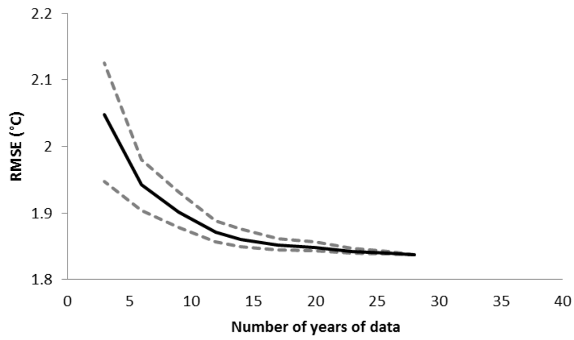
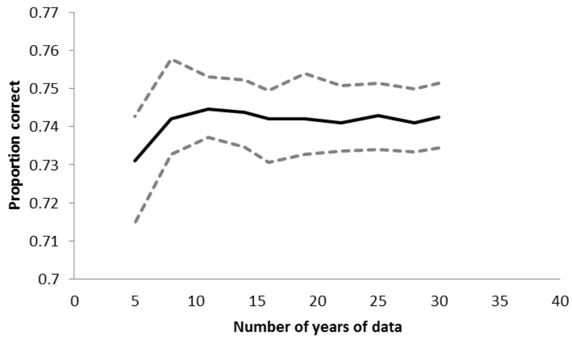
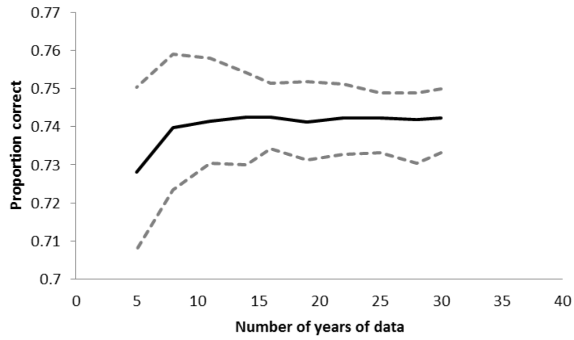


Figure 5 Effects of missing data on the Root Mean Squared Error (RMSE) of downscaled daily mean temperature depending on whether random days or blocks of years are omitted for a,b) Charlottetown, Prince Edward Island, Canada and for c,d) Tunis, Tunisia. Each plot shows the range (dashed lines) and median (solid line) RMSE based on 100 simulations.

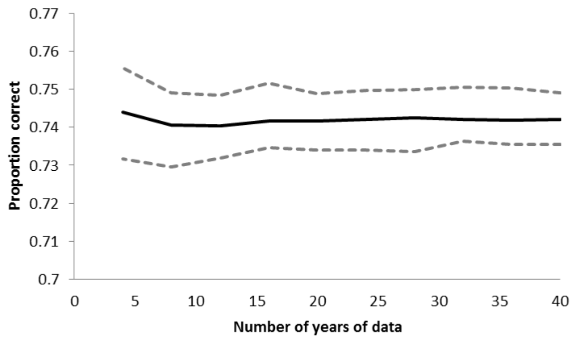
a) Addis Ababa (day)



b) Addis Ababa (year)



c) Chang wu (day)



d) Chang wu (year)

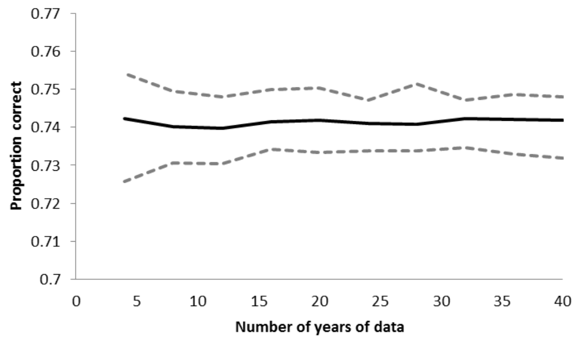


Figure 6 Effects of missing data on the proportion correct wet-day occurrence (PCW) depending on whether random days or blocks of years are omitted for a,b) Addis Ababa, Ethiopia and for c,d) Chang wu, China.

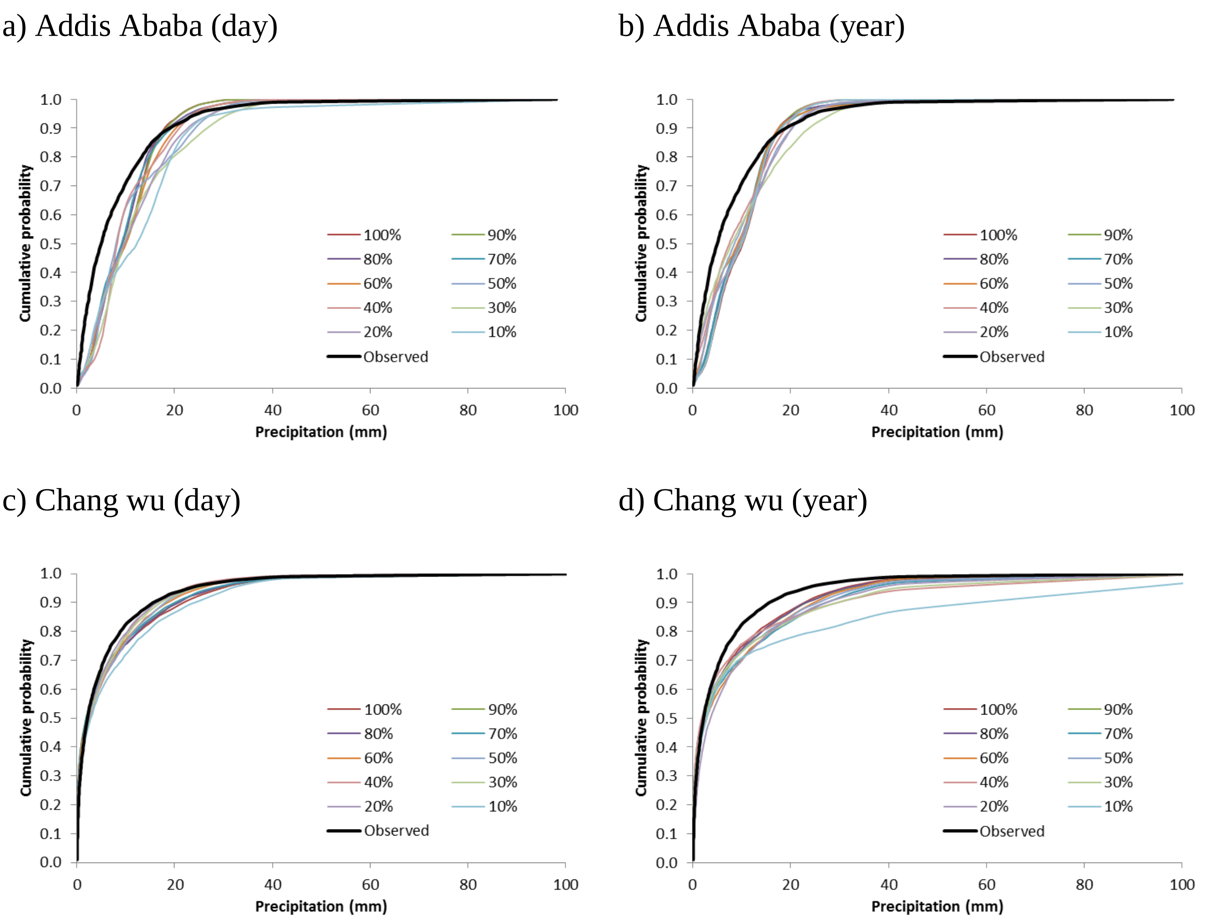
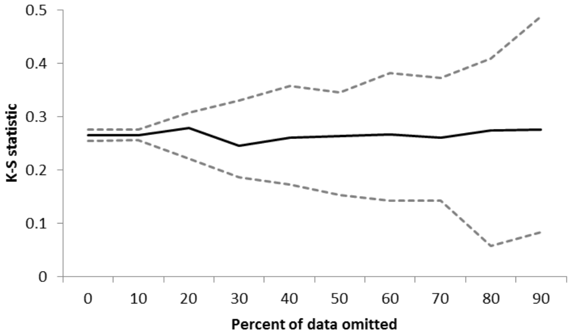
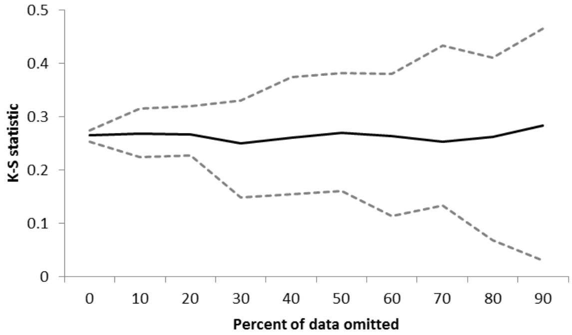


Figure 7 Sensitivity of downscaled daily precipitation distributions to percent of data omitted by random day (left) or year (right) removal for Addis Ababa (upper) and Chang wu (lower)..

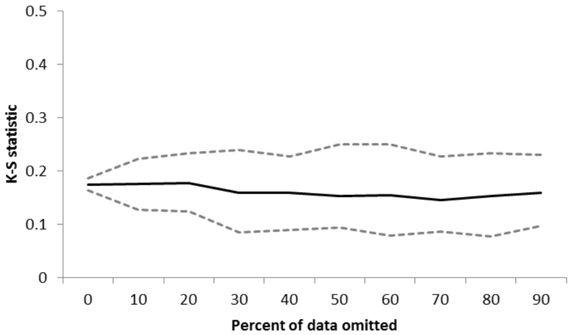
a) Addis Ababa (day) [0.8%]



b) Addis Ababa (year) [1.2%]



c) Chang wu (day) [22.3%]



d) Chang wu (year) [22.3%]

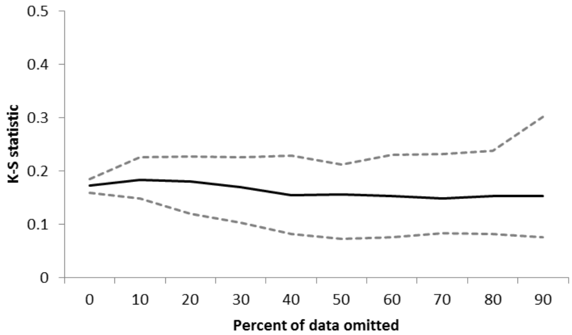


Figure 8 Sensitivity of the Kolmogorov-Smirnov statistic to percent of data omitted by random day (left) or year (right) removal for Addis Ababa (upper) and Chang wu (lower). The percent of simulations with KS < Dcrit (0.14 at p=0.05) is given [in brackets].

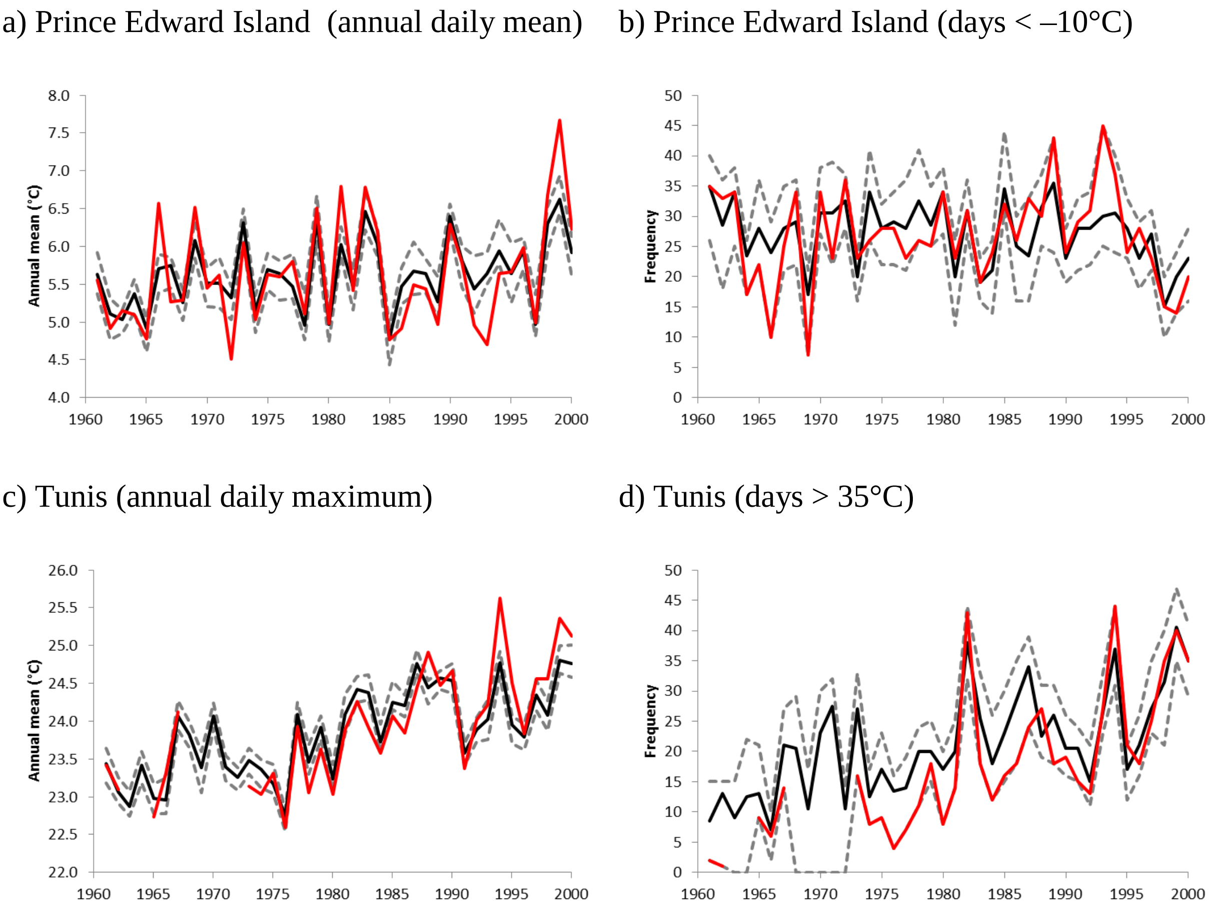


Figure 9 Reconstructed and in-filled (solid black line) temperatures compared with observations (red line) for a, b) Prince Edward Island, Canada and c,d) Tunis, Tunisia. Dashed lines show the downscaled ensemble range.

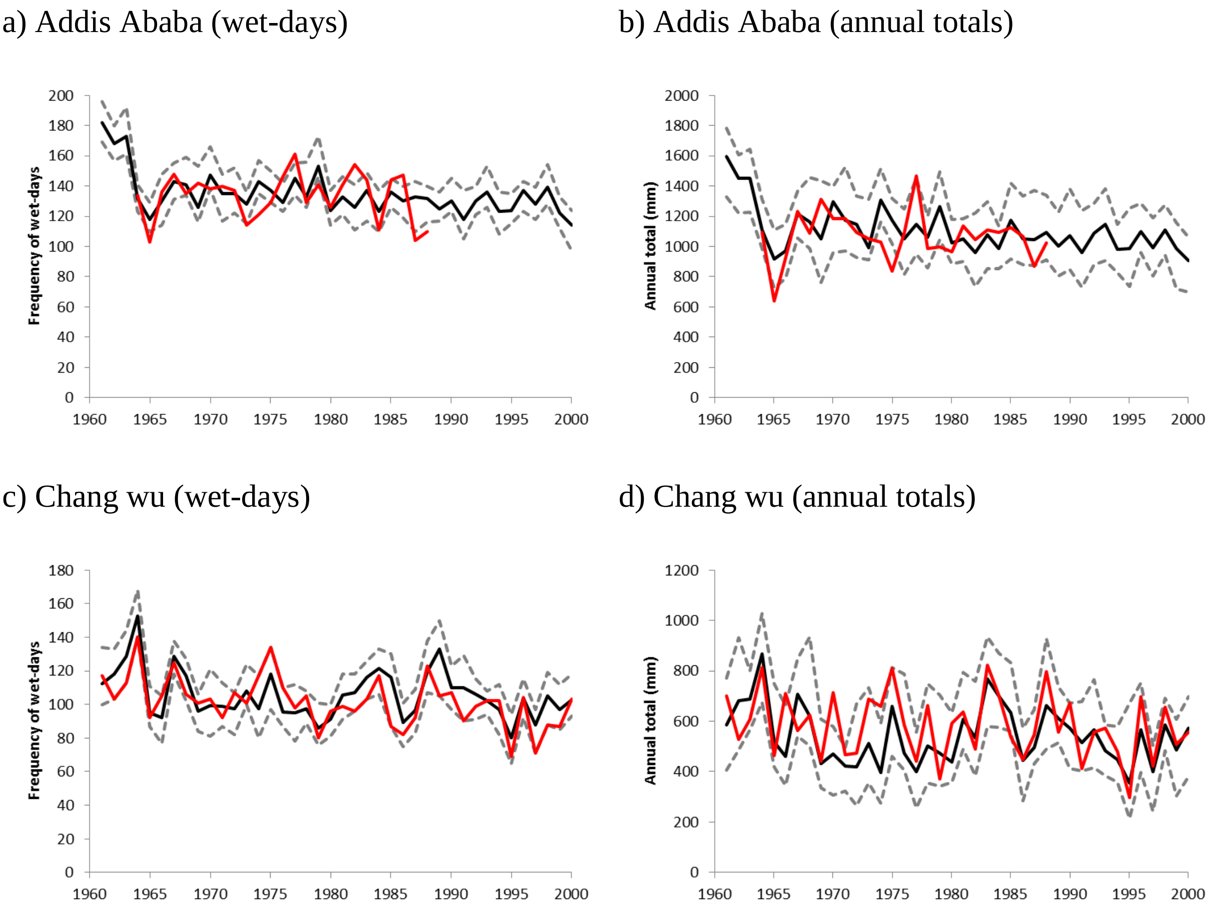


Figure 10 Reconstructed wet-day frequencies and annual precipitation totals for a,b) Addis Ababa, Ethiopia and c,d) Chang wu, China.

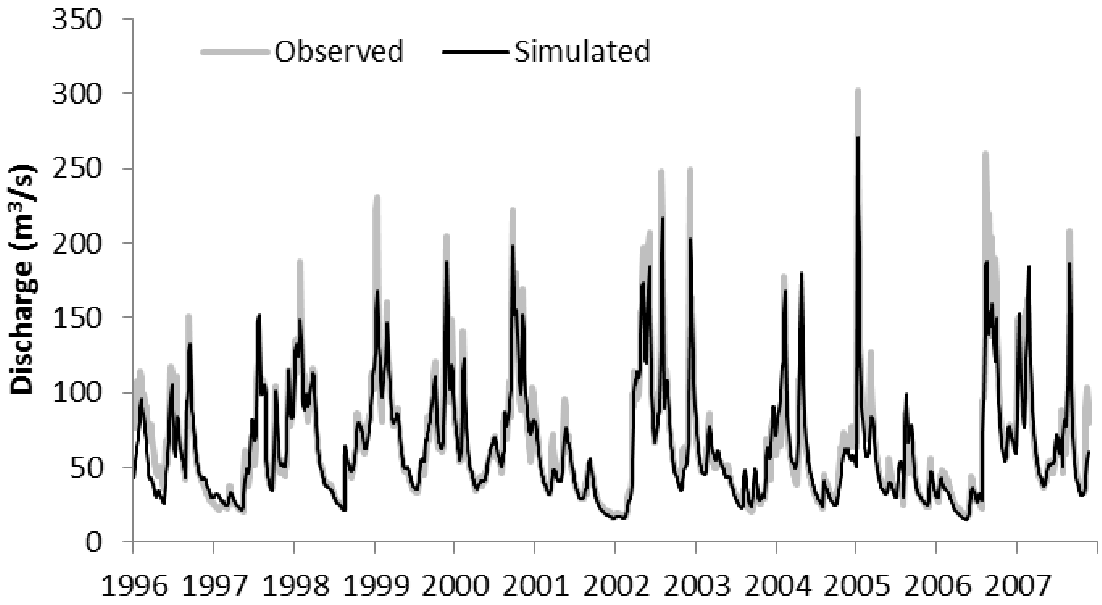


Figure 11 Comparison of observed (grey line) and HYSIM (black line) simulations of winter daily flows in the River Boyne for the verification period 1997-2007.

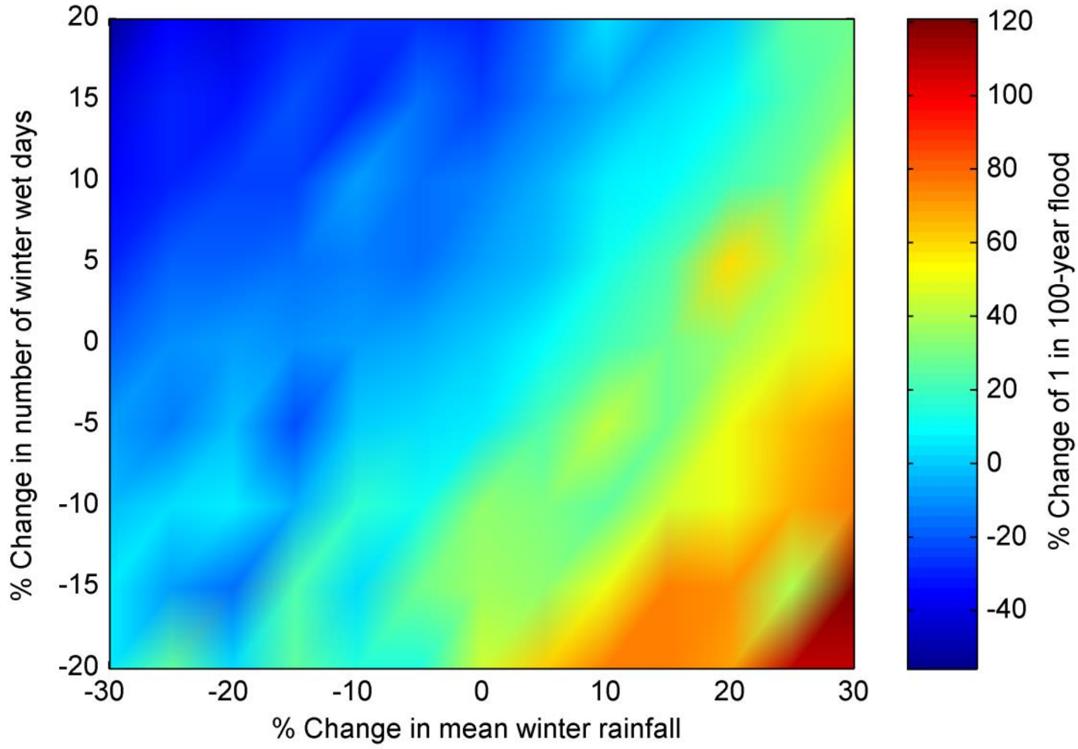


Figure 12 Response surface representing the sensitivity of percent changes in the magnitude of the winter 1-in-100 year flood to changes in mean winter rainfall and occurrence of winter wet days.

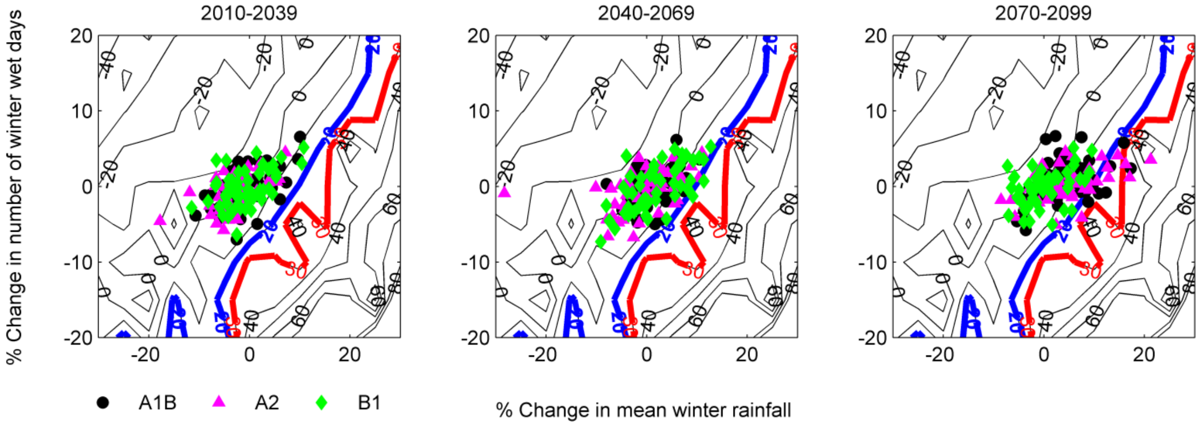


Figure 13 Sensitivity of precautionary allowances to projected changes in climate during winter months (DJF). Contours representing allowances of an additional 20 and 30 % of design flow (1-in-100 year flood) are highlighted in blue and red respectively. Climate change projections (Bastola et al., 2011) represent a sample of 17 GCMs from the CMIP3 project forced with the A1B, A2 and B1 SRES emissions scenarios for the 2020s (2010-39), 2050s (2040-69) and 2080s (2070-99).