Bias correction of daily precipitation simulated by a Regional Climate Model: A comparison of methods

Thomas Lafon[1,2,3], Simon Dadson[1], Gwen Buys[1,4], and Christel Prudhomme[1]

[1] Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, OX10 8BB. UK

[2] Now at Oxford Brookes University, Headington Campus, Gipsy Lane, Oxford OX3 0BP, UK

[3] Fundación Entropika, Apartado Aéreo N° 20,

Leticia, Amazonas, Colombia

[4] Now at British Antarctic Survey, High Cross, Madingley Road,

Cambridge, CB3 0ET, UK

Corresponding author: tlafon@entropika.org

Thomas Lafon

Abstract

Quantifying the effects of future changes in the frequency of precipitation extremes is a key challenge in assessing the vulnerability of hydrological systems to climate change, but is difficult as climate models do not always accurately simulate daily precipitation. This paper compares the performance of four published techniques used to reduce the bias in a Regional Climate Model (RCM) precipitation output: (i) linear, (ii) non-linear, (iii) gamma-based quantile mapping and (iv) empirical quantile mapping. Overall performance and sensitivity to the choice of calibration period were tested by calculating the errors in the first four statistical moments of generated daily precipitation time series and using a cross validation technique. The study compared the 1961-2005 precipitation time series from the Regional Climate Model HadRM3.0-PPE-UK (unperturbed version) with gridded daily precipitation time series derived from rain gauges for seven catchments spread throughout Great Britain. We found that whilst the first and second moments of the precipitation frequency distribution can be corrected robustly, correction of the third and fourth moments of the distribution is much more sensitive to the choice of bias-correction procedure and to the selection of a particular calibration period. Overall, our results demonstrate that, if both precipitation datasets can be approximated by a gamma distribution, the gamma-based quantile-mapping technique offers the best combination of accuracy and robustness. In circumstances where precipitation datasets cannot adequately be approximated using a gamma distribution, the non-linear method is more effective at reducing the bias but the linear method is least sensitive to the choice of calibration period. The empirical quantile mapping method can be highly accurate, but results were very sensitive to the choice of calibration time period. However, it should be borne in mind that bias correction introduces additional uncertainties, which are greater for higher-order moments.

Key words: Regional climate model, bias correction, daily precipitation, downscaling, cross-validation, UK.

Sponsors: Natural Environment Research Council (UK); Environment Agency (UK); DEFRA (UK), UK Water Industry Research (UKWIR).

1. Introduction

The impact of climate change on the hydrological cycle is of great interest to environmental and water resource managers (Arnell, 2001, Bates et al., 2008). Quantifying the effects of future changes in the frequency of daily precipitation extremes is a key challenge in assessing the vulnerability of hydrological systems to climate change. Nevertheless, whilst the accuracy of Global Climate Models (GCMs) in simulating the large-scale atmospheric circulation has improved markedly in recent years, global models have difficulty resolving the processes that govern local precipitation. The most common problem associated with GCM simulations of precipitation is that, at a daily time-scale, precipitation occurs more frequently than observed, but often with a lower intensity (e.g., Sun et al., 2006).

In order to make simulations at hydrologically-relevant spatial and temporal scales, downscaling is necessary. Downscaling techniques that have been reviewed in the literature include statistical downscaling, which uses empirical relations between climate model outputs and historical observed data, and dynamical downscaling, which involves the use of a Regional Climate Model (RCM) (see Fowler et al., 2007 for a detailed review). RCMs offer a more physically-realistic approach to GCM downscaling than statistical downscaling because they provide an explicit representation of the mesoscale atmospheric processes that produce heavy precipitation. When nested within a GCM, RCMs provide regional detail that is not only consistent with the parent GCM, but which is spatially-coherent. That is, a degree of spatial persistence of large-scale atmospheric features is automatically ensured, because the model generates these features dynamically. This property of RCM simulations is important in producing realistic forcing data for hydrological models because many floods and droughts are caused by spatially- and temporally-persistent precipitation patterns. Two major studies of the accuracy of RCM precipitation estimates used daily extreme precipitation statistics to compare the performance of several different 50 km RCMs nested within both ECMWF ERA-15 reanalysis data (Frei et al., 2003), and within the Hadley Centre HadAM3 GCM (Frei et al., 2006). They found that the RCMs were capable of reproducing important mesoscale patterns of observed precipitation, particularly during autumn and winter, and demonstrated that the effects of topography on precipitation could be much better represented at regional scales. Nevertheless, Frei et al. (2006) found that appreciable model biases could occur in summer when convective precipitation is common. Simulations using two Hadley Centre RCMs (HadRM2 and HadRM3) nested within HadCM2 demonstrated that UK precipitation statistics were well reproduced by the model, and that biases in the rainfall extremes (on timescales from one to thirty days and return periods of two to twenty years) were approximately equivalent in magnitude to those in the mean (Buonomo et al., 2007).

Evidence of bias in GCM and RCM precipitation data has prompted many investigators to avoid direct use of climate model precipitation outputs for hydrological climate change impact analysis. One very common approach is to calculate the difference between future and baseline climate from the GCM/RCM outputs, and apply this 'factor of change' to historical observed time series to generate synthetic time series assumed to be possible realisations of the future (e.g., Arnell et al., 2003). However, this approach does not change any of the temporal structure of the time series. Another approach is to generate synthetic precipitation time series using a stochastic weather generator, where the parameters in the generator are changed according to estimated changes in the climate from the GCM/RCM outputs (e.g. Kilsby et al., 2007, Fatichi et al., 2011). However, most weather generators produce local time series, and if applied to different locations, do not generate spatially consistent data, limiting their use to local studies. For analysis where spatio-temporal structure of precipitation is important, or for regional or global studies, the use of climate model outputs provides this spatial consistency and incorporates any changes in the temporal structure of precipitation. Techniques to correct the biases in the climate model outputs are therefore used to improve the realism of GCM/RCM precipitation time series, based on statistical properties obtained from observed data taken from the same baseline period.

The bias in GCM and RCM daily precipitation simulations may not be limited to monthly means, but may also affect precipitation variability and other derived measures that are of hydrological importance (Arnell et al., 2003, Diaz-Nieto and Wilby, 2005, Fowler et al., 2007). These effects include differences in variability and extreme daily precipitation, and differences between the modelled and observed distribution of dry days, and periods of dry days. When correcting for biases in climate model output, it is also important that changes in the frequency distribution of climatic variables are correctly represented. Several techniques for adjusting biases in GCM and RCM daily precipitation have been published. These techniques can be grouped into the following four families: (i) linear (e.g., by using a seasonally- and spatially-varying change factor; Lenderink et al., 2007); (ii) non-linear (e.g., by using a seasonally- and spatially-varying change factor and exponent; Leander and Buishand, 2007); (iii) distribution-based quantile mapping (e.g., gamma distribution, Hay et al., 2002, Piani et al., 2010); and (iv) distribution-free quantile mapping (e.g., empirical distribution, Wood et al., 2002, Wood et al., 2004, Ashfaq et al., 2010).

We have considered a range of established bias-correction techniques to determine which is the most effective and robust method to use when correcting daily precipitation simulated by a RCM for subsequent use in a hydrological model. This paper discusses the accuracy of the four techniques in detail when applied to the Exe-Culm river basin in south-west England. This is followed by summary results for another six basins spread throughout Great Britain, each representative of different climate conditions; Scotland is the coolest and wettest part of the United Kingdom throughout most of the year, while the South is usually warmer with England being the driest region. Observed daily gridded precipitation was obtained at 1 km horizontal resolution by interpolation from Met Office rain gauges (Keller et al., 2006). RCM daily precipitation at 0.22º horizontal resolution (approximately 25 km) was obtained from the HadRM3.0-PPE-UK unperturbed model from the ensemble of perturbed-physics experiments employed in the UK Climate Impacts Programme study UKCP09 (Jenkins et al., 2009). Prior to the bias-correction, the RCM time series where regridded to 1 km horizontal resolution and modified so that the frequency of wet days was the same as that in the 1 km gridded observed dataset. Each bias-correction method was then calibrated using the 1 km modified RCM time series for each month independently.

This paper is organized as follows. The first part describes the bias-correction methods that form the subject of this study. The statistical methods used to evaluate the bias corrections are then outlined. This is followed by a description of the data and study regions, before the results of the evaluation procedures are presented. The paper concludes with a discussion of the findings.

2. Bias-correction methods

2.1. Linear correction method

When using the linear correction method, RCM daily precipitation amounts, P, are transformed into P^* such that $P^* = aP$, using a scaling factor, $a = \bar{O}/\bar{P}$, wherein $\bar{O}$ and $\bar{P}$ are the monthly mean observed and RCM precipitation for that 1 km grid point, respectively. Here, the monthly scaling factor is applied to each uncorrected daily observation of that month, generating the corrected daily time series. The linear correction method belongs to the same family as the 'factor of change' or 'delta change' method (Hay et al., 2000). This method has the advantage of simplicity and modest data requirements: only monthly climatological information is required in order to calculate monthly correction factors. However, correcting only the monthly mean precipitation can distort the relative variability of the inter-monthly precipitation distribution, and may adversely affect other moments of the probability distribution of daily precipitations (Arnell et al., 2003, Diaz-Nieto and Wilby, 2005).

2.2. Non-linear correction method

Noting that a linear scaling factor adjusts the mean but not the standard deviation of monthly precipitation, Shabalova et al. (2003) and Leander and Buishand (2007) advocate the use of a power-law correction such that $P^* = aP^b$, where b is a scaling exponent. The constants a and b are calculated in two stages: (i) the scaling exponent, b, is calculated iteratively so that, for each grid box in each month, the coefficient of variation of the RCM daily precipitation time series matches that of the observed precipitation time series. Here, this is achieved using Brent's method (Press et al., 1986); (ii) the prefactor, a is then calculated so that the mean of the transformed precipitation values is equal to the observed mean. Finally, monthly constants a and b are applied to each uncorrected daily observations corresponding to that month in order to generate the corrected daily time series.

In common with the linear method, this approach has the advantage that it requires monthly observed statistics but, in addition to the mean, it needs information on the coefficient of variation of precipitation. This approach results in the mean and the standard deviation of the daily precipitation distribution becoming equal to those of the observed distribution. Biases in higher order moments are not removed by the non-linear method; however these will be affected to a certain degree by the correction procedure.

2.3. Gamma distribution correction method

The gamma distribution-based correction method assumes that the probability distributions of both observed and RCM daily precipitation datasets can be approximated using a gamma distribution, for example:

$$f(P; k, \theta) = P^{k-1} \frac{\exp(-P/\theta)}{\Gamma(k)\theta^k}, \qquad [5]$$

where $k > 0$ and $\theta > 0$ are the form and scaling parameters of a gamma distribution, respectively, and where P represents RCM daily precipitation. Here, parameters k and θ were estimated for each grid box for each month, using the method of moments:

$$k = \left(\frac{\bar{P}}{\sigma_P}\right)^2, \qquad [6]$$

$$\theta = \frac{{\sigma_P}^2}{\bar{P}}, \qquad [7]$$

where $\bar{P}$ and σ_P are the sample mean and standard deviation of P, respectively.

In order to perform the bias-correction, each daily RCM precipitation amount was expressed as a quantile, q, calculated as the inverse of the cumulative distribution function, F, where:

$$F(P; k, \theta) = \frac{\gamma(k, P/\theta)}{\Gamma(k)}, \qquad [8]$$

wherein γ is the lower incomplete gamma function (e.g., Press et al., 1986), and k and θ are the parameters of the gamma distribution fitted to the RCM simulated precipitation. This quantile was then used to generate a bias-corrected precipitation time series by replacing the RCM precipitation amount P by its value resampled from the gamma distribution fitted to the observations and associated with the same quantile. This method is designed to remove biases in the first two statistical moments and similar methods were found to perform well when used on GCM outputs at global and European scales (Vidal and Wade, 2008a, b, Piani et al., 2010).

2.4. Empirical distribution correction method

The correction method based on empirical distributions follows the same approach as the gamma distribution method, with the RCM distribution transformed to match the observed distribution through a transfer function. Unlike the gamma distribution method, the empirical method does not make any *a priori* assumptions about the precipitation distribution (Figure 1).

To implement the empirical distribution correction method, the ranked observed precipitation distribution is divided into a number of discrete quantiles. For each quantile division, a linear correction factor was calculated by dividing the mean observation in that quantile by the RCM simulated mean precipitation in the same quantile; this being the transfer function. The number of quantile divisions controls the accuracy of the method: using fewer quantiles might smooth out the information contained within the observed record, while using too many quantiles might result in over-fitting of the model to the data. A method of the same family has been shown to perform well in the correction of RCM precipitation forecasts for use as variables of interest for hydrologic simulations and climate change studies (Wood et al., 2002, Wood et al., 2004, Themeßl et al., 2010).

3. Evaluation methodology

Although an overall assessment of the performance of each bias-correction method can be obtained by comparing differences between observed, RCM and corrected datasets, the quantification of robustness is more complex. We have evaluated the robustness of each method by quantifying the sensitivity of its performance to the choice of time period used for calibration. Prior work has noted that properties of the extreme tail of the precipitation distribution can be difficult to estimate robustly using short (e.g., 30-year) RCM simulations, particularly where natural variability of the climate has a significant component varying on decadal time-scales (Kendon et al., 2008). The robustness of the bias-correction techniques is important because it determines the degree to which the correction procedure is sensitive to the differences in natural climate variability between the observed data and the output of the RCM (and its driving GCM).

Here, a cross-validation technique similar to the jack-knife (Bissell and Ferguson, 1975) is used, where measures of performance are evaluated using a sample which was not included in the calibration of the correction procedure. This approach offers the possibility of quantifying sample spread, defined as the difference between the re-sampled time series, and frequency of error reduction. To quantify the sample spread, given a total length of observed record N (in the present case N = 40 years), we repeatedly removed m consecutive years (here m = 10 years), calibrated the parameters of the correction method on the N-m sample, and performed the bias-correction of RCM precipitation for the remaining m-year period. This procedure generated N-$m+1$ sets of bias-corrected precipitation time series. To evaluate the robustness of the correction procedure, we calculated the average of the absolute value of the relative differences (ARD, defined as $|X' - X|/X$, where X and X' are statistics from observed and bias-corrected precipitation, respectively) between the N-$m+1$ sets of corrected and observed precipitation data over the m-year period that was not used to calibrate the bias-correction method. We also calculated the frequency of error reduction, defined as the proportion of bias-corrected time series where ARD was smaller than that calculated from the 1 km RCM-driven data before bias correction. The higher this number, the more certain one can be that the correction method will improve the match between observed and simulated data. A similar procedure has been used to evaluate the robustness of a gamma-based quantile mapping technique in Northern Eurasia (Li et al., 2010).

4. Data

4.1. Study regions

It is important that daily precipitation bias-correction methods are capable of correcting over the full extent of the spatial area of interest. Great Britain has a wide range of annual average precipitation and topography; hence, a method which successfully corrects biases in one region may not necessarily be as effective in another. To explore this question more comprehensively, the four bias-correction methods were each applied to seven test regions.

The seven regions were chosen by comparing river catchments from the National River Flow Archive (NRFA) Hydrometric Register (Marsh and Hannaford, 2008). The starting point for choosing the catchments was the NRFA Benchmark Network, which is a subset of the NRFA stations which have a nearly natural flow regime and little impact from human activity (Bradford and Marsh, 2003). Box-bound regions or multiple catchments are used where single catchment areas are too small to be worth comparing or not large enough to investigate spatial pattern of the results (i.e. East Anglia, North Scotland and Exe-Culm). The catchments were compared using statistics such as elevation range and mean annual average rainfall, so that they describe the range of climatic conditions found in Great Britain. Details of the seven study regions are given in Table 1 and their location is shown in Figure 2.

4.2. Observed data

Our observed precipitation dataset is the 1 km daily precipitation data associated with the Continuous Estimation of River Flows (CERF) model (Keller et al., 2006). It was generated from rain-gauge observations from the UK Met Office using the triangular planes method followed by a normalisation based on average annual precipitation (Jones, 1983). The dataset extends over

England, Scotland and Wales and covers the period 1961–2008, however only data for the period 1961-2000 were used in this study. In total, data from 17,812 rain-gauges were used to derive the CERF dataset.

4.3. RCM data

Daily precipitation outputs from the Met Office Hadley Centre Regional Model Perturbed Physics Ensemble simulations for the 21st Century for the UK domain, HadRM3-PPE-UK (http://badc.nerc.ac.uk/data/hadrm3-ppe-uk/) were considered here. The model was run for the period 1950–2100 at 0.22° horizontal resolution (approximately 25 km), with lateral boundary conditions taken from the HadCM3 GCM using the SRES A1B emission scenario (rapid, regionally-convergent growth with a balance of fossil and non-fossil fuels). Here, only the outputs of the unperturbed version of the HadRM3.0 model for the period 1961-2000 were used.

Previous studies have shown that climate models often simulate precipitation time series in which the frequency of wet days with low precipitation is higher than observed (the so-called drizzle effect, e.g., Sun et al., 2006). This means that the sequencing of wet and dry days, which is vital for the generation of hydrological extremes (both floods and droughts), is not well reproduced. To reduce this issue, the RCM daily precipitation outputs were modified so that the monthly frequency of rain days matched that of the observed record. As observations were available at 1 km, and because a fine spatial distribution of precipitation is important in the accurate representation of runoff-generating processes in hydrological models, the procedure was applied for each 1 km grid cell by setting the precipitation to zero for any day with precipitation lower than a threshold calculated so that the monthly frequency of wet day becomes identical to that of the observed time series for this grid cell. In the following, ‘simulated data’ refers to the 1 km RCM-driven time series after wet frequency correction and not to the original RCM precipitation outputs. The wet-day correction is regarded as essential for most hydrological applications (Weedon et al., 2011) and so it is used in association with all of the bias correction methods described here.

5. Results for the Exe-Culm catchment

In total, seven bias-correction methods were compared, each belonging to one of the four families of methods described earlier. For the empirical distribution correction approach, four different methods were considered where the transfer functions were defined using 25, 50, 75 and 100 quantiles. All methods were constructed monthly for each grid cell of the catchments. Results are presented at the catchment level (i.e. average performance within a catchment) and from 1 km maps of ARD in the four first statistical moments.

First, overall performance of each model is calculated on the calibration period 1961–1990, aiming to investigate how well high moments can be reproduced. Second, the robustness of each model is tested using a cross-validation methodology on the period 1961–2000, to evaluate how well the models perform outside their calibration period. Calibration periods of 30 years are used in both steps as this time-scale is considered by the World Meteorological Organization (WMO, 1983) as being a good compromise for capturing natural variability.

5.1 Overall performance

For each grid cell, ARD associated with the four first moments are calculated and corresponding catchment average ARD derived. The ARD, being calculated over a period of 30 years, will here be referred to as "ARD30". As a baseline measure, ARD30 from the RCM-driven precipitation time series are also calculated for all considered statistics (Figure 3). Note that, although linear and non-linear methods force the first moment (the mean) of the corrected distribution to be equal to that of the observed distribution, the resulting impact on the other moments is not known *a priori*.

For all methods, ARD30 is lower or equal after correction (Figure 3): biases in the daily precipitation distribution are effectively reduced by the bias-correction methods. While, by design, the bias-corrected mean (first moment) is equal to the observed when using the linear and non-linear methods (ARD30 equal to 0), this is not the case for the distribution-based methods. This is because these methods aim to reproduce as well as possible all statistical moments of the distribution and not only the first. Consequently, the ARD30 values for all moments are reduced but none is completely removed. The non-linear method is the only one which reproduces the standard deviation precisely, as it is the only method which imposes this constraint. The ARD30 on standard deviation is slightly reduced using the linear method, but the error remains large, suggesting that the linear method cannot entirely correct bias in precipitation variability. Amongst the distribution-based methods, the gamma distribution is associated with the lowest ARD30 on standard deviation while ARD30 reduction for the empirical distribution method increases with the number of quantiles used. ARD30 values for the third (skewness) and fourth (kurtosis) moments are not reduced by the linear method, but are reduced by about one fifth by the non-linear method and halved by the gamma distribution method. Finally, the empirical distribution method provides the most consistent reduction of ARD30 across all the four moments (reduction on average of 0.014 in ARD30 for every increase of 25 quantiles), likely to be a result of the amount of information used for its calibration: the lowest ARD30 is achieved with 100 quantiles, but as the number of quantiles employed falls, the empirical distribution method becomes less effective. Note that when calibrated using 25 quantiles, the empirical distribution method results in ARD30 smaller or equal to those obtained from the linear method for all moments considered here, except the mean. The greatest errors in uncorrected simulated precipitation are found in between May and July when rainfall is mainly convective. Monthly variation evens out over all moments using the gamma distribution and is reduced, even if only slightly, using the other correction methods. However, the empirical distribution method seems to consistently induce an error peak in July. In view of the complexity of the empirical distribution method (here using a minimum of 25 parameters), model performance is arguably best with the two-parameter gamma distribution method, which has smaller ARD30 than the non-linear method (which also uses two parameters) and an overall acceptable performance.

5.2 Robustness

Using the cross-validation methodology presented above, the seven bias-correction methods (linear, non-linear, gamma, empirical with 25, 50, 75 and 100 quantiles) were calibrated by constructing daily artificial time series generated from the observed data by removing ten consecutive years of data in turn. In total, 31 samples representative of the observations, each of 30 years duration, were created. The robustness of each of the seven bias-correction techniques was evaluated by calculating ARD between the bias-corrected precipitation time series and the 10-year observations removed from the calibration sample. The resulting ARD, being calculated over a 10-year period, will be referred to as "ARD10".

5.2.1. Sample spread

The spread of the ARD10 partly reflects the uncertainty associated with each bias-correction method. The seasonal means of the 31 ARD10 obtained from the cross-validation procedure are plotted on each 1 km grid (Figures 4–8, depending on the statistic considered for the evaluation) to investigate whether spatial and seasonal patterns emerge, where each method is described in a separate row. The first row (SIM) illustrates the difference between the uncorrected simulated precipitation distribution and the observed precipitation distribution. Catchment mean ARD10 is shown in the top right of each plot. Small mean ARD10 indicate a good reproduction of the statistical properties of the observed daily precipitation distribution by the bias-corrected time series, and suggests robust model performance outside its calibration period. In contrast, large errors suggest that the bias-correction procedure is sensitive to the choice of calibration period. This suggests that caution must be taken if the bias-correction procedure is applied to a time period different from that which was used for calibration, as this is the case for example when using the correction procedure for future periods.

Three conclusions emerge from our study of the first moment of the distribution (Figure 4): (i) errors remaining after bias correction are smaller than those present in the uncorrected RCM data; (ii), for all bias-correction procedures, the correction of the mean of the precipitation distribution is relatively insensitive to the choice of calibration time-period, although this result is sensitive to the change of season, with spring and summer exhibiting more variability than other seasons. This could be a consequence of the occurrence of convective precipitation in this season, which has little spatial or temporal structure, and is therefore very difficult to reproduce using any bias-correction method; and (iii) applying a bias-correction technique using observed data at a finer scale than that of the RCM (i.e., 1 km instead of 0.22º) introduces a detailed spatial distribution of precipitation within the catchment. This can be advantageous in spatially-distributed hydrological applications if the spatial pattern of precipitation plays an important role in runoff generation. However, such an approach assumes that rainfall values at 1 km can be compared with RCM estimates at a coarser resolution and ignores for the sake of simplicity the possibility that wide-area rainfall averages are systematically different from finer-resolution rainfall estimates (Rodriguez-Iturbe and Mejía, 1974).

Figures 5 and 6 show the seasonal and spatial patterns of mean ARD10 in the standard deviation and coefficient of variation. While mean ARD10 in the standard deviation is largely reduced by most models, the spatial patterns of mean ARD10 vary between seasons, and between methods particularly in the summer. Again, this result is likely due to the strong convective activity that generates much of the summer precipitation in southern England, which is both local and episodic, but treated differently by the different methods. The coefficient of variation (CV, defined as the ratio of the standard deviation to the mean) is a dimensionless measure of variability. For all methods, mean ARD10 in the CV is increased for all seasons, suggesting that these methods are not robust to the choice of the calibration period when correcting the CV. Note however that mean ARD10 in the CV reduces in spring when bias correction is done using the empirical distribution method. The standard deviation and CV show a large difference in robustness, which could be attributed to the combination of the corrections of both the average and the standard deviation, as the CV is the result of dividing the latter by the former.

The impact of the calibration period on the results is highest for ARD10 in the third and fourth moments (Figures 7 and 8 respectively) where all methods show an increased catchment mean ARD10 (top-right corner of the maps). The non-linear method shows the greatest increase in catchment ARD10, while ARD10 is smaller for the gamma distribution and linear correction methods than for the empirical distribution method. This suggests that, while the greatest accuracy is achieved by an empirical distribution method defined by at least 25 quantiles (i.e., the overall error from the same calibration-evaluation period is smallest), results are also most sensitive to the chosen calibration period. This is likely due to an over-fitting of the parameters of the correction method to a particular set of data, particularly noticeable in summer when precipitation values are more variable in this catchment.

5.2.2 Frequency of error reduction

Table 2 provides a quantitative assessment of the frequency with which the application of each of the bias-correction procedures actually resulted in improved precipitation statistics when evaluated against data from a time-period which was different from that over which the bias-correction procedures had been calibrated. For the mean, ARD10 is reduced between 86% (non-linear method) to 89% (empirical distribution method with 25 and 50 quantiles) of the time. For the standard deviation, the gamma distribution method achieves a reduction in ARD10 82% of the time while it is true only 76% and 77% of the time for the non-linear method and empirical distribution method with 25 and 50 quantiles. For the CV, this frequency drops to 35% for the gamma distribution method and 27% for the linear and non-linear methods, suggesting that the combination of errors in the mean and standard deviation is greater than their individual errors. For the higher moments, the frequency of error reduction further decreases to 11% (gamma distribution and linear methods) and 6% (empirical distribution method with 25 quantiles) for skewness, and to 11% (linear method and empirical distribution method with 75 and 100 quantiles) methods to 7% (non-linear method) for kurtosis.

We note from this analysis that no method gives a frequency of error reduction greater than 11% for the third and fourth moments (skewness and kurtosis), i.e., calibrated bias-correction methods succeed in generating daily time series with high-order moments more similar to the observed only slightly more often than one in ten times when applied to time periods not included in the calibration. Moreover, the gamma distribution method and empirical distribution method with 75 and 100 quantiles give the best chances for reducing the bias. Performance also varies with seasons, with an ARD10 reduced more often in spring and less often in summer. The overall frequency of error reduction suggests that the choice of correction technique must be made very carefully with an awareness of the additional uncertainties that may be introduced through the use of bias-correction techniques.

6 Results over Great Britain

The robustness of the methods for the six remaining catchments is assessed by considering how the performance of each correction method varies with location and climatic characteristics using the methodology described above.

6.1 Sample spread

The results of the robustness tests are shown in Table 3. The ARD, being calculated over a 10-years period, will here be referred to as "ARD10". It can be seen that all bias-correction methods in all catchments are effective in reducing ARD10 in at least the lower order statistical moments. The linear method consistently improves the average but rarely improves the higher order moments. The non-linear and gamma methods show similar performance in most catchments, with a reduction of errors achieved at higher order moments. However, both of these methods struggle to improve the higher order moments in the summer season (Newton et al., 2010), and in particular the non-linear method. This is most apparent in the East Anglian catchment where summer precipitation is dominated by convective storms. The empirical approach shows the best results for the higher order moments, however its performance can be erratic and can result in high ARD10 (mean ARD10 > 1) even in the lower order moments in all catchments.

The transfer functions calculated for each percentile of the distribution were analysed to understand the variability in performance of the empirical approach. In some months, in localised areas of the catchments, the transfer functions for some quantiles (generally the lower quantiles) were unrealistically high. These high values of the transfer function occur when, for some quantiles, the simulated precipitation is significantly lower than the observed precipitation. This low precipitation intensity, known as a weakness of climate models (Sun et al., 2006), leads to mid to low simulated precipitation quantile values which are significantly lower than those of the observed precipitation data. The bias correction procedure attempts to generate those higher precipitation values through very high transfer function values.

6.2 Frequency of error reduction

Table 4 provides a quantitative assessment of the frequency with which the application of each of the bias-correction procedures actually resulted in improved precipitation statistics when evaluated against data from a time-period which was different from that over which the bias-correction procedures had been calibrated. For the mean, ARD10 is reduced between 77% (empirical distribution method with 75 quantiles) to 84% (linear method) of the time. For the standard deviation, the gamma distribution method achieves a reduction in ARD10 79% of the time while it is true only 58% of the time for the empirical distribution method with 25 quantiles. For the CV, this frequency is 64% for the gamma distribution method and drops to 36% for the linear method. For the skewness and kurtosis, the highest frequency of error reduction is achieved by the linear method (61% and 62%, respectively), while the lowest frequency is obtained using the empirical distribution method with 50 quantiles (24% and 18%, respectively).

We note from this analysis that the frequencies of the third and fourth moments (skewness and kurtosis) do not drop as rapidly as over the Exe-Culm basin and that the gamma distribution method gives on average the best chance of correcting for the bias. Performance also varies with season, with ARD10 reduced more often in winter and less often in spring. The overall frequency of error reduction is slightly higher than that over the Exe-Culm basin, although this does not change the suggestion that the choice of correction technique must be made very carefully with an awareness of the additional uncertainties that may be introduced through the use of bias-correction techniques.

7 Discussion and conclusion

In this paper, we have compared four bias-correction techniques to determine which is the most effective and robust method to use when correcting daily precipitation simulated by a RCM for subsequent use in a hydrological model. These published techniques can be grouped into the following four families: (i) linear, (ii) non-linear, (iii) distribution-based quantile mapping and (iv) distribution-free quantile mapping. The robustness of a method has been tested using a cross-validation technique, similar to that of the jack-knife, carried out on seven catchments spread throughout Great Britain, each representative of different climate conditions.

The linear method showed the weakest correction as it is designed to alter only the mean, while the non-linear method corrects up to the second statistical moment of the frequency distribution. Using a gamma distribution affects all statistical moments, but this bias-correction technique uses only two parameters, and so its success is contingent on the underlying precipitation data being drawn from a gamma distribution. The most comprehensive correction was achieved by using the empirical quantile-mapping methods, which incorporate information from the frequency distributions of modelled and observed precipitation. The greater the number of quantile divisions used to represent the underlying frequency distributions, the better the correction.

However, the effectiveness of bias-correction was found to be sensitive to the time-period for which the bias-correction procedures have been calibrated. As the complexity of the bias-correction technique increases, the amount of information that it requires from the baseline observed climatological dataset also increases. It should therefore be expected that the accuracy of the bias-correction procedure will increase as more information from the observed record is taken into account. At the same time, the potential to over-calibrate the bias-correction procedure to a particular set of reference data increases as more and more observed data are used to calculate the correction parameters.

Taken together, our results indicate that there is a high probability of correcting RCM precipitation data up until the second moment, but the ability to correct higher moments falls quickly, owing to the problem of over-tuning the correction procedure to the particular observed dataset employed. Our results also support the assertion of Li et al. (2010) that the frequent extrapolation and interpolation that is required with empirical quantile mapping methods can lead to unsatisfactory results, particularly for extreme quantiles (Themeßl et al., 2010).

We conclude by stating that whilst the first and second moments of the precipitation distribution can be corrected robustly, none of the methods evaluated here could robustly correct the third and fourth moments of the precipitation frequency distribution. The empirical quantile mapping method with 100 quantile divisions was highly accurate but its sensitivity to the choice of time period was higher than that of methods which used fewer parameters and which were, therefore, less vulnerable to over-tuning. Overall, the correction method based on a gamma distribution offers the best combination of accuracy and robustness, but it is valid only when the observed and modelled precipitation data are gamma distributed. In circumstances where precipitation datasets cannot adequately be approximated using a gamma distribution, the linear and non-linear correction methods were most effective at reducing bias across all moments tested here (Table 3). The non-linear method is more effective at reducing the bias but the linear method is least sensitive to the choice of calibration period. However, it should be borne in mind that bias correction introduces additional uncertainties, due to the choice of bias correction method and the choice of calibration period, which are greater for higher-order moments (Table 4).

Acknowledgements

This work was funded by the UK Department for Food and Rural Affairs via the Environment Agency, and by the UK Natural Environment Research Council (NE/011969/1). We thank Alison Kay for advice on cross-validation methods.

References

Arnell, N., Hudson, D. A. & Jones, R. G. (2003) Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa. *Journal of Geophysical Research,* 108**,** 4519. DOI: 10.1029/2002JD002782.

Arnell, N., Liu, C., Conpagnucci, R., da Cuhna, L., Hanaki, K., Howe, C., Mailu, G., Shicklomanov, I., Stakhiv, E. (2001) Hydrology and Water Resources. IN Mccarthy, J. J., Canziani, O. F., Leary, N. A., Dokken, D. J., White, K. S. (Ed.) *IPCC Climate Change 2001: Impacts, Adaptation & Vulnerability, The Third Assessment Report of Working Group II of the Intergovernmental Panel on Climate Change (IPCC).* Cambridge, UK., Cambridge University Press.

Ashfaq, M., Bowling, L. C., Cherkauer, K., Pal, J. S. & Diffenbaugh, N. S. (2010) Influence of climate model biases and daily-scale temperature and precipitation events on hydrological impacts assessment: A case study of the United States. *Journal of Geophysical Research,* 115**,** D14116. DOI:10.1029/2009JD012965.

Bates, B. C., Kundzewicz, Z. W., Wu, S. & Palutikof, J. P. (2008) Climate Change and Water. Geneva, IPCC Secretariat.

Bissell, A. F. & Ferguson, R. A. (1975) The jackknife-toy, tool, or two-edged weapon? *Statistician,* 24**,** 79-100.

Bradford, R. B. & Marsh, T. J. (2003) Defining a network of benchmark catchments for the UK. *Water & Maritime Engineering,* 156**,** 109-116. DOI: 10.1680/wame.2003.156.2.109.

Buonomo, E., Jones, R., Huntingford, C. & Hannaford, J. (2007) On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations. *Quarterly Journal of the Royal Meteorological Society,* 133**,** 65-81. DOI: 10.1002/qj.13.

Diaz-Nieto, J. & Wilby, R. L. (2005) A comparison of statistical downscaling and climate change factor methods: impacts on low flows in the River Thames, United Kingdom. *Climatic Change,* 69**,** 245-268. DOI: 10.1007/s10584-005-1157-6.

Fatichi, S., Ivanov, V. Y. & Caporali, E. (2011) Simulation of future climate scenarios with a weather generator. *Advances in Water Resources,* 34**,** 448-467. DOI: 10.1016/j.advwatres.2010.12.013.

Fowler, H. J., Blenkinsop, S. & Tebaldi, C. (2007) Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology,* 27**,** 1547-1578. DOI: 10.1002/joc.1556.

Frei, C., Christensen, J. H., Déqué, M., Jacob, D., Jones, R. G. & Vidale, P. L. (2003) Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. *Journal of Geophysical Research,* 108**,** 4124. DOI:10.1029/2002JD002287.

Frei, C., Schöll, R., Fukutome, S., Schmidli, J. & Vidale, P. L. (2006) Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *Journal of Geophysical Research,* 111**,** D06105. DOI:10.1029/2005JD005965.

Hay, L. E., Clark, M. P., Wilby, R. L., Gutowski, W. J., Leavesley, G. H., Pan, Z., Arritt, R. W. & Takle, E. S. (2002) Use of regional climate model output for hydrolgic simulations. *Journal of Hydrometeorology,* 3**,** 571-590. DOI: 10.1175/1525-7541.

Hay, L. E., Wilby, R. L. & Leavesley, G. H. (2000) 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. DOI: 10.1111/j.1752-1688.2000.tb04276.x.
Jenkins, G. J., Murphy, J. M., Sexton, D. S., Lowe, J. A., Jones, P. & Kilsby, C. G. (2009) *UK Climate Projections: Briefing report,* Exeter, UK, Met Office Hadley Centre.
Jones, S. B. (1983) The Estimation of Catchment Average Point Rainfall Profiles (IH Report Number 87). *Report.* Wallingford, Institute of Hydrology.
Keller, V., Young, A. R., Morris, D. & Davies, H. (2006) Environment Agency R & D Project W6-101 Continuous Estimation of River Flows (CERF) Technical Report: Estimation of Precipitation Inputs. Wallingford, Centre for Ecology and Hydrology.
Kendon, E. J., Rowell, D. P., Jones, R. G. & Buonomo, E. (2008) Robustness of Future Changes in Local Precipitation Extremes. *Journal of Climate,* 21**,** 4280-4297. DOI: 10.1175/2008JCLI2082.1.
Kilsby, C. G., Jones, P. D., Burton, A., Ford, A. C., Fowler, H. J., Harpman, C., James, P., Smith, A. & Wilby, R. L. (2007) A daily weather generator for use in climate change studies. *Environmental Modelling & Software,* 22**,** 1705-1719. DOI: 10.1016/j.envsoft.2007.02.005.
Leander, R. & Buishand, T. A. (2007) Resampling of regional climate model output for the simulation of extreme river flows. *Journal of Hydrology,* 332**,** 487-496. DOI:10.1016/j.jhydrol.2006.08.006.
Lenderink, G., Buishand, A. & van Deursen, W. (2007) Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. *Hydrology and Earth System Sciences,* 11**,** 1145-1159. DOI:10.5194/hess-11-1145-2007.
Li, H., Sheffield, J. & Wood, E. F. (2010) Bias correction of monthly precipitation and temperature fields from Intergovernmental Panel on Climate Change AR4 models using equidistant quantile matching. *Journal of Geophysical Research,* 115**,** D10101. DOI: 10.1029/2009jd012882.
Marsh, T. J. & Hannaford, J. (2008) UK Hydrometric Register (Hydrological Data UK Series). Wallingford, Centre for Ecology and Hydrology.
Newton, G., Dadson, S., Lafon, T. & Prudhomme, C. (2010) Precipitation Downscaling and Bias-Correction (SC090016/PN3). Wallingford, UK, Centre for Ecology and Hydrology.
Piani, C., Haerter, J. O. & Coppola, E. (2010) Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology,* 99**,** 187-192. DOI: 10.1007/s00704-009-0134-9.
Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. (1986) *Numerical Recipes in FORTRAN: The Art of Scientific Computing,* Cambridge, CUP.
Rodriguez-Iturbe, I. & Mejía, J. M. (1974) On the transformation of point rainfall to areal rainfall. *Water Resources Research,* 10**,** 729-735. DOI: 10.1029/WR010i004p00729.
Shabalova, M. V., van Deursen, W. & Buishand, T. A. (2003) Assessing future discharge of the river Rhine using regional climate model integrations and a hydrological model. *Climate Research,* 23**,** 233-246. DOI: 10.3354/cr023233.
Sun, Y., Solomon, S., Dai, A. & Portmann, R. W. (2006) How often does it rain? *Journal of Climate,* 19**,** 916-934. DOI: 10.1175/JCLI3672.1.
Themeßl, M. J., Gobiet, A. & Heinrich, G. (2010) Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology,* 31**,** 1530–1544. DOI: 10.1002/joc.2168.
Vidal, J. P. & Wade, S. D. (2008a) A framework for developing high-resolution multi-model climate projections: 21st century scenarios for the UK. *International Journal of Climatology,* 28**,** 843-858. DOI: 10.1002/joc.1593.
Vidal, J. P. & Wade, S. D. (2008b) Multimodel projections of catchment-scale precipitation regime. *Journal of Hydrology,* 353**,** 143-158. DOI: 10.1016/j.jhydrol.2008.02.003.
Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Osterle, H., Adam, J. C., Bellouin, N., Boucher, O. & Best, M. (2011) Creation of the WATCH Forcing Data and its

use to assess global and regional reference crop evaporation over land during the twentieth century. *Journal of Hydrometeorology,* 12**,** 823-848. DOI: 10.1175/2011JHM1369.1.

WMO (1983) Guide to Climatological Practices (WMO-No. 100). Geneva, Switzerland, World Meteorological Organization.

Wood, A. W., Leung, L. R., Sridhar, V. & Lettenmaier, D. P. (2004) Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change,* 62**,** 189-216. DOI: 10.1023/B:CLIM.0000013685.99609.9e.

Wood, A. W., Maurer, E. P., Kumar, A. & Lettenmaier, D. P. (2002) Long-range experimental hydrologic forecasting for the eastern United States. *Journal of Geophysical Research,* 107**,** 4429. DOI: 10.1029/2001JD000659.

Table 1. Information about the seven test regions used in the study. NRFA catchment ID (followed by a B if the catchment is in the benchmark network), mean annual rainfall and elevation range all come from Marsh and Hannaford (2008). The easting and northing distances are the number of 1 km grid squares required in each direction to encompass the region. The type indicates whether the region is delimited by a catchment outline (C) or a rectangle (R).

Name	NRFA ID	Mean Annual Rainfall (mm/yr)	Elevation Range (m)	Easting Distance (km)	Northing Distance (km)	Location	Type
Exe-Culm	45001	1295	488	56	53	South West England	C
	45003	985	249				
Northern Scotland	95001 B	2201	1409	130	20	Northern Scotland	R
Tay	15006 B	1461	1184	147	75	Eastern Scotland	C
Ribble	71001 B	1345	678	60	70	North West England	C
Conway	66011	2183	1028	39	30	North Wales	C
Severn	54001	924	809	115	76	Midlands of England	C
East Anglia	33019 B	641	60	61	69	South East England	R
	36008	609	92				

Table 2. Frequency of error reduction, in percent, for four statistical moments over the whole year and for each season over all four statistical moments of each of the frequency distributions of observed, simulated and corrected daily precipitation datasets over the whole of the Exe-Culm basin, where the probability is calculated out of the 31 10-year slices of the cross-validation study. See description of Figure 3 for table acronyms.

Moment	**Frequency of error reduction (%)**						
	Lin	Nln	Gam	Emp			
				025Q	050Q	075Q	100Q
Mean	88	86	86	89	89	88	88
Standard deviation	79	77	82	76	77	79	80
Coefficient of variation	27	27	35	30	33	34	33
Skewness	11	7	11	6	7	8	8
Kurtosis	11	7	9	7	10	11	11
Season							
Winter	50	45	49	48	49	49	49
Spring	56	50	57	57	59	59	59
Summer	29	24	27	20	22	26	26
Autumn	38	44	45	41	42	43	43

Table 3. Catchment mean absolute relative differences (ARD10) in mm day^{-1} for four statistical moments over the whole year between observed, simulated and bias-corrected daily time series resulting from the 10 years cross-validation methodology. Mean ARD10 > 1 corresponds to large differences ranging from 3.60 to 1.82 x 10^6. See description of Figure 3 for table acronyms.

Moment	**Catchment**	**Mean absolute relative difference (mm/day)**							
		Sim	Lin	Nln	Gam	Emp			
						025Q	050Q	075Q	100Q
Mean	*N. Scotland*	0.26	0.04	0.04	0.04	> 1	> 1	> 1	> 1
	Tay	0.26	0.02	0.02	0.02	0.04	0.09	> 1	0.05
	Ribble	0.18	0.02	0.03	0.03	0.03	0.03	> 1	0.02
	Conway	0.28	0.03	0.04	0.04	0.04	0.04	0.64	0.03
	Severn	0.17	0.03	0.03	0.03	0.03	0.03	> 1	0.05
	East Anglia	0.24	0.01	0.02	0.02	0.03	0.08	0.02	0.11
		0.23	0.03	0.03	0.03	> 1	> 1	> 1	> 1
Standard deviation	*N. Scotland*	0.20	0.18	0.04	0.04	> 1	> 1	> 1	> 1
	Tay	0.19	0.13	0.03	0.03	0.07	0.52	> 1	0.21
	Ribble	0.17	0.10	0.02	0.02	0.06	0.05	> 1	0.04
	Conway	0.24	0.12	0.04	0.04	0.07	0.05	> 1	0.04
	Severn	0.12	0.08	0.03	0.03	0.07	0.05	> 1	0.20
	East Anglia	0.13	0.12	0.03	0.02	0.07	0.37	0.05	0.61
		0.18	0.12	0.03	0.03	> 1	> 1	> 1	> 1
Coefficient of variation	*N. Scotland*	0.16	0.16	0.02	0.02	0.07	0.09	0.13	0.12
	Tay	0.14	0.13	0.02	0.03	0.03	0.02	0.02	0.02
	Ribble	0.10	0.12	0.01	0.03	0.03	0.02	0.02	0.02
	Conway	0.12	0.14	0.01	0.05	0.03	0.02	0.01	0.01
	Severn	0.09	0.09	0.02	0.03	0.04	0.03	0.02	0.02
	East Anglia	0.09	0.13	0.02	0.03	0.04	0.03	0.03	0.03
		0.12	0.13	0.02	0.03	0.04	0.04	0.04	0.04
Skewness	*N. Scotland*	0.14	0.14	0.15	0.11	0.09	0.13	0.18	0.18
	Tay	0.17	0.14	0.10	0.07	0.07	0.07	0.06	0.06
	Ribble	0.15	0.15	0.08	0.06	0.06	0.07	0.07	0.07
	Conway	0.15	0.15	0.09	0.09	0.05	0.06	0.06	0.06
	Severn	0.10	0.09	0.09	0.07	0.07	0.07	0.07	0.07
	East Anglia	0.10	0.10	0.20	0.12	0.05	0.07	0.08	0.08
		0.14	0.13	0.12	0.09	0.07	0.08	0.09	0.09
Kurtosis	*N. Scotland*	0.14	0.14	0.14	0.11	0.09	0.13	0.16	0.16
	Tay	0.17	0.14	0.09	0.07	0.08	0.08	0.07	0.07
	Ribble	0.16	0.16	0.07	0.07	0.07	0.07	0.06	0.06
	Conway	0.17	0.17	0.08	0.10	0.04	0.06	0.06	0.06
	Severn	0.11	0.10	0.09	0.07	0.07	0.08	0.07	0.07
	East Anglia	0.11	0.11	0.18	0.10	0.05	0.06	0.07	0.07
		0.14	0.14	0.11	0.09	0.07	0.08	0.08	0.08

Table 4. Frequency of error reduction, in percent, for four statistical moments over the whole year and for each season over all four statistical moments of each of the frequency distributions of observed, simulated and corrected daily precipitation datasets averaged over the whole of the 6 catchments, where the probability is calculated out of the 31 10-year slices of the cross-validation study. See description of Figure 3 for table acronyms.

Moment	**Frequency of error reduction (%)**						
	Lin	Nln	Gam	Emp			
				025Q	050Q	075Q	100Q
Mean	84	82	83	82	82	77	79
Standard deviation	69	72	79	58	62	64	64
Coefficient of variation	36	59	64	53	57	60	60
Skewness	61	42	60	25	24	26	26
Kurtosis	62	36	53	24	18	19	20
Season							
Winter	72	75	82	57	56	55	57
Spring	54	48	58	44	46	48	49
Summer	56	45	59	45	48	51	51
Autumn	68	65	70	48	45	43	43

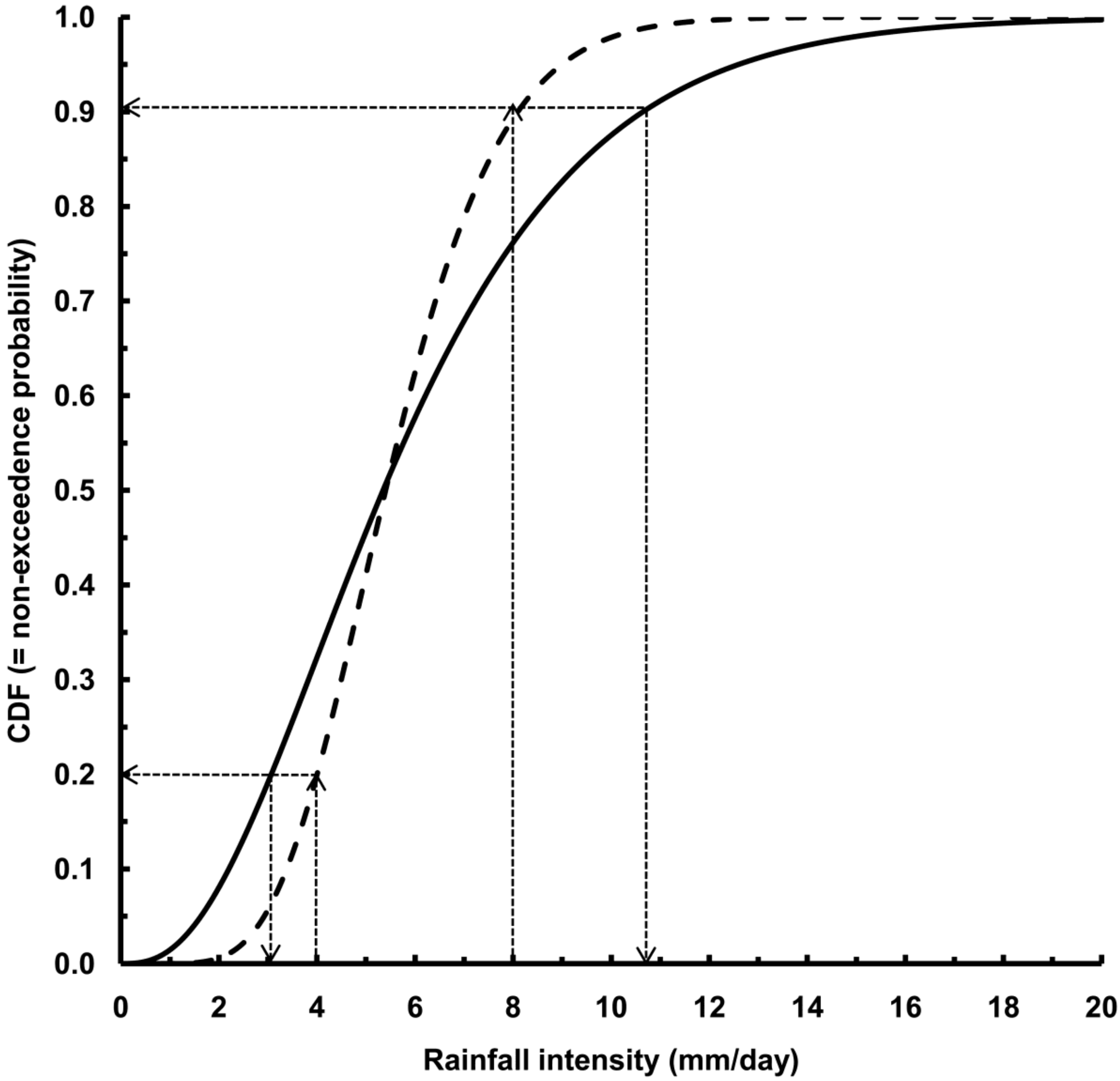

Figure 1. Schematic of the empirical distribution correction approach with observed (solid) and simulated (dashed) cumulative density function of daily precipitation and the transfer function (dotted lines) used to correct the simulated precipitation intensity. The corrected intensity of a simulated precipitation of a given quantile is found by resampling from the observed distribution the precipitation intensity with the same quantile value.

Figure 2. Location of the seven catchments in the UK. For the rectangular regions the catchments from which the statistics in Table 1 derive are shown inside the region.

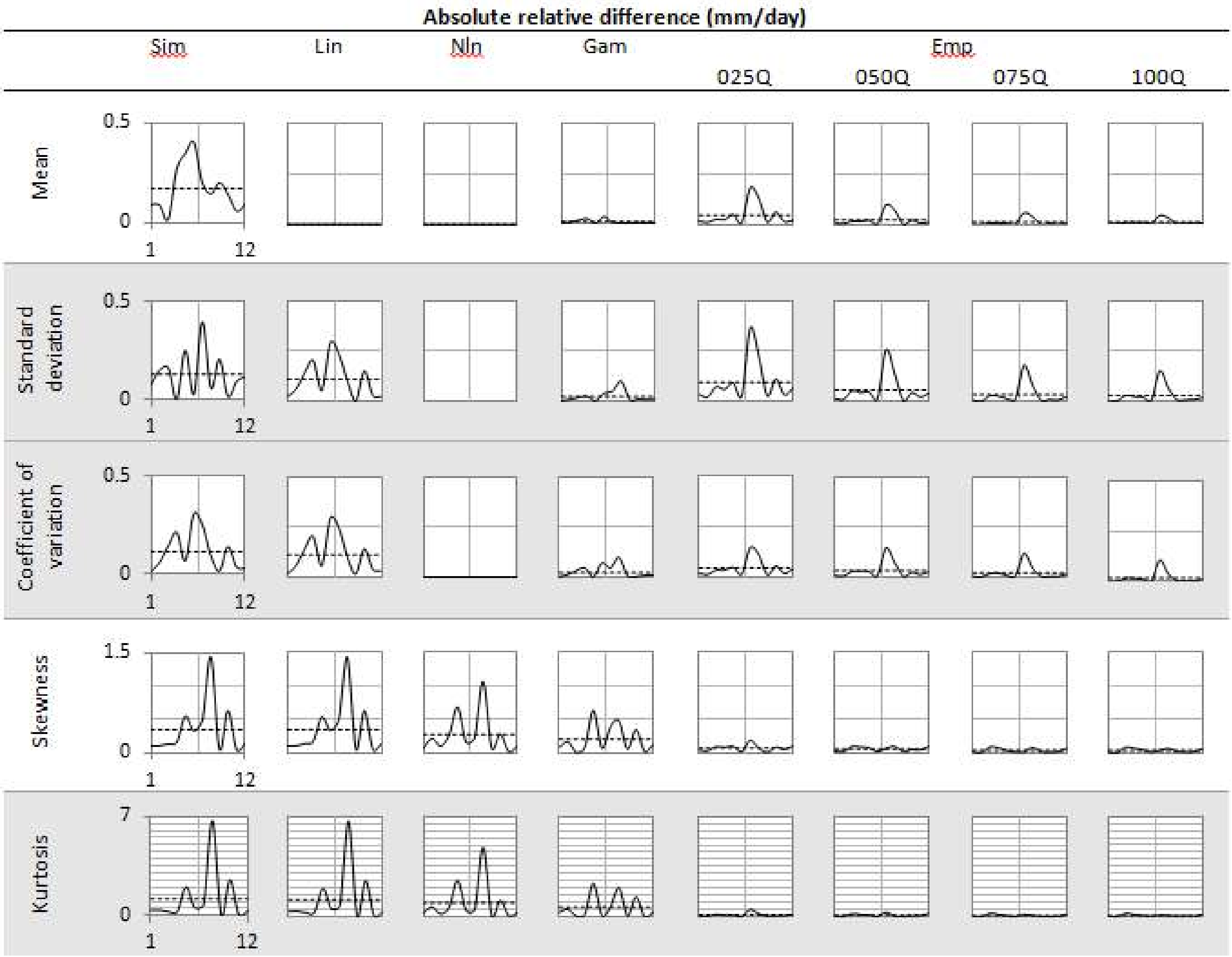

Figure 3. Monthly catchment average absolute relative differences (ARD30) in mm day^{-1} for four statistical moments of each of the frequency distributions of simulated (Sim) and corrected (Lin for linear, Nln for non-linear, Gam for gamma and Emp 025Q, 050Q, 075Q and 100Q for empirical with 25, 50, 75 and 100 quantiles, respectively) daily precipitation datasets over the whole of the Exe-Culm basin, where correction factors are calculated over 1961–1990 and applied to the same period. Monthly values of the catchment average ARD30 are shown by the solid line, while the dotted line illustrates the annual mean value of the catchment average ARD. The Y-axis and X-axis represent the absolute relative difference and corresponding month, respectively.

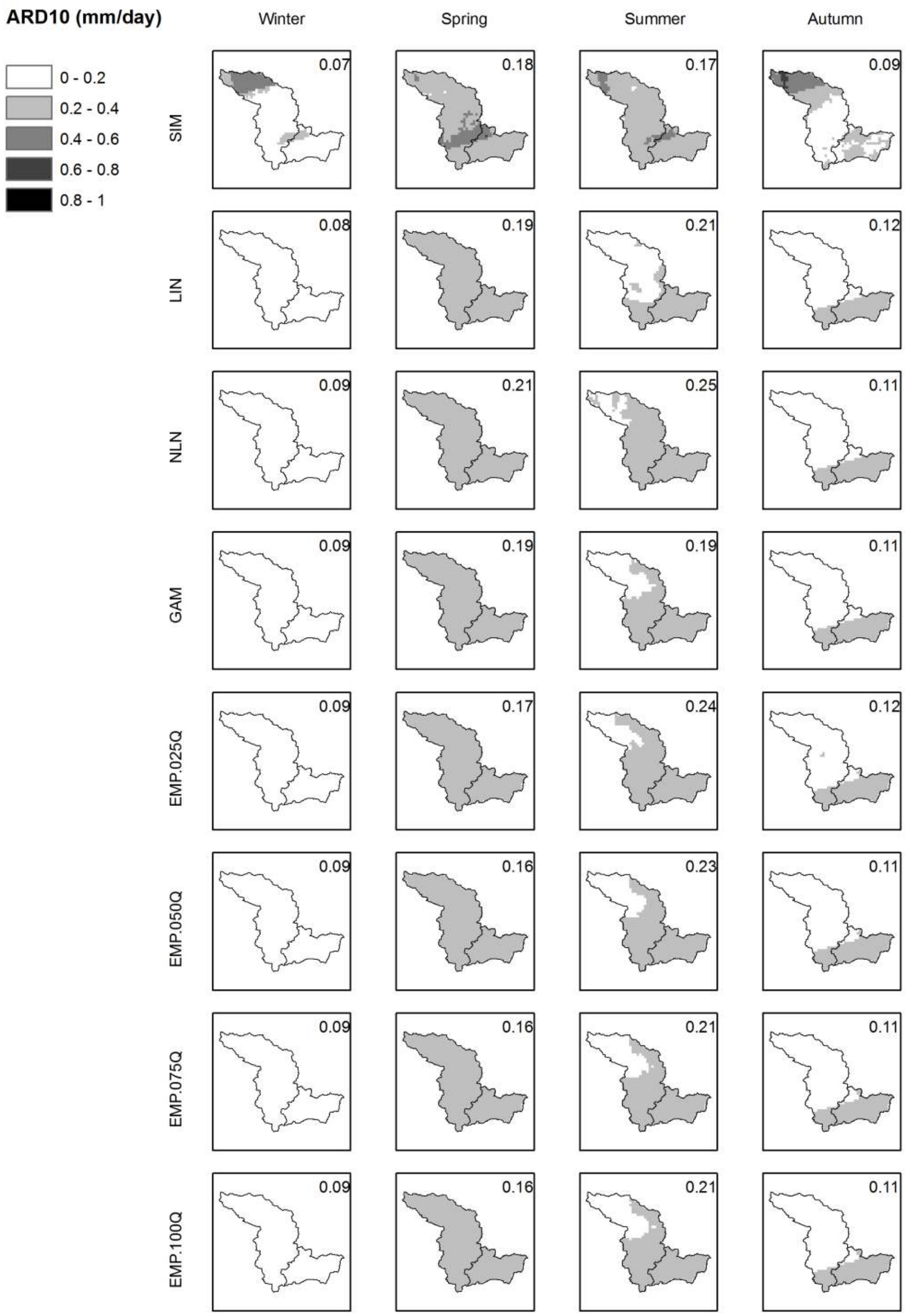

Figure 4. Seasonal absolute relative differences (ARD10; in mm/day) in the mean between observed, simulated (SIM) and each bias-corrected data obtained from the 10 years cross-validation procedure (top-right corner: catchment average), for the Exe-Culm river basin. Bias correction methods are linear (LIN), non-linear (NLN), gamma distribution (GAM), empirical distribution with 25, 50, 75 and 100 quantiles (EMP.025Q, EMP.050Q, EMP.075Q and EMP.100Q respectively).

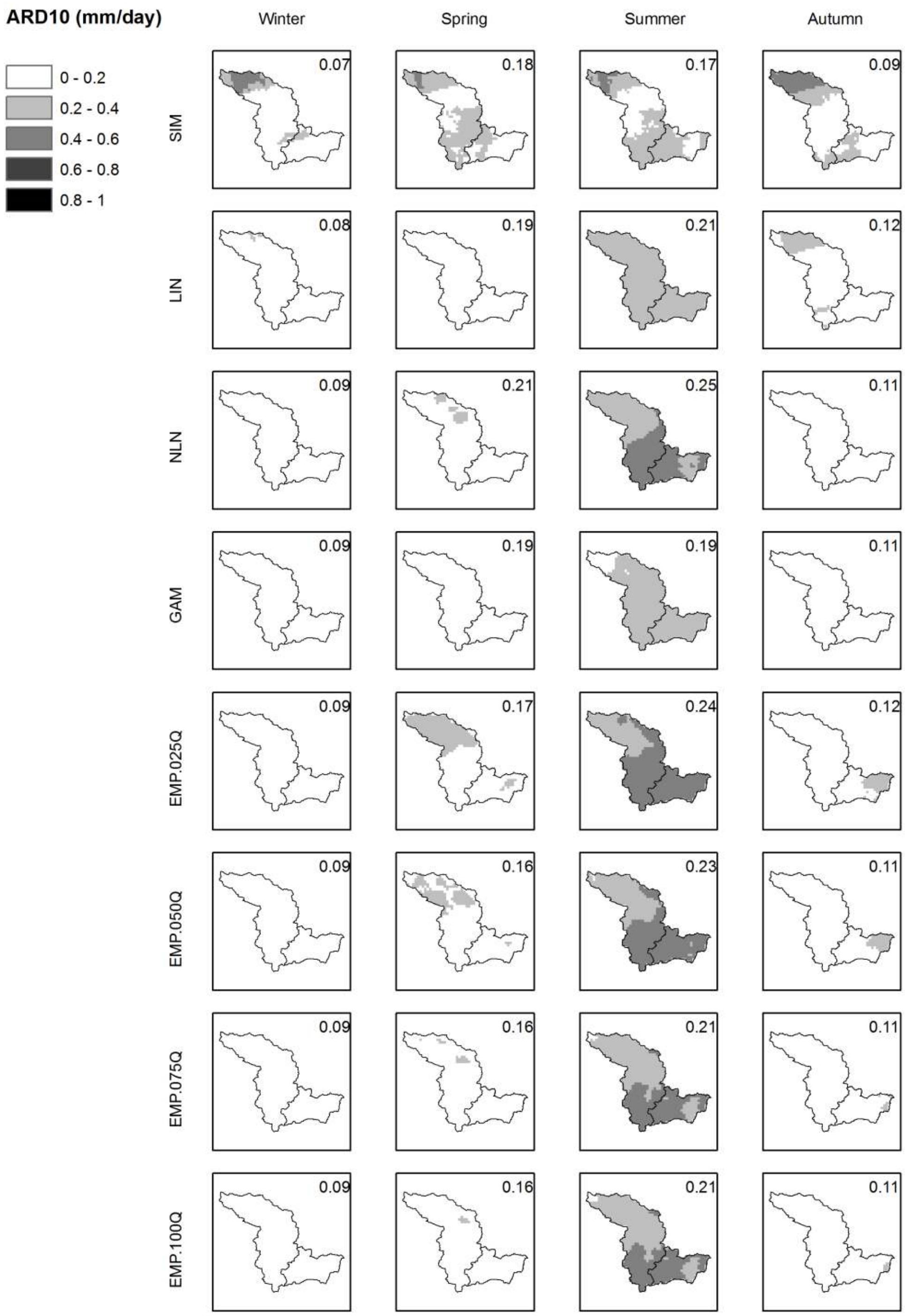

Figure 5. Same as Figure 4 for the standard deviation.

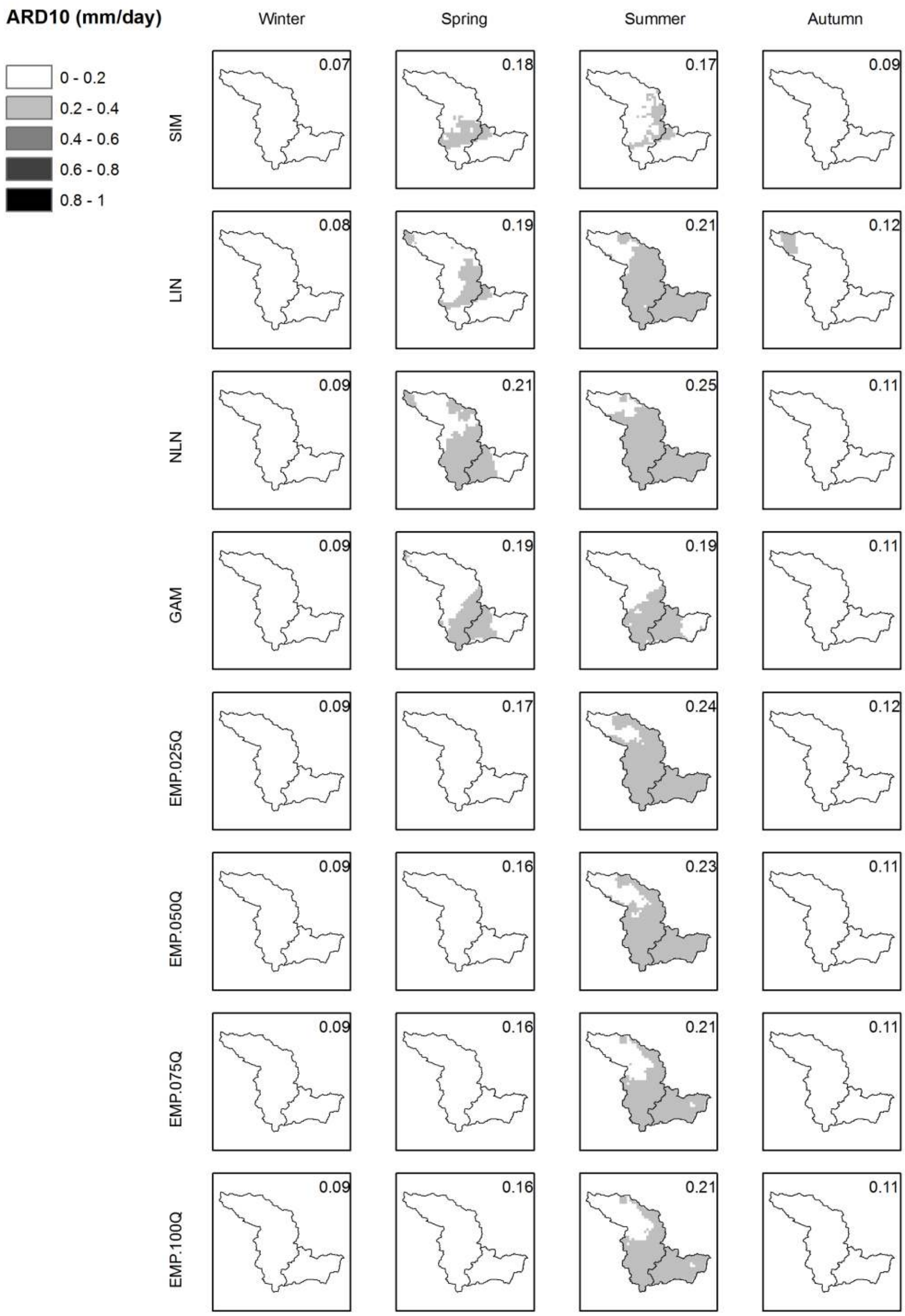

Figure 6. Same as Figure 4 for the coefficient of variation.

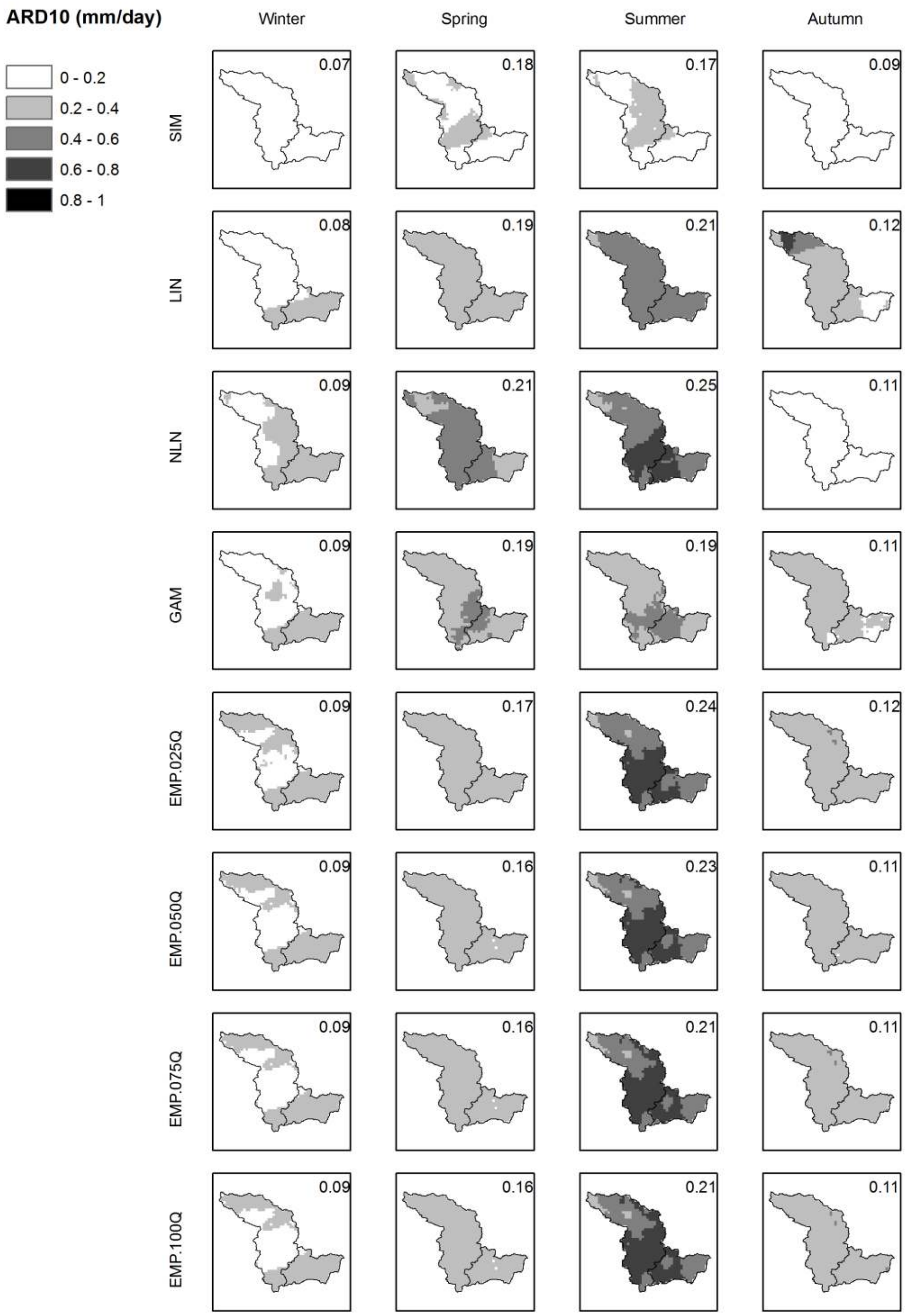

Figure 7. Same as Figure 4 for the skewness.

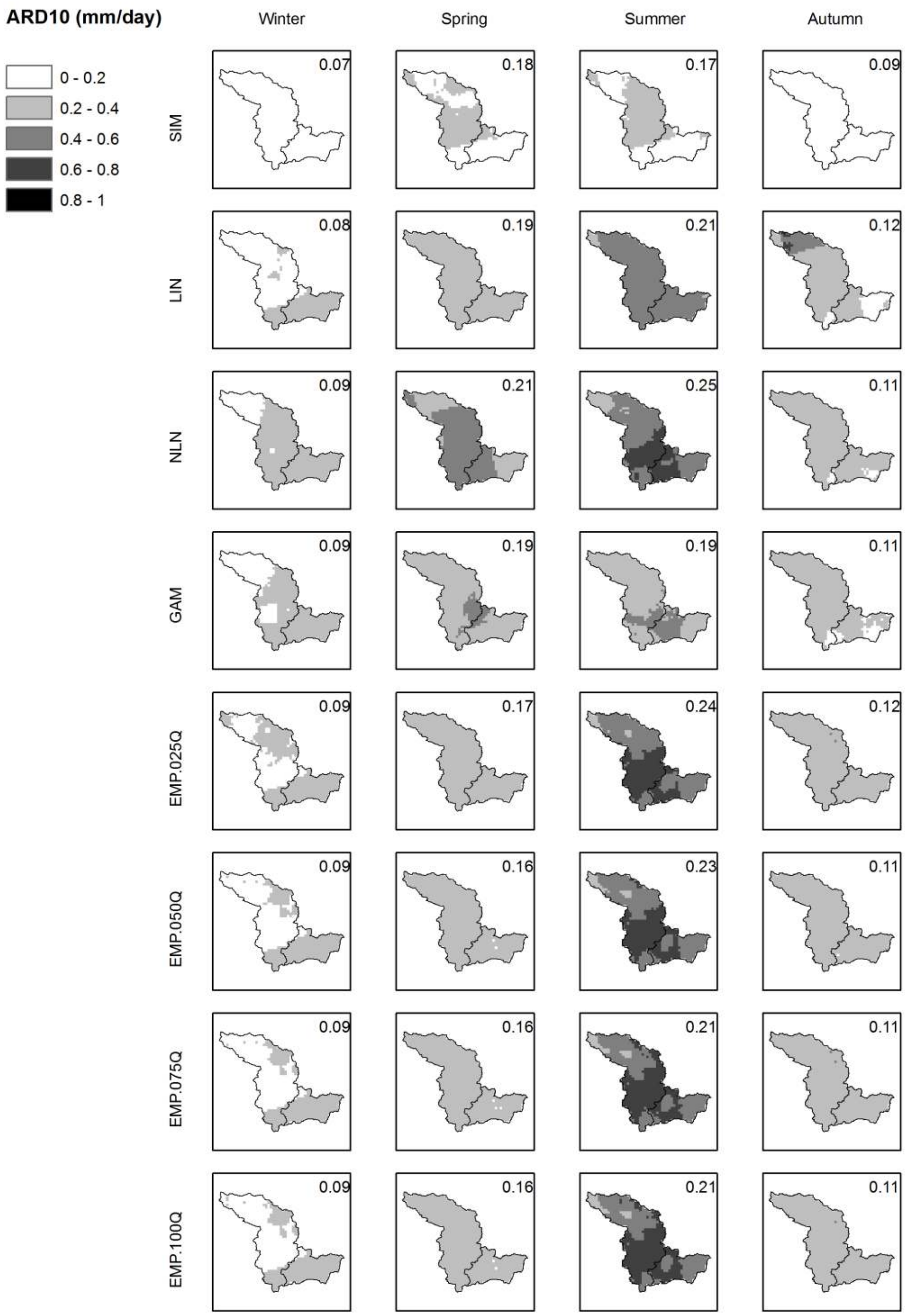

Figure 8. Same as Figure 4 for the kurtosis.