Department for Environment Food & Rural Affairs

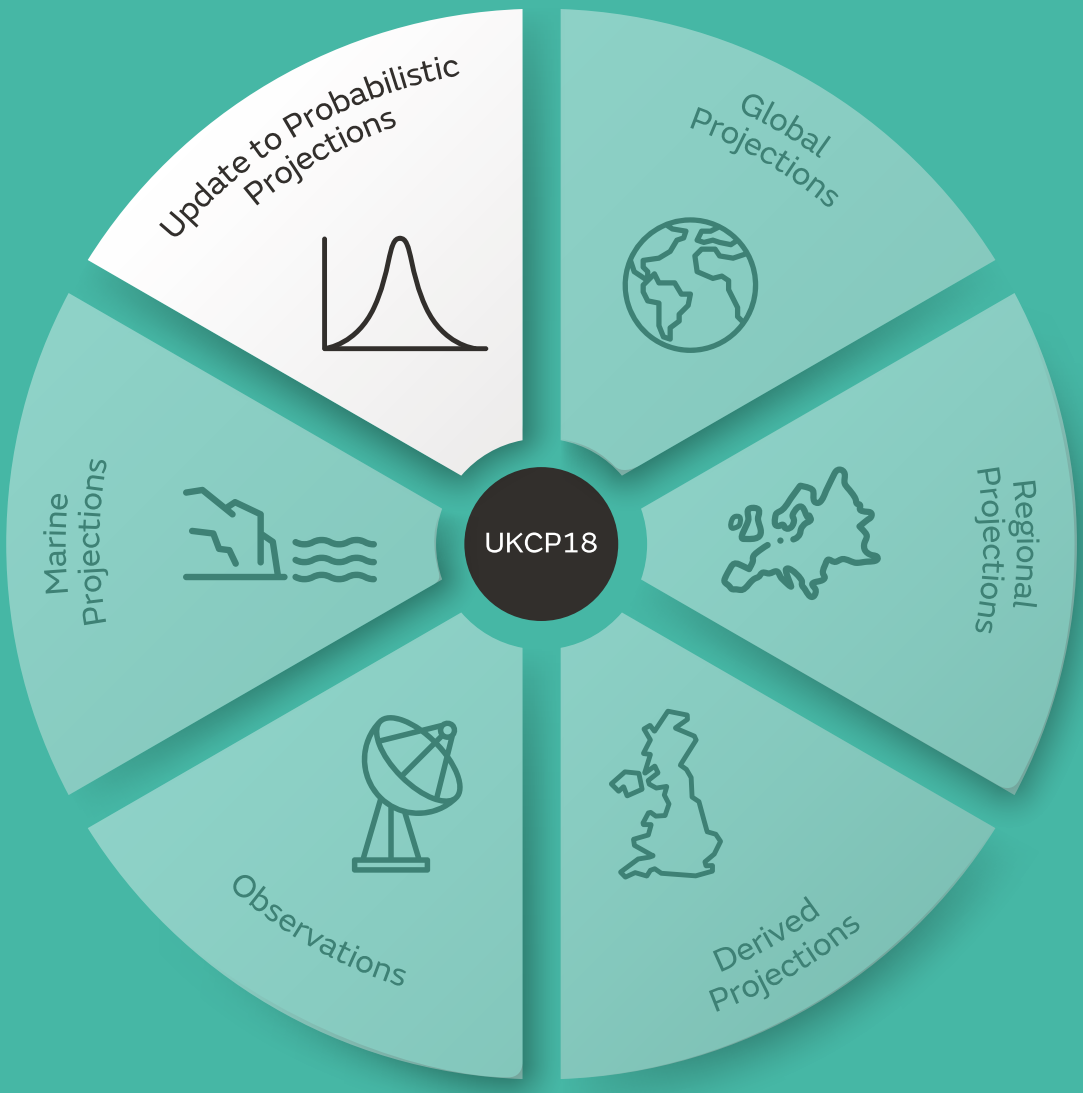
Department for Business, Energy & Industrial Strategy

Met Office Hadley Centre

Environment Agency

Update to UKCP probabilistic projections

July 2022



G.R. Harris, J.M. Murphy, J.S.R. Pirret, D.M.H. Sexton
Met Office Hadley Centre, Exeter, U.K.

Version 1.0

Contents

Summary

- The UKCP18 probabilistic projections consist of 3000 individual realizations of time-dependent climate change plus a set of probability distributions, derived from the realizations and presented as probability density functions (pdfs) and cumulative distribution functions (cdfs).
- These products are provided on a 25km national grid and for 43 aggregated regions. They have been updated to incorporate five developments. These improve consistency between the probabilistic projections and the climate model data on which they are based and make the projections easier to use.
- The results cited in this Summary refer mainly to projected changes for 2070-2089 relative to 1981-2000 under RCP85[1] emissions. This scenario and period is selected to illustrate the impacts of the developments most clearly, because it gives the strongest anthropogenic forcing and largest changes available in the probabilistic projections data. Where differences between the new and original projections are quoted as absolute numbers (e.g., in °C or percentage precipitation anomaly), users should note that smaller differences will be found in earlier 21st century periods or in the other four emissions scenarios for which pdfs are provided, since these correspond to lower greenhouse gas emissions and so have weaker forcing than RCP85.

What are the developments?

- Two of the developments fix software errors found during analysis of the original UKCP18 results.
- The first resolves an error in the downscaling section of the probabilistic calculations, strengthening and tightening relationships between global and regional climate model simulations used to add spatial detail to the projected changes. This fix applies to all the probabilistic projection variables. Correcting the downscaling error has its largest impacts when uncertainties in projected changes are large in comparison to the median signals of projected change. This is because the error affects the component of projected responses that varies between the climate model simulations used to determine the downscaling relationships, but not the component common to all the simulations.
- The second development fixes a software error affecting daily maximum and minimum surface temperature (Tmax and Tmin) results from one of the ensembles of global climate model simulations used in the projections. Correcting this error tightens relationships between projected changes in Tmax, Tmin and daily mean surface temperature (Tmean), improving physical consistency and substantially reducing the scatter found in the original UKCP18 results.
- A third development improves the treatment of extreme changes in precipitation. This involves merging the original version of the statistical calculations with an alternate version that uses a logarithmic transform to avoid statistical generation of unrealistic negative precipitation amounts. Merging improves the credibility of projected changes that are either extremely dry or extremely wet. In the new projections it is no longer necessary to truncate the 3000 realizations to remove instances of negative precipitation outcomes.

[1] In this report we cite results from the RCP2.6 and 8.5 emissions scenarios of Moss et al. (2010), two of the five scenarios for which pdfs are provided. These are referred to as RCP26 and RCP85.

- The fourth development removes biases in the centring of projected climate anomalies relative to the baseline period of 1981-2000. This development removes the potential need for users to recentre the 3000 realizations of time-dependent climate change themselves. The biases corrected by recentring are generally minor but can be larger in some individual realizations. While this development also affects the pdfs and cdfs, its effects on these are negligible.
- The fifth development improves the representation of climate variability in the annual probability distributions. This is done by pooling annual variability from the 3000 realizations over a longer time window, to achieve better sampling of alternative phases of variability in the annual pdfs. This reduces decadal undulations that were present in the time-evolving quantiles of the original UKCP18 pdfs, without affecting the underlying realizations. The pdfs of long-term average changes are unaffected by this update because these are derived from 20- and 30-year averages of the realizations, so further time-smoothing is not required.
- The first, second, third and fourth developments affect both the 3000 realizations and the pdfs and cdfs, while the fifth affects only the pdfs and cdfs of annual changes.

How do the developments affect the probabilistic projections?

- The original and new distributions are compared by assessing differences in future changes projected at the 10th (low, P10), 50th (median, P50) and 90th (high, P90) percentiles of the pdfs. Uncertainties are quantified as the P90 – P10 range, referred to as spread. The numbers cited below apply to changes for 2070-2089 relative to 1981-2000 under RCP85 emissions, unless stated otherwise.
- Additional maps and tables for Tmean and precipitation changes are provided in a companion document. See Harris et al. (2022), available from https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/guidance-science-reports. This extends the RCP85 information shown in this report to include maps for all four seasons and provides a set of maps for changes in 2040-2059 relative to 1980 under RCP45, a case with lower greenhouse gas emissions.
- Overall, the downscaling correction has the largest influence on the new probabilistic projections. Its main impacts apply to the 25km pdfs, leading typically to significantly increased spread in the new projections for precipitation (27%), sea-level pressure (55%), total cloud cover (28%), downward (35%) and net (32%) surface short-wave radiation and net surface long-wave radiation (25%). Figures in brackets give the fractional changes in spread between the new and original pdfs. The numbers are averaged across all spatial locations and all 17 period definitions (months, seasons and annual mean).
- These increases in spread arise because the downscaling correction strengthens the relationships between large-scale and local changes in the calculations (acting to increase spread), which offsets corresponding reductions in residual uncertainties (which act to reduce spread).
- In the new 25km pdfs spread typically reduces for Tmean (17% fractional change on average), Tmax (15%), Tmin (10%) and specific humidity (9%). The downscaling correction contributes to reduced spread for these variables, because reduced residual uncertainties outweigh strengthening of downscaling relationships. In the cases of Tmax and Tmin, correcting the Tmax/Tmin software error increases the reductions in spread.

- Differences in the median response are typically smaller. On average, the 25km-scale differences in P50 amount to 13% of the spread in the new projections for sea-level pressure, and to 6% or less for other variables.
- These differences vary with season and location. For example, the differences in P50 for Tmax are largest in July (amounting to 1.24°C on average in the 25km-scale pdfs, 18% of the average P90 – P10 spread), because correcting the Tmax/Tmin error has its largest impact in that month.
- For precipitation, the largest increases in spread occur in spring and autumn. In spring, the average P10 response changes from -9.4% to -20% in the new 25km projections, while the average P90 response increases from 13.7% to 21.3%.
- In summer, the new projections reinforce existing advice (including from other components of UKCP18) that strong drying is possible according to the P10 changes, especially over much of England and Wales. Compared to the original results, the improved treatment of precipitation extremes generally reduces the intensity of dry-end responses, while the downscaling correction increases the intensity. The net effects are modest reductions in levels of P10 drying over Wales and most parts of southern England, and modest increases in drying over most of Scotland and Northern Ireland. On average, changes in P10 amount to 4% in magnitude.
- For aggregated regions the downscaling correction has a smaller impact in general, although it does affect the new projections in cases where the selected large-scale predictor variable (taken from global climate model simulations) changes relative to the original calculations. On average, fractional differences in the median response are 6% or less for all variables. Average fractional differences in spread are largest for Tmax (12%) and Tmin (11%). Values are also close to 10% for total cloud cover and both surface short-wave radiation variables but well below 10% for other variables. These effects on spread are invariably smaller (on average) for aggregated-regions than for their 25km-scale counterparts, due to the smaller impact of the downscaling correction.
- There is generally more similarity between aggregated-region pdfs and co-located 25km pdfs in the new projections, since correcting the downscaling error removes a source of inconsistency that affected the original results.
- Five of the aggregated regions are national definitions for England, Scotland, Wales, Northern Ireland and the entire UK. For these, we assessed winter and summer responses for 2070-2089 relative to 1981-2000 in the RCP26 and RCP85 scenarios. For Tmean the new projections are mostly similar to the original results, in both scenarios. Larger differences do occur for Northern Ireland and Scotland in winter, particularly under RCP85. For Scotland, P50 decreases from 2.68°C to 2.13°C and P90 from 4.64°C to 4.02°C.
- In general, differences between the new and original precipitation pdfs of national changes are also small. Under RCP85 differences in specific percentiles amount to 4% or less in most cases, compared with P90 – P10 ranges of typically 40-50%. Larger differences occur for Northern Ireland in winter (where P50 increases from 15% to 21% and P90 from 34% to 43%), and for Wales and England in summer where P10 changes from -62% to -55% and from -61% to -54%, respectively.

Implications of the new probabilistic projections for users

- The extent to which the new projections affect existing user studies will depend on: (a) how large the differences are for the input data of interest; (b) how the pdfs are used to determine impacts, risks or decisions in specific studies.
- As a general guide, we suggest that conclusions:
  - Are *unlikely* to be affected if the differences in median (P50) response, or in the spread (P90 – P10 range) of responses are smaller than 10% of the spread in the new projections.
  - *May* be affected if either or both metrics lie in the range 10-20%.
  - Are *likely* to be affected if either or both metrics exceed 20%.
- Below we provide advice for the variables and region-definitions provided in the probabilistic projections. This is based mainly on average values of the fractional differences in median and spread between the new and original projections (see above), derived specifically from RCP85 results for 2070-2089 relative to 1981-2000.
- However, we also encourage users to compare the new and original results for their individual applications, since the fractional differences may vary according to the specific regions, emissions scenarios, seasons and periods of interest. Furthermore, some applications may be more sensitive to differences between the new and original projections measured in units of the relevant variable (e.g. as °C, Wm^{-2} or percentage precipitation anomaly), rather than as the fractional metrics defined above.
- In most applications that use the aggregated-region pdfs of Tmean and/or precipitation, we anticipate that the new results will not imply changes in decisions or conclusions obtained from the original UKCP18 results. Potential exceptions may occur in cases where larger differences occur, such as the national changes for winter in Scotland, mentioned above.
- Some applications using aggregated-region pdfs of Tmax and/or Tmin may be affected, since the average change in spread for these variables amounts to 12% for Tmax and 11% for Tmin. Changes in the median are generally below 10%, apart from July where the median response for RCP85 in the 2080s increases by 1.39°C in aggregated regions (20% of the spread) on average. Studies of monthly heat extremes in July are therefore likely to be influenced by the differences between the new and original projections.
- For other aggregated-region variables, average differences in P50 are well below 10%. This indicates that most studies sensitive to shifts in the median response are unlikely to be significantly affected, though there could be specific exceptions. In the cases of Tmin, total cloud cover and the surface short-wave radiation variables, average differences in spread are close to 10%, suggesting that some studies sensitive to changes in high- or low-end responses may be affected. For sea-level pressure, specific humidity and net surface long-wave radiation, average differences in spread are well below 10%, indicating that most existing studies are unlikely to be significantly affected.
- Regarding the 25km-scale projections, we expect that most decisions derived from the original pdfs of Tmean changes will be robust in the light of the new projections presented here, based on the average fractional differences in the median and spread of responses.

- For Tmax and Tmin some studies using the 25km-scale data may be affected. The differences between the new and original projections are qualitatively similar to those described above for aggregated-regions, noting that the average changes in spread (17% for Tmax and 16% for Tmin) are somewhat larger compared to their aggregated-region counterparts.
- It is likely that some applications using the 25km precipitation data will be significantly affected by the differences, particularly in cases where the low- and/or high-end changes play a key role in determination of impacts, decisions and risks. This is because the average difference in spread amounts to 27%, with increases occurring at most locations. For example, in winter the average P90 response increases from 40.5% to 46.6% in the new projections, though the difference varies substantially with spatial location.
- Applications using the 25km results for sea-level pressure, total cloud cover and surface radiation variables are also likely to be affected by the changes. This is due mainly to increased spread in the new projections, amounting to 20% or more (considerably larger than for the corresponding aggregated-region results).
- For sea-level pressure the average change in median response is 13%, suggesting that shifts in the pdfs may also contribute to the impacts. For other 25km-scale variables, average changes in P50 responses are well below 10%.
- We suggest also reading the accompanying Frequently Asked Questions (Fung et al, 2022) available from https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/guidance-science-reports. This provides general advice for users of the probabilistic data.
- This update does not affect other UKCP18 projection products.

1. Introduction

The probabilistic projections (pdfs) are a component of the land projections in UKCP18, providing the most comprehensive information on uncertainties in future climate (Murphy et al., 2018). A set of improvements has been developed for these projections. Two improvements arise from fixing software errors found by the UKCP18 science team following further analysis of the original UKCP18 results prompted by feedback from a user. The first error affected daily maximum (Tmax) and daily minimum (Tmin) surface air temperature. The second error affects the "downscaling" aspect of the probabilistic methodology (see Section 2.5), in which relationships between global and regional climate model simulations are used to derive UK projections expressed on a 25km grid and for three sets of aggregated regions (Fung et al., 2018).

For the revised projections, we have taken the opportunity to implement three additional improvements to the methodology. Two of these apply to all variables, improving the representation of climate variability on the decadal time scale and removing minor biases in the centring of projected anomalies relative to the historical baseline period. The third improves the representation of extreme outcomes in the precipitation pdfs.

The effects of fixing the Tmax/Tmin bug are described in Section 2.1, followed by the three methodological improvements in Sections 2.2-2.4. In these Sections, the impacts of the changes are assessed relative to the original UKCP18 projections. The effects of fixing the downscaling bug are described in Section 2.5. In this case, we isolate the specific effects of the downscaling fix by comparing the new projections (containing all five updates) with projections containing the other four updates with the downscaling fix withheld. This is followed in Section 2.6 by a summary of how the different updates affect each climate variable for which probabilistic projections are provided. Section 3 then describes the effects of these updates relative to the original UKCP18 pdfs, for a representative selection of UK variables. Appendix A presents further technical detail for these updates, while Appendix B provides a recap of the methodology used to produce the probabilistic projections, as context for interested readers.

2. Methodological Updates

The five updates are described in turn, in Sections 2.1-2.5 below. Section 2.6 provides a table summarising how each update influences the pdfs and realizations for the set of climate variables provided. In Sections 2.1-2.4 we combine the Tmax/Tmin bug fix and the improvements to precipitation extremes, climate variability and baselining to make a set of results referred to as "interim" probabilistic projections. These are compared with the UKCP18 results (termed the "original" projections), to illustrate the impacts of these four developments[2]. In Section 2.5 we compare projections containing all five developments (referred to as the "new" projections) against the interim projections that contain all developments except the downscaling fix. This allows us to isolate the effects of the latter, which (overall) has the largest impact on the new probabilistic projections.

2.1. Software Error affecting Tmin and Tmax

In raw climate model output, future changes in Tmean are normally close to the average of changes in Tmax and Tmin. The pdfs are constructed from 3000 statistical realizations of time-dependent climate anomalies that are designed to broaden the representation of uncertainties in the pdfs, by emulating the responses of a larger set of climate model simulations than it is feasible to run (Appendix B). While their time series are intended to mimic the characteristics of climate model output as closely as possible, the use of statistical emulation and sampling methods inevitably places limits on the level of spatial and inter-variable coherence in the realizations. However, a user reported instances of larger-than-expected differences between Tmax, Tmin and Tmean in the realizations, implying physical inconsistencies in their projections. This has been traced to an indexing error in a standard software package (Iris, 2010-2022) used to access Tmax and Tmin data from some of the global climate model (GCM) simulations used in the pdfs. The result was a one-month shift in the assignment of months to the variables. This meant, for example, that Tmax data for January was wrongly interpreted as applying to February in the software used to convert the climate model data into pdfs. This error, affecting only Tmax and Tmin, has now been fixed. The revised projections show much-improved coherence between Tmax, Tmin and Tmean.

As an example of the effects of incorrect indexing on the original projections, Figure 2.1a shows one projected realization of Tmin, Tmax and Tmean for the London grid-point. Differences between Tmean and the average of Tmin and Tmax should be small (Equation (1)).

$$Tdif = 0.5 * (Tmin + Tmax) - Tmean \quad (1)$$

In the GCM simulations underlying the probabilistic projections, such differences are typically in the range 0.1-0.2°C. However, this is not the case in Figure 2.1a and average differences of 2.5°C (black curve) are obtained by the end of the century, with considerably larger differences in some years. This indexing error clearly leads to loss of coherence and inconsistent projections for the temperature variables. Figure 2.1b shows a typical realization of Tmin, Tmax and Tmean when the error is corrected, with much reduced values for Tdif and much improved inter-variable coherence.

[2] The Figures in Sections 2.1-2.4 use results from aggregated regions. Although these are derived from GCM data using downscaling, the downscaling error does not affect aggregated region results directly. This is because the error occurred only in downscaling relationships calibrated using 25km data (see Section 2.5).

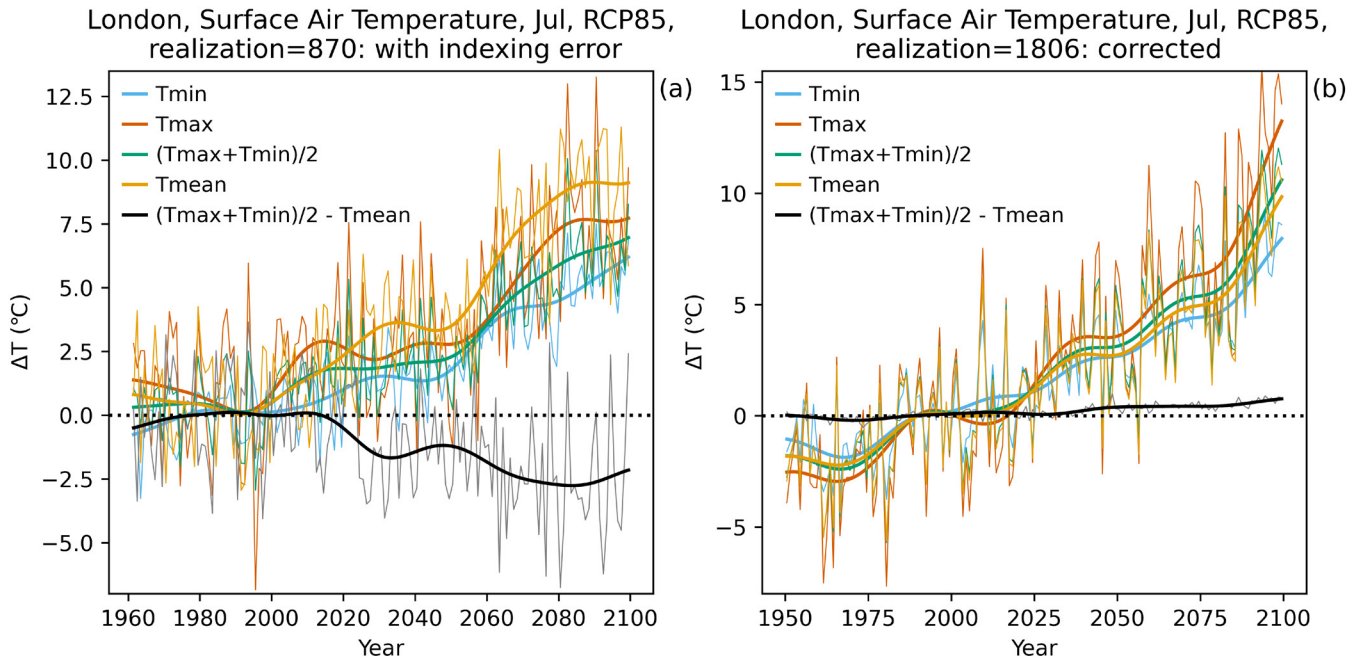


Figure 2.1. Examples of projected realizations for Tmin (blue curve), Tmax (red curve) and Tmean (orange curve) for the 25km resolution London grid-point in response to RCP85 forcing, relative to the 1981-2000 mean. Annual data (thin lines), and smoothed data (thick lines) using a low pass filter with a 30-year cut-off are shown. Green curves show the average of Tmin and Tmax, while the black curve Tdif is the difference between this and Tmean. (a) shows a typical UKCP18 realization influenced by the indexing error (see text), with up to a 2.5°C difference in the black curve. This demonstrates the lack of coherence in the projections. (b) shows a typical realization following correction of the error, with much reduced values for Tdif.

Figure 2.2 illustrates this further, plotting individual realizations of July Tmax for the Southern England GCM grid-point as a function of Tmean, for the year 2090. Actual GCM values for these variables are plotted for comparison, obtained from transient climate change simulations used to underpin the pdfs (Appendix B). These consist of 57 Earth System Perturbed Parameter Ensemble projections (green) and 12 CMIP5 Earth System Model projections (black). Their results are much more highly correlated than the original UKCP18 realizations obtained with incorrect indexing (blue points). The standard deviation in Tdif for the blue points is 2.52°C, compared to 0.25°C for the CMIP5 data. The orange points show equivalent data, but for realizations with the indexing error corrected.

The corrected realizations show that July Tmax increases more rapidly than Tmean for changes in Tmean exceeding ~5°C. The linear regression coefficient for the corrected realizations is 1.26, and the relationship between Tmax and Tmean changes is much clearer. Their correlation is now much closer to that obtained for the GCM data, while the standard deviation in Tdif is reduced to 0.72°C. Nevertheless, the statistical processing required to produce the projections will reduce the inter-variable coherence somewhat compared to the underlying GCM data, even in the corrected projections. This is because the pdfs are made using statistical representations of plausible outcomes generated through emulation, scaling and sampling of internal variability and residual regression errors (see Appendix B). These techniques produce approximations of climate model output, and therefore do not fully reproduce the level of coherence found in GCM or RCM simulations. However, this statistical processing is an essential component of the method, that allows us to produce pdfs derived from a much larger sample of possible outcomes than is directly available from climate model output.

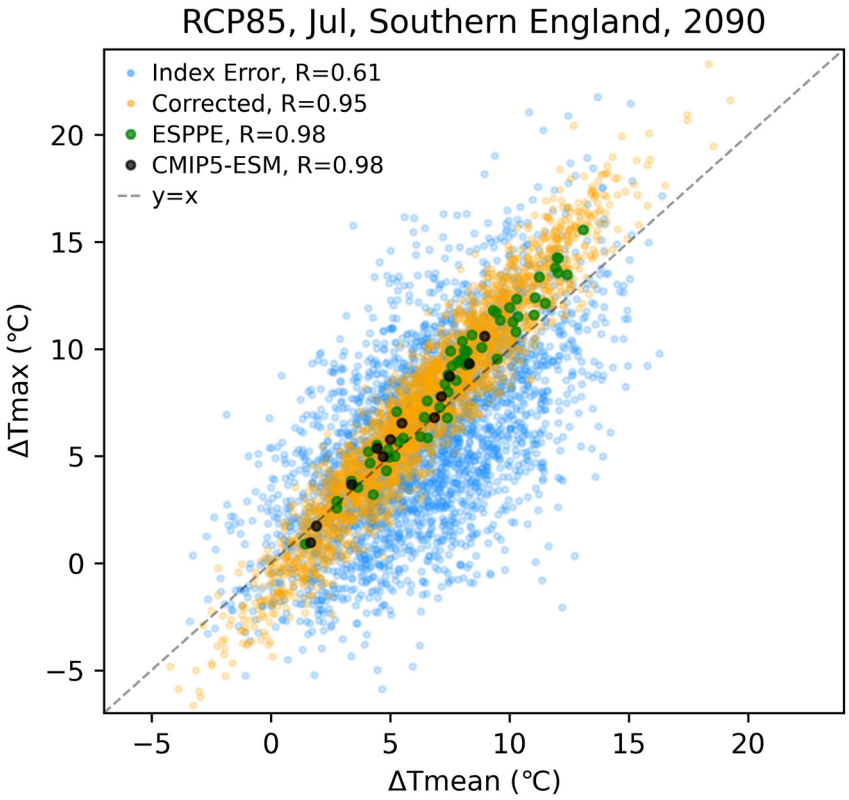


Figure 2.2. Realizations of July maximum surface temperature change for the year 2090 (relative to the 1981-2000 mean) in response to RCP85 emissions for the Southern England GCM grid-point (y axis) compared to corresponding realizations for July mean surface temperature change (x axis). Blue points correspond to the original UKCP18 projections, while orange points are for interim projections with the Tmax/Tmin software error corrected. Green and black points correspond to actual GCM values for these variables from the underlying Earth System Perturbed Parameter Ensemble projections and the CMIP5 Earth System Model projections respectively.

Correcting the indexing error affects multi-year averages of the projected changes, as well as improving the relationships between Tmean, Tmax and Tmin in individual years. As an example, Figure 2.3 compares the seasonal cycle in the future response of the P10 (low), P50 (median) and P90 (high) quantiles for Tmax and Tmin for the original and interim projections. These correspond to outcomes with a 10%, 50% and 90% chance of being exceeded. The same probability levels were used to characterise uncertainty ranges in the UKCP18 Science report (Murphy et al., 2018). The values shown are averages of P10, P50 and P90 values calculated for each UK administration region. Although the prediction for a summer peak in the response of Tmax is retained in the interim projections, there are some month-by-month differences. In particular, the correction brings forward the summer peak from August to July. The median July response for Tmax is therefore 1.4°C higher, while P90 is 1.8°C higher. The changes in P90 in February and October are also substantial, falling by 1.4°C and 1.5°C respectively. This leads to substantial reductions in spread, measured as P90 – P10: 2.0°C less for October, and 1.3°C less for February. July and December are the only months for which spread increases noticeably, by about 1.0°C in both cases.

The differences between the interim and original projections of P10, P50 and P90 are typically smaller for Tmin than those obtained for Tmax. This is because the seasonal cycle in Tmin changes is less pronounced than for Tmax, so the one-month offset caused by the indexing error generally has a smaller impact. For example, the root-mean-square differences (across all administration regions and months) for the median response are 0.35°C for Tmin, and 0.57°C for Tmax. In July, the increase in the median response for Tmin is

smaller than the increase for Tmax. P90 is substantially reduced in February and March, and also in October and November. The spread in Tmin is generally reduced a little, except for December, January and July.

In summary, the update to correct the Tmax/Tmin software error substantially improves the coherence of the surface temperature projections and also has an impact on the magnitude and spread of the Tmin and Tmax changes, since the seasonal cycle is now correctly identified. Notable increases in both response and spread are obtained in July. The chances of an extremely warm response over the autumn period are reduced.

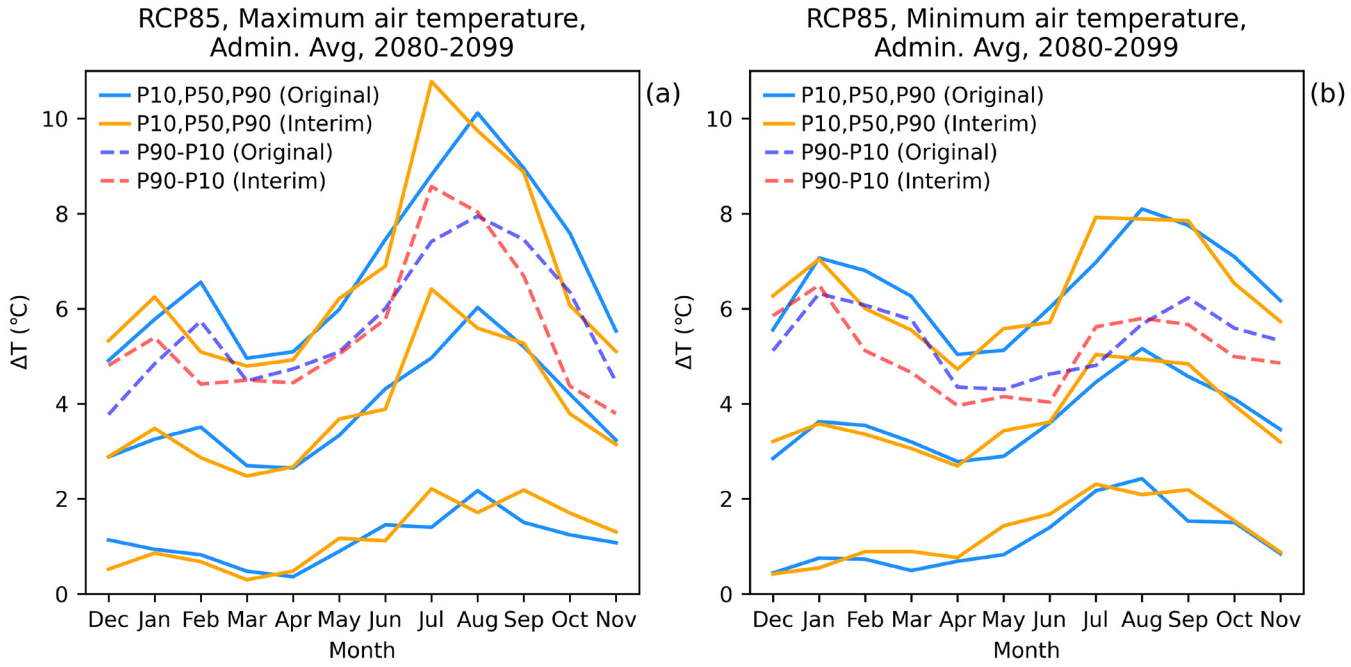


Figure 2.3. Comparison of the seasonal cycle for the P10, P50 and P90 quantiles (solid lines) for (a) monthly mean daily maximum surface temperature response Tmax, and (b) monthly mean daily minimum surface temperature response Tmin to RCP85 forcing for the original and interim projections for the 20-year mean period 2080-2099. Dashed lines represent the P90-P10 measure of spread. All quantities shown are averages of regional P10, P50 and P90 values calculated for each UKCP18 administration region. A 1981-2000 baseline is used.

2.2. Update to Smoothing Used to Estimate Pdfs

Production of the pdfs involves the use of statistical downscaling (Section 2.5), employing linear regression to link changes simulated by global climate models at the 300km scale to regional changes simulated by regional climate model (RCM) simulations at 25km scale. When the pdfs are made, the temporally-varying regression residual errors are also sampled. These capture both uncertainties in the relationship between long-term global and regional climate change, and also local RCM anomalies in individual years arising from simulated natural climate variability. These residuals are statistically noisy across time, since only 11 RCM simulations were available to estimate them. To produce the pdfs for UKCP18, data was therefore pooled over an 11-year sliding window. This reduced the effects of the residual noise, but not sufficiently to prevent undulations in quantiles of the pdfs on a 10-30 year time scale (e.g., Figure 2.4).

These undulations are a result of the limited sampling of phases of variability available from the 11 RCM simulations. Therefore, they do not represent physically meaningful variations in the spread of the pdfs. We have instead employed a 31-year sliding window in the new projections, to better characterise evolution of uncertainties on 1-30 year time scales. Using this larger window, the percentiles in Figure 2.4 are

considerably smoother, and better reflect the longer-term trends in the underlying realizations. Note that for both pooling windows, the median response for July precipitation in Southern England in Figure 2.4 does not equal zero over the 1981-2000 base climatology period, although the mean response during 1981-2000 is centred on zero. This difference arises because the precipitation distribution is positively-skewed, and therefore asymmetric. The change in pooling interval only affects the variability apparent in the annual pdfs and has no impact on variability in the 20-year and 30-year multi-year pdfs. This update is applied to all variables (see Section 2.6) and has a similar impact to those shown in Figure 2.4 (see Section 3.2 for further examples).

Physically-based contributions to the spread in the annual pdfs arise from uncertainties in long-term climate response, and interannual variability (derived from the RCM simulations). The smoothing shown in Figure 2.4 arises from a more robust estimate of the contribution from interannual variability, but does not result in a systematic reduction in its magnitude. For precipitation, it remains the case that including interannual variability drives a considerable increase in the spread of the annual pdfs, in comparison to pdfs of 20- or 30-year average changes (e.g., Murphy et al. (2018), Figure 2.9).

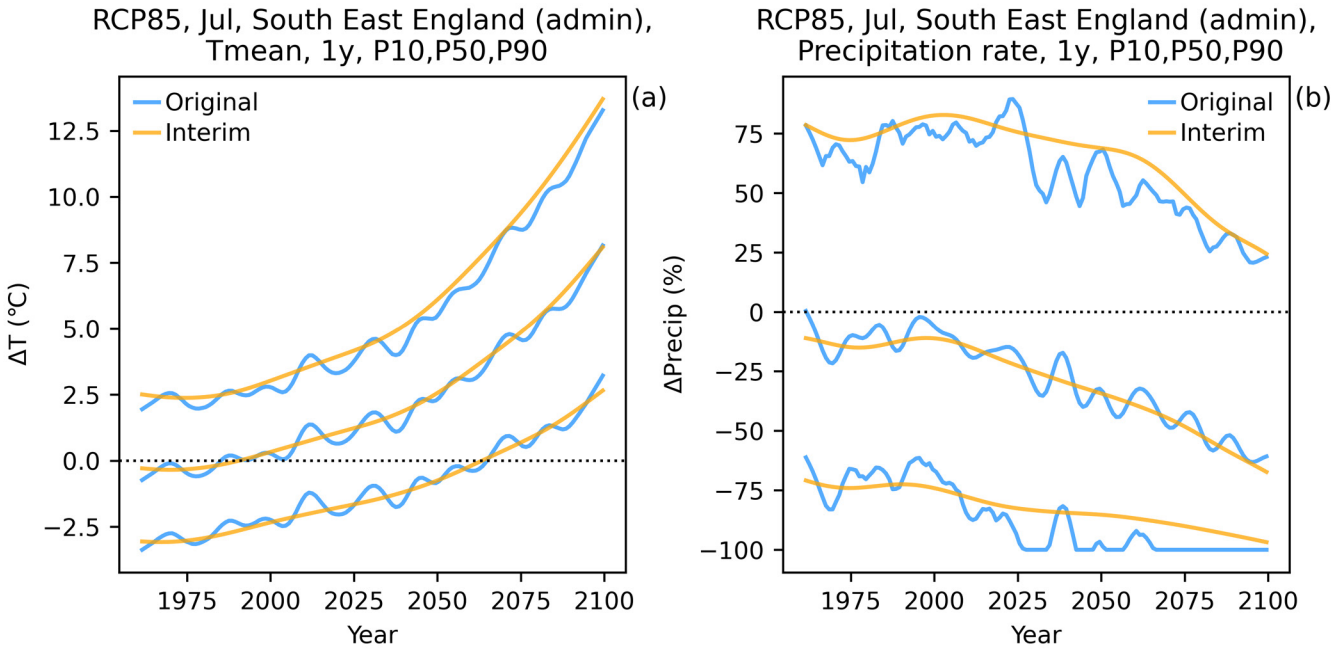


Figure 2.4. (a) Comparison of P10, P50 and P90 values from the pdfs of projected annual changes relative to 1981-2000 for Tmean in July (°C) for the South East England region, in response to RCP85 emissions. Blue curves correspond to the original UKCP18 pdfs constructed using an 11-year centred pooling window, while orange curves use a 31-year pooling window. (b) As (a) but for precipitation change (%). Orange curves show the interim projections, containing all developments except the downscaling correction.

2.3. Improved Estimation of Precipitation Response in the Tails of the Distribution

For precipitation, an additional change is applied to the calculation of these residuals. This arises from the bounded nature of the variable: changes are expressed as percentages and cannot go below -100%, since this indicates zero precipitation. However, the statistical application of scaling and sampling techniques to responses expressed as percentage changes can occasionally lead to the lower bound being exceeded (see Appendix A). The blue curve in Figure 2.5a shows an example of one such realization for projected July precipitation changes in response to RCP85 forcing for the London administration region. It produces a mean drying of -71% for the period 2070-2099, and nine individual years have a predicted drying exceeding 100%. The blue curves in Figure 2.6 correspond to the probability distribution for the full sample of realizations in Figure 2.5. A substantial fraction (16%) of unphysical seasonal anomalies occurs.

For precipitation, these negative values were removed from the original realizations and pdfs in UKCP18, as part of the clipping applied to extreme values (UKCP18 Technical Note, 2019). This happened in a minority of cases, notably in southern locations during summer under strong forcing scenarios. An alternative approach is to apply a logarithmic transformation to precipitation data prior to all statistical processing (Watterson, 2008), followed by application of the inverse transform to later recover projections expressed as percentage change:

$$\Delta P\ (\%) = 100(e^{y} - 1),\ \ y = lnP - ln\langle P\rangle \qquad (2)$$

where $\langle P\rangle$ represents the temporally-averaged precipitation P over the baseline period.

Two projections are made for precipitation at each grid point and season, one using percentage change and one using the scaled log transform approach. The red curve in Figure 2.5a shows the realization corresponding to the blue curve when the scaled log transform is used. By construction this avoids negative precipitation events. Although similar, the noise characteristics of the two time-series can vary in some details due to the non-linear properties of the log transform. However, use of the log transform for all realizations can present other difficulties, as illustrated in Figure 2.5b. This shows the same variable as in Figure 2.5a, but for a different sampled realization. In this instance, two individual years are predicted to be above 500%. Across all 3000 sample realizations a few individual months with changes as high as 1800% are obtained with the scaled log approach. These occasional large positive anomalies at the wet end of the distribution are generated statistically and are implausible, as they are not supported by the underlying RCM data. The largest positive precipitation anomaly found in the RCM simulations was about 310% and is shown as the dashed line in Figure 2.5b. The statistical generation of these large positive anomalies is explained further in Appendix A. The red curves in Figure 2.6 show the probability distribution for the full sample of scaled log projections. A long tail in the cdf is clearly visible, representing the implausible, wet end responses. At the dry end of the distribution, no realizations below -100% are permitted when using the log transform. However, the probability (although small) is not identically equal to zero. This arises since we use kernel density smoothing (Virtanen et al, 2020) to estimate the pdf, which introduces a small amount of additional statistical spread (see below for further discussion).

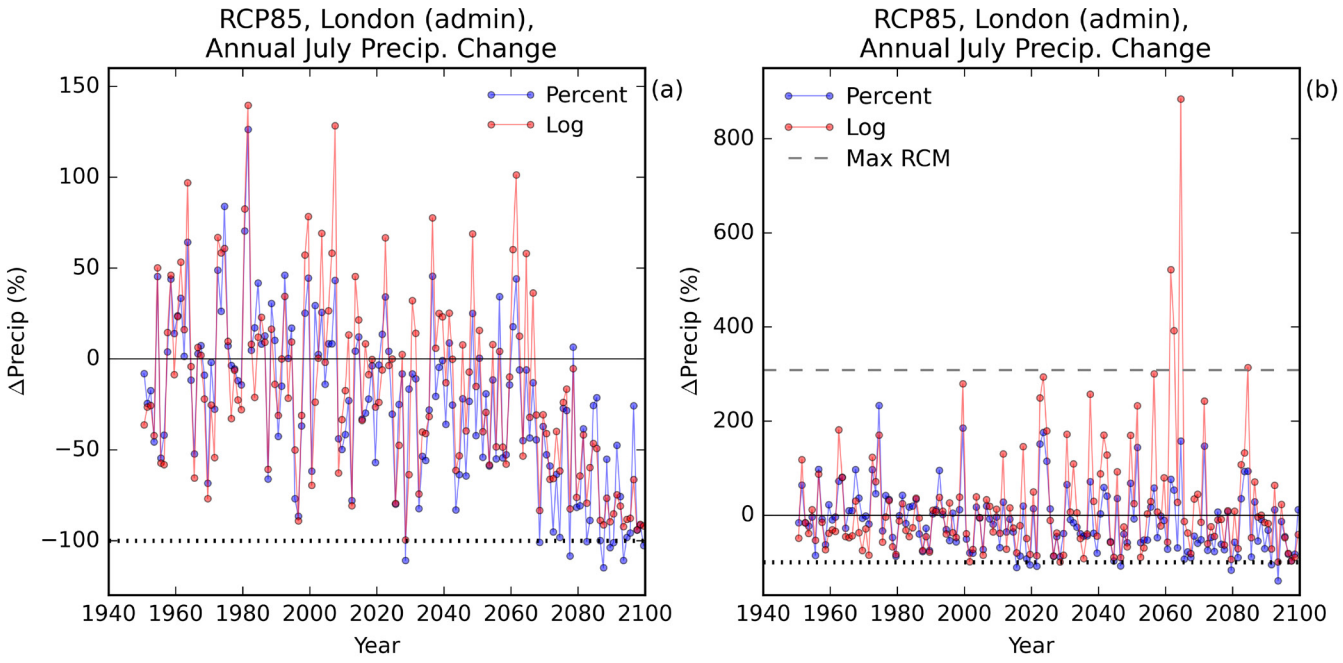


Figure 2.5. (a) One realization of projected percentage change in July precipitation (relative to the 1981-2000 mean) in response to RCP85 forcing for the London administration region. The blue and red curves use the scaled percentage and scaled log approaches respectively. This realization shows an example of strong drying and demonstrates the potential for "unphysical" values below -100% with the scaled percentage approach. (b) A different realization for the same variable in (a), this time selected to demonstrate the possibility for generation of implausible annual wet events much greater than the maximum value in the underlying RCM data.

These issues in the tails of the precipitation distribution occur mainly in the summer months when strong drying is associated with strong forcing (Appendix A). To circumvent them, a new mixed approach has been developed. The log transform is employed for those realizations with strong drying, while for realizations with weak drying or projected increases in precipitation, linear scaling of percentage change is used. For example, the log transform is selected for the realization in Figure 2.5a, to avoid the negative precipitation events encountered when the scaled percent method is used. Conversely, for the realization in Figure 2.5b use of the log transform inflates the spread beyond the range of the RCM data, so in this case the scaled percent (blue curve) is selected. However, there are still a few instances with annual values below -100% in this realization, so they are reset to the equivalent annual values based on the scaled log-transform projection. The maximum permissible value for the scaled log projections is defined as the maximum projected anomaly in the scaled percent projections. Further details of the selection algorithm for the method are given in Appendix A.

The green curves in Figure 2.6 show the probability distribution obtained by applying this merging method to the full sample of projections corresponding to the individual realizations of Figure 2.5. A pooling window of 31 years is used. As designed, the merged distribution for very dry responses tends toward that obtained for the pure scaled log projections (red). At the wet end the converse is true, and the merged distribution approaches that for the pure scaled percent projections (blue). In these strong drying cases, the overall shape of the distribution is therefore closer to that of pure scaled log projections. P10 for the scaled percent distribution is an unphysical -111% (values like this were clipped to 100% for UKCP18). The corresponding P10 for the merged distribution is -93%. The median for the merged distribution of -56% is close to that for the scaled log distribution, and 6% less than the value of -50% obtained for the scaled percent distribution. P90 for the scaled log distribution is 26% higher than that obtained for the scaled percent distribution, demonstrating that the statistical generation of implausible wet extremes also has an effect at the P90 level. For the merged distribution, P90 is generally close to that for the scaled percent distribution, and here

equals +37%. For smaller precipitation changes (in the range ~±30%), the merged distribution produces probabilities that transition between the scaled percent and scaled log choices.

Even though instances of negative precipitation are now removed completely from the realizations, the kernel density smoothing used to estimate the pdfs can still generate small cumulative probabilities for negative precipitation. In the example of Figure 2.6a (green curve) this equals 3.0%. However, in the final pdfs released to users, clipping at this level is applied to remove these statistical outcomes in the low-end tails. This clipping of the precipitation pdfs is in addition to the clipping at the 1% and 99% probability levels that is applied for all variables to both realizations and pdfs (see UKCP18 Technical Note (2019)). In the present update, the clipping of negative precipitation tails in the pdfs (typically at a probability level of ~2%) is only required over the summer months from June to September, mostly for southern parts of England and Wales. These negative precipitation tails occur mainly during the second half of the 21st century, under scenarios of high warming. For 2069-99 under RCP85 emissions, the negative precipitation clip is applied on average to 44% of all UK grid-points during June to September.

In summary, for distributions with strong drying the improvement to the handling of extremes often leads to a general sharpening of the distribution, with less extreme values at the dry end of the distribution, and somewhat drier median responses. For annual July projections in response to RCP85 emissions, this update makes the median changes typically ~5% drier compared to the original results, for 25km grid-points.

Figure 2.7a shows an interim precipitation pdf for the full summer season for Wales. Compared to the smaller-scale London example, the Wales pdf shows a smaller range and a weaker median reduction of -37%. The most notable effect of merging is a P10 value of -71%, 9% less dry than obtained for the pure scaled percent distribution. Figure 2.7b compares autumn distributions for the Northumbria river basin regional mean, which has no substantial median signal for precipitation change (just a +5% increase). Compared to the scaled percent distribution, the merged transform approach results in P10 increasing by 4% to -32%; otherwise, there is no significant impact, the P50 and P90 values being almost identical. These examples demonstrate that the updated methodology is successful in both avoiding the generation of realizations below -100% and improving credibility for wet extremes. Figure 3.3 in Section 3.2 shows an example of how the merged transform approach influences the temporal evolution of summer precipitation changes in the new projections.

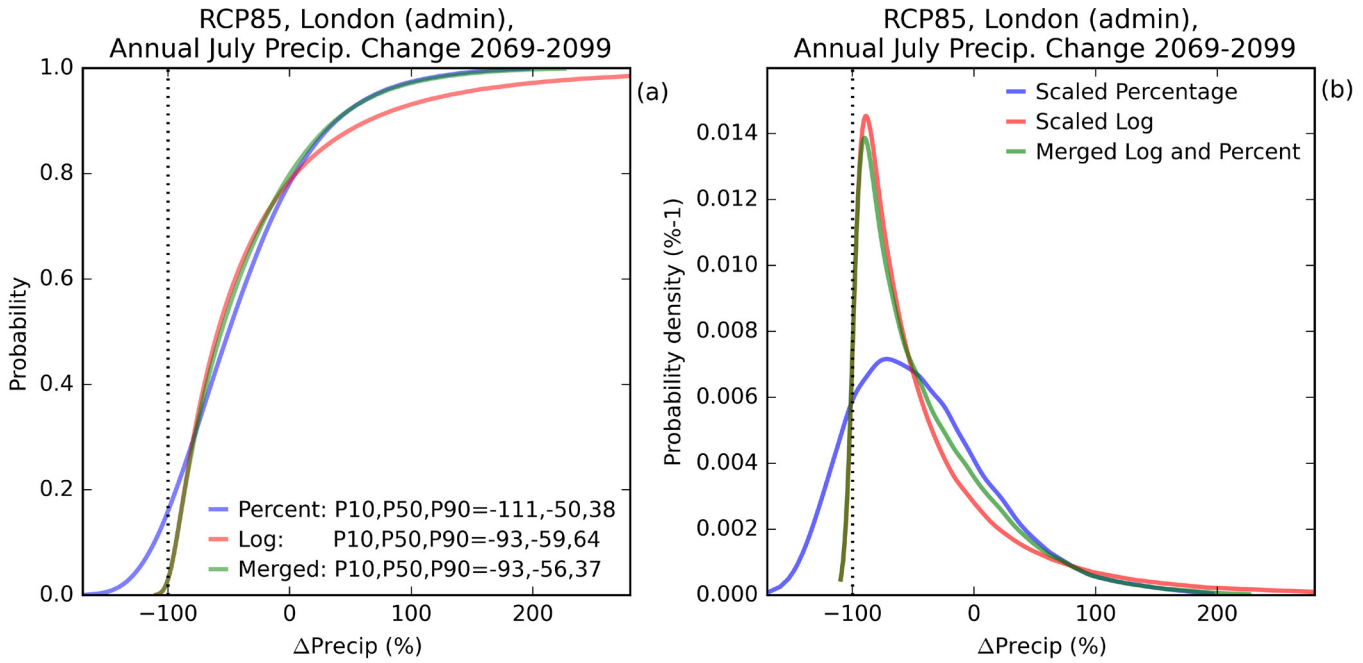


Figure 2.6. Cdfs (a) and pdfs (b) for the London administration region, constructed from the full 3000-member sample of realizations (one of which is shown in Figure 2.5). These show projected percentage change in July precipitation for individual years during 2069-2099 (relative to the 1981-2000 mean) in response to RCP85 forcing. The blue and red curves correspond to the scaled percentage and scaled log approaches respectively. The green curves show the probability distribution obtained by merging scaled percentage and scaled log realizations, as described in Section 2.3. The legend in (a) gives values for the P10, P50 and P90 quantiles for the three distributions (units for these are % change).

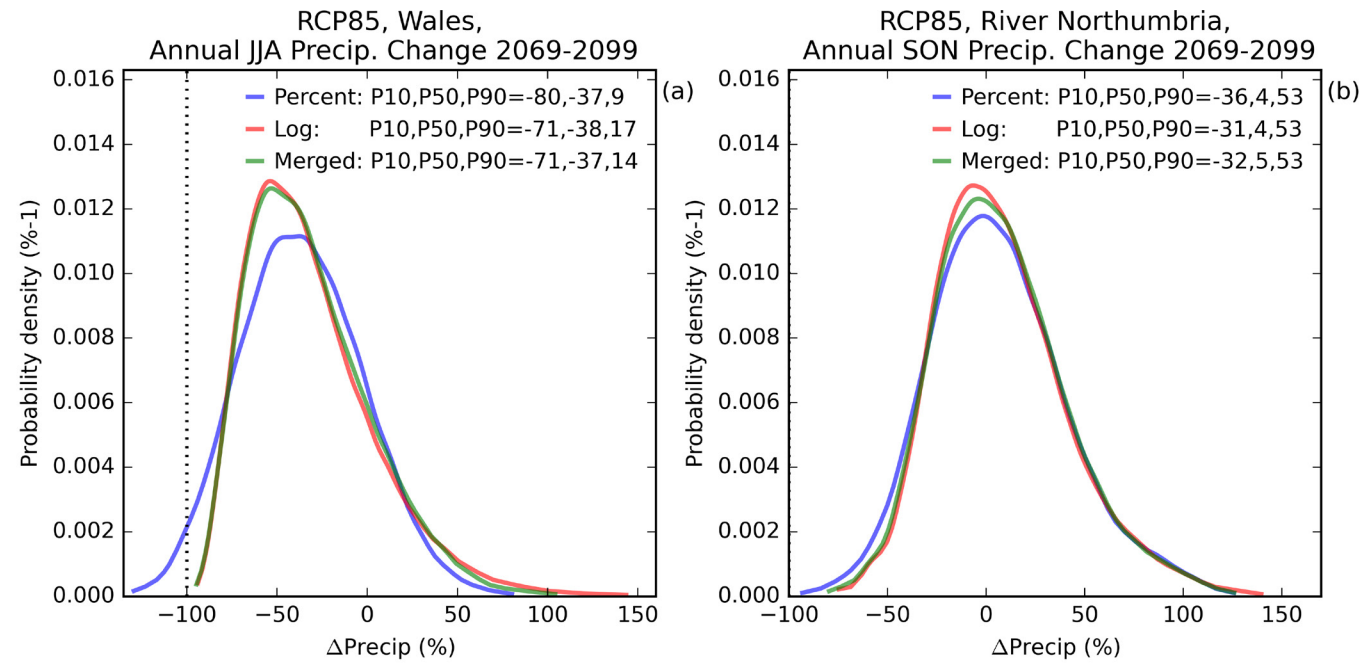


Figure 2.7. Pdfs for projected percentage change in precipitation for individual years for (a) Wales in summer (JJA), and (b) the Northumbria river basin region in autumn (SON). The pdfs show pooled annual changes for 2069-2099 relative to the 1981-2000 mean, in response to RCP85 forcing.

2.4. The Effect of Filtering on Baseline Centering

The UKCP18 pdfs are presented for all variables as anomalies relative to a baseline period of 1981-2000. By definition, the average anomaly during the baseline period should be zero. However, due to application of temporal filtering in the statistical calculations, average anomalies during 1981-2000 were not exactly zero in the original UKCP18 pdfs. While the average offsets were typically very small, they are larger in some individual realizations, implying that for some user applications bias corrections could be needed. This was documented in a UKCP18 Technical Note (2019). For the revised projections here, the methodology has been reformulated to apply the filtering at a different point in the calculations, ensuring that all 3000 realizations for a given variable are precisely centred on the 1981-2000 baseline.

The effect of this update is demonstrated in Figure 2.8, using projected annual changes in Tmean for Wales in January as an example. Although individual realizations may be up to 0.5°C warmer or cooler when the baseline is recentred, the effect on the distribution is barely discernible in Figure 2.8b, with a median reduction of 0.1°C. January is the month for which the baseline recentring has greatest impact. For Tmean, monthly adjustments show a typical magnitude of 0.03°C on average over all UK regions. These effects are small compared to both the spread in response, and to the effects of the other updates (Section 2.1-2.3 and 2.5). Similarly, small effects are obtained in the case of precipitation. Baseline recentring leads to typical mean adjustments of about 0.3% (positive or negative) for projected monthly precipitation change. The largest adjustments occur in December, with a reduced precipitation response of 0.8% averaged over the administrative and river-basin regions.

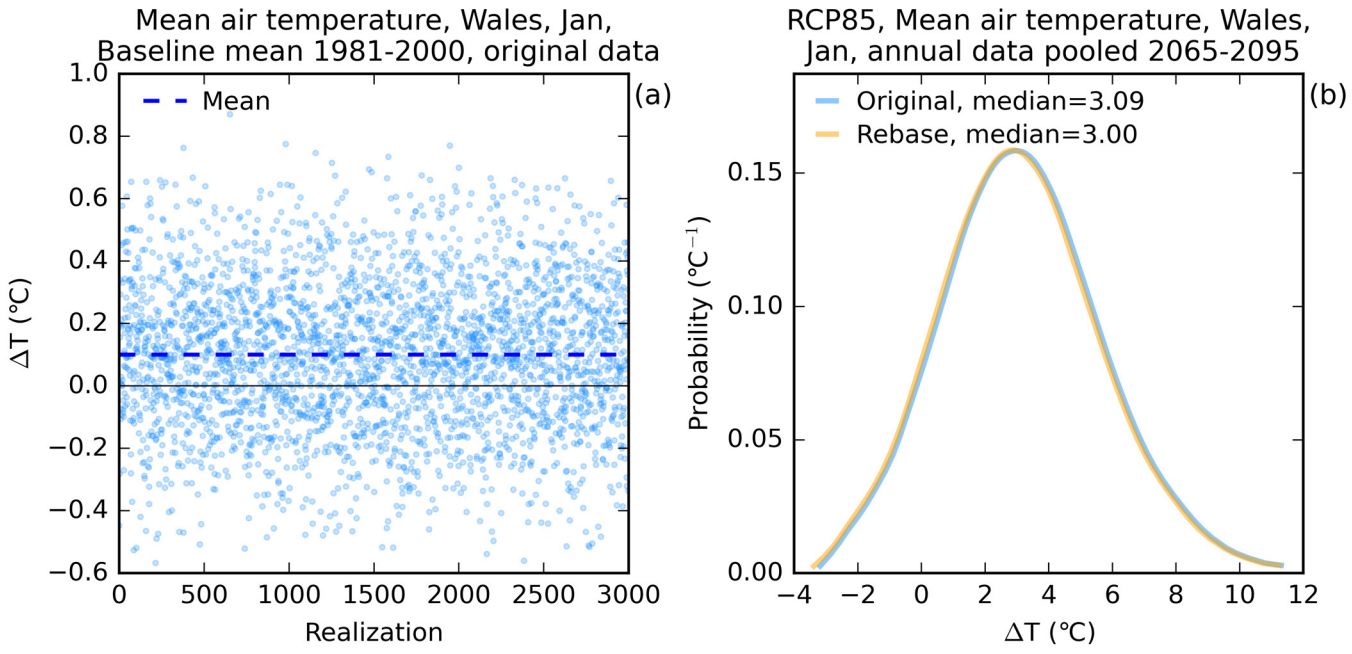


Figure 2.8. (a) 20 year average values for the baseline period 1981-2000 for the 3000 sample realizations of January Tmean in Wales, from the original UKCP18 data. The mean value for the sample is 0.10°C, with standard deviation 0.23°C. (b) Corresponding pdfs of future change for individual Januaries between 2065 and 2095, in response to RCP85 emissions. The new pdf with corrected baseline (orange) is almost indistinguishable from the original UKCP18 version (blue).

2.5 Software Error affecting Downscaling Calculations

A software error was recently discovered in the downscaling calculations used in the original UKCP18 probabilistic projections. Below, we provide a recap of the downscaling method. This is followed by a description of the error and the impacts of correcting it on the new projections.

a. Recap of downscaling methodology

The methodology underpinning the probabilistic projections involves several steps applied to results from GCM simulations, leading to a set of 3000 equally-likely realizations of time-dependent monthly, seasonal and annual anomalies for each year from 1961-2100 (Appendix B) at the spatial resolution of the GCM. The final step is to downscale the GCM time series to a 25km version of the OSGB British National Grid and to three sets of aggregated regions (see Fung et al., 2018). These calculations use an 11-member PPE of HadRM3, the RCM configuration of HadCM3 (Murphy et al. 2009). Downscaling is achieved using univariate linear regression relationships that link projected changes in the RCM simulations with changes from one of five UK land points in the driving GCM simulations. The latter consists of a corresponding PPE of the HadCM3 model using the SRES A1B emissions scenario. The GCM boxes correspond approximately to Scotland, Ireland, Northern England, Southern England and Wales (Harris et al., 2010). The regression relationships adjust climate change signals from the GCM to account for the finer scale information available from the RCM. Regression residuals include signals of change not captured by the linear relationships, and uncertainties due to interannual variability. The residuals are resampled and included in the data provided to users, consisting of the 3000 realizations of downscaled anomalies and corresponding pdfs.

For each OSGB grid box, a single regression relationship is calculated by pooling data from the eleven pairs of RCM and GCM simulations (e.g. Figure 2.9). The data are expressed as anomalies relative to 1981-2000, and the regressions are calibrated using data from 1990-2099 to focus on future climate change signals. Prior to calibration the GCM anomalies are low-pass filtered to remove variability on time scales shorter than 20 years (see Figure 2.5 (right) of Murphy et al., 2018) and hence isolate long-term changes. The RCM data is left unfiltered, to ensure that the regression residuals capture interannual variability in the RCM simulations.

The relationships use GCM data for the same variable as the target RCM variable, so RCM precipitation is always paired with GCM precipitation, etc. The nearest GCM grid box is used as the predictor variable, unless a different GCM box explains at least 10% additional variance of the RCM anomalies. Separate relationships are derived for each 25km grid box, and for each month or season. Therefore, the GCM predictor used for (say) Bristol precipitation in June can be different from that used for an adjacent 25km box in June, or from that used for Bristol precipitation in July.

Downscaling for the aggregated regions was done in a similar way, except that the GCM predictor box for each aggregated region was prescribed as that selected most often in its constituent 25km boxes. This was done to reduce the risk of inconsistencies between the aggregated region and OSGB results.

b. Downscaling software error

To calculate the regression relationships correctly, data from each of the eleven RCM simulations must be paired with the corresponding GCM simulation used to drive it with prescribed time series of sea surface

temperature and lateral atmospheric boundary conditions (of surface pressure, wind, temperature and moisture values). In this way, the regressions can precisely isolate the downscaling impacts of finer spatial resolution in the RCM. Unfortunately, this pairing was recently found to be incorrect in the 25km downscaling calculations made for the original UKCP18 results. The error arose from an Iris function[3] (Iris, 2010-22) used to load RCM data on the OSGB grid, resulting in incorrect indexing of RCM ensemble members and hence misassignment of GCM-RCM pairs.

This indexing error did not directly affect the downscaling relationships for aggregated regions in UKCP18, as a different Iris function was used to load the RCM data for these. However, some of the aggregated-region results were affected indirectly by the error in the 25km data. This is because fixing the error can lead to a different selection of GCM predictor for some 25km boxes. If sufficient 25km boxes within an aggregated region change predictor, this can in turn lead to a different choice of GCM predictor for the aggregated-region downscaling (see above).

c. How does the downscaling error affect GCM-RCM regression relationships?

The downscaling software error affects all the probabilistic projection variables. On average, across all variables, OSGB grid boxes and period definitions (months, seasons and annual mean) used in the probabilistic projections, fixing the downscaling error leaves the selection of GCM predictor box unchanged in 75% of cases. Figure 2.9 gives three examples in which the same GCM predictor is selected in the original and corrected calculations. Panel a shows the original regression relationships for Tmean for the London box in October, including the downscaling error. Each GCM simulation is colour coded. The black dashed line shows the overall regression (calculated across all eleven GCM-RCM pairs), while the coloured lines (provided for illustration) show individual regression lines for each pair. The latter show a wide spread of slopes: For example, the light blue (pink) lines in Fig. 2.9a show an example where a GCM simulation with relatively high (low) future warming is incorrectly paired with an RCM simulation showing a low (high) warming. On fixing the downscaling error (Fig. 2.9b) in the new projections, the spread in these individual regressions becomes much smaller since the GCM and RCM simulations are now correctly paired. The overall slope becomes slightly larger in the corrected regression (0.96 in Fig. 2.9b cf 0.90 in Fig. 2.9a), better reflecting the true physical relationships between the RCM and GCM changes.

Figure 2.9c-f show further examples, relating to precipitation for Glasgow in September and sea-level pressure (slp) for Leeds in winter. In these cases, correcting the indexing error leads to larger increases in the regression slope. The impact on slope varies because the correction removes a spurious contribution to the cross-ensemble scatter in the GCM-RCM relationships but does not affect the relationship between the ensemble-mean changes. For variables (such as Tmean, Tmax and Tmin) where the ensemble-mean changes provide a substantial fraction of the total variance of projected anomalies in the GCM and RCM ensembles, correcting the indexing error increases the regression slope only modestly. In contrast, larger increases in slope occur for variables in which the projected ensemble-mean changes are relatively small, compared with the ensemble ranges of change. For this reason, slp shows the largest increases in slope, from 0.12 to 0.94 when averaged across all 25km locations and period definitions. Substantial increases also occur for precipitation (from 0.41 to 0.78 on average), total cloud cover (from 0.30 to 0.63 on average) and the surface radiation variables.

[3] The load function in the Iris analysis software library used in 2018 (version 1.9.2) randomized the input under certain use conditions and metadata settings. Current Iris versions no longer do this.

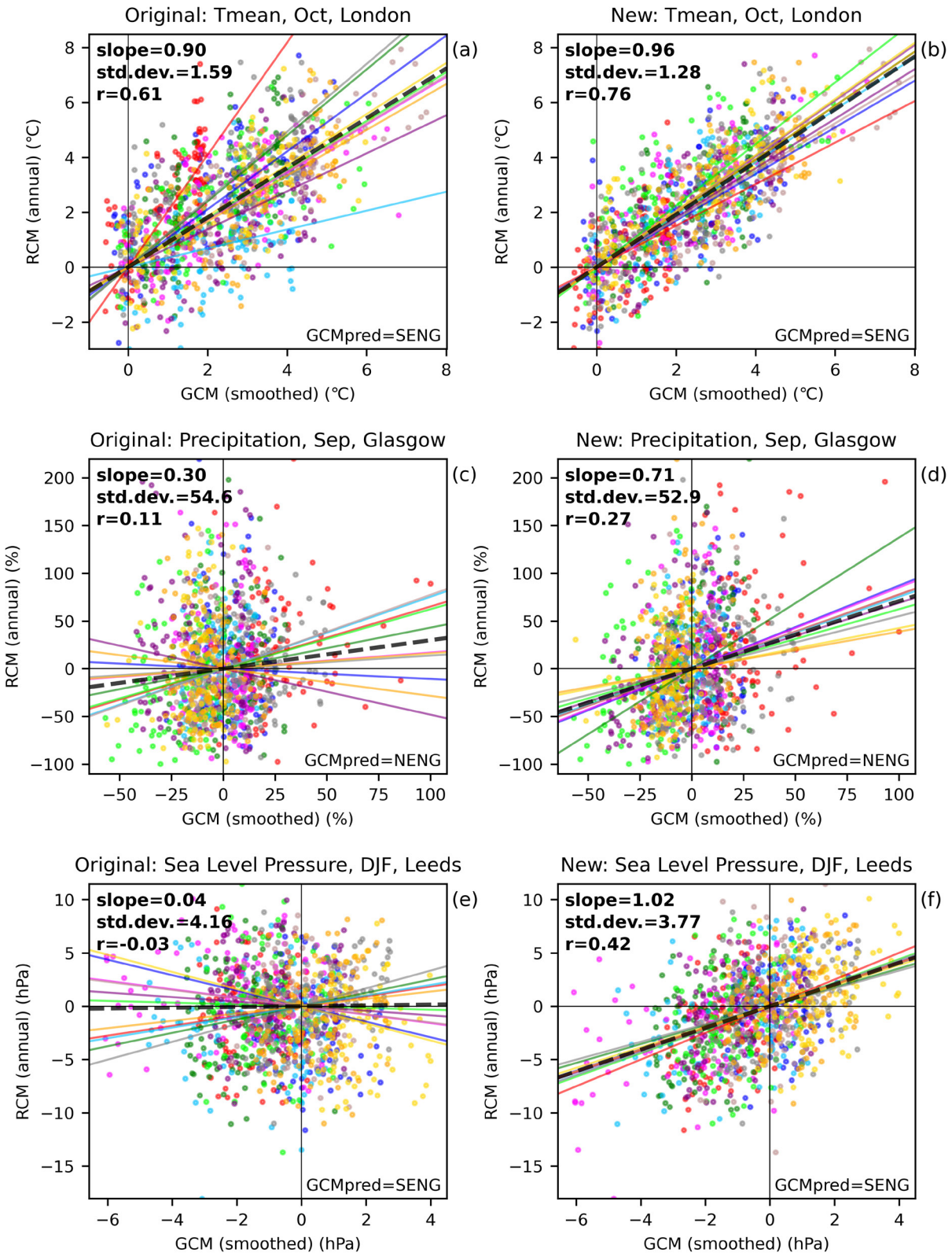


Figure 2.9. Scatter plots showing relationships between the eleven pairs of GCM and RCM simulations used to calibrate linear regression relationships for downscaling to 25km OSGB boxes in the UKCP Probabilistic Projections. Panels on the left show original results containing the downscaling software error, with corrected results on the right. Top, middle and bottom rows show Tmean for London in October (a, b), precipitation for Glasgow in September (c, d), and sea-level pressure (slp) for Leeds in winter (e, f). Anomalies relative to 1981-2000 are shown for each year during 1990-2099, with GCM anomalies (but not RCM anomalies) smoothed to remove variability on time scales shorter than 20 years. Black dashed lines show the regression relationships used to form pdfs of local changes, derived by pooling all the data. Each dot is coloured to denote a specific GCM ensemble member. The coloured lines show relationships from individual GCM-RCM pairs, included to show the effects of fixing the incorrect pairings on the left. Bold figures show the (pooled) regression slopes, the standard deviations of associated residuals and the correlations (r) between GCM and RCM anomalies. In all cases, the GCM predictor box remains unchanged in the new regressions.

In Figure 2.9, the residual downscaling variances reduce by 35%, 6% and 18% in the Tmean, precipitation and slp examples respectively. This is a second benefit of removing the spurious component of uncertainty caused in the original results by the incorrect indexing. Increases in slope act to broaden the spread in the probabilistic projections, since stronger downscaling relationships are more effective at transmitting the ranges of climate response found in the GCM projections to local scales. This is particularly evident in the slp example (Fig. 2.9e, f), where the original regression finds no relationship between the GCM and RCM anomalies whereas the corrected regression finds a slope close to unity. In contrast, reductions in residual variance act to reduce the spread in the pdfs, since smaller residuals imply smaller downscaling uncertainties. These competing influences on spread are discussed further in Section 2.5d.

For precipitation and slp the original calculations show small correlations (0.11 and 0.03 respectively) between GCM and RCM anomalies in Fig. 2.9. In the corrected calculations, correlations increase to 0.27 for precipitation and 0.42 for slp but remain modest because there is still considerable scatter in the corrected relationships. The Tmean example also shows significant scatter in the corrected results, although the correlation is higher in this case (0.76). The scatter includes the influence of interannual variability in the RCM anomalies, with relatively low correlations occurring in cases where interannual variability provides a high proportion of the uncertainties present in the RCM projections. The contribution of RCM interannual variability is included in the downscaling relationships by design (section 2.5a), so that interannual variability can be accounted for when sampling the downscaling residuals to construct the 3000 realizations from which local pdfs are derived.

Figures 2.10 and 2.11 show regression slopes and residuals for Tmean in summer and precipitation in autumn, for all OSGB locations. In some places the selected GCM predictor box changes when the downscaling error is corrected (see panels g and h in both Figures), although 89% of the selections remain unchanged in Fig. 2.10 and 73% in Fig. 2.11. The regression slopes nearly always increase (modestly for Tmean and more strongly for precipitation) while the residual variances nearly always reduce. The only exceptions occur in (some of) the places where correcting the error leads to selection of a different GCM predictor. For example, the Tmean slope reduces slightly at a few locations in eastern Scotland and northern England, due to changes in GCM predictor from Scotland to North England in the former case and from North England to Wales in the latter.

For aggregated regions, the downscaling regressions between the GCM predictor box (once selected) and the aggregated RCM data were calculated correctly in the original UKCP18 results (Section 2.5b). However, the software error did potentially influence the choice of GCM box since this was derived from OSGB regressions for all 25km boxes within the relevant region (Section 2.5a). Therefore, correcting the indexing error influences the aggregated region results only in cases where this leads to a different selection of GCM predictor. Across all periods, variables and regions, this occurs on 34% of occasions for Tmean and 27% of occasions for precipitation.

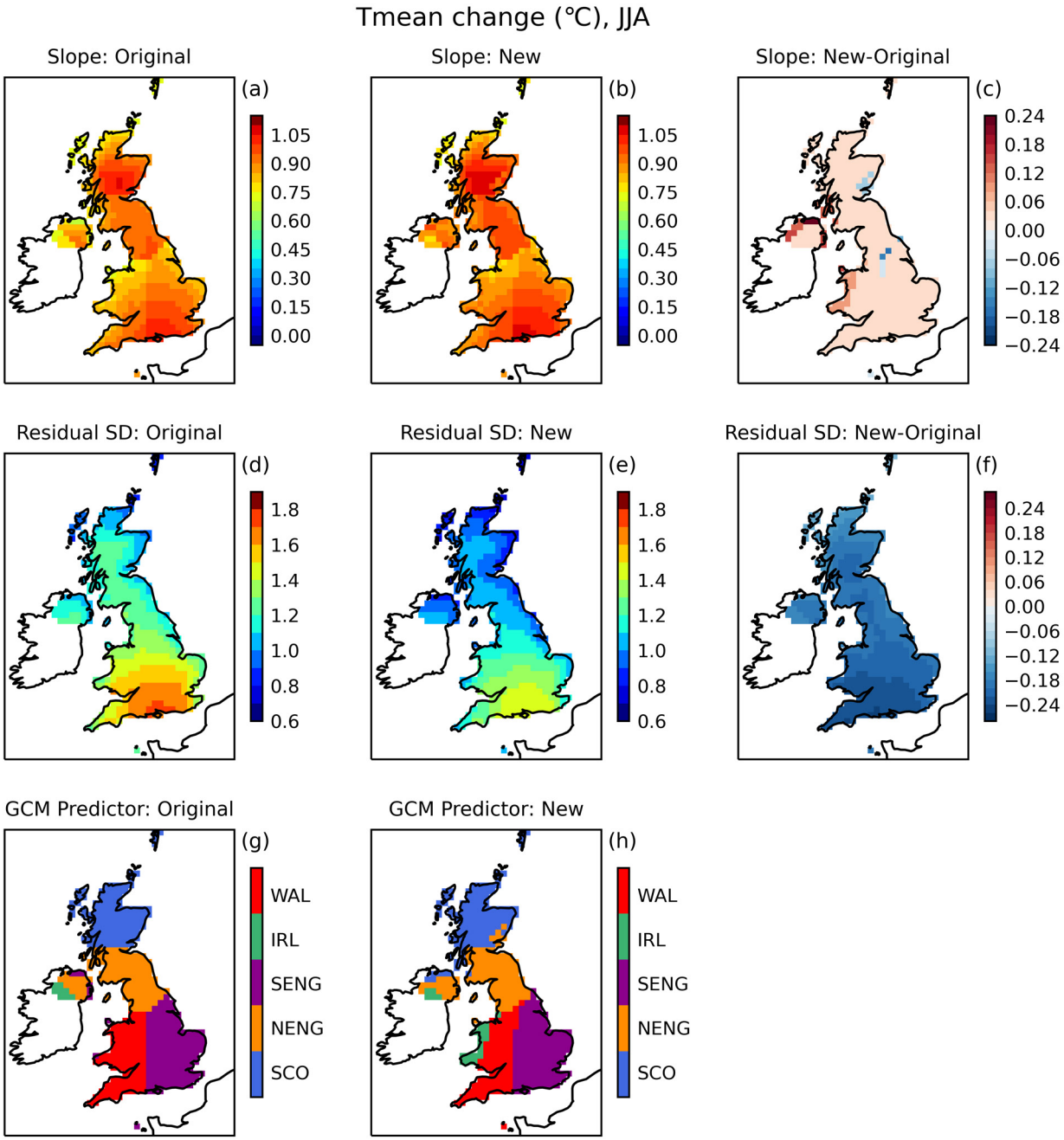


Figure 2.10. Downscaling slopes and residual standard deviations at OSGB 25km locations for anomalies in Tmean in summer, for the original regressions (left, (a) and (d)), the new regressions with the software error fixed (middle, (b) and (e)), and differences (right, (c) and (f)). The GCM boxes selected for downscaling are shown in (g) and (h), for the original and new regressions respectively.

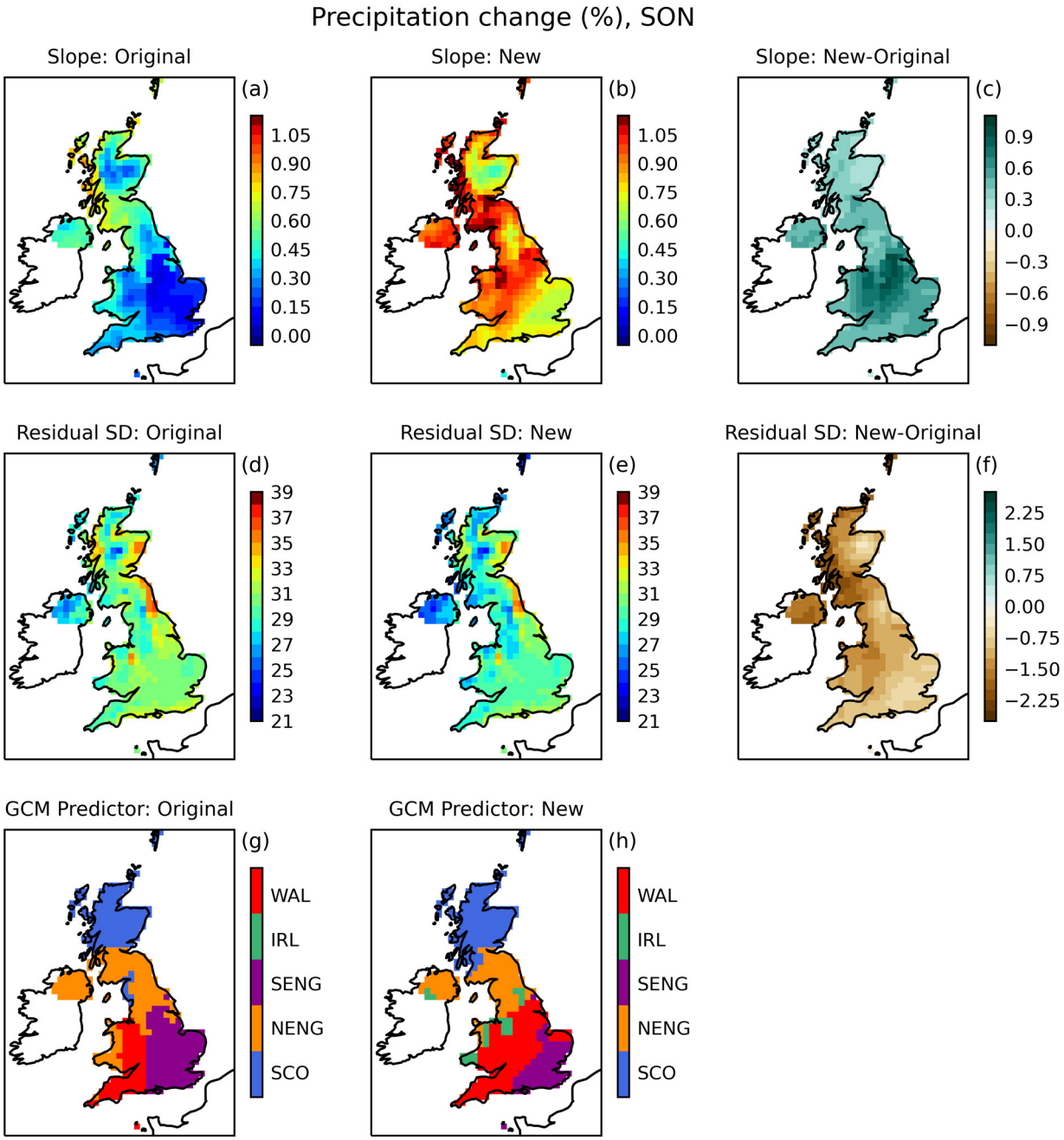


Figure 2.11. As Figure 2.10, for slopes, residual standard deviations and GCM predictor selections in the original (left) and new (middle) downscaling regressions for precipitation anomalies in autumn.

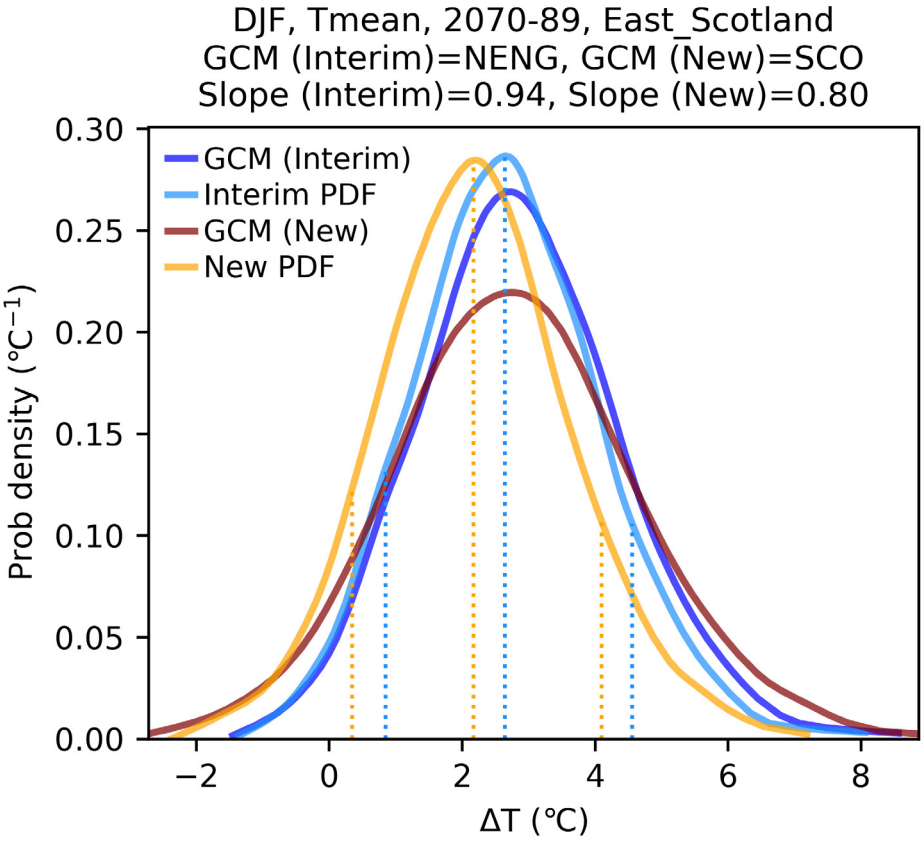


Figure 2.12. Probability distributions for changes in Tmean for 2070-2089 relative to 1981-2000 in response to the RCP85 scenario, for the East Scotland aggregated region in winter. Light blue curve shows the pdf from the interim projections, including the developments of Sections 2.1-2.4 but not the downscaling correction. The orange curve shows the new projections containing all updates including the downscaling correction. The dark blue curve shows the pdf for North England, the GCM box used to provide downscaling for the interim East Scotland pdf. The maroon pdf is for the Scotland GCM box, used for downscaling in the new projections. Dashed lines show the P10, P50 and P90 values of the East Scotland pdfs.

Figure 2.12 shows an example, for Tmean in East Scotland in winter. The selected GCM box was North England in the original downscaling and Scotland in the corrected version. The impact of changing the GCM predictor is isolated by comparing the new East Scotland pdf that includes all five developments (orange curve) against the interim pdf including all developments except the downscaling correction (light blue curve). Projected changes are shown (as for all the examples discussed in this section) for 2070-2089 relative to 1981-2000 using the RCP85 scenario. The P10, P50 and P90 values in the new distribution are shifted approximately 0.5°C cooler. The Scotland GCM box (maroon curve) shows a similar median but a broader spread than the North England box (dark blue curve), exploring warmer and cooler outcomes at either extreme. This reduces the regression slope in the new projections (0.80 cf 0.94 in the interim projections), because a flatter relationship is required to fit the warmest and coolest GCM responses to the target RCM data when Scotland is substituted for North England as the predictor variable (Figure 2.13). The reduction in slope explains the cool shift in the new regional pdf. The spread in the new and interim East Scotland pdfs is similar because the impact of sampling the broader GCM pdf for Scotland is offset by the reduced regression slope.

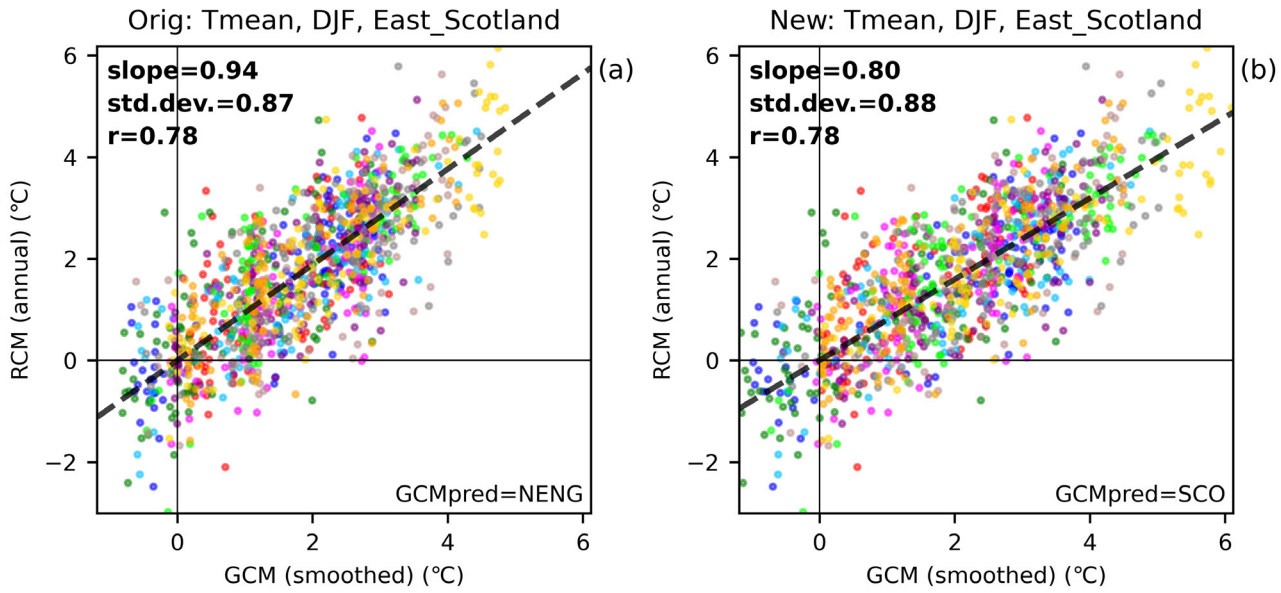


Figure 2.13. Scatter plots showing relationships between the eleven pairs of GCM and RCM simulations (colour coded as in Fig. 2.9) to calibrate regression relationships (black dashed lines) for downscaling to East Scotland for Tmean in winter. Anomalies relative to 1981-2000 are shown for each year during 1990-2099, with GCM anomalies (but not RCM anomalies) smoothed to remove variability on time scales shorter than 20 years. Panel (a) shows the original regression used in the interim regional pdf (light blue curve in Figure 2.12). Panel (b) shows the new regression used to make the new regional pdf (orange curve in Figure 2.12). Correction of the OSGB downscaling error leads to selection of the Scotland GCM box as the predictor variable in (b), cf the original selection of the North England box (a). Bold figures show the (pooled) regression slopes and the standard deviations of associated residuals, and the correlation (r) between the GCM and RCM anomalies.

d. How does the downscaling error affect the probabilistic projections?

Figure 2.14 shows three examples of the impact on the 25km-scale pdfs of correcting the downscaling indexing error. These cases correspond to the scatter plots of Fig. 2.9, recalling that the GCM predictor box remains unchanged in these examples. As in Fig. 2.12, we isolate the effect of the downscaling correction by comparing the new projections against interim projections containing the developments of Sections 2.1-2.4 but lacking the downscaling correction. For Tmean in London, the new October pdf (Fig. 2.14a) is similar to the interim pdf but shows slightly less spread. This is because the effect of reduced residual error variance in the corrected downscaling outweighs the small increase in slope (Fig. 2.9a, b). Since the corrected slope is only slightly below unity, the new downscaled pdf also follows closely the pdf of changes for South England, the GCM predictor box (compare grey and orange curves in Fig. 2.14a).

In the other two cases (Figs. 2.14b, c) the ranges in the local pdfs increase when the downscaling correction is included, because the effects of larger slopes exceed those of reduced residuals. In the precipitation example (Glasgow in September, Fig. 2.14b), the dry tail in the new downscaled pdf (orange curve) therefore extends to larger reductions compared to the pdf obtained when the downscaling correction is withheld (blue curve). However, the new pdf still shows a less marked dry tail than that of the GCM predictor distribution. This occurs because the corrected downscaling slope of 0.71, while larger than the 0.30 obtained originally, still implies some attenuation of the long-term climate change signals in the GCM data.

In the sea level pressure example (Leeds in winter, Fig. 2.14c), the downscaling error effectively removed the relationship between the GCM and RCM circulation anomalies in the original calculations, indicated by a slope close to zero. The uncertainties captured by the GCM predictor pdf were therefore excluded from the downscaled pdf, which was consequently much narrower. In contrast the new local pdf follows that of the GCM predictor quite closely because the corrected slope is ~1.0. Uncertainties in the GCM-scale responses are thus fully reflected in the new downscaled distribution.

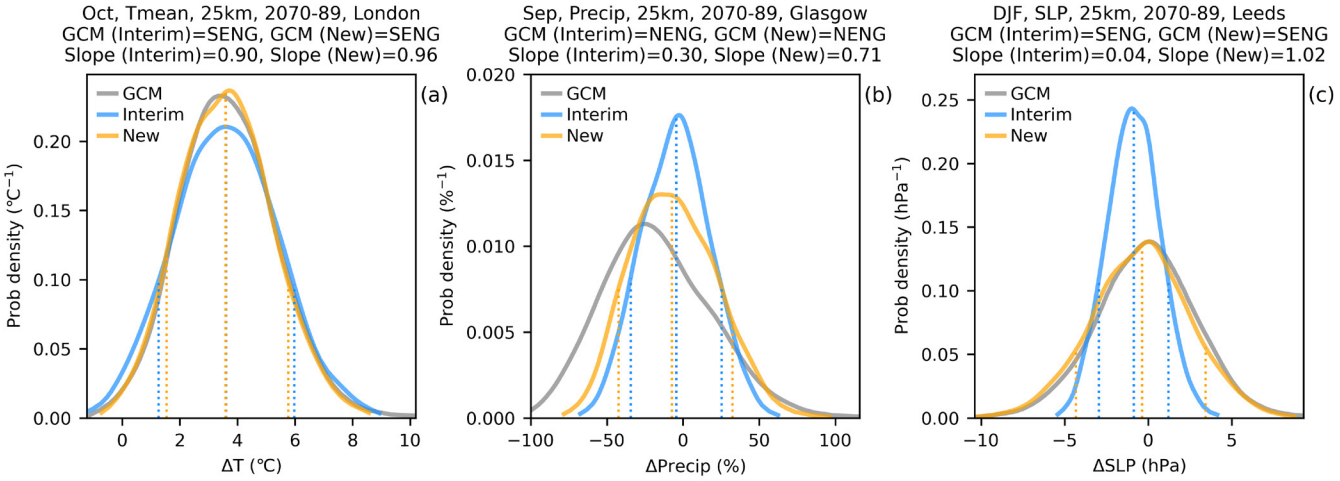


Figure 2.14. Probability distributions showing interim (blue) and new (orange) projections of changes for 2070-2089 relative to 1981-2000 for RCP85 at specific OSGB (25km) locations, for (a) Tmean (°C) in October for London, (b) precipitation (%) for Glasgow in September and (c) sea-level pressure (hPa) for Leeds in winter. Grey curves show corresponding pdfs for the GCM predictor boxes used for downscaling to the target locations. These are South England (a and c) and North England (b). In each case the same boxes are selected in the interim projections containing the downscaling software error, and the new projections in which the error is fixed. Dashed lines show P10, P50 and P90 values.

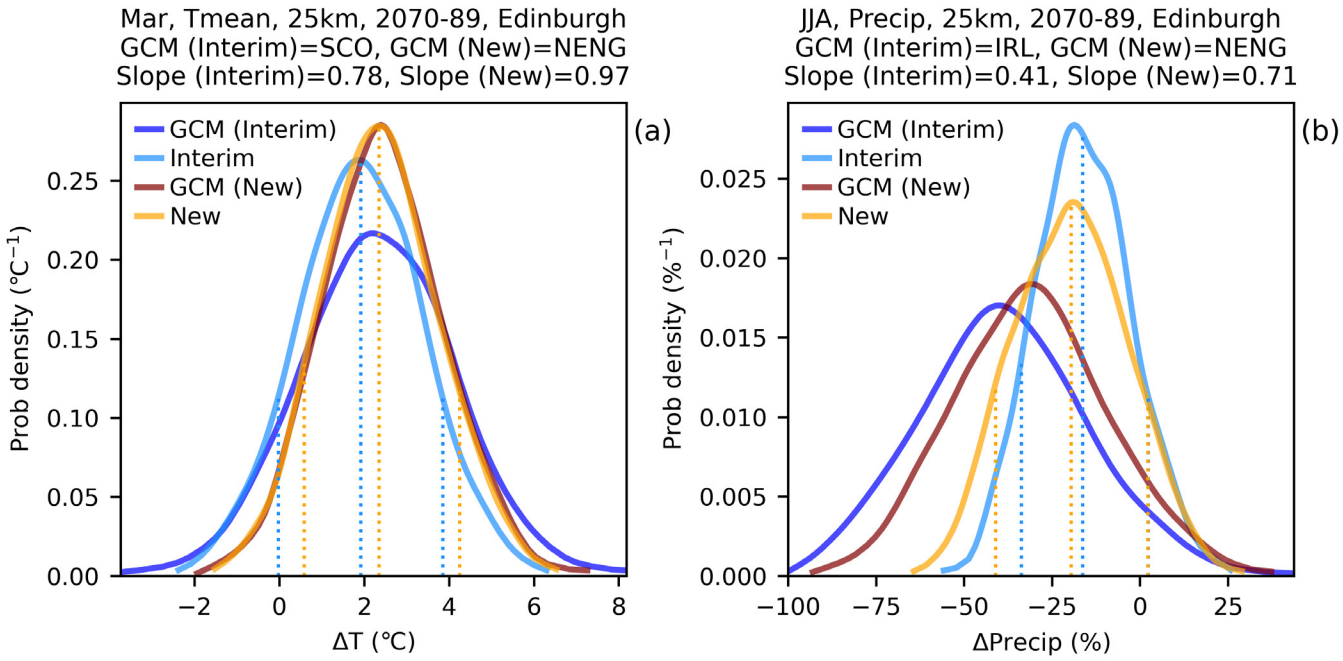


Figure 2.15. As Figure 2.14, for two cases at a 25km box (Edinburgh) where the GCM predictor box changes on correcting the downscaling software error. For Tmean in March, the predictor box changes from Scotland to North England (a), and for precipitation in summer (b) it changes from Ireland to North England. The original and corrected GCM selections are shown in dark blue and maroon.

In cases where a different GCM predictor is selected in the corrected downscaling, this introduces an additional factor in understanding differences between the new and interim 25km-scale pdfs. Two examples from the Edinburgh OSGB box are shown in Figure 2.15, to illustrate how changing the predictor can influence the local results from a methodological standpoint. For Tmean in March (a), correcting the downscaling error shifts the new pdf warmer compared to its interim counterpart, due mainly to the increase in regression slope from 0.78 to 0.97. The North England GCM predictor selected in the corrected downscaling shows a narrower P10-P90 range than the Scotland box chosen when the error is retained (maroon cf dark blue curves). This combines with the effect of smaller regression residuals to offset the broadening influence of increased slope in the corrected calculations, leading to a slightly reduced P10-P90 range in the new pdf for Edinburgh compared with the interim version (orange cf light blue curves).

For Edinburgh precipitation in summer, the GCM predictor changes from Ireland to North England (Fig. 2.15b). In both the interim and new pdfs, the GCM predictors show a more pronounced drying than the downscaled pdfs. However, the difference between the GCM and local pdfs is smaller in the new projection, because correcting the downscaling error increases the regression slope from 0.41 to 0.71. Consequently, the new pdf for Edinburgh has a stronger dry tail than the interim pdf, due to the increase in regression slope. However, the difference is mitigated somewhat by the change of GCM predictor, since the North England pdf shows slightly less drying than the Ireland pdf selected in the presence of the downscaling error. Other examples of the effects of changing the GCM predictor (which can vary in both sign and magnitude) are discussed below (see Figs. 2.18 and 2.19), and also in Section 3.3 which compares maps of future response in the original and new projections.

In analysis of regional changes, users can choose either the aggregated region pdfs or the 25km-scale pdfs. This raises the question of how correcting the downscaling error influences the level of consistency between the two products. We assess this by comparing pdfs for aggregated regions (in which co-located 25km RCM data is spatially averaged prior to downscaling), with distributions created by aggregating co-located 25km-scale pdfs. These are made by averaging the probability densities found for each individual OSGB box contained within a given aggregated region. We refer to these as OSGB-AGG distributions.

As an example, Figure 2.16a shows precipitation changes for the East Midlands region in April, showing aggregated-region and OSGB-AGG distributions for the interim and new projections. In both cases, the South England GCM box is used as the predictor variable for all the constituent OSGB boxes, as well as for the aggregated-region pdfs. Three of the four distributions are very similar. The exception is the OSGB-AGG distribution (dark blue) for the interim projections, which is the only distribution influenced by the downscaling indexing error. In this distribution, the error leads to an erroneously small regression slope, resulting in less spread than in the other distributions.

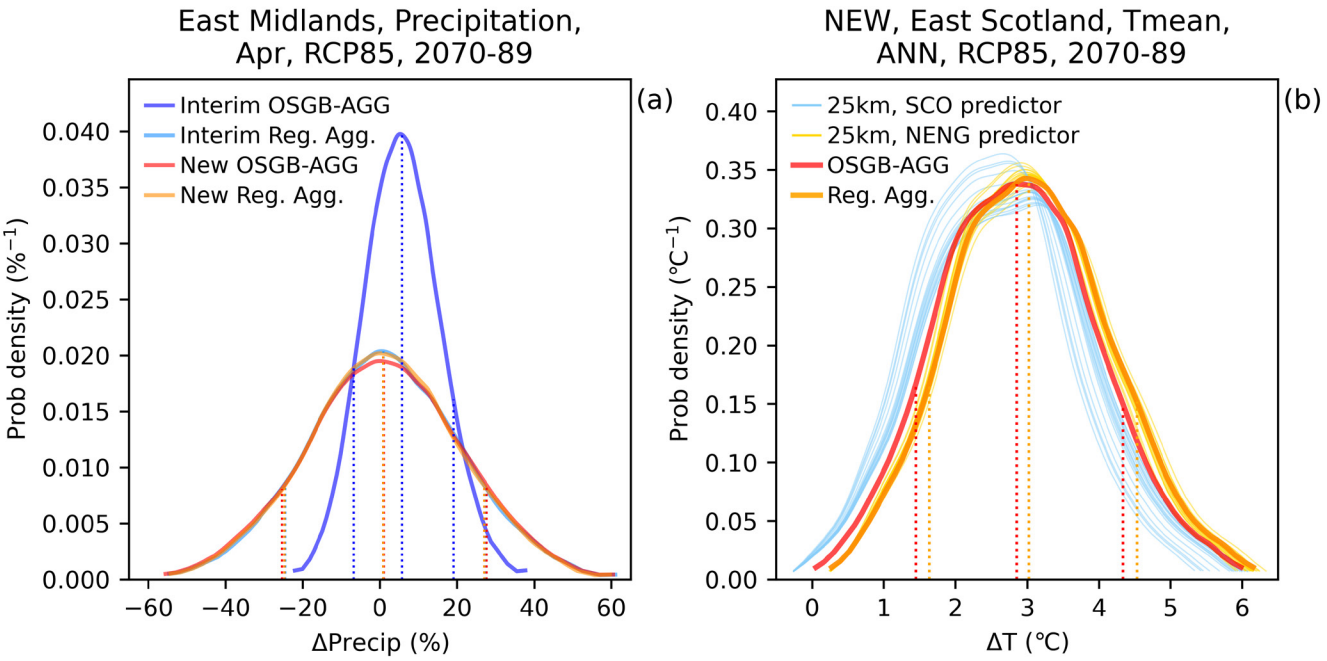


Figure 2.16. (a) Probability distributions from the new (orange) and interim (light blue) projections for changes in precipitation (%) in the East Midlands aggregated region. The dark blue (interim projections) and red (new projections) curves show corresponding 25km-scale results (OSGB-AGG), obtained by averaging probability densities across pdfs for all OSGB grid boxes contained within the East Midlands region. Dashed lines show P10, P50 and P90 values in each case. (b) Changes in annual Tmean for the East Scotland aggregated region, for the same period and scenario. Only the new projections are shown in this case. The thin lines show pdfs for every OSGB box within East Scotland, for cases where the GCM predictor is selected as Scotland (light blue) or North England (light orange). The thick orange curve shows the pdf for the aggregated region, and the red curve shows OSGB-AGG probability densities derived from all OSGB boxes within East Scotland.

Figure 2.17a extends this comparison of April precipitation changes to a set of aggregated regions. Differences between the OSGB-AGG and aggregated-region distributions are shown for P10, P50 and P90. In the interim projections containing the downscaling error, differences for P10 and P90 show substantial positive and negative values respectively. This demonstrates that the narrower spread found at the 25km-scale in East Midlands (Fig. 2.16a) is a consistent feature across different parts of the UK. On correcting the downscaling error in the new projections, differences between P10, P50 and P90 values for the OSGB-AGG and aggregated-region pdfs consistently become much smaller, due mainly to increased spread in the 25km distributions. The examples of Figs. 2.16a and 2.17a show that correcting the error in the new projections brings the OSGB-AGG precipitation distributions into much closer agreement with the aggregated-region pdfs, significantly improving consistency between these and the 25km-scale pdfs.

A similar comparison is shown in Figure 2.17b, for annual Tmean. In the interim projections, the OSGB-AGG distributions show slightly higher P90 values than the aggregated-region pdfs and slightly smaller P10 values, indicating larger spread. In the new projections, correcting the downscaling error reduces the variance of regression residuals (see Fig. 2.10), which outweighs the effects of larger slopes and hence reduces slightly the spread in the 25km distributions. The brings the OSGB-AGG distributions into better agreement with the aggregated-region pdfs, percentile differences being generally reduced to less than 0.1 °C.

While the P10, P50 and P90 differences are smaller in the new projections, minor differences remain between the OSGB-AGG results and the aggregated-region pdfs. This is expected, partly because the characteristics of climate variability vary with spatial scale (e.g., Osborn and Hulme, 1997). Another factor is the univariate nature of our downscaling method, which requires selection of a single GCM predictor box for both individual OSGB locations and for aggregated-region variables (Section 2.5a). In some cases (e.g., Fig. 2.16a), the same GCM predictor point is selected for downscaling every OSGB box within a given

aggregated region, and (therefore) for downscaling the aggregated variable itself. In other cases, different GCM predictors are chosen at different OSGB locations and the aggregated variable is downscaled using the GCM box chosen most often for the 25km-scale pdfs.

One such case is annual Tmean for East Scotland. In this example, the Scotland and North England GCM boxes are both used in the 25km downscaling. Consequently, the OSGB pdfs fall into two distinct clusters in the new projections (Fig. 2.16b), distinguished by the choice of GCM predictor. The aggregated-region pdf (thick orange curve) is made using the North England predictor, and hence replicates closely the corresponding OSGB cluster (thin orange curves). However, it is shifted warm compared to the OSGB cluster using the Scotland predictor (light blue curves) and is also slightly warmer than the OSGB_AGG distribution (red curve), which lies between the two clusters. For this reason, the P10, P50 and P90 values in OSGB-AGG are ~0.2°C cooler than those of the aggregated-region pdf, somewhat larger than the differences found for other aggregated regions (Fig. 2.17b).

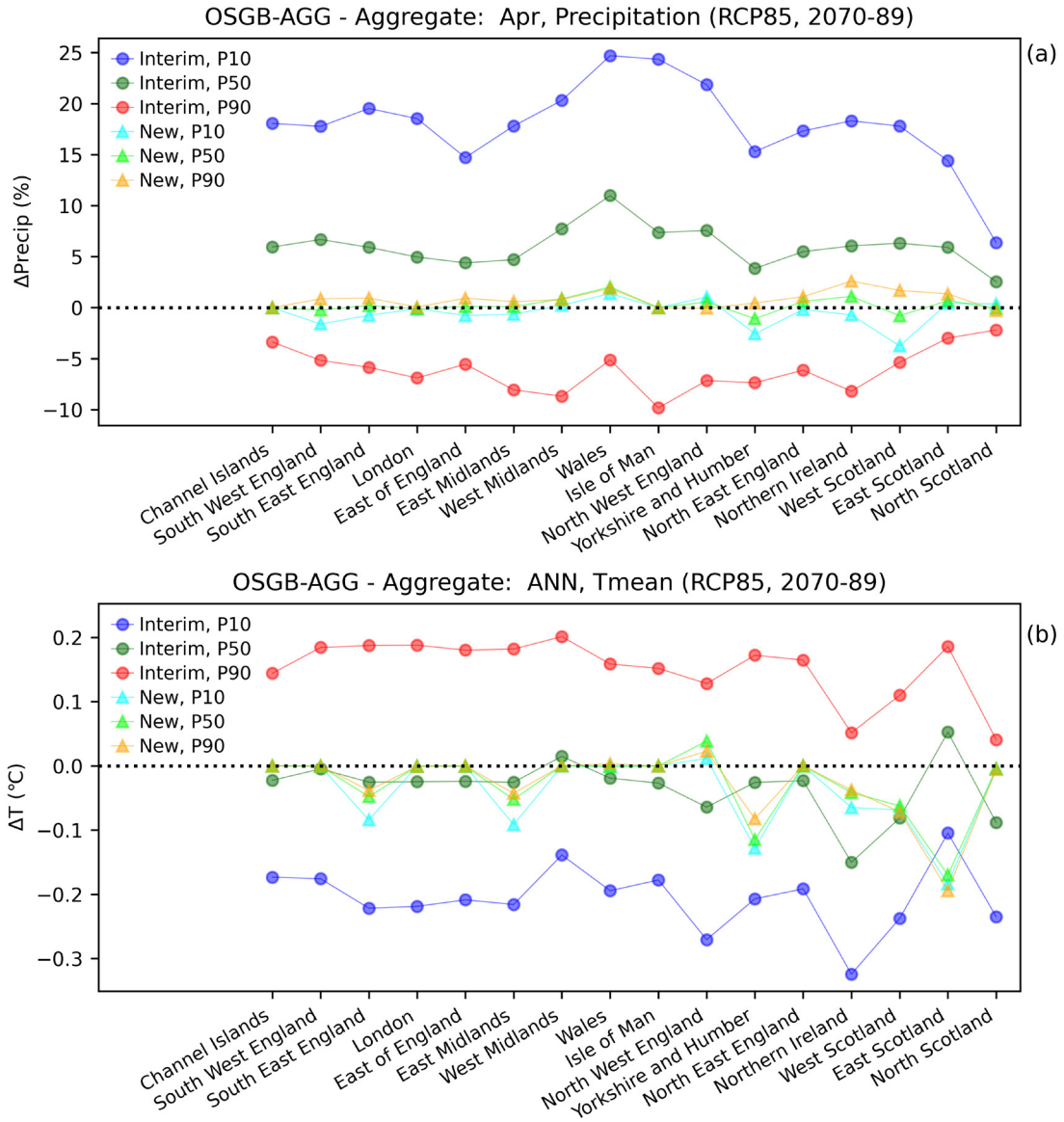


Figure 2.17. Differences between P10, P50 and P90 changes for OSGB-AGG 25km-scale distributions within a given aggregated region, and the corresponding aggregated-region pdfs. The OSGB-AGG distributions are obtained by averaging probability densities across local pdfs for all 25km grid boxes within the relevant region. Differences between anomalies for 2070-2089 relative to 1981-2000 under RCP85 are shown for the interim and new projections, which differ through inclusion of the downscaling error in the former. (a) April precipitation (%) and (b) Annual Tmean (°C), with differences for the interim (new) projections in dark (light) blue for P10, dark (light) green for P50 and red (orange) for P90.

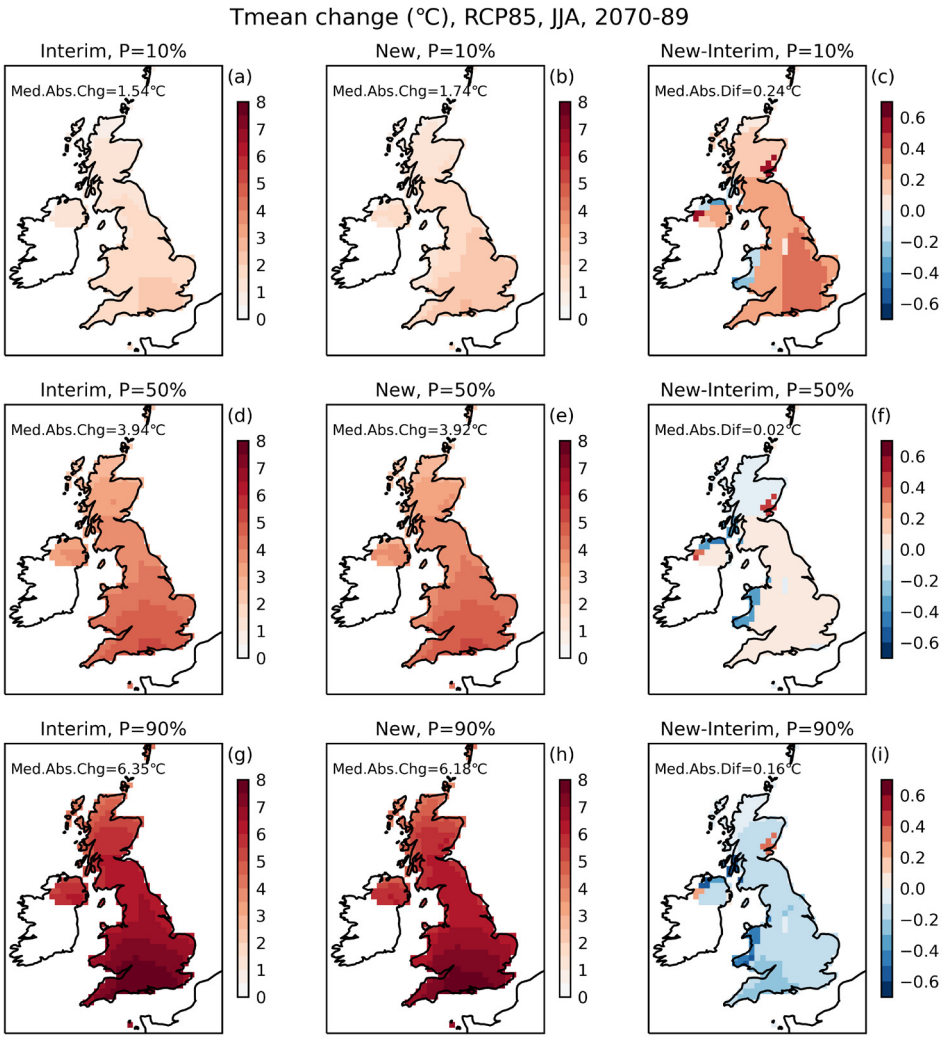


Figure 2.18. Changes in Tmean (°C) in summer for 2070-2089 relative to 1981-2000 under RCP85, showing P10, P50 and P90 for the interim (left) and new (centre) projections. Differences in response (right) show the impact of correcting the downscaling error.

For Tmean on the OSGB grid, a nationwide view of the impact of correcting the downscaling error is shown in Figure 2.18, using summer changes as an example. The impacts on the P50 responses are consistently small. The P10 (low-end) changes become slightly warmer in most places, while the P90 (high-end) changes generally become slightly cooler. Median absolute differences are 0.24°C (P10), 0.02°C (P50) and 0.16°C (P90), relatively small compared with the broad P90 – P10 ranges found in the pdfs. The results confirm a widespread but small reduction in spread in the new projections. This is driven (as discussed above) by reduced regression residuals in the corrected calculations. In a few locations the impacts of the downscaling correction are either larger than elsewhere, or of opposite sign. These occur where the correction leads to changes in the selected GCM predictor (Fig. 2.10g, h), including locations in East Scotland and Northern Ireland, and along the Welsh coastline. A change in GCM predictor[4] occurs at 11% of locations. However, the maximum impacts remain relatively modest compared with the projected changes, lying within the range ±0.6°C. Qualitatively similar conclusions apply to Tmean changes in other seasons (not shown). The P50 results of Fig. 2.18 can be compared with Fig. 3.8 in section 3.3, which compares summer Tmean responses in the new and original projections.

[4] The occurrence of changes in the GCM predictor box depends strongly on the variable and period of interest. In a few cases (sea level pressure in February, April or December), no changes occur at OSGB locations. On the other hand, changes occur at more than 70% of locations for annual downward short-wave radiation and for specific humidity in February and May. On average (across locations, months, seasons and the annual mean), frequencies of a change in the selected GCM box range from 13% (sea-level pressure) to 41% (specific humidity). The average frequencies for Tmean and precipitation are 32% and 25%.

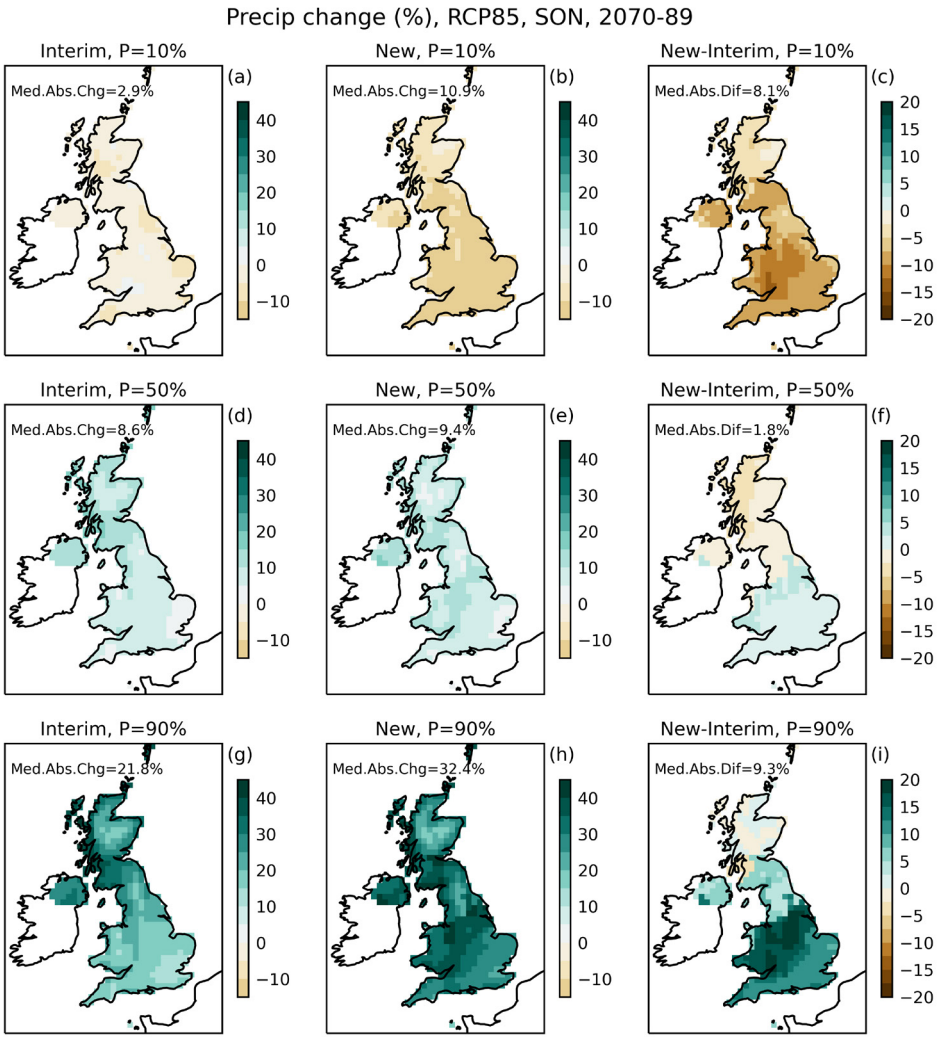


Figure 2.19. Changes in precipitation (%) in autumn for 2070-2089 relative to 1981-2000 under RCP85, showing P10, P50 and P90 for the interim (left) and new (centre) projections. Differences in response (right) show the impact of correcting the downscaling error.

Maps of local precipitation changes are shown in Figure 2.19, taking autumn as the example in this case. Correcting the downscaling error leads to a stronger dry-end (P10) response everywhere, while the wet-end (P90) responses increase over England, Wales, southern Scotland and Northern Ireland, indicating increased spread. As explained above, thjs is driven by increases in regression slope (Fig. 2.11), which (in contrast to the Tmean results) exert a larger impact than the accompanying reductions in downscaling residuals. Differences in the P50 response are generally smaller than those in P10 or P90, the downscaling error correction leading to slight reductions in the median response over Scotland, with small changes of mixed sign elsewhere.

Correcting the downscaling error leads to selection of a different GCM predictor at 27% of locations (Fig. 2.11g, h). However, differences between the new and interim pdfs are not necessarily at their largest in these places (in contrast to the Tmean results of Fig. 2.17). This is because the changes in regression slope are much larger in the case of precipitation and dominate the impacts of correcting the downscaling error to a greater degree. Effects of correcting the error in other seasons (not shown) are qualitatively similar, being characterised by consistent increases in spread accompanied by shifts in the median change of varying sign, dependent on location and season.

Summary of effects of correcting downscaling error on probabilistic projections at the 25km-scale

In Table 1 we summarise the impacts of correcting the downscaling error on the 25km-scale probabilistic projections, covering all the variables provided. The results show typical low (P10), median (P50) and high (P90) responses for the interim and new projections in winter and summer, for 2070-2089 relative to 1981-2000 under RCP85. These are derived by averaging P10, P50 and P90 values found at each OSGB location on the 25km grid. Therefore, the results represent typical local responses in the UK, rather than percentiles of a single pdf of aggregated UK changes.

Table 1. Average 25km-scale changes for 2070-2089 relative to 1981-2000 under RCP85 from the interim and new probabilistic projections.

Variable	Season	Interim projections			New projections			Fractional difference in median response	Fractional difference in spread
		P10	P50	P90	P10	P50	P90		
Daily maximum temperature (Tmax, °C)	Winter	0.98	2.69	4.49	0.84	2.51	4.28	(?) 0.03	(-) 0.06
	Summer	1.37	4.26	7.14	1.53	4.27	7.02		
Daily minimum temperature (Tmin, °C)	Winter	0.86	2.88	5.18	0.71	2.72	4.98	(?) 0.02	(-) 0.07
	Summer	1.55	3.66	5.74	1.79	3.64	5.55		
Daily mean temperature (Tmean, °C)	Winter	0.97	2.73	4.63	0.82	2.55	4.42	(?) 0.02	(-) 0.07
	Summer	1.53	3.92	6.32	1.74	3.92	6.16		
Precipitation (%)	Winter	-0.4	17.4	38.5	-1.7	20.1	46.6	(?) 0.05	(+) 0.25
	Summer	-47.4	-24.3	2.7	-50.6	-25.8	2.3		
Specific humidity (%)	Winter	3.3	19.6	36.4	3.6	19.5	35.9	(-) 0.03	(-) 0.08
	Summer	4.5	19.0	34.5	6.3	18.9	33.1		
Total cloud amount (%)	Winter	-2.9	0.2	3.1	-4.5	-0.4	4.0	(?) 0.05	(+) 0.26
	Summer	-23.5	-10.3	2.5	-26.4	-10.9	3.9		
Sea-level pressure (hPa)	Winter	-3.08	-0.97	1.13	-5.14	-0.90	3.19	(+) 0.13	(+) 0.54
	Summer	-1.23	0.10	1.43	-0.85	1.55	4.28		
Total downward short-wave radiation (surface, Wm^{-2})	Winter	-2.7	-0.8	1.1	-4.3	-1.0	2.1	(?) 0.04	(+) 0.33
	Summer	-3.5	13.1	31.3	-7.9	13.8	36.9		
Net downward long-wave radiation (surface, Wm^{-2})	Winter	-1.4	1.2	3.7	-2.0	1.1	4.3	(?) 0.04	(+) 0.23
	Summer	-8.6	-1.9	4.6	-10.0	-1.6	6.3		
Net downward short-wave radiation (surface, Wm^{-2})	Winter	-1.8	-0.2	1.4	-3.3	-0.6	2.1	(?) 0.05	(+) 0.30
	Summer	-3.0	11.9	28.4	-6.2	11.7	31.1		

P10, P50 and P90 values are calculated at each 25km box on the OSGB grid, and then averaged to provide the values shown in the coloured columns. The white columns show non-dimensional metrics F_{med} and F_{spread}, which measure the typical magnitude of the differences in P50 and the P10 – P90 range respectively, that result from correcting the downscaling error. See text and equations (3) and (4) for details. The P10, P50 and P90 changes are supplied for winter and summer, while the F_{med} and F_{spread} values are averages over all OSGB boxes and all 17 periods (monthly, seasonal and annual) for which pdfs are provided. The F_{med} and F_{spread} values are accompanied by symbols in brackets indicating the typical signs of local fractional differences in P50 and spread. The plus and minus signs indicate that positive or negative differences occur in more than two-thirds of cases, with "?" marked otherwise.

We also use two non-dimensional metrics to quantify the magnitude of the differences in response caused by correcting the downscaling error. The first diagnostic is the fractional difference in median response, defined as:

$$F_{med} = \frac{|\,P50_{new} - P50_{interim}\,|}{(P90\text{-}P10)_{new}} \quad (3)$$

This measures the absolute difference in the median (P50) response, normalised by the spread of responses in the new pdf. The second diagnostic is the fractional difference in spread, defined as:

$$F_{spread} = \frac{|\,(P90\text{-}P10)_{new} - (P90\text{-}P10)_{interim}\,|}{(P90\text{-}P10)_{new}} \quad (4)$$

This measures the absolute difference in spread, again normalised by the spread in the new pdf. We calculate local values of F_{med} and F_{spread} at each OSGB box, for each of our seventeen period definitions (12 months, four seasons plus the annual mean). These are then averaged to produce a single number for each variable in Table 1, that characterises the typical impact of the downscaling correction in a form comparable across variables.

The largest changes in P50 occur for sea-level pressure, for which correction of the downscaling error allows the regressions to pick up large-scale signals from the GCM simulations that were previously obscured (see Figs. 2.9 and 2.14, and related discussion). In summer, the average P50 response changes from 0.1 to 1.55 hPa in the new pdfs. This is consistent with a predominant shift to the positive phase of the summer North Atlantic Oscillation in the majority of the GCM projections used in the methodology (Murphy et al., 2018). For other variables the differences in P50 are (on average) 5% or less compared to the P90 – P10 ranges found in the new pdfs, and generally of variable sign. Since the differences in the median response vary with location (e.g., Figs. 2.18 and 2.19), the impact of the downscaling correction will be larger in some places and seasons than shown by the average values of Table 1.

The spread in the pdfs is changed substantially by the downscaling correction for several variables. The largest changes occur for sea-level pressure, with average fractional differences of 54% (Table 1). The changes are predominantly increases. They occur because the new pdfs now reflect uncertainties found in the GCM circulation changes (e.g., Fig. 2.14), as well as the median signals of change. In summer, the average P10 and P90 values are -0.85 and 4.28 hPa respectively, consistent with ranges of GCM responses that span reductions in the summer North Atlantic Oscillation to increases considerably larger than the median value (Murphy et al., 2018).

Fractional differences in spread reach 30% or more (on average) for net and downward surface short-wave radiation, and exceed 20% for precipitation, cloud cover and net downward surface long-wave radiation. In all these variables, the corrected spread increases in most cases. The increases arise from significantly stronger regression slopes driven (as in the precipitation examples discussed above) by the corrected

representation of GCM-RCM relationships that were previously partially masked by the downscaling error. The impacts on spread vary with location and season (not shown). For example, the average F_{spread} for precipitation in summer is 0.09, indicating smaller increases in spread compared to the average value of 0.25 obtained across all periods (Table 1).

For Tmean, Tmax, Tmin and specific surface humidity, changes in spread are smaller, and dominated by reductions. These variables show impacts qualitatively similar to the Tmean examples discussed above, in which reduced downscaling residuals exert the main influence on the change in spread. These variables all show relatively small increases in slope, because (like the Tmean examples shown in Figs. 2.9 and 2.13) the GCM and RCM projections show substantial ensemble-mean changes that explain a large fraction of the anomaly variance in the training datasets used to calibrate downscaling relationships. The ensemble-mean component of variance is not affected by the indexing error (see discussion of Fig. 2.9), leading to smaller changes in slope when the error is corrected.

Summary of effects of correcting downscaling error on probabilistic projections for aggregated regions

For aggregated regions, the differences between the interim and new projections are generally smaller, because the downscaling error in the original calculations only affected the choices of GCM predictor (Section 2.5b). In particular, average differences in spread are considerably smaller because these were mainly driven in the 25km-scale projections by correction of the indexing error, whereas the indexing was correct in the original regressions for aggregated-region variables. For example, the values of F_{spread} (averaged across all 43 aggregated regions and across the seventeen period definitions) are 0.01 for sea-level pressure, 0.04 for precipitation and 0.05 for total downward short-wave radiation, compared with 0.54, 0.33 and 0.25 respectively in the corresponding 25km-scale results. Average values of F_{med} are also smaller for aggregated regions, never exceeding 0.03. For sea-level pressure, the average value drops from 0.13 for the 25km results (Table 1) to 0.01 for aggregated regions.

The average P10, P50 and P90 values for aggregated regions (not shown) are very similar to their 25km-scale counterparts in the new projections, whereas values in the interim projections differ more markedly. For example, in the new projections the average P90 for slp in summer is 4.28hPa in both the aggregated-region and 25km results. In the interim projections, the P90 values are 4.27hPa for the aggregated regions (almost identical to the value in the new projections), but only 1.43hPa in the 25km results, due to the lack of spread caused by the indexing error. In the case of precipitation, Fig. 2.16a shows how the indexing error can drive a clear distinction between the interim and new pdfs at 25km scale, while the corresponding aggregated-region pdfs are essentially identical if the selected GCM predictor does not change. In winter, the average P90 (wet-end) changes are similar in the interim and new projections for aggregated regions (42.9% cf 44.6%) but differ more at 25km-scale (38.5% cf 44.6% - see Table 1).

2.6. Which Updates affect which Probabilistic Projections Variables?

Table 2 provides a summary showing how the updates of Sections 2.1-2.5 affect the variables for which probabilistic projections are provided. The larger effects are shown in normal font, with smaller effects in italics. For example, baseline centring affects all variables, and affects both the 3000 realizations and the pdfs. However, the effects on quantiles of the pdfs are small (compared with the spread in the original pdfs), whereas the effects on individual realizations can be somewhat larger (Fig. 2.8a cf 2.8b).

The downscaling software error described in Section 2.5 affects all variables, changing both the pdfs and the 3000 realizations.

The software error described in Section 2.1 affects Tmax and Tmin directly, so its impact on these variables can be relatively large in some cases (Figure 2.3). However, other variables can be affected indirectly. This is because (a) the calculations are based on a perturbed parameter ensemble (PPE) of simulations using a single GCM, and (b) a statistical emulator, trained on the available GCM outputs, is used to predict results for combinations of parameter values for which no GCM simulation is available (Appendix B). Since the emulator links multiple parameters to multiple climate variables, a change in one or more input GCM variables (in this case Tmax and Tmin) can affect emulated estimates of others through inter-variable relationships. These indirect effects will typically be small, since emulated estimates for a given variable (say precipitation) will be determined mainly by precipitation changes simulated by PPE members.

In practice variables are split into two batches, to make the statistical calculations computationally tractable. The first batch includes Tmax, Tmin, Tmean and precipitation, with the other six variables (Table 2) in a second batch. This means that the Tmax/Tmin software error can have indirect impacts on Tmean (e.g., Figure 3.8, discussed in Section 3) and precipitation, reflected in the relevant entries in Table 2. However, fixing this error has no impact on (say) cloud cover, since this variable is processed in the other batch. Similarly, the update to the treatment of precipitation extremes (Section 2.3) has an indirect effect on Tmax, Tmin and Tmean (as they are in the same batch as precipitation), but none on the other variables.

The improved time-smoothing (Section 2.2) applies to the pdfs and cdfs of annual variability and change for all variables, although this update does not affect the realizations. It also has no effect on the pdfs of 20- and 30-year average changes since these are constructed from time-averaged realizations without pooling. In cases where other updates affect the pdfs, the impacts apply both to the annual distributions and those of multi-year average changes.

Table 2. Impacts of updates on specific variables in the realizations and probability distributions of the UKCP18 probabilistic projections. Main effects in normal font, secondary effects in italics (see text for details).

Variable	Update				
	Software error (downscaling)	Software error (Tmax and Tmin)	Improved time smoothing	Improved precipitation extremes	Centred baseline
Daily maximum temperature (Tmax)	Realizations and distributions	Realizations and distributions	Annual distributions only	*Realizations and distributions (minor)*	Realizations; *Distributions (minor)*
Daily minimum temperature (Tmin)	Realizations and distributions	Realizations and distributions	Annual distributions only	*Realizations and distributions (minor)*	Realizations; *Distributions (minor)*
Daily mean temperature (Tmean)	Realizations and distributions	Realizations and distributions (minor)	Annual distributions only	*Realizations and distributions (minor)*	Realizations; *Distributions (minor)*
Precipitation	Realizations and distributions	Realizations and distributions (minor)	Annual distributions only	Realizations and distributions	Realizations; *Distributions (minor)*
Specific humidity	Realizations and distributions	No impact	Annual distributions only	No impact	Realizations; *Distributions (minor)*
Total cloud amount	Realizations and distributions	No impact	Annual distributions only	No impact	Realizations; *Distributions (minor)*
Sea-level pressure	Realizations and distributions	No impact	Annual distributions only	No impact	Realizations; *Distributions (minor)*
Total downward short-wave radiation (surface)	Realizations and distributions	No impact	Annual distributions only	No impact	Realizations; *Distributions (minor)*
Net downward long-wave radiation (surface)	Realizations and distributions	No impact	Annual distributions only	No impact	Realizations; *Distributions (minor)*
Net downward short-wave radiation (surface)	Realizations and distributions	No impact	Annual distributions only	No impact	Realizations; *Distributions (minor)*

3. Comparison of New Projections with UKCP18

In this section we compare the new projections (including all five updates) against the original UKCP18 projections, using a selection of statistics and examples from the UKCP18 regions and variables. Section 3.1 compares the results in tabular form, considering the 25km OSGB projections and the aggregated national projections. Section 3.2 presents a selection of time-dependent pdfs ("plumes"), comparing the P10 (low), P50 (median) and P90 (high) quantiles for the original and new projections for different regions, scenarios, variables and meaning periods. Finally, Section 3.3 compares maps of response for the 25km gridded data. We focus on projections for winter and summer, considering mainly changes for 2070-2089 relative to 1981-2000 under RCP85. Additional information for Tmean and precipitation is provided in a companion document. See Harris et al. (2022), available from https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/guidance-science-reports. This extends the RCP85 information cited in this report to include spring and autumn and includes maps and tables for changes in 2040-2059 relative to 1980 under RCP45 emissions, a case with lower greenhouse gas emissions.

3.1. Tables comparing Original and New Projections

This section includes a set of tables providing quantitative comparisons between the new and original projections. Sub-section a considers the local pdfs on the OSGB (25km) national grid, and sub-section b provides corresponding statistics for the set of 43 aggregated regions. In both cases, the statistics are compiled as average results across all UK locations. In sub-section c we provide tables for the national aggregated-region definitions.

a. Probabilistic projections at the 25-km scale

We start by discussing UK averages of projected P10, P50 and P90 changes at 25km-scale boxes for 2070-2089 relative to 1981-2000 under RCP85. Average values from the original and new projections are provided for all probabilistic projection variables in Table 3. We also provide the F_{med} and F_{spread} diagnostics of equations (3) and (4). These measure respectively the difference in the P50 response and the P90 – P10 spread of responses, expressed in both cases relative to the spread in the new projections. In this case, results from the original UKCP18 projections replace results from the interim projections in the equations. The results from Table 3 show the combined influences of all five developments. These can be compared with results from Table 1, which isolates the contribution of correcting the downscaling error (this being the only difference between the interim and new projections). As in Table 1, the fractional impact metrics represent averages over all OSGB grid boxes and all 17 period definitions (monthly, seasonal and annual) for which the probabilistic projections are provided.

Table 3. Average 25km-scale changes for 2070-2089 relative to 1981-2000 under RCP85 from the original and new probabilistic projections.

Variable	Season	Original projections			New projections			Fractional difference in median response	Fractional difference in spread
		P10	P50	P90	P10	P50	P90		
Daily maximum temperature (Tmax, °C)	Winter	0.90	2.76	4.72	0.84	2.51	4.28	(?) 0.06	(-) 0.17
	Summer	1.32	4.23	7.31	1.53	4.27	7.02		
Daily minimum temperature (Tmin, °C)	Winter	0.76	2.76	5.02	0.71	2.72	4.98	(?) 0.05	(-) 0.16
	Summer	1.56	3.48	5.69	1.79	3.64	5.55		
Daily mean temperature (Tmean, °C)	Winter	0.96	2.86	4.79	0.82	2.55	4.42	(?) 0.04	(-) 0.09
	Summer	1.38	3.82	6.47	1.74	3.92	6.16		
Precipitation (%)	Winter	-0.5	18.7	40.5	-1.7	20.1	46.6	(?) 0.05	(+) 0.27
	Summer	-49.6	-25.5	-0.4	-50.6	-25.8	2.3		
Specific humidity (%)	Winter	3.9	19.6	36.9	3.6	19.5	35.9	(-) 0.03	(-) 0.10
	Summer	3.1	18.7	35.4	6.3	18.9	33.1		
Total cloud amount (%)	Winter	-2.8	0.1	2.9	-4.5	-0.4	4.0	(?) 0.05	(+) 0.28
	Summer	-23.1	-10.3	2.5	-26.4	-10.9	3.9		
Sea-level pressure (hPa)	Winter	-3.23	-1.04	1.07	-5.14	-0.90	3.19	(+) 0.13	(+) 0.55
	Summer	-1.19	0.12	1.42	-0.85	1.55	4.28		
Total downward short-wave radiation (surface, Wm^{-2})	Winter	-2.6	-0.8	1.0	-4.3	-1.0	2.1	(?) 0.05	(+) 0.35
	Summer	-3.3	13.3	30.7	-7.9	13.8	36.9		
Net downward long-wave radiation (surface, Wm^{-2})	Winter	-1.3	1.2	3.7	-2.0	1.1	4.3	(?) 0.05	(+) 0.25
	Summer	-8.2	-1.7	4.5	-10.0	-1.6	6.3		
Net downward short-wave radiation (surface, Wm^{-2})	Winter	-1.7	-0.2	1.4	-3.3	-0.6	2.1	(?) 0.05	(+) 0.32
		-2.9	11.4	26.9	-6.2	11.7	31.1		

P10, P50 and P90 values are calculated at each 25km box on the OSGB grid, and then averaged to provide the values shown in the coloured columns. The white columns show the non-dimensional metrics F_{med} and F_{spread}, which measure the typical magnitude of the differences in P50 and the P10 – P90 range respectively, between the original and new projections. See Section 2.5d, noting that here the original percentile values replace the "interim" percentile values in the application of equations (3) and (4). The P10, P50 and P90 changes are supplied for winter and summer, while the F_{med} and F_{spread} values are averages over all OSGB boxes and all 17 periods (monthly, seasonal and annual) for which pdfs are provided. The F_{med} and F_{spread} values are accompanied by symbols in brackets indicating the typical signs of the fractional differences in P50 and spread. The plus and minus signs indicate that positive or negative differences occur in more than two-thirds of cases, with "?" marked otherwise.

For the three surface temperature variables, differences in the P50 climate changes are typically modest (~5% of those in the original projections), and of varying sign. The spread reduces in most cases, with average changes of 17% for Tmax, 16% for Tmin and 9% for Tmean. For Tmax and Tmin these differences are driven mainly by correction of the Tmax/Tmin indexing error (Section 2.1), while correcting the downscaling error typically contributes about a third of the difference (compare Tables 1 and 3). The downscaling error is the main driver of differences for Tmean, which is affected only indirectly by the Tmax/Tmin error (Section 2.6). In winter, the average upper-end (P90) warming reduces from 4.72°C to 4.28°C for Tmax. In summer the average P90 response reduces from 7.31°C to 7.02°C while the average P10 response increases from 1.32°C to 1.53°C. This reduces the uncertainty ranges in the new projections, although the spread values remain considerable.

In general, the main impact on precipitation is increased spread (amounting to a 27% difference on average), with differences in the P50 response typically 5% and of variable sign. Overall, the downscaling correction is the dominant driver of the larger ranges of change in the new projections (Table 3 cf Table 1). The increases in spread vary through the annual cycle, with seasonally-specific values of F_{spread} amounting to 0.15 in winter, 0.44 in spring, 0.11 in summer and 0.30 in autumn. In spring, the average value of P10 changes from -9.4% in the original projections to -20.0% in the new results. The average value of P90 changes from 13.7% to 21.0% (Harris et al., 2022).

The values are smaller in the solstitial seasons because these have the largest median changes in precipitation (strong increases in winter and strong reductions in summer), whereas average P50 changes in spring and autumn are smaller (-0.5% and 9.2% respectively, in the new projections). As explained in Section 2.5c, fixing the downscaling error has its largest impacts when uncertainties in projected changes are large in comparison to the median signals of projected change. This is because the error only affected the component of change that varies across members of the GCM and RCM ensembles used to calibrate the downscaling relationships.

In summer, the improved treatment of precipitation extremes (Section 2.3) also plays a role, tending to offset increases in spread driven by the downscaling correction. In the absence of the downscaling correction, this update drives an increase in the average P10 from -49.6% in the original projections to -47.4% (Table 1), indicating slightly reduced drying at the low ends of the 25km distributions. Adding the downscaling correction then increases the low-end drying. The net result is a slightly drier average value of P10 in the new projections (-50.6%, Table 3), compared to the original results.

For the other six variables, the typical differences between the new and original 25km-scale pdfs are dominated by the effects of the downscaling correction (Table 3 cf Table 1). The average magnitude of differences in the P50 response is largest for sea-level pressure (13%) and modest in other cases, amounting to 3% for specific humidity and 5% otherwise. Average differences in spread reach 55% for sea-level pressure and exceed 20% for total cloud amount and the surface radiation variables, with increases dominating.

Specific humidity is an exception, since there are typically substantial increases in P50 (on average 19.5% in winter and 18.9% in summer in the new projections – Table 3), driven by the increase in moisture-holding capacity of the atmosphere resulting from warming temperatures (e.g., Held and Soden, 2000). In this case, the spread changes follow those for Tmean, amounting to 10% on average and being dominated by reductions.

Correcting the downscaling error leads to two effects on spread: (a) increases in the regression slope linking large-scale GCM changes to local RCM changes, which increase spread; (b) reductions in the associated residual errors, which reduce spread. For the temperature variables and specific humidity, the reductions in residual uncertainty dominate because increases in regression slope are relatively small (Fig. 2.9 and discussion in section 2.5c). For other variables the increases in regression slope are larger, because signals of future climate change common to all simulations (and hence unaffected by the downscaling indexing error) are smaller in comparison to the uncertain component of the changes (which was obscured in the original projections by the downscaling error). Therefore, the variables for which increases in spread are typically large are those in which increases in regression slope dominate the effects of smaller residuals. In such cases, correcting the downscaling error allows uncertain responses simulated in the GCM to be properly reflected in the new local pdfs (e.g., Fig. 2.14c, discussed in Section 2.5d).

b. Probabilistic projections for aggregated regions

The downscaling error was present in the original regressions only for 25km-scale variables. It affects the downscaling for aggregated-region variables only where correcting the 25km-scale regressions leads to a different choice of GCM predictor. This occurs in some cases, because the selection of GCM predictor is based on the most frequent choice across 25km boxes contained within the relevant region (Section 2.5a). The frequency of selection of different GCM predictors varies widely, from 2% for sea-level pressure in summer, to 70% for the surface temperature variables in winter. In summer, the temperature selections change much less often (9% of regions), whereas for precipitation there are more changes in summer (63% of regions, cf 28% in winter).

Table 4 provides results for aggregated regions in the same format as the 25km-scale results of Table 3, considering the corresponding period (2070-2089 relative to 1981-2000) and emissions scenario (RCP85). Differences between the new and original projections are generally smaller for the aggregated regions, compared with the 25km-scle pdfs. Average values of F_{spread} are smaller than at 25km-scale for all variables, the differences being particularly marked for those variables possessing F_{spread} values exceeding 20% in their 25km-scale results.

For aggregated regions, the largest average F_{spread} values occur for Tmax (12%) and Tmin (11%), compared with 17% and 16% respectively in the 25km-scale results. Changes in spread for aggregated regions are driven mainly by correction of the Tmax/Tmin software error, as in the case of the 25km-scale results. Average values of F_{spread} for aggregated regions are also close to 10% for total cloud cover and the surface short-wave radiation variables, but smaller for other variables. In the cases of Tmax, Tmean and precipitation, the differences are not dominated by systematic increases or reductions, in contrast to the 25km results.

The average difference in median response between the new and original projections is much smaller for sea-level pressure for aggregated regions (F_{med} being 2%, cf 13% for the 25km-scale pdfs). For other variables, average values of F_{med} are 6% or smaller for both sets of regions, the values for aggregated regions never exceeding the corresponding 25km-scale values (Table 3 cf Table 4).

Table 4. Changes averaged across the 43 aggregated regions for 2070-2089 relative to 1981-2000 under RCP85, from the original and new probabilistic projections.

Variable	Season	Original projections			New projections			Fractional difference in median response	Fractional difference in spread
		P10	P50	P90	P10	P50	P90		
Daily maximum temperature (Tmax, °C)	Winter	1.18	2.73	4.45	0.84	2.51	4.27	(?) 0.06	(?) 0.12
	Summer	1.73	4.23	6.91	1.55	4.28	7.01		
Daily minimum temperature (Tmin, °C)	Winter	0.86	2.75	5.06	0.71	2.72	4.99	(?) 0.05	(-) 0.11
	Summer	1.69	3.50	5.53	1.81	3.64	5.53		
Daily mean temperature (Tmean, °C)	Winter	1.07	2.83	4.74	0.82	2.55	4.41	(+) 0.04	(?) 0.03
	Summer	1.69	3.86	6.11	1.76	3.93	6.15		
Precipitation (%)	Winter	-1.6	20.2	46.0	-1.1	19.6	44.6	(?) 0.03	(?) 0.07
	Summer	-51.5	-26.0	-1.1	-49.7	-25.5	1.8		
Specific humidity (%)	Winter	4.2	19.9	36.8	3.7	19.5	35.8	(?) 0.03	(-) 0.05
	Summer	5.0	19.0	34.1	6.3	18.8	32.8		
Total cloud amount (%)	Winter	-3.9	-0.2	3.5	-4.3	-0.3	3.8	(?) 0.03	(+) 0.10
	Summer	-23.5	-10.1	3.1	-26.2	-10.8	3.9		
Sea-level pressure (hPa)	Winter	-5.19	-1.11	2.87	-5.14	-0.91	3.18	(+) 0.02	(+) 0.04
	Summer	-0.94	1.44	4.29	-0.84	1.55	4.28		
Total downward short-wave radiation (surface, Wm^{-2})	Winter	-3.5	-0.9	1.7	-4.1	-1.0	2.0	(?) 0.03	(+) 0.09
	Summer	-6.6	13.3	34.8	-7.4	13.9	36.6		
Net downward long-wave radiation (surface, Wm^{-2})	Winter	-1.6	1.1	3.8	-1.7	1.2	4.2	(-) 0.03	(+) 0.07
	Summer	-9.3	-1.7	5.6	-9.8	-1.7	5.9		
Net downward short-wave radiation (surface, Wm^{-2})	Winter	-2.8	-0.5	1.8	-3.1	-0.5	2.0	(?) 0.03	(+) 0.09
	Summer	-4.0	11.9	29.0	-5.8	11.7	30.8		

P10, P50 and P90 values are calculated for each aggregated region and then averaged to provide the values shown in the coloured columns. The white columns show the non-dimensional metrics F_{med} and F_{spread}, which measure the typical magnitude of the differences in P50 and the P10 – P90 range respectively, between the original and new projections. See Section 2.5d, noting that here the original percentile values replace the "interim" percentile values in the application of equations (3) and (4). The P10, P50 and P90 changes are supplied for winter and summer, while the F_{med} and F_{spread} values are averages over all aggregated regions and all 17 periods (monthly, seasonal and annual) for which pdfs are provided. The F_{med} and F_{spread} values are accompanied by symbols in brackets indicating the typical signs of the fractional differences in P50 and spread. The plus and minus signs indicate that positive or negative differences occur in more than two-thirds of cases, with "?" marked otherwise.

Are the differences large enough to influence user applications?

The impacts on users of the differences between the new and original projections will depend on how big the differences are, and on how the pdfs are used in specific applications. For example, a user study focused on low-probability high-impact events may be most sensitive to differences in low- or high-end extremes of the relevant pdfs, whereas a study focused on (say) exceedance of a threshold could be sensitive to a shift in the distribution typified by a different P50 response, even in the absence of a change in spread.

For a given climate variable, the differences will vary with region and season. For example, differences in Tmax are largest in July (Figure 2.3). Under RCP85, the P50 responses for aggregated regions are (on average) 1.39°C warmer in the new projections, for 2070-2089 relative to 1981-2000. This amounts to an average F_{med} of 20%, cf 6% when F_{med} is averaged across all period definitions (Table 4). However, the average F_{med} is only 1% in the seasonal pdfs for summer, because differences in July are balanced by differences of opposite sign in adjacent months (Fig. 2.3a).

Therefore, we encourage users to compare the new and original results for their variables, regions, emissions scenarios, seasons and periods of interest. As a general guide, the average results of Tables 3 and 4 can be used to indicate variables and region definitions for which decision-relevant differences are more or less likely to occur, based on the F_{med} and F_{spread} diagnostics. To enable this, we assume that:

- Differences of less than 10% indicate that decisions, or assessed risks, are *unlikely* to be significantly affected by repeating the relevant study using the new projections.
- Differences in the range 10-20% indicate that the conclusions *may* be significantly affected by the new results.
- Differences exceeding 20% indicate that the conclusions are *likely* to be significantly affected by the new results.

For aggregated regions, Table 4 suggests that some studies using Tmax or Tmin may be affected, in cases where the decision or risk is influenced by the spread of the probability distributions (average F_{spread} being 12%). In studies influenced by changes in the median, the chance of a significant impact from the new projections is generally quite small, but potentially large in July (see F_{med} values quoted above). For other aggregated-region variables, the average F_{med} and F_{spread} values suggest that most studies sensitive to changes in the median response are unlikely to be significantly affected (since average F_{med} is invariably 5% or less), perhaps with a few specific exceptions. However, some studies sensitive to spread may be affected in the cases of Tmin, total cloud cover and the surface short-wave radiation variables, since average F_{spread} is close to 10%. In many applications, differences in both the P50 response and the spread (P90 – P10 range) may influence the relevant conclusions.

For the 25km-scale pdfs, applications sensitive to spread are (in general) likely to be affected by the new pdfs in studies involving sea-level pressure, precipitation, total cloud cover or the surface radiation variables, given that their average F_{spread} values exceed 20%. Some studies using Tmax and/or Tmin may also be affected, since average F_{spread} is 17% and 16% respectively. In general, studies influenced mainly by a shift in the 25km-scale pdfs are less likely to be affected (given that average F_{med} is 6% or less), but there could be specific exceptions. Many impacts studies use two or more climate variables (often including Tmean and precipitation). In these cases, users will need to assess the combined effects of differences in the relevant marginal pdfs.

Consistency between the 25-km scale and aggregated-region results

In general, the 25km results are more consistent with their aggregated-region counterparts in the new projections than in the original results. This can be seen by comparing average percentile values between Tables 3 and 4. For example, the new projections give average high-end (P90) changes in winter precipitation of 46.6% at 25km scale and 44.6% for aggregated regions. The difference of 1.8% is smaller than the corresponding difference of 6% in the original results, which gave average values of 40.5% at 25km scale and 46.6% for aggregated regions. For sea-level pressure in winter, the original projections showed an average difference of 2.0 hPa in the low-end (P10) changes, with -3.2hPa at 25km scale cf -5.2hPa for aggregated regions. In the new projections the average P10 values are identical, amounting to -5.1hPa in both the 25km and aggregated region results. For total downward short-wave radiation in summer, the average P90 changes in the 25km-scale and aggregated region results are only 0.3Wm^{-2} different (36.9 cf 36.6Wm^{-2}), whereas they differed by 4.1Wm^{-2} in the original results (30.7 cf 34.8Wm^{-2}). Occasional exceptions to this improved consistency occur for specific variables and regions, in cases where the selection of GCM predictor is a choice between two or more subsets of co-located 25km boxes in which different GCM boxes are chosen with roughly equal frequency (see Fig. 2.16b and related discussion).

c. Results for national aggregated regions

Tables 5-8 provide winter and summer projections for the five national aggregated regions, for Tmean and precipitation. The results compare the P10, P50 and P90 quantiles of the distributions for the original and new projections for the 20-year period 2070-89 in response to RCP26 and RCP85 emissions. Changes are calculated relative to the 1981-2000 baseline.

For Tmean, differences in P50 between the original and new projections are generally small. For the RCP85 summer response the difference is 0.11°C for Wales, and smaller in other cases. In winter, the RCP85 differences are smaller than 0.1°C for the UK and England, but larger for Wales (0.2°C), Northern Ireland (0.32°C) and Scotland (0.55°C). These larger changes are driven mainly by changes in the GCM predictor used for downscaling. For example, the predictor for Northern Ireland changes from South England in the original calculations to North England in the new projections, while that used for Scotland changes from North England to Scotland. The F_{med} values for Scotland and Northern Ireland are 0.15 and 0.10 respectively, showing that the impacts of the new developments can be larger than the average values discussed above, in specific aggregated regions.

For RCP26 the levels of warming are smaller (due to the smaller radiative forcing applied in this scenario), and the magnitudes of difference between the new and original projections are also smaller. The largest differences occur in winter (as for the RCP85 results), in cases where changes in GCM predictor occur.

The spread in the Tmean pdfs reduces in the new projections under both emissions scenarios, except for Scotland in summer under RCP85. However, all the differences are small, so the new projections show considerable uncertainties in projected warming, in common with the original UKCP18 results.

For winter precipitation differences in the national projections are small, apart from Northern Ireland (Table

7). Here, differences between the new and original projections result mainly from a change in the GCM predictor for downscaling, from Southern England to Northern England. For RCP85 this leads to a P50 increase of 21% (cf 15% in the original projections). This is accompanied by a 43% increase at P90 (cf 36% originally), resulting in enhanced spread in the new projections that encompass larger increases at the wet end of the probability distribution. In other nation regions, quantile-specific differences between the new and original projections are 2% or less.

In summer, the GCM predictor used for precipitation downscaling changes in all regions except Scotland. However, in most cases the new projections differ only modestly from the original results (Table 8), showing that changes in GCM predictor do not always drive substantial changes in the precipitation projections (see also Section 2.5d). Under RCP85, larger difference do occur in the P10 values for England (drying of 54% cf 61% originally) and Wales (drying of 55% cf 62% originally). Differences in the P50 response are 4% or smaller in all cases. Under RCP85, the P90 – P10 spread reduces from 61% to 52% for England (F_{spread} = 0.17) and increases from 40% to 46% for Northern Ireland (F_{spread} = 0.13), with smaller changes in other cases.

Table 5. Original and new projections for the national aggregated regions of changes in DJF Diurnal Mean Air Temperature (Tmean), for 2070-2089 relative to 1981-2000 under the RCP26 and RCP85 emissions scenarios. P10, P50, and P90 denote the 10th, 50th (median) and 90th percentiles of the relevant probability distributions.

Country	Version	RCP26 DJF Mean Surf Air Temp				RCP85 DJF Mean Surf Air Temp			
		P10	P50	P90	P90-P10	P10	P50	P90	P90-P10
United Kingdom	Original	0.04	1.17	2.35	2.31	1.13	2.91	4.83	3.70
	New	0.02	1.13	2.27	2.25	1.11	2.83	4.75	3.63
England	Original	0.04	1.17	2.35	2.31	1.13	2.93	4.86	3.72
	New	0.02	1.12	2.27	2.25	1.12	2.84	4.77	3.65
Scotland	Original	-0.00	1.09	2.22	2.22	0.94	2.68	4.64	3.70
	New	-0.28	0.80	1.91	2.19	0.34	2.13	4.02	3.64
Wales	Original	0.03	1.14	2.29	2.26	1.10	2.84	4.71	3.62
	New	-0.06	1.03	2.18	2.24	0.91	2.64	4.46	3.55
Northern Ireland	Original	0.04	1.09	2.19	2.15	1.03	2.68	4.45	3.41
	New	-0.04	0.95	1.97	2.02	0.78	2.36	4.06	3.28

Table 6. Original and new projections for the national aggregated regions of changes in JJA Tmean, for 2070-2089 relative to 1981-2000 under the RCP26 and RCP85 emissions scenarios.

Country	Version	RCP26 JJA Mean Surf Air Temp				RCP85 JJA Mean Surf Air Temp			
		P10	P50	P90	P90-P10	P10	P50	P90	P90-P10
United Kingdom	Original	0.71	1.73	2.85	2.14	1.99	4.09	6.24	4.24
	New	0.69	1.68	2.75	2.06	2.07	4.14	6.25	4.18
England	Original	0.76	1.85	3.06	2.30	2.14	4.40	6.71	4.57
	New	0.74	1.80	2.97	2.23	2.22	4.46	6.74	4.52
Scotland	Original	0.17	1.20	2.31	2.14	1.15	3.09	5.13	3.98
	New	0.14	1.17	2.23	2.09	1.19	3.13	5.18	4.00
Wales	Original	0.69	1.79	2.97	2.28	1.80	4.18	6.71	4.92
	New	0.69	1.75	2.88	2.19	1.89	4.29	6.69	4.80
Northern Ireland	Original	0.42	1.46	2.51	2.09	1.50	3.54	5.65	4.15
	New	0.44	1.39	2.46	2.02	1.60	3.62	5.71	4.11

Table 7. Original and new projections for the national aggregated regions of changes in DJF Precipitation Rate (%) for 2070-2089 relative to 1981-2000, for the RCP26 and RCP85 emissions scenarios.

Country	Version	RCP26 DJF Precipitation Rate				RCP85 DJF Precipitation Rate			
		P10	P50	P90	P90-P10	P10	P50	P90	P90-P10
United Kingdom	Original	-4	7	19	23	-2	16	38	40
	New	-3	7	18	21	-1	16	35	36
England	Original	-4	8	22	26	-2	19	44	46
	New	-3	8	21	24	-1	18	41	42
Scotland	Original	-5	7	21	26	-1	20	47	48
	New	-5	6	19	24	-2	18	43	45
Wales	Original	-6	7	22	28	-2	20	45	47
	New	-4	7	21	25	-2	18	42	44
Northern Ireland	Original	-4	7	18	22	-2	15	34	36
	New	-3	7	18	20	-1	14	31	32

Table 8. Original and new projections for the national aggregated regions of changes in JJA Precipitation Rate (%) for 2070-2089 relative to 1981-2000, for the RCP26 and RCP85 emissions scenarios.

Country	Version	RCP26 JJA Precipitation Rate				RCP85 JJA Precipitation Rate			
		P10	P50	P90	P90-P10	P10	P50	P90	P90-P10
United Kingdom	Original	-26	-14	-1	26	-47	-26	-5	42
	New	-26	-12	2	28	-43	-22	-2	41
England	Original	-34	-15	4	39	-61	-30	0	61
	New	-33	-15	3	36	-54	-29	-2	52
Scotland	Original	-24	-12	2	27	-44	-18	7	51
	New	-23	-10	3	26	-41	-19	6	47
Wales	Original	-36	-18	-1	35	-62	-34	-7	56
	New	-34	-16	2	36	-55	-31	-3	52
Northern Ireland	Original	-23	-10	3	27	-41	-21	-2	40
	New	-25	-10	6	31	-44	-22	2	46

3.2. Comparison of time-dependent uncertainties in future regional projections

Figure 3.1 compares uncertainty plumes for Tmean for the London administration region assuming RCP85 emissions. The downscaling correction does not influence this example, because the GCM predictor is unchanged in the new projections and the indexing error did not affect the calibration of the associated regression relationship. Panel (a) compares pdfs of annual Tmean data for the original projections (that uses an 11-year pooling window), with the new projections (that use the new 31-year pooling interval). The longer interval reduces the effects of sampling noise in the underlying model data and hence smooths time variations in the pdfs (Section 2.2), removing the spurious multidecadal variability present in the original projections. Panel (b) shows pdfs for 20-year average Tmax data. These are affected by the Tmax/Tmin software error and (to a minor degree) the updates concerning baselining and precipitation extremes, but not the smoothing update (Table 2). The new distribution shifts towards warmer changes, P50 being about 1.4°C warmer by the end of the century. This is because the cooler June response, previously misassigned to July by the Tmax/Tmin error, has now been corrected. The new Tmean responses in Figure 3.1a do not replicate the systematic upward shifts shown for Tmax in Figure 3.1b, because Tmean was affected only indirectly by the Tmax/Tmin error in the original results.

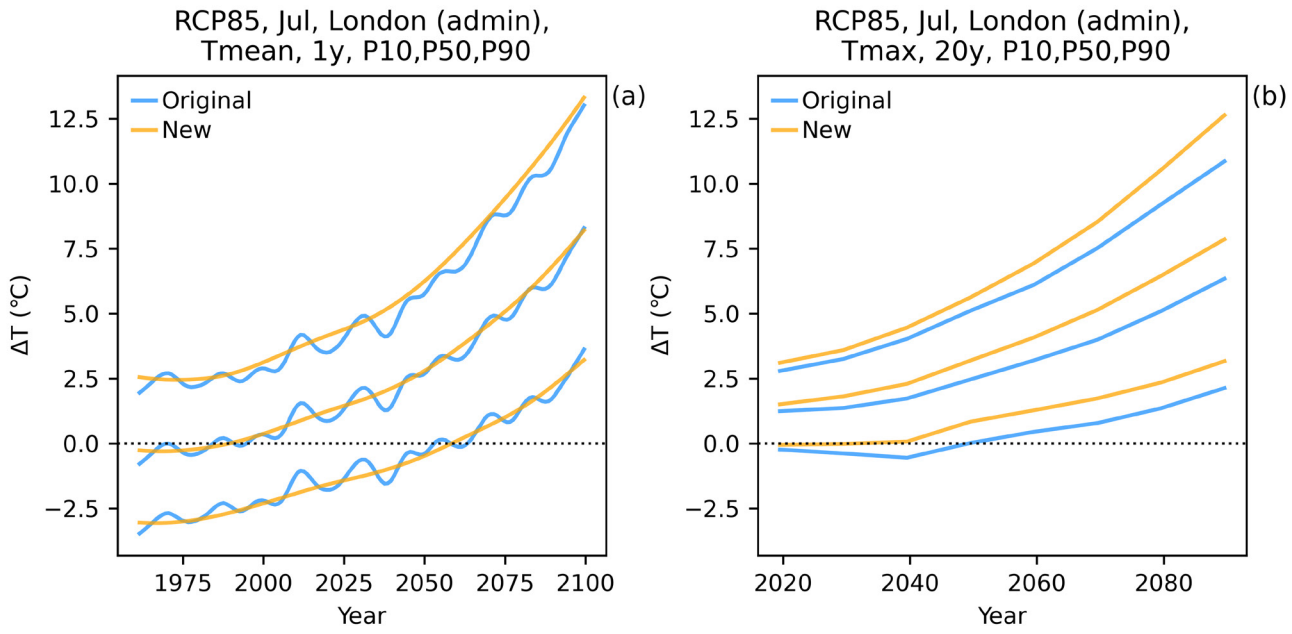


Figure 3.1. Comparison of the P10, P50 and P90 quantiles for the original UKCP18 pdfs (blue curves) and the new projections (orange curves) relative to the 1981-2000 baseline. (a) Change in annual values for daily mean surface air temperature (Tmean) in July for the London administration region in response to RCP85 emissions. Blue curves use an 11-year pooling window to construct the pdfs, orange curves use a 31-year window. (b) Change in 20-year mean July maximum temperature (Tmax) for the London administration region in response to RCP85 emissions. No pooling is required for 20-year means, so the time smoothing update affects only panel (a).

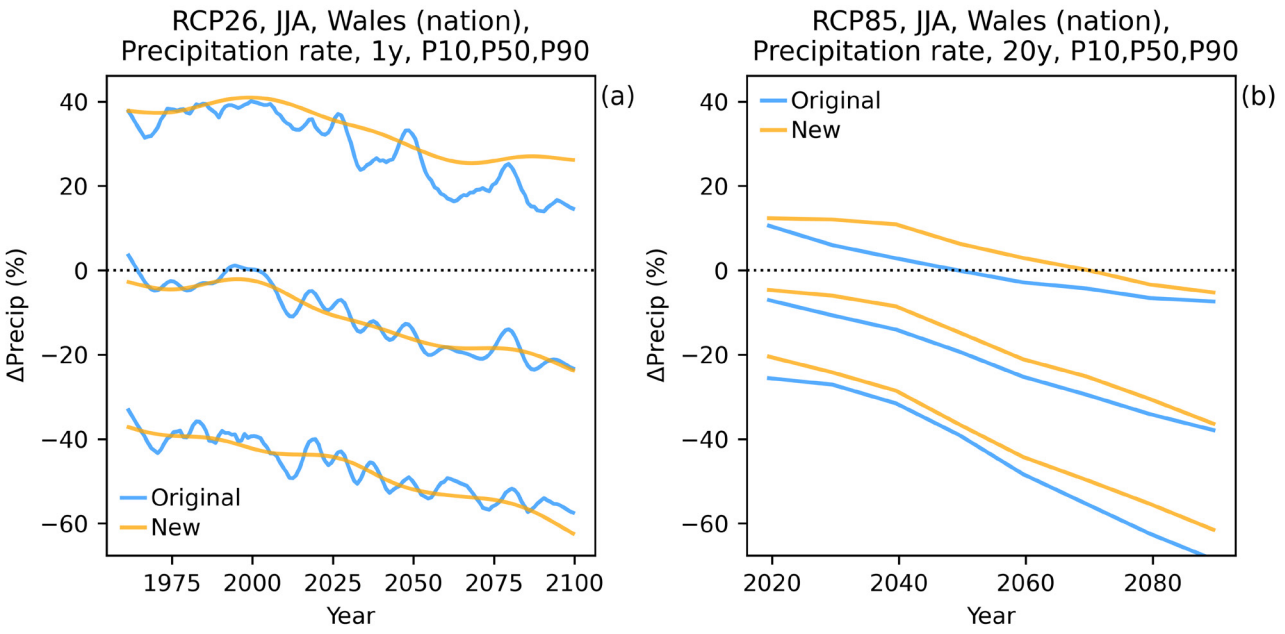


Figure 3.2. Comparison of the P10, P50 and P90 quantiles of summer precipitation changes for Wales in the original UKCP18 pdfs (blue curves) and the new projections (orange curves) relative to the 1981-2000 baseline. (a) Change in annual values in response to RCP26 emissions. Blue curves use an 11-year pooling window to construct the pdfs, orange curves use a 31-year window. (b) Change in 20-year mean values in response to RCP85 emissions.

A comparison of summer precipitation changes for Wales is presented in Figure 3.2. Panel (a) compares annual responses for the RCP26 emissions scenario, while panel (b) shows the 20-year mean responses for RCP85. The differences in the RCP26 pdfs mainly reflect the impact of a larger pooling window, which smooths time variations in the pdfs (as in Fig. 3.1a discussed above). There are also small systematic increases in P90, beyond 2040. For the 20-year mean RCP85 projections all quantiles shift systematically towards a less dry response, more so for the lower quantiles than for P90 (see also Table 8, discussed above). While the GCM predictor used in downscaling changes from Ireland to Wales for the new projections, the differences in P10 arise mainly from the improved treatment of precipitation extremes (Section 2.3). The differences in P50 are small relative to the uncertainty in response. For example, P50 for 2070-2089 increases by 3% (from -34% to -31%) while the P90-P10 range in the new distribution is 52%.

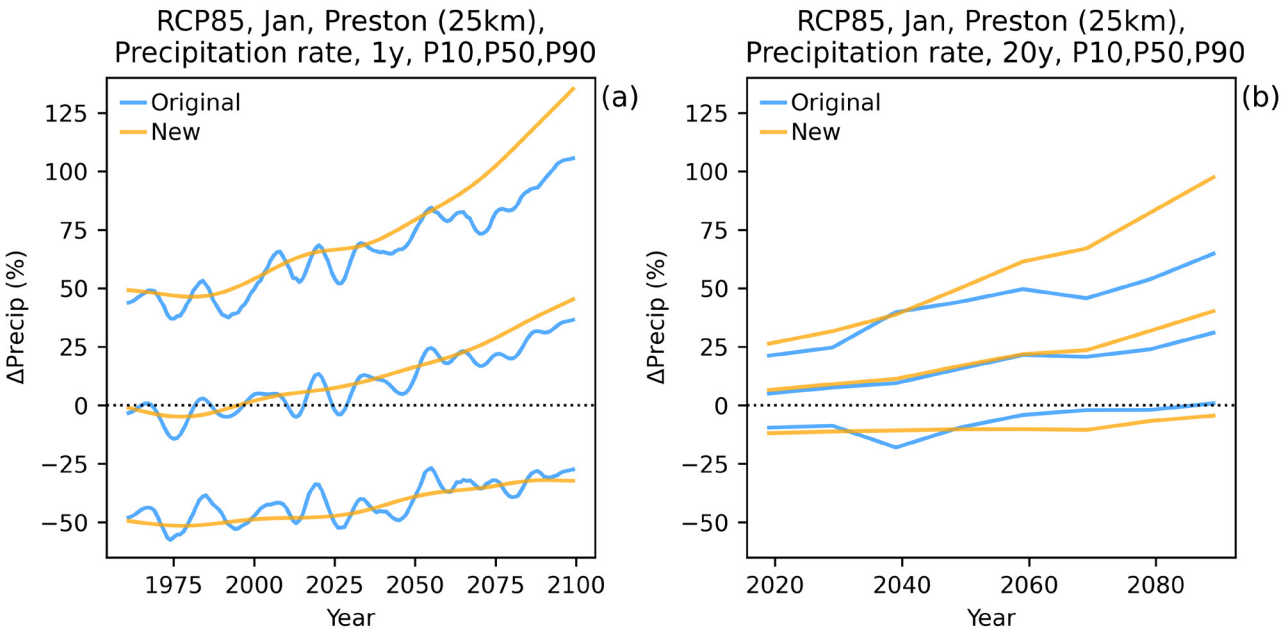


Figure 3.3. Comparison of the P10, P50 and P90 quantiles for the original UKCP18 pdfs (blue curves) and the new projections (orange curves) relative to the 1981-2000 baseline. (a) Change in annual values for January precipitation for the Preston 25km grid-point in response to RCP85 emissions. Blue curves use an 11-year pooling window to construct the pdfs, orange curves use a 31-year window. (b) As (a), but for 20-year mean projections.

Figure 3.3 compares original and new precipitation changes for January under RCP85 for the 25km Preston grid-point. The first panel compares the responses for annual data, while the second shows the corresponding 20-year mean responses. The effects of the longer time-smoothing window are again clear in the annual data (cf Figs. 3.1 and 3.2). In the 20-year mean data, differences from the original projections are dominated by the effects on the regression relationship of correcting the downscaling error (noting that the GCM predictor does not change in this example). The new plumes show increased spread arising from a stronger regression slope (see Section 2.5), with substantial increases in the P90 response beyond 2050 in comparison to the original projections. The median response also increases beyond 2050, while the P10 response becomes slightly drier in the new projections. These differences also influence the annual time series, with P90 values growing well beyond 100% by 2100 in the new projections.

Figure 3.4 shows changes in total cloud cover in summer under RCP85, comparing 20-year mean changes at (a) a 25km location (Exeter) with (b) changes for the aggregated region containing Exeter (South-West England). The most significant impact stems from the downscaling correction, which leads to a progressive broadening of the Exeter distribution after 2050 in the new projections, compared to the original results. The new aggregated-region results show the same general characteristics as the Exeter results, with median reductions of ~20% by the end of the 21st century accompanied by broad uncertainty ranges. These encompass small increases in cloud at P90 to major reductions exceeding 50% at P10.

However, the new and original projections are more similar for South-West England than for Exeter, especially at P10 and P50. In both cases, the GCM predictor used for downscaling remains unchanged in the new results. For Exeter, correcting the indexing error in the 25km regression leads to a stronger slope and hence increased spread in the new projections. In contrast, the absence of the indexing error in the original aggregated-region regression allowed the original downscaling to reflect uncertainties faithfully, hence the new pdfs match the original results more closely.

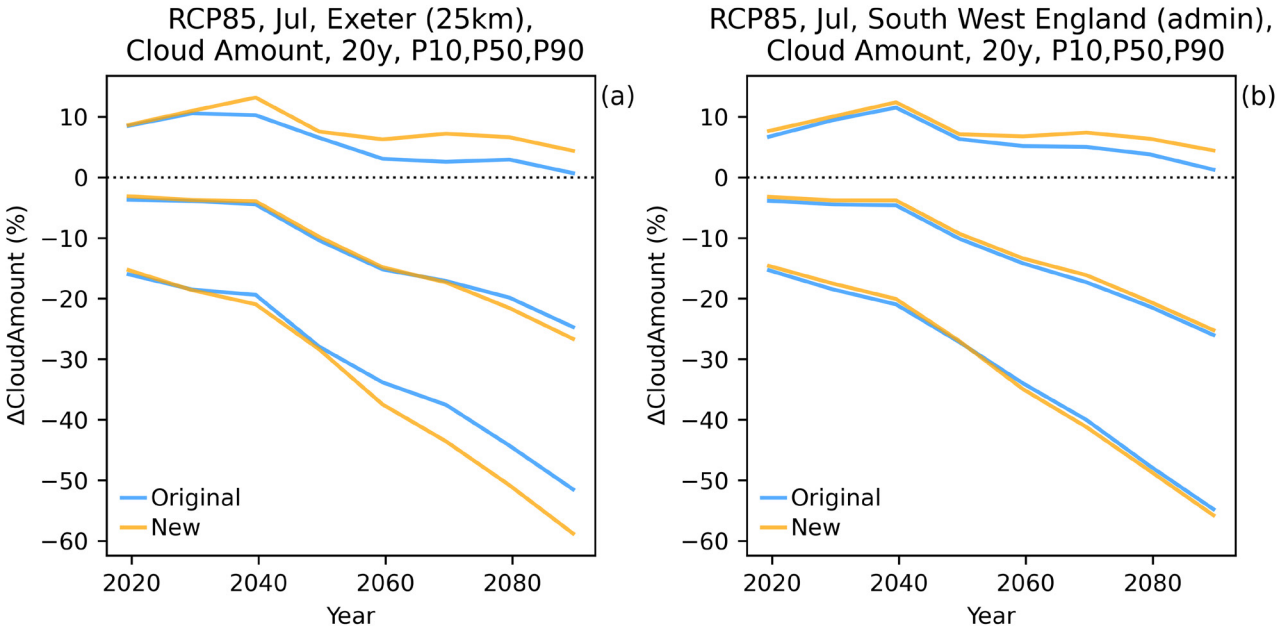


Figure 3.4. Comparison of the P10, P50 and P90 quantiles for the original UKCP18 pdfs (blue curves) and the new projections (orange curves) for summer cloud amount. (a) 25km OSGB box for Exeter, (b) South-West England aggregated region. These projections are for 20-year mean changes in response to RCP85 emissions relative to 1981-2000.

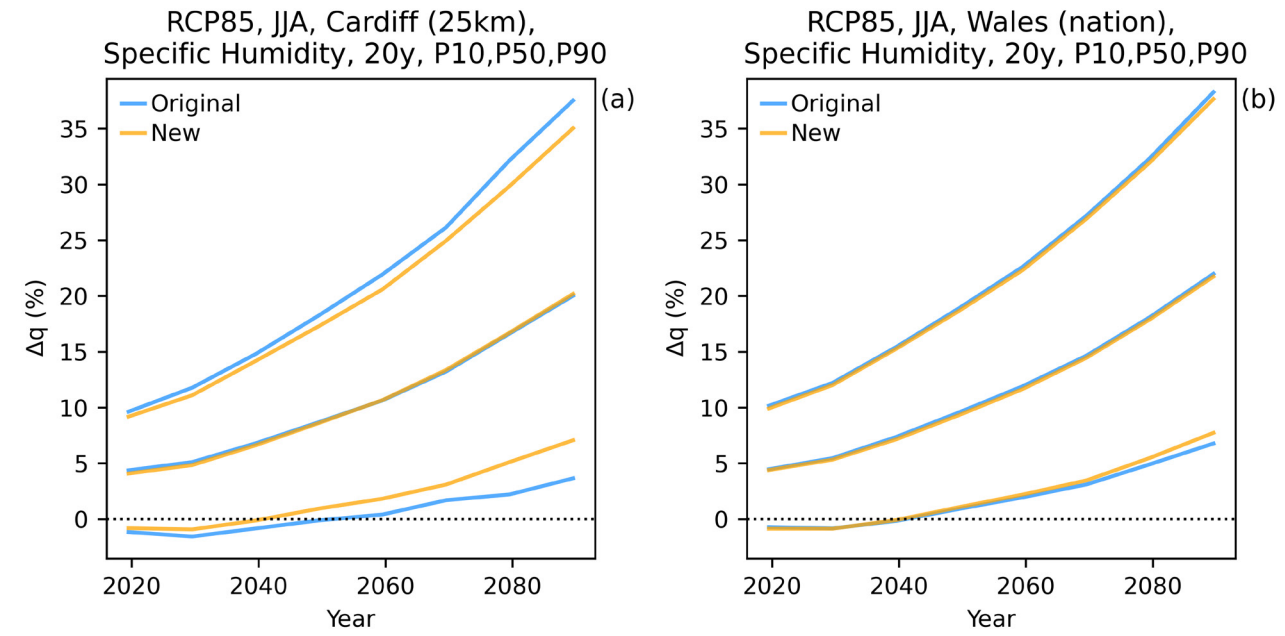


Figure 3.5. Comparison of the P10, P50 and P90 quantiles for the original UKCP18 pdfs (blue curves) and the new projections (orange curves) for summer specific humidity. (a) 25km OSGB box for Cardiff, (b) Wales aggregated region. These projections are for 20-year mean changes in response to RCP85 emissions relative to 1981-2000.

Another comparison between 25km and aggregated-region projections is shown in Fig. 3.5, comparing 20-year mean specific humidity changes in summer under RCP85 for (a) the 25km Cardiff box and (b) the Wales aggregated region. During the second half of the 21st century, the medians of the projected changes in specific humidity become larger (in comparison to the associated P90 – P10 ranges) than for variables such as total cloud cover, sea-level pressure or the radiation variables. In such cases, correcting the downscaling error leads to reduced rather than increased uncertainties in the new 25km projections

(Section 3.1), because the effects of reduced downscaling residuals outweigh the impact of slightly increased regression slopes (Section 2.5c). The spread in the Cardiff projections is therefore reduced compared to the original results, in contrast to the cloud cover projections of Fig. 3.4. Since the GCM predictor does not change and the downscaling error does not affect the aggregated-region regression, the new pdfs for Wales are very similar to the original projections (Fig. 3.5b), differing only through minor impacts of the other developments included in the new projections.

3.3. Maps comparing Original and New Projections

In this section we discuss selected maps showing typical effects of the updates on the 25km gridded data. Figure 3.6 compares the P50 responses of Tmax in July to RCP85 emissions, for the 20-year mean period 2070-2089. The difference map is shown in the third panel. The median response increases everywhere in the new 25km projections, consistent with corresponding aggregated-region results (e.g., Fig. 3.1). The mean absolute difference (MAD) is 1.24°C. In the interim projections (containing all updates apart from the downscaling correction) the MAD is 1.07°C, driven mainly by increases arising from correcting the Tmax/Tmin error. The remaining component of the MAD arises from the downscaling correction.

The coarse blocky nature of the difference map reflects the spatial resolution of the underlying GCM data used in the statistical downscaling (Section 2.5a). This is expected, since the Tmax/Tmin software error only affected GCM data. The GCM predictor used for downscaling changes at 22% of locations. These changes generally increase the difference between the new and original projections (notably in parts of South-West England, Northern England, and parts of Scotland), enhancing the MAD as described above. The average value of F_{med} is 0.18, showing that the increases in P50 in the new 25km projections for July typically represent a significant fraction of the associated P90 – P10 uncertainty ranges. However, July is the month most affected by correction of the Tmax/Tmin error (see Figure 2.3a), so the impact of this update is smaller in other months. The average F_{med} across all period definitions is 0.06 (Table 3).

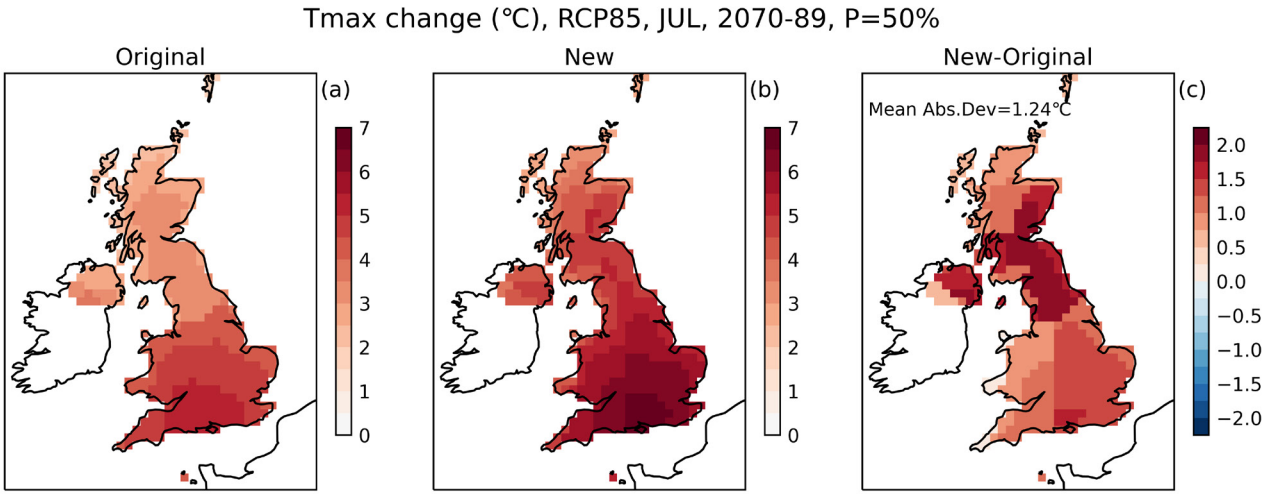


Figure 3.6. Median (P50) changes in July Tmax under RCP85 emissions for gridded pdf data for the 20-year mean period 2070-2089, relative to the 1981-2000 baseline. Panel (a) shows the original UKCP18 projections, (b) shows the new projections, while (c) shows the differences. The mean absolute difference over the UK is 1.24°C.

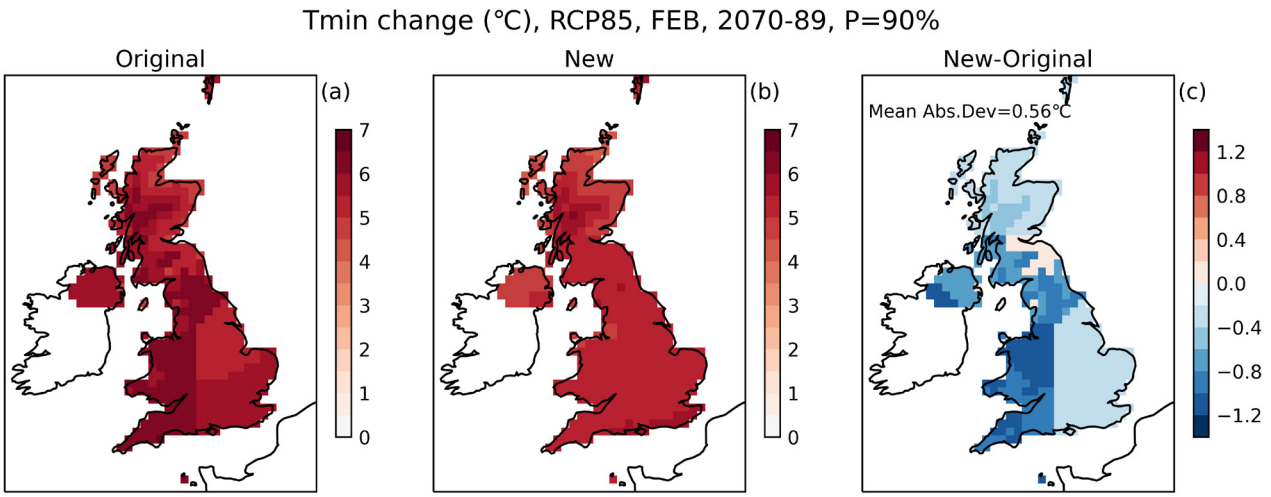


Figure 3.7. Upper-end (P90) changes in February Tmin under RCP85 emissions for gridded pdf data for the 20-year mean period 2070-2089, relative to the 1981-2000 baseline. Panel (a) shows the original UKCP18 projections, (b) shows the new projections, while (c) shows the differences. The mean absolute difference over the UK is 0.56°C.

Similarly, Figure 3.7 compares projections for Tmin, but this time P90 for February is shown. At most locations the increases in Tmin are smaller in the new projections, with a MAD of 0.56°C. The differences arise mainly from correcting the Tmax/Tmin error, which drives a MAD of 0.41°C in the absence of the downscaling correction. The cooling impact of the Tmax/Tmin correction is consistent with the aggregated-region results of Figure 2.3b, noting that these reveal February as the month with the largest differences in P90. The GCM predictor used for downscaling changes at 25% of locations. These changes generally exert a modest cooling influence in the new projections, particularly over Wales, South-West England and parts of Northern England and Scotland.

Figure 3.8 compares median summer 20-year Tmean projections, a surface temperature variable not directly impacted by the Tmax/Tmin error. Differences are typically ~0.1°C, with local exceptions where differences of ±0.3°C occur. Most of these larger differences arise from changes in the GCM predictor used in downscaling (see Fig. 2.18f). Other updates also contribute to the pattern of differences, notably through indirect impacts of the revisions to Tmin and Tmax data acting through their relationships with Tmean (Section 2.6). These generally increase the differences slightly, whereas the downscaling correction has little impact, or cools the responses slightly, in places where no change in GCM predictor occurs (Fig. 2.18f).

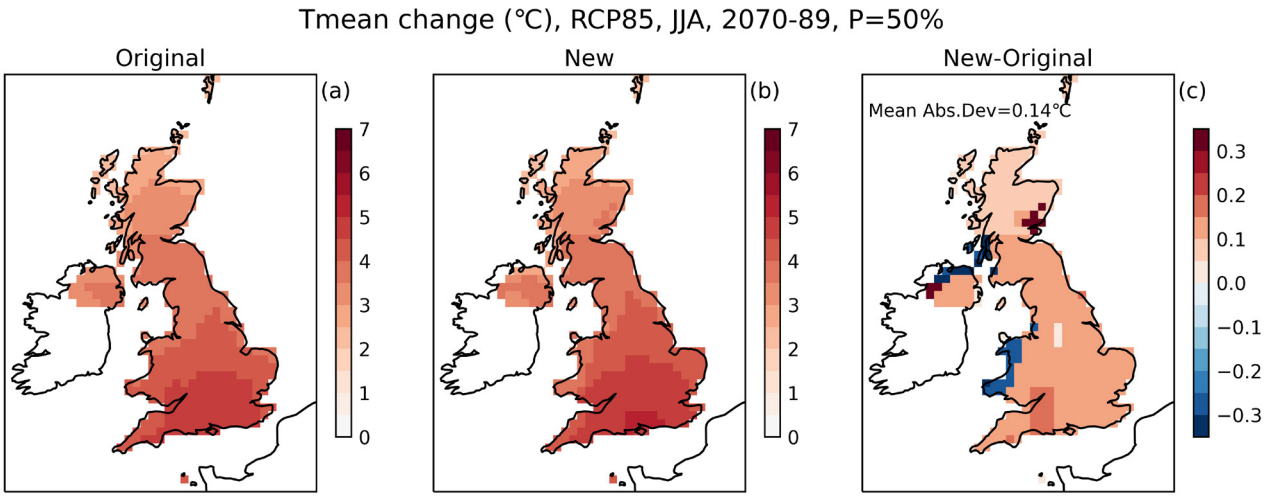


Figure 3.8. Median (P50) changes in summer Tmean under RCP85 emissions for gridded pdf data for the 20-year mean period 2070-2089, relative to the 1981-2000 baseline. Panel (a) shows the original UKCP18 projections, (b) shows the new projections, (c) shows the differences. The mean absolute difference averaged over the UK is 0.14°C.

Figure 3.9 makes a similar comparison for median summer precipitation change, the difference pattern showing increases in most of England and Wales and reductions in most of Scotland and Northern Ireland. The downscaling correction and the improved methodology for constructing precipitation distributions make the dominant contributions to the differences in the new projections. The MAD is 2.2%, dropping to 1.5% when the downscaling update is excluded. The increased P50 values over England and Wales are driven mainly by the updated precipitation calculations (Section 2.3), while the impact of the downscaling correction is to reduce P50 slightly in most places, because increased regression slopes (Section 2.5) tend to increase the median downscaled signals of change, as well as increasing their spread. Thus, the downscaling correction tends to offset the increases in P50 in the south, while strengthening the reductions to the north. The largest differences reach approximately ±4%, over parts of North-East England and Eastern Scotland respectively. However, the typical changes are modest compared with the P90 -P10 ranges in the pdfs, with an average F_{med} of 0.04.

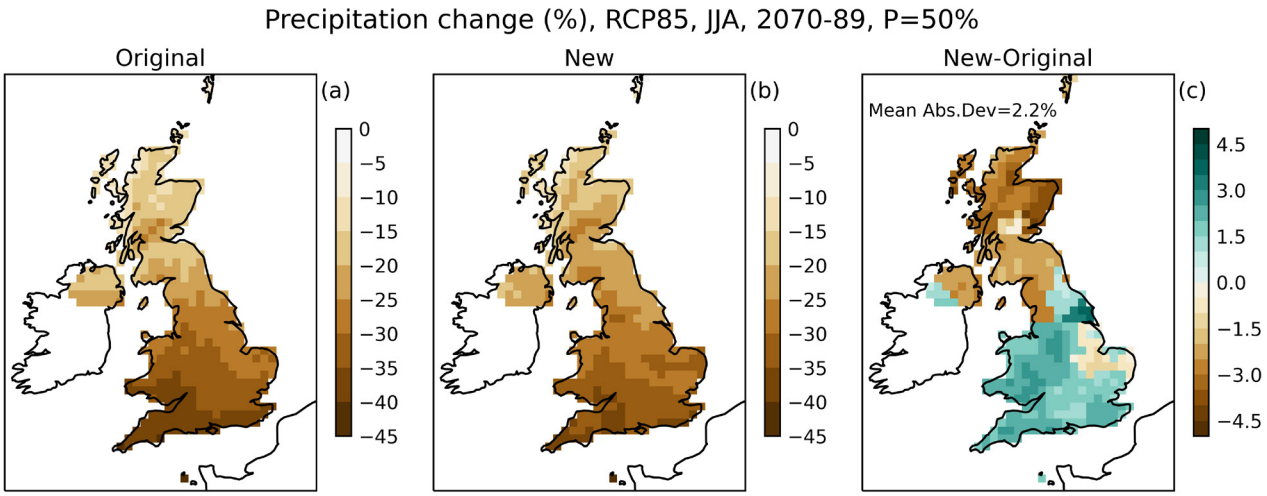


Figure 3.9. Median (P50) changes in summer precipitation (%) in response to RCP85 emissions for the 25km gridded pdf data for the 20-year mean period 2070-2089 relative to the 1981-2000 baseline. Panel (a) shows the original UKCP18 projections, (b) shows the new projections, (c) shows the differences. The mean absolute difference averaged over the UK is 2.2%.

Figure 3.10 shows corresponding results for the P10 (dry-end) changes. Increases occur in the new projections relative to the original results, over Wales and most parts of southern England and the Midlands. The largest differences occur over South Wales and South-West England, resulting in somewhat less intense drying compared to the original UKCP18 projections. In contrast, P10 values reduce over most of Scotland and Northern Ireland, leading to enhanced drying in the new projections. The MAD is 4.0% with all developments included, cf 2.3% when the downscaling correction is excluded. The improved treatment of precipitation extremes drives increases in P10 in all areas except northern Scotland, whereas the downscaling correction drives reductions almost everywhere. These influences combine to produce the net differences shown in Fig. 3.10.

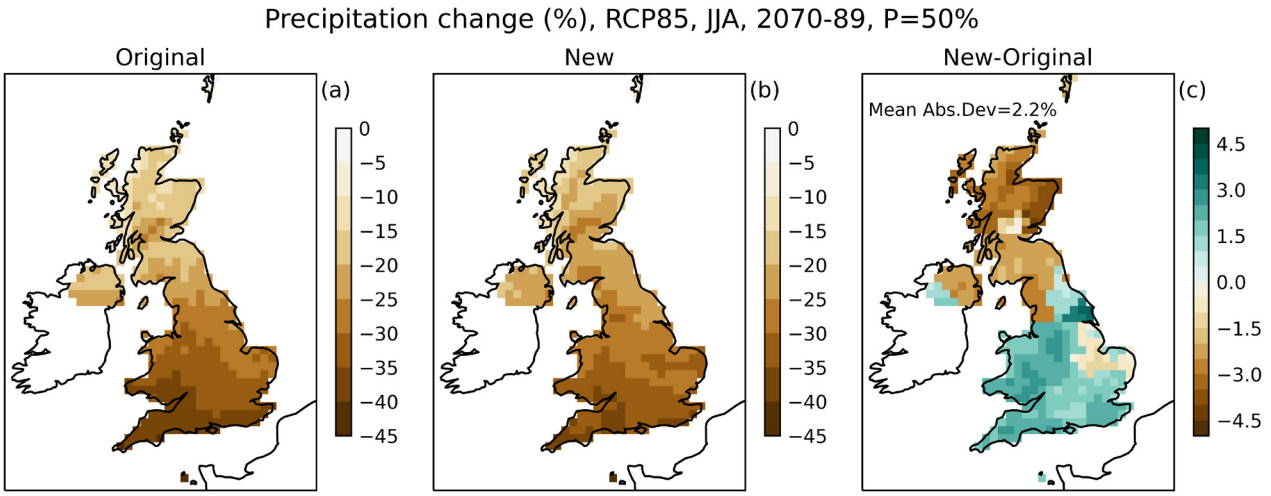


Figure 3.10. Low-end (P10) changes in summer precipitation (%) in response to RCP85 emissions for the gridded pdf data for the 20-year mean period 2070-2089 relative to the 1981-2000 baseline. Panel (a) shows the original UKCP18 projections, (b) shows the new projections, (c) shows the differences. The mean absolute difference averaged over the UK is 4.0%.

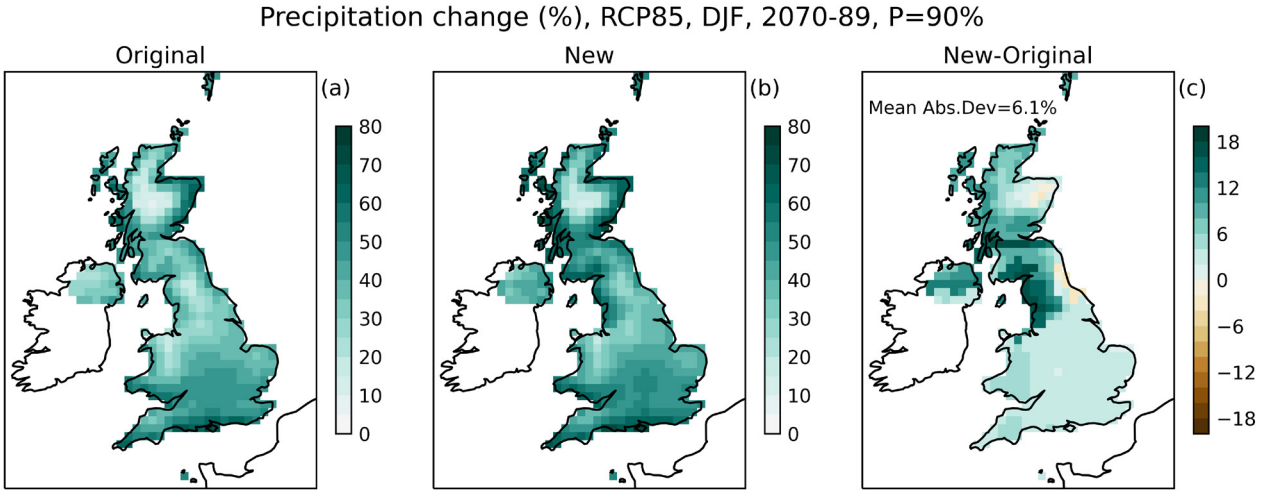


Figure 3.11. High-end (P10) changes in winter precipitation (%) in response to RCP85 emissions for the gridded pdf data for the 20-year mean period 2070-2089 relative to the 1981-2000 baseline. Panel (a) shows the original UKCP18 projections, (b) shows the new projections, while (c) shows the differences. The mean absolute difference averaged over the UK is 6.1%.

The final precipitation comparison in Figure 3.11 shows the pattern of P90 (wet-end) changes for the winter season. The new projections show larger P90 responses almost everywhere, primarily due to the corrected downscaling relationships. The MAD is +6.1%, although greater differences are obtained in parts of North-West England and Southern Scotland where changes in GCM predictor augment the increase in spread. For example, P90 for Glasgow is 52%, an increase of 16%. By contrast, for a few grid-squares in North-East Scotland and England a change in GCM predictor offsets the general tendency towards increased spread arising from stronger downscaling relationships (Sections 2.5c, d), leading to small net reductions in P90 at these locations.

The difference maps in Figs. 3.9-3.11 are less blocky than those of Figs. 3.6 and 3.7. This is partly because fixing the Tmax/Tmin software error directly affects the temperature projections from the GCM-based pdfs sampled when constructing the downscaled local projections (Section 2.5), while the corresponding GCM-based precipitation pdfs are affected only indirectly, through inter-variable correlations with the

temperature variables (Section 2.6). Secondly, the effects of the downscaling correction on patterns of change are dominated by increases in regression slope in the case of precipitation, whereas changes in GCM predictor are more dominant for surface temperature variables (see discussion of Fig. 2.19, Section2.5d). Therefore, boundaries between regions dominated by different choices of GCM predictor do not dominate the patterns of difference in Figs. 3.9-3.11, even though the GCM predictor changes at 49% of OSGB grid squares in summer, and 25% in winter.

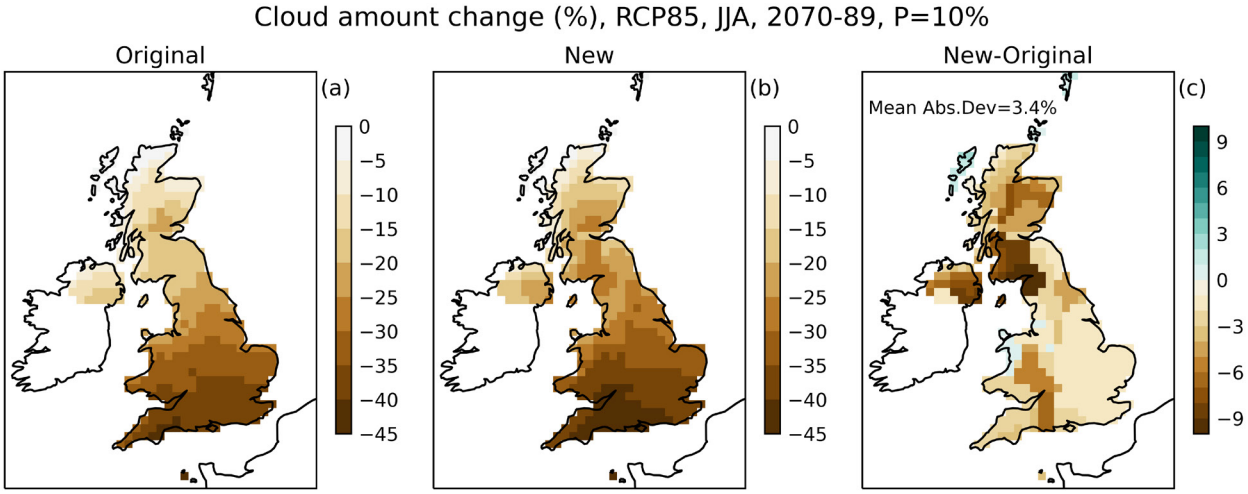


Figure 3.12. Low-end (P10) changes in summer cloud cover (%) in response to RCP85 emissions for the gridded pdf data for the 20-year mean period 2070-2089 relative to the 1981-2000 baseline. Panel (a) shows the original UKCP18 projections, (b) shows the new projections, and (c) the differences. The mean absolute difference averaged over the UK is 3.4%.

Finally, Figure 3.12 shows the P10 changes in total cloud cover that accompany the precipitation changes of Fig. 3.10. These changes drive corresponding high-end increases in downward net short-wave radiation (not shown), that exceed 50 Wm^{-2} in parts of southern England. In the new projections, the average slope in the corrected downscaling regressions for cloud cover is 0.68, cf 0.46 in the original results. This leads to increases in spread that intensify the cloud cover reductions projected at the low ends of the local pdfs. The MAD is 3.4%, with maximum differences amounting to 6-9% in the Lake District, South-West Scotland, and parts of Northern Ireland. Changes in GCM predictor occur at 34% of OSGB points, also contributing to the differences. However, the slope changes exert the main influence on the patterns of difference, as for precipitation (Fig. 3.10).

The Summary provides a recap of the updates described in Section 2, and of the main conclusions arising from the comparison in this Section of the original and new probabilistic projections. Users can also obtain general advice regarding use of the probabilistic projections data from the accompanying Frequently Asked Questions (Fung et al, 2022) available from https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/guidance-science-reports.

Acknowledgements

We thank Prof. Nigel Arnell, Dr Fai Fung, and Prof. Jason Lowe for their reviews of this report.

Appendix A

A.1 Occurrence of implausible precipitation projections

In the precipitation realizations released in 2018, statistical processing gave rise to the occurrence of a small fraction of unphysical negative precipitation events (projected drying exceeding 100%), particularly for locations in southern UK during summer with strong forcing scenarios. This arose from a linear scaling employed in the method (see Appendix B), applied to percentage precipitation changes normalised by changes in global mean surface temperature (GMST). Negative precipitation amounts occasionally resulted, since this scaling is not mathematically constrained to remain above -100%. The potential for unphysical values increases in cases where large GMST responses combine with a strong normalised drying signal, and when sampled uncertainties associated with natural climate variability, emulation, scaling and structural errors (Appendix B) are large. Since we make annual projections, the temporal variability can be particularly large (especially at the 25km scale). For realizations featuring a strong mean drying, individual years with reductions exceeding -100% therefore become possible. For the original UKCP18 projections these unphysical realizations were clipped at -100%.

An alternative is to apply a logarithmic transform to the precipitation variable, prior to statistical processing. Although this eliminates occurrences of drying exceeding -100%, it can give rise to the generation of implausible wet extremes for individual years in some realizations. Figure 2.5b shows one example. This arises from a combination of factors. Firstly, variability in the projections is simulated by resampling the residual error obtained for the downscaling regression relationships used to map the GCM response to regional scales (Section 2.5). This uncertainty is then sampled independently of the mean projected change, which can occasionally lead to combinations of high variability and weak response that are not necessarily present in the underlying climate model data. In principle, this residual variability in log coordinates could depend weakly on response, but with only eleven RCM simulations available (Appendix B.2) we are unable to quantify such relationships. In addition, the exponential operator, applied to invert the log transform when making the final projections, is very sensitive for positive values. When applied at the wet end of the distribution, statistical uncertainty can therefore inflate to implausible values when converting to percentage change. Below, we describe a method of combining use of transformed and untransformed calculations, to circumvent the issues outlined above.

A.2 Merging scaled log and scaled percent projections

To avoid statistically generated unphysical values below -100% in the annual precipitation realizations, scaled log projections are preferred, while to avoid implausible wet extremes the converse applies and scaled percent projections are favoured. To resolve these issues, we implement a conditional selection of the two transforms to create final merged distributions. Two projections are made for precipitation at each grid point and season, one using percentage change and one using the scaled log transform approach. The scaled log or scaled percent projection is selected separately for each statistical realization using the following set of rules.

- For the scaled log realizations, individual annual anomalies are first inverted (using Eq.2) to recover percentage values.
- For the scaled percent projections, any individual years showing drying exceeding -100% are then reset to the inverse scaled log value for the corresponding year.

- The mean percentage change in precipitation for the period 2070-2099 (Tend) is calculated for each of the scaled percent realizations.
- If Tend<-30%, scaled log is selected; if Tend>+30%, scaled percent is selected.
- For Tend in the range [-30, +30] the two choices are randomly sampled, assuming the probability for scaled percent increases linearly from 0 to 1 within this range[5].
- For each scaled log realization selected, if any of its individual annual anomalies are substantially outside the range of RCM data, the scaled percent projection is used instead. The maximum permissible value is defined as the maximum projected anomaly in the scaled percent projections. This is somewhat higher than the range of RCM data, since it includes sampled statistical uncertainty (Appendix B).

Figure 2.5a shows an example of a single realization in which the scaled log projection is selected, due to a strong reduction in precipitation of -55% by the end of the century. Figure 2.5b shows a different realization for the same grid-point, month and emissions scenario. In this case the scaled percent projection is selected, even though the end-of-century mean reduction is -28%. This is because several years in the scaled log projection show statistically generated wet extremes outside the range of the underlying data.

[1] For example, a realization with mean projected change of 0%, midway between the limits [-30, +30], has an equal chance of being selected from either the scaled log projections or the scaled percent projections.

Appendix B

The methodology for the UKCP18 probabilistic projections is based on the Bayesian statistical framework that underpinned a previous set of UK Climate Projections (UKCP09, Murphy et al. 2009). Detailed scientific descriptions are given in Sexton et al. (2012) and Harris et al. (2013). Subsequent updates are described in Murphy et al. (2018), which documents in detail the version used to produce UKCP18. Below we summarise the UKCP18 method and its elements in a shorter format, to provide context for the updates described in this report.

B.1 Short summary of method

UKCP18 uses a Bayesian approach to produce pdfs for the future climate response of the UK for different emissions scenarios. To achieve this, we perform a large Monte-Carlo integration over defined ranges of uncertain model parameters from a single GCM (HadCM3), weighting the sampled outcomes by relative likelihood (estimated using a set of specified observables) to produce observationally constrained pdfs. To make the estimation computationally feasible, predicted outcomes are sampled from a fast statistical emulation of the equilibrium response to doubled CO_2, calibrated to the response of a large, perturbed parameter ensemble (PPE) of HadCM3 simulations. A scaling approach, calibrated to the transient response of a second PPE of earth system model variants, allows the transient response to be inferred. This also enables sampling of other uncertainties associated with the carbon cycle, ocean, and aerosol components of the climate system. A third ensemble of CMIP5 simulations is also used to add sampling of structural model errors to the parametric uncertainties sampled using the PPE approach.

B.2 Elements of the method

Modelling uncertainty is explored using the PPE approach (Collins et al. 2006, Murphy et al. 2007), where key parameters controlling processes in the atmosphere, surface, ocean, aerosol, and land carbon cycle components are varied within expert-defined “prior” parameter spaces for a single climate model. Here, variation in historical climate and the equilibrium response to a doubling of CO_2 concentrations is explored using a relatively large 280 member PPE based on the atmosphere-mixed-layer (hereafter SLAB) configuration of the HadCM3 model (Pope et al. 2000, Johns et al. 2003). Multivariate regression relationships trained on the 280 PPE variants are then used to construct a statistical emulator (Sexton et al, 2012) for the historical and future equilibrium responses. In contrast to use of multi-model ensemble (MME) output, emulation based on a PPE provides a systematic and comprehensive sampling of climate response for untried parameter combinations and makes use of a Bayesian approach both possible and practicable.

Users are impacted by the transient climate response to scenarios of climate forcing rather than the equilibrium response, so a second emulation stage is implemented. Assuming the climate response is proportional to GMST change (Santer et al. 1990), scaling techniques are used to translate the emulated equilibrium response into estimates of past and future climate response for any point in parameter space. Given emulated predictions for the climate feedbacks, a Simple Climate Model (SCM) (Harris et al. 2013, Suppl. Mat.) is used to predict the transient global temperature response and perform the scaling. A substantial development to the scaling methodology for UKCP18 was the extension to provide the

probabilistic projections for individual years, rather than the 30-year averages of UKCP09. This is based on the method of Sexton and Harris (2015). For the new pdfs presented here, the scaling approach follows that of UKCP18, apart from the changes described in Appendix A for precipitation. The revised precipitation approach has been previously used to produce pdfs for regional precipitation change for Europe (Brunner et al, 2020).

Uncertainties in modelling the rate of CO_2 uptake by the land and ocean biogeochemical systems contribute substantially to uncertainties in projections of GMST, which in turn influence regional responses (Knutti et al. 2008, Booth et al. 2012). We therefore use an emission-driven approach (rather than concentration-driven) to take fuller account of known limitations in the current modelling of earth system processes. To this end, a 57-member PPE of transient climate change simulations, based on variants of the HadCM3C Earth System Model (ESM) has been used to explore modeling uncertainty, accounting for interactions between different earth system components (Lambert et al. 2013, Murphy et al. 2014). This ensemble (hereafter the ESPPE) is used to calibrate the SCM and provide prior distributions for key earth system processes, such as ocean heat uptake and climate-carbon feedbacks, as well as providing adjustment of potential differences between SLAB and transient regional response.

To produce probability distributions for climate response, Monte-Carlo integration is performed over the prior space of HadCM3 and SCM parameters. Emulation is used to estimate responses, allowing large sample sizes of order 10^6 to be produced and hence improving coverage of parameter space. Some PPE members perform better than others when comparing, for example, their simulations of historical climate. Within the Bayesian framework, we include a specified multivariate set of observables in the set of simulated variables and estimate a "likelihood" for each parameter set (model variant), given the observed data. Each variant is weighted by relative likelihood in the Monte-Carlo integration, to produce updated posterior probability distributions for future climate response (Harris et al. 2013). The observations used to assess likelihood include the same set of 12 seasonal climatological spatial fields used for UKCP09 (Sexton et al. 2012), and historical trends for several climate indicators. These include the Braganza indices based on GMST (Braganza et al. 2003), heat content change in the top 700m of the oceans and change in atmospheric CO_2 concentration over a recent 45-year period (Booth et al. 2017). Data sources and references for climatological and historical trend data are listed in Tables B.1 and B.2 of Murphy et al. (2018).

An important component of the method is the recognition that models are imperfect. Given that some parameter choices lead to better simulations of real-world earth system processes than others, we assume that there exists a "best input" set of parameter choices that provides the best simulation of true climate. However, due to model imperfection, even the best-input variant will possess an irreducible structural error (termed "discrepancy" in Sexton et al. 2012). We estimate structural errors by using output from other independent climate model simulations, searching the prior space for parameter sets that best reproduce output for selected CMIP5 models (Taylor et al. 2012). Since we account for carbon-cycle modeling uncertainty, emission-driven CMIP5-ESMs are used for this purpose. This structural uncertainty component, which is estimated at GCM spatial scales and included in the projections, is specified as a multivariate Gaussian distribution, sampling of which serves to broaden and adjust the prior distribution. This method accounts for differences in structural errors found in alternative climate models, but not for systematic errors common to all models.

The above procedures are combined to produce 10^6 statistical realizations of UK climate change of time-dependent variability and change, at the HadCM3 spatial resolution of 2.5°×3.75°. These include representation of statistical uncertainties by sampling residual misfit terms associated with several steps outlined above, including equilibrium response emulation error, conversion of equilibrium responses to characteristic transient responses and scaling error associated with estimating time-dependent changes.

These outcomes are then resampled by likelihood to produce a smaller sample of 3000 equally-likely realizations convenient for analysis of impacts. Finally, these realizations are downscaled to a set of UKCP regions (Fung et al., 2018) to provide projections at scales suitable for user applications. The downscaling method is discussed in Section 2.5.

References

Booth BBB, Jones CD, Collins M, Totterdell IJ, Cox PM, Sitch S, Huntingford C, Betts RA, Harris GR, Lloyd J (2012), **High sensitivity of future global warming to land carbon cycle processes**, Environ. Res. Lett., 7, 24002, doi:10.1088/1748-9326/7/2/024002

Booth BBB, Harris GR, Murphy JM, House JI, Jones CD, Sexton DMH, Sitch S (2017), **Narrowing the range of future climate projections using historical observations of atmospheric CO_2**, J. Clim. 30, 3039-3053, https://doi.org/10.1175/jcli-d-16-0178.1

Braganza K, Karoly DJ, Hirst AC, Mann ME, Stott P, Stouffer RJ, Tett SFB (2003), **Simple indices of global climate variability and change: Part I — variability and correlation structure**, Clim Dyn 20:491–502.

Brunner L, McSweeney C, Ballinger AP, Befort DJ, Benassi, M, Booth B, Coppola E, de Vries H, Harris G, Hegerl GC, Knutti R, Lenderink G, Lowe J, Nogheretto R, O'Reilly C, Qasmi S, Ribes A, Stocchi P, Undorf S (2020), **Comparing methods to constrain future European climate projections Using a consistent framework**, J. Climate 33, 8671-8692. https://doi.org/10.1175/JCLI-D-19-0953.1

Collins M, Booth BBB, Harris GR, Murphy JM, Sexton DMH, Webb MJ (2006), **Towards quantifying uncertainty in transient climate change**, Clim. Dyn., 27, 127-147. doi:10.1007/s00382-006-0121-0.

Fung F, Stephens A, Wilson A (2018), **UKCP18 Guidance: Data availability, access and formats**, Available from https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/download-data (updated July 2021).

Harris GR, Collins M, Sexton DMH, Murphy JM, Booth BBB (2010), P**robabilistic projections for 21st century European climate**, Nat. Hazards Earth Syst. Sci. 10, 2009-2020.

Harris GR, Sexton DMH, Booth BBB, Collins M, Murphy JM (2013), **Probabilistic projections of transient climate change**, Clim Dyn 40: 2937. https://doi.org/10.1007/s00382-012-1647-y

Fung F, Pirret JSR, Fuller EA, Harris GR, Murphy JM, Sexton DMH (2022), **Probabilistic Projections Update – FAQs**, Met Office. Available from https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/guidance-science-reports.

Harris GR, Murphy JM, Pirret JSR, Sexton DMH (2022), **Update to UKCP18 probabilistic projections: Maps of projected changes in surface temperature and precipitation**, Met Office. Available from https://www.metoffice.gov.uk/research/approach/collaboration/ukcp/guidance-science-reports.

Held IM, Soden BJ (2000), **Water vapor feedback and global warming**, Annu. Rev. Energy Environ. 25: 441–475.

Iris: A Python package for analysing and visualising meteorological and oceanographic data sets (2010-2022), Met Office, Exeter, Devon. http://scitools.org.uk/

Johns TC, Gregory JM, Ingram WJ, Johnson CE, Jones A, Lowe JA, Mitchell JFB, Roberts DL, Sexton DMH, Stevenson DS, Tett SFB, Woodage MJ (2003), **Anthropogenic climate change for 1860 to 2100 simulated with the HadCM3 model under updated emissions scenarios**, Clim Dyn 20:583–612

Knutti R, Allen MR, Friedlingstein P, Gregory JM, Hegerl GC, Meehl GA, Meinshausen M, Murphy JM, Plattner G-K, Raper SCB, Stocker TF, Stott PA, Teng H, Wigley TML (2008), **A review of uncertainties in global temperature projections over the twenty-first century**, J. Clim. 21: 2651-2663.

Lambert FH, Harris GR, Collins M, Murphy JM, Sexton DMH, Booth BBB (2013), **Interactions between perturbations to different Earth system components simulated by a fully-coupled climate model**, Clim Dyn 41: 3055. https://doi.org/10.1007/s00382-012-1618-3

Moss RH, Edmonds JA, Hibbard KA, Manning MR, Rose SK, van Vuuren DP, Carter TR, Emori S, Kainuma M, Kram T, Meehl GA, Mitchell JFB, Nakicenovic N, Riahi K, Smith SJ, Stouffer RJ, Thomson AM, Weyant JP, Wilbanks TJ (2010), **The next generation of scenarios for climate change research and assessment**, Nature 463: 747-756.

Murphy JM, Booth BBB, Collins M, Harris GR, Sexton DMH, Webb MJ (2007), **A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles**, Philos Trans R Soc A 365:1993–2028

Murphy JM, Sexton DMH, Jenkins GJ, Boorman PM, Booth BBB, Brown CC, Clark RT, Collins M, Harris GR, Kendon EJ, Betts RA, Brown SJ, Howard TP, Humphrey KA, McCarthy MP, McDonald RE, Stephens A, Wallace C, Warren R, Wilby R, Wood RA (2009), **UK Climate Projections Science Report: Climate change projections**, Met Office Hadley Centre, Exeter, U.K., https://webarchive.nationalarchives.gov.uk/20181204111026/http://ukclimateprojections-ukcp09.metoffice.gov.uk/22530#projections

Murphy JM, Booth BBB, Boulton CA, Clark RT, Harris GR, Lowe JA, Sexton DMH (2014), **Transient climate changes in a perturbed parameter ensemble of emissions-driven earth system model simulations**, Clim Dyn 43: 2855. https://doi.org/10.1007/s00382-014-2097-5

Murphy JM, Harris GR, Sexton DMH, Kendon EJ, Bett PE, Clark RT, Eagle KE, Fosser G, Fung F, Lowe JA, McDonald RE, McInnes RN, McSweeney CF, Mitchell JFB, Rostron JW, Thornton HE, Tucker S, Yamazaki K (2018), **UKCP18 Land Projections: Science Report**, Met Office Hadley Centre, Exeter, U.K., https://www.metoffice.gov.uk/pub/data/weather/uk/ukcp18/science-reports/UKCP18-Land-report.pdf

Myhre G, Shindell D, Bréon F-M, Collins W, Fuglestvedt J, Huang J, Koch D, Lamarque J-F, Lee D, Mendoza B,Nakajima T, Robock A, Stephens G, Takemura T, Zhang H (2013), **Anthropogenic and natural radiative forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change** [Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 659–740, doi:10.1017/CBO9781107415324.018.

Osborn TJ, Hulme M (1997), **Development of a relationship between station and grid-box rain day frequencies for climate model evaluation**, J. Clim. 10: 1885-1908.

Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000), **The impact of new physical parametrizations in the Hadley Centre climate model - HadAM3**, Clim Dyn 16:123-146.

Rougier J (2007), **Probabilistic inference for future climate using an ensemble of climate model evaluations**. Climatic Change 81:247, https://doi.org/10.1007/s10584-006-9156-9

Santer BD, Wigley TML, Schlesinger ME, Mitchell JFB (1990), **Developing climate scenarios from equilibrium GCM results**, Max-Planck Institute for Meteorology Report Number 47, Hamburg

Sexton DMH, Harris GR (2015), **The importance of including variability in climate change projections used for adaptation**, Nature Climate Change 5, 931-936, https://doi.org/10.1038/nclimate2705

Sexton DMH, Murphy JM, Collins M, Webb MJ (2012), **Multivariate probabilistic projections using imperfect climate models, Part I: Outline of methodology**, Clim Dyn, 38: 2513, https://doi.org/10.1007/s00382 011 1208 9

Taylor KE, Stouffer RJ, Meehl GA (2012), **An overview of CMIP5 and the experiment design**, Bull. Am. Met. Soc. 93: 485-498.

UKCP18 Technical Note: Clipping and Baseline advice on Land Strand 1 data in UKCP18. https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp18-technical-note-clipping-and-baseline-guidance-on-land-strand-1-data-in-ukcp18.pdf

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0 Contributors. (2020), **SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python**. Nature Methods, 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2 For Gaussian kernel density estimation: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.gaussian_kde.html

Watterson IG (2008), **Calculation of probability density functions for temperature and precipitation change under global warming**, J Geophys Res 113:D12106. doi: 10.1029/2007JD009254