Kendon EJ, Roberts NM, Fowler HJ, Roberts MJ, Chan SC, Senior CA.
Heavier summer downpours with climate change revealed by weather forecast resolution model.
***Nature Climate Change* 2014, 4(7), 570-576**

DOI link to article:

http://dx.doi.org/10.1038/NCLIMATE2258

Heavier summer downpours with climate change revealed by weather forecast resolution model

Elizabeth J. Kendon *
Met Office Hadley Centre, Exeter, UK

Nigel M. Roberts
MetOffice@Reading, Reading, UK

Hayley J. Fowler
School of Civil Engineering and Geosciences, Newcastle University, UK

Malcolm J. Roberts
Met Office Hadley Centre, Exeter, UK

Steven C. Chan
School of Civil Engineering and Geosciences, Newcastle University, UK

Catherine A. Senior
Met Office Hadley Centre, Exeter, UK

April 16, 2014

The intensification of precipitation extremes with climate change[1] is of key importance to society due to the large impact through flooding. Observations show that heavy rainfall is increasing on daily timescales in many regions,[2] but how changes will manifest themselves on sub-daily timescales

---

[*] *Corresponding author address:* Elizabeth Kendon (née Kennett), Met Office Hadley Centre, Fitzroy Road, Exeter, EX1 3PB, UK.
E-mail: elizabeth.kendon@metoffice.gov.uk

remains highly uncertain. Here we perform the first climate change experiments with a very high resolution (1.5km grid spacing) model more typically used for weather forecasting, in this instance for a region of the UK. The model simulates realistic hourly rainfall characteristics including extremes,[3,4] unlike coarser resolution climate models,[5,6] giving us confidence in its ability to project future changes at this timescale. We find the 1.5km model shows increases in hourly rainfall intensities in winter, consistent with projections from a coarser 12km resolution model and previous studies at the daily timescale.[7] However, the 1.5km model also shows a future intensification of short-duration rain in summer, with significantly more events exceeding high thresholds indicative of serious flash flooding. We conclude that accurate representation of the local storm dynamics is an essential requirement for predicting changes to convective extremes; when included we find for the model here that summer downpours intensify with warming.

Few studies have examined changes in rainfall on hourly timescales due to sparse sub-daily observations and the inability of climate models to reliably simulate sub-daily rainfall. The studies so far suggest greater increases in hourly compared to daily rainfall extremes,[5,8] but due to model deficiencies we have low confidence in these projections. This is of concern since it is short duration convective extremes which tend to be responsible for flash flooding events, such as the Boscastle flood in August 2004,[9] particularly important in urban environments and for small or steep river catchments.

The Clausius-Clapeyron (CC) relation describes the rate of change of saturated water vapour pressure with temperature as approximately 7%/°C, and sets a scale for change in precipitation extremes.[1] Increasing evidence from observational studies suggests intensities of sub-daily precipitation extremes increase more rapidly with temperature than for daily extremes; above the CC rate, at least in some regions.[8,10] This appears to be a property of convective precipitation[10] and may be explained by latent heat released within storms invigorating vertical motion, leading to greater increases in rainfall intensity. However, the extent to which this scaling may apply over the longer-term with global warming is uncertain.

Global and regional climate models (with typical grid spacings of 60-300km and 10-50km respectively) rely on a convective parameterisation scheme to represent the average effects of convection. This simplification is a known source of model error, and leads to deficiencies in the diurnal cycle of convection[11] and the inability (by design) to produce hourly precipitation extremes.[5,6,8] Very high resolution models (order 1km grid spacing), on the other hand, can represent deep convection explicitly without the need for such a parameterisation scheme.[3,12] Such models are termed 'convection-permitting' because larger storms and meso-scale convective organisation are permitted (largely resolved) but convective plumes and small showers are still not represented.

Convection-permitting models are commonly used in short-range weather forecasting. They give a much more realistic representation of convection and are able to forecast localised extreme events not captured at coarser resolutions.[13] However, there are few examples of convection-permitting resolutions being applied in climate studies, due to their high computational cost. Previous studies have been limited to small domains and often just a single season[12,14,15] or selected events.[16,17] Some studies have built up multi-year climatologies through a sequence of seasonal[18,19] or shorter[20] simulations. However, long continuous simulations are needed to represent long-term memory in the soil and its feedbacks with precipitation.[21] We recently[3] carried out the first extended (20-year) length climate simulation with a convection-permitting (1.5km) model over a region of the UK. Here we use the same model to examine future changes. To our knowledge this is the first time that continuous multi-year simulations at such high resolutions have been carried out to study rainfall change for a future climate scenario. Climate change experiments have been carried out at 4km resolution over the western US,[22] but this resolution is not high enough to adequately represent typical convection over the UK.[13]

We compare future changes in hourly rainfall in the 1.5km model with results from a 12km regional climate model (RCM) over the southern UK. The models are run for 13-year present-day (1996-2009) and 13-year future ($\sim$2100, under the Intergovernmental Panel on Climate Change RCP 8.5 scenario) periods, driven by a 60km global climate model (GCM). Model biases for the present-day have been assessed by comparison with gridded hourly observations from radar, available for 2003-2012.[23] Since radar tends to systematically underestimate heavy rain,[24] we apply a bias correction using daily gauge observations.

On hourly timescales, rainfall is heavier over the southern UK in summer than in winter (Figures 1 and 2). Model biases compared to radar data are also larger in summer. In particular, the 12km-RCM significantly underestimates heavy rainfall in summer, whilst the 1.5km model tends to provide an overestimate, particularly in the south-east. The tendency for heavy rain to be too intense in small convective cores in the 1.5km model is understood and is a current inherent weakness of a ‘convection-permitting’ model.[13] Smaller showers are not properly resolved, with some showers having updrafts on the wrong scale with insufficient turbulent sub-grid mixing. Nevertheless, the 1.5km model gives a much better representation of hourly rainfall characteristics including extremes[3,4] than the 12km model, and extensive testing within numerical weather prediction trials at the Met Office has shown considerable benefit from the 1.5km model leading to its operational implementation as a replacement for the previous 4km and 12km models.

We find that both models show future increases in heavy rainfall in winter, consistent with studies on daily timescales.[7] The 1.5km model also shows significant increases in heavy rain intensities in summer, which are not seen in the 12km-RCM. Previous studies relying on coarser models have shown large uncertainties in changes in summertime extremes.[5,7] The summertime increases in the 1.5km model are 36% for the southern UK on average which, given a surface warming of about 4-5°C for heavy rain days, are consistent or possibly higher than CC-scaling.

Looking at the heaviest 50 events (averaged to the 12km grid) in the 1.5km model in summer in the future simulation, about half of these are larger-scale storms (embedded convection within a front, Mesoscale Convective Systems or squall lines) with the remainder being individually smaller storms (often clustered). The events appear physically plausible, with realistic evolution and the model responding to the environment in a sensible way. In particular, nearly all events are associated with cyclonic flow and hot humid conditions (high 850hPa wet-bulb potential temperature, θ_w). A recent observed event (27th July 2013) with similar conditions is shown in Supplementary Figure 1, illustrating the ability of the 1.5km model to capture rainfall accumulations associated with an intense squall-line. In the present-day control run, a similar proportion of the heaviest events are larger organised systems, but these are associated with lower hourly accumulations (only 6 events have values exceeding 30mm/h over a 12km grid box compared to 22 in the future). We find that present-day θ_w is considerably lower ($3 - 4$°C less for the heaviest events), suggesting that future increases in heavy rainfall intensities are associated with the warmer moister environment (see Supplementary Material).

Future changes in the number of exceedances above high thresholds are useful because of the relationship to flood risk. Here we examine the frequency of episodes exceeding present-day percentiles of wet hourly precipitation (Figure 3), which removes any issues with model or radar absolute bias. Here an episode is a continuous period of exceedance at a given grid point, counted separately for different grid points; whereas an event refers to the rainfall field at a given time. For the 12km-RCM, there are too few episodes exceeding low thresholds in both seasons (blue asterisks) and so individual episodes must be too persistent. By comparison the number of episodes is well captured in the 1.5km model. Both models show a significant future increase in the number of episodes exceeding high thresholds in winter (red asterisks). In summer, both models show a significant decrease in the number of episodes exceeding low thresholds, consistent with it raining less often in future, but only the 1.5km model shows an increased frequency of episodes exceeding the higher thresholds. The interplay between fewer rainfall episodes overall and an increasing intensity of heavy rain means that for a high enough threshold (99.999th percentile of wet hours for the present-day, corresponding to 28 mm/h) the 1.5km model shows a significant increase in the number of episodes in the future summer (24 episodes exceed 28mm/h in the present-day and 117 in the future).

The duration-intensity characteristics of rainfall are of key importance for flooding. The 12km-RCM has too little short-duration high-peak intensity rain and too much long-duration low-intensity rain. This is shown for summer in Figure 4, with similar biases seen in winter (Supplementary Figure 4). These biases are considerably larger than possible radar error (Supplementary Figure 3). The 1.5km model, by contrast, is able to capture the observed characteristics with significantly reduced biases. Thus this model allows us to project in a much more reliable way how these characteristics may change in the future. In winter, both models show a very similar change, with significant increases in peak rainfall intensity across all durations (Figure 4). Importantly in summer, the two models show quite different signals of change. The 1.5km model shows an intensification of short-duration rain which is not seen in the 12km-RCM. This intensification of short-duration rain is not inconsistent with many of the heaviest events in future being large scale storms (see Supplementary Material).

It is perhaps surprising that the two models show consistent changes in winter, given that the biases in the 12km-RCM appear to be responsible for different changes compared to the 1.5km model in summer. We explain this difference by examining composites of the heaviest 50 events in the two seasons, where each event centre is relocated to a common point to produce a 'typical' heavy event (Supplementary Figure 5). In winter, both models show an intensification of the whole event, whereas in summer the increase in the 1.5km model is confined to the peak of the event. Since high peak intensities in summer are linked to convection (or convective enhancement within large scale storms), this points to deficiencies in the convection parameterisation in the 12km-RCM being the cause of the different summertime change. In winter, the more widespread increase is consistent with enhanced large-scale moisture convergence, and thus deficiencies in the convective enhancement of rain in the coarser-scale model (apparent from the control biases) do not impact on the future change.

In conclusion, future projections of changes to UK winter rainfall are robust from coarser to higher resolution models. However, in summer, deficiencies in the convective parameterisation scheme in coarse resolution models seriously impact on projections of rainfall change. Using a convection-permitting model we find evidence of an intensification of hourly rainfall in summer not seen in a coarser 12km resolution model. This is of major importance for flooding; in particular, an accumulation threshold of 30mm/h is used by the Met Office/Environment Agency Flood Forecasting Centre as guidance to indicate likely flash flooding, and results here suggest this may be exceeded more often, over a wide area (12km x 12km), in the future. These results are based on one climate model, and so we cannot assess modelling uncertainty. However, the intensification of summertime convective events in a warmer moister environment is consistent with both theoretical expectations of super-Clausius-Clapeyron scaling and the limited observational studies of sub-daily rainfall to date. We conclude that accurate representation of the local storm dynamics is essential for predicting future change in convective storms (along with accurate representation of changes in the larger-scale environment inherited from the driving GCM). This implies that previous interpretations of future regional climate change scenarios should be revisited as changes in convective rain events could have been underestimated.

Methods

The models used here are all configurations of the Met Office Unified Model (MetUM).[25] The 1.5km model spans the southern UK and is driven by the 12km-RCM, which spans Europe and is in turn driven by the 60km-GCM. The 1.5km model is as described previously,[3] with some upgrades to the model physics, particularly an improved microphysical parameterisation of drizzle and fog.[26] The 12km-RCM and 60km-GCM both have UM Global Atmosphere 3.0 configuration,[25] and have similar model physics to that in the 1.5km model except that, at 1.5km resolution, the convection scheme has been switched off and Smagorinsky-Lilly turbulence diffusion is applied.

For the present-day control runs, monthly sea surface temperature (SST) and sea-ice forcings from the Program for Climate Model Diagnosis and Intercomparison were used. Other forcings follow the Atmospheric Model Intercomparison Project II (AMIP-II) protocols, excepting that Shine-Li ozone climatology[27] was used for the 1.5km and 12km RCMs. For the future runs, SST was configured as time-varying monthly SST from the control run plus the (multi-year) mean SST change for each month between 1990-2010 and 2090-2110 in the HadGEM2-ES runs[28] under the IPCC RCP8.5 scenario. Carbon dioxide, methane, nitrous oxide, CFC and HFC concentrations were adjusted accordingly, but do not vary with time. Sea-ice comes directly from the HadGEM2-ES integration, as a repeat monthly cycle. Ozone and aerosol forcings were not changed between the present-day and future runs. In the 60km-GCM and 12km-RCM, aerosol mass mixing ratios provide the cloud droplet number for autoconversion. In the case of the 1.5km model, however, autoconversion limits are based on droplet number assumptions.[29] Since aerosol forcings are not changed for the future simulations, this is not expected to have a large impact on the climate change results.

Soil moisture evolves freely using the Joint UK Land Environment Simulator (JULES).[30] Soil moisture in the 1.5km model is initialised from the 12km-RCM, and takes a few months to 'spin-up' (except potentially in the very deepest layer, where it can take several years to fully reach equilibrium). Thus the first few months of the simulation were discarded (the simulations were actually 13 years 7 months), and the analysis here only uses 13 years of model data from December 1996 to November 2009 for the present-day and similarly for the future runs. We note that a key benefit of long-continuous simulations (rather than seasonal slices) is that long-memory land-surface feedbacks can be represented.

The radar data used here are 5km hourly data from the Nimrod database.[23] Radar data offer good spatial coverage and are available for an extended period (2003 - 2012). However, the radar tends to underestimate heavy rain because of beam attenuation,[24] and so we apply a bias correction using daily rain gauge observations (further details of the observational datasets are provided in the Supplementary Material). In particular, at times when hourly rainfall is a major contributor to the daily total, it is possible to estimate an upward correction to the hourly radar intensity by comparing daily radar totals with daily gauge totals. Specifically we identify heavy hourly rainfall amounts in the radar data, and identify when these are $> 0.3\times$ daily radar total. If this criterion is met, and the daily radar total $<$ daily gauge total for that grid point on that day, then we upscale the radar hourly amount as follows:

$$P_{adjusted} = P_{Hourly\_radar} \times \frac{P_{Daily\_gauge}}{P_{Daily\_radar}} \tag{1}$$

If the daily radar total $>$ daily gauge total, no correction is applied. This corresponds to the situation of a heavy localised shower missed by the gauges, for which the radar provides the best (although potentially biased) estimate. The sensitivity of the bias correction to the selection criterion and the impact of the correction on the results are discussed in the Supplementary Material (see Supplementary Figures 2 and 3).

All analysis here is carried out at the 12km scale, with the hourly precipitation fields for the 1.5km model and 5km radar being first aggregated onto the 12km-RCM grid. Bootstrap resampling is used to assess the significance of model biases and future changes with respect to year-to-year variability. 1000 bootstrap samples are produced for the model (radar) data by selecting 13 (10) years from the full dataset randomly with replacement. These are used to produce 1000 estimates of the difference between the model and radar, or the future and present-day model runs, allowing a confidence interval for the difference to be calculated.

References

[1] Trenberth, K. E., Dai, A., Rasmussen, R. M. & Parsons, D. B. The changing character of precipitation. *Bull. Am. Meteorol. Soc.* **84**, 1205–1217 (2003). doi:10.1175/BAMS-84-9-1205.

[2] Min, S.-K., Zhang, X., Zwiers, F. W. & Hegerl, G. C. Human contribution to more-intense precipitation extremes. *Nature* **470**, 376–379 (2011).

[3] Kendon, E. J., Roberts, N. M., Senior, C. A. & Roberts, M. J. Realism of rainfall in a very high resolution regional climate model. *J. Climate* **25**, 5791–5806 (2012). doi:10.1175/JCLI-D-11-00562.1.

[4] Chan, S. C. *et al.* The value of high-resolution Met Office regional climate models in the simulation of multi-hourly precipitation extremes. *J. Climate* (in press).

[5] Hanel, M. & Buishand, T. A. On the value of hourly precipitation extremes in regional climate model simulations. *J. Hydrol.* **393**, 265–273 (2010). doi:10.1016/j.jhydrol.2010.08.024.

[6] Gregersen, I. B. *et al.* Assessing future climatic changes of rainfall extremes at small spatio-temporal scales. *Clim. Change* **118**, 783–797 (2013). doi:10.1007/s10584-012-0669-0.

[7] Fowler, H. J. & Ekström, M. Multi-model ensemble estimates of climate change impacts on UK seasonal precipitation extremes. *Int. J. Climatol.* **29**, 385–416 (2009). doi:10.1002/joc.1827.

[8] Lenderink, G. & van Meijgaard, E. Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geosci.* **1**, 511–514 (2008). doi:10.1038/ngeo262.

[9] Burt, S. Cloudburst upon Hendraburnick Down: The Boscastle storm of 16 August 2004. *Weather* **60**, 219–227 (2005). doi:10.1256/wea.26.05.

[10] Berg, P., Moseley, C. & Haerter, J. O. Strong increase in convective precipitation in response to higher temperatures. *Nature Geosci.* **6**, 181–185 (2013). doi:10.1038/ngeo1731.

[11] Brockhaus, P., Lüthi, D. & Schär, C. Aspects of the diurnal cycle in a regional climate model. *Meteorol. Z.* **17**, 433–443 (2008). doi:10.1127/0941-2948/2008/0316.

[12] Hohenegger, C., Brockhaus, P. & Schär, C. Towards climate simulations at cloud-resolving scales. *Meteorol. Z.* **17**, 383–394 (2008). doi:10.1127/0941-2948/2008/0303.

[13] Lean, H. W. *et al.* Characteristics of high-resolution versions of the Met Office Unified Model for forecasting convection over the United Kingdom. *Mon. Weather Rev.* **136**, 3408–3424 (2008). doi:10.1175/2008MWR2332.1.

[14] Langhans, W., Schmidli, J., Fuhrer, O., Bieri, S. & Schär, C. Long-term simulations of thermally driven flows and orographic convection at convection-parameterizing and cloud-resolving resolutions. *J. Appl. Meteorol. Clim.* **52**, 1490–1510 (2013). doi:10.1175/JAMC-D-12-0167.1.

[15] Prein, A. F. *et al.* Added value of convection permitting seasonal simulations. *Clim. Dyn.* **41**, 2655–2677 (2013). doi:10.1007/s00382-013-1744-6.

[16] Mahoney, K., Alexander, M., Scott, J. D. & Barsugli, J. High-resolution downscaled simulations of warm-season extreme precipitation events in the Colorado Front Range under past and future climates. *J. Climate* **26**, 8671–8689 (2013). doi:10.1175/JCLI-D-12-00744.1.

[17] Attema, J. J., Loriaux, J. M. & Lenderink, G. Extreme precipitation response to climate perturbations in an atmospheric mesoscale model. *Environ. Res. Lett.* **9** (2014). doi:10.1088/1748-9326/9/1/014003.

[18] Wakazuki, Y., Nakamura, M., Kanada, S. & Muroi, C. Climatological reproducibility evaluation and future climate projection of extreme precipitation events in the Baiu Season using a high-resolution non-hydrostatic RCM in comparison with an AGCM. *J. Meteorol. Soc. Jpn* **86**, 951–967 (2008).

[19] Knote, C., Heinemann, G. & Rockel, B. Changes in weather extremes: Assessment of return values using high resolution climate simulations at convection-resolving scale. *Meteorol. Z.* **19**, 11–23 (2010). doi:10.1127/0941-2948/2010/0424.

[20] Trapp, R. J., Robinson, E. D., Baldwin, M. E., Diffenbaugh, N. S. & Schwedler, B. R. J. Regional climate of hazardous convective weather through high-resolution dynamical downscaling. *Clim. Dyn.* **37**, 677–688 (2011). doi:10.1007/s00382-010-0826-y.

[21] Hohenegger, C., Brockhaus, P., Bretherton, C. S. & Schär, C. The soil moisture-precipitation feedback in simulations with explicit and parameterized convection. *J. Climate* **22**, 5003–5020 (2009).

[22] Pan, L.-L. *et al.* Influences of climate change on California and Nevada regions revealed by a high-resolution dynamical downscaling study. *Clim. Dyn.* **37**, 2005–2020 (2011). doi:10.1007/s00382-010-0961-5.

[23] Golding, B. W. Nimrod: A system for generating automated very short range forecasts. *Meteorol. Appl.* **5**, 1–16 (1998). doi:10.1017/S1350482798000577.

[24] Harrison, D. L., Driscoll, S. J. & Kitchen, M. Improving precipitation estimates from weather radar using quality control and correction techniques. *Meteorol. Appl.* **7**, 135–144 (2000). doi:10.1017/S1350482700001468.

[25] Walters, D. N. *et al*. The Met Office Unified Model global atmosphere 3.0/3.1 and JULES global land 3.0/3.1 configurations. *Geosci. Model Devel.* **4**, 919–941 (2011). doi:10.5194/gmd-4-919-2011.

[26] Wilkinson, J. M. *et al*. Improved microphysical parametrization of drizzle and fog for operational forecasting using the Met Office Unifed Model. *Q. J. R. Meteorol. Soc.* **139**, 488–500 (2013). doi:10.1002/qj.1975.

[27] Li, D. & Shine, K. P. A 4-dimensional ozone climatology for UGAMP models. Technical Report 35, UGAMP (April 1995).

[28] Collins, W. J. *et al*. Development and evaluation of an Earth-System model - HadGEM2. *Geosci. Model Devel.* **4**, 1051–1075 (2011). doi:10.5194/gmd-4-1051-2011.

[29] Bower, K. N. & Choularton, T. W. A parametrisation of the effective radius of ice free clouds for use in global climate models. *Atmos. Res.* **27**, 305–339 (1992). doi:10.1016/0169-8095(92)90038-C.

[30] Best, M. J. *et al*. The Joint UK Land Environment Simulator (JULES), model description - Part 1: Energy and water fluxes. *Geosci. Model Devel.* **4**, 595–640 (2011). doi:10.5194/gmdd-4-595-2011.

Corresponding author

Correspondence to be addressed to Elizabeth Kendon.

Acknowledgements

Thanks to colleagues across the UK Met Office for their help setting up the 1.5km and 12km model experiments, especially Chang Wang, Jorge Bornemann and Wilfran Moufouma-Okia. Also thanks to Jonathan Wilkinson, Paul Field, Charlie Pilling and Humphrey Lean for useful discussions. We gratefully acknowledge funding from the Joint Department of Energy and Climate Change (DECC) and Department for Environment Food and Rural Affairs (Defra) Met Office Hadley Centre Climate Programme (GA01101). This work also forms part of a joint UK Met Office and Natural Environment Research Council (UKMO-NERC) funded project on Convective Extremes (CONVEX, NE/1006680/1).

Author contributions

E.J.K. carried out the 1.5km and 12km model experiments and wrote the paper. N.M.R. analysed performance of the 1.5km model from weather forecasts, produced Supplementary Figure 1, and along with H.J.F. extensively contributed to the manuscript. M.J.R. ran the 60km global model experiments. All authors discussed the results and commented on the manuscript.

Competing financial interests

The authors declare that they have no competing financial interests.

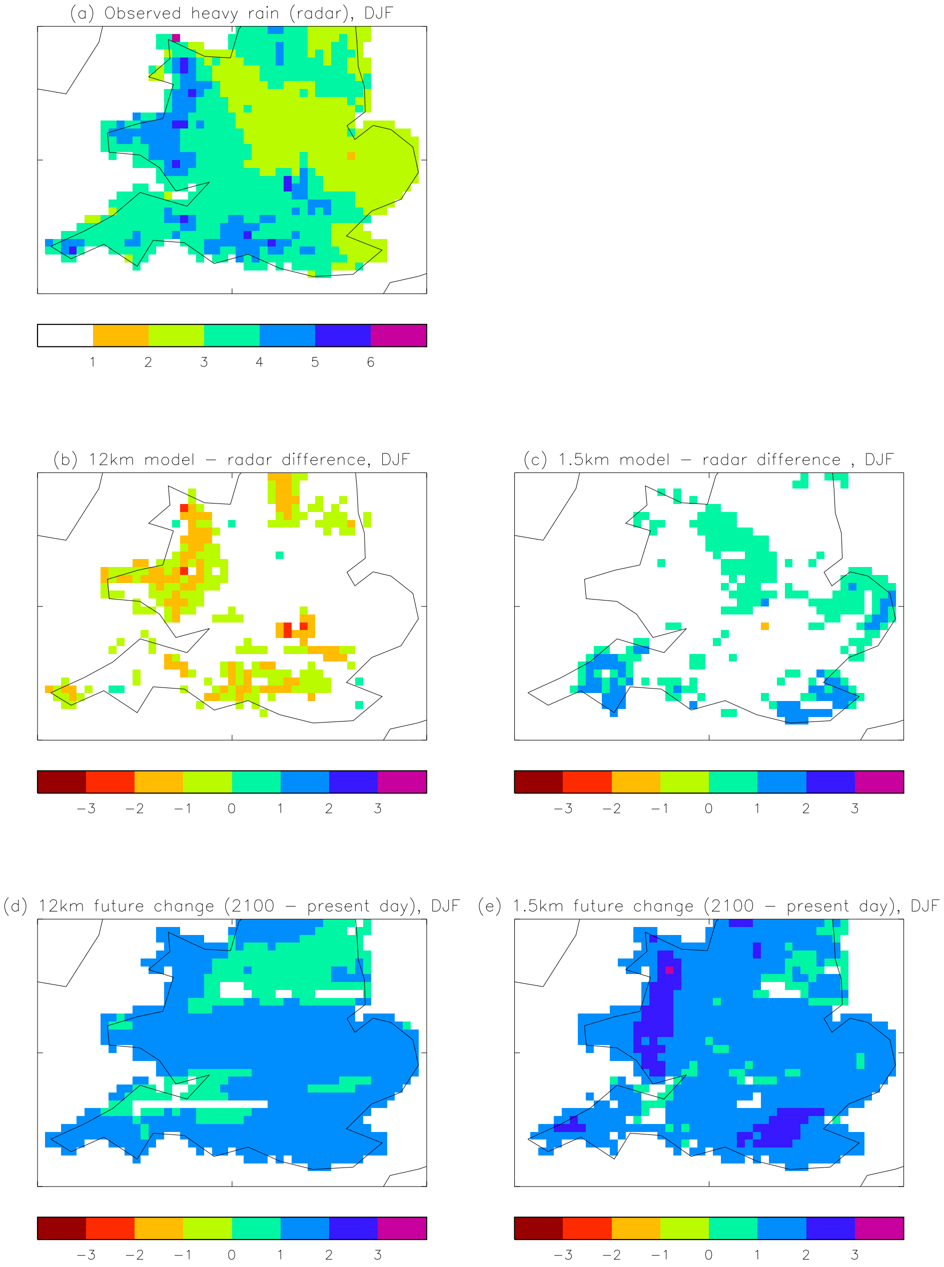


Figure 1: Heavy rainfall on hourly timescales (mm/h) in winter (December-January-February or DJF) in (a) radar, and (b,c) model-radar differences and (d,e) future changes for the 12km and 1.5km models. Heavy rainfall is defined as the mean of the upper 5% of wet values (>0.1mm/h). White indicates differences or future changes not significant at the 1% level compared to year-to-year variability. The radar data has been bias corrected using daily rain gauge data.

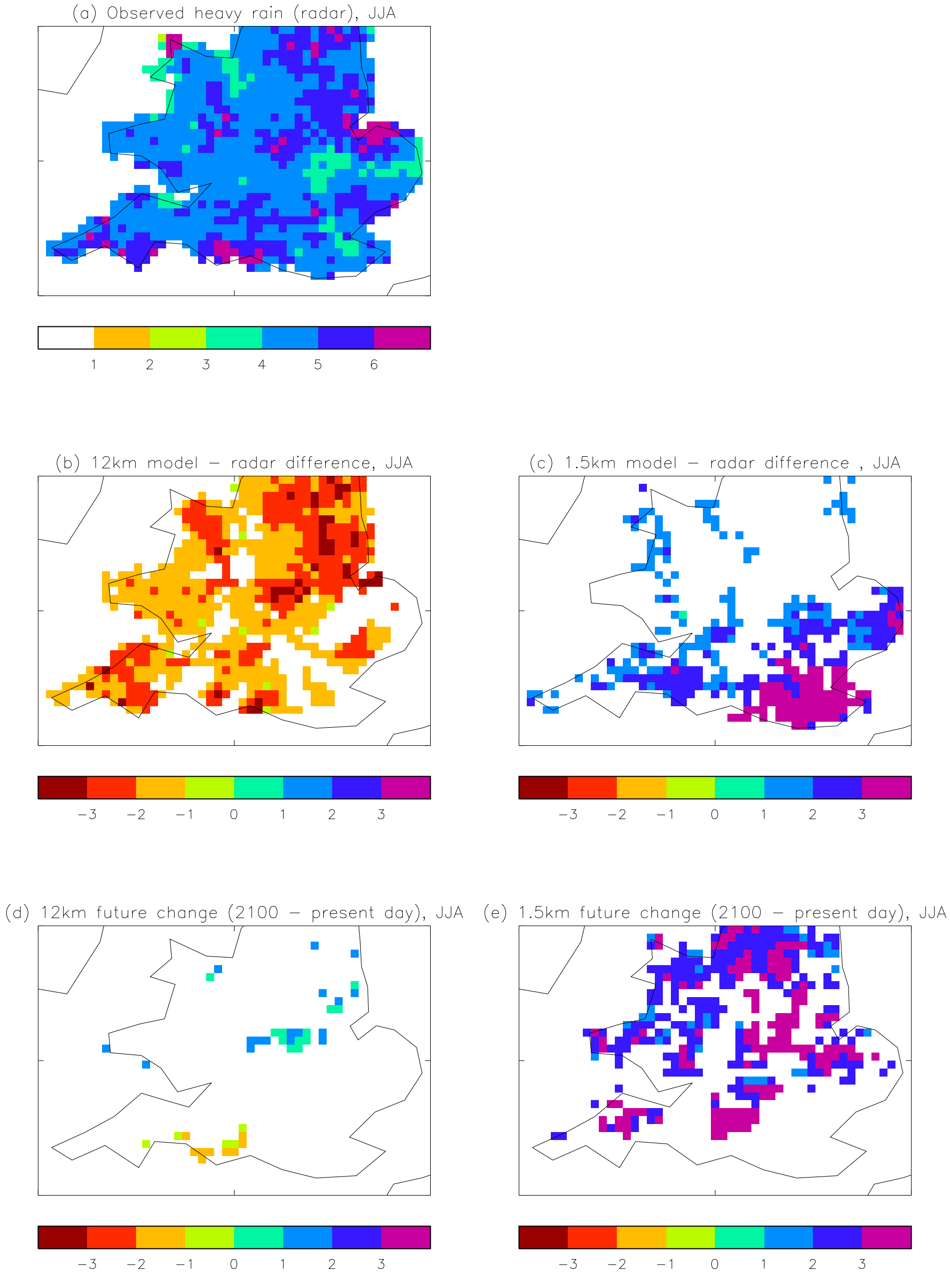


Figure 2: As in Figure 1 but for summer (June-July-August or JJA).

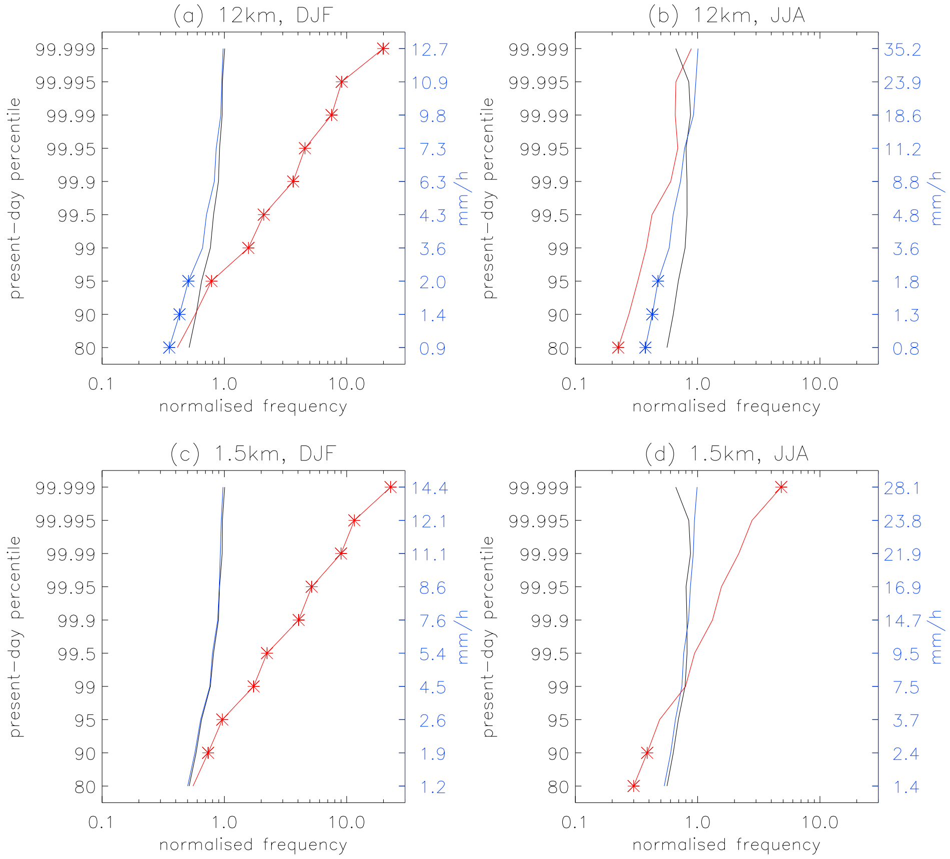


Figure 3: Frequency of episodes across the southern UK exceeding present-day percentiles of wet hourly precipitation. Results are shown for radar (black), present-day (blue) and future (red) model runs, for (a,b) 12km and (c,d) 1.5km models, for DJF and JJA. Frequency is measured as the number of episodes (continuous periods of exceedence above threshold) over all land points normalised by number of hours exceeding the threshold in the present-day. Percentile values (calculated for all grid boxes spatially pooled) are shown for the models on the right axis in blue. Blue (red) asterisks indicate model biases (future changes) significant at the 1% level.

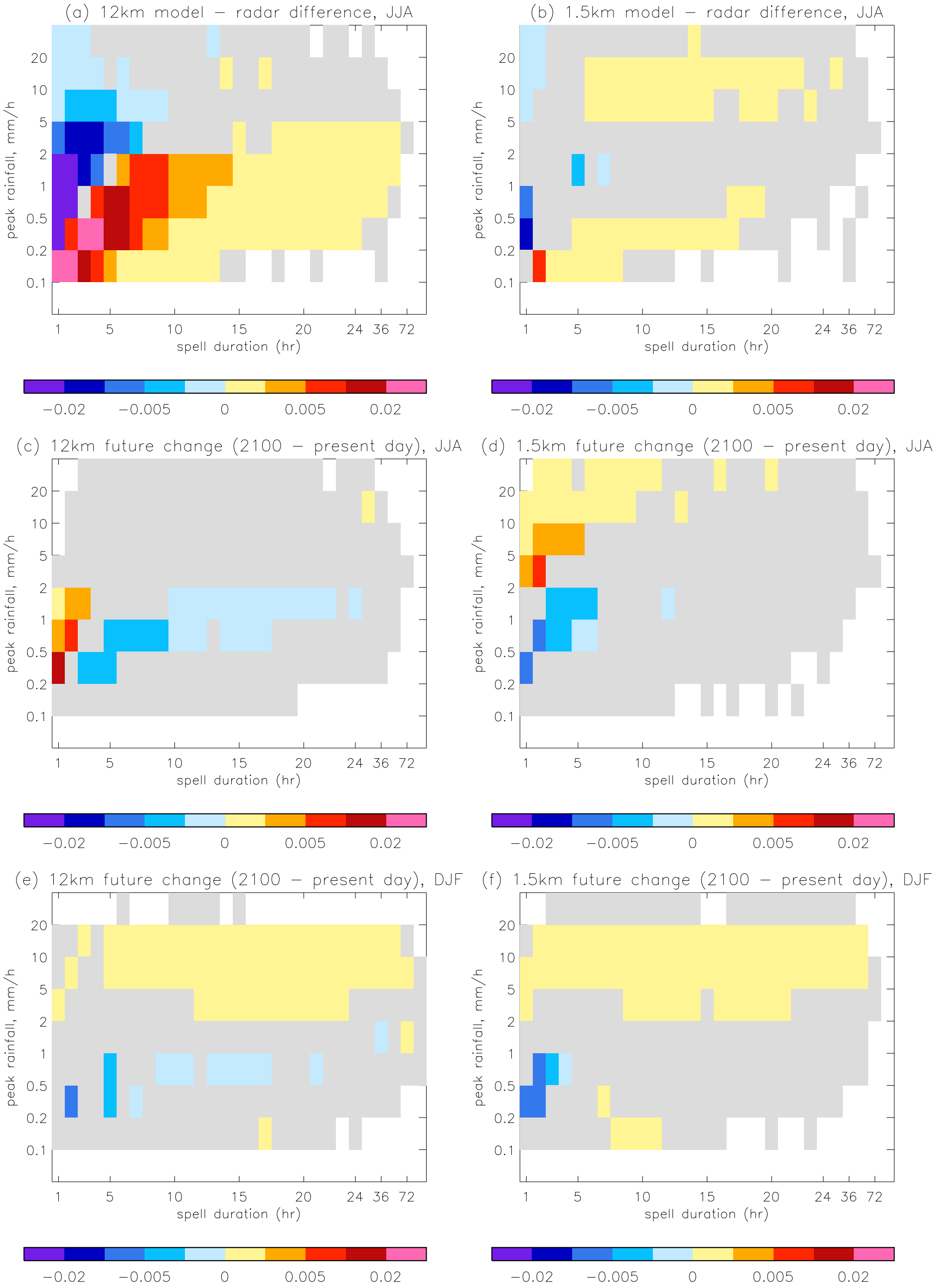


Figure 4: Model biases and future changes in the joint probability distribution of wet spell duration versus peak intensity over the southern UK. Shown are (a,b) model-radar differences for JJA and future changes for (c,d) JJA and (e,f) DJF, for the 12km and 1.5km models. Wet spells are defined as continuous periods when rain exceeds 0.1 mm/h. Peak intensity from the radar has been bias corrected using daily rain gauge data. Differences that are not significant at the 1% level are masked in grey.

Heavier summer downpours with climate change revealed by weather forecast resolution model

Supplementary Material

An example of a recent observed heavy rainfall event, similar to some events we see in the future in the 1.5km model, is shown in Supplementary Figure 1. This event on the 27th July 2013 was a very intense storm over France, which later moved north and became a wider area of rain over the UK. It can be seen that the high rainfall accumulations associated with the intense squall-line were well captured by the 1.5km forecast model (UKV), which has the same formulation as the 1.5km model used in this study except for its larger domain. By comparison the 12km forecast model is unable to capture the high rainfall accumulations associated with this event. The improvement in the 1.5km forecast model is typical for the simulation of convective storms. For this event, wet-bulb potential temperature (θ_w) was high over France (> 19°C at 850hPa), conditions which are seen over the UK much more often in the future (according to the simulations here driven by the 60km GCM). The fact that the 1.5km model responds well in these conditions, as illustrated by this observed event, gives us confidence in its ability to project the occurrence of such events in the future.

Wet bulb potential temperature (θ_w) is a combined measure of temperature and humidity. Very warm and humid southerlies for example would have a typical θ_w at 850hPa in excess of 16°C. For the heaviest (50) summertime rainfall events in the 1.5km model, θ_w at 850hPa is $3 - 4$°C higher in the future ($\sim$2100) compared to the present-day (1996-2009) simulation. This is the result of future increases in temperature (~ 4°C higher at 850hPa) and specific humidity (almost 30% higher), with relative humidity remaining approximately constant. For those events which exceed 30mm/h over a 12km grid box in the 1.5km model, θ_w at 850hPa is $> 17°$C (in both the present-day and future) suggesting that this may be around the minimum value required for the majority of the most extreme events over the UK. Although we note that such a threshold may not apply everywhere.

We note that averaging of the 1.5km model data to the 12km scale, carried out here to allow direct comparison with the 12km-RCM, will favour the selection of larger systems as the heaviest events. Small storms, even if very intense, may not reach high totals when averaged over 12km as low-intensity pixels are also included in the averaging. Nevertheless, it is appropriate to filter out grid-scale information from convection-permitting models. A benefit of averaging to the 12km scale is that it focusses on the events that produce high rainfall totals over a wide area (a large volume of water) that can lead to major and extensive flooding. The number of events (required to be at least one day apart to ensure independence) in summer with values exceeding 30mm/h over a 12km grid box is 6 in the present-day and 22 in the future for the 1.5km model, and 8 in the present-day and 4 in the future for the 12km-RCM. There are no events with values exceeding such a high threshold over a 12km grid box in either model in winter.

Radar data are used here to assess model biases in hourly rainfall for the present-climate. They provide the only currently available gridded hourly rainfall observations for the UK.[1] However, there are many issues with radar (clutter, anaprop, bright band, beam attenuation), and in particular radar data are known to systematically underestimate heavy rainfall amounts. The Met Office calibrates radar against rain gauges and employs algorithms to take account of known issues[2] but some problems cannot be fully rectified. One of these is that the hourly gauges used in the calibration are relatively sparse, and thus are not able to fully correct for locally-varying effects such as attenuation.

Despite possible large measurement errors, daily rainfall extremes from the radar are found to be comparable to those from gridded daily gauge data.[3] Additionally, the rate of increase of return levels with increasing return period (the 'growth curve' commonly used in hydrology) is comparable between radar data and hourly gauge station estimates.[3] This suggests that radar data can provide useful information for assessing model-simulated precipitation extremes. However, in this paper, where we are assessing hourly rainfall intensities we apply a bias correction to the radar data to try and account for errors due to beam attenuation.

The bias correction (see Methods) is a simple scaling approach similar to that used operationally, except we use daily (instead of hourly) gauge data as these have much better spatial coverage and are available as a gridded product.[4] The bias correction is dependent on the selection criterion (hourly amount is required to be $> x\times$ daily total in the radar, where x is chosen to be 0.3). If x is set too high (in the extreme case you could require the hourly amount = daily total in the radar, in which case the daily gauge is actually providing an exact and independent measure of the hourly amount), then the correction would only be applied to very short-duration events. However, if x is set too low, then the correction may be a poor estimation of the true radar bias for a given hour, since another heavy event may occur in the same day that dominates the bias in the daily total.

The impact of the bias correction on the results here is relatively small. The differences in average heavy rain intensity (upper 5% of wet hours) between corrected and uncorrected (raw) radar are largely insignificant (Supplementary Figure 2). There is some impact of the bias correction on peak rain intensity, with the corrected radar showing more frequent high peak intensity rain at short-durations (Supplementary Figure 3). By comparison, biases in the 12km-RCM are considerably larger and unaffected by any radar bias correction; whilst for the 1.5km model biases are of a similar magnitude and to a large degree are removed (particularly for peak-intensities at short-durations) with the radar correction applied (Figure 4 and Supplementary Figure 4).

We fully acknowledge that the correction here is far from perfect. Heavy localised showers are often missed by the gauges, and thus for these events we are unable to correct for radar attenuation. In addition, the requirement that the hourly amount is a major contributor to the daily total (x=0.3, necessary because we only have daily gauge data) means that a correction will not be applied to long-duration heavy rain. Nevertheless, the correction does give an indication of the likely contribution of radar error to apparent model biases.

Future intensification of rain in the 1.5km model in summer is found for durations of 1-5 hours (Figure 4). Duration here is examined at a given point in space, rather than in the frame of a moving storm. This approach is clearly of relevance to flooding, but has the disadvantage that a short spell of rain may result from either a small slow-moving shower or a larger fast-moving storm. In particular, when averaged to 12km, 1-5 hours of rain could come from a 24km rain area moving slowly or a 200km rain area moving quickly. Thus the fact that we see an intensification of short-duration rain is not inconsistent with many of the heaviest events in future being large scale storms.

Composites of the heaviest 50 events in the 1.5km and 12km models are shown in Supplementary Figure 5. These events correspond to those times with the 50 heaviest hourly rainfall accumulations in a 12km grid box somewhere in the southern UK, with selected events required to be at least one day apart to ensure independence. Each event centre is relocated to a common point to produce a 'typical' heavy event. It can be seen that increases in winter are widespread, whereas in the 1.5km model in summer increases are confined to the peak of the event. This points to different processes being responsible for increases in heavy rain intensity in the two seasons. In particular, increases in peak rainfall intensity in summer are linked to convection, whereas the more widespread increase in winter is consistent with enhanced large-scale moisture convergence. This explains why changes may be robust from coarser to higher model resolution in winter, but not in summer when deficiencies in the convective parameterisation scheme in coarse resolution models seriously impact projections of rainfall change.

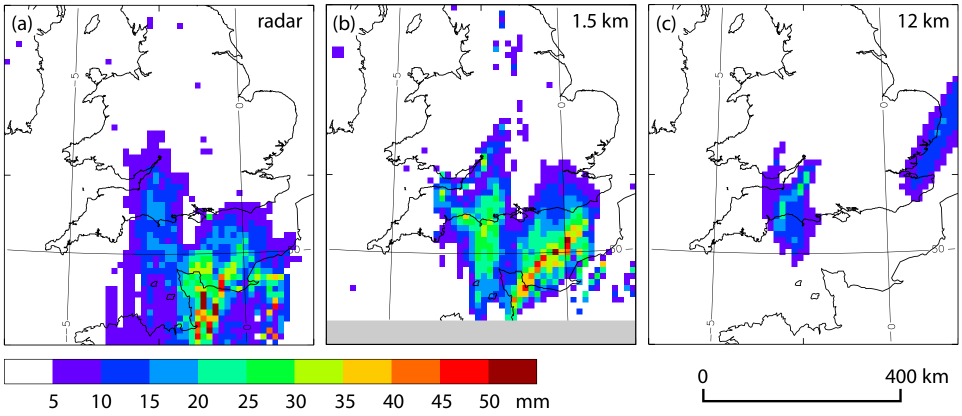
References

[1] Golding, B. W. Nimrod: A system for generating automated very short range forecasts. *Meteorol. Appl.* **5**, 1–16 (1998). doi:10.1017/S1350482798000577.

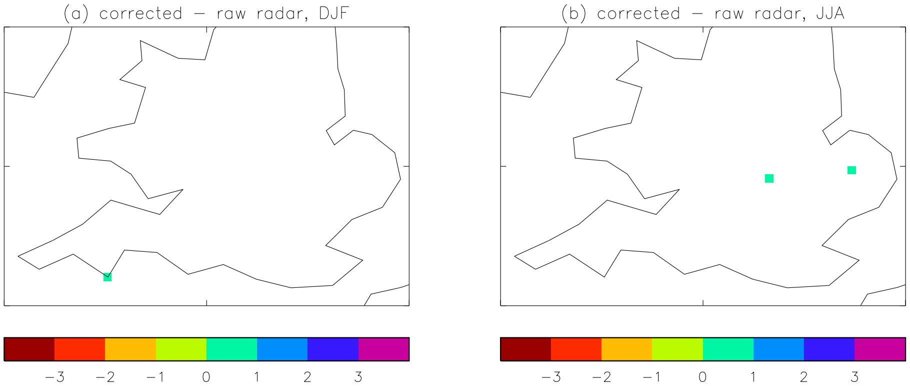
[2] Harrison, D. L., Driscoll, S. J. & Kitchen, M. Improving precipitation estimates from weather radar using quality control and correction techniques. *Meteorol. Appl.* **7**, 135–144 (2000). doi:10.1017/S1350482700001468.

[3] Chan, S. C. *et al*. The value of high-resolution Met Office regional climate models in the simulation of multi-hourly precipitation extremes. *J. Climate* (in press).

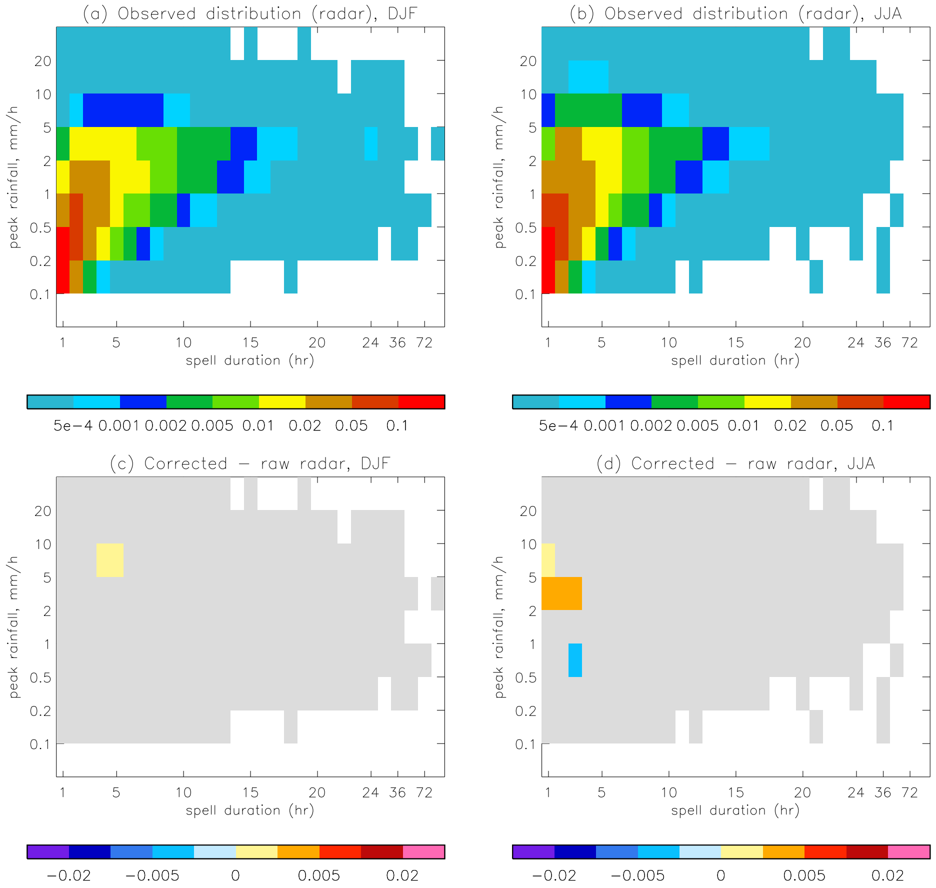
[4] Perry, M., Hollis, D. & Elms, M. *The Generation of Daily Gridded Datasets of Temperature and Rainfall for the UK*. Met Office National Climate Information Centre, FitzRoy Road, Exeter, Devon EX1 3PB, United Kingdom (June 2009).



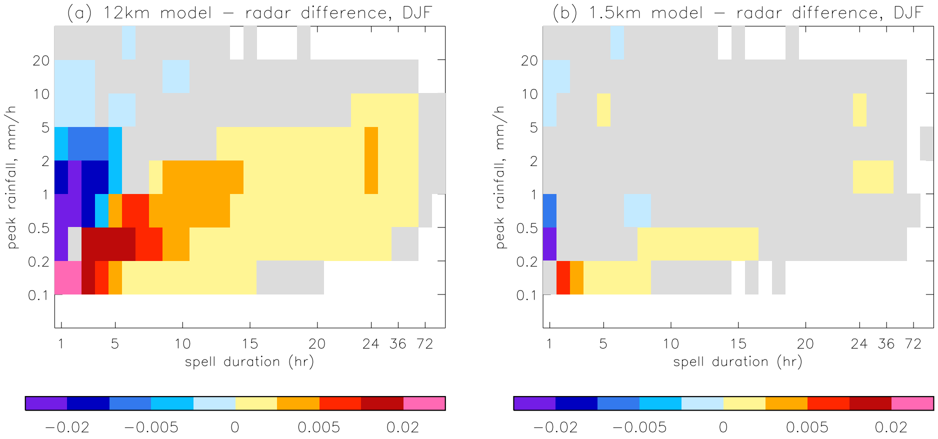
Supplementary Figure 1: Rainfall accumulations (mm) for the 5 hour period 13 to 18 UTC on 27th July 2013 for the (a) radar, (b) 1.5km UKV forecast model and (c) 12km forecast model. Forecasts were initiated at 21 UTC on the 26th (UKV) and 18 UTC on the 26th (12km). Earlier or later start times give similar forecast differences between the two models.



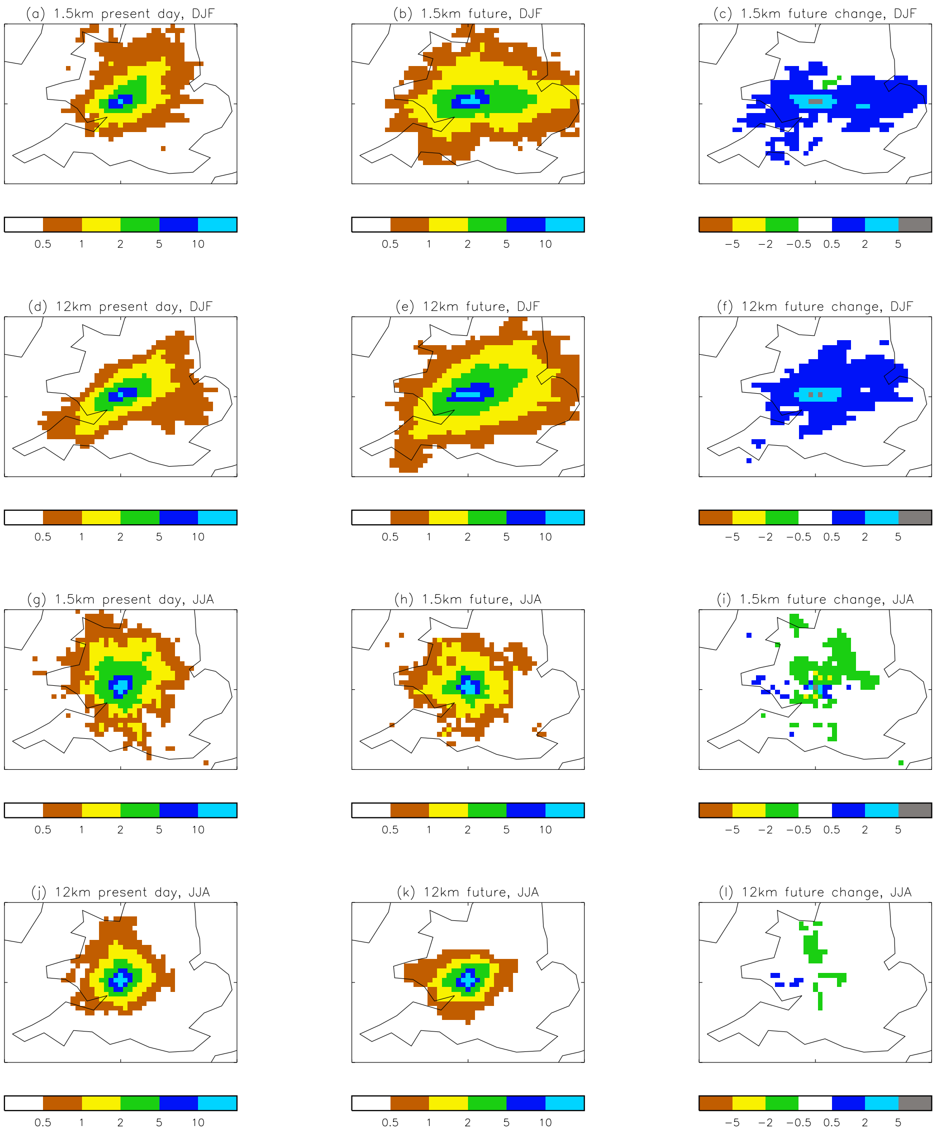
Supplementary Figure 2: Impact of bias correction on heavy rainfall (mm/h) in the radar, for (a) DJF and (b) JJA. Heavy rainfall is defined as the mean of the upper 5% of wet hours (>0.1mm/h). White indicates differences not significant at the 1% level compared to year-to-year variability.



Supplementary Figure 3: Joint probability distribution of wet spell duration versus peak intensity over the southern UK in the radar. Shown are (a,b) the observed distribution for the bias corrected radar data and (c,d) the impact of the bias correction on this distribution, for DJF and JJA. Wet spells are defined as continuous periods when rain exceeds 0.1 mm/h. Differences that are not significant at the 1% level are masked in grey.



Supplementary Figure 4: Model biases in the joint probability distribution of wet spell duration versus peak intensity over the southern UK, in the 12km and 1.5km models, for DJF. Wet spells are defined as continuous periods when rain exceeds 0.1 mm/h. Peak intensity from the radar has been bias corrected using daily rain gauge data. Differences that are not significant at the 1% level are masked in grey.



Supplementary Figure 5: Composites of the heaviest 50 events (mm/h) for the (a,d,g,j) present day and (b,e,h,k) future simulations and (c,f,i,l) their difference, in the 1.5km and 12km models, for DJF and JJA. Each event centre is relocated to a common point (at the centre of the 1.5km model domain) to produce a typical heavy event.