Validating UKCP18 model data against observations

1 Background

UKCP18 ~~contains both~~ ~~a~~ regional climate model (RCM) at 12km resolution and ~~a~~ convection-permitting model (CPM) at 2.2km resolution. This is the first ensemble of convection-permitting climate simulations run over the UK (Fosser et al., 2020).

RCMs are generally effective at simulating some aspects of the climate, such as wintertime precipitation intensity, or the seasonal average rainfall. However for areas of the climate where dynamics of fine scale convective storms are important, such as short duration rainfall extremes and summertime rainfall intensity, RCMs are less skillful (Figure 1). This is partly because the spatial scale of RCMs is generally larger than the size of convective storms, and so they must rely on convection parameterisation schemes to represent the net effect of convection on each grid cell. Parameterising deep convection accurately is difficult because it ~~involves~~ complex interactions between processes at scales from the microscale to the synoptic scale and representing its triggering, organisation and maintenance is complex. Consequently, RCMs with convection parameterisation schemes often poorly represent intense and localised convective precipitation. For instance, they have poor accuracy in capturing sub-daily precipitation, and the diurnal cycle often peaks too early in the day (Maraun et al., 2010). Rainfall biases are common, with light rainfall events occurring too often (the 'drizzle effect') ~~and~~ dry days too rarely and the intensity of extreme precipitation events frequently underestimated (Prein et al., 2015).

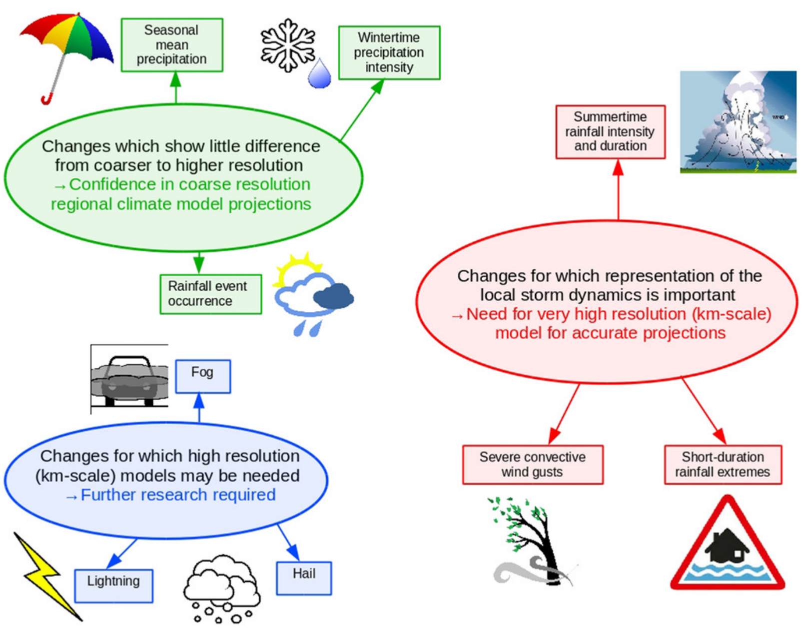
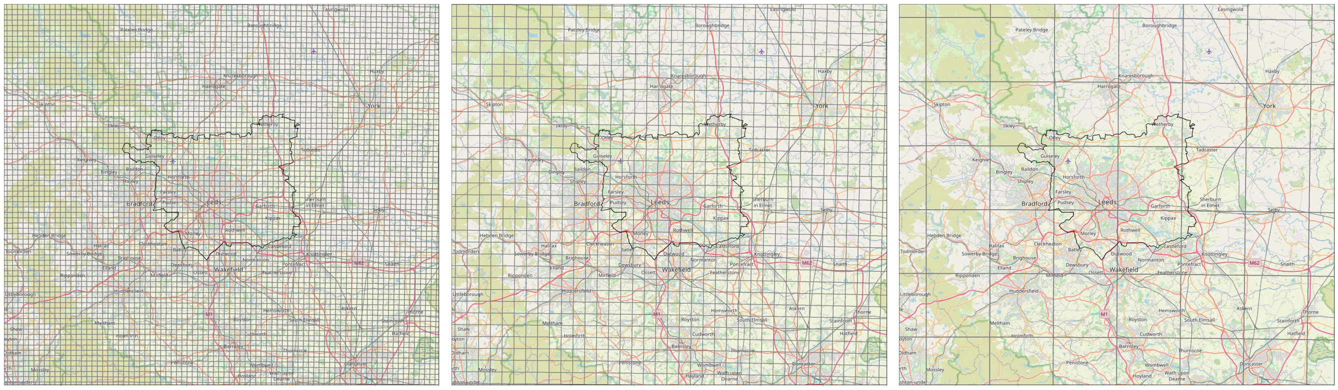


Figure 1: Schematic summarizing where we have confidence in coarse-resolution RCM projections and where very high-resolution (kilometer scale) models are needed for accurate projections (Kendon et al., 2017)

CPMs have finer horizontal resolutions (from around 5km down to 1 or 2km). This allows convective clouds to begin to be represented explicitly, and convection parameterisation schemes to be turned off. However, they are termed 'convection-permitting' because whilst they largely resolve larger storms and mesoscale convective organisations, convective plumes and small showers are still not represented. These models allow projections to be made for localised sub-daily extremes, and research on their use in Numerical Weather Prediction (NWP) indicates that the enhanced spatial resolution adds value in terms of the representation of the temporal and spatial distribution of precipitation. As convection is still not fully resolved, CPMs still do not represent convection perfectly and have a



((a)) CEH-GEAR ((b)) UKCP18 2.2km ((c)) UKCP18 12km

Figure 2: Positioning of grid cells over region around Leeds. The black outline is the Leeds City Council boundary

tendency to simulate heavy rainfall as too intense

2 UKCP precipitation projections for Leeds

This analysis aims to validate the local credibility of precipitation projections from UKCP18's CPM at 2.2km resolution over Leeds, and to compare them to projections from the RCM at 12km. Hourly precipitation projections from the models are compared to a set of 1km gridded hourly observations known as CEH-GEAR. These gridded observations are compiled by the Environment Agency through interpolation of quality-controlled rain gauge data. Hourly estimates of rainfall are derived through applying nearest neighbour interpolation to the rain gauge data, and then these hourly estimates are used to temporally disaggragate the CEH-GEAR daily rainfall dataset. They are on a 1km grid and cover the period of 1990-2014.

2.1 Differences in hourly precipitation intensities

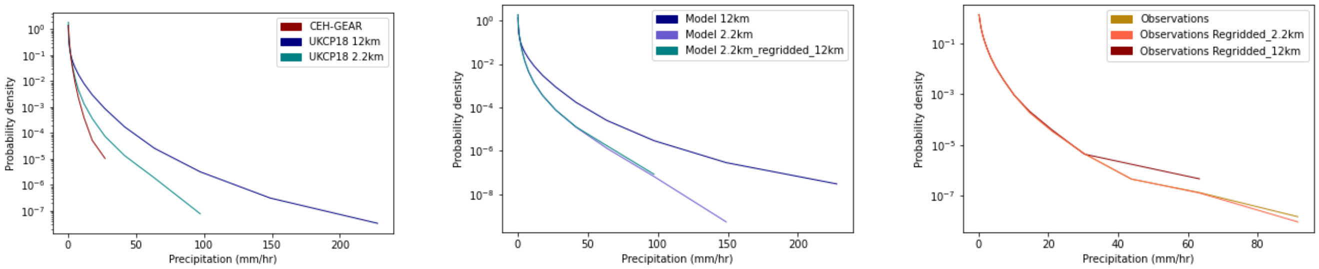
The data from all of the grid cells over the Leeds region are combined for CEH-GEAR, and for all of the grid cells and for all twelve ensemble members for UKCP18 2.2km and UKCP18 12km, to create three cumulative datasets for the region. These data sets are trimmed to include only the overlapping time period found in all three data sets (1990-2000), and only values in summertime, defined as June, July and August (JJA) are included. To allow ~~comparison between the range of precipitation intensities in the~~ 12km RCM, 2.2km CPM and 1km observations, all data sets were first regridded onto the 12km grid of UKCP18 12km.

2.1.1 High intensities

Probability distribution functions (PDFs) showing the proportion of hourly precipitation values which are of various intensities are plotted. The probability density is plotted on a log scale, which allows easier comparison of the probability of the highest precipitation values. Figure 3(a) compares the distributions of hourly precipitation intensities between CEH-GEAR, UKCP18 2.2km and UKCP18 12km. This demonstrates that UKCP18 2.2km has a wet-bias, predicting the heaviest events to be too intense; however, this wet-bias is considerably more pronounced in UKCP18 12km

For both CEH-GEAR and UKCP18 2.2km, the distribution of hourly precipitation intensities in the regridded data is compared to the data at the original resolution (Figure 3(b) and 3(c)). This provide assurance that the regridding process itself is not altering the distribution of values in either case.

Hourly precipitation accumulations corresponding to various percentile thresholds, for CEH-GEAR and both UKCP18 2.2km and UKCP18 12km, are shown in Figure 4. Figure 4(c) show that both models have higher values than CEH-GEAR from around the 90th percentile and above, indicating a wet-bias at higher precipitation intensities, and that this is particularly pronounced in UKCP18 12km.



((a)) CEH-GEAR ((b)) UKCP18 2.2km ((c)) UKCP18 12km

Figure 3: PDFs of hourly precipitation intensities. For UKCP18, data from the twelve ensemble members are combined and considered as one. Results correspond to years 1990-2001 which are the overlapping years between the model baseline period and CEH-GEAR-1km data

2.1.2 Lower intensities

Figure 4(a) and 4(b) shows the hourly precipitation intensities equivalent to just the lower percentiles, to allow the differences at lower intensities to be seen more clearly, The proportion of hourly precipitation intensities above and below various thresholds are also displayed in Table 1. In CEH-GEAR, a very high proportion (>90%) of hourly precipitation accumulations are 0mm/hr, compared to a much lower proportion (<10%) in UKCP18 2.2km and an extremely low proportion (<1%) in UKCP18 12km. The threshold often used to define 'wet-hours' is 0.1mm/hr. Considering this threshold, the proportion of 'dry' hours with <0.1mm/hr is the same (92.4%) in both CEH-GEAR and UKCP18 2.2km (Table 1). Consequently, at all percentiles up to around the 90th percentile, CEH-GEAR has a value of 0mm/hr, and UKCP18 2.2km has values very close to 0. In UKCP18 12km, a much smaller proportion (38.3%) of hours are classified as dry, and consequently the values associated with lower percentiles are much higher, being 0.28mm/hr and 5.72mm/hr for the 50th and 90th percentiles respectively (Figure 4). The proportion of hours with >1mm/hr is fairly similar in CEH-GEAR (2.2%) and UKCP18 2.2km (2.7%), and is quite different in UKCP18 12km (33.5%) (Table 1). Taken together, this all demonstrates that the UKCP18 12km exhibits a strong 'drizzle effect', with a much higher number of hours with light rainfall than in the observations, and that this effect is reduced significantly in UKCP18 2.2km.

The differences between CEH-GEAR, UKCP18 2.2km and UKCP18 12km in the distribution of hourly precipitation intensities <2mm/hr can be seen in more detail in Figure 5.

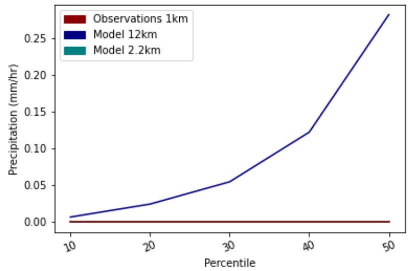
Table 1: Proportion of hours for various thresholds for CEH-GEAR, UKCP18 12km and UKCP18 2.2km

	0mm/h	>0 & <0.11mm/hr	>0 & <0.51mm/hr	>0 & <1mm/hr	>1mm/hr
CEH-GEAR (1km)	91.4%	1.0%	4.7%	6.2%	2.2%
UKCP18 (2.2km)	8.6%	83.8%	87.2 %	88.7%	2.7%
UKCP18 (12km)	0.4%	38.3%	57.1%	66.1%	33.5%

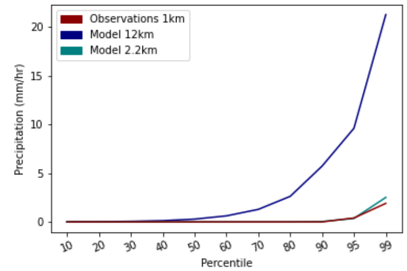
2.2 Regional differences in hourly precipitation statistics

Percentage differences between mean values from across the twelve UKCP18 2.2km model ensemble members and values from CEH-GEAR observations dataset, regridded to 2.2km, are calculated for various statistics of present day JJA rainfall, and plotted over the region around Leeds (Figure 6). The location of the rain gauges used to validate the CEH-GEAR gridded data are also included for reference.

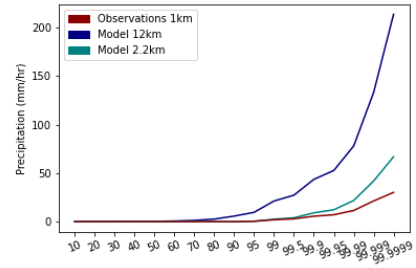
For the 99.5th, 99.75th and 99.9th percentiles and the mean and maximum precipitation intensities, the modelled values are higher than the observations across the whole of the region around Leeds. The maximum percentage difference between these values gets higher at higher percentiles, up to a maximum difference of 800% in the maximum values. The higher bias in the extremes compared to the mean may indicate that the wet bias is due to



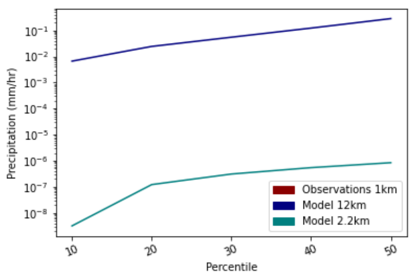
((a)) P10-P50



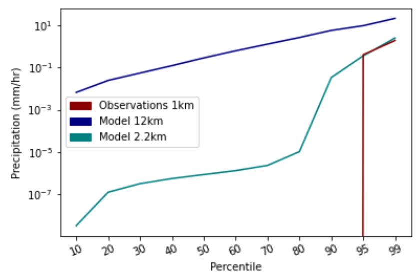
((b)) P10-P99



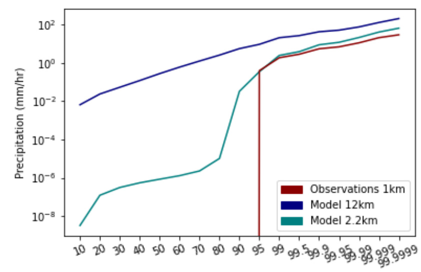
((c)) P10-P99.9999



((d)) P10-P50 (log scale)

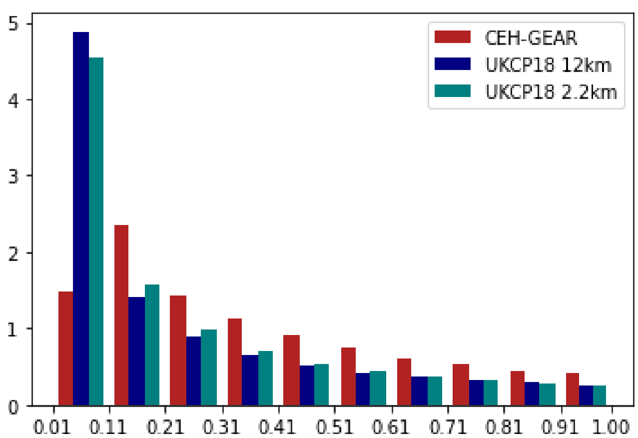


((e)) P10-P99 (log scale)

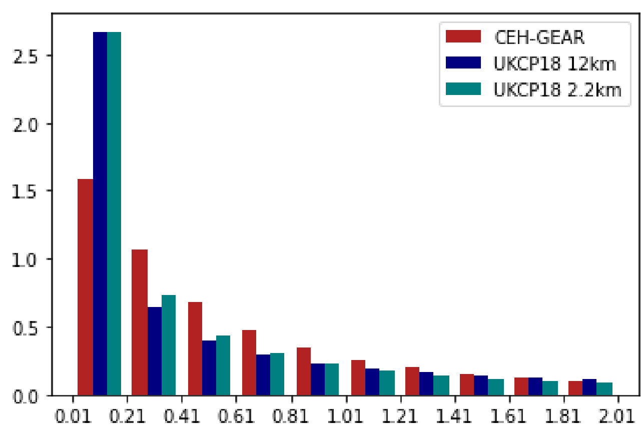


((f)) P10-P99.9999 (log scale)

Figure 4: Hourly rainfall accumulations (mm/hr) for given percentile thresholds using data from grid boxes across area surrounding Leeds shown in Figure 2 for CEH-GEAR, UKCP18 RCM and CPM. For UKCP18, data from the twelve ensemble members are combined and considered as one. Results correspond to years 1990-2001 which are the overlapping years between the model baseline period and CEH-GEAR-1km data



((a)) 0.01-1.01 (0.1 intervals)



((b)) 0.01-2.01 (0.2 intervals)

Figure 5: Histograms

the observations not capturing extremes due to smoothing, rather than a genuine wet bias in the model. However, if this was the case, it would be expected that rain gauge locations would match up with regions with the least bias, which does not seem to be the case.

The lower percentiles (P95 and P970) exhibit the opposite trend seen in the higher percentiles, with the model predicting lower precipitation values than found in the observations. This could be because whilst the model is predicting heavier hourly rainfall intensities during hours where it is raining, it may also contain a greater number of dry hours. As such, the model may be associated with higher hourly rainfall accumulations at higher percentiles, but lower percentiles might correspond to lighter rain or drier days than the observations. In contrast, the mean

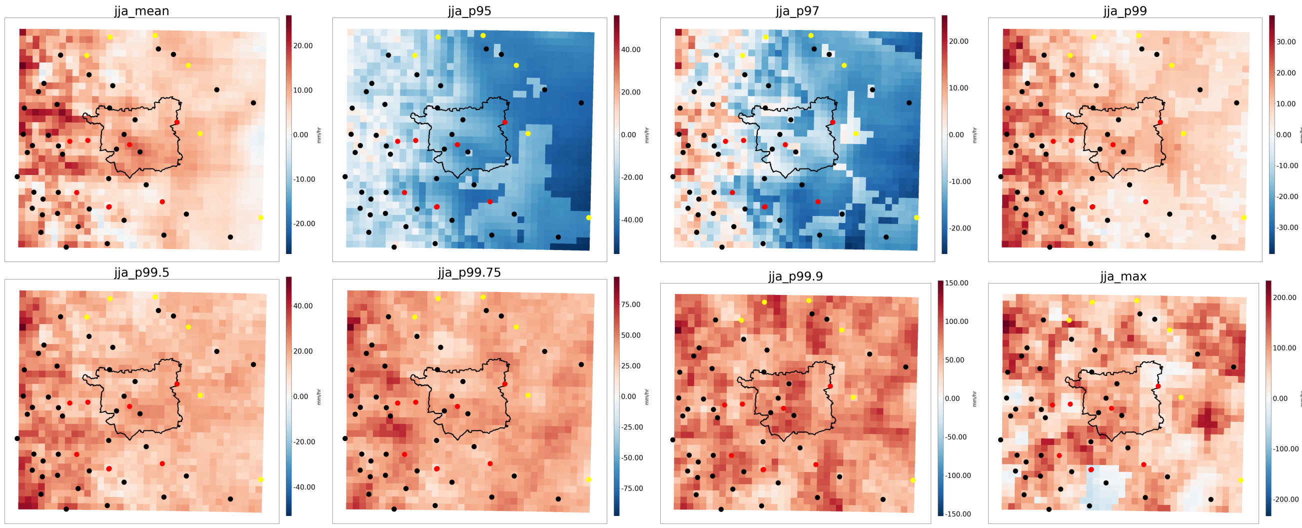


Figure 6: The percentage difference between UKCP18 2.2km local and CEH-GEAR observations regridded to 2.2km for various precipitation statistics. Percentage difference is calculated for each grid cell as ((model-obs)/obs)*100.

may equate to a much higher percentile, if the total rainfall amount is primarily contributed to by heavy rainfall events.

References

Fosser, Giorgia, Elizabeth Kendon, Steven Chan, Adrian Lock, Nigel Roberts, and Mike Bush. 2020. "Optimal configuration and resolution for the first convection-permitting ensemble of climate projections over the United Kingdom." *International Journal of Climatology* 40:3585–3606.

Kendon, Elizabeth J, Nikolina Ban, Nigel M Roberts, Hayley J Fowler, Malcolm J Roberts, Steven C Chan, Jason P Evans, Giorgia Fosser, and Jonathan M Wilkinson. 2017. "Do convection-permitting regional climate models improve projections of future precipitation change?" *Bulletin of the American Meteorological Society* 98:79–93.

Maraun, Douglas, F Wetterhall, AM Ireson, RE Chandler, EJ Kendon, Martin Widmann, S Brienen, HW Rust, T Sauter, M Themeßl, et al. 2010. "Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user." *Reviews of geophysics* 48.

Prein, Andreas F, Wolfgang Langhans, Giorgia Fosser, Andrew Ferrone, Nikolina Ban, Klaus Goergen, Michael Keller, Merja Tölle, Oliver Gutjahr, Frauke Feser, et al. 2015. "A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges." *Reviews of geophysics* 53:323–361.