Validating CEH-GEAR gridded observations against rain gauge data

1 Observation Datasets

1.1 CEH-GEAR 1km Gridded Observations Dataset

The CEH-GEAR 1km Gridded Observations "dataset contains 1km gridded estimates of hourly rainfall for Great-Britain for the period 1990-2014. The estimates are derived by applying the nearest neighbour interpolation method to a national database of hourly rain gauge observations collated by Newcastle University and the Centre for Ecology Hydrology (CEH). These interpolated hourly estimates were then used to temporally disaggregate the CEH-GEAR daily rainfall dataset. The estimated rainfall on a given hour refers to the rainfall amount accumulated in the previous hour" (Lewis et al., 2018)

1.2 Rain Gauge Data

An hourly rainfall dataset for the UK has been constructed using rain gauge data (Blenkinsop et al., 2017). This is based on around 1600 rain gauges from three different data sources, namely the Met Office Integrated Data Archive System (MIDAS), the UK Environment Agency (EA) and the Scottish Environmental Protection Agency (SEPA). The data from the gauges has been collated and subjected to additional quality-control procedures. The raw data used to produce this data set has been acquired for rain gauges within the Yorkshire region. There are six gauges within the Leeds city region, and Table 1 displays the time periods they cover, and the number of missing hourly values.

## 2 Validating CEH-GEAR 1km against raw rain gauge data

2.1 Using individual gauges within Leeds city boundaries

The CEH-GEAR dataset is interpolated and so it is likely that within it extreme values may have been smoothed. To investigate this, hourly precipitation intensities from the six rain gauges within the Leeds city boundary are compared against the values from the CEH-GEAR grid cells from within which they are found. The time period covered by each of these six rain gauges and the period of time which the rain gauge overlaps with the CEH-GEAR data set is shown in Table 1.

Table 1: Rain Gauges within Leeds City region, the time period they cover and the time period they cover which overlaps with CEH-GEAR. The number of hourly precipitation values in each of these time periods is also shown for both gauge and CEH-GEAR grid cell

|	Time period		Rain Gauge			CEH-GEAR	
	Full	Overlapping	Missing Values	Full	Overlapping	Full	Overlapping
Bramham	18/02/1987 - 20/06/2018	1990-2014	6404	268277	215879	219134	219134
Farnley Hall	12/10/1987 - 18/06/2018	1990-2014	1436	266129	215879	219134	219134
Otley	31/07/1985 - 20/06/2018	1990-2014	17575	270708	205649	219134	219134
Eccup	13/08/1986 - 21/06/2018	1990-2014	6026	273217	215652	219134	219134
Headingley	25/01/1996 - 20/06/2018	1996-2014	1078	195290	164894	219134	165974
Knostrop	13/08/1986 - 18/06/2018	1990-2014	16737	262447	204549	219134	219134

Figure 1 contains PDFs for each rain gauge within the Leeds City region using both the raw rain gauge data, and the data from the CEH-GEAR grid box within which the rain gauge is found. The left hand figures show the
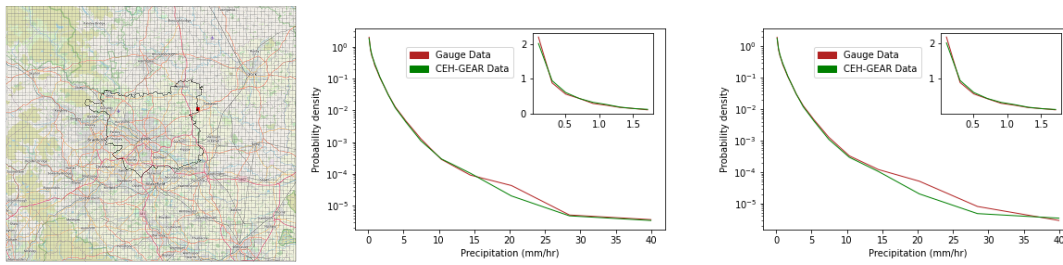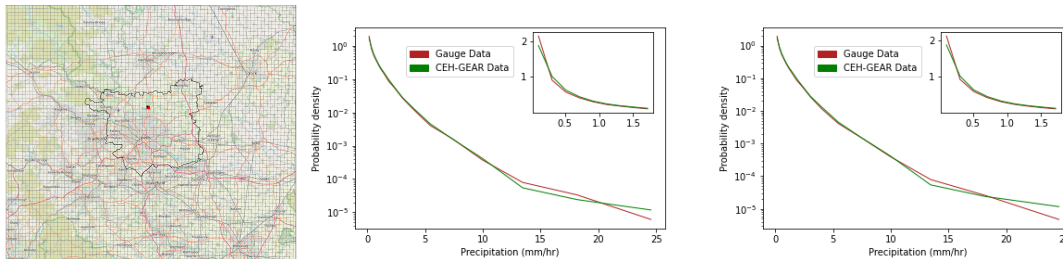
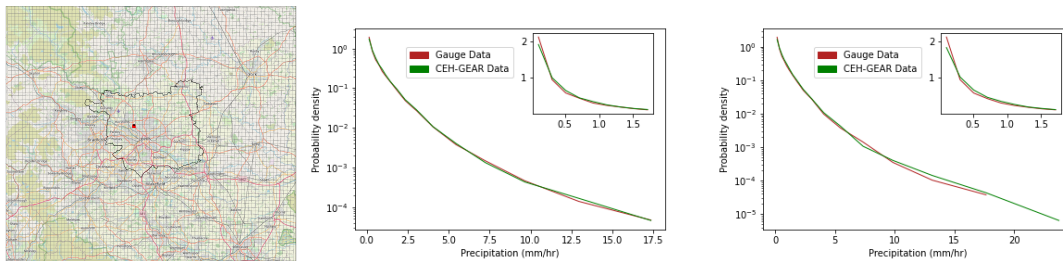(a) Bramham, gauge: 18/02/1987 - 20/06/2018 and CEH-GEAR: 1990-2014
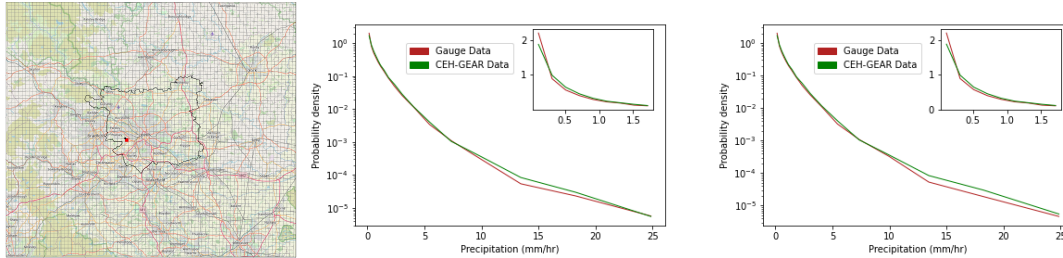
(b) Farnley Hall, gauge: 12/10/1987 - 18/06/2018 and CEH-GEAR: 1990-2014
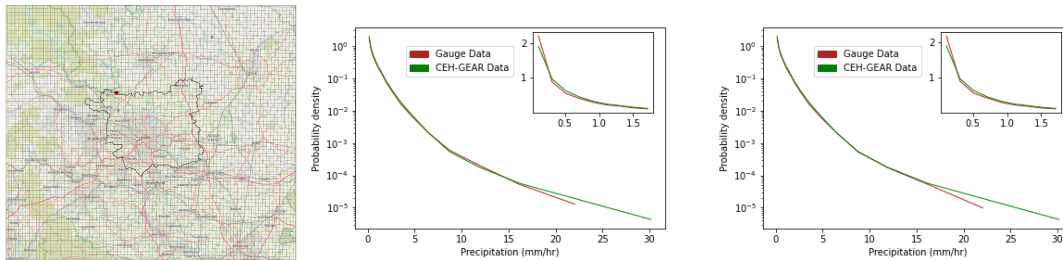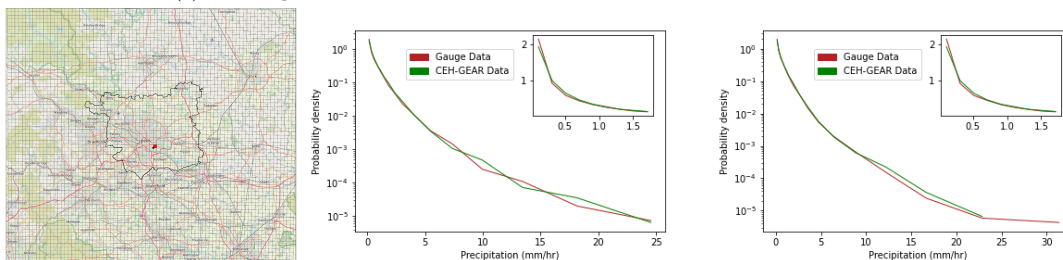

(c) Otley, gauge: 31/07/1985 - 20/06/2018 and CEH-GEAR: 1990-2014

(d) Eccup, gauge: 13/08/1986 - 21/06/2018 and CEH-GEAR: 1990-2014

(e) Headingle, gauge: 25/01/1996 - 20/06/2018 and CEH-GEAR: 1990-2014

(f) Knostrop, gauge: 13/08/1986 - 18/06/2018 and CEH-GEAR: 1990-2014

Figure 1: Location of six rain gauges within Leeds city region, and highlighted in red the CEH-GEAR grid cell they are found within. PDFs of hourly rainfall intensity, using the full time period (right) and the overlapping time period (middle)

locations of the rain gauges and the CEH-GEAR grid cells from which the data has been drawn. In the far right figure all the hourly precipitation values from both the rain gauge and CEH-GEAR grid cell are included. In the middle figure only precipitation values from within the overlapping time period are used.

Across the six rain gauges, there is a good degree of similarity between the rain gauge and the corresponding CEH-GEAR grid cell. The maximum intensity value in the gauge and the CEH-GEAR grid cell is roughly the same for the gauges at Bramham, Farnley Hall, Eccup, Headingley and Knostrop, whereas at Otley the CEH-GEAR grid cell contains a higher maximum intensity value. At Bramham the gauge generally has more precipitation values of higher intensities than the grid cell, whereas at Farnley Hall, Otley and Headingley the opposite is true, and at Eccup and Knostrop it switches over between the gaue and the grid cell at different precipitation values.

2.2 Combining data from all gauges across leeds-at-centre region

The period of time covered by individual rain gauges and their corresponding CEH-GEAR grid cells is generally too short to include many extreme values (Table 2), and thus to be instructive on how well the interpolated CEH-GEAR data set captures extremes.

A different approach is to pool the data from across all the rain gauges, and the CEH-GEAR grid cells they are found within, in the region surrounding Leeds (Figure 2(a)). This is a total of 41 rain gauges. All of the rain gauge records end in June 2018; however, they have variable lengths with some starting in the mid 1980s, and others only in the early 2000s. For each rain gauge the overlapping time period with the corresponding CEH-GEAR grid cell is calculated, and the hourly precipitation values from within this period for both the gauge and grid cell are added to the pools of data from across all 41 gauges. These pooled data sets contain a greater number of higher precipitation values (Table 3). A PDF of the pooled gauge data against pooled CEH-GEAR data is in Figure 2(b).

Table 2: Number of precipitation values over thresholds of 10, 15 and 20 mm/hr for each rain gauge and its corresponding CEH-GEAR grid cell

	Over 10mm/h		Over 15mm/h		Over 20mm/h	
	Gauge	Grid cell	Gauge	Grid cell	Gauge	Grid cell
Full time period						
Headingley	15	24	4	6	0	1
Bramham	43	22	15	7	4	3
Farnley Hall	21	19	4	5	2	1
Knostrop	20	24	4	5	2	1
Otley	29	24	5	4	2	1
Overlapping						
Headingley	15	9	4	4	0	0
Eccup	17	21	5	5	1	2
Bramham	2	3	1	0	0	0
Farnley Hall	16	19	4	5	2	1
Knostrop	15	24	3	5	1	1
Otley	25	24	4	4	2	1

Table 3: Number of precipitation values over thresholds of 10, 15 and 20 mm/hr for each rain gauge and its corresponding CEH-GEAR grid cell

	Over 10mm/h		Over 15mm/h		Over 20mm/h	
	Gauge	Grid cell	Gauge	Grid cell	Gauge	Grid cell
Full time period	920	1083	213	243	63	69
Overlapping	732	772	165	183	44	48

2.3 Investigating smoothing of extremes in CEH-GEAR away from gauges

The previous sections allow the extent to which the CEH-GEAR data set matches up with the raw rain gauge data from the locations where there are rain gauges to be explored. However, most of the smoothing is in fact likely to
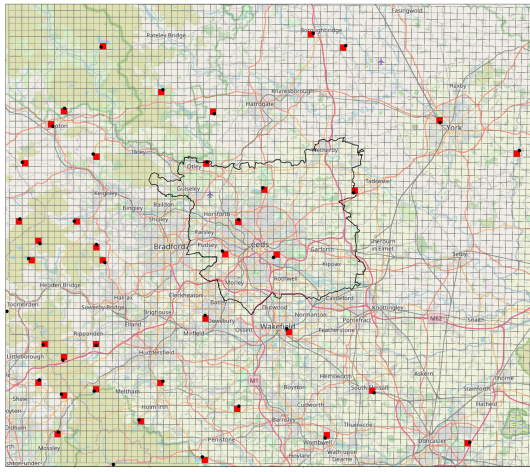
((a)) Rain gauges and corresponding CEH-GEAR grid cells
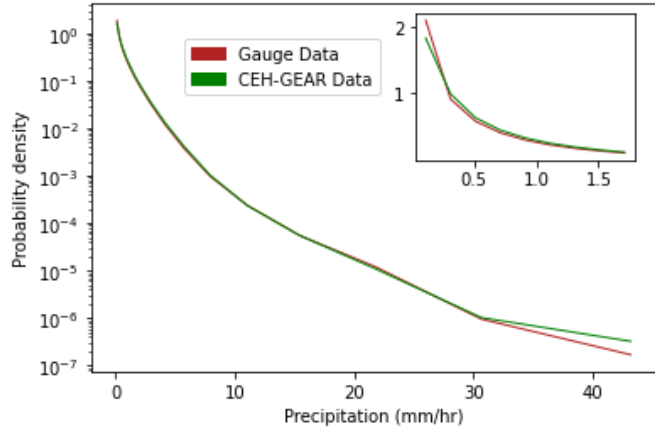
((b)) PDF

Figure 2: Pooled data from across all 41 gauges in Leeds region

occur in grid cells between gauges, where some kind of average will be taken from the values of the nearest rain gauges.

2.3.1 Comparing gauge locations to statistics of hourly precipitation intensities

To investigate whether the CEH-GEAR dataset suffers from smoothing out of extremes away from the gauges, values of hourly precipitation statistics are plotted and the location of the gauges overlain. This is to test whether there are higher values clustered around the station locations, and lower values in the gaps between gauge stations.

2.3.1.1 Defining gauge locations

The UKCP18 data includes a variable 'distance to gauge', which when plotted, allows the location of the gauges to be inferred. At different time slices, different gauges are used (Figure 3). The location of Environment Agency (EA) gauges, for which raw rain gauge data is available, are overlain in black, and Met Office (MO) gauges are overlain in red. These plots also indicate that there are some gauges additional to the MO and EA gauges which are used in constructing the CEH-GEAR data set. The locations of these are defined manually and are shown in yellow

((a)) 1990-01-09

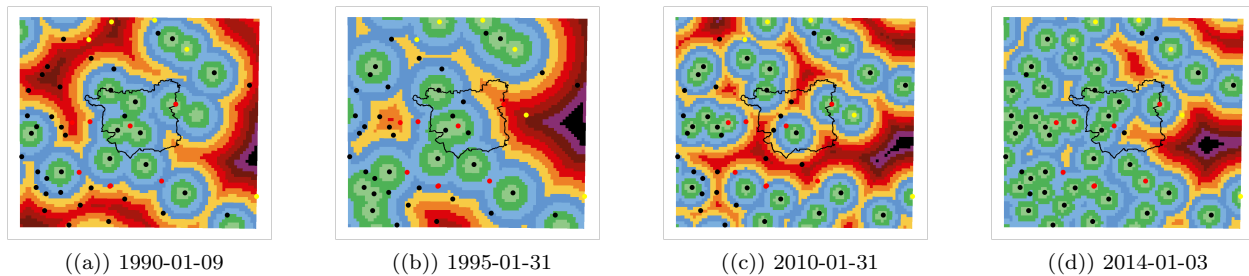
((b)) 1995-01-31

((c)) 2010-01-31

((d)) 2014-01-03

Figure 3: Locations of rain gauges, including EA gauges (black), Met Office gauges (red) and locations of other gauges included (yellow)

2.3.1.2 Plots of hourly precipitation statistics with gauge locations overlain

In Figure 4 the locations of the gauges defined in Section 2.3.1.1 are overlain over three different statistics of hourly precipitation intensity calculated from CEH-GEAR. Across the three plots there are no clear trends of lower precipitation intensities away from the gauges. There are some patterns apparent, particularly in the maximum plot, which are likely derived from the interpolation process used in CEH-GEAR (natural neighbour, Voronoi diagrams?)
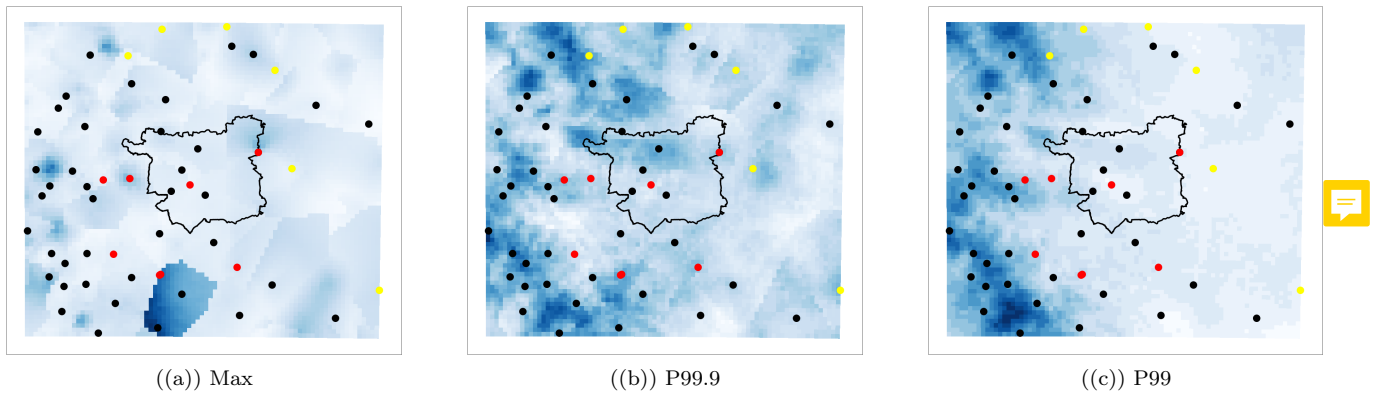
|((a)) Max|((b)) P99.9|((c)) P99|

Figure 4: Statistics of hourly precipitation intensity from CEH-GEAR, overlain with locations of rain gauges, including EA gauges (black), Met Office gauges (red) and locations of other gauges (yellow)

2.3.2 Hourly precipitation intensities at grid cells far from gauges

Another approach to determining the extent to which there is smoothing out of extremes at locations further away from gauges, is to select grid cells which are further away from any gauges and plot a PDF of hourly precipitation intensities for that grid cell, and for its 3 closest rain gauges. Analysis of Figure 5 indicates that values at CEH-GEAR grid cells not close to gauges are not consistently lower than the values at the gauges closest to them.
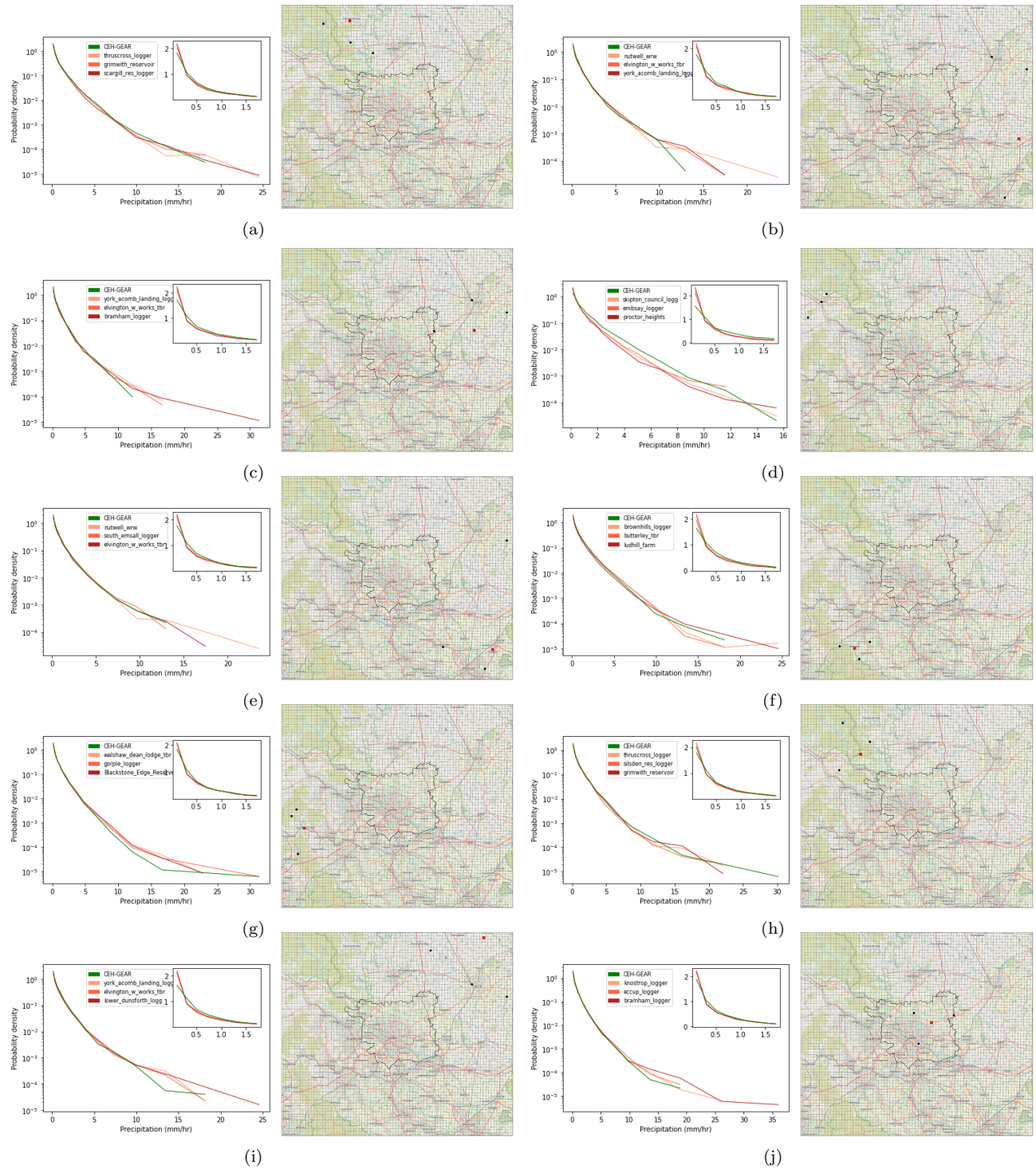
Figure 5: PDFs of hourly precipitation intensity for CEH-GEAR grid cells which are not located near a rain gauge, and the three closest rain gauges. The location of the grid cells is marked in red, and the gauges in black, in the right hand plot

References

Blenkinsop, Stephen, Elizabeth Lewis, Steven C Chan, and Hayley J Fowler. 2017. "Quality-control of an hourly rainfall dataset and climatology of extremes for the UK." *International Journal of Climatology* 37:722–740.

Lewis, Elizabeth, Niall Quinn, Stephen Blenkinsop, Hayley J Fowler, Jim Freer, Maliko Tanguy, Olivia Hitt, Gemma Coxon, Paul Bates, and Ross Woods. 2018. "A rule based quality control method for hourly rainfall data and a 1 km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr." *Journal of hydrology* 564:930–943.