A global sub-daily rainfall dataset

Dr Elizabeth Lewis
(Elizabeth.lewis2@ncl.ac.uk)

Prof Hayley Fowler, Dr Stephen Blenkinsop, Dr Renaud Barbero, Dr Selma Guerreiro, Dr Xiaofeng Li

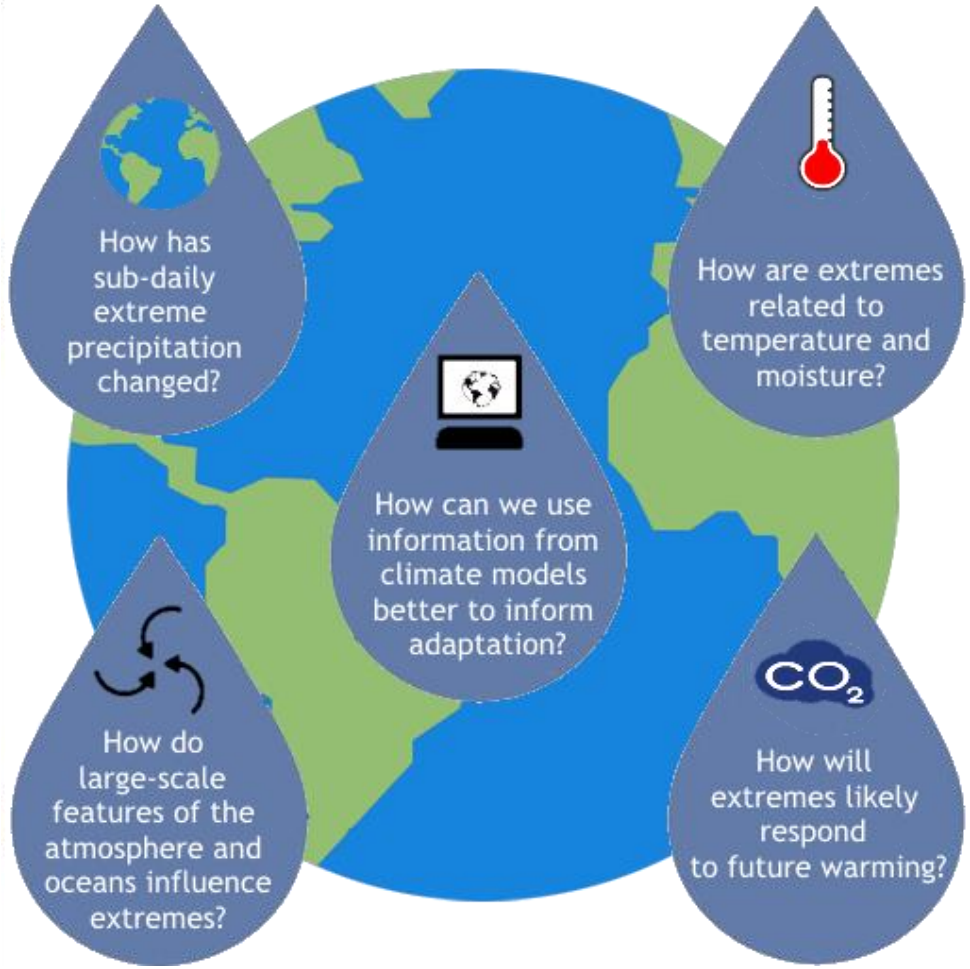
For more details visit: http://research.ncl.ac.uk/intense/

Our main aim…

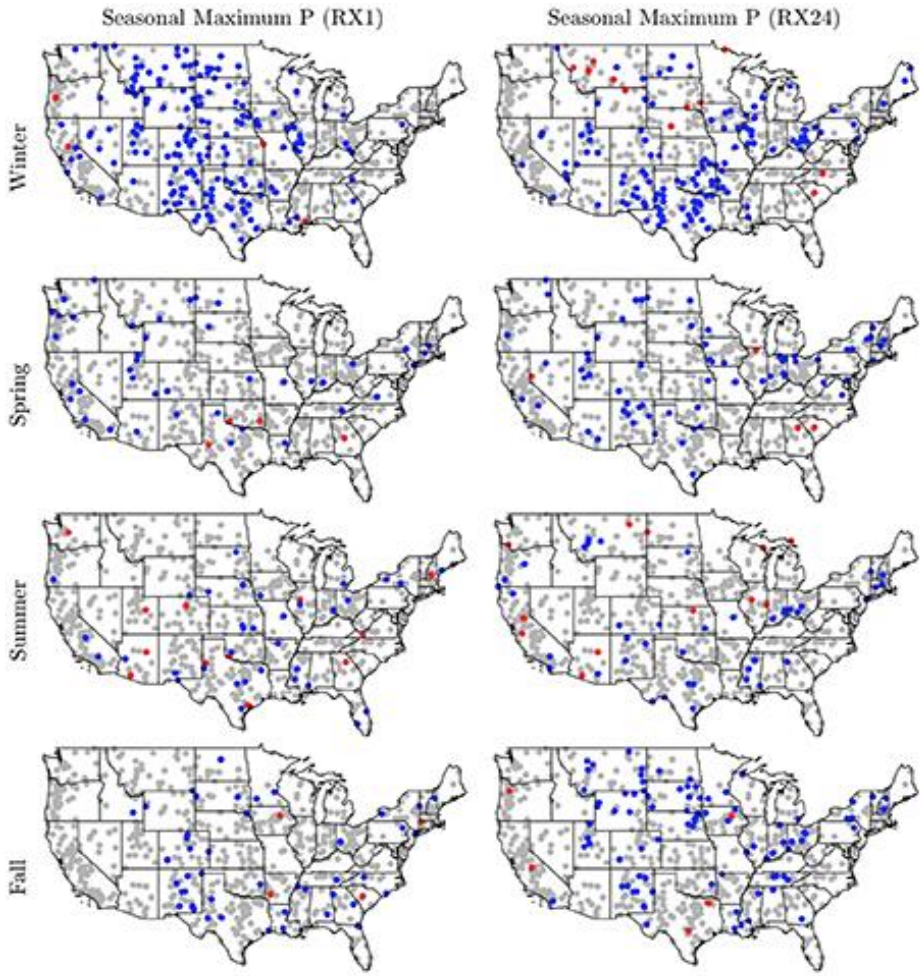
INTENSE aims to understand the nature and drivers of extreme sub-daily rainfall

INTENSE: INTElligent use of climate models for adaptatioN to non-Stationary hydrological Extremes

INTENSE aims to answer these questions…

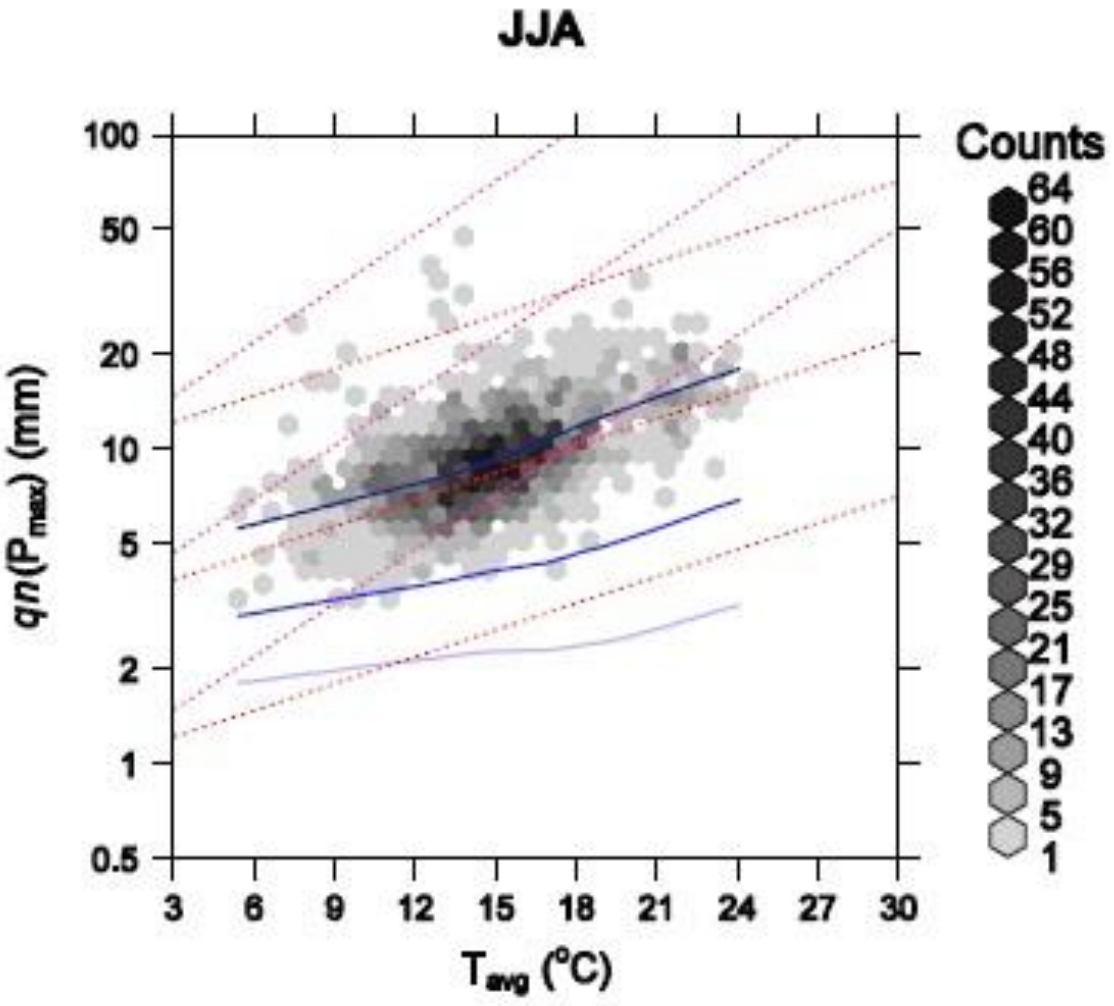


We are looking for trends in extreme rainfall…



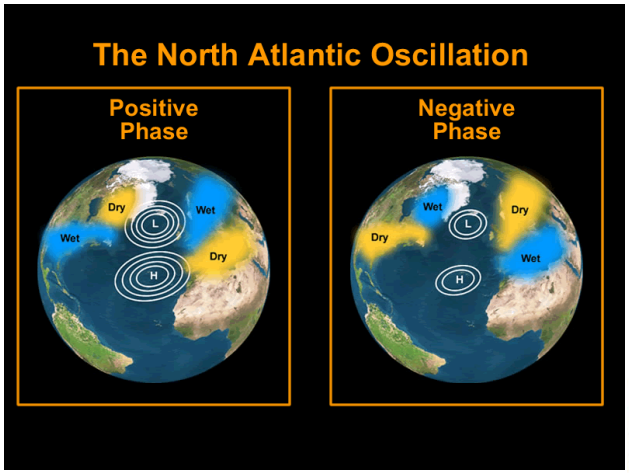
The blue (red) dots indicate stations with statistically significant increasing (decreasing) trends at the 5% level according to the Mann-Kendall test. The grey circles refer to the location of the stations that did not experience statistically significant changes at the 5% level.

Temperature scaling of extreme precipitation…

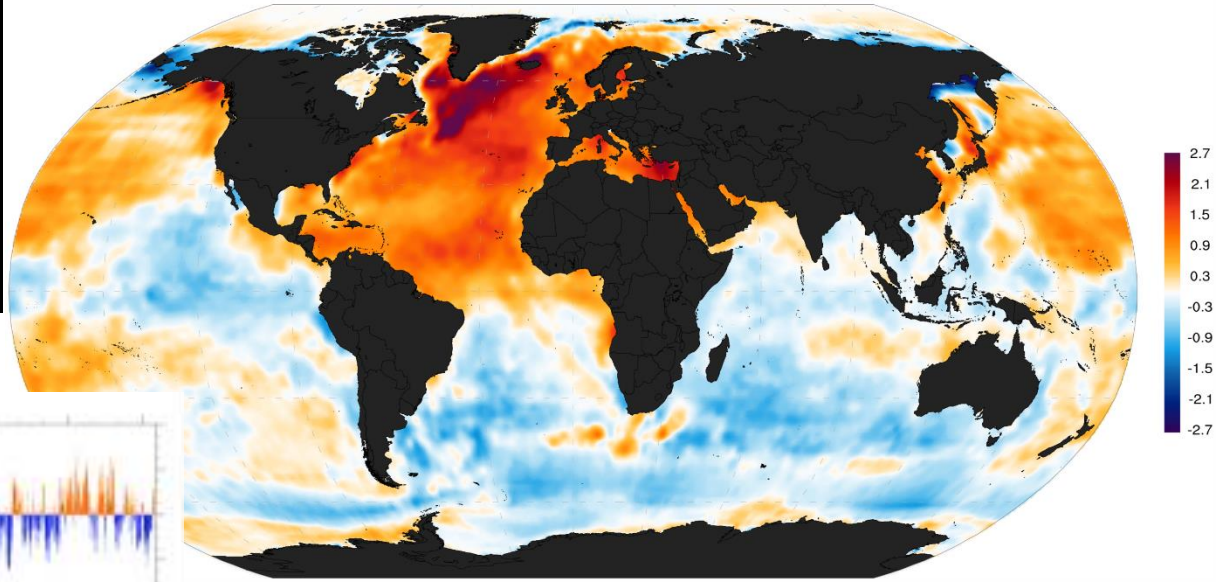


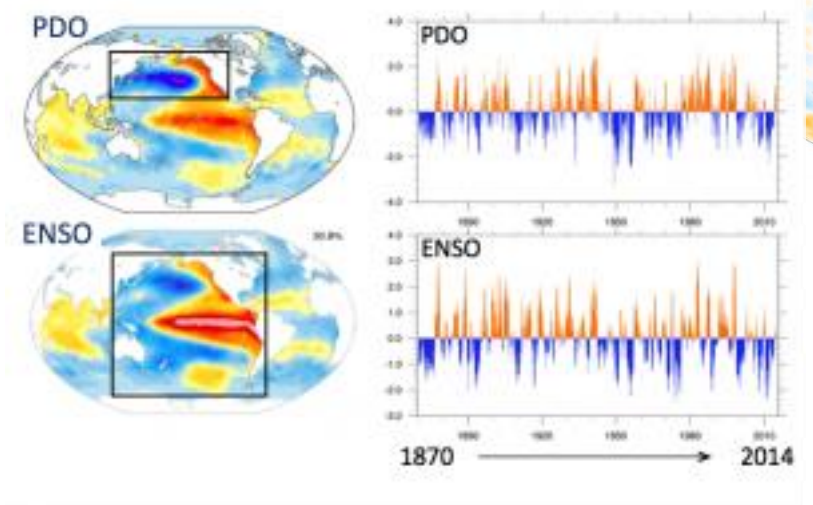
*The density plot shows the relationship between mean daily temperature - T_{avg} and the n*th quantile of *maximum hourly rainfall for events binned by temperature - $qn(P_{max})$ for UK gauges (JJA). Blue lines indicate relationships for n=99, 75 and 50.*

And the impact of large scale circulation patterns on extreme rainfall…



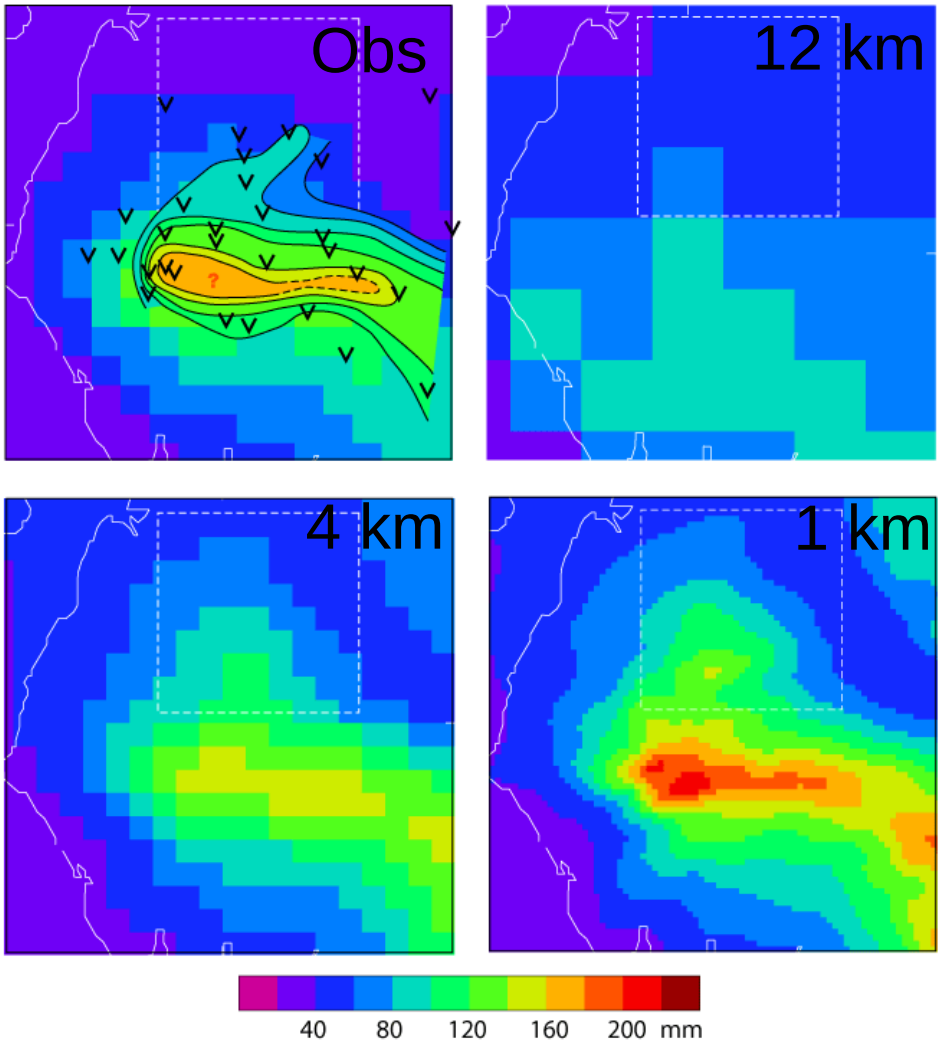
Atlantic Multi-decadal Oscillation(AMO)



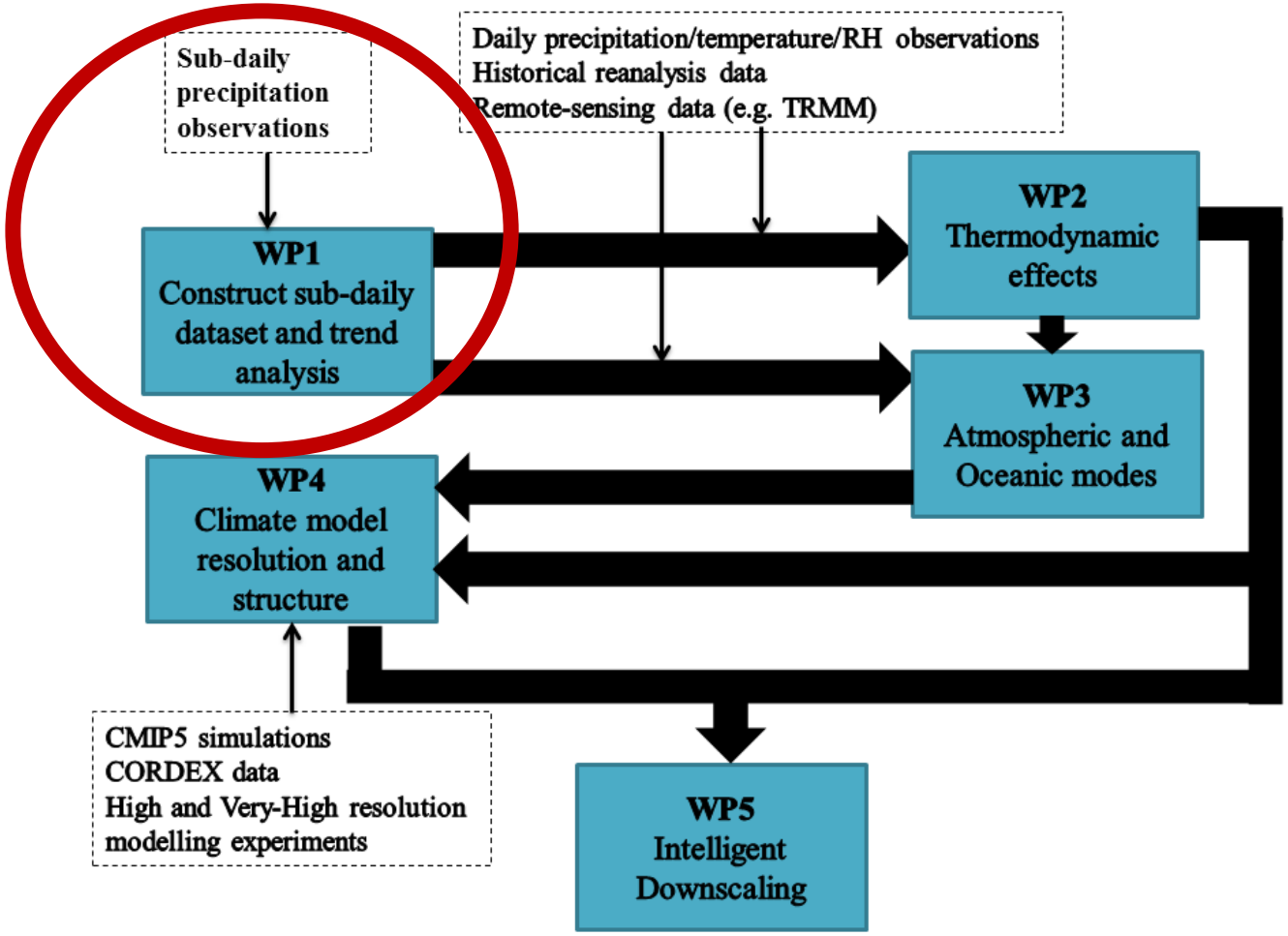


Pacific decadal Oscillation (PDO)

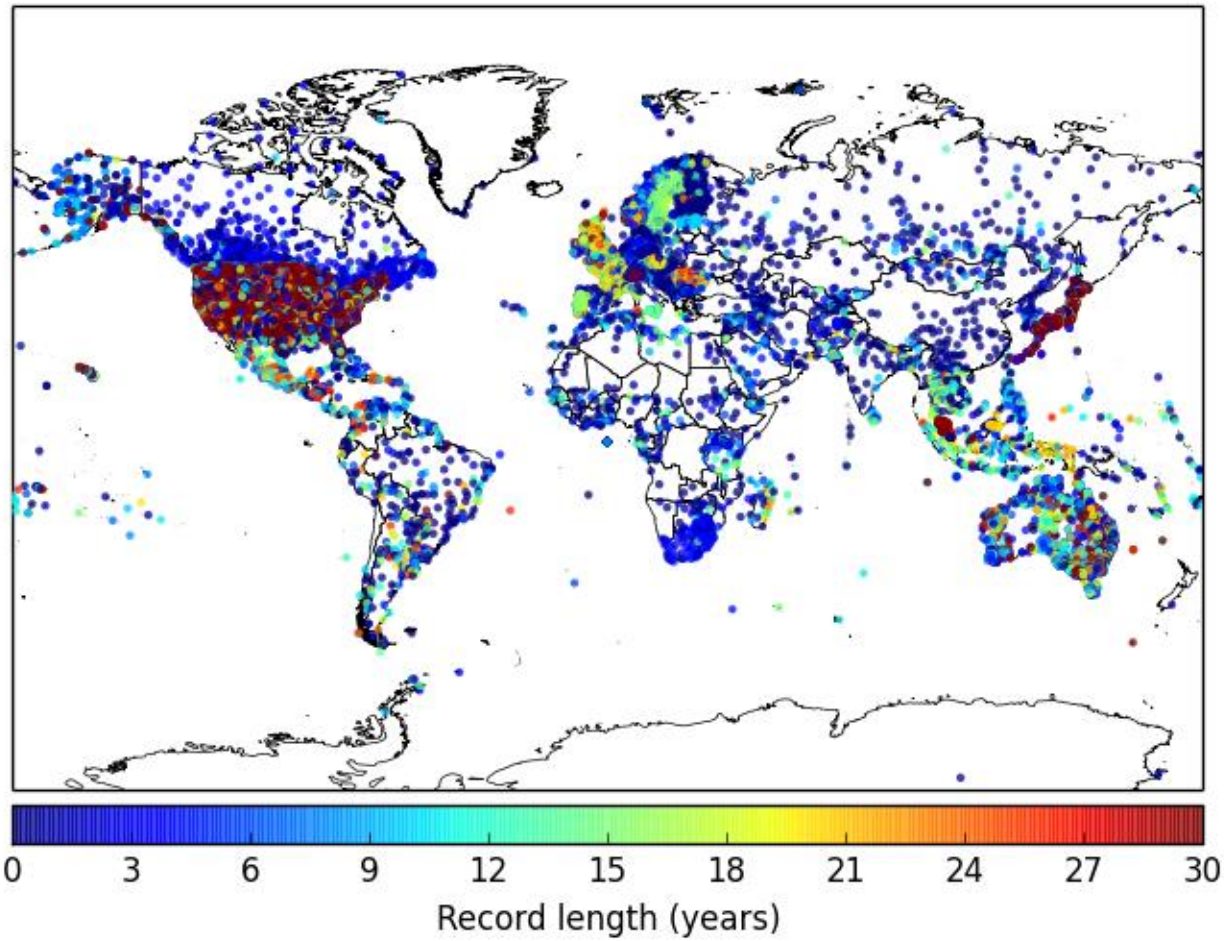
We will use this knowledge to improve/validate high resolution climate models…



We are following this work plan…

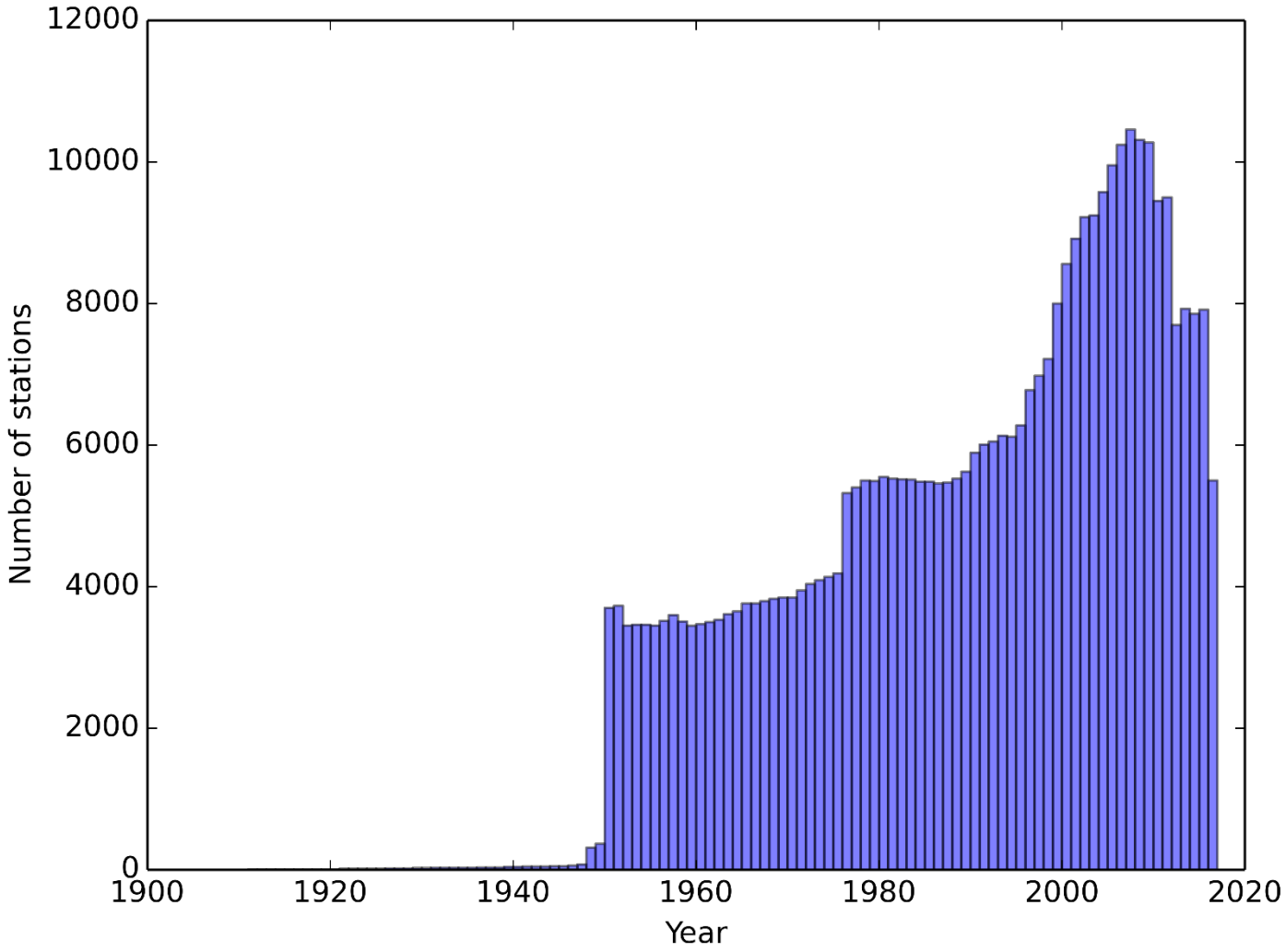


So far, have collected data from ~25,000 stations…

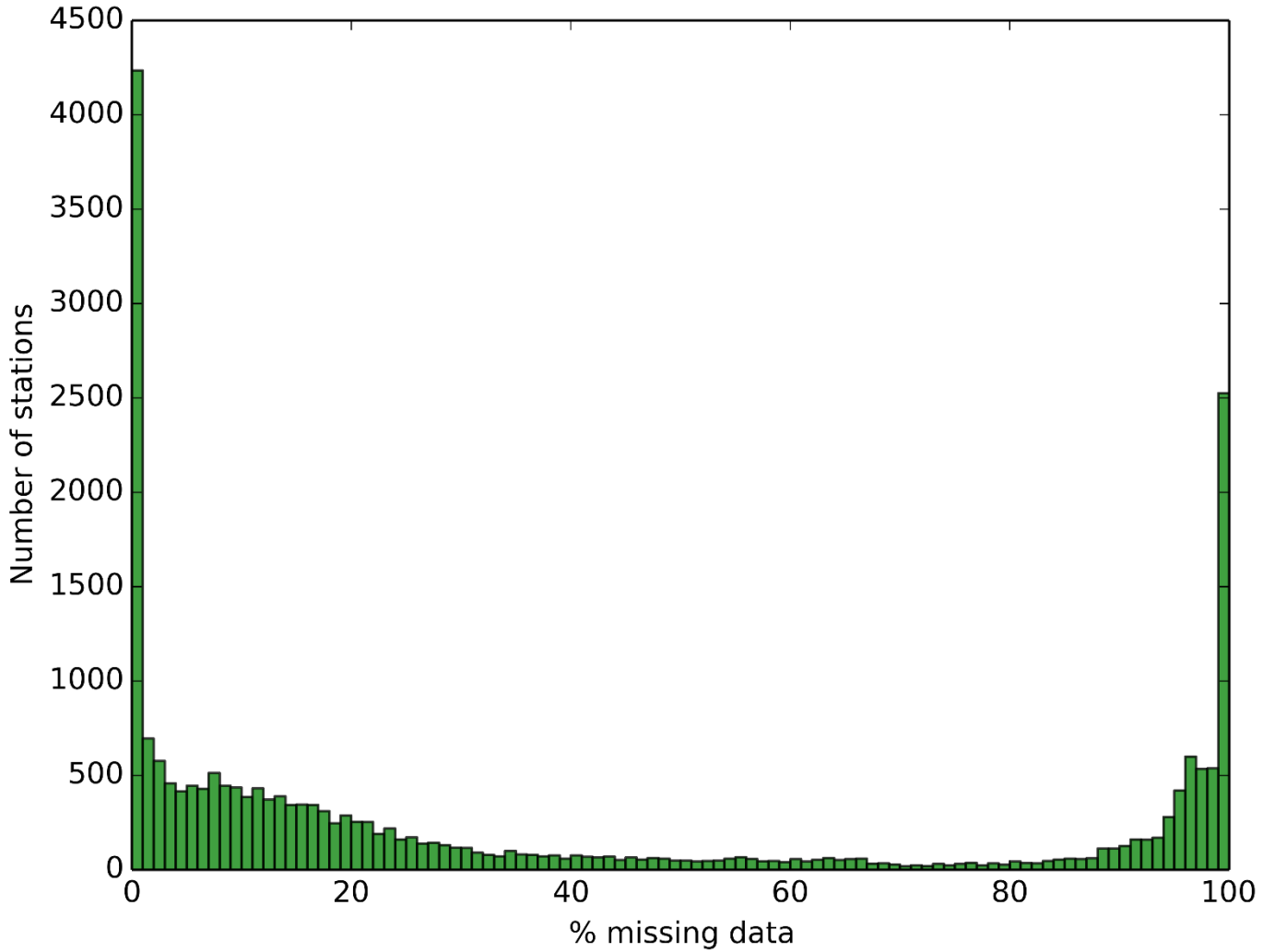


- Got: UK, US, Canada, Brazil, France, Germany, Spain, Portugal, Italy, Philippines, Norway, Sweden, The Netherlands, Finland, Australia, Kenya, Indonesia, Slovenia, Costa Rica, Switzerland, Austria, Hungary, Panama, Ireland, Japan, Malaysia, Singapore, Dominica, Trinidad & Tobago
- Getting: Spain, India, Argentina, Ecuador, Colombia, Bahamas, Mexico, New Zealand, China?
- Global datasets: ISD

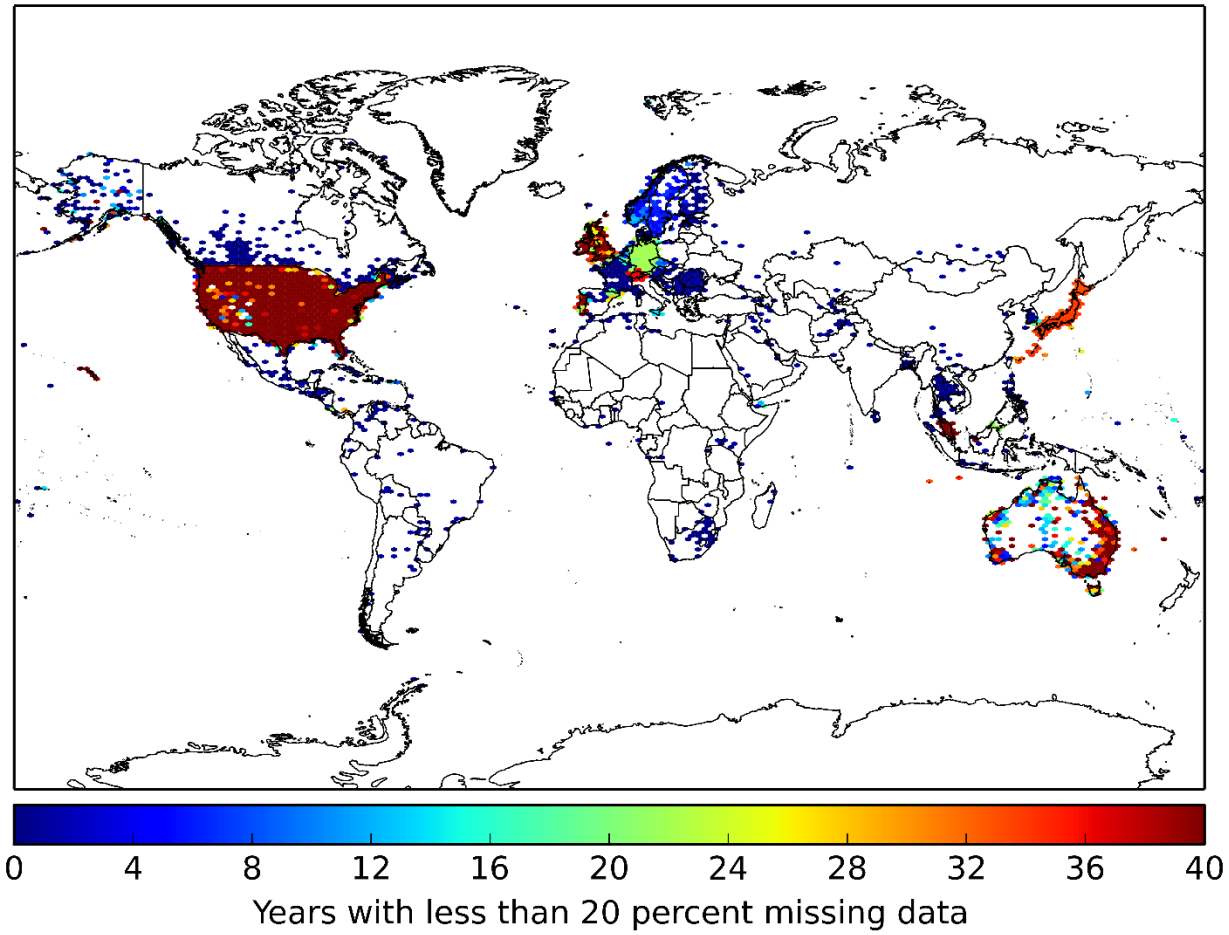
Stations come online at different times…



The data is of varying completeness…



Some analysis requires a year to have less than 20% missing data…



We want to create useful products

What data products would be useful to you?

We have developed quality control algorithms…

Newcastle University
UK | Malaysia | Singapore

Site specific tests

- rain gauge metadata,
- implausible large values (1h & 24h records)
- Monthly maximum 1-day precipitation
- long dry periods due to gauge malfunction
- accumulated totals (often at 9am)
- repeated values
- Change in resolution
- Duplicate records

Nearby gauge comparisons

- Statistical test of consistency with nearby gauges but problematical for extremes in summer/autumn therefore only partially applied

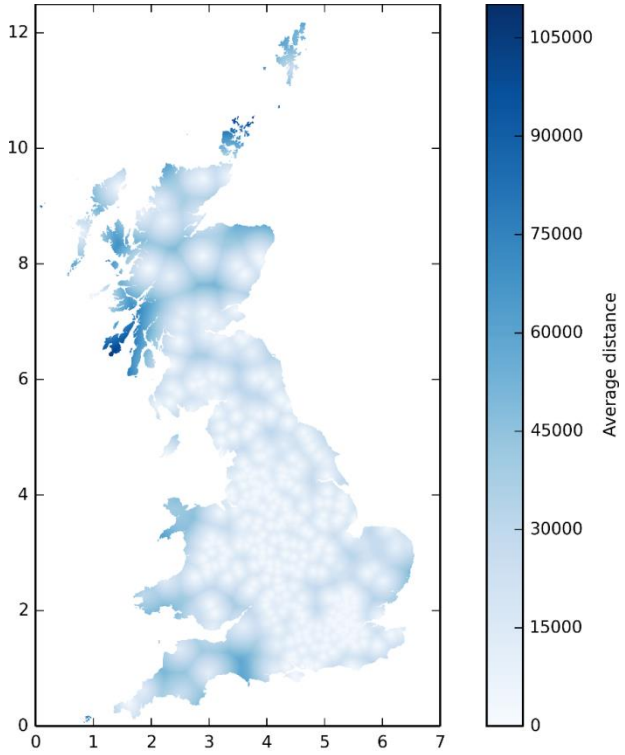
Multiple QC flags applied to each hour for each test

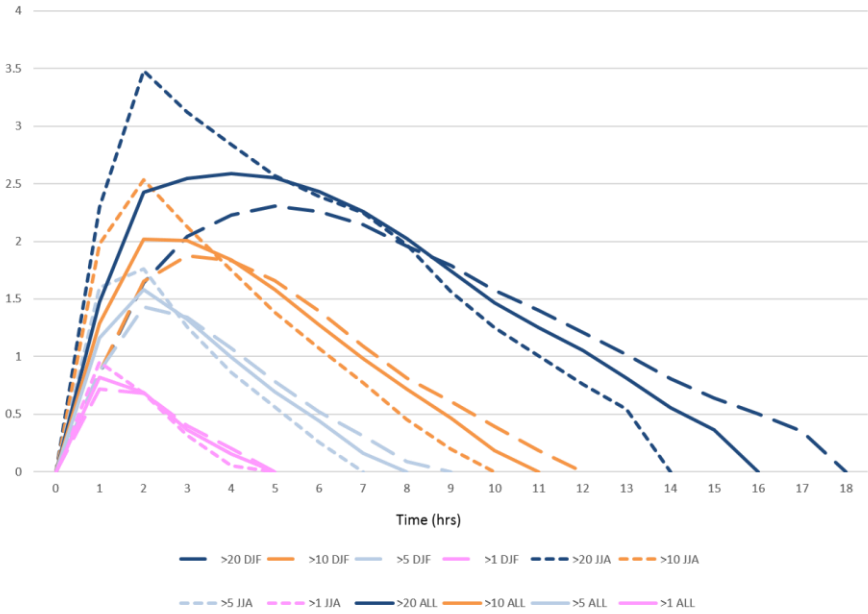
Automated rule base to define exclusions

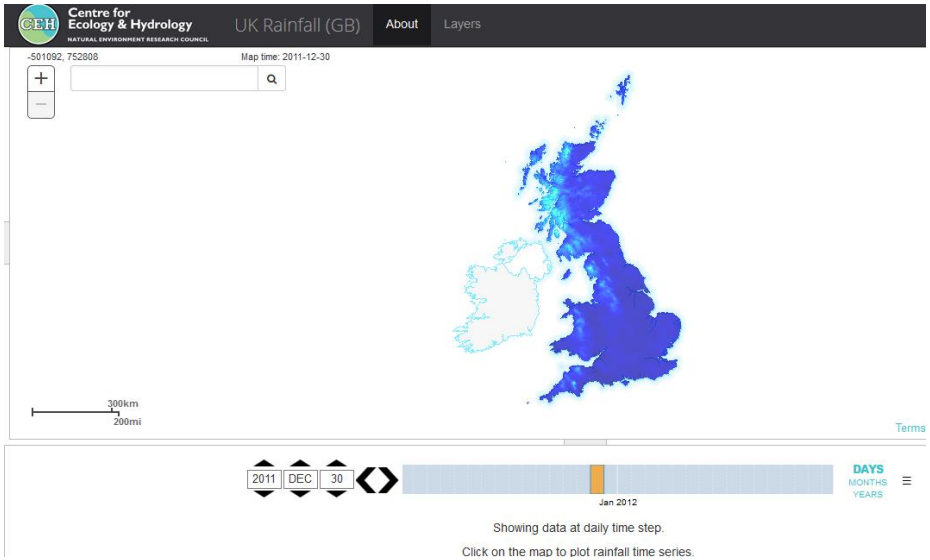
For example:

- all implausible hourly totals
- "large" hourly totals if in winter at 9am after ≥23 dry hours
- "large" hourly totals if after gauge non-operation (long dry spell)

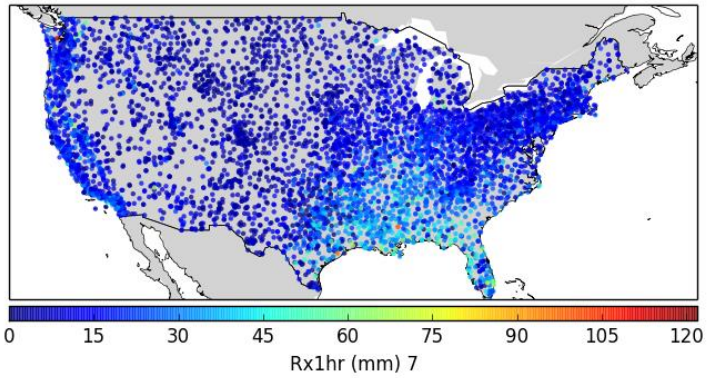
A gridded hourly dataset for the UK…

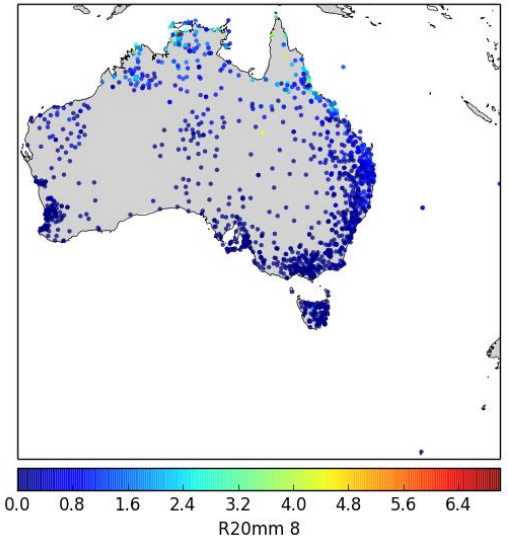


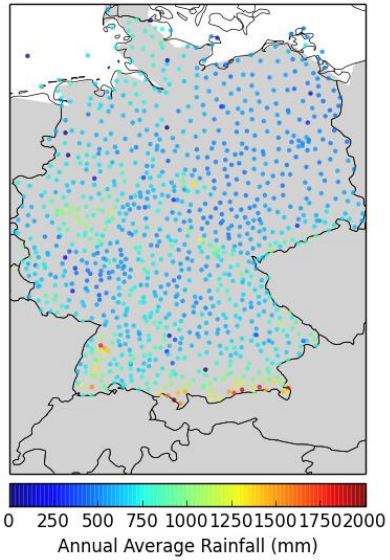




We want to use this data to calculate useful indices…







The indices we have decided to calculate are…

Monthly maximum indices

- Monthly maximum 1-hour precipitation
- Monthly maximum 3-hour precipitation
- Monthly maximum 6-hour precipitation
- Percent of daily total that fell in the Monthly maximum 1-hour precipitation

Diurnal cycle indices

- Monthly likely wettest hour within a day
- Monthly likely driest hour within a day
- Dispersion around Monthly likely wettest hour within a day
- Simple hourly precipitation intensity index
- Maximum length of wet spell, maximum number of consecutive hours with RR ≥ 1mm

Frequency/threshold indices

- Monthly count of hours when hourly PRCP ≥ 10mm
- Monthly count of hours when hourly PRCP ≥ 20mm
- Annual count of hours when hourly PRCP ≥ nnmm, nn is a user defined threshold

General indices

- Annual total precipitation in wet hours

Monthly maximum indices…

- **Rx1hr,** *Monthly maximum 1-hour precipitation* :
  Let RR_{ij} be the hourly precipitation amount on hour *i* in period *j*. The maximum 1-hour value for period *j* is:
  $Rx1hr_j = \max (RR_{ij})$

- **Rx3hr,** *Monthly maximum 3-hour precipitation* :
  Let RR_{ij} be the 3 hour precipitation amount on 3 hour period *i* in period *j*. The maximum 3-hour value for period *j* is:
  $Rx3hr_j = \max (RR_{ij})$

- **Rx6hr,** *Monthly maximum 6-hour precipitation* :
  Let RR_{ij} be the 6 hour precipitation amount on 6 hour period *i* in period *j*. The maximum 6-hour value for period *j* is:
  $Rx6hr_j = \max (RR_{ij})$

Monthly maximum indices…

- **Rx1hrP,** *Percent of daily total that fell in the Monthly maximum 1-hour precipitation*:

  Let RR_{ij} be the hourly precipitation amount on hour *i* in period *j*. Let *RD* be the total rainfall for the day containing $RR_{ij.}$:

  $$Rx1hr_jP = \frac{Rx1hrj}{RD}$$

Diurnal cycle indices…

- **LW1H,** *Monthly likely wettest hour within a day*:
  Let RR_{ij} be the hourly precipitation amount on hour *i* in period *j*. Let $MWHI_h$ be the mean monthly wet hour intensity for hour *h where* $RR_{ij} \geq 0.2$mm

  $LWH = h$ *where max* $(MWHI_h)$

- **LD1H,** *Monthly likely driest hour within a day* :
  Let RR_{ij} be the hourly precipitation amount on hour *i* in period *j*. Let $MWHI_h$ be the mean monthly wet hour intensity for hour *h where* $RR_{ij} \geq 0.2$mm

  $LDH = h$ *where min* $(MWHI_h)$

- **DLW1H,** *Dispersion around Monthly likely wettest hour within a day*:
  Let *DT* be the dispersion threshold. Let $MWHI_h$ be the mean monthly wet hour intensity for hour *h where* $RR_{ij} \geq 0.2$mm

  $DT = \frac{max(MWHIh) - min(MWHIh)}{2}$

  Count the number of hours where $MWHI_h > DT$

Diurnal cycle indices…

- **S1HII,** *Simple hourly precipitation intensity index* :

  Let RR_{wj} be the hourly precipitation amount on wet hours, *w (RR ≥ 0.2mm)* in period *j*. If *W* represents number of wet hours in *j*, then:

$$SHII_j = \frac{\sum_{w=1}^{W} RR_{wj}}{W}$$

- **CW1H,** *Maximum length of wet spell, maximum number of consecutive hours with RR ≥ 0.2mm*:

  Let RR_{ij} be the hourly precipitation amount on hour *i* in period *j*. Count the largest number of consecutive hours where $RR_{ij} \geq 0.2mm$

Frequency over threshold indices…

- **R10mm1hr,** *Monthly count of hours when hourly PRCP ≥ 10mm*:

  Let RR_{ij} be the hourly precipitation amount on hour *i* in period *j*. Count the number of hours where:

  $RR_{ij} \geq 10mm$

- **R20mm1hr,** *Monthly count of hours when hourly PRCP ≥ 20mm*:

  Let RR_{ij} be the hourly precipitation amount on hour *i* in period *j*. Count the number of hours where:

  $RR_{ij} \geq 20mm$

- **Rnnmm1hr,** *Annual count of hours when hourly PRCP ≥ nnmm, nn is a user defined threshold*:

  Let RR_{ij} be the hourly precipitation amount on hour *i* in period *j* . Count the number of hours where:

  $RR_{ij} \geq nnmm$

General indices…

- **PRCPTOT1hr,** *Annual total precipitation in wet hours* :

  Let RR_{ij} be the hourly precipitation amount on hour *i* in period *j*. If *i* represents the number of hours in *j*, then:

$$PRCPTOT_j = \sum_{i=1}^{I} RR_{ij}$$

Newcastle University
UK | Malaysia | Singapore

What indices would be useful to you?

A gridded hourly rainfall dataset for the UK

Elizabeth Lewis[1], Stephen Blenkinsop[1], Niall Quinn[2],
Jim Freer[2], Gemma Coxon[2], Ross Woods[2], Paul Bates[2], Hayley Fowler[1]

A gridded hourly dataset is a useful resource…

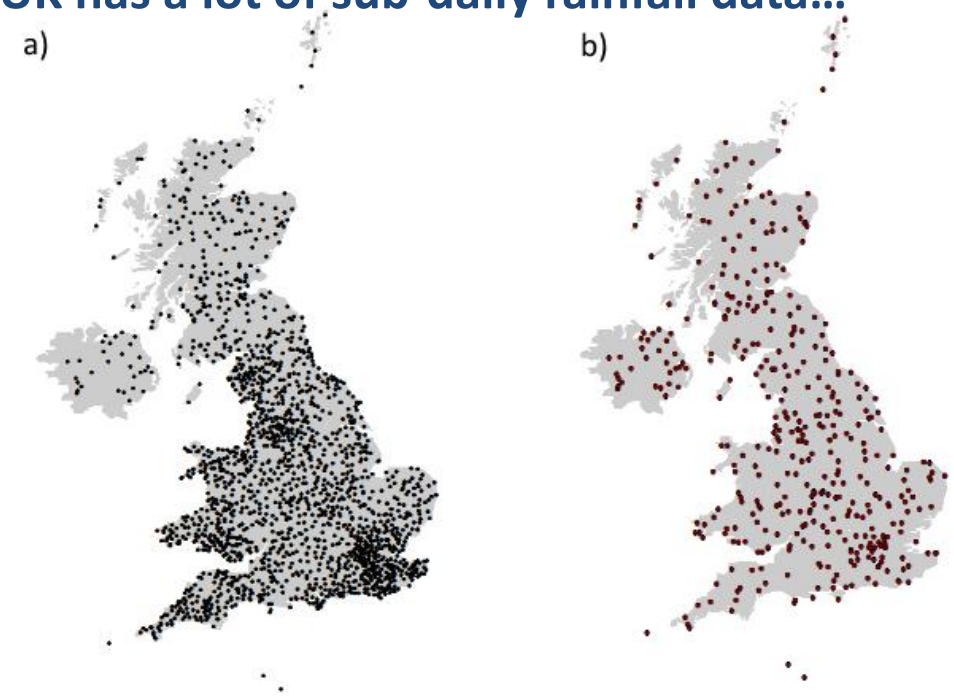
Newcastle University
UK | Malaysia | Singapore

- More detailed input to hydrological models
- Flash flooding
- Hourly climatology
- Validate/inform high resolution climate models

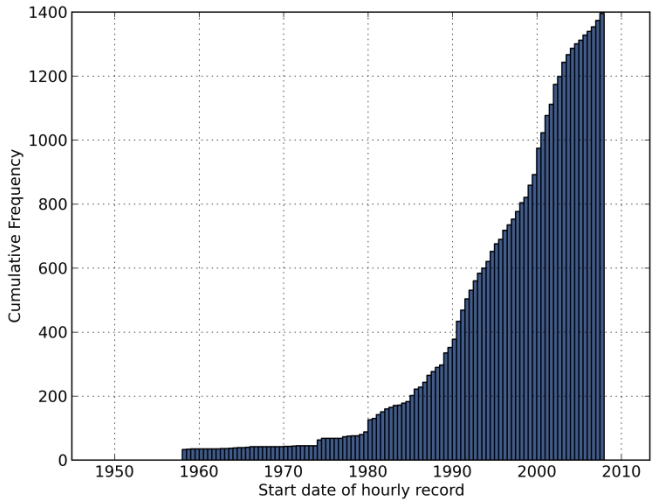
The UK has a lot of sub-daily rainfall data…

Newcastle University
UK | Malaysia | Singapore

All

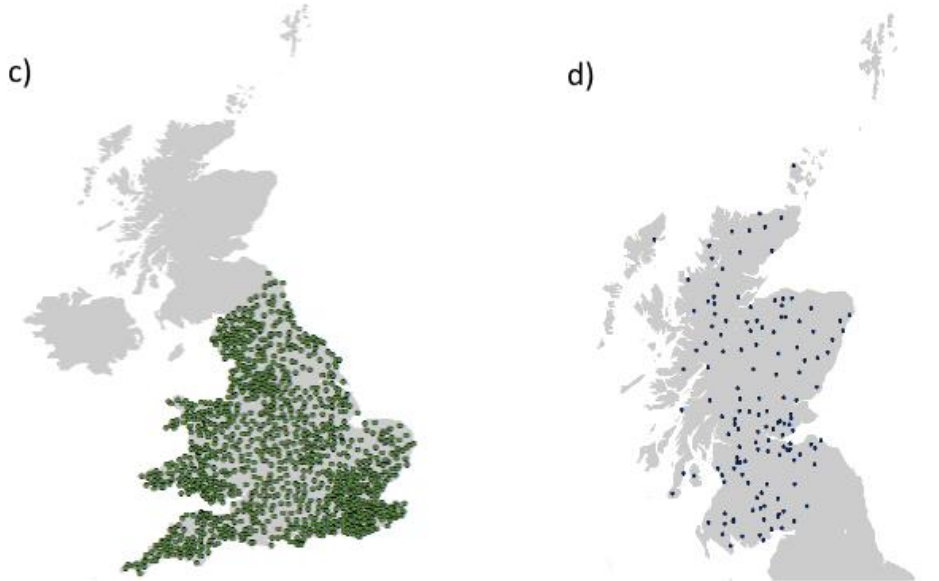


MIDAS



1990-2014

EA + NRW

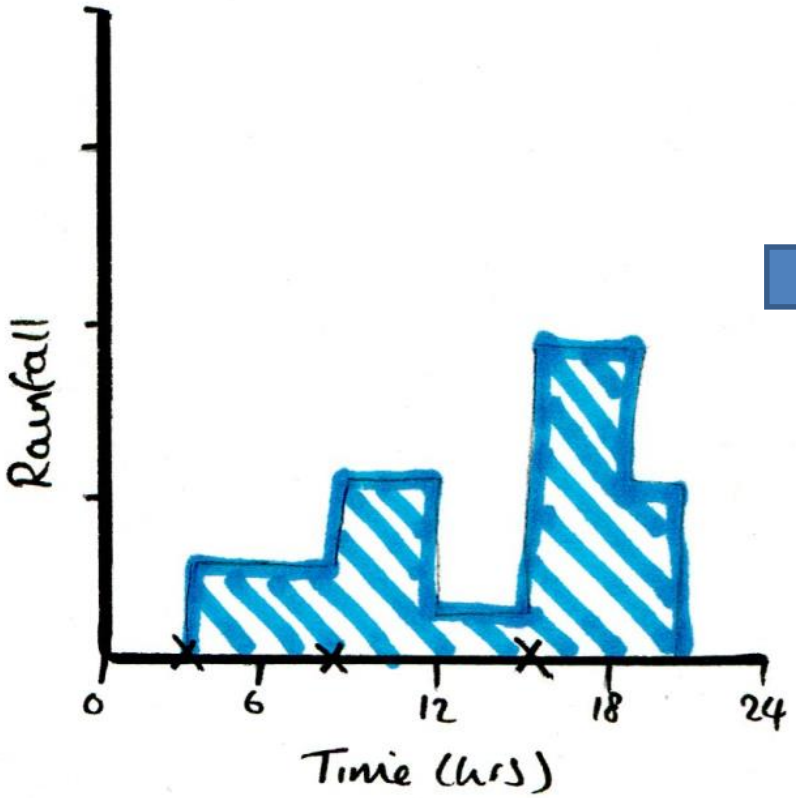


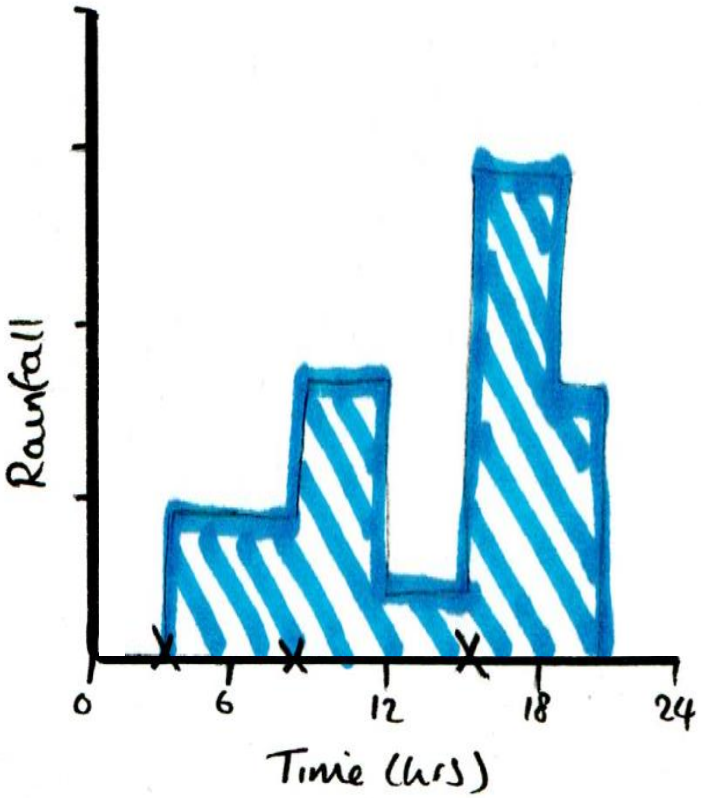
SEPA

And a new gridded 1km daily rainfall dataset…

CEH-GEAR: Centre for Ecology and Hydrology Gridded estimates of daily and monthly areal rainfall for the United Kingdom (1890-2014)

We wanted to make a gridded hourly dataset consistent with CEH-GEAR…





Use the shape of the hourly record

Preserve the daily total rainfall

The first step was to quality control the hourly data…

Acronym	Description
QC1	1h record exceeded by ≥ 20%, 33%, or 50%
QC2	80% of 1h record exceeded in April-October period
QC3	24h record exceeded by ≥ 20%, 33%, or 50%
QC4	Suspect daily accumulations at 0900 or 1200 flagged where a recorded rainfall amount at these times is preceded by 23 hours with no rain. Thresholds of (2x) the mean wet hour amount for the corresponding month are applied to maximise the chance of identify accumulated values.
QC5	Suspect consecutive daily accumulations at 0900 or 1200 flagged recorded rainfall amounts at these times are preceded by 23 hours with no rain on consecutive days with no threshold to the wet hour amount applied.
QC6	Suspect monthly accumulations. Identified where only one hourly value is reported over a period of a month and that value exceeds the mean wet hour amount for the corresponding month.
QC7	Total consecutive 'large' values. Large values are defined as those exceeding 2x the mean wet hour amount for the corresponding month. This threshold was applied as genuine duplicate small to moderate rainfall amounts occur legitimately.
QC8	Frequent tipping (only applied to EA TBR data using Upton & Rahimi (2003) algorithm)
QC9	'Terminal' dry spell at start/end of gauge record ≥ 31d duration
QC10	Dry period ≥ 31d duration, 45d duration
QC11	Neighbourhood check – dry period > 14 days. Summer, winter 90th , 95th, 99th bounds
QC12	Neighbourhood check – daily accumulation. Summer, winter 90th, 95th, 99th bounds

Once suspect data are flagged, we needed to decide which values to remove…

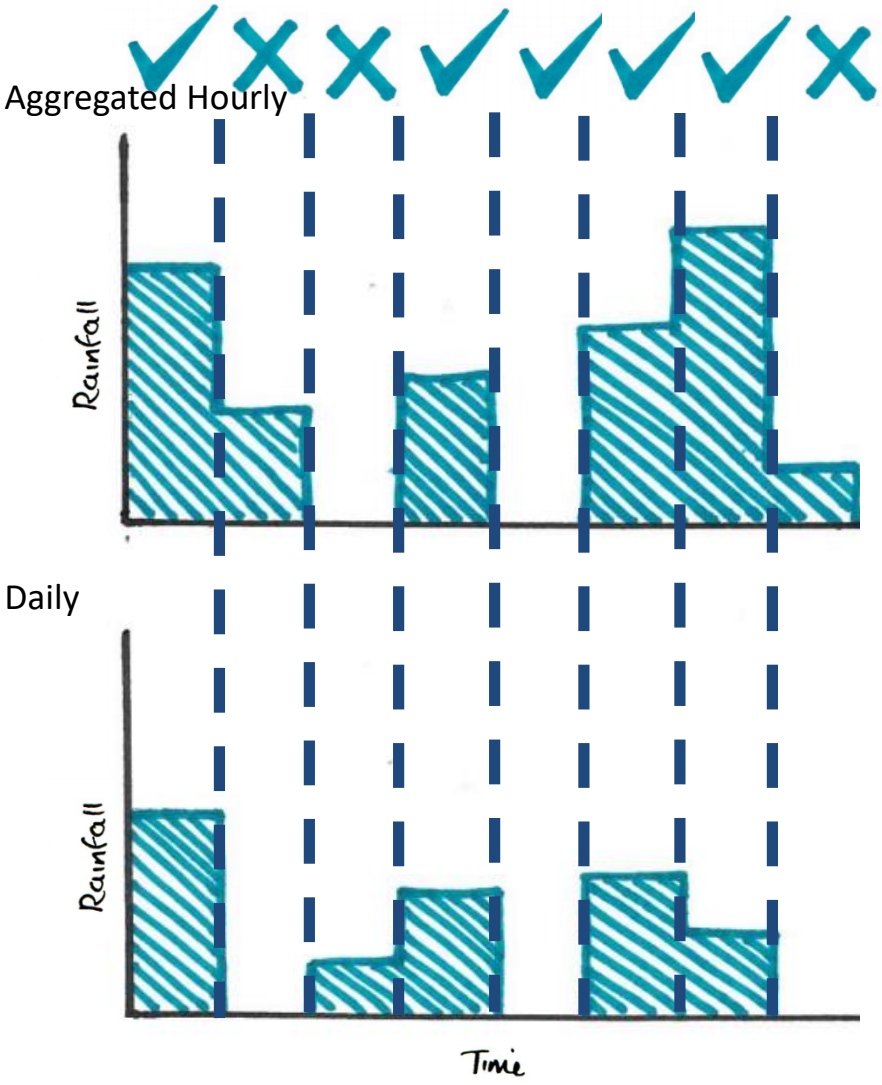
Rule Base 1 (Least severe)

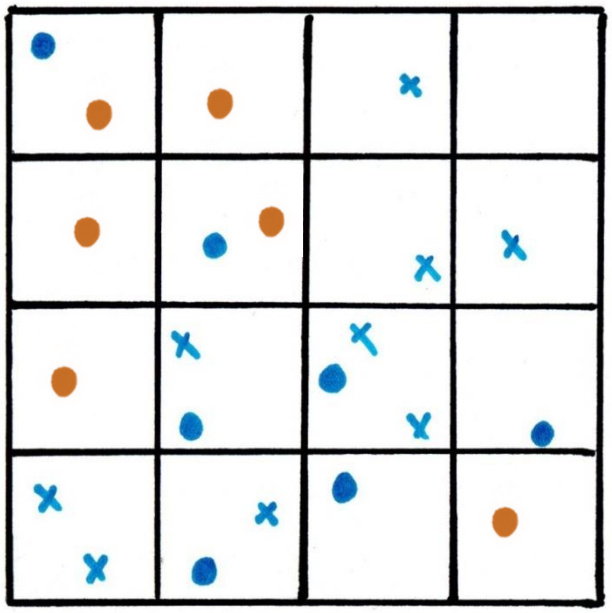
Rule base 2

Rule base 3 (most severe)

And test which rule base best captures errors…

Validate the rule base against UKCP09 5km gridded daily rainfall dataset



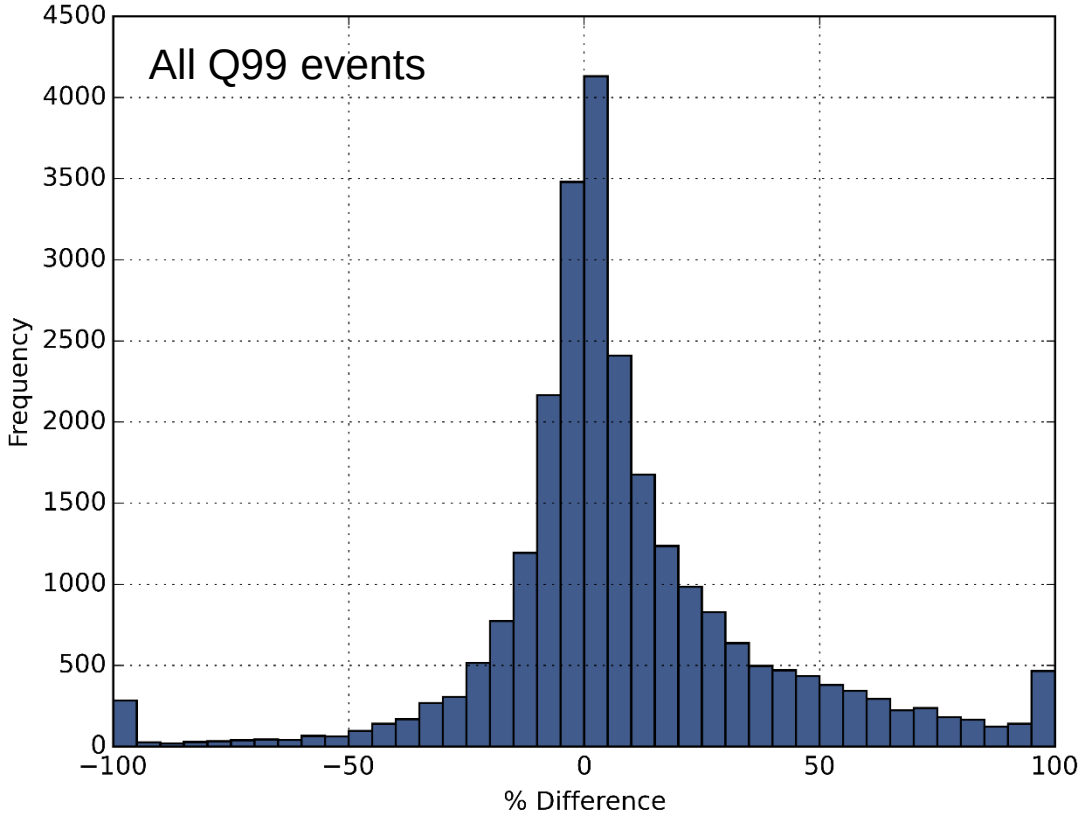


- Hourly gauge with good coverage and correlation
- Hourly gauge with average coverage and correlation
- Hourly gauge with poor coverage and correlation

The most severe rule base correctly excluded the most erroneous high values…

High Values (Q99)

- % difference between hourly (24hr total around Q99) and daily rainfall
- “correct” refers to % diff < +/- 65%

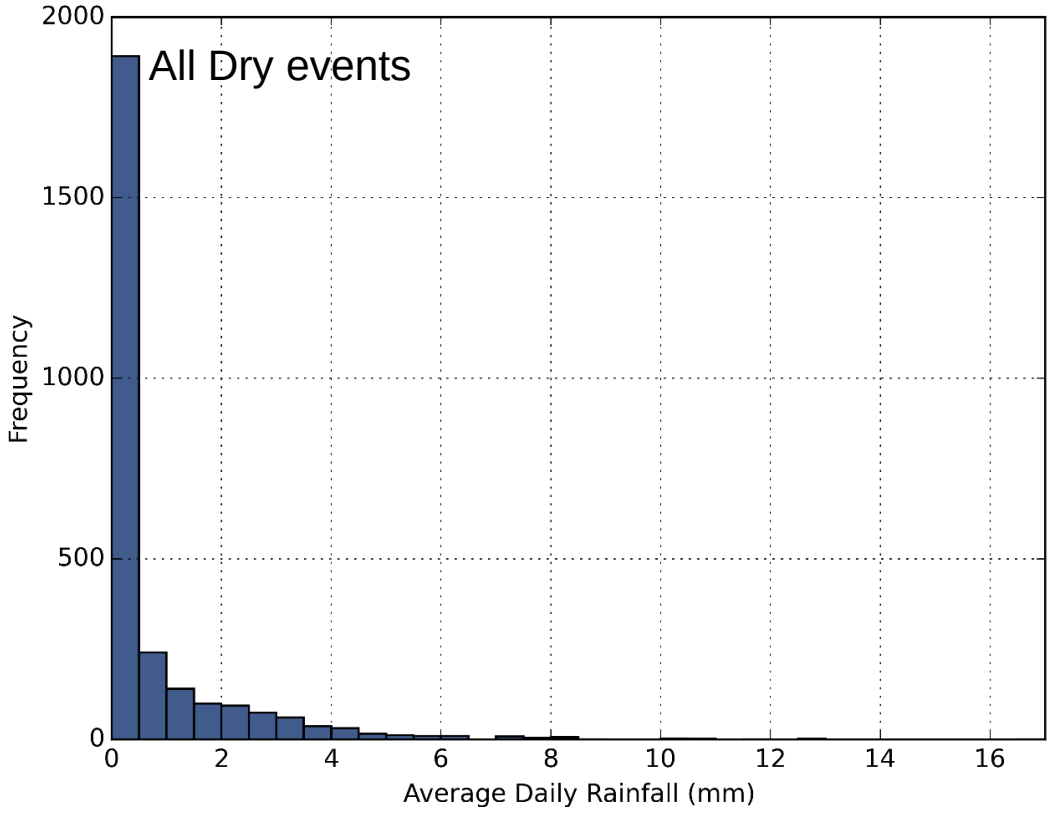


	rb1	rb2	rb3
% incorrectly flagged- ALL	0.63	1.19	1.61
% incorrectly flagged- JJA	0.69	1.14	1.68
% incorrectly flagged- DJF	0.5	1.34	1.63
% correctly flagged- ALL	12.84	14.53	16.72
% correctly flagged- JJA	9.05	10.44	12.33
% correctly flagged- DJF	23.46	26.75	30.45

And the most erroneous dry spells…

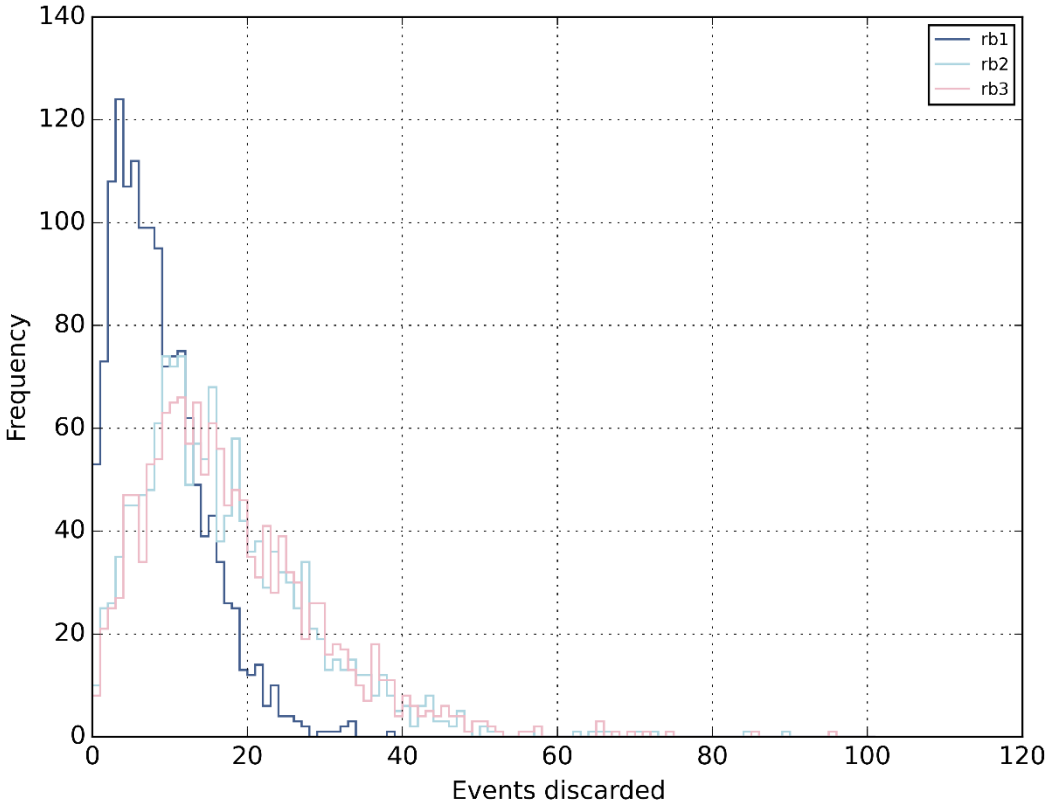
Dry Spells (>20 days)

- Average daily rainfall over the dry spell
- “correct” refers to average daily rainfall < 1mm



	rb1	rb2	rb3
% incorrectly flagged	9.47	9.8	9.99
% incorrectly flagged- JJA	10	10.11	10.11
% incorrectly flagged-DJF	19.53	21.09	21.88
% correctly flagged	66.99	78.59	79.9
% correctly flagged- JJA	72.6	73.97	76.03
% correctly flagged-DJF	62.42	84.24	84.85

The most severe rule base correctly excludes the most data…



	rb1	rb2	rb3
avg events discarded per gauge	8	31	32
min num events discarded per gauge	0	0	0
max num events discarded per gauge	38	7063	7073
total num events discarded per gauge	11679	44530	45966
Average % hours discarded per gauge	2.10%	2.60%	2.6%
total hours discarded	4102687	5030850	5071378

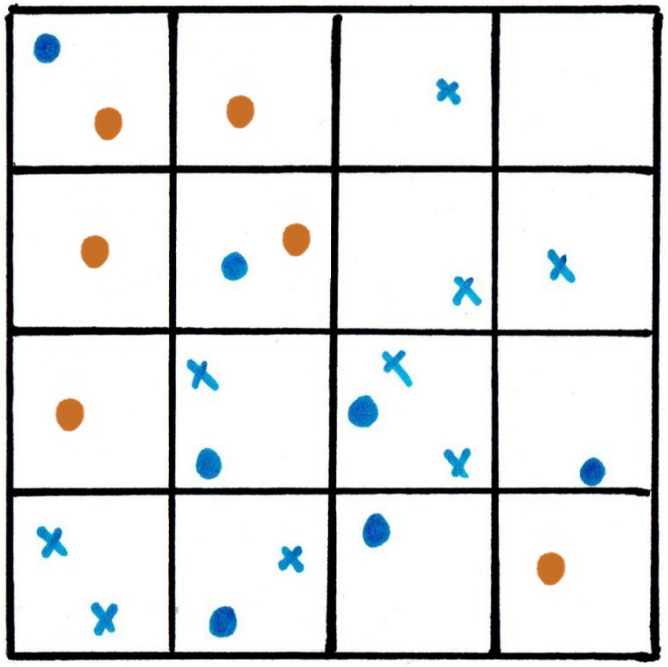
The rule bases were also checked against known extreme rainfall events…

Newcastle University
UK | Malaysia | Singapore

Location	Date	time (approx start)	reported rainfall amount (mm)	Reported duration (hr)	highest recorded hourly value	flaged	Hour of highest hourly value
Bristol	23/05/1992	18:00	25	1	22.7	N	20
Ashow (warks)	20/08/1996	20:30	24.5	0.5	28.0	N	21
Plumpton (East Sussex)	11/10/2000	19:00	143.8	24	8.2	N	23
Beaufort Park, Bracknell (Berks)	7/5/2000	18:30	180	1	49.4	N	20
Winterbourne (Birmingham)	26/09/1998	18:00	33.1	2	33.0	N	20
Boscastle	16/08/2004	11:00 ?	200	4	46.0	Y	14
Hazelrigg (Lancashire)	20/08/2004	3:00	23.4	1	26.6	N	5
Ottery st Mary (east Devon)	30/10/2008	03:00 ?	200	2	18.4	N	2
Hawnby (north york moors)	19/06/2005	17:00	59.8	1	59.8	N	17
Seathwaite (Cumbria)	20/11/2009	9:00	253	24	17.6	N	6
Wallington hall (Northumberland)	6/8/2011	18:45	34.6	6	17.2	N	15

Examples taken from ‘Weather’ articles

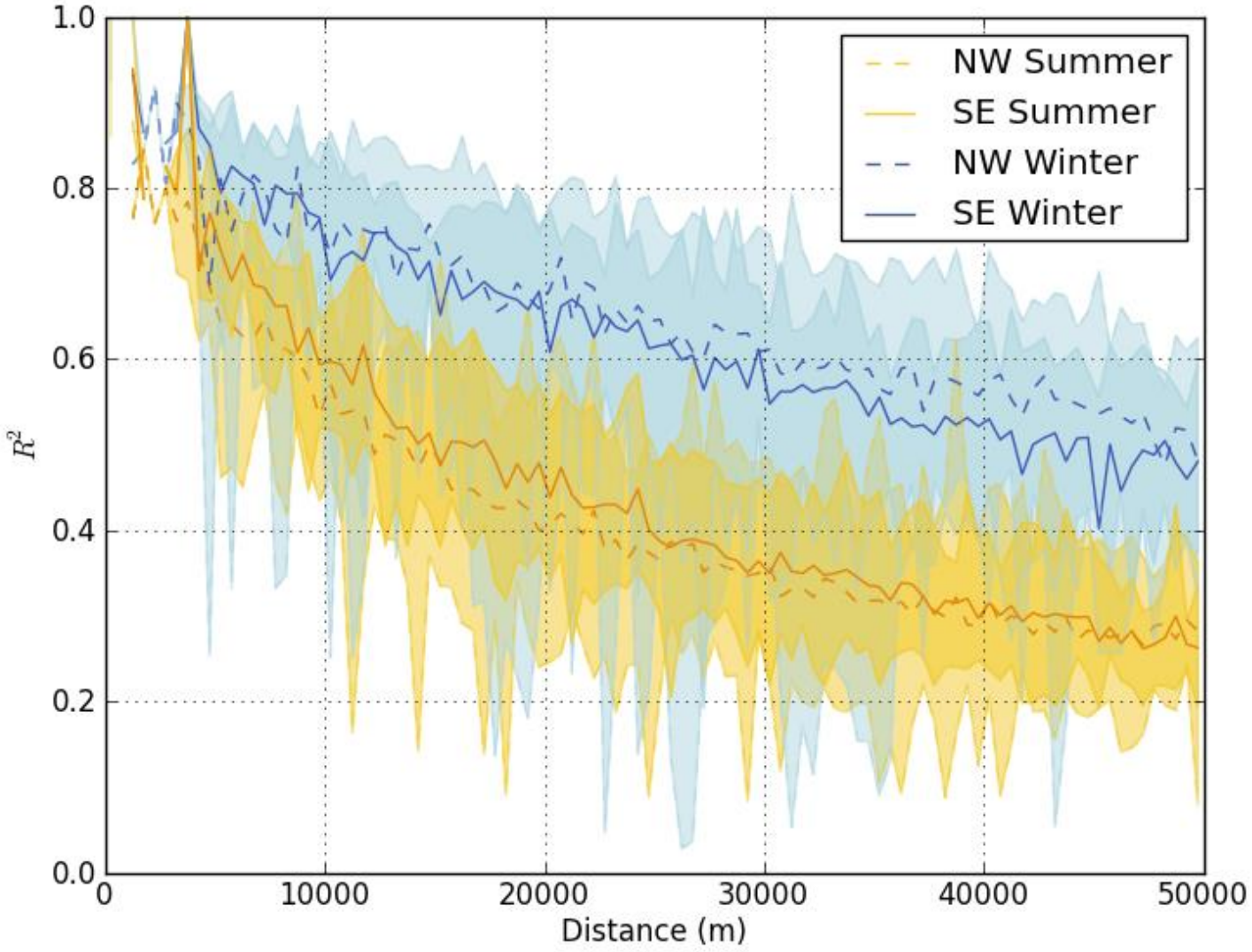
We can now use the quality controlled dataset to disaggregate the daily data…



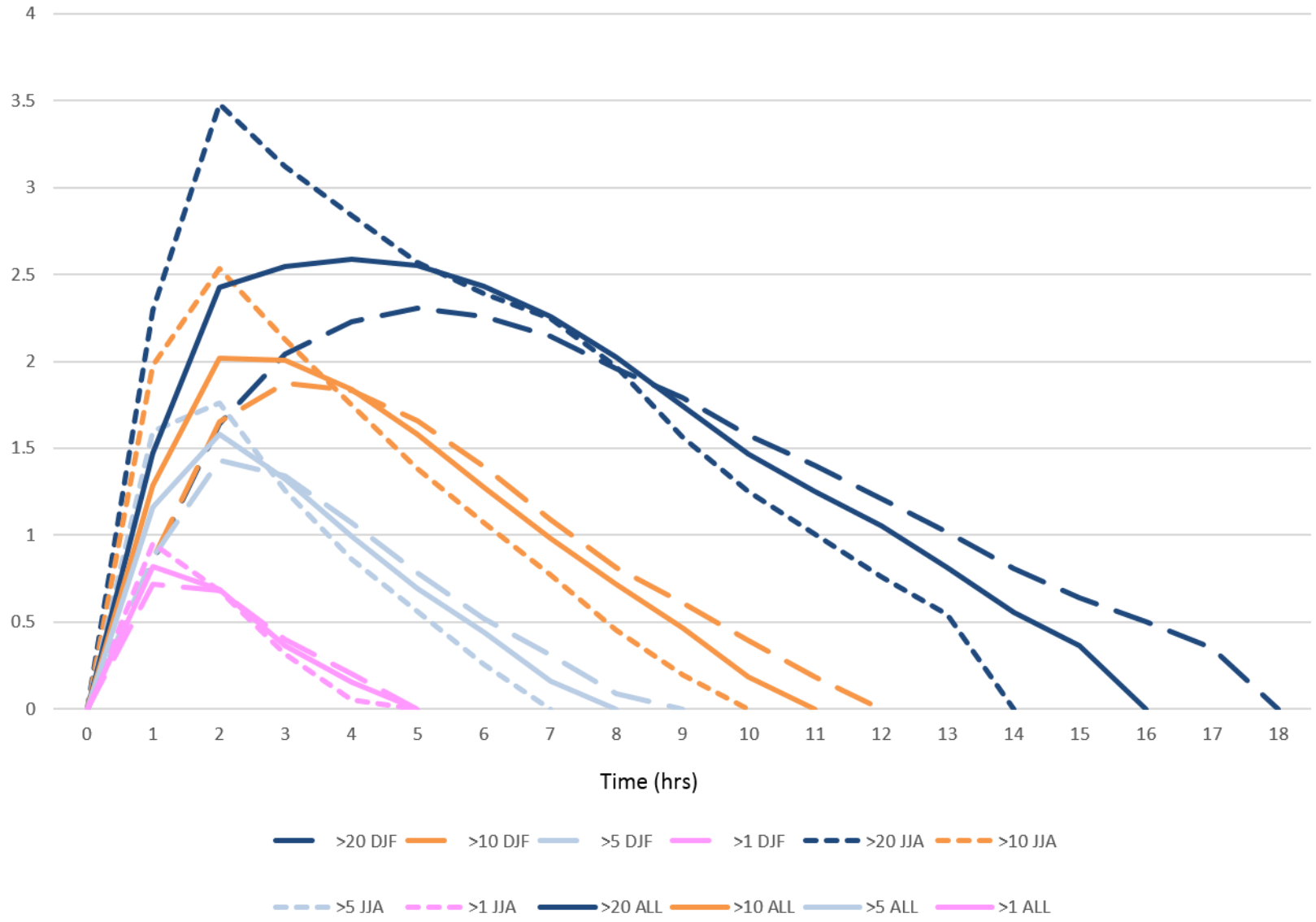
- Hourly gauge with good coverage and correlation
- Hourly gauge with average coverage and correlation
- Hourly gauge with poor coverage and correlation

The distance to search for a gauge is limited…

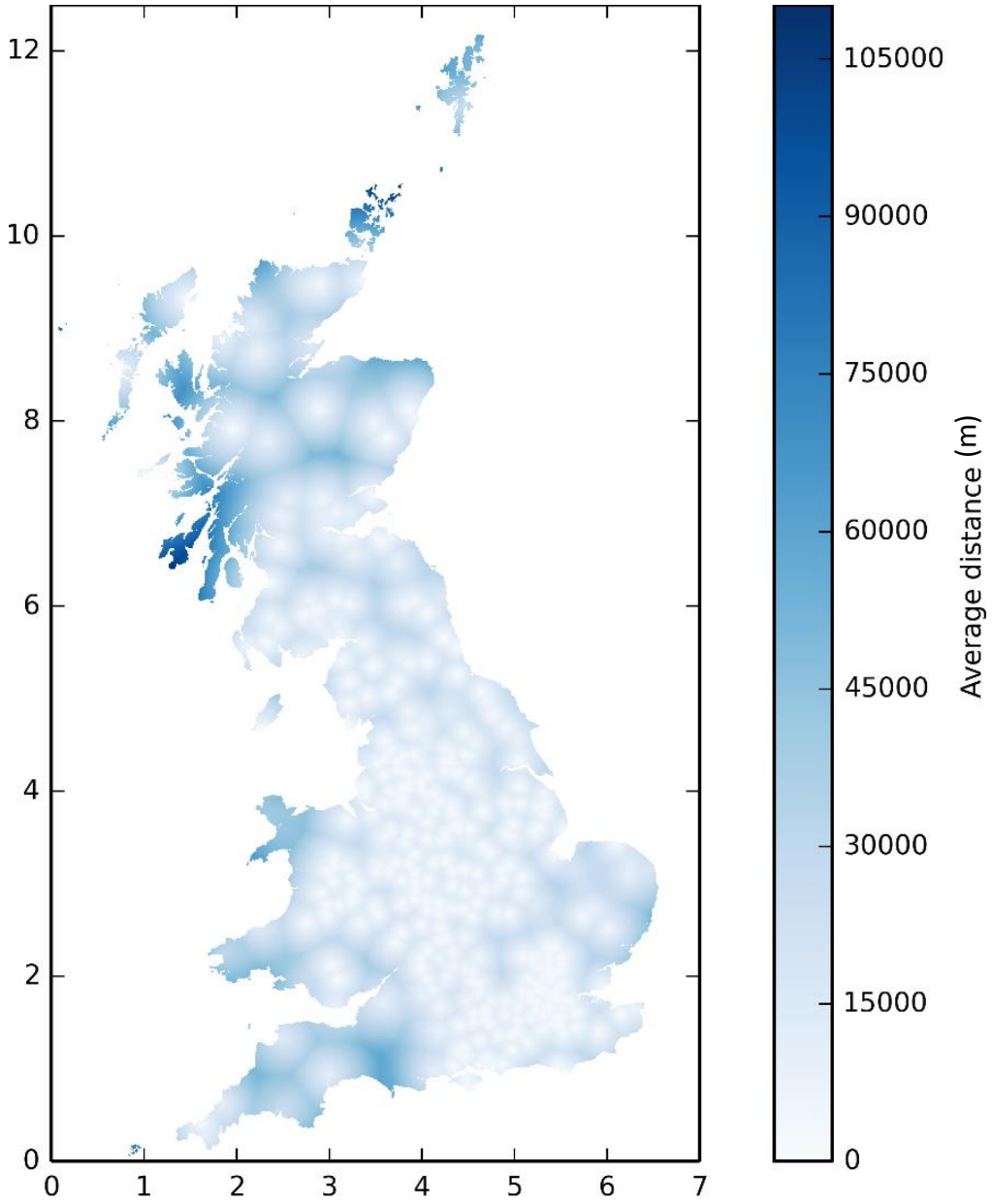
Newcastle University
Singapore



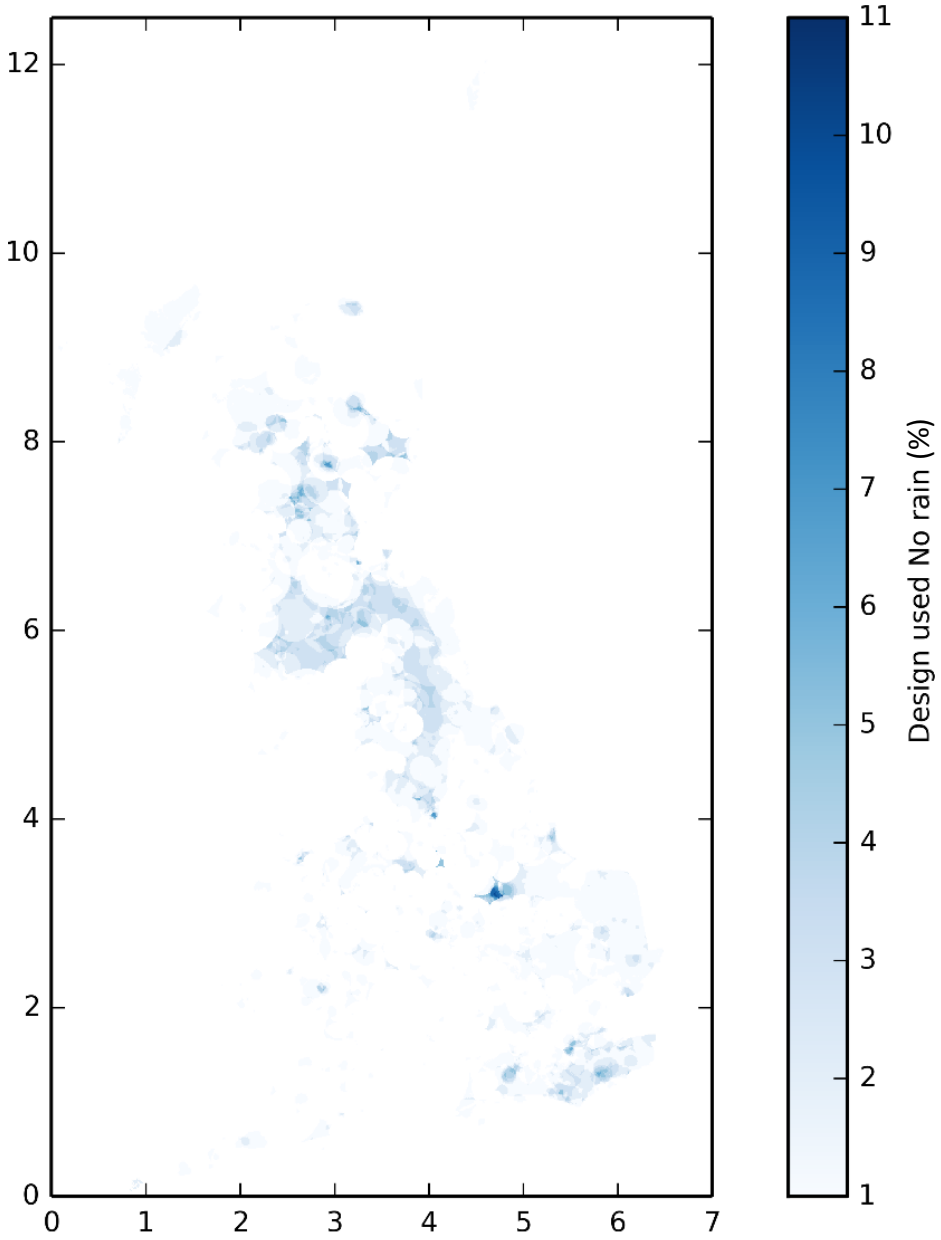
When there is rain in the daily record but not in the hourly record we use a design storm…



Average distances used are low…



And use of the design storm is typically low…



Conclusions

- We are collating a global sub-daily rainfall dataset
- And are developing a quality control algorithm
- We are producing indices to describe sub-daily rainfall characteristics
- For the UK, we have created a gridded hourly dataset which will be freely available and hosted by CEH
- We want to hear your feedback about what indices would be useful for your work.

Elizabeth Lewis (Elizabeth.lewis2@ncl.ac.uk)