# The link between catchment precipitation forecast skill and spread to that of downstream ensemble hydrological forecasts

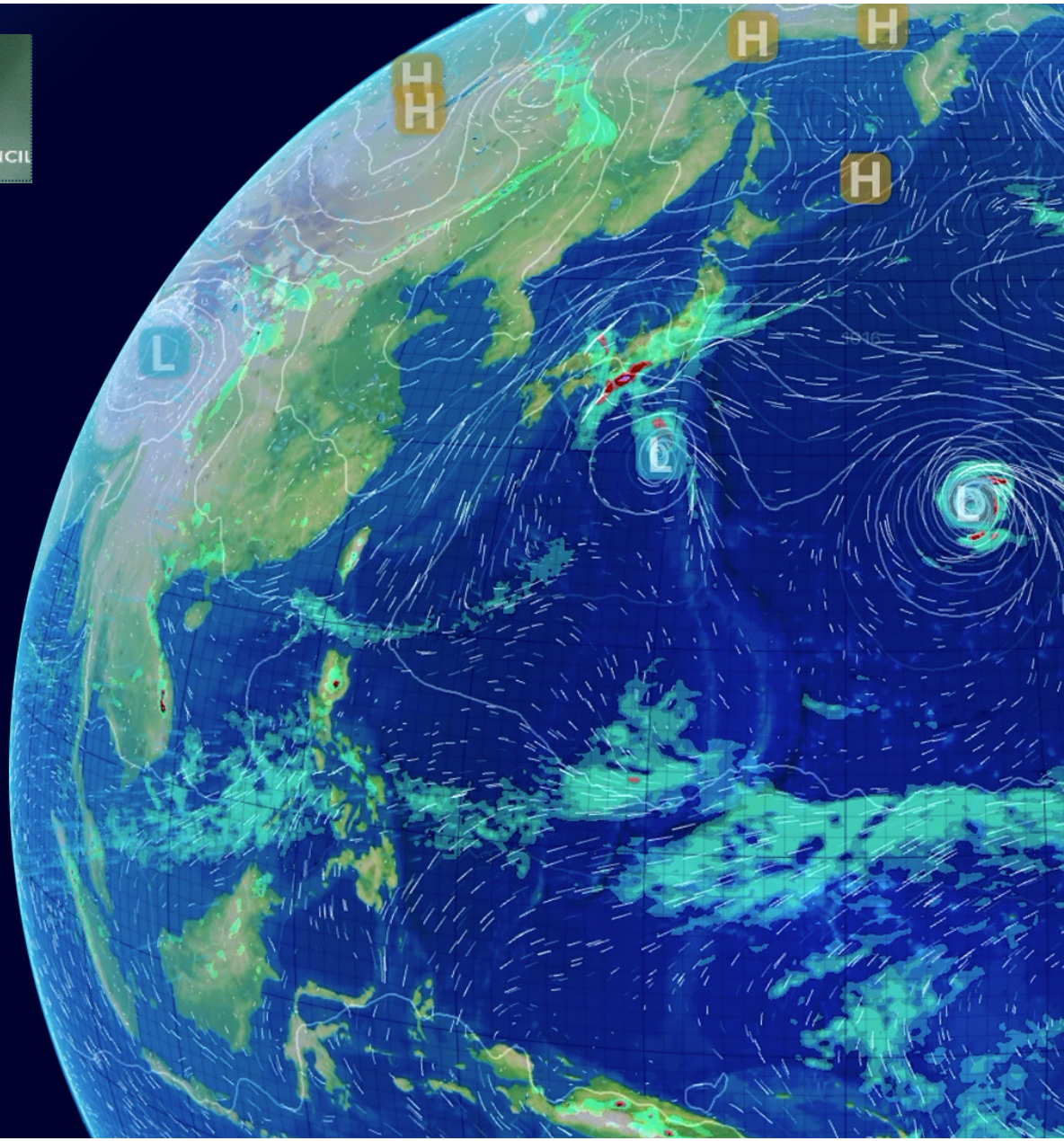
Gabriella Csima and **Marion Mittermaier**

*Met Office*

Seonaid Dey, Bob Moore and

Steve Cole

*Centre for Ecology and Hydrology*



# Outline

1. Project aims, methodology and data sets
2. Results
   - Sensitivity to observation type
   - National and regional variations in precipitation and river flow forecast attributes
   - Spatial variations in skill
   - Mapping precipitation skill onto river flow skill by catchment
3. Conclusions

# Project outline and data sets

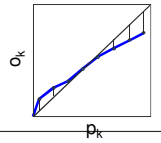
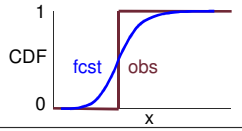
**Table 2.1 Metrics to be used for NWP and Hydrological forecast verification.**

| Common between NWP and Hydrology | NWP only |
|---|---|
| • Continuous Rank Probability Score (CRPS) with decomposition<br>• Brier Score (BS) with decomposition<br>• Continuous Rank Probability Skill Score (CRPSS)<br>• Brier Skill Score (BSS) with decomposition<br>• Reliability Diagram<br>• Relative Operating Characteristic Diagram and Area Under Curve Skill Score (ROCSS)<br>• Relative Economic Value (REV)<br>• Rank Histogram | • Mean Error (ME, a measure of bias) of areal mean precipitation per member<br>• Root Mean Squared Error (RMSE) of areal mean precipitation per member |

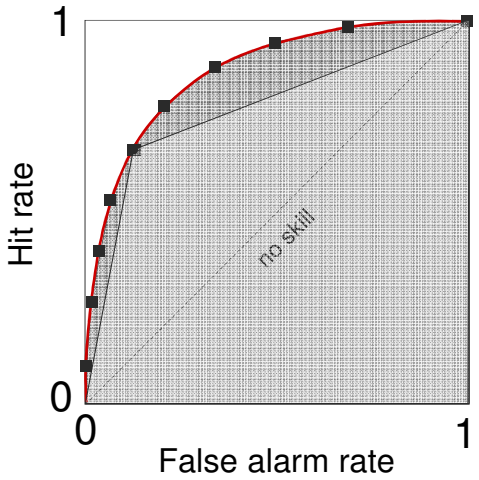
- Aim to develop a verification framework for catchment-scale ***ensemble*** precipitation which drives a river-flow ***ensemble***.
- Understand how ***uncertainty*** (affecting skill) is propagated from precipitation to river flow.
- Understand the ***impact of truth type*** and ***accumulation length*** on precipitation verification.

Met Office

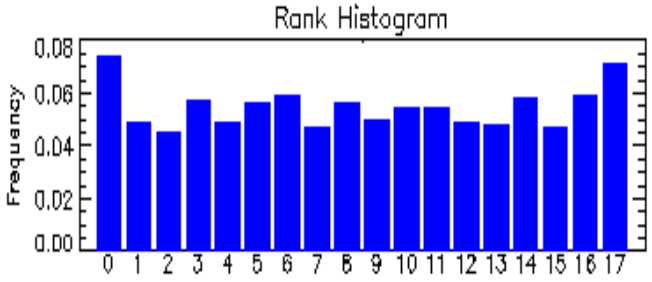
# Selected framework components

| **Deterministic** | **Probabilistic** (suggestions) | **Visual aid** |
|---|---|---|
| Mean bias | Reliability term of BS $\frac{1}{N}\sum_{k=1}^{K} n_k (p_k - \bar{o}_k)^2$ |  |
| RMS error | Brier score (square root) $\sqrt{BS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - o_i)^2}$ | |
| Mean absolute error | CRPS $\int (P_{fcst}(x) - P_{obs}(x))^2 dx$ |  |





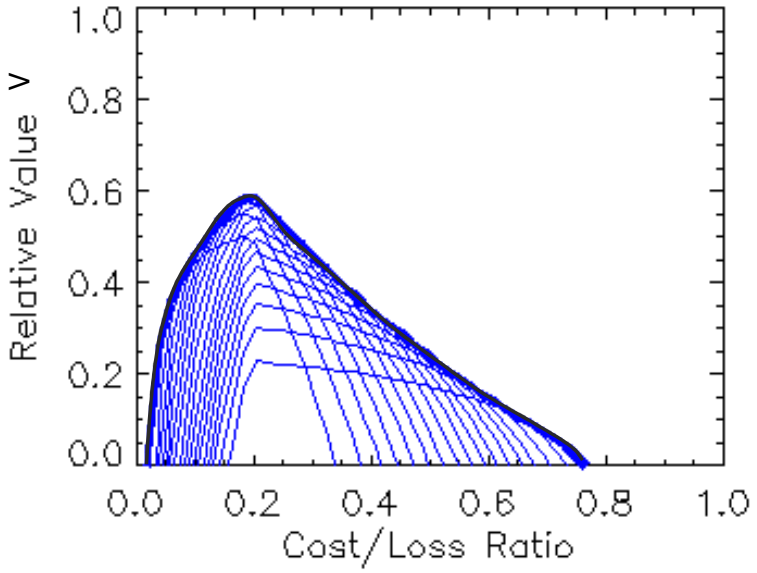

- ROC curve is independent of forecast bias – "potential skill"
- Area under curve ("ROC area") is a useful summary measure of forecast skill
- Perfect:1;No skill: 0.5




- Measures how well the ensemble spread of the forecast represents the true variability (uncertainty) of the observations
- Count where the verifying observation falls with respect to the ensemble forecast data, which is arranged in increasing order at each grid point.
- Perfect = flat; u-shape=under spread, many obs falling outside bounds of ensemble

# Relative value score

Measures the relative improvement in economic value as a function of the cost/loss ratio C/L for taking action based on a forecast as opposed to climatology




$$V = (1-F) - \left(\frac{1-C/L}{C/L}\right)\left(\frac{\bar{o}}{1-\bar{o}}\right)(1-H) \quad \text{if } C/L < \bar{o}$$

$$V = H - \left(\frac{C/L}{1-C/L}\right)\left(\frac{1-\bar{o}}{\bar{o}}\right)F \quad \text{if } C/L > \bar{o}$$

where $H$ is the hit rate and $F$ is the false alarm rate

- The relative value is a skill score of expected expense, with climatology as the reference forecast.
- Range: -∞ to 1. Perfect score: 1
- Plot V vs C/L for various probability thresholds. The envelope describes the potential value for the probabilistic forecasts.

# Data sets

Fig. 1 The UK weather radar network

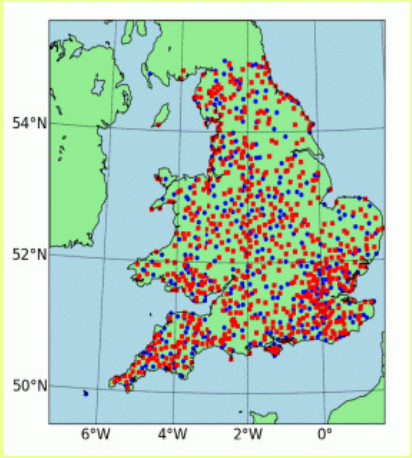
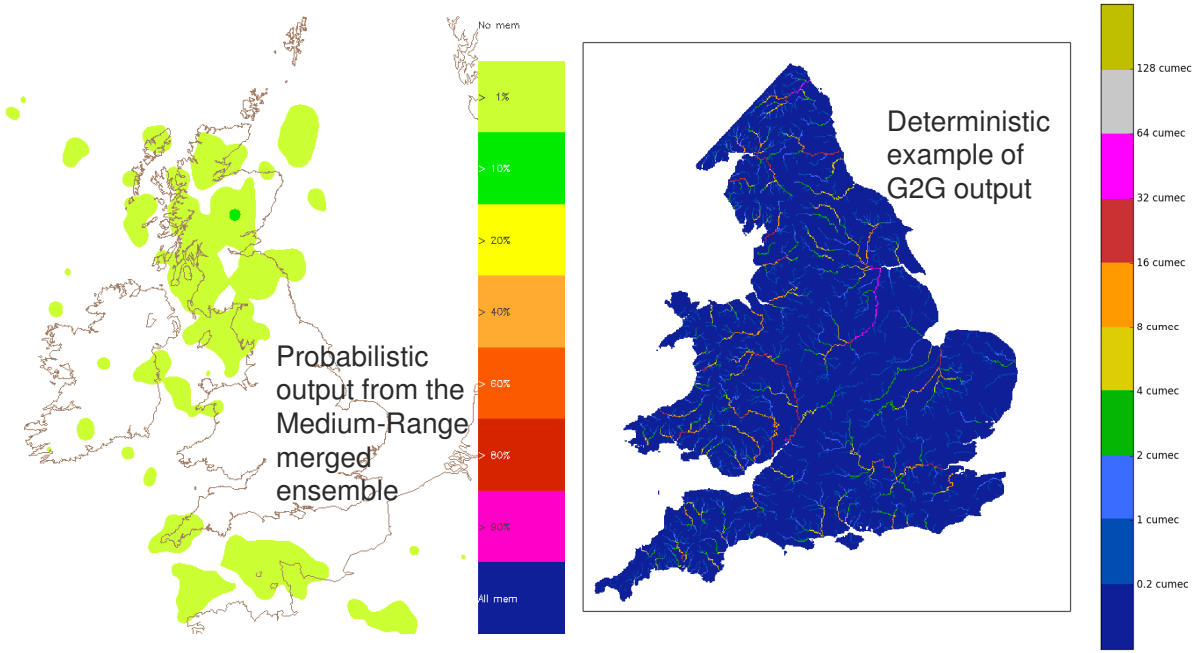



Fig. 2 UK real-time rain-gauge network.
● = Met Office ● = Environment Agency

- Radar-only (gauge-adjusted) analyses
- Gridded gauge-only analyses

Two periods considered: winter Nov-Dec 2015 & summer May-Jun 2016.

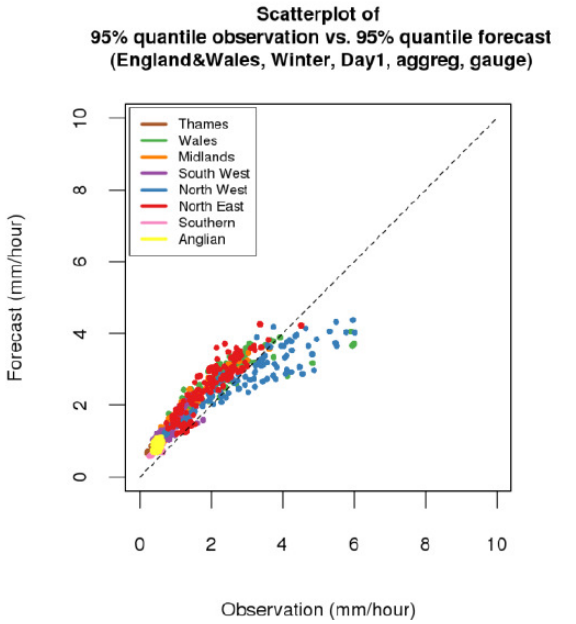


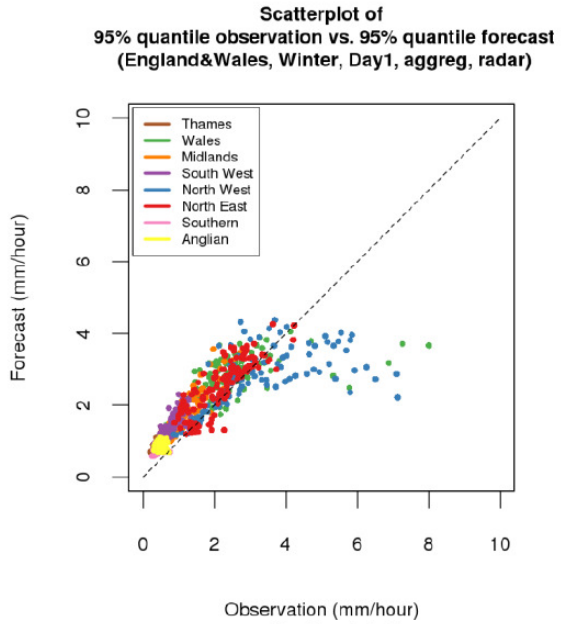

- Merged nowcast (STEPS – t+7h) and MOGREPS-UK (2.2 km – ~t+32h) and MOGREPS-G (32 km – ~t+144h) precipitation ensemble out to 6 days, on a 1 km grid. Output is at 15 min resolution.
- River-flow ensemble (G2G) output at 15 min intervals

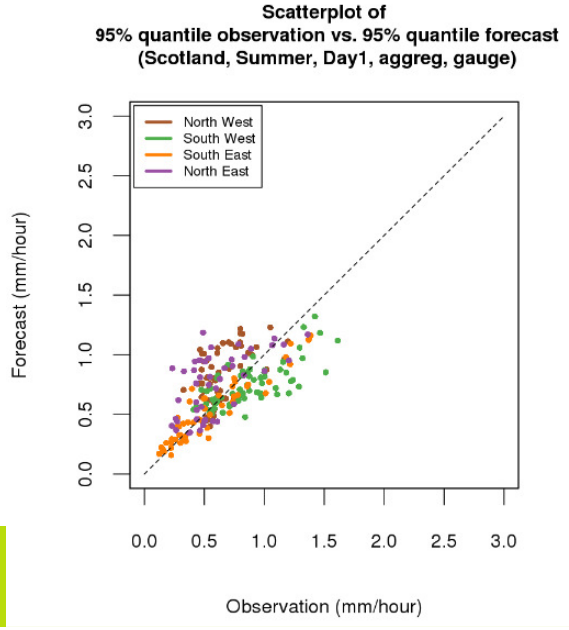
# Radar v gauge

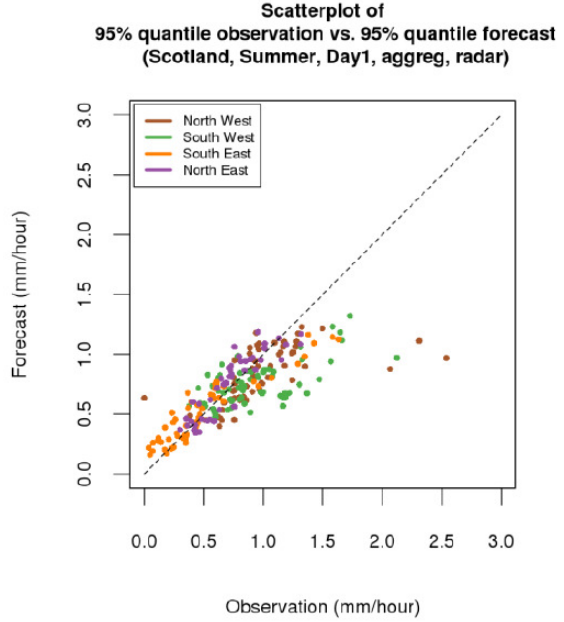
- Only showing results for Day 1.
- Precipitation was evaluated at the hourly and daily scale, aggregated over 4 runs per day

• Centiles in good general agreement.
• Radar shows more large totals, which may be due to the detailed spatial sampling radar can provide. [Interpolation can not “create” large totals between gauges, i.e. the largest totals may be missed if they don’t occur precisely over a gauge. More acute problem in convective conditions]
• Tendency for model to produce too much light rain but under-estimating larger totals.



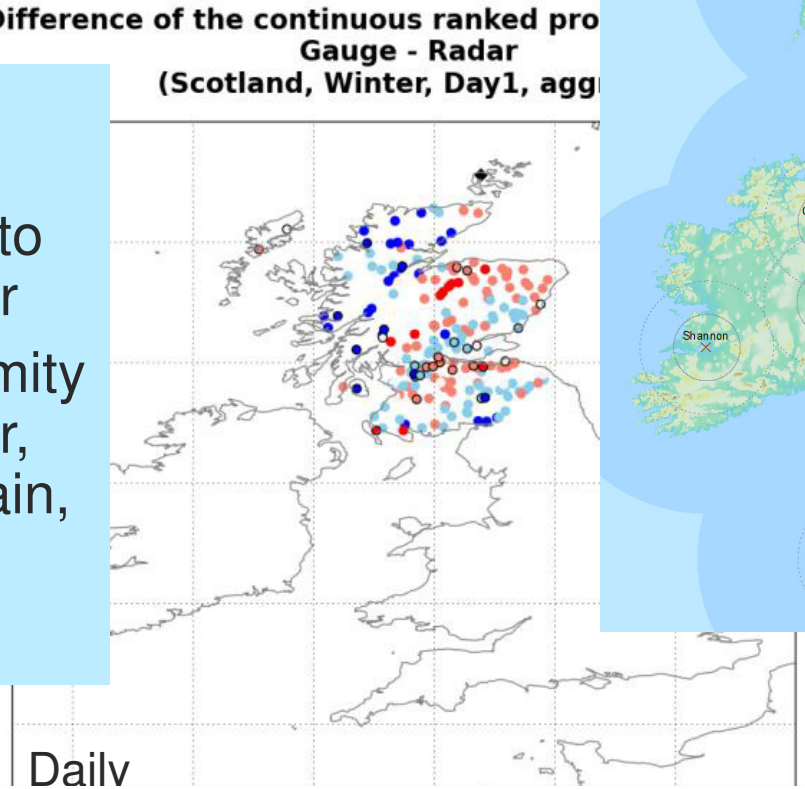



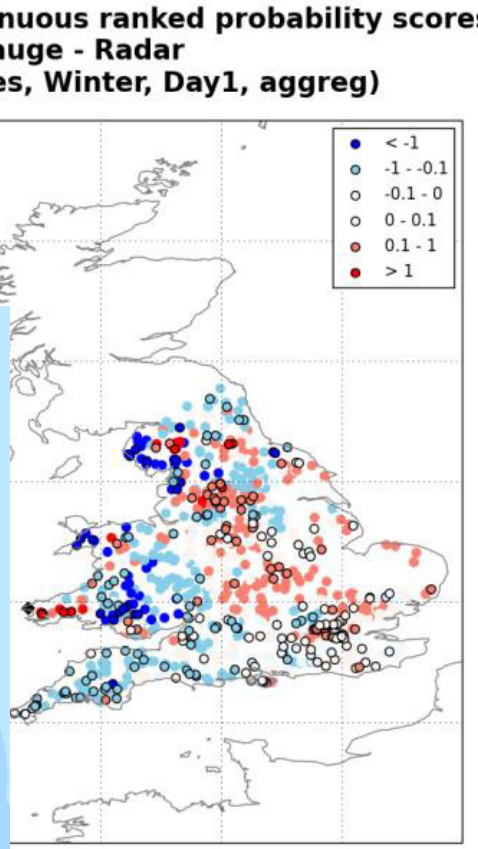



# Sensitivity of CRPS to observation type: gauge v radar

- No thresholds
- Impact of radar coverage is likely to be dominant factor
- Affected by proximity of location to radar, complexity of terrain, orographic rain correction



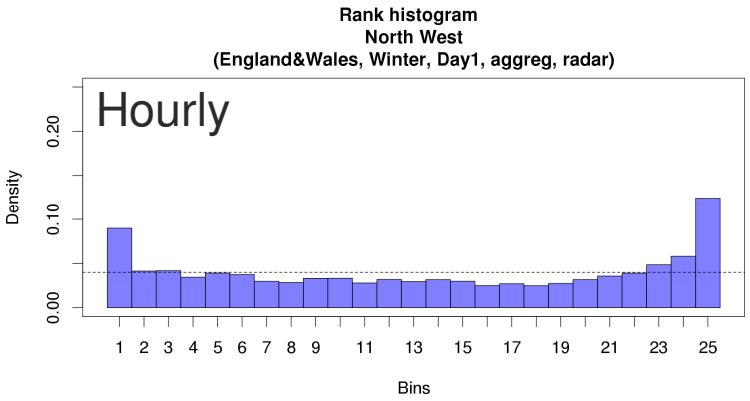


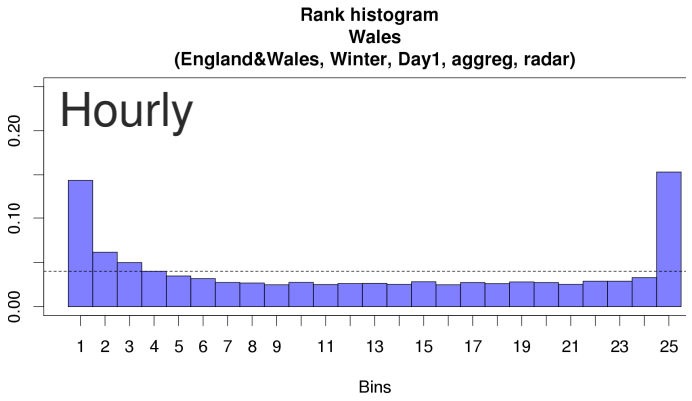


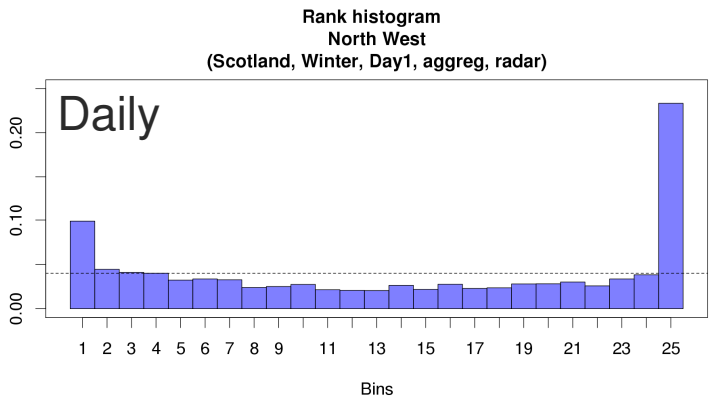

Radar has noticeably worse scores over complex terrain

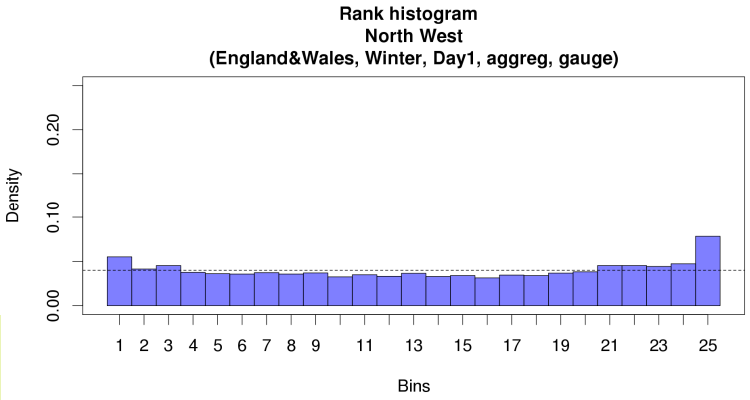
# Sensitivity of rank histogram to observation type: gauge v radar

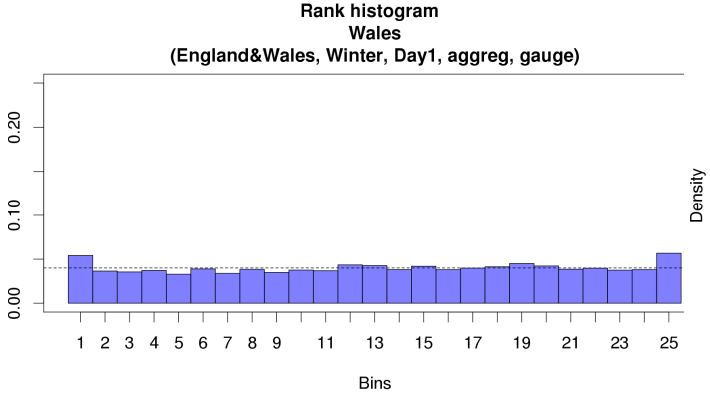
- Rank Histograms are found to be very sensitive to the underlying distribution, which affects interpretation of ensemble spread.
- Strong regional variations which are also be related to radar coverage and the spatial sampling capabilities of radar compared to the point measurements from the gauges.

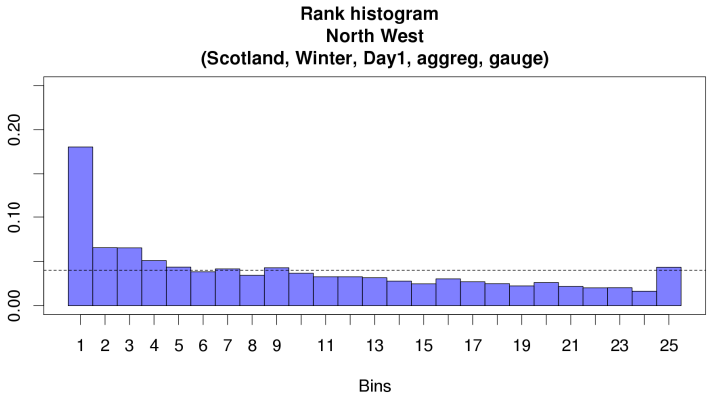


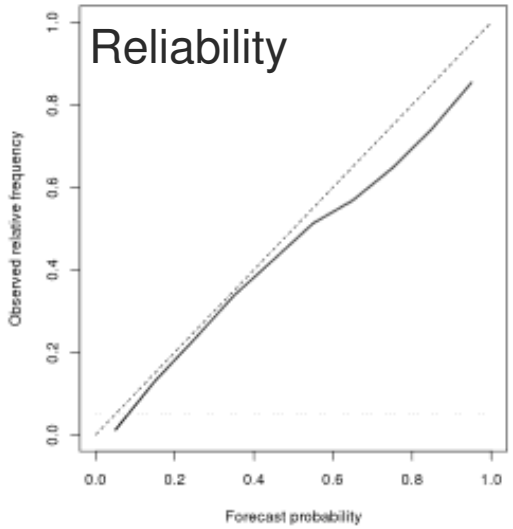



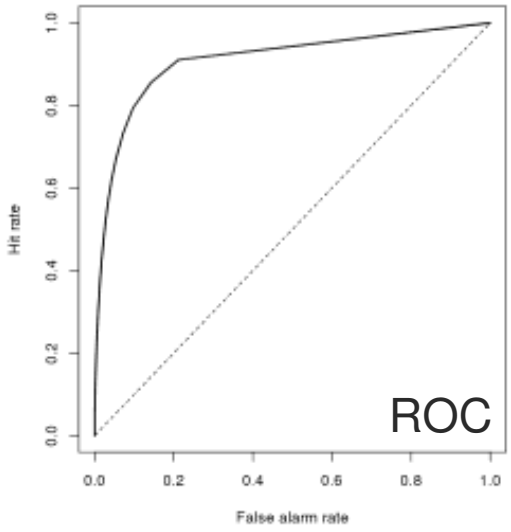
Met Office | Centre for Ecology & Hydrology, Natural Environment Research Council

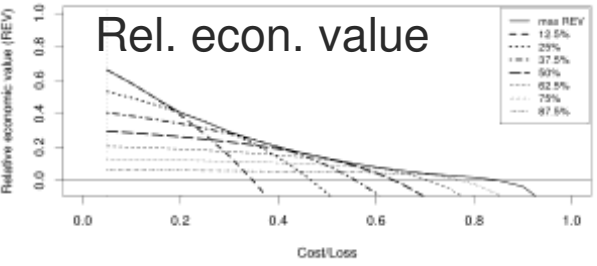
**Day 1 forecasts, England & Wales**

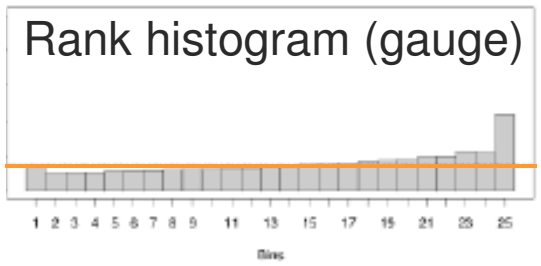
# National scale variation

## Rainfall *hourly* accumulation 95%ile (gauge)

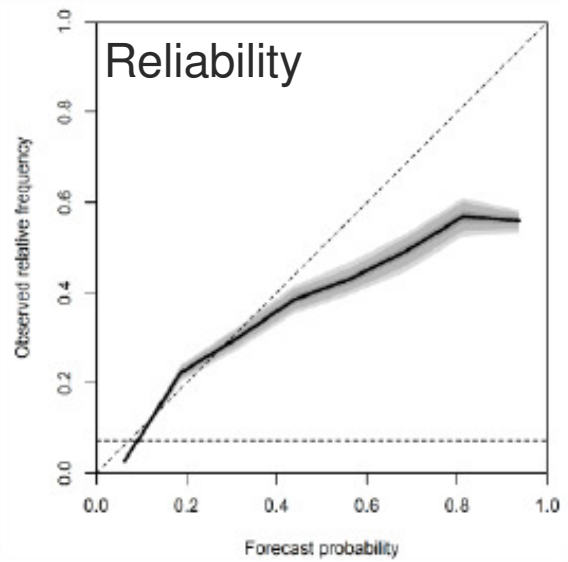


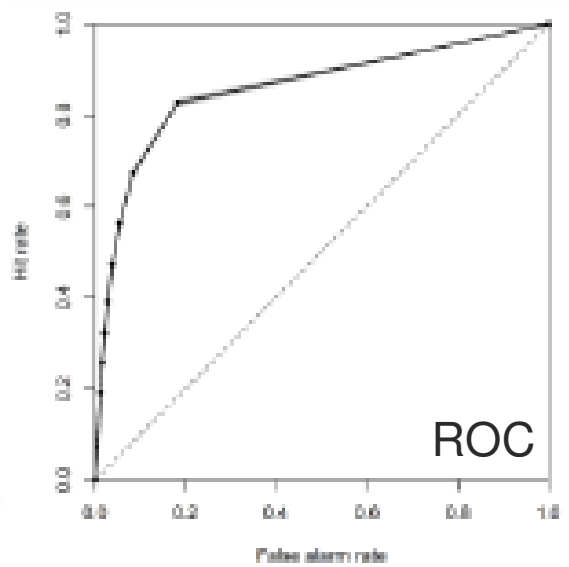



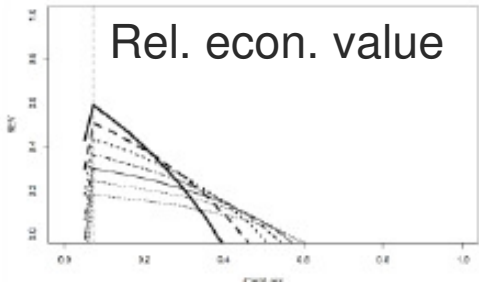



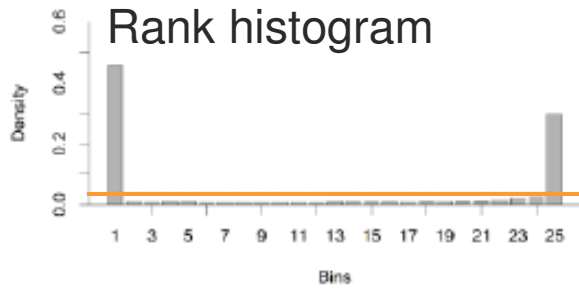




## River flow threshold: Q(2)/2



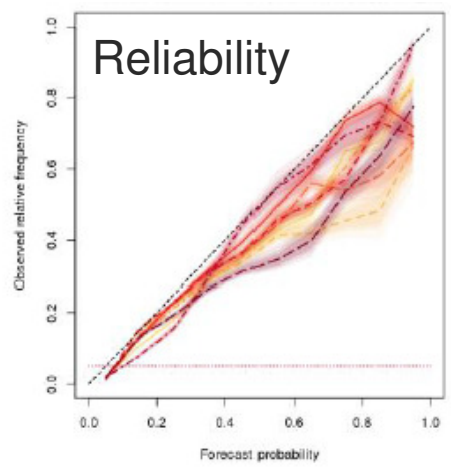



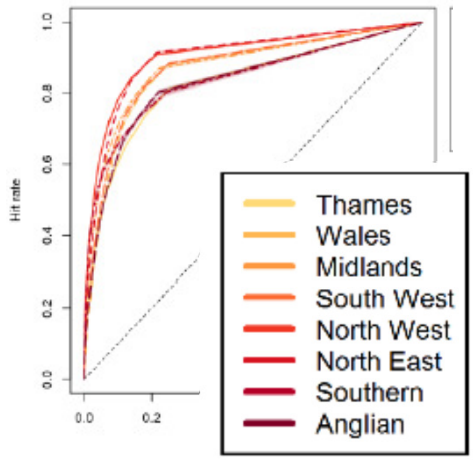



# Regional scale variation

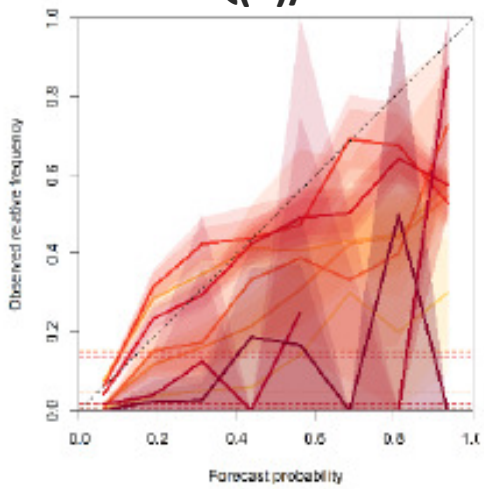
- Day 1 results for England and Wales winter.
- River flow forecasts generally **over-confident**.
- **Sampling issues** for the river flow even for very modest thresholds, leading to very large variations.
- Precipitation forecasts also **over-confident** but fairly reliable with good potential skill (Area under ROC).

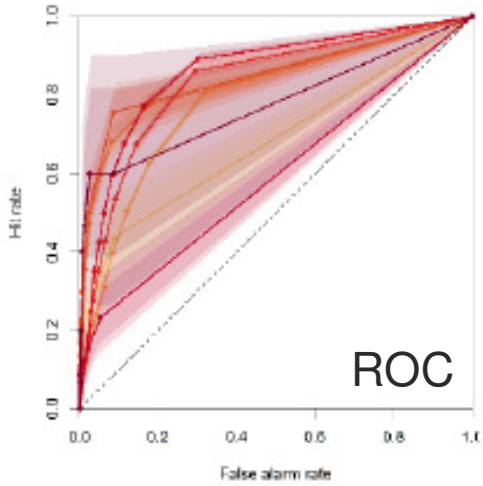
**Rainfall *hourly* accumulation 95%ile (gauge)**




**River flow threshold: Q(2)/2**



# Spatial distribution of the CRPSS

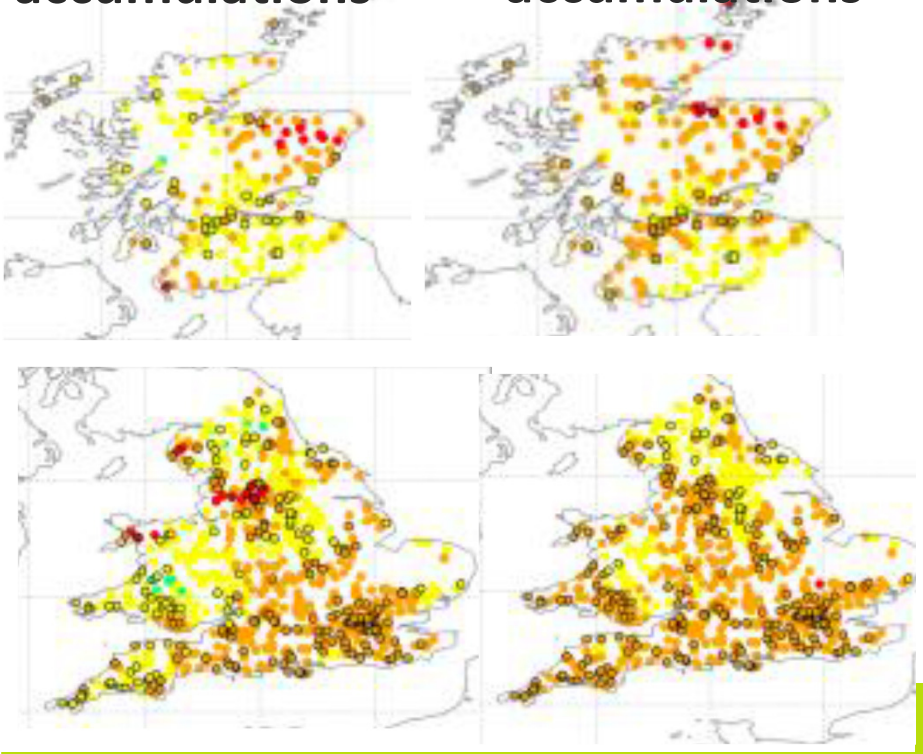
Use sample climatology
Bigger is better

Spatially noisy. Large decrease in skill from day 1
Very little skill by day 4

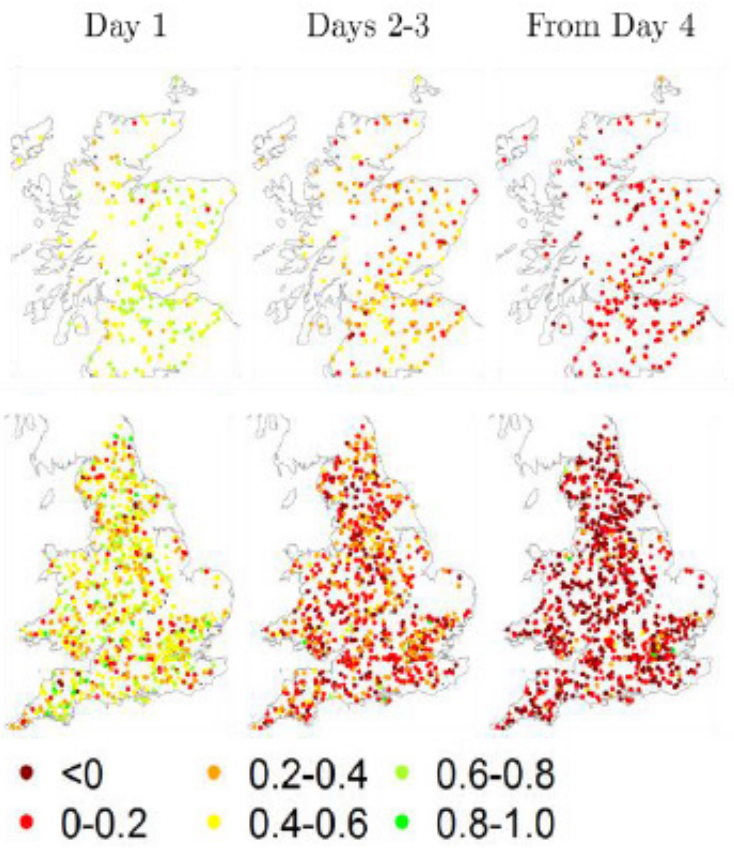
## Rainfall Day 1 (gauge)

daily accumulations

hourly accumulations

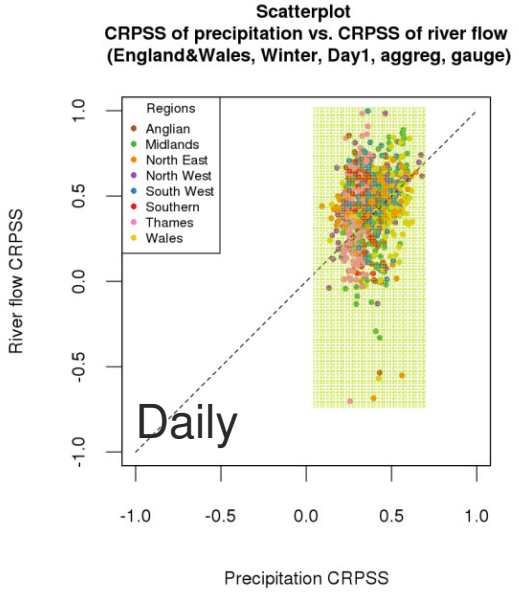


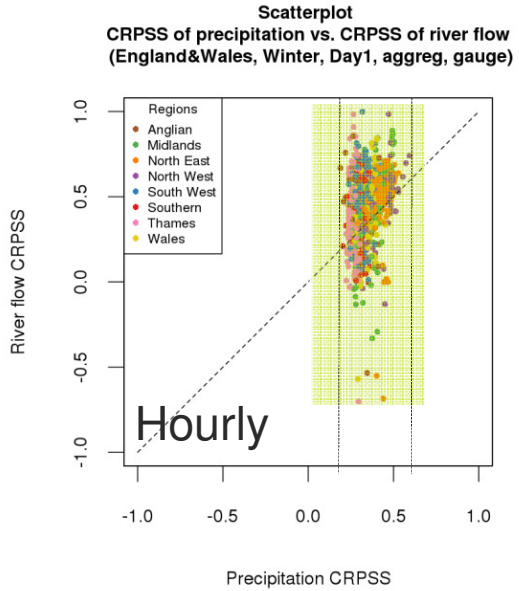

## River flow

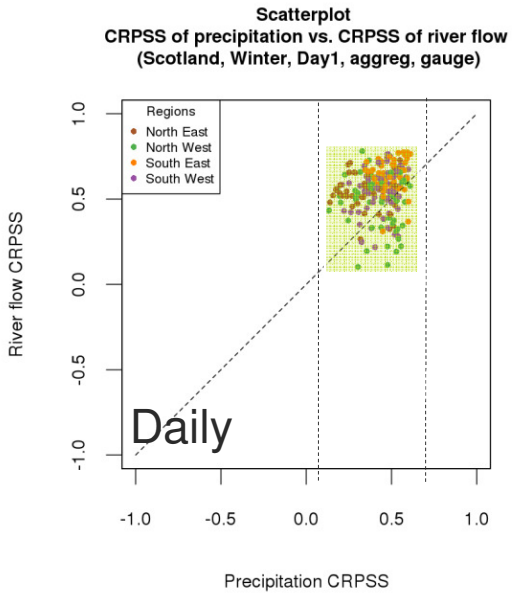


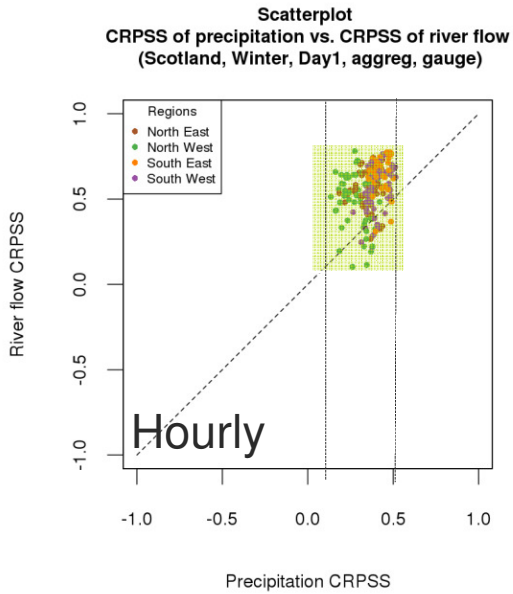

Spatial uniformity
Impact of snow
Little variation between 1h and 24h

# Rainfall v river flow skill




- Larger spread in river flow scores, but a greater proportion of river flow CRPSS are higher than the rainfall CRPSS for the same catchment.
- Strong regional dependence/clustering.
- Reduction in spread of rainfall scores for hourly accumulations.

# Conclusions

- **Sampling uncertainty** is the major consideration to obtain robust, meaningful verification information for ensembles in particular.
- **Observation uncertainty** can have a noticeable impact on the verification, especially on the rank histogram and the evaluation of ensemble spread.
- **Probabilistic forecasts derived from both river flow and precipitation accumulation ensembles tended to be over-confident**, with over-confidence increasing with forecast probability, threshold and lead-time. This is true particularly for river flow forecasts.
- The **river flow ensemble was found to be severely under-spread according to the Rank Histogram**. This suggests that unaccounted-for uncertainties in the hydrological modelling process may be important for forecast accuracy.
- Both river flow and precipitation ensembles showed **good potential skill.**
- Threshold-based verification scores were found to be **regionally dependant.**
- **Daily and hourly precipitation accumulations lead to similar overall conclusions**

Met Office

Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

FLOODFORECASTINGCENTRE
a working partnership between Environment Agency | Met Office

SEPA
Scottish Environment Protection Agency
Buidheann Dìon Àrainneachd na h-Alba

# Questions?

Work commissioned by:
Flood Forecasting Centre (EA)
Scottish Flood Forecasting Service (SEPA)