Met Office verification: Recent activities

Marion Mittermaier et al.

Met Office

Outline

Comparing hourly NWP with nowcast

Conditional (process-based) verification

Making SEEPS more relevant for sub-10-km models with a TRMM-based climatology

Catchment-scale precipitation and river-flow ensemble verification

Launch of the 2nd verification challenge on the “best new user-relevant metric” using *non-conventional observations*

Met Office

Nowcast-forecast comparison

FSS comparison between STEPS control and NWP (hourly cycling UKV)

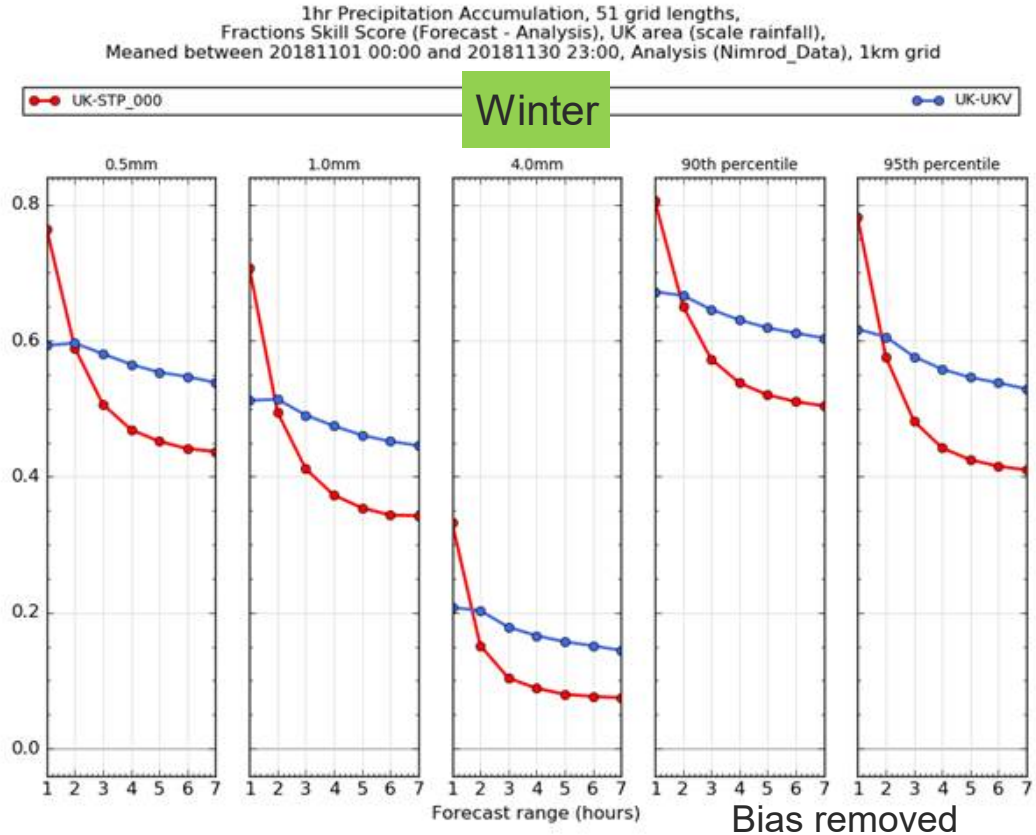
This evaluation is appropriate for “raw” forecasts.

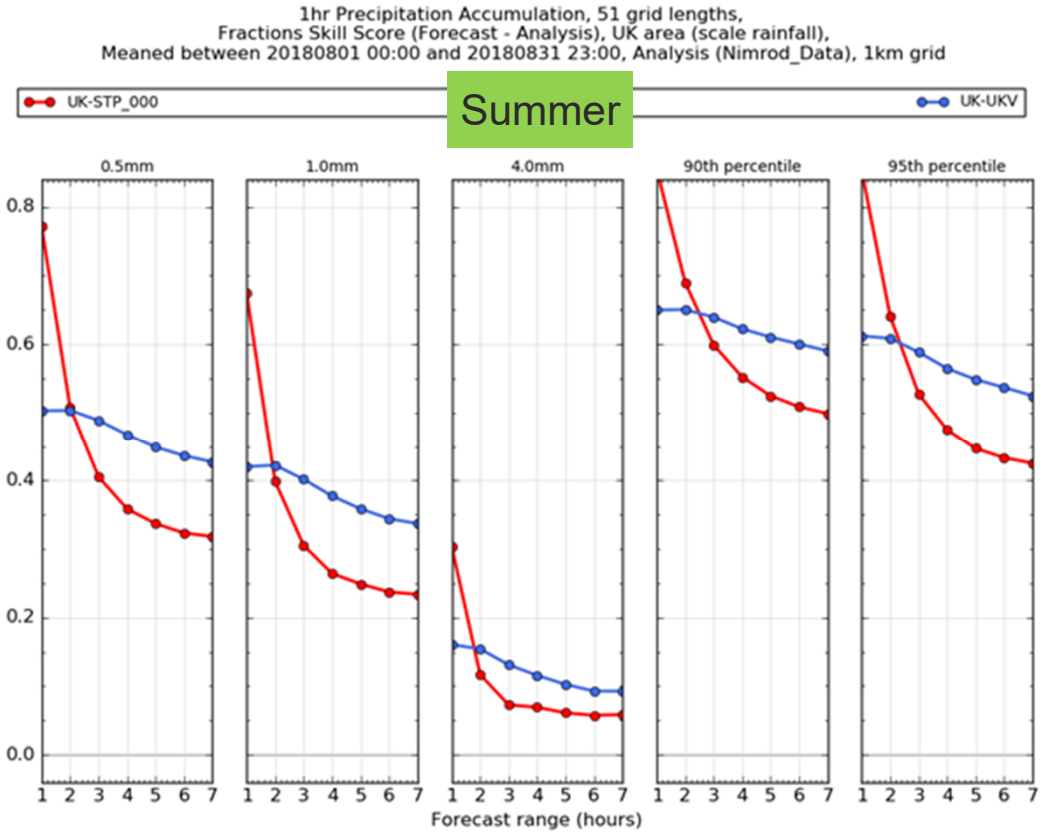
UKV hourly cycling compared to STEPS control

Before hourly cycling UKV cross over was between t+2h and t+3h

Now between t+1h and t+2h, especially for higher thresholds

Note: 51 km neighbourhood may not be enough to show useful skill for hourly precip.

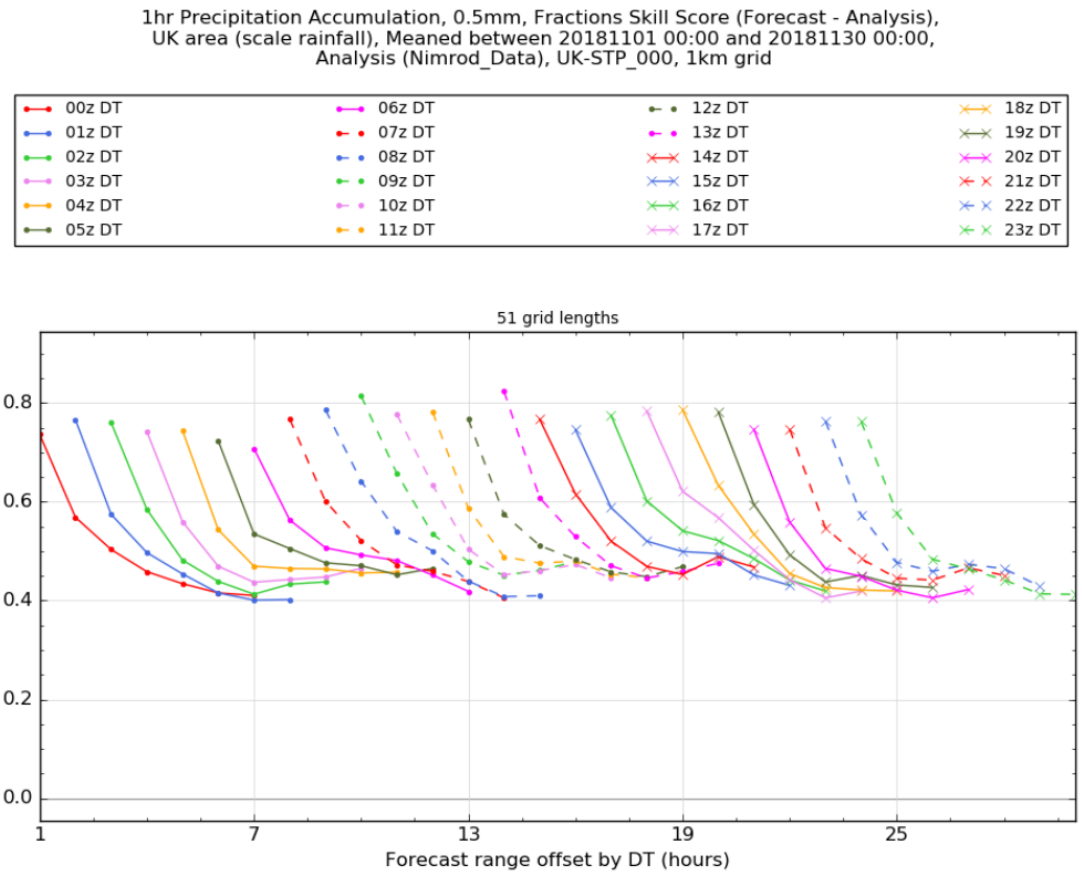


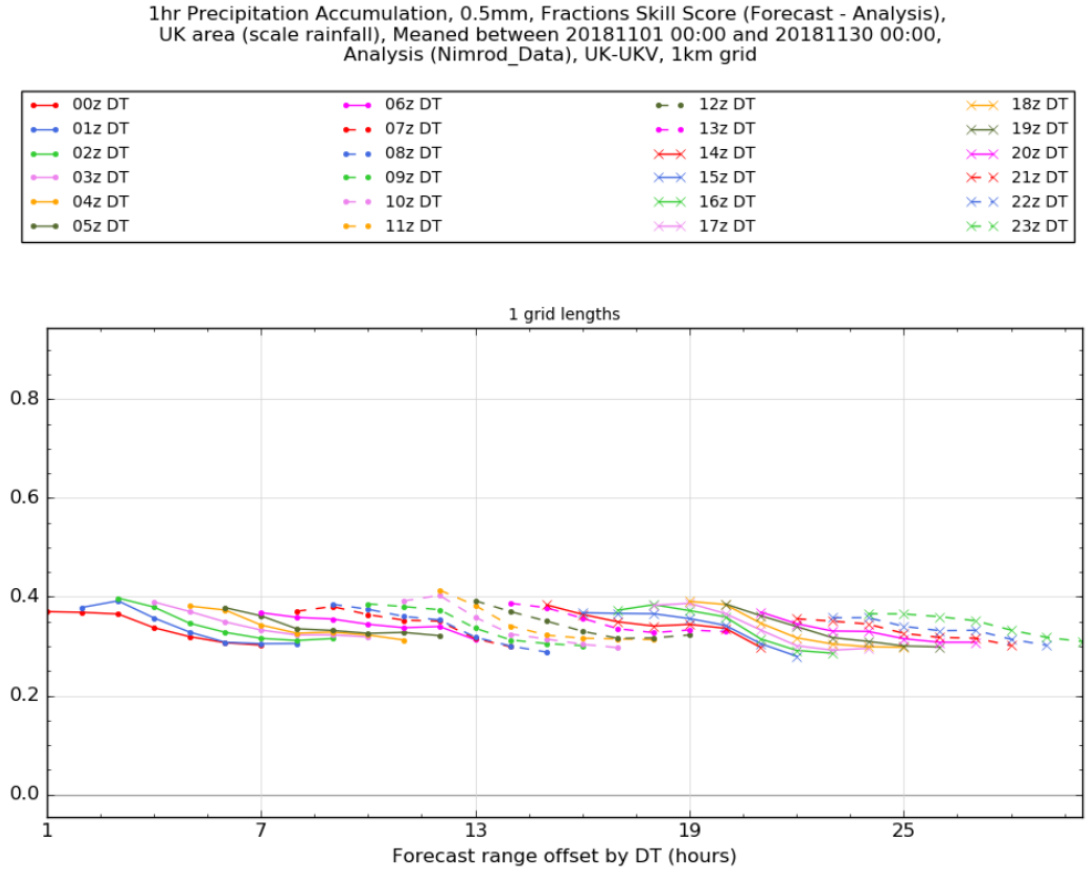


Bias removed

All precip verified

FSS diurnal variations at 51 km





Reduction in skill in nowcast much more dramatic in first 3 hours. NWP far more muted.

Met Office

Conditional verification

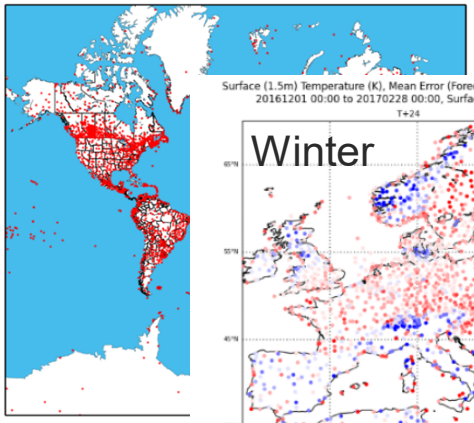
Stratification of TCA errors and biases by CBH and CTP

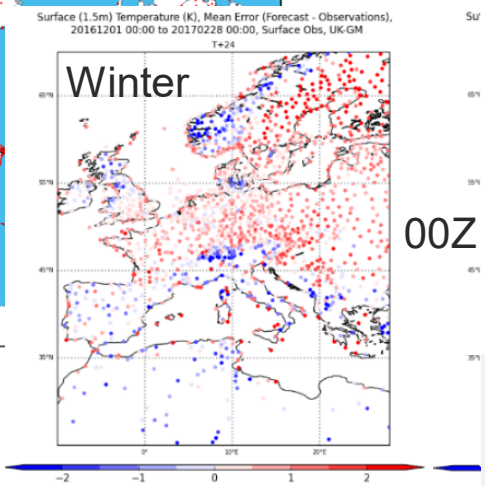
Stratification of temperature biases by location and land-surface type

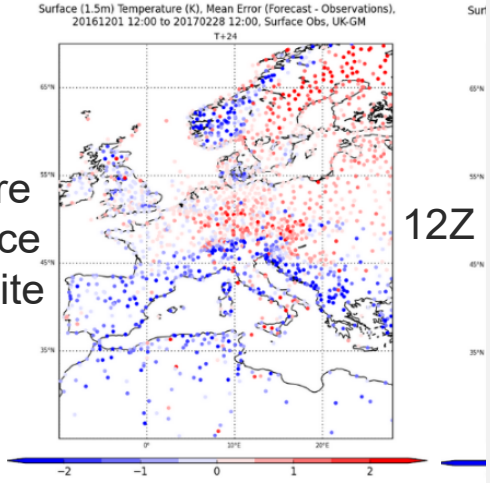
Csima and Mittermaier, 2019

Met Office Aggregating temperature scores

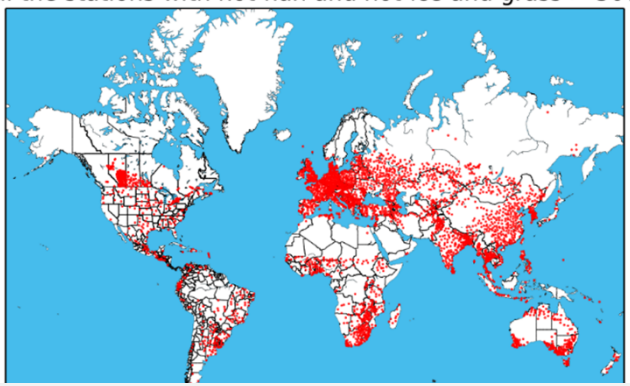
All the stations







All the stations with not-nan and not-ice and grass > 50%

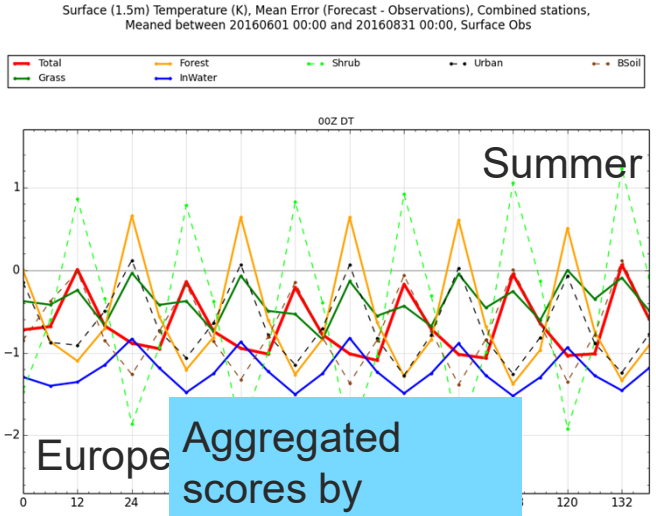


Local temperature performance may be quite different.

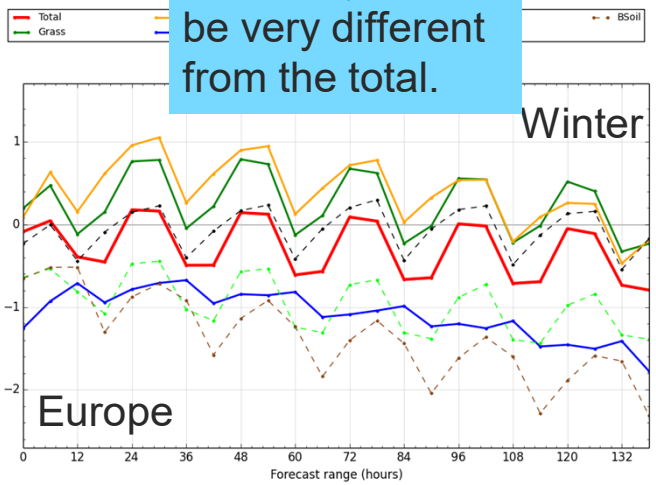
Observing sites are expected to be grass enclosures, unless this isn't possible (rock, sand, snow, ice).
This may also not be the case during the cold season in many mid-latitude locations (snow, ice).

Diagnosed forecast 1.5m T is a weighted average of temperatures for different land surface types.

We have 9 (sub-)tiles

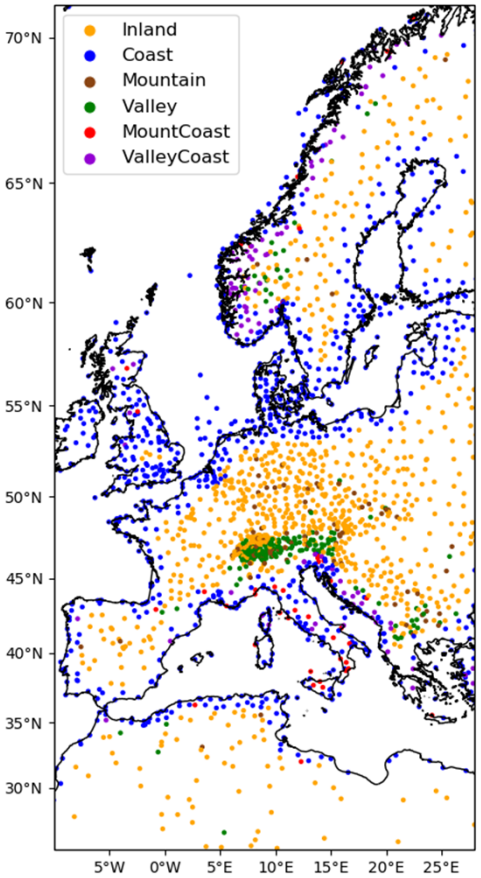


Aggregated scores by dominant land surface type can be very different from the total.

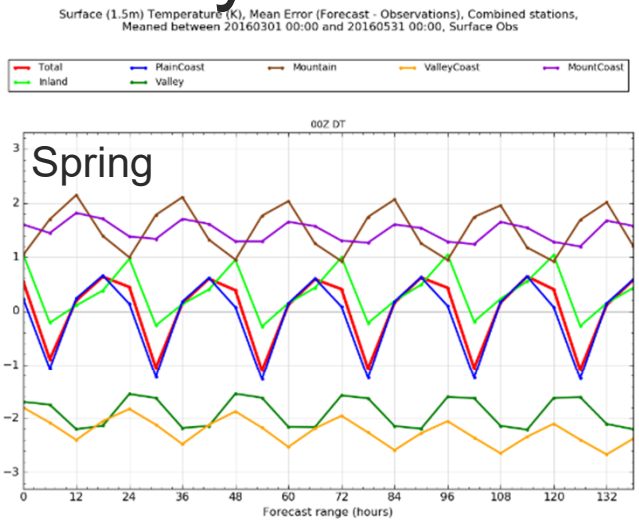


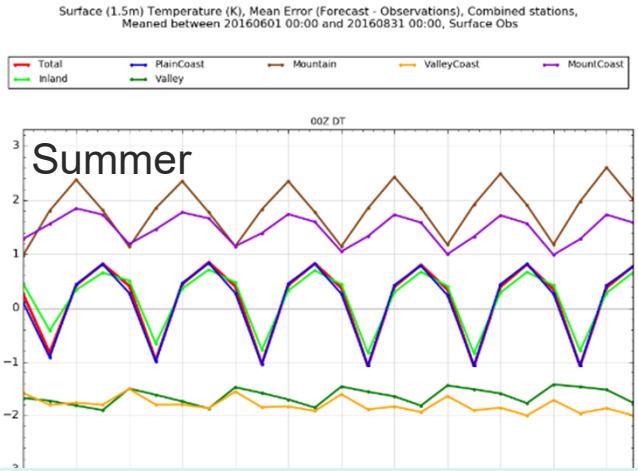
Stratification by location

Csima and Mittermaier, 2019

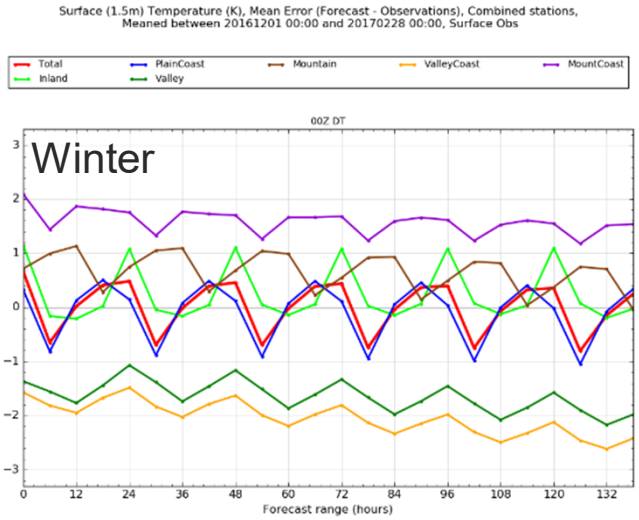


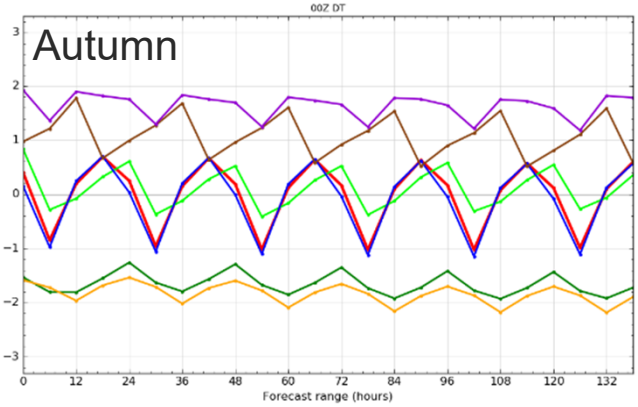
Inland	Coast	Mountain	Valley	Mount Coast	Valley Coast
4282	3424	156	528	98	239





Flat coastal and inland sites dominate the total bias (they are the most numerous).

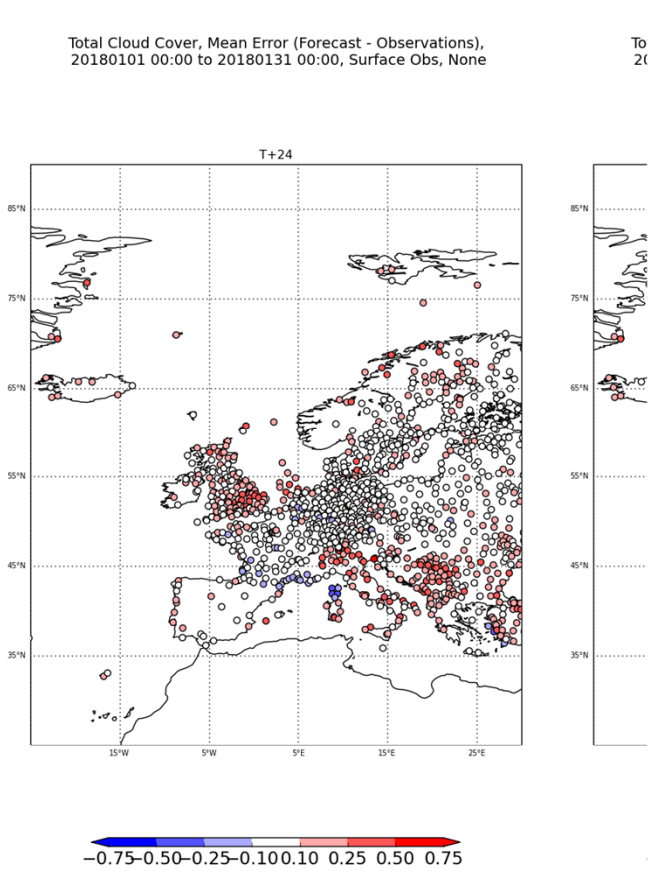


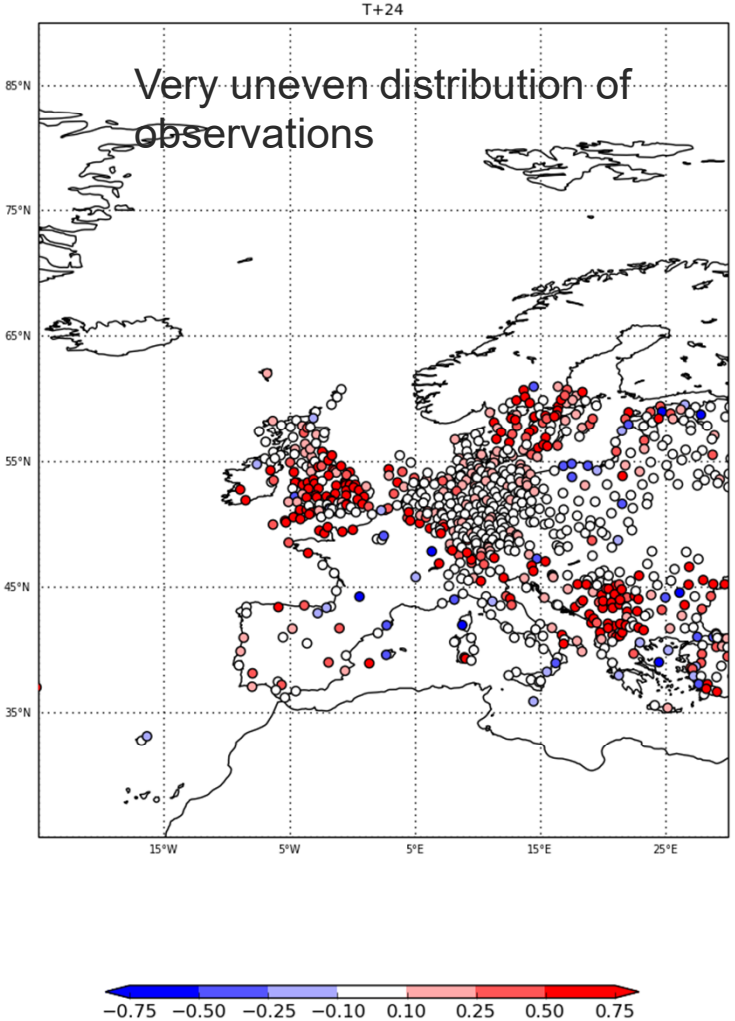


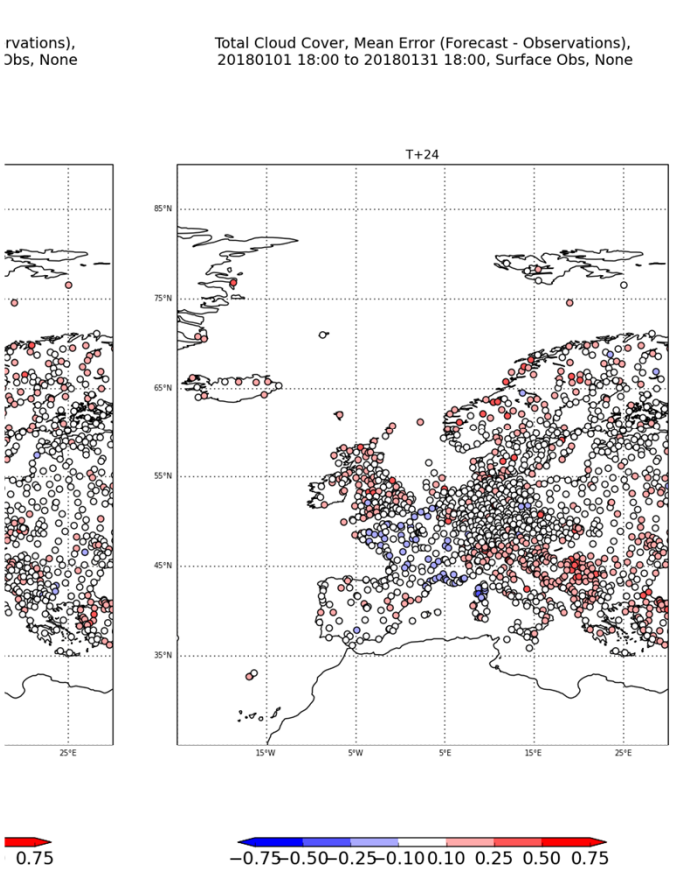
Met Office Diurna

Ric Crocker, 2019

- When equalising over time of day many locations lost







Observing practices
Observation density

Met Office

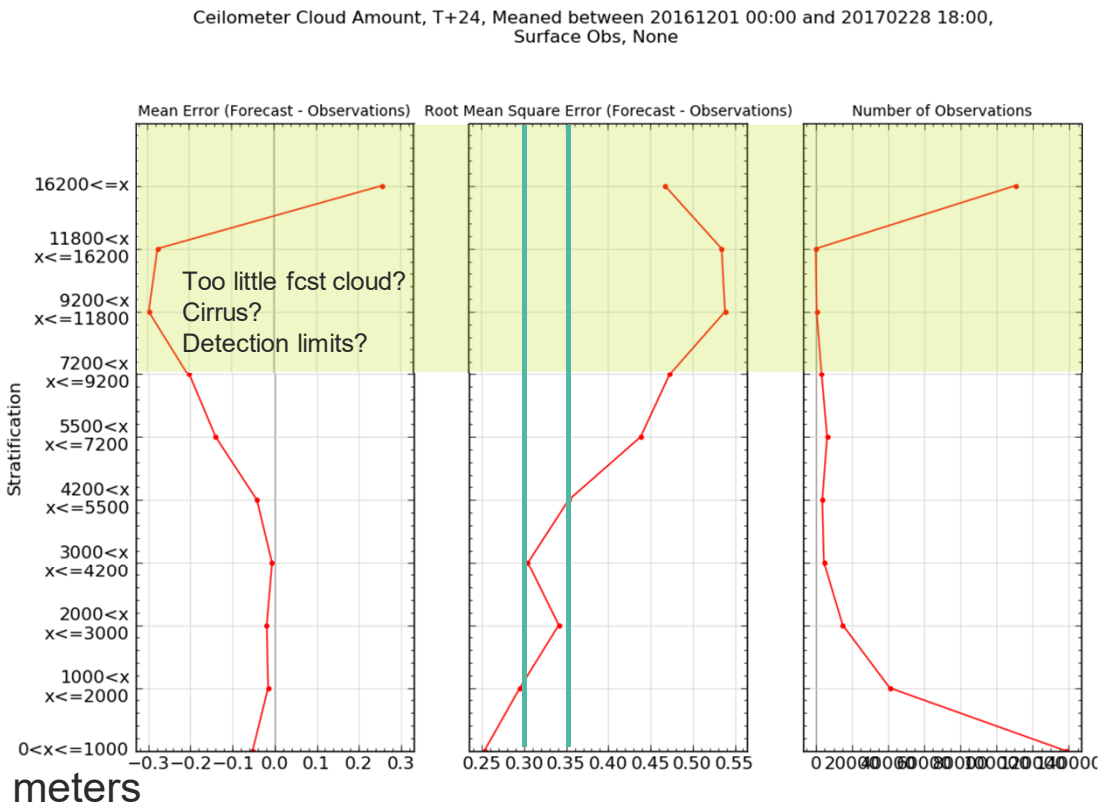
“Vertical” stratification

Ric Crocker, 2019

In approximately equalised bins

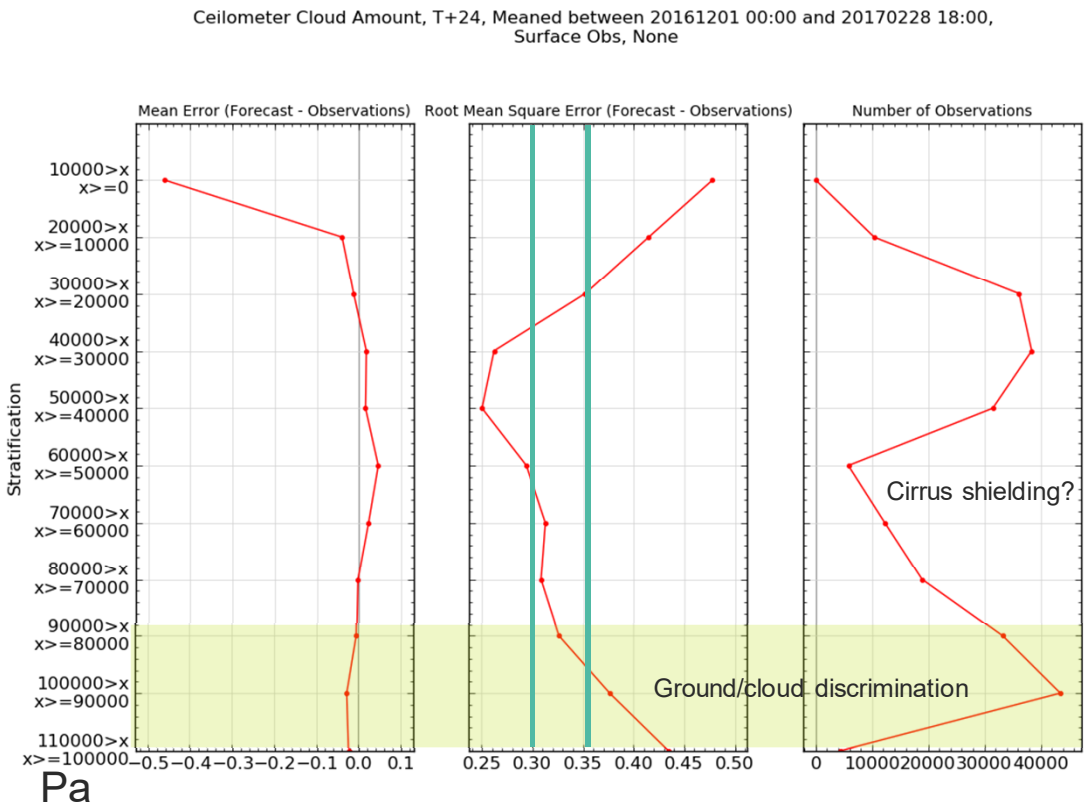
Using capped at 6.5 km Total Cloud Amount diagnostic

Bottom up



Cloud Base Height

Top down



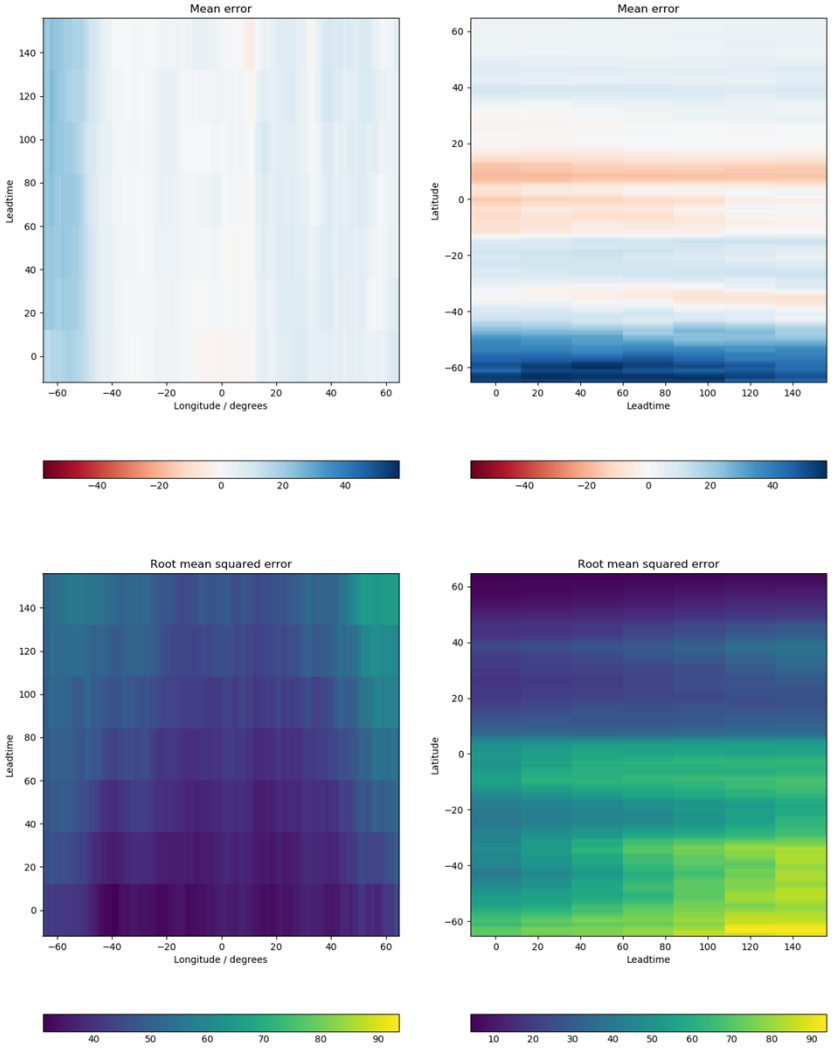
Cloud Top Pressure

Europe

Surface shortwave radiation (in W/m^2)

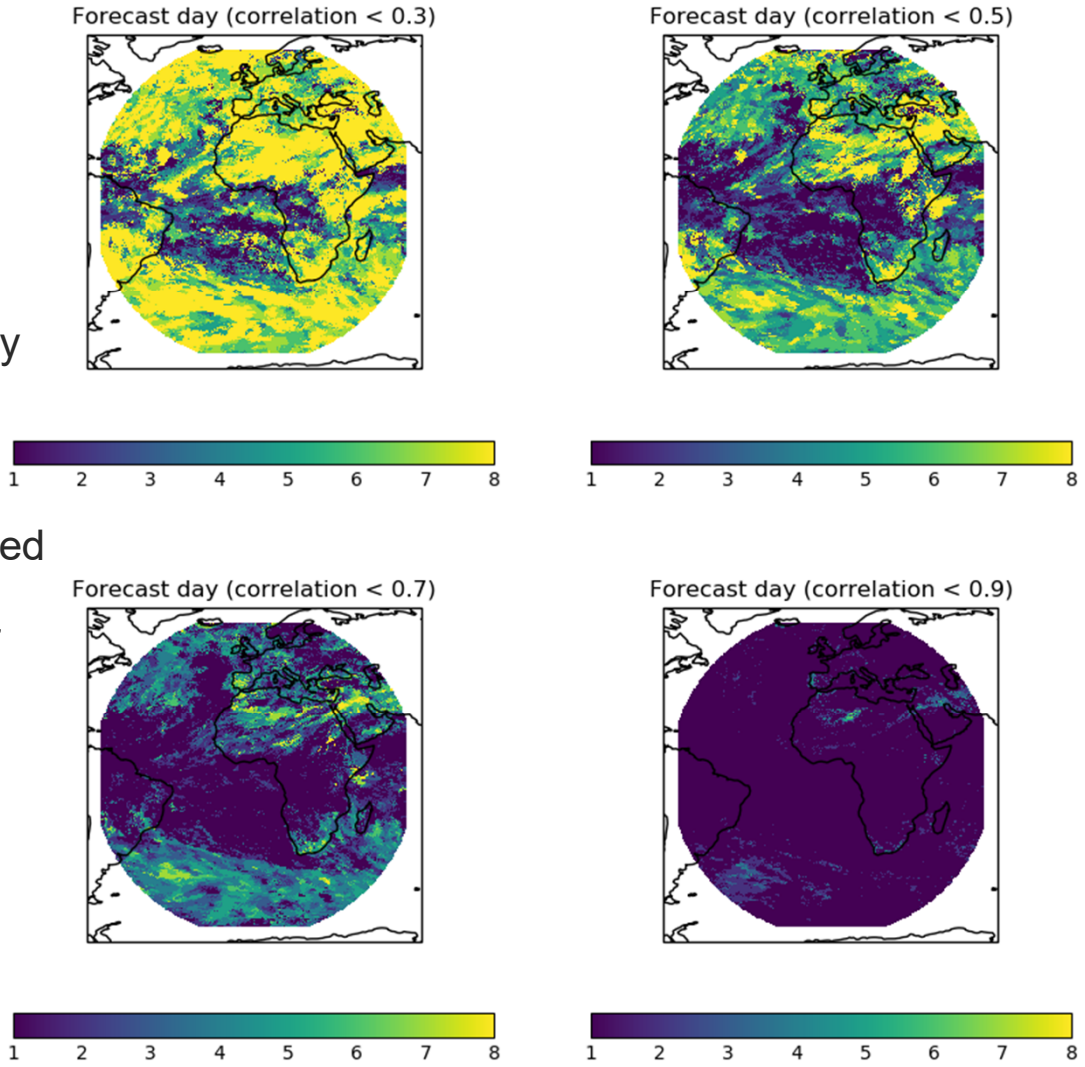
Jan 2019 - Forecast day where correlation < threshold

Ric Crocker, 2



Use as proxy for cloud amount

Not influenced by model diagnosis of cloud in column



Met Office

Stable Equitable Error in Probability Space (SEEPS)

A verification metric that was designed for monitoring model precipitation skill using a climatology derived from rain gauges to provide a climatologically “aware” assessment

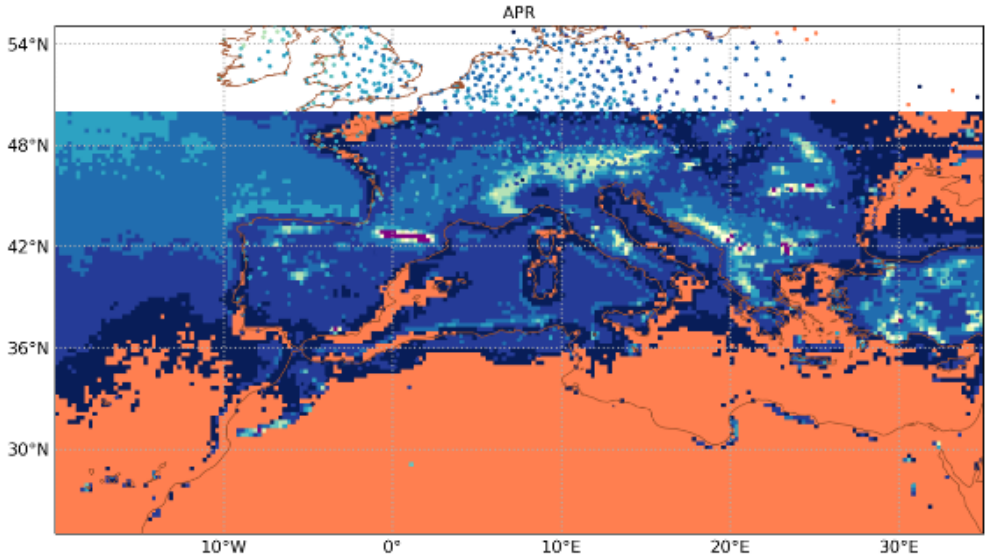
See Rodwell et al. (2010), Haiden et al. (2012) for details.

TRMM climatology
Daily scores

TRMM climatology

Compiled 1998-2015
Verification on 25 km grid, enables land-sea split

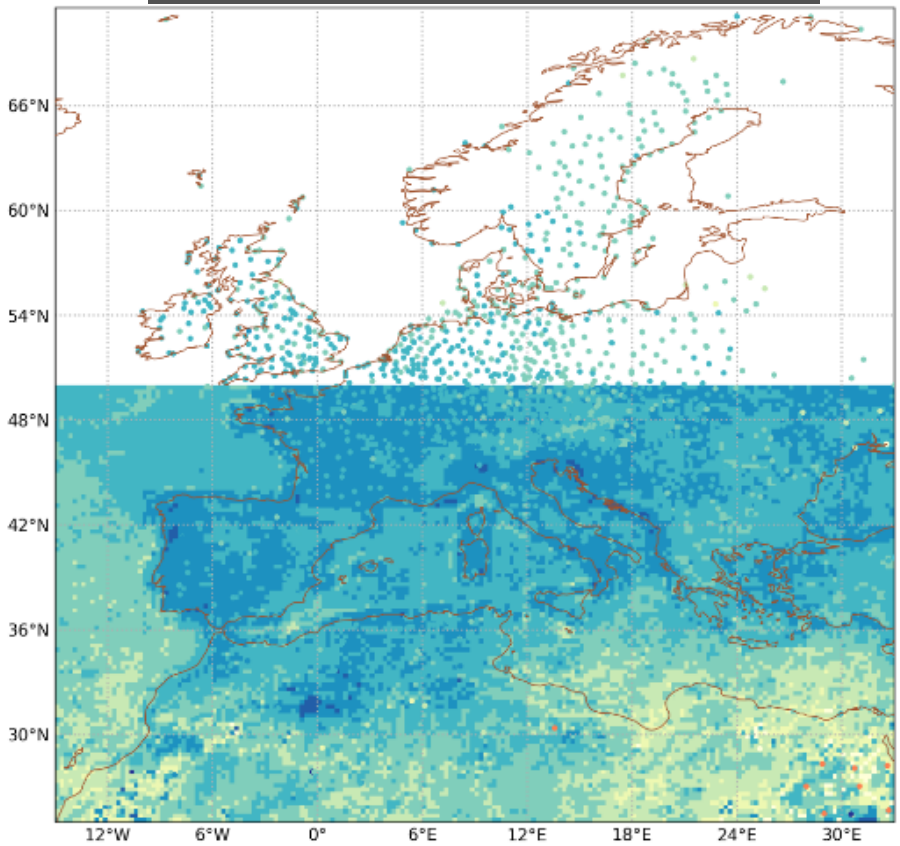
p1 – probability of dry (< 0.2 mm)



Very wet — very dry

April

t2 – threshold between light/heavy

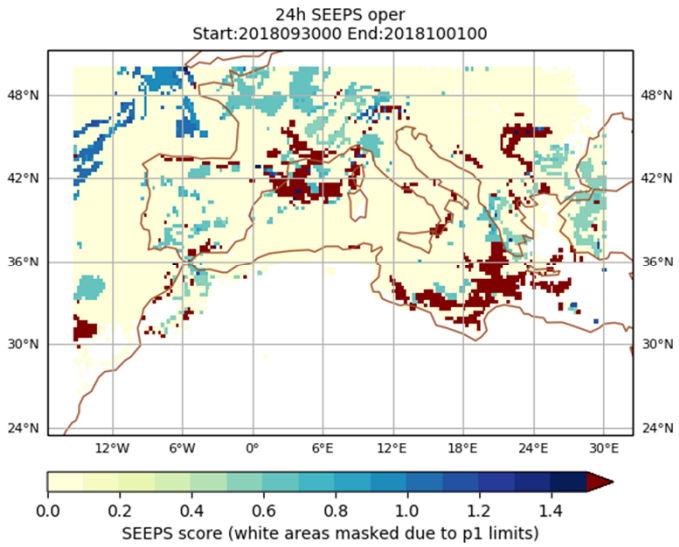


mm/day

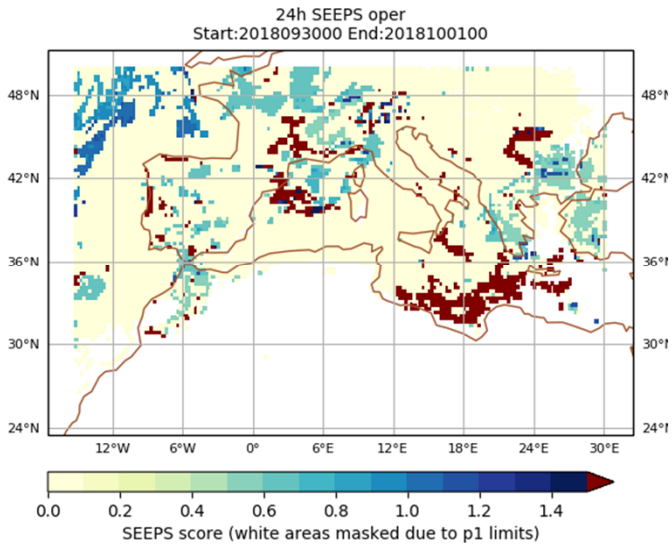
North, Mittermaier and Milton, 2019

Daily SEEPS: Europe

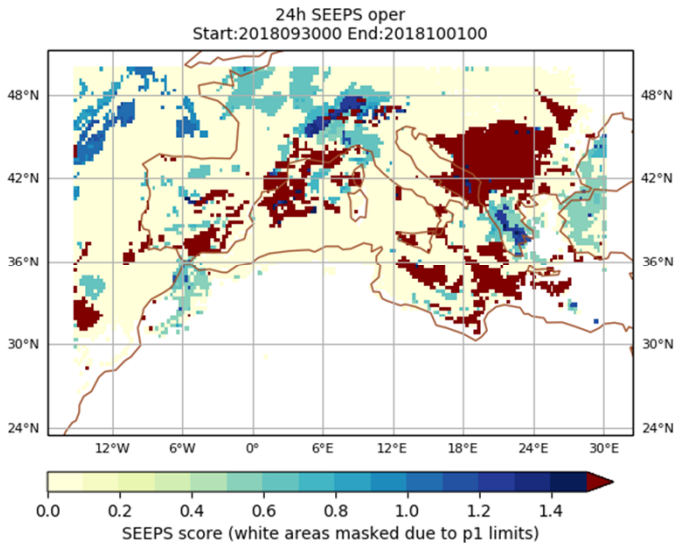
Day 1



Day 3



Day 5



= 0 is perfect
> 1 considered poor

North, Mittermaier and Milton, 2019

Errors follow synoptic features/systems
Regions of “gross errors”
Could work for km-scale as upscaling

Catchment scale ensemble verification

“Best Medium Range” (BMR) ensemble used to drive river flow ensemble based on the distributed G2G hydrological model. BMR is a “stitched together in time” ensemble providing output to 144h.

BMR includes STEPS ensemble at the start. STEPS is an extrapolation nowcast based on radar data.

BMR will be replaced by IMPROVER at some stage.

Evaluating accuracy and skill at the catchment level is being addressed in a joint project between the Met Office and the Centre for Ecology and Hydrology (CEH). Phase 2 started in December.

Some Phase 1 results are shown here. The observation type can be very influential and may skew interpretation of verification results.

Met Office

Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

Datasets

Fig. 1 The UK weather radar network

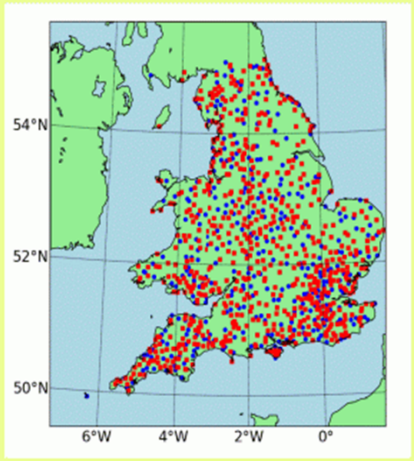
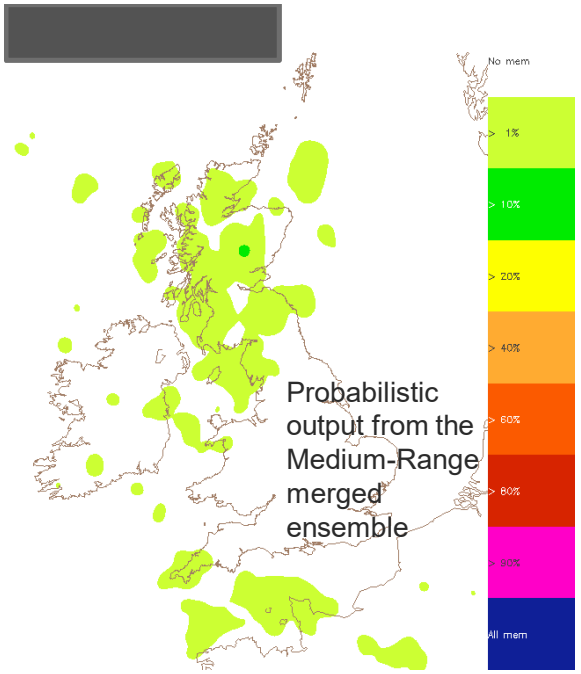


Fig. 2 UK real-time rain-gauge network.
● = Met Office ● = Environment Agency

- Radar rainfall analyses
- Gridded raingauge rainfall analyses

Two periods considered:
Winter Nov-Dec 2015 & Summer May-Jun 2016



Probabilistic output from the Medium-Range merged ensemble

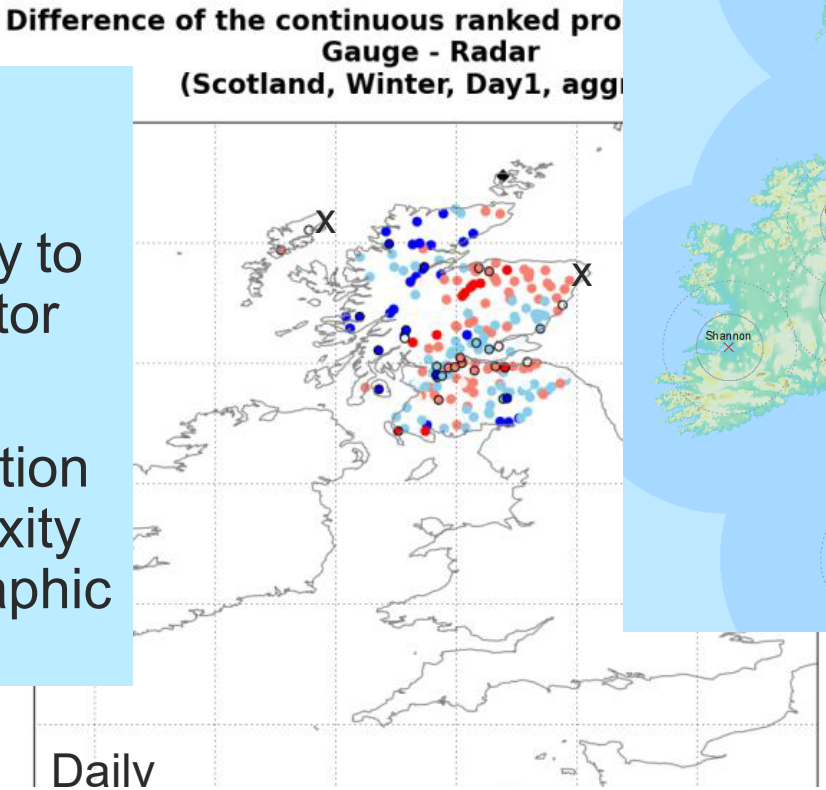
Deterministic example of G2G output

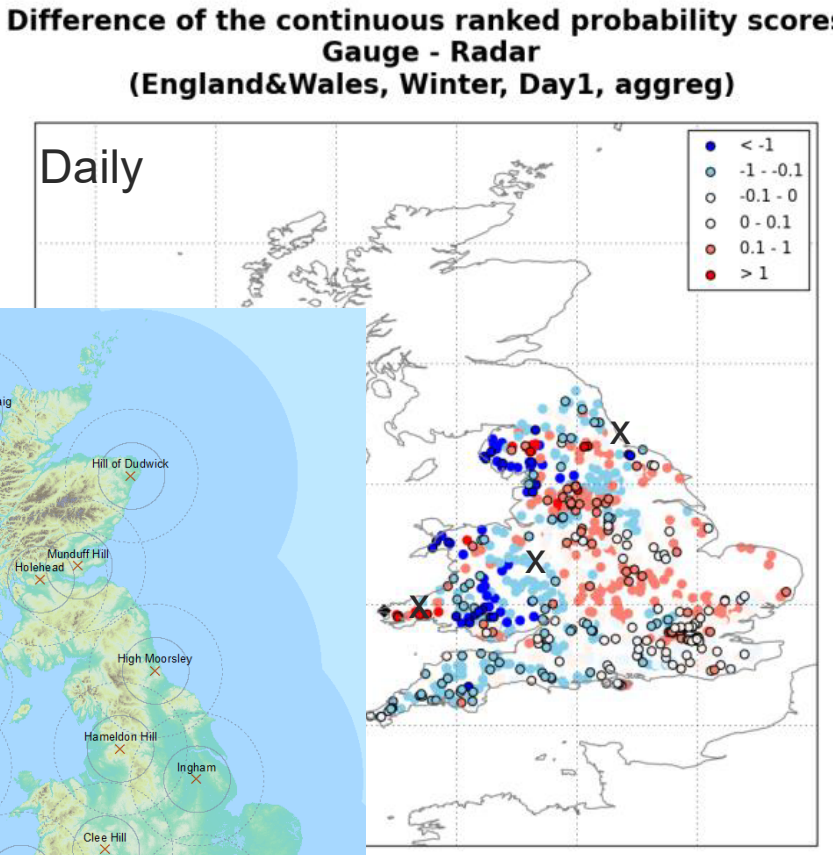
- Medium Range Ensemble:
  - **Nowcast (STEPS, t+7h)**
  - MOGREPS-UK (2.2km, ~t+32h)
  - MOGREPS-G (32km, ~t+144h/6 days)
- Precipitation output: 15 min, 2km
- River-flow ensemble (G2G) output: 15 min,1km

Anderson et al 2019, J. Hydrol

Sensitivity of CRPS to observation type: raingauge 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





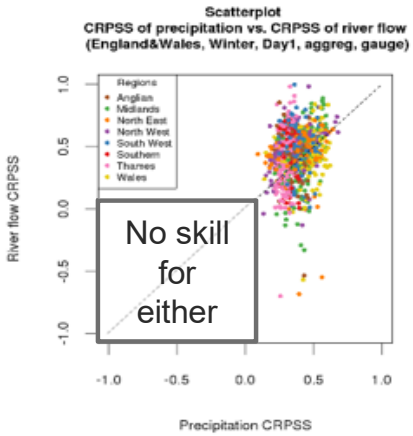
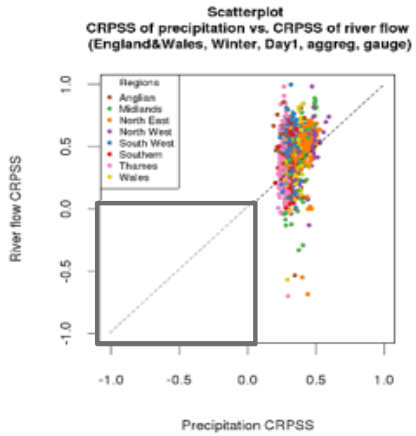
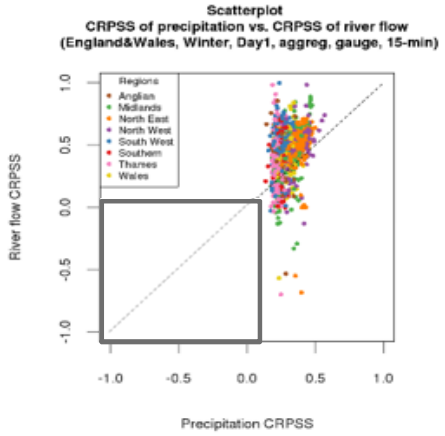
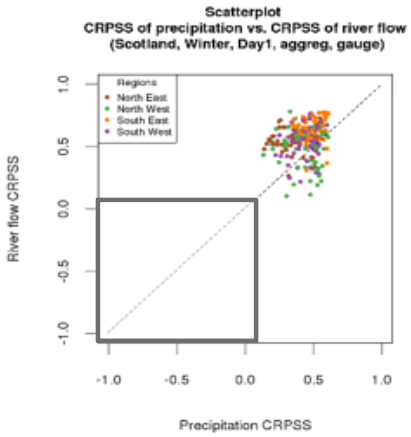
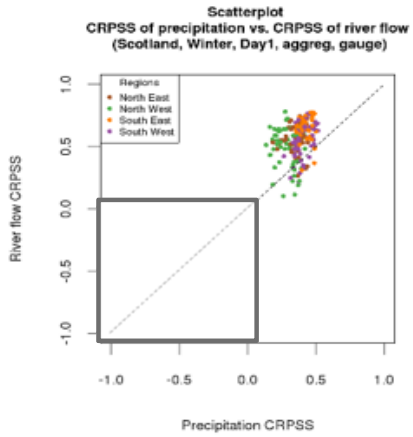
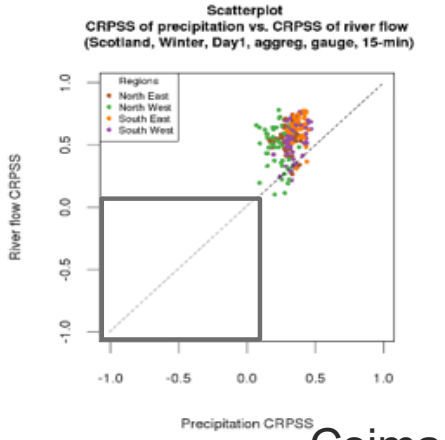
Radar has noticeably worse scores over complex terrain

Anderson et al. 2019 (J. Hydrol)

Comparison of precipitation and river flow CRPSS for raingauge, Winter period (Day 1)

Met Office

Centre for Ecology & Hydrology
NATURAL ENVIRONMENT RESEARCH COUNCIL

	Daily	Hourly	15-min
E&W			
Scotland			

Precise matching

Much larger range of scores for river flow

No negative scores for Precipitation

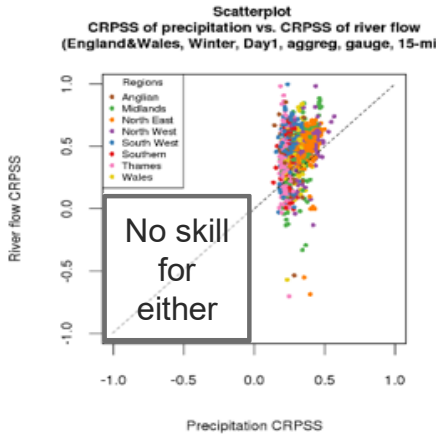
River flow scores are based on 15 min throughout

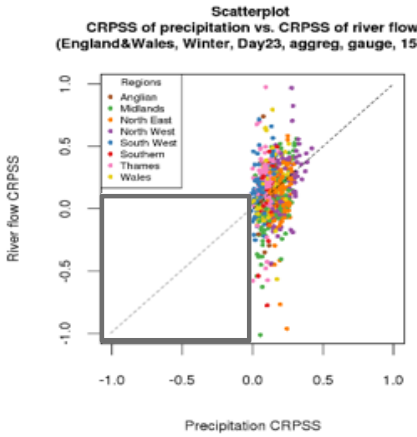
Csima and Mittermaier, 2018, 2019

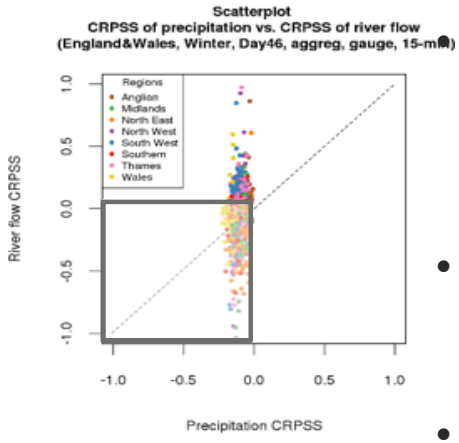
Comparison of precipitation and river flow CRPSS for raingauge, Winter (15 min)

Day1 | Day23 | Day46

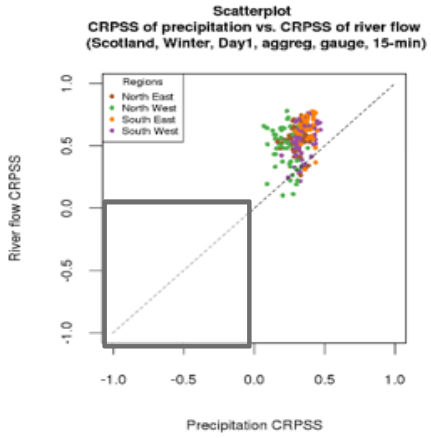
E&W

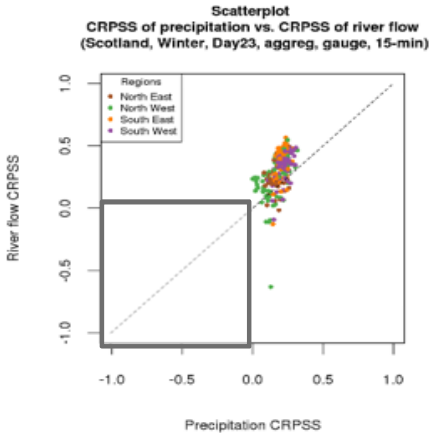


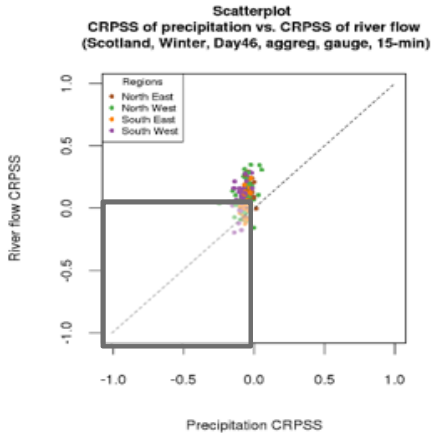




Scotland



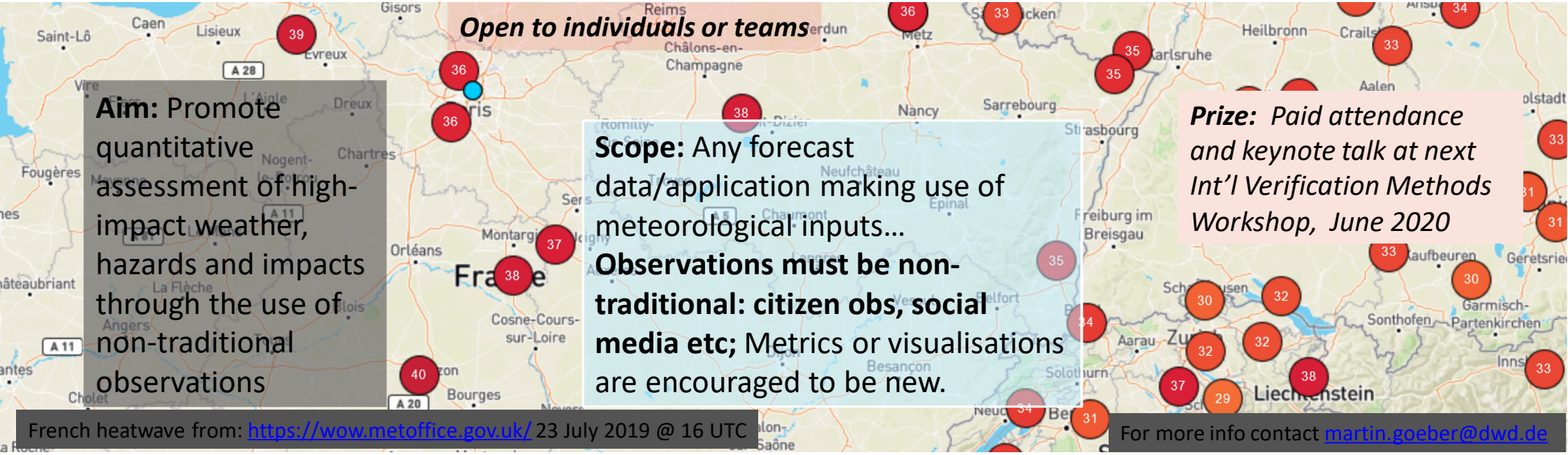




Precise matching

- Reduction in skill with lead-time for both river flow and precipitation
- Scatter for river flow increases with lead-time
- Scatter for precipitation narrows with lead-time
- Differences primarily because the range of CRPS is greater for longer accumulation periods

2nd Challenge to develop and demonstrate the best new forecast verification metric *using non-traditional observations*

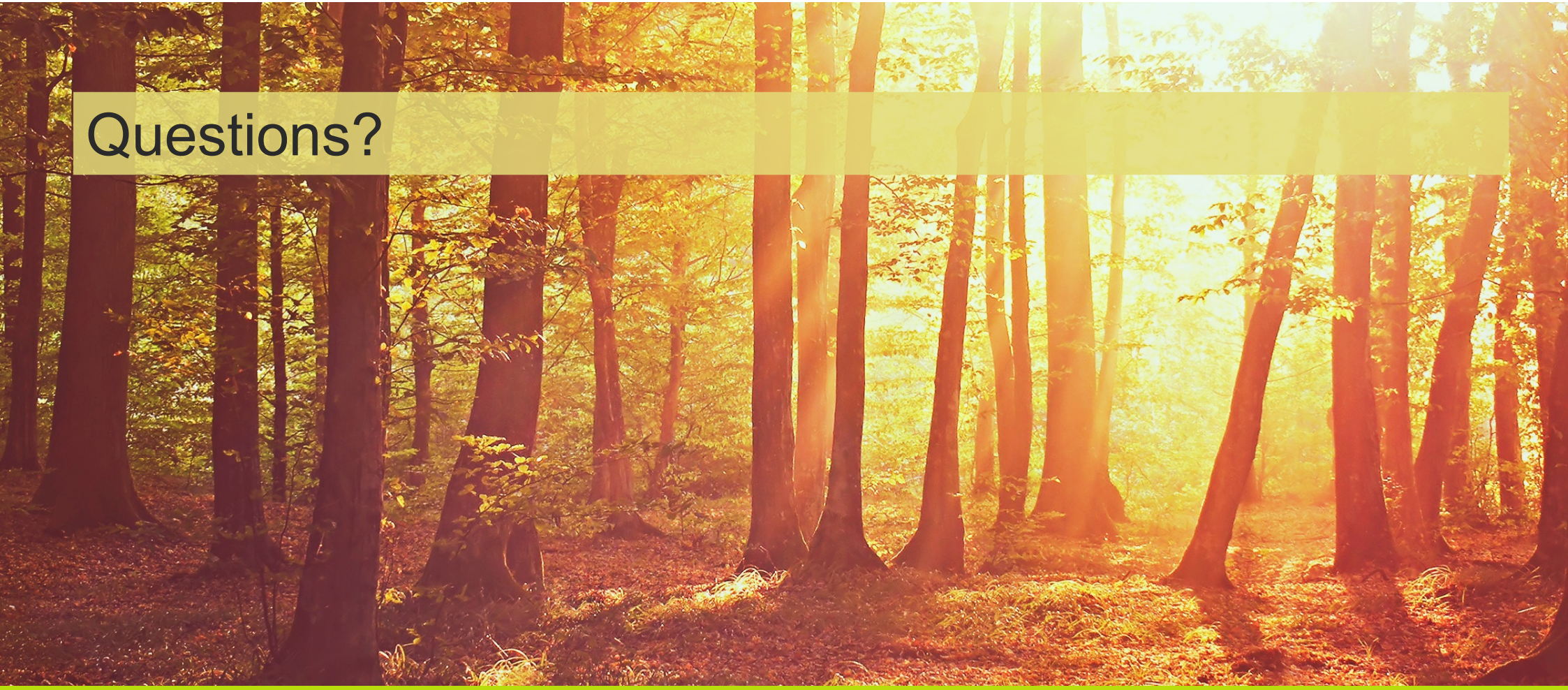


Timeline :

- Launch, EMS, Copenhagen, September 2019
- Deadline for entries : 15 February 2020
- Announcement of winner : end March 2020

Run by WMO Joint Working Group on Forecast Verification Research in support of WWRP HiWeather Project

Met Office



Questions?