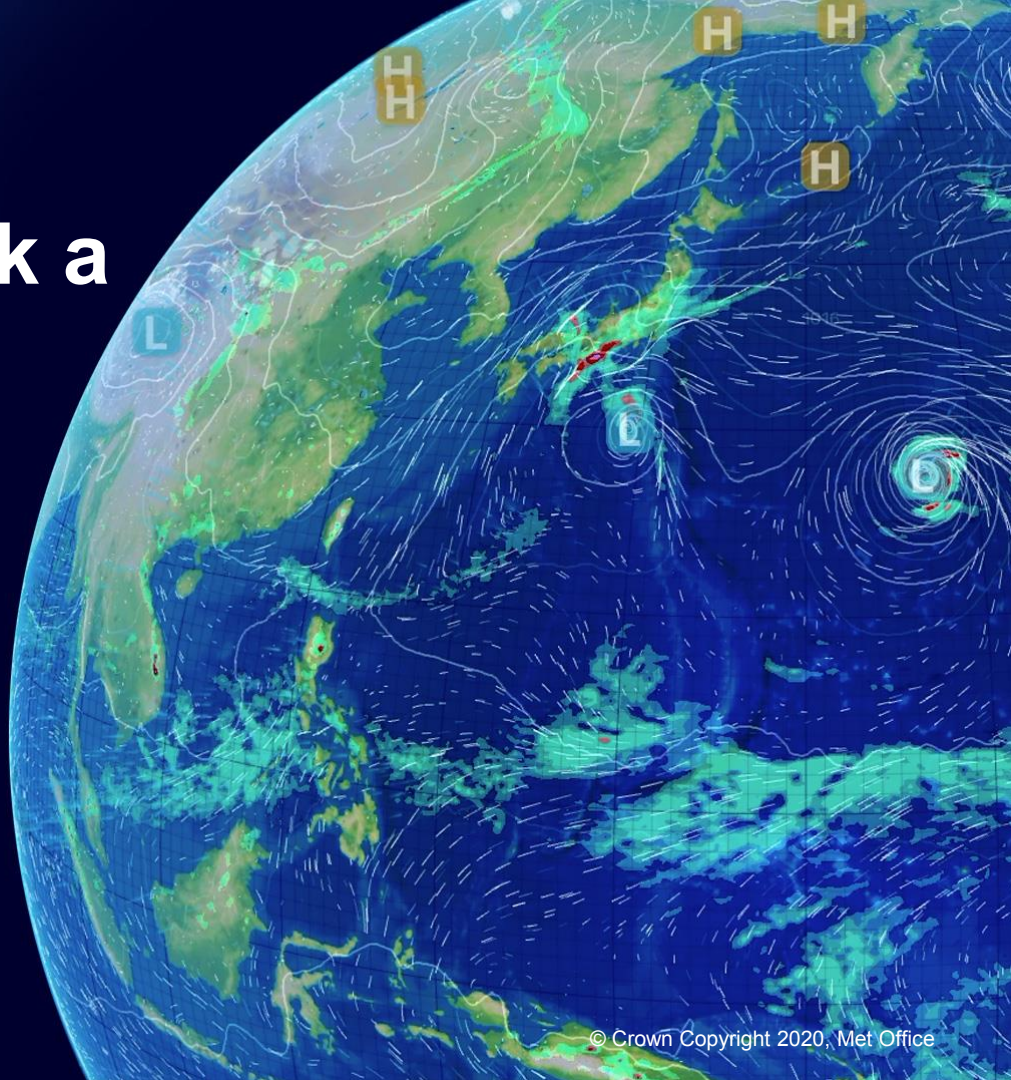
Met Office

Verifikációs projektek a Met Office-ban

Csima Gabi

2020/12/03



Tartalom

- A Met Office tudományos részlegei
- Az elmúlt kb. 6 év tudományos munkái
- 2 érdekesebb kutatási téma kicsit részletesebben

A Met Office tudományos részlegei (research – without services)

Weather science

- Data assimilation
- Ensemble forecasting
- Atmospheric dispersion and air quality
- Verification, impacts and post-processing
- Research to operations
- Ocean forecasting
- Satellite applications
- Observations research and development

Climate science

- Climate, cryosphere and oceans
- Understanding climate change
- Climate impacts
- Climate monitoring and attribution
- Earth system science
- Monthly to decadal prediction
- UK climate maps and data

Foundation science

- Dynamics research
- Global modelling
- Observation-based research
- Atmospheric processes and parametrization
- Regional model evaluation and development

Applied science

- Aviation applications
- Climate information for international development
- Climate security
- Climate services
- Defence applications
- Post-processing applications
- Science for impacts, resilience and adaptation
- Science health strategy

Verification, impacts and post-processing

Feladat:

- A numerikus időjárás előrejelzések outputjainak rendszerszerű és tudományos feldolgozása (előrejelzés és verifikáció készítése a Public Weather Service számára)

Csapatok:

- 2 **verification** (model diagnostics & operational)
- 2 post processing (gridded & site specific)
- Weather impacts
- System team

Az elmúlt évek tudományos munkái

- **Villámlás**
  - WAFC Cb horizontal extent előrejelzések verifikációja
    - Obs (1): villámlás (földi megf. pl. ATDNet) & OT
    - Obs (2): villámlás (GLMs - Geostationary Lightning Mapper – GOES16 & 17 műholdakon)
  - Villám előrejelzések verifikálása (Obs: Earth Networks data)
    - Viktória tó és környéke
    - India
- **Csapadék**
  - Csapadék és vízhozam ensemble előrejelzés verifikációja UK vízgyűjtőire (radar & csap. mérő áll. adatok)
  - Nagy csapadékú riasztások verifikációja Skóciára (radar)
- **Hőmérséklet**
  - Megfigyelő állomások geográfiai, orográfiai és domináns talajfajta alapján történő objektív rétegződése szerinti 2m-es hőmérséklet verifikáció
- **UV Index**
- **Összetett verifikációk**
  - HiRA (High Resolution Assessment)

Az elmúlt évek tudományos munkái

- **Villámlás**
  - WAFC Cb horizontal extent előrejelzések verifikációja
    - Obs (1): villámlás (földi megf. pl. ATDNet) & OT
    - Obs (2): villámlás (GLMs - Geostationary Lightning Mapper – GOES16 & 17 műholdakon)
  - Villám előrejelzések verifikálása (Obs: Earth Networks data)
    - Viktória tó és környéke
    - India
- **Csapadék**
  - Csapadék és vízhozam ensemble előrejelzés verifikációja UK vízgyűjtőire (radar & csap. mérő áll. adatok)
  - Nagy csapadékú riasztások verifikációja Skóciára (radar)
- **Hőmérséklet**
  - Megfigyelő állomások geográfiai, orográfiai és domináns talajfajta alapján történő objektív rétegződése szerinti 2m-es hőmérséklet verifikáció
- **UV Index**
- **Összetett verifikációk**
  - HiRA (High Resolution Assessment)

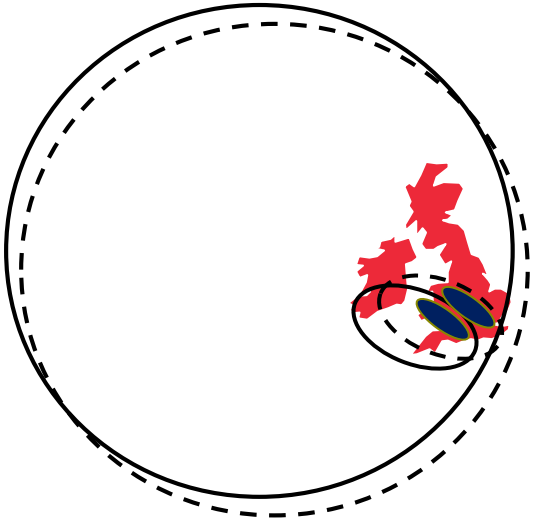
HiRA (High Resolution Assessment)

1. How to consistently demonstrate skill in increasingly higher resolution models? (*)
2. How to demonstrate skill of a high resolution ensemble forecast over a high resolution deterministic model? (How can we quantify the benefit of a km-scale ensemble forecast (MOGREPS-UK) over a similar resolution deterministic forecast (UKV)?)

*Mittermaier M.P., 2014: A Strategy for Verifying Near-Convection-Resolving Model Forecasts at Observing Sites. Wea. Forecasting

What is the problem with the traditional "pointwise" methods?

Small uncertainty at large scales ⟷ Large uncertainty at small scales



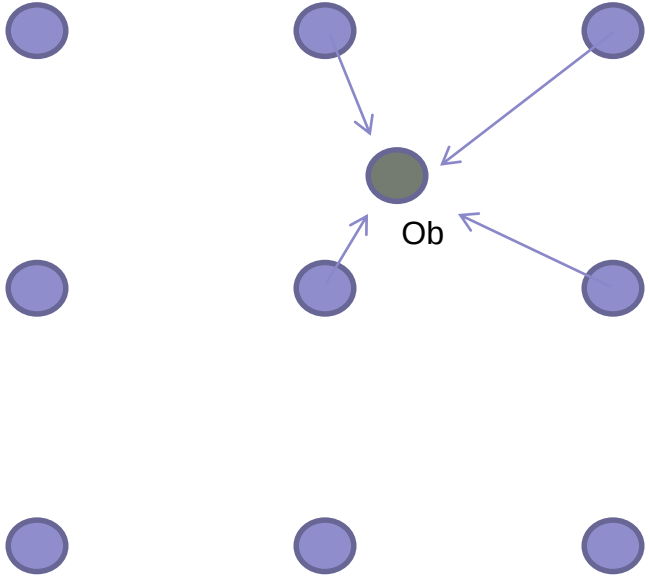
In small scales (high res. models):
timing & spacing errors, double penalty effect, ...etc.

Need to use **spatial verification method** and **probabilistically** handling

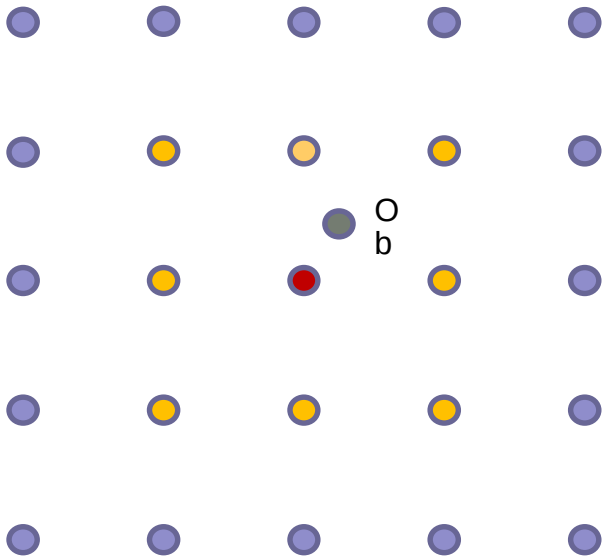
HiRA

High
Resolution
Assessment

Traditional verification method – “pointwise” comparing



Single-observation-neighbourhood-forecast approach (SO-NF) (Ebert, 2008)



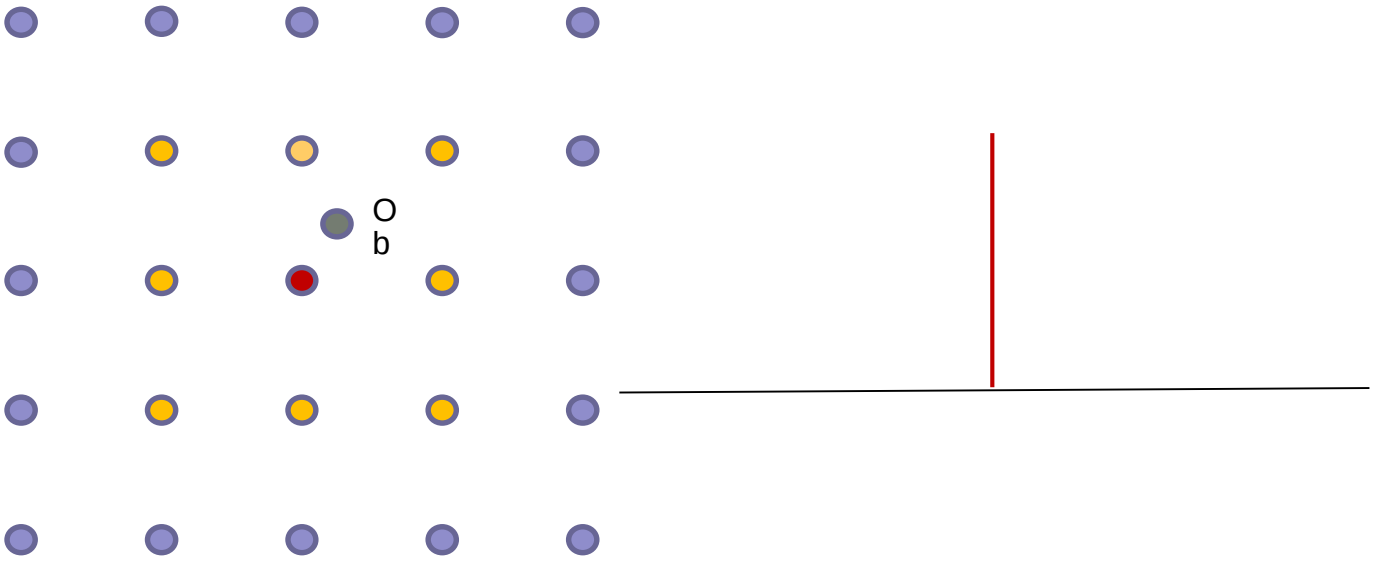
Neighbourhood sizes:

1*1, 3*3, 5*5, ...

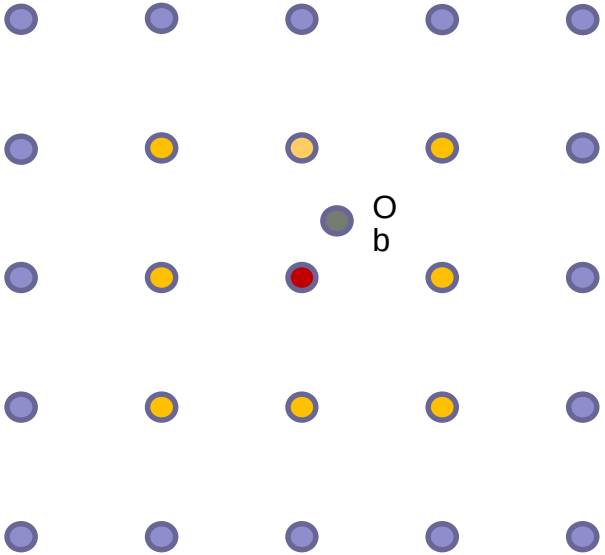
“Pseudo” ensemble members:

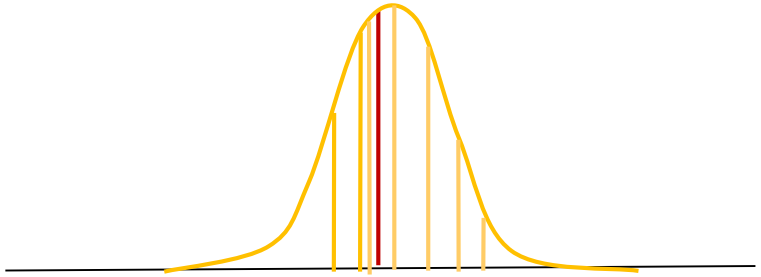
Number of the gridpoints (1, 9, 25, etc.)

Single-observation-neighbourhood-forecast approach (SO-NF)

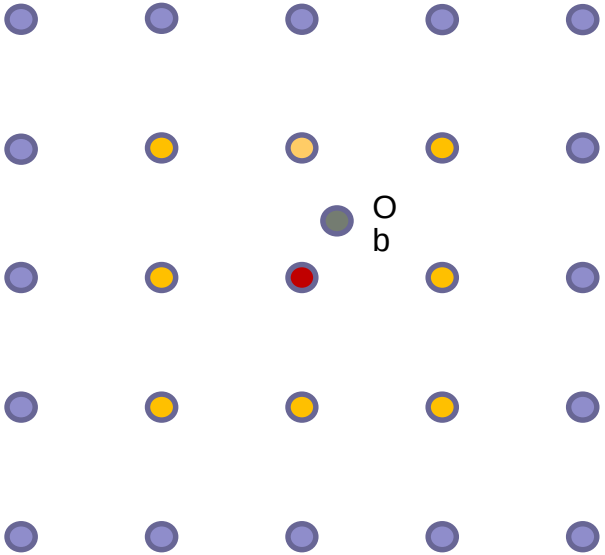


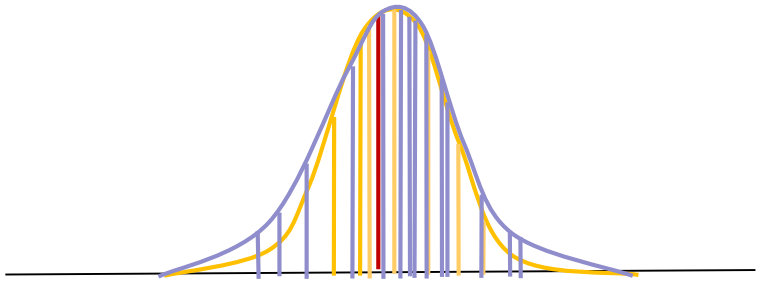
Single-observation-neighbourhood-forecast approach (SO-NF)



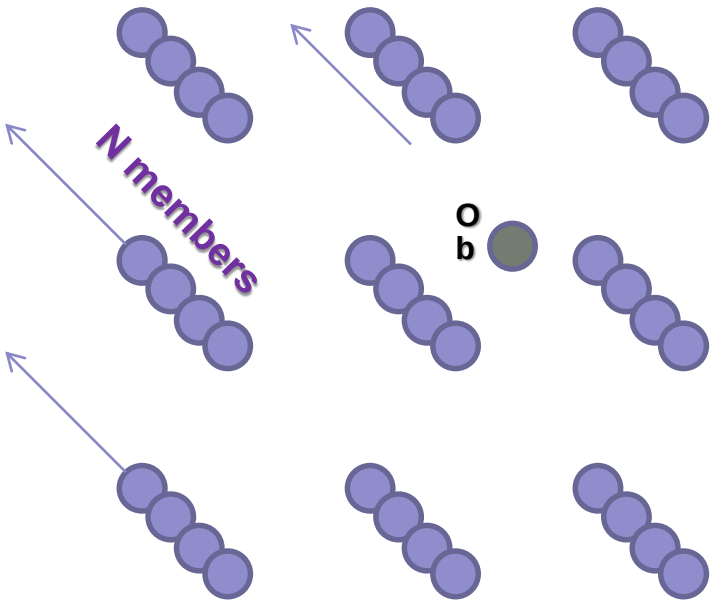


Single-observation-neighbourhood-forecast approach (SO-NF)





Single-observation-neighbourhood-forecast approach (SO-NF) for ensemble forecasts



Neighbourhood sizes:

1*1, 3*3, 5*5, ...

“Pseudo” ensemble members:

N * (number of gridpoints)
(N, 9*N, 25*N, etc.)

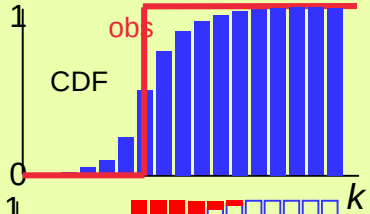
HiRA framework ...

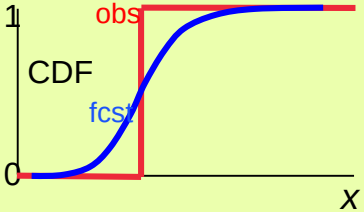
- is be able to compare
  - deterministic vs. deterministic at different resolutions
  - deterministic vs. deterministic at the same resolution in trials (physics)
  - deterministic vs. ensemble
  - ensemble vs. ensemble
- uses
  - standard synoptic observations and averaged daily forecasts
  - 24h **persisted obs**ervations as reference (for skill scores)
  - **Brier Score** (BS)
  - **(Continuous) Ranked Probability Score** (RPS; CRPS)
  - Associated skill scores (BSS, (C)RPSS)
- tests whether differences are statistically significant (Wilcoxon signed rank test)

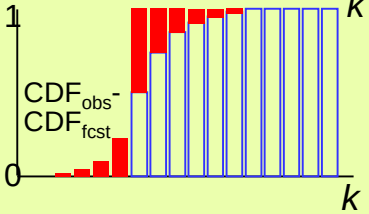
Model-specific vs. user-relevant

Probabilistic Forecast / Binary Observation

Modeller interested overall performance or pdf







$$CRPS = \int_{-\infty}^{\infty} \left(P_{fcst}(x) - P_{obs}(x)\right)^2 dx$$

Perfect score = 0

$$RPS = \frac{1}{K-1} \sum_{k=1}^{K} \left(CDF_{fcst,k} - CDF_{obs,k}\right)^2$$

K = discrete threshold categories

Perfect score = 0

Users interested in specific thresholds: e.g. defining a decision point or hazard

$$BS = \frac{1}{N} \sum_{i=1}^{N} (p_i - o_i)^2$$

p_i = forecast probability
o_i = observed occurrence (0 or 1)

Perfect score = 0

Model-specific vs. user-relevant thresholds

Variable	Thresholds
Wind speed (BF)	2.06 3.60 5.65 8.74 11.31 14.39 17.48 21.07 24.67 60 m/s
1h precipitation	0.25 0.5 1.0 2.0 4.0 8.0 16.0 32.0 64.0 mm/h
Total Cloud	0.0625 0.1875 0.3125 0.4375 0.5625 0.6875 0.8125 0.9375
Cloud Base Height	50. 100. 200. 500. 1000. 1500. 2000. 2500. 5000. m
Visibility	50. 100. 200. 500. 1000. 2000. 3000. 4000. 5000. 7500. m

Variable	Threshold
Temperature	0 ºC (-2, 2 and 25, 27, 30)
Wind speed	17.48 m/s (BF7)
1h precipitation	4 mm/h
Total Cloud	0.8125
Cloud Base Height	500 m
Visibility	1000 m

Neighbourhood sizes

`	UKV		MOGREPS-UK		
Neighbourh'd	km	member	km	member	Neighbourh'd
1 x 1	1.5	1	2.2	12	**1 x 1**
3 x 3	4.5	9	6.6	108	**3 x 3**
5 x 5	7.5	25	11.0	300	**5 x 5**
7 x 7	10.5	49	15.4	588	**7 x 7**
9 x 9	13.5	81	19.8	972	9 x 9
11 x 11	16.5	121	24.2	1452	**11 x 11**
13 x 13	19.5	169	-	-	-
15 x 15	22.5	225	-	-	-
17x 17	25.5	289	-	-	-

Equalise on area or ensemble size?

Equalise on areas

for deterministic vs. deterministic / ensemble ctrl

`	UKV		MOGREPS-UK		
Neighbourh'd	km	member	km	member	Neighbourh'd
1 x 1	1.5	1	2.2	12	**1 x 1**
3 x 3	**4.5**	9	**6.6**	108	**3 x 3**
5 x 5	7.5	25	**11.0**	300	**5 x 5**
7 x 7	**10.5**	49	**15.4**	588	**7 x 7**
9 x 9	13.5	81	**19.8**	972	9 x 9
11 x 11	**16.5**	121	**24.2**	1452	**11 x 11**
13 x 13	**19.5**	169	-	-	-
15 x 15	22.5	225	-	-	-
17x 17	**25.5**	289	-	-	-

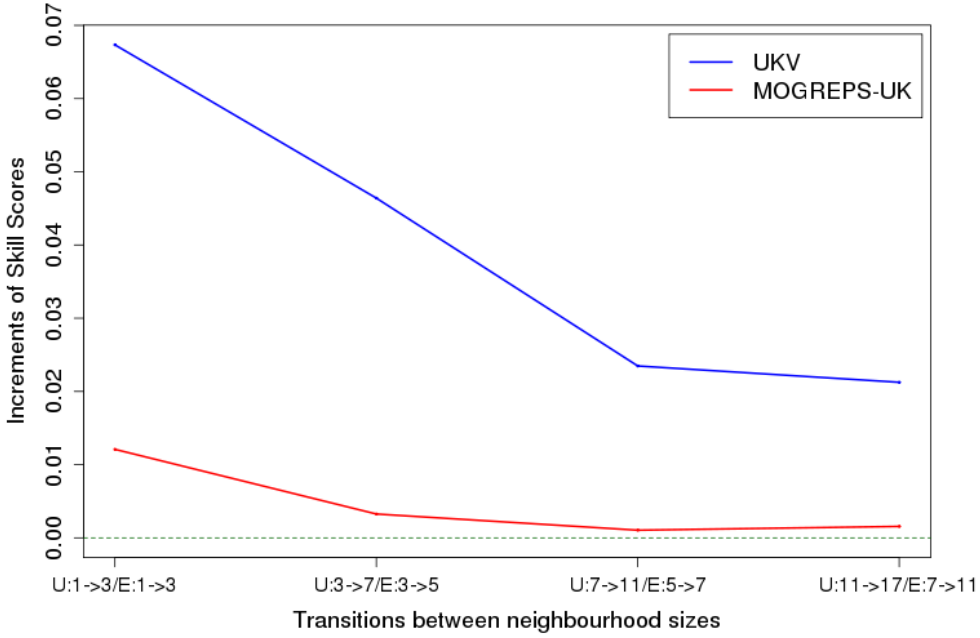
Equalise on ensemble members

for deterministic vs. ensemble

`	UKV		MOGREPS-UK		
Neighbourh'd	km	member	km	member	Neighbourh'd
1 x 1	1.5	1	2.2	**12**	**1 x 1**
3 x 3	4.5	**9**	6.6	**108**	**3 x 3**
5 x 5	7.5	25	11.0	**300**	**5 x 5**
7 x 7	10.5	49	15.4	588	**7 x 7**
9 x 9	13.5	81	19.8	972	9 x 9
11 x 11	16.5	**121**	24.2	1452	**11 x 11**
13 x 13	19.5	169	-	-	-
15 x 15	22.5	225	-	-	-
17x 17	25.5	**289**	-	-	-

Average increments in the function of the neighbourhood size

- Averaging on all lead times and all parameters
- UKV: large benefit at the first two increasing steps
- MOGREPS-UK: small positive benefit for the second (3*3) neighbourhood



Equalise on ensemble members

for deterministic vs. ensemble

`	UKV		MOGREPS-UK		
Neighbourh'd	km	member	km	member	Neighbourh'd
1 x 1	1.5	1	2.2	**12**	**1 x 1**
3 x 3	4.5	**9**	6.6	**108**	**3 x 3**
5 x 5	7.5	25	11.0	**300**	**5 x 5**
7 x 7	10.5	49	15.4	588	**7 x 7**
9 x 9	13.5	81	19.8	972	9 x 9
11 x 11	16.5	**121**	24.2	1452	**11 x 11**
13 x 13	19.5	169	-	-	-
15 x 15	22.5	225	-	-	-
17x 17	25.5	**289**	-	-	-

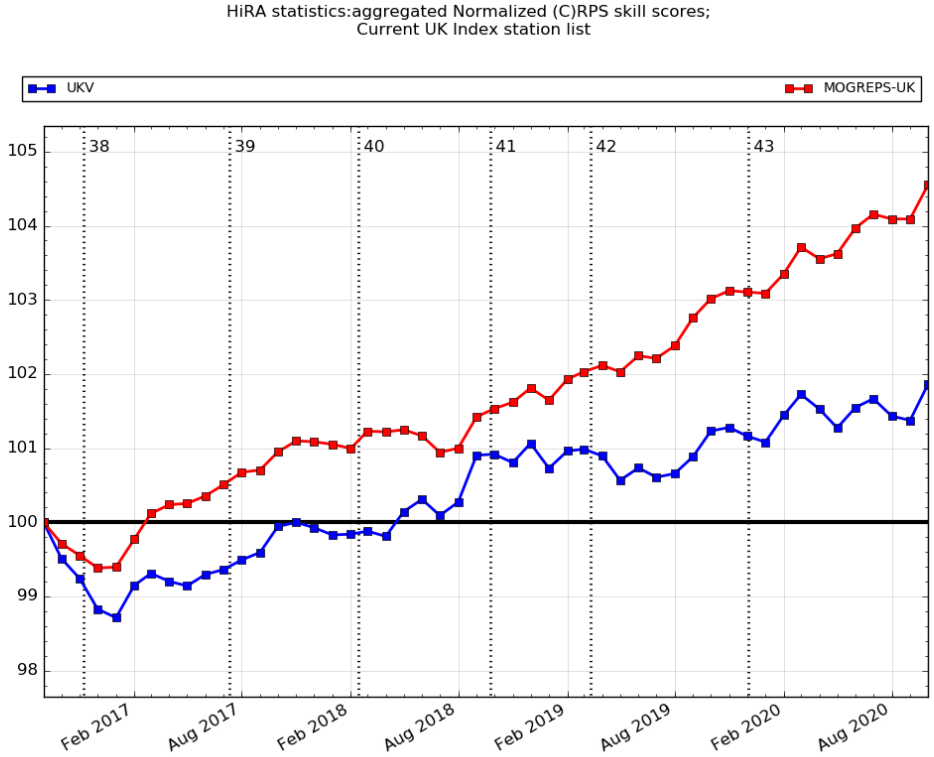
Chosen pair:

- **UKV – 11*11 neigbourhood size**
- **MOGRES-UK – 3*3 neigbourhood size**

`	UKV		MOGREPS-UK		
Neighbourh'd	km	member	km	member	Neighbourh'd
1 x 1	1.5	1	2.2	**12**	**1 x 1**
3 x 3	4.5	**9**	6.6	**108**	**3 x 3**
5 x 5	7.5	25	11.0	**300**	**5 x 5**
7 x 7	10.5	49	15.4	588	**7 x 7**
9 x 9	13.5	81	19.8	972	9 x 9
11 x 11	16.5	**121**	24.2	1452	**11 x 11**
13 x 13	19.5	169	-	-	-
15 x 15	22.5	225	-	-	-
17x 17	25.5	**289**	-	-	-

Jelen…

- Havi redszerességgel (quasi-operatív) fut a rendszer
- UKV vs. MOGREPS-UK illetve GM vs UKV
- Külön a 6 paraméterre ill. a paraméterek aggreglásából kapott értékre (csak skill score) + **normalizált értékek (48 hónapos futó átlag)**
- 2021 (2022?): ”Key Performance Indicator” (KPI) lesz, de csak a MOGREPS-UK HiRA Index



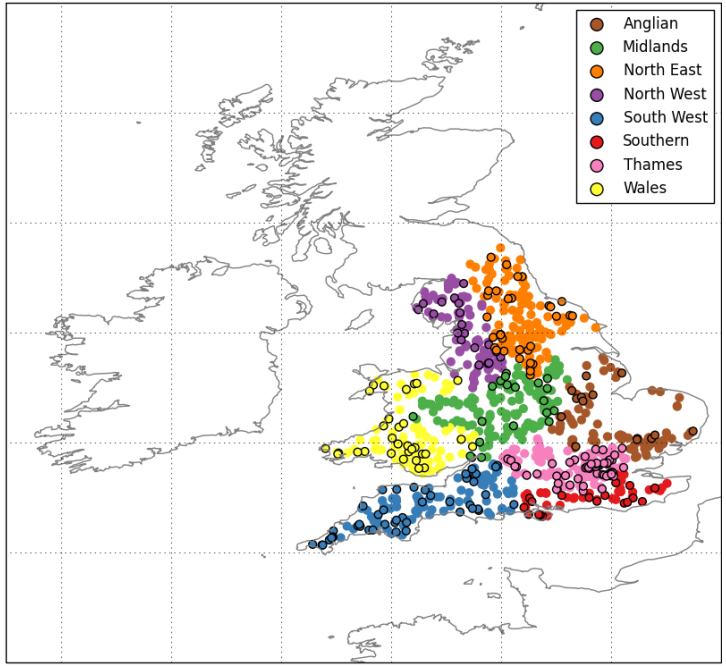
Megfigyelési adatok bizonytalanságának becslése

- Paraméter: vízgyűjtők órás csapadék-átlaga
- Terület: Anglia + Wales & Skócia
- Előrejelzés: első 24 óra órás előrejelzése
- Megfigyelés: csapadékmérő (gauge) & radar

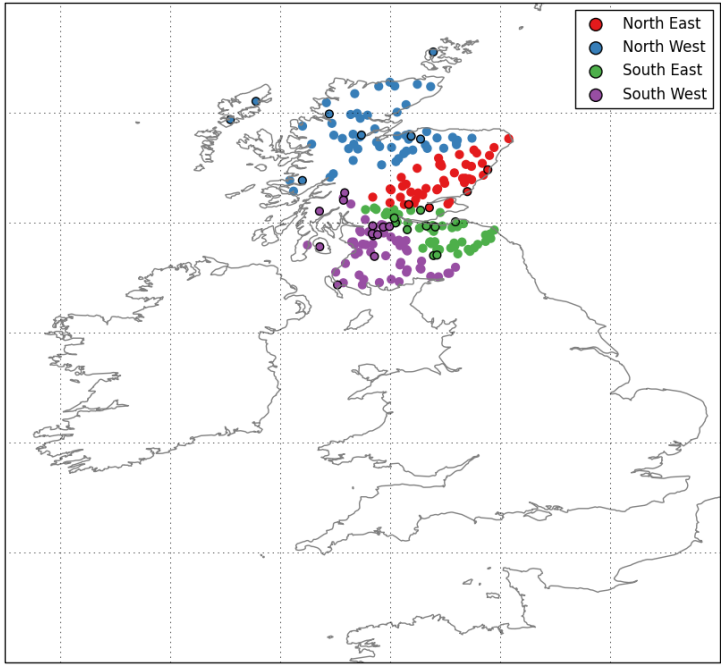
Catchments, Regions and Countries

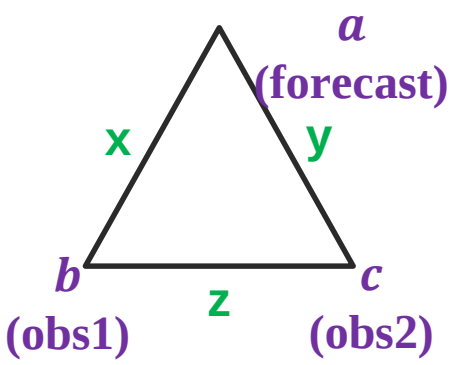
Used the *catchment averages* as the forecast and observation dataset

England & Wales (E&W)



Scotland





f_a : actual values of forecast
f_b : actual values of obs1
f_c : actual values of obs2
t : true values

x: MSE of the forecast & obs1
y: MSE of the forecast & obs2
z: MSE of the obs1 & obs2

$$x = E(f_a - f_b)^2 = E[(f_a - t) - (f_b - t)]^2 = E(\delta_a^2) + E(\delta_b^2) - \cancel{2E(\delta_a \delta_b)}$$

Assumption:
the two errors are not correlated

$$E(f_a - f_b)^2 = x = E(\delta_a^2) + E(\delta_b^2)$$

$$E(f_a - f_c)^2 = y = E(\delta_a^2) + E(\delta_c^2)$$

$$E(f_b - f_c)^2 = z = E(\delta_b^2) + E(\delta_c^2)$$

$$\Rightarrow$$

$$E(\delta_a^2) = \frac{x+y-z}{2}$$ **forecast error**

$$E(\delta_b^2) = \frac{x-y+z}{2}$$ **obs1 error**

$$E(\delta_c^2) = \frac{-x+y+z}{2}$$ **obs2 error**

Observation error estimation

- Determined an adjacent bin series:

  BIN1: **[0,0.1)**
  BIN2: **[0.1,1)**
  BIN3: **[1, 2)**
  BIN4: **[2, 4)**
  BIN5: **[4, ∞)**

Other binning strategy gave similar results

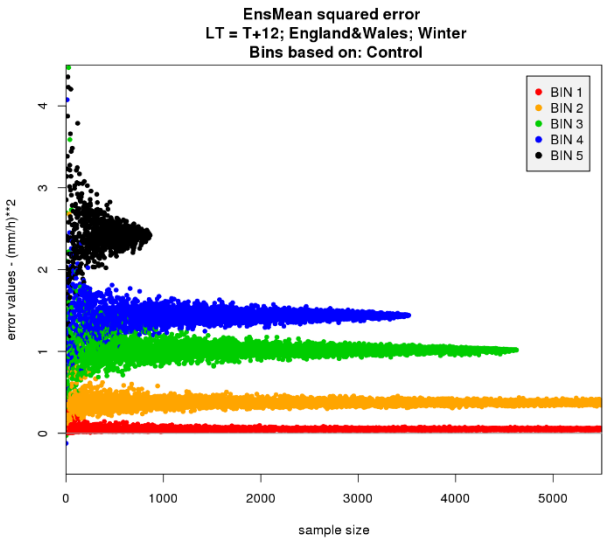
The distribution of the forecast (ensemble mean & control) and the observations based on the bins

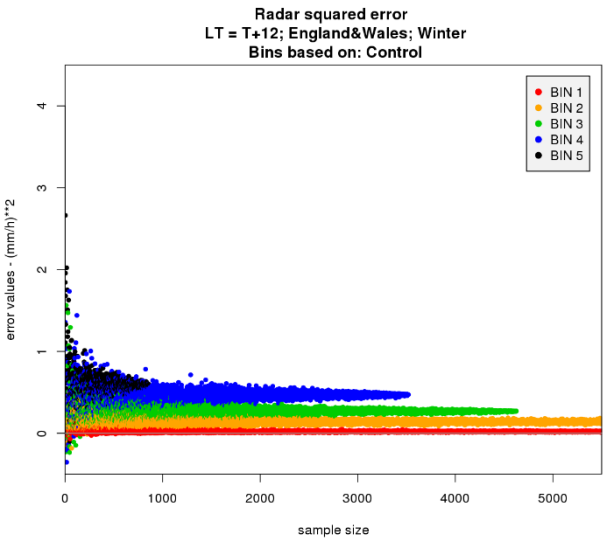
%	BIN 1	BIN 2	BIN 3	BIN 4	BIN 5
Ensemble mean	59.4	31.4	6.3	2.7	0.2
Radar	74.5	17.8	4.3	2.6	0.8
Gauge	71.1	21.4	4.3	2.5	0.7
Control forecast	72.1	17.7	5.5	3.7	1.0

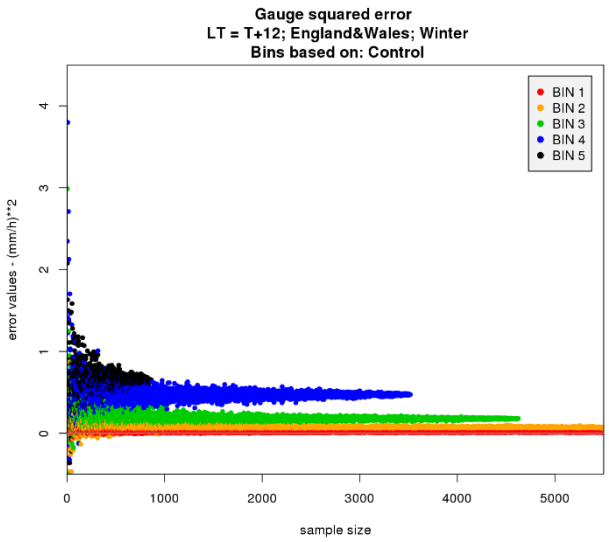
Observation error estimation

Squared error plots as a function of the sample size

- Binning based on the **Control**
- Forecast error were calculated from **Ensemble mean**



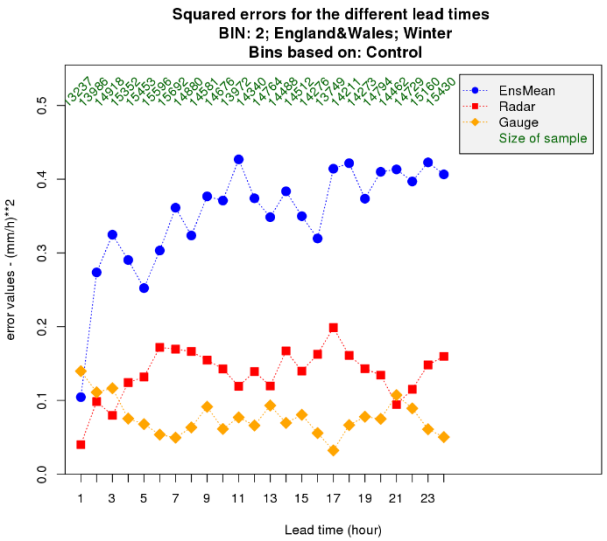


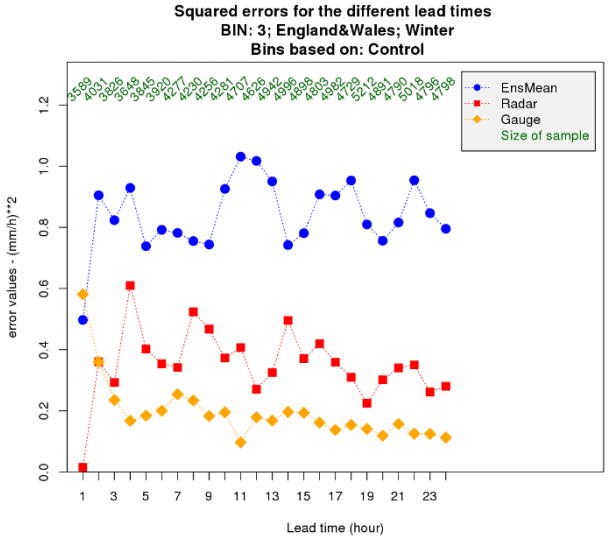


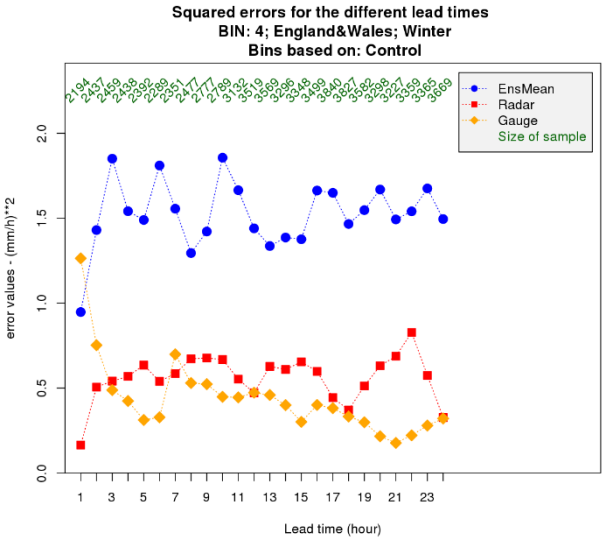
Observation error estimation

Squared error plots as a function of the lead time

- Observation errors are **not changing** with the lead time
- The forecast error is **increasing** with the lead time

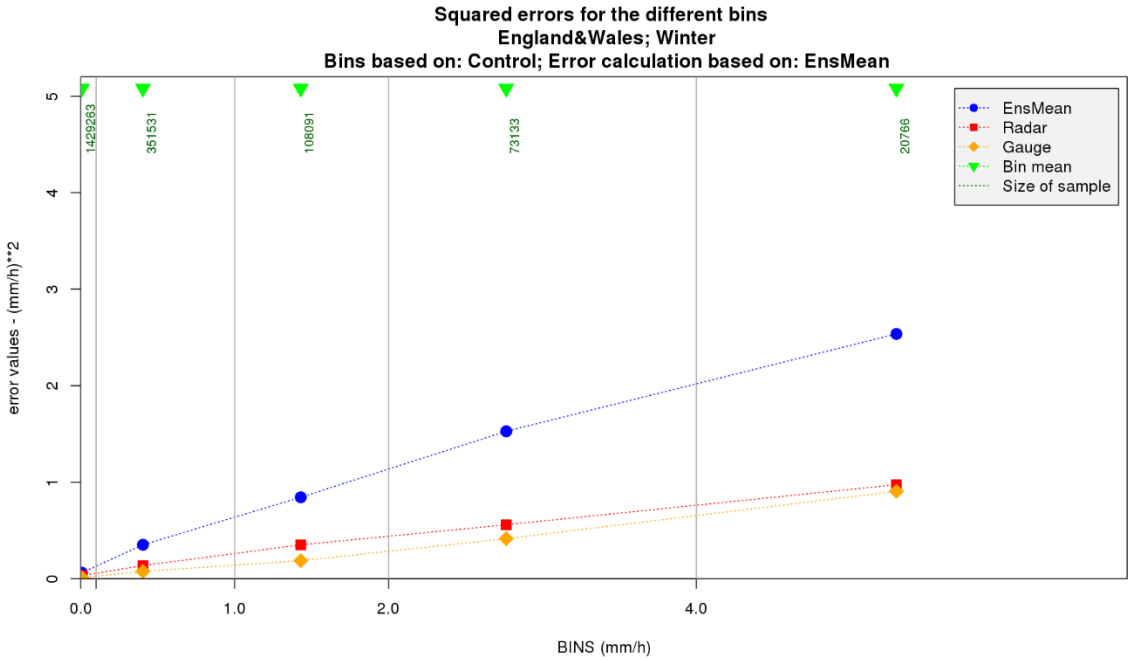






Observation error estimation

Squared error values as a function of the bin-mean values of the forecast

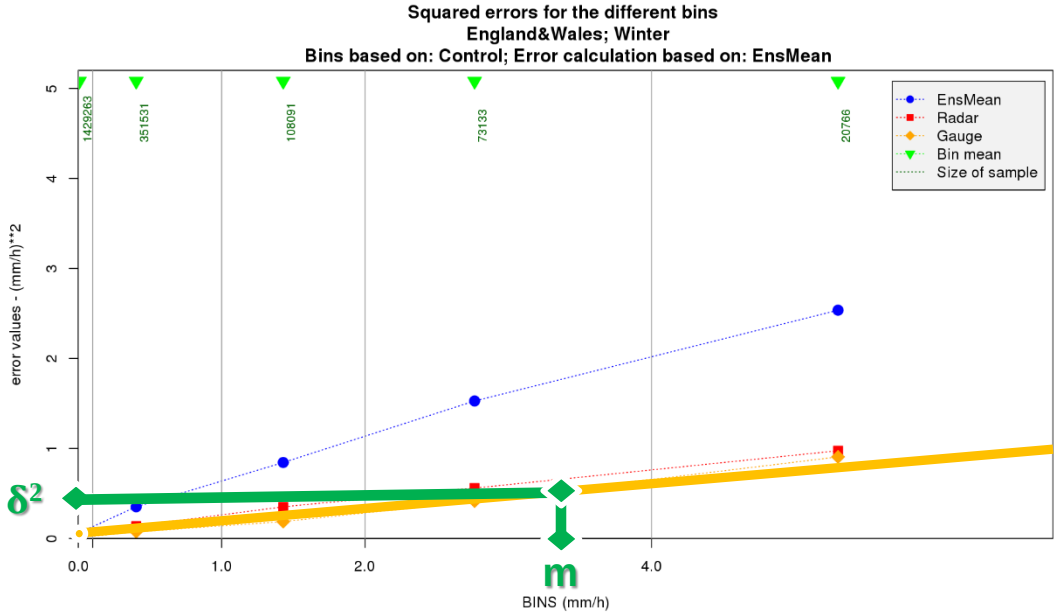


The functions are roughly log linear

=>

Can fit linear models to them

Using the estimated observation error in the verification



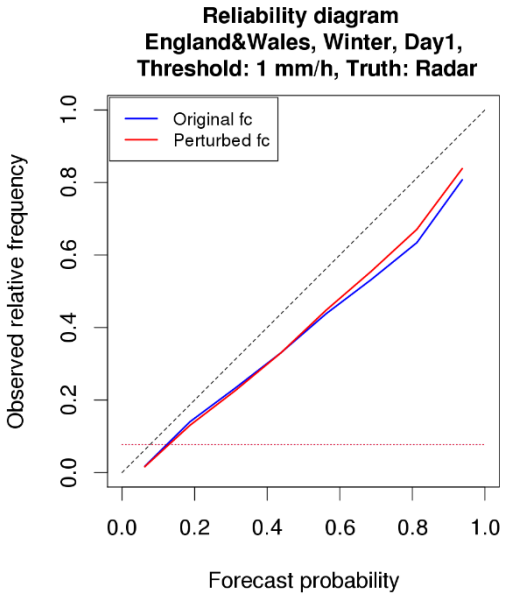
For each forecast case:

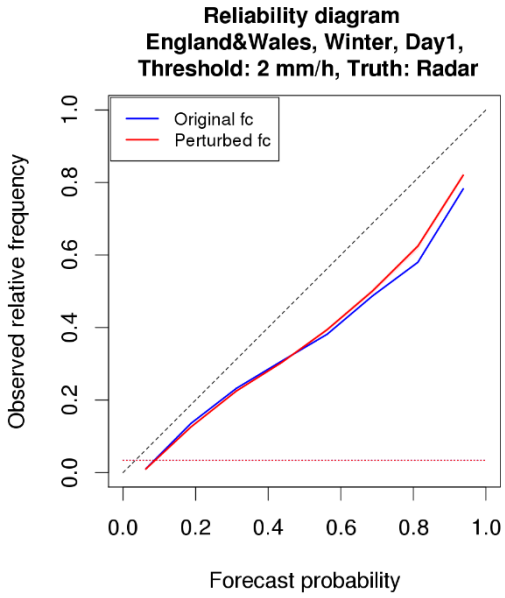
1. Based on the forecast magnitude (m) and the linear model => estimate the observation error (δ)
2. Random sample from N(0, δ^2)
3. Add the random sample to given forecast case magnitude

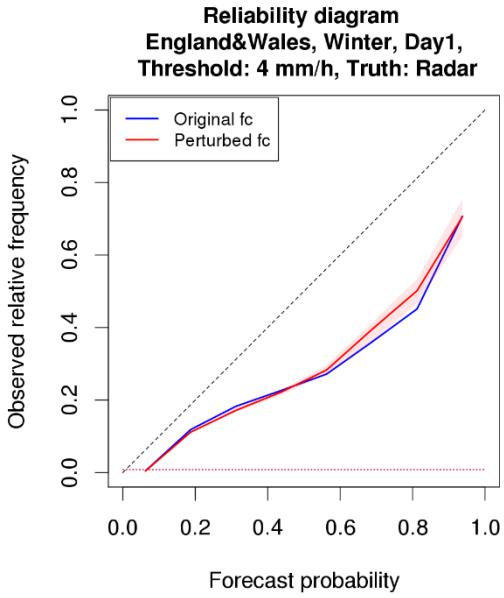
Using the estimated observation error in the verification

Reliability diagrams

Radar
Thresholds: 1, 2, 4 mm/h



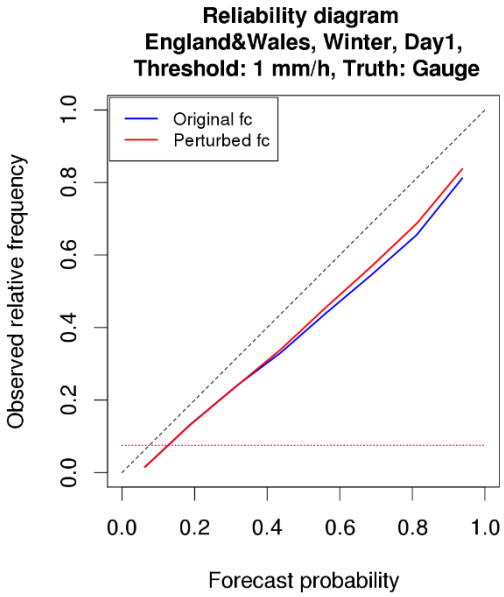


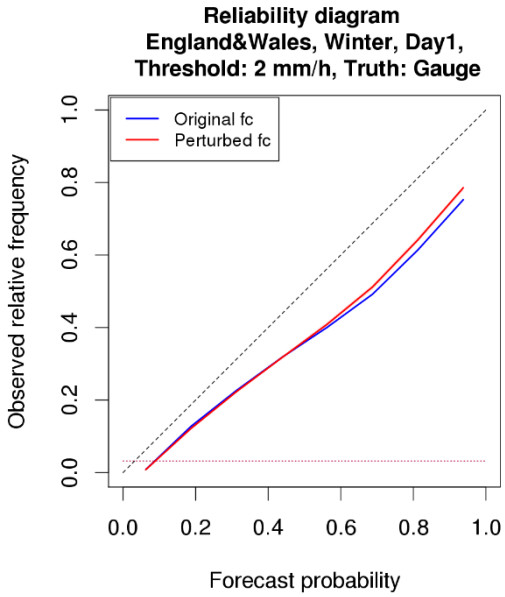


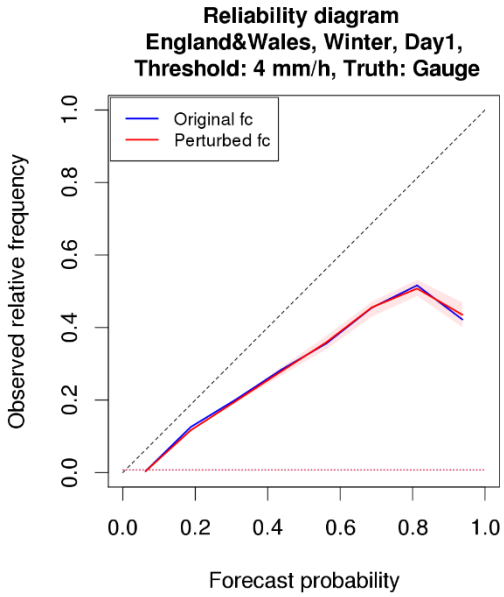
Using the estimated observation error in the verification

Reliability diagrams

Rain gauges
Thresholds: 1, 2, 4 mm/h







Köszönöm a figyelmet! Kérdés, hozzászólás?