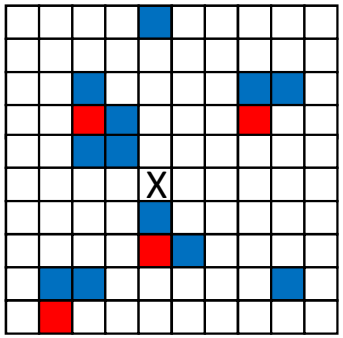
Idealised investigation of how a neighbourhood applied to a forecast, but not the observation can reward under-forecasting (using the RPS / Brier Score)

Neighbourhoods are used to alleviate the double penalty problem, but we could make it worse!

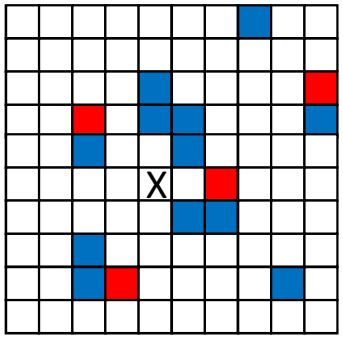
Application of a neighbourhood to the forecast, but not observation

Observed

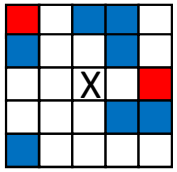


Observation

Forecast

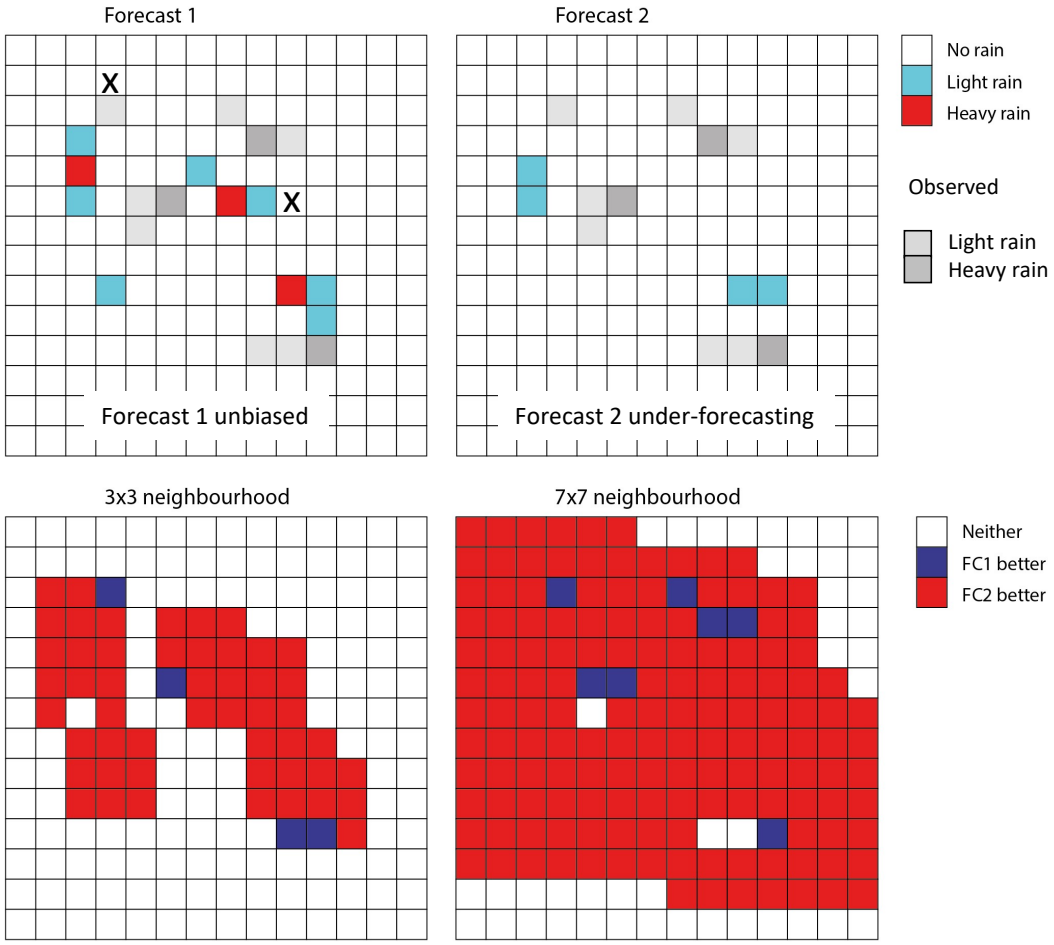


5x5 neighbourhood



Neighbourhood gives increased chance of rain categories in the forecast (at that location)
-> score gets worse! Is this general or just at that location?

Idealised situation using the RPS with observed points and forecast neighbourhoods



Forecast 1 unbiased

Forecast 2 under-forecasting (biased)

Similar spatial error

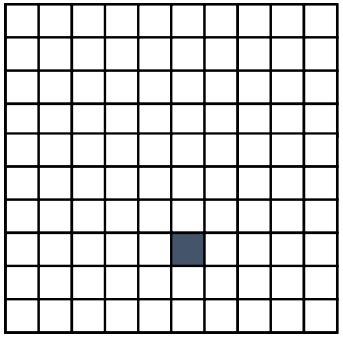
Both completely wrong at the grid scale (double penalty)

Relative RPS score for each grid square to account for all possible rain gauge locations

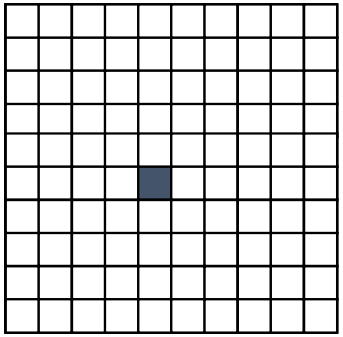
The biased forecast scores better in many more possible gauge locations (more red in 3x3 and 7x7)

The total RPS over all locations is better (lower) for the biased forecast
- More gauges (better sampling) won't help!

Idealised random rain pixel



Observed



Forecast

Examine an idealised random placement of 1 rain pixel in 100 pixels (possible gauge locations).

Have the randomly positioned rain pixel in both the observed and forecast grids (no bias).

Here we only have two categories, so RPS = Brier Score.

For any given forecast, the chances of the four possible outcomes, at a gauge location are:

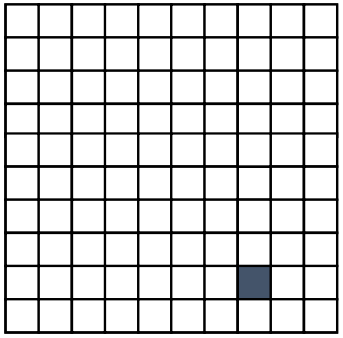
Chance of rain and rain	(RPS=0)	= 0.01 x 0.01	= 0.0001
Chance of rain and no rain	(RPS=1)	= 0.01 x 0.99	= 0.0099
Chance of no rain and rain	(RPS=1)	= 0.99 x 0.01	= 0.0099
Chance of no rain and no rain	(RPS=0)	= 0.99 x 0.99	= 0.9801

			1.0000

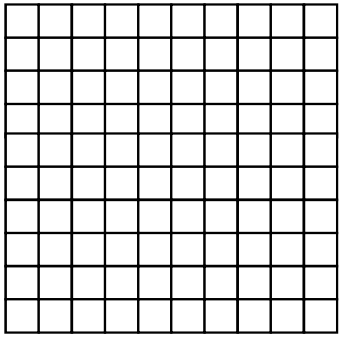
98.02% chance that RPS = 0 (perfect forecast)
1.98% chance that RPS > 0 (not perfect forecast)

Therefore, the expected total RPS over 100 events = **1.98**

RPS for a zero-rain forecast and observed field with 1% coverage



Observed



Forecast

Now consider a forecast system that never forecasts any rain

For any given forecast the chances of the possible outcomes are:

Chance of rain and rain	(RPS=0)		= 0.00
Chance of rain and no rain	(RPS=1)	= 0.01 x 1.00	= 0.01
Chance of no rain and rain	(RPS=1)		= 0.00
Chance of no rain and no rain	(RPS=0)	= 0.99 x 1.00	= 0.99

			1.00

99.0% chance that RPS = 0
1.0% chance that RPS > 0

Expected total RPS over all permutations = **1.00**

Compare with **1.98** for an unbiased forecast (lower value is better skill)

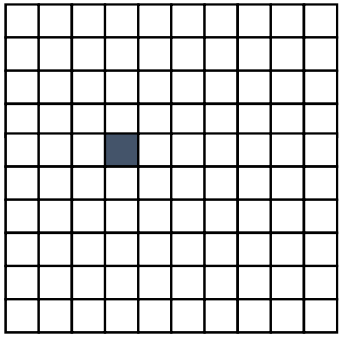
Forecasting nothing means less chance of a wrong forecast and improves skill (double penalty)

Met Office

Now apply a 3x3 neighbourhood to the forecast rain

This makes no difference to the forecast with zero rain

Apply a 3x3 neighbourhood to the forecast with rain



A neighbourhood is not applied to the observed field because we only know the value at the square being sampled (where the rain guage is) and can't construct a neighbourhood

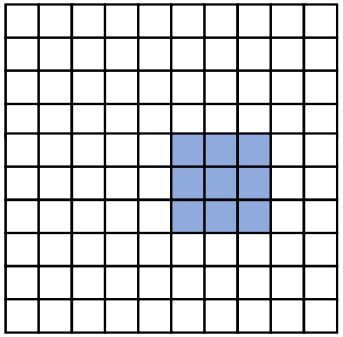
For any given forecast the chances of the four possible outcomes are:

Chance of rain and rain	(RPS=64/81)	= 0.01 x 0.09	= 0.0009
Chance of rain and no rain	(RPS=1)	= 0.01 x 0.91	= 0.0091
Chance of no rain and rain	(RPS=1/81)	= 0.99 x 0.09	= 0.0891
Chance of no rain and no rain	(RPS=0)	= 0.99 x 0.91	= 0.9009

			1.0000

90.09% chance RPS = 0

9.91% chance RPS > 0



Nine forecast squares can have probability > 0 (probability = 1/9)

Chance of rain and rain	RPS=64/81	x 0.0009	= 0.00071
Chance of rain and no rain	RPS=1	x 0.0091	= 0.00910
Chance of no rain and rain	RPS=1/81	x 0.0891	= 0.00110
Chance of no rain and no rain	RPS=0	x 0.9801	= 0.00000

			0.01091

Total RPS = **1.00**

Overall findings

Zero rain forecasts	Rain no neighbourhood	3x3 neighbourhood	5x5 neighbourhood
1.0% chance that RPS > 0	**1.98%** chance that RPS > 0	**9.91%** chance that RPS > 0	**25.75%** chance that RPS > 0

The use of a neighbourhood greatly increases the chance of a forecast with rain scoring worse than a no-rain forecast

Zero rain forecasts	Rain no neighbourhood	3x3 neighbourhood	5x5 neighbourhood
Total RPS = **1.0**	Total RPS = **1.98**	Total RPS = **1.091**	Total RPS = **1.02**

The use of a neighbourhood scores worse on average than a no-rain forecast, but less than with no neighbourhood

Overall – the use of a neighbourhood means that favouring under-forecasting is more likely

Final comments

Idealised scenarios suggest the use of a neighbourhood can favour under-forecasting (using RPS / Brier Score)

Other idealised configurations show the same (if coverage < 50%)

Worse if rain coverage is small or the neighbourhood does not span the spatial error

An ensemble will have the same effect as a neighbourhood (because it increases forecast coverage)

These are idealised studies and further investigation using more realistic or real cases is needed to confirm whether there is an issue