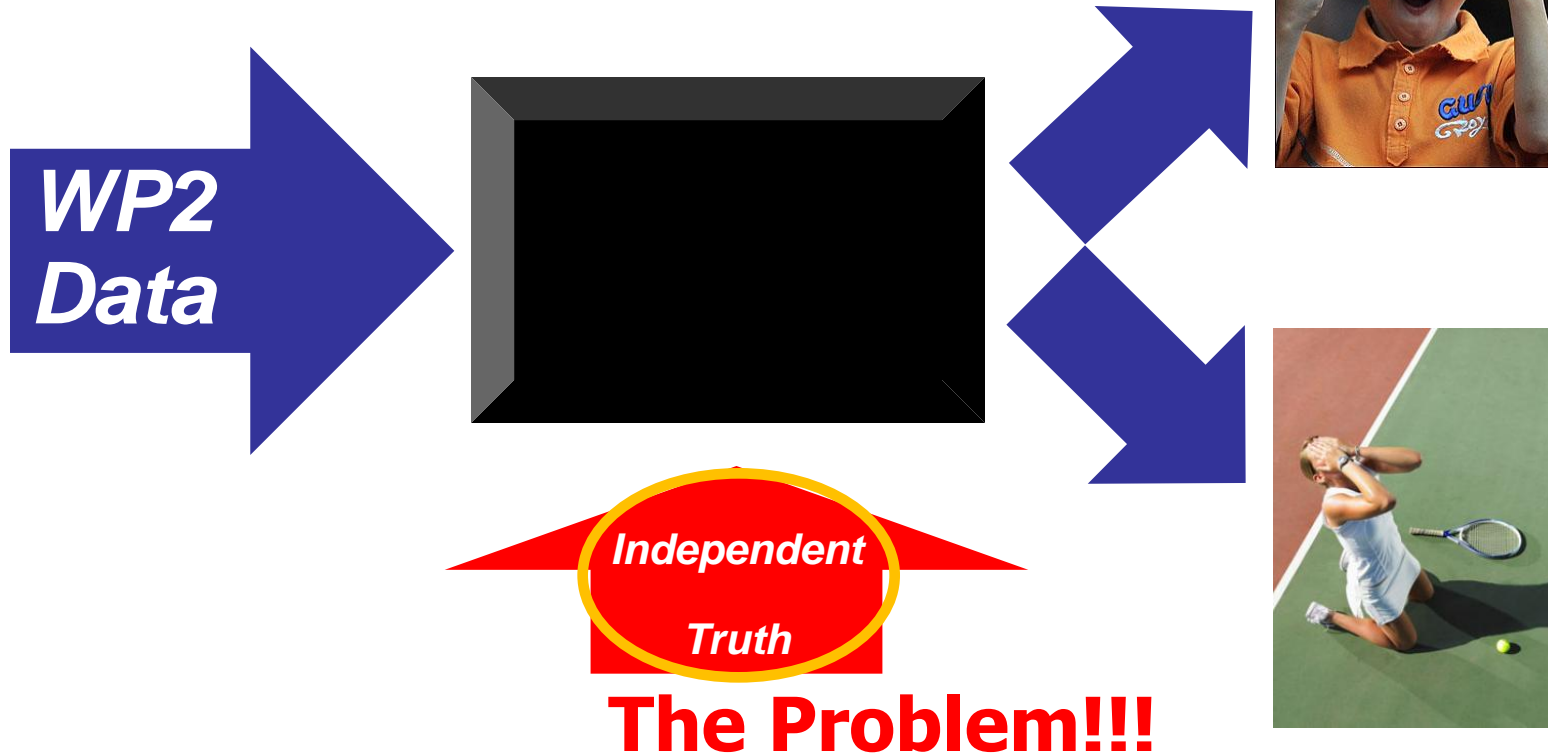


Validation & Independent Obs.

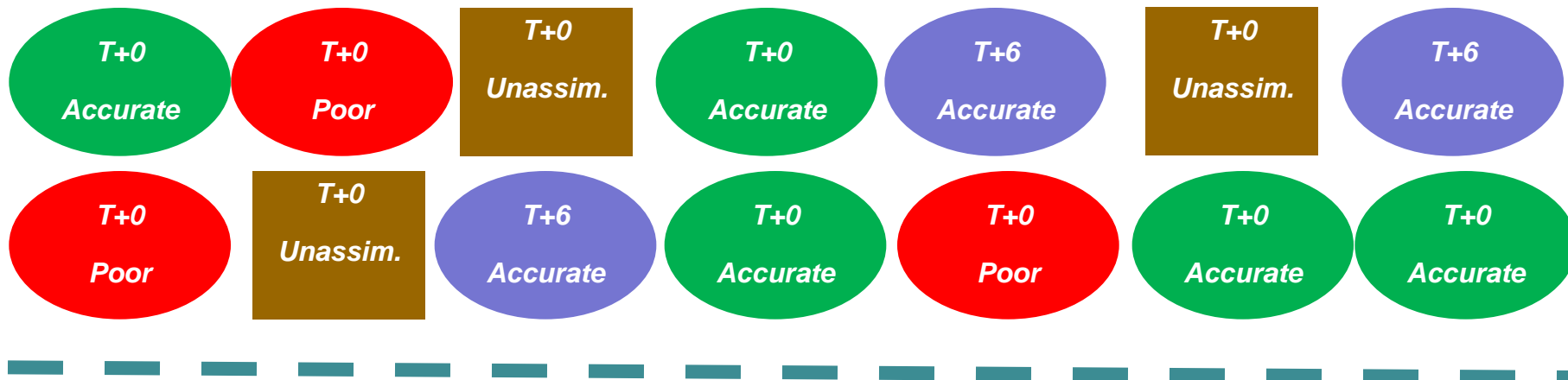
A WP2 perspective

Peter Jermey

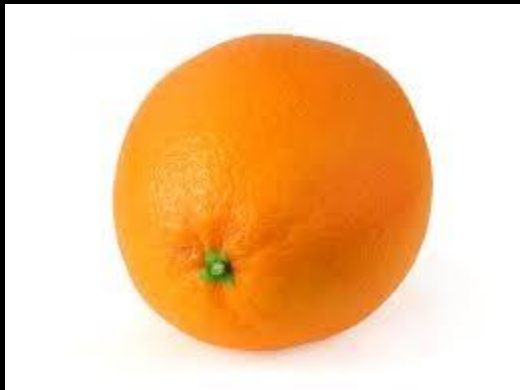
What we want...



Independent Truth



Deterministic reanalysis



Ensemble of reanalyses





Met Office

Validation (Deterministic)

- Lots of possible measures: RMSE, skill, ETS, Pierce, FSS, SEEPS etc
- All of these attempt to estimate the distance between the truth and our reanalysis

$$\epsilon^* = x - x^*$$



Met Office











O-A's, O-B's and Forecast Error



Forecast Model

DA

Forecast Model

O-A: Obs – analysis: comparing  with 
O-B: Obs – background: comparing  with 
Forecast Error: comparing  with   
or  with 

Not Independent Truth

Independent Truth

NB O-B's are a special case of forecast error



Met Office

T+6

Accurate

$$\epsilon^* = x - x^*$$

Time Shifted 'Truth'



- Straight-forward
- Assimilated observation types
- Useful for comparing different models
- Gives an upper bound on the reanalysis error
- Difficult to estimate actual reanalysis error
- Includes model error (and obs error)

$$M_t(x) - y_t \approx Q_t + x_t - x_t^* + \epsilon_y$$





Met Office

Unassimilatable Obs.

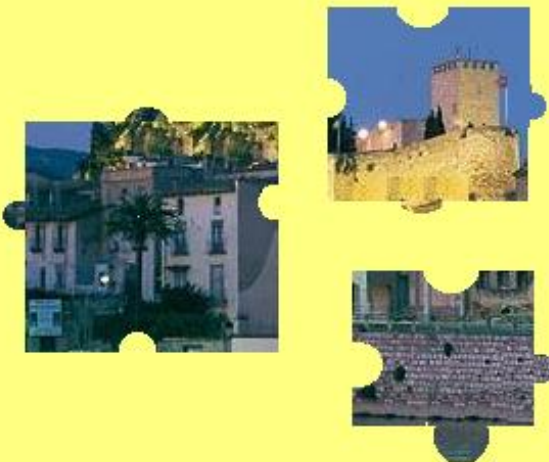
$T+0$

Unassim.

$$\epsilon^* = x - x^*$$

- Could be for technical reasons or downstream/irrelevant reasons
 - **Limited** observation types/positions
 - Useful for estimating error in reanalysis
 - Includes obs error
 - Downstream operator may contain errors

$$H(x) - y \approx R + x - x^* + \epsilon_y$$





Met Office

Poor Quality Obs.

- Observation error dominates ... No confidence in results!!

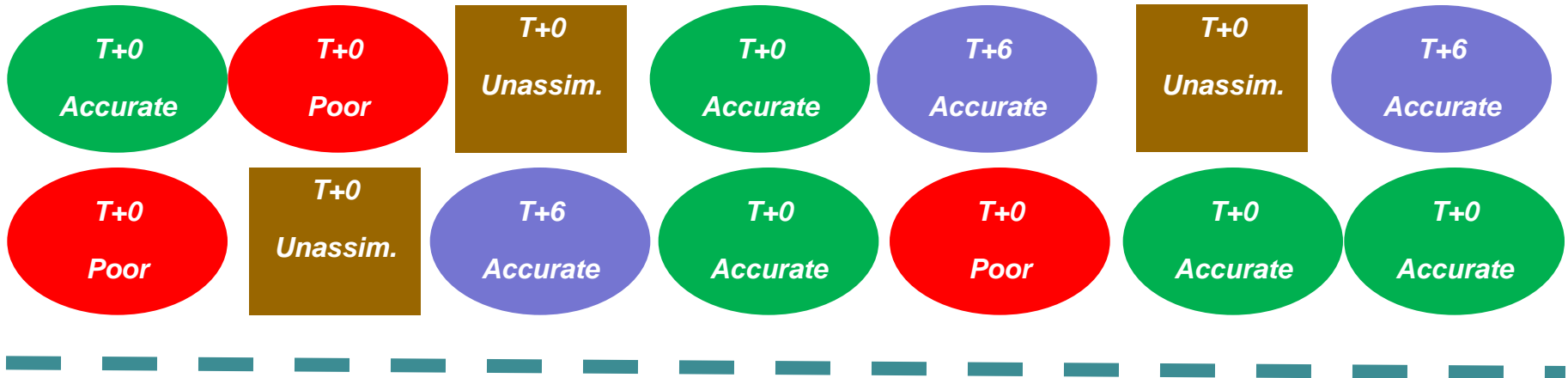
T+0

Poor

$$\epsilon^* = x - x^*$$



Degrade Reanalysis...



Deterministic reanalysis



Ensemble of reanalyses



Ensemble Primer



Ensembles represent uncertainty

We want the ensemble to represent the span of all possible reanalyses fixing everything except the random errors

$$\text{RMSE}(\bar{x}_j) \approx \text{Std.Dev.}(x_{ij})$$

What does a good ensemble look like..?



1. Each member is equally likely
2. Mean error < Control error
3. Ensemble spread = **Mean** error
4. Modelled freq. = Observed freq.



Met Office

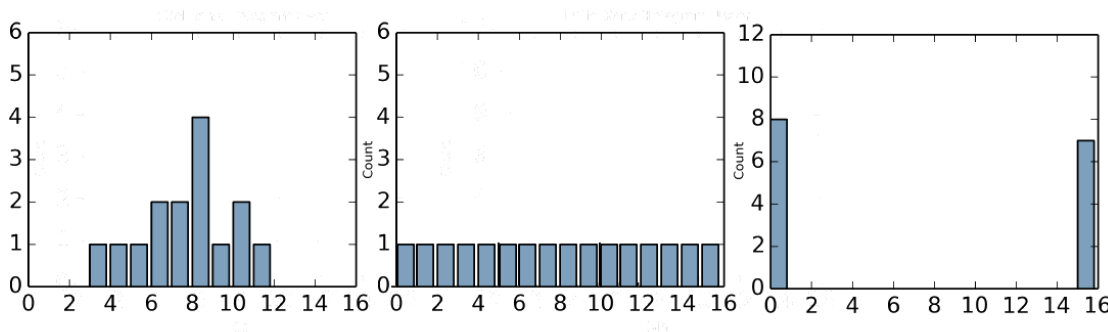
Validating Ensembles



1. Each member is equally likely

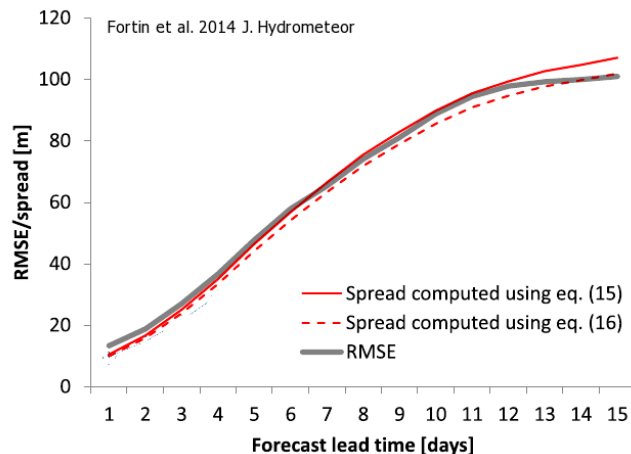
2. Mean error < Control error

- Rank Histogram



3. Ensemble spread = Mean error

- Plot ensemble RMSE and spread



4. Modelled freq. = Observed freq.

$$BSS = 1/N \sum (f_t - o_t)^2$$

$$IGN = -\log_2(p_e((ob)))$$

$$RPS = 1/(K - 1) \sum (CDF_{obs} - CDF_{en})^2$$



Met Office

Truth for Ensembles

- Particularly spread vs error

T+6

Accurate

- Spread increases with forecast time (but could look at trend!)

T+0

Unassim.

- Limited variables, downstream modelling issues (as before)

T+0

Accurate

- Not independent unless held back from assimilation

- **Should do a bit of everything!!**



Gridded Truth

Gridded Truth

- For some (deterministic) scores we require gridded truth
- Better than points for comparing like-with-like
- Problem! If available at higher quality than the reanalysis (truth) then ...
why are we doing the reanalysis?!
- Might be available in certain variables/scales
- Some datasets are based on high res obs that are not directly available

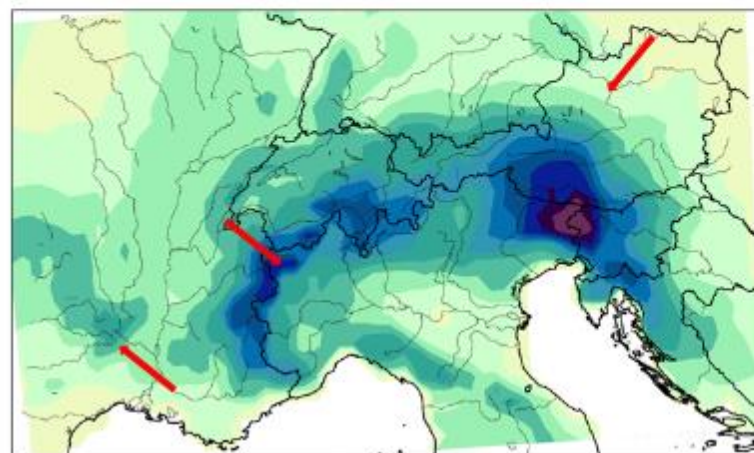
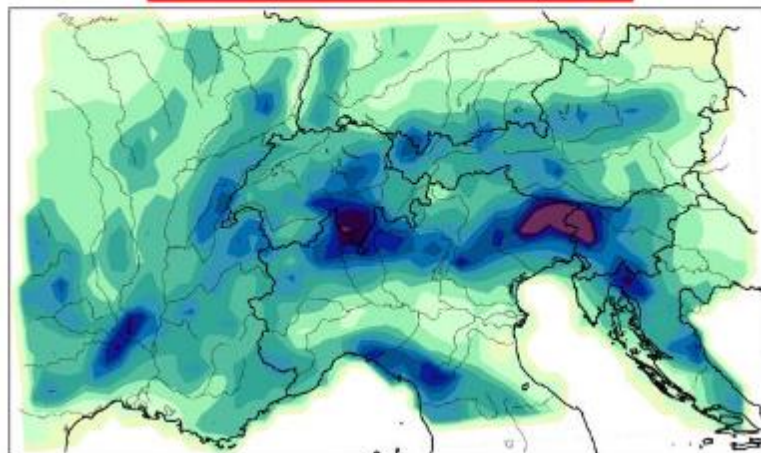


Mean annual precipitation 2008



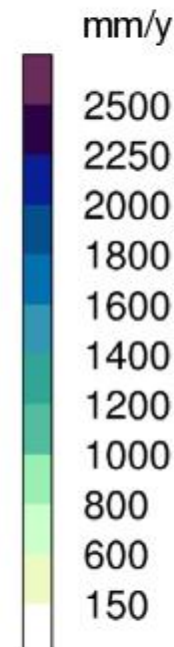
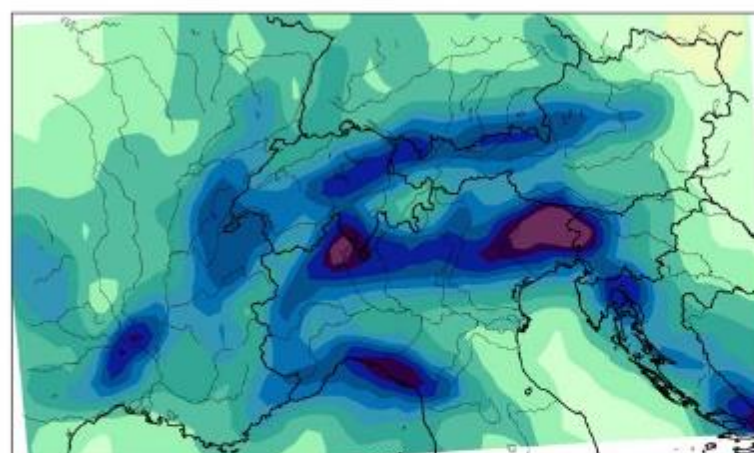
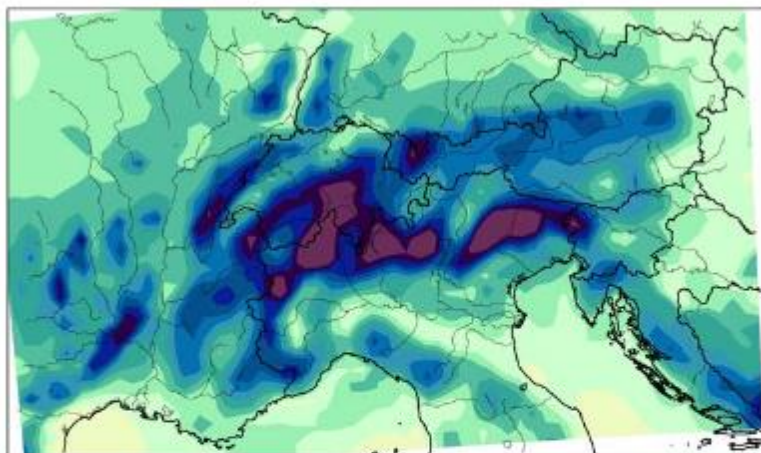
APGD (upscaled ~22 km)

E-OBS



HIRLAM

UK Met Office (upscaled ~22 km)



Summary

- As producers we want to have a measure of accuracy for our product
- We also need a measure of how well our ensemble estimates uncertainty
- Both of these require a measure of truth independent from the reanalyses
- Any 'truth' has associated problems ... So it is best to use lots of different truths

Thank you for listening...

...Questions?

Ensemble Primer



We use ensembles to predict uncertainty

We want the ensemble to represent the span of all possible reanalyses fixing everything except the random errors

We want the ensemble to spread to be a measure of uncertainty in our deterministic/mean reanalysis.

$$\text{MSE}(\bar{x}_j) = \frac{1}{J} \sum_{j=1}^J (\bar{x}_j - y_j)^2$$

$$\text{MSE}(\bar{x}_j) \approx \text{Var}(x_{ij})$$

$$\text{RMSE}(\bar{x}_j) \approx \text{Std.Dev.}(x_{ij})$$

$$\frac{1}{J} \sum_{j=1}^J (\bar{x}_j - y_j)^2 \approx \frac{1}{J} \sum_{j=1}^J (\bar{x}_j - x_{ij})^2$$

$$\frac{1}{I} \sum_{i=1}^I \frac{1}{J} \sum_{j=1}^J (\bar{x}_j - y_j)^2 \approx \frac{1}{I} \sum_{i=1}^I \frac{1}{J} \sum_{j=1}^J (\bar{x}_j - x_{ij})^2$$