UERRA Regional Reanalysis:
System Design

Peter Jermey

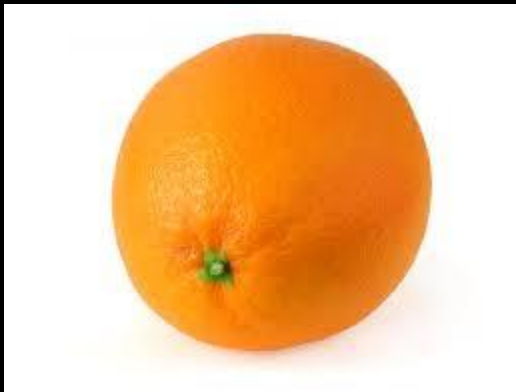Richard Renshaw, Jemma Davie, Dale Barker

Requirements

*Development and production of a satellite-era (**1978-present**) high resolution European **ensemble regional reanalysis** dataset, based on **ensemble-variational data assimilation**.*

UERRA   D.O.W.

- EDA based deterministic reanalysis
    - 12km CORDEX-EU grid

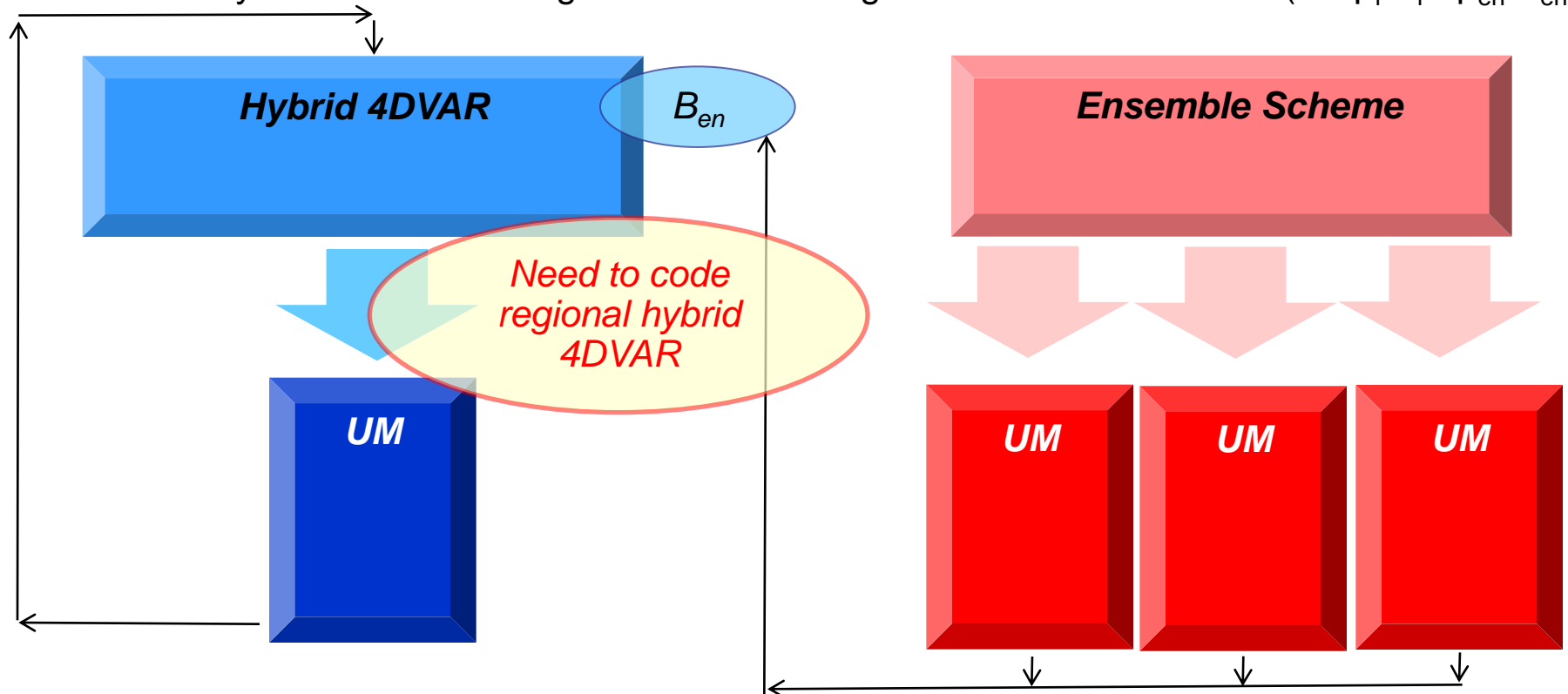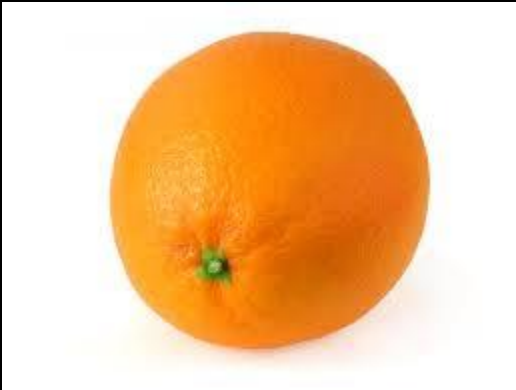- Ensemble of reanalyses
    - 20 members

System Overview

- EURO4M - MO is a 12km European deterministic regional reanalysis

- 4DVAR assimilation scheme at 24km – fixed background error covariances ($B = B_f$)

- UERRA – MO will include a 12km European deterministic regional reanalysis

- EDA – "hybrid" 4DVAR – weighted sum of background error covariances ($B = \beta_f B_f + \beta_{en} B_{en}$)
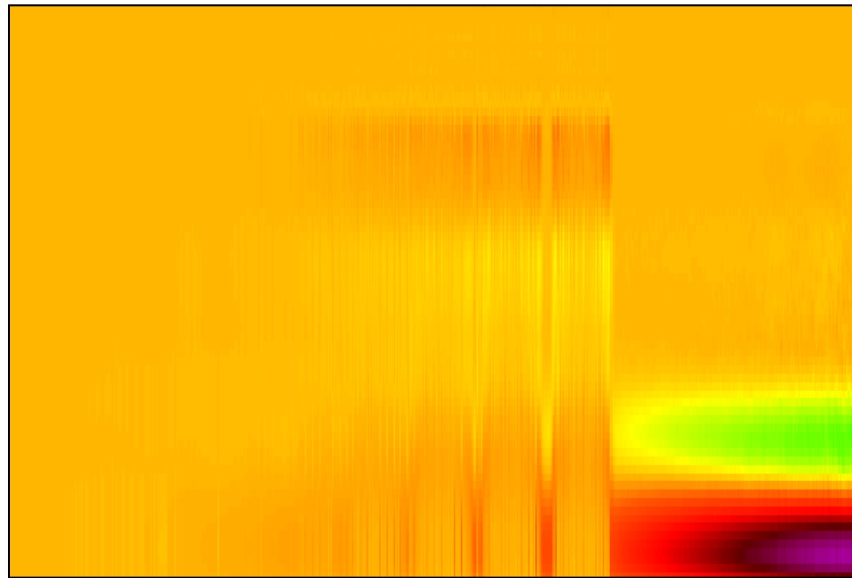
Hybrid 4DVAR

B_{en}

Ensemble Scheme

Need to code regional hybrid 4DVAR

UM

UM   **UM**   **UM**

Deterministic System

Estimate state of atmosphere using obs & model

Background Error Cov

- "fixed" B_f will be updated on a regular basis

- B_f is a smooth idealised model with parameters tuned by training data
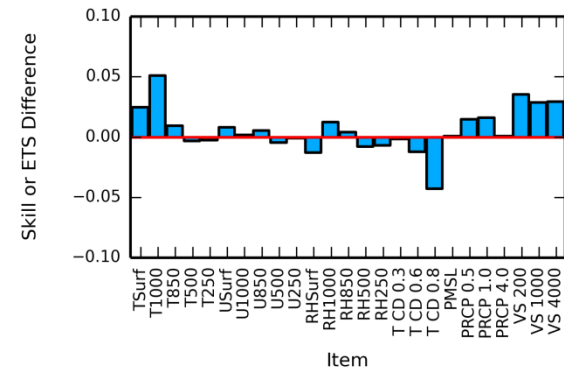
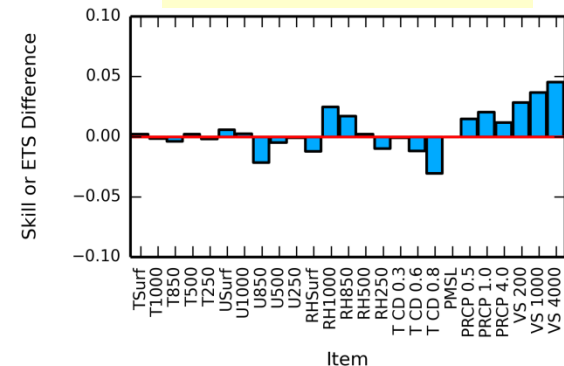- B_{en} from ensemble – "Errors of the day"

Forecast Model: ENDGAME

- New dynamical core

- Improved

  coupling with parameterisations

  handling of Rossby waves

  stability

  orographic effects

  mass conservation

- **Improved precip. & visibility**

- Trials at 30km (2.5 x reanalysis grid)

Winter T+12

Summer T+12

Ensemble System

Estimate uncertainty in det. system
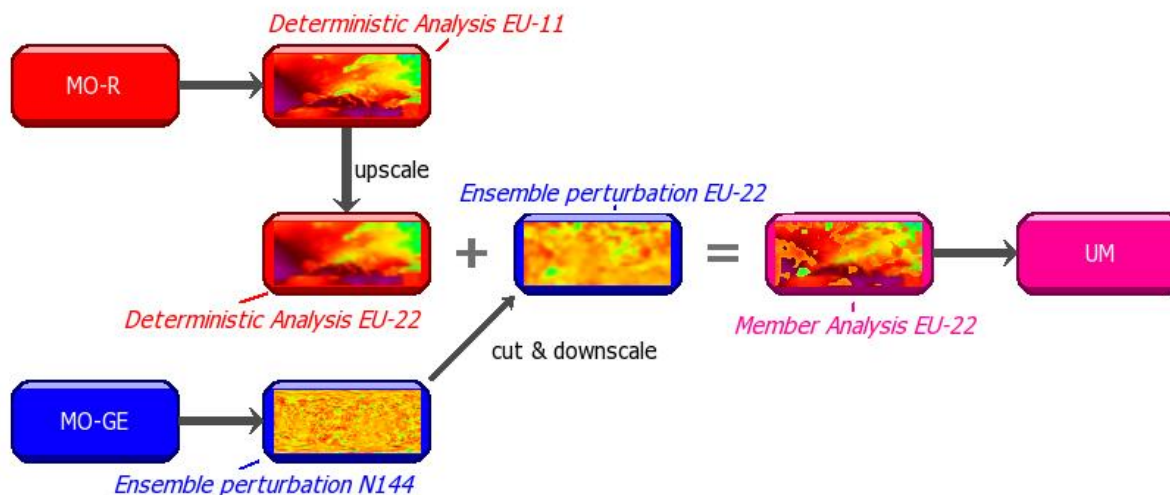
Ensemble Candidate: ETKF

- Operational since 2005, ensemble (MOGREPS) is based on ETKF

- Requires inflation

- Known to be under-spread even after inflation

- Regional version **requires** global version

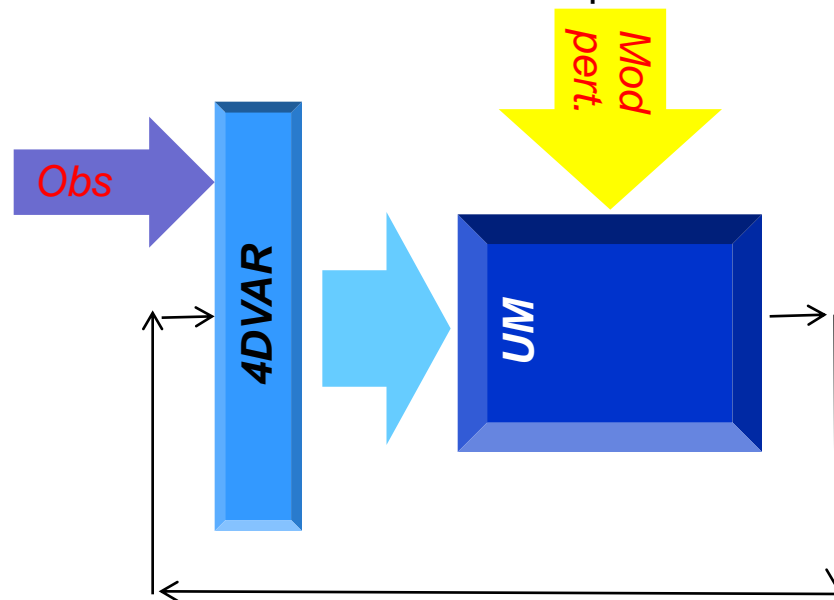- Also requires two-way coupling with deterministic system

Ensemble Candidate: En4DVAR

- Ensemble of deterministic system can be achieved by perturbing obs and model
- En4DVAR has been previously used to produce estimates of model error
- A similar system is used for operational EDA at ECMWF
- Closely follows det. system, i.e. Spread should be closer to det. uncertainty.
- Regional version **does not require** global version
- Also requires archive of deterministic analysis increments
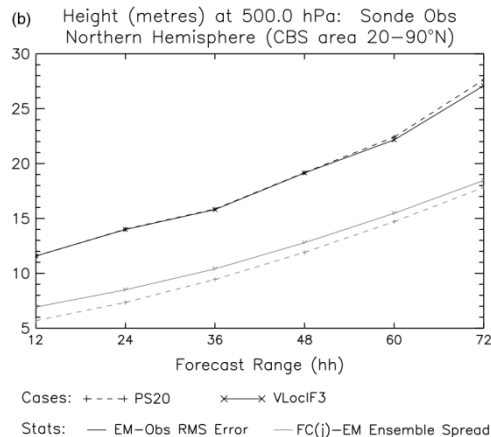- Each member runs a 4DVAR – much more expensive than ETKF

Obs

Mod pert.

4DVAR

UM

- En4DVAR is technically more flexible than ETKF

Candidate Comparison

- NH 500hPa at T+6:   **spread(En4DVAR) ≈ RMSE**   **spread(ETKF) ≈ 0.5RMSE**
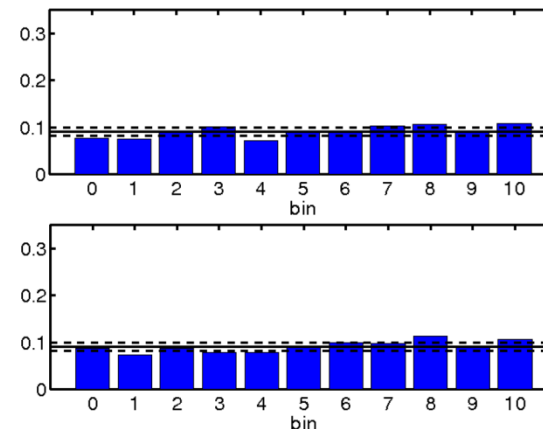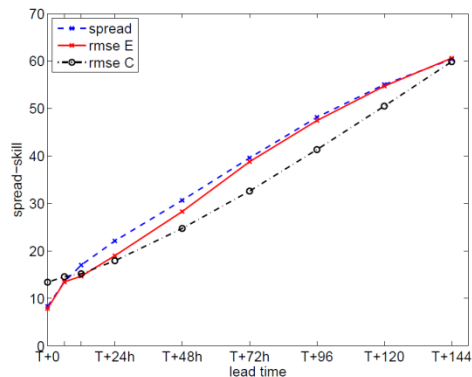
Spread Vs RMSE(mean)

Rank Histogram

ETKF

En4DVAR

En4DVAR

- Perturb...
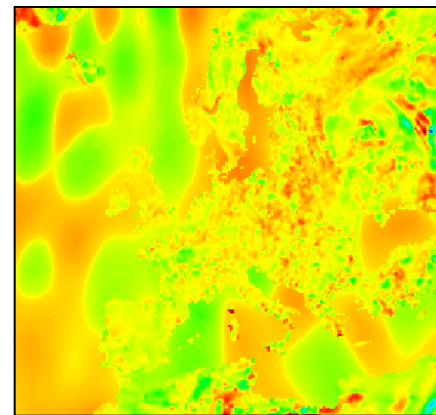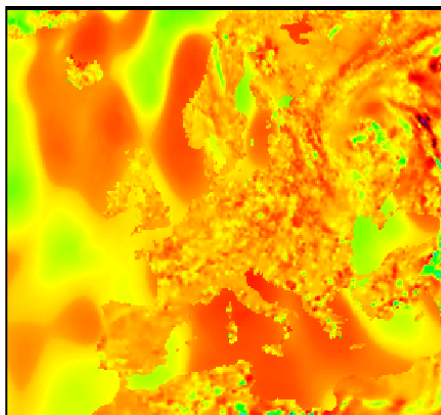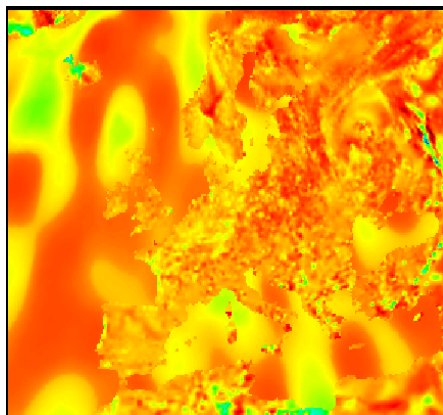 - **Observations**   **Random from error distribution**
  - **Model**   **Random archived analysis increment**
  - **External BCs**   **Ensemble datasets**
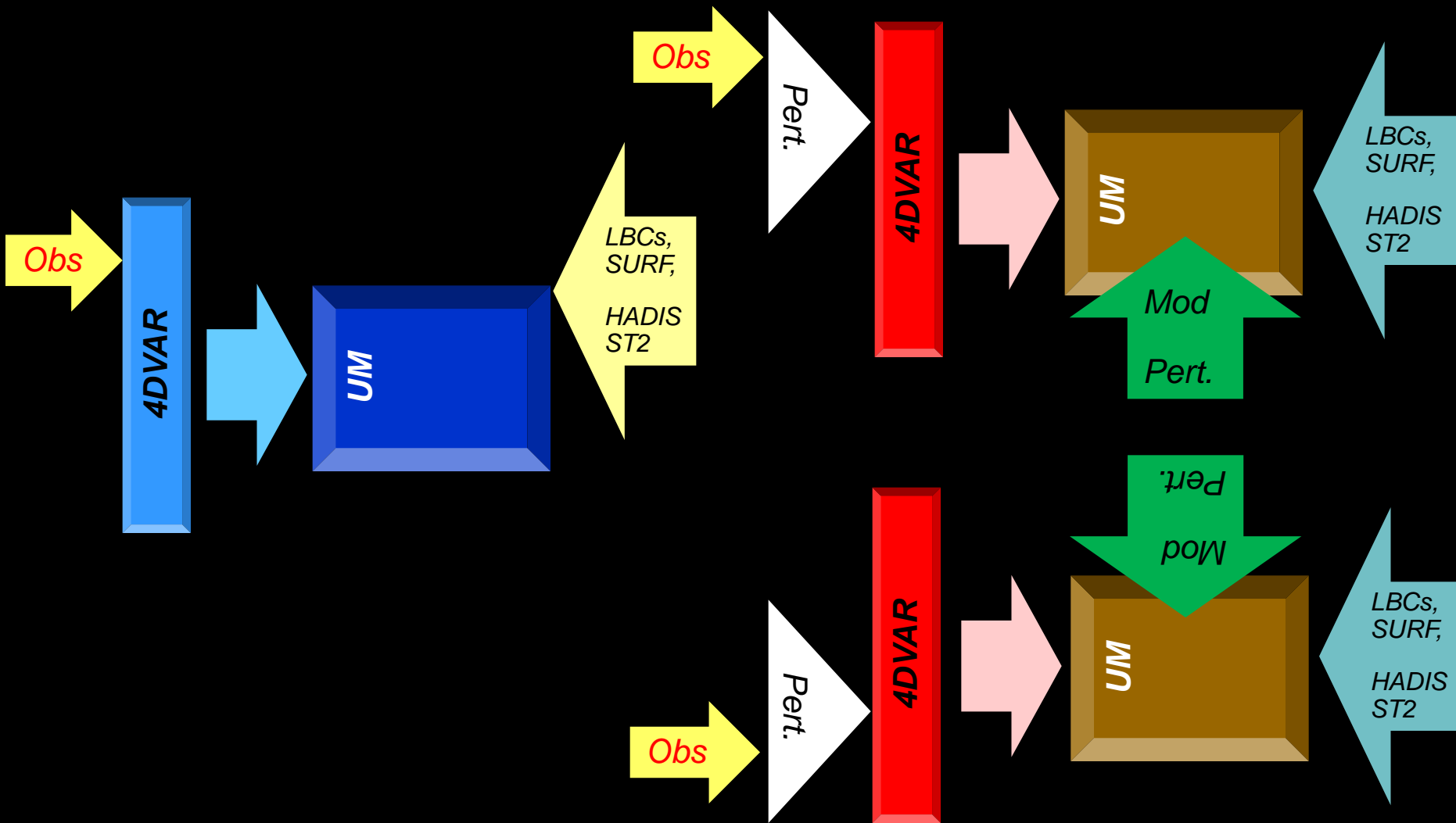
Ensemble Members

- System is not complex - members are independent of each other & deterministic system

Cost

- En4DVAR ≈ 16 x Det.
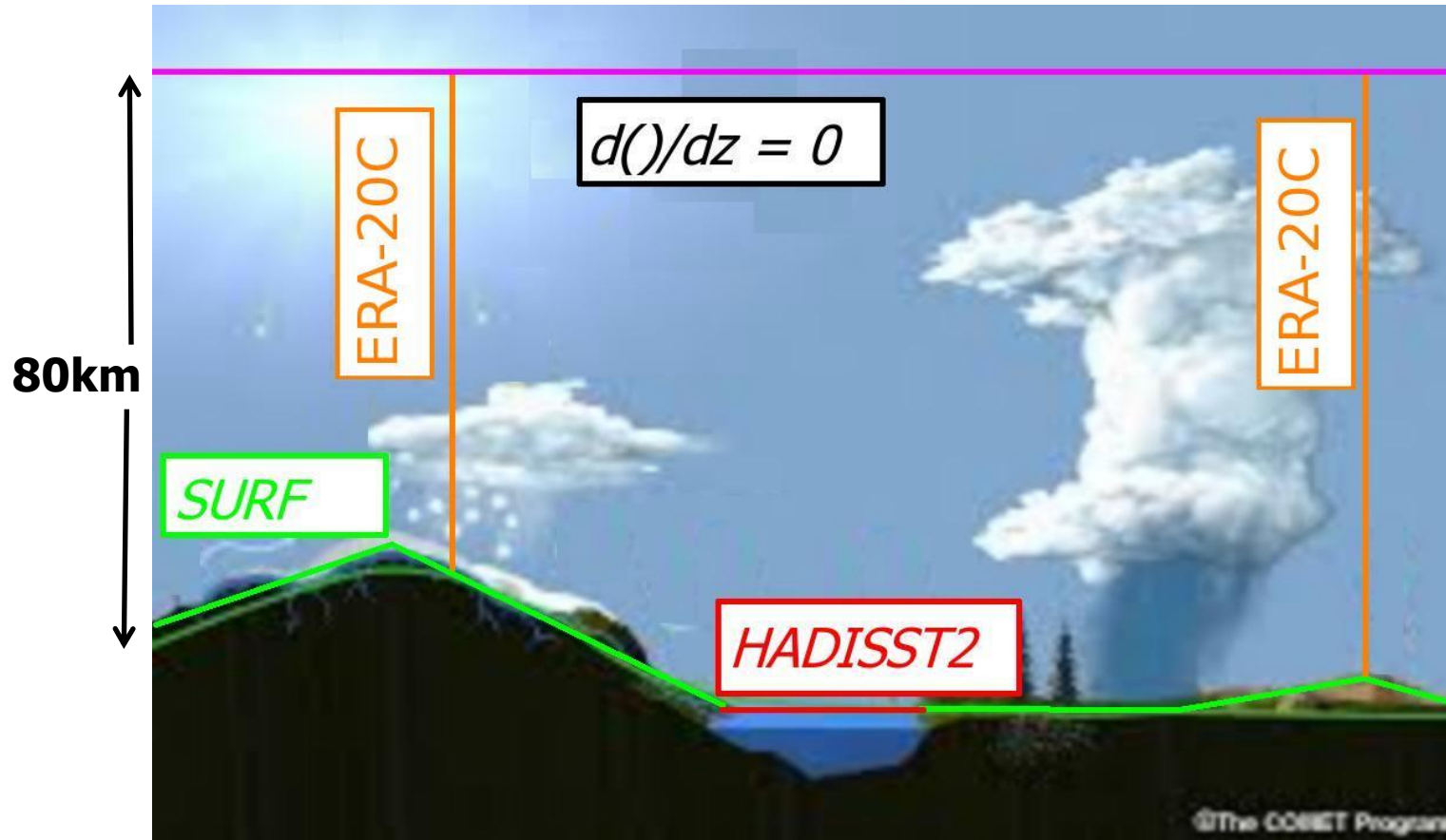
- En4DVAR ≈ 2 x ETKF
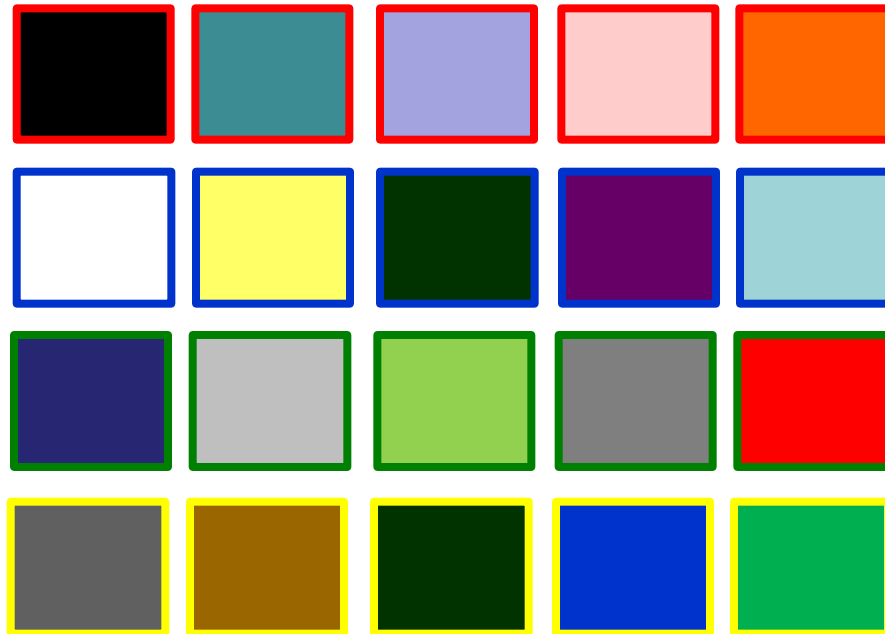
- Reduce resolution

- Other efficiencies...

Boundaries

Lateral Boundary Conditions

- Ensemble BCs from ERA5 (max x10)

- Multiple members using the same BC – to include large scale and small scale uncertainty

Summary

- Deterministic system will use hybrid 4DVAR (EDA)

- En4DVAR is expected to give a better estimate of uncertainty than ETKF

- En4DVAR is technically more flexible than ETKF

- En4DVAR may cost up to twice as much as ETKF, but costs can be reduced

- Ensemble includes representation of uncertainties in ...

  - Observations
  - LBCs
  - Surface (land/ocean)
  - Model

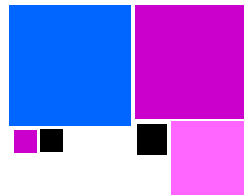# Thank you for listening...

...Questions?

Cost

- Cost of deterministic system 90 PEdays/day

- Cost of ETKF estimated to be 600 PEdays/day

- Cost of full res. En4DVAR estimated to be 1400 PEdays/day

- Cost of 1/2 res. En4DVAR estimated to be 400 PEdays/day

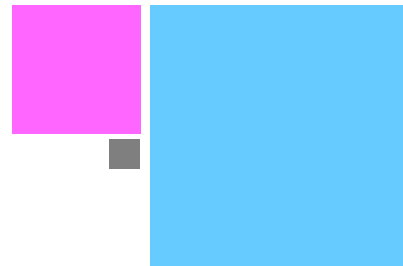- Further reduce cost via efficient use of ensembles

UM DA OPS

Det.   ETKF   En4DVAR full res.   En4DVAR 1/2 res.

Forecast Model: Uncertainty

$$x_i^a = x_i^b + x_i' \qquad x_i' = x_i^a - x_i^b$$

$$x_i^b = M\left(x_{i-1}^a\right)$$

$$Q_i = x_i^* - M\left(x_{i-1}^*\right)$$

$$x_i^a - M\left(x_{i-1}^a\right) = x_i^a - x_i^b = x_i'$$

$$\boxed{\left\{x_j'\right\} \approx \left\{Q_j\right\}}$$
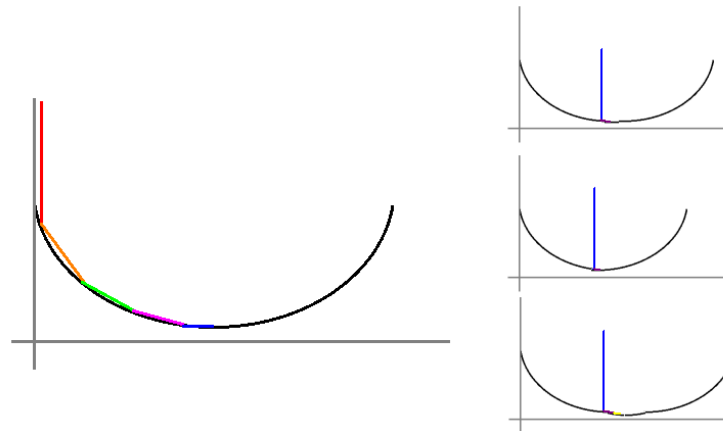
Observations: Uncertainty

- Our assimilation scheme uses estimates of observation error ObsErr

- For a certain ensemble member, for each observation:

    - obtain a random number (usually [-0.5,0.5], Gaussian), Rand

    - ObsPert = ObsVal + Rand x ObsErr

Efficient Ensemble Use

- Ensemble spread estimates uncertainty in ensemble mean

- Ensemble mean should be more accurate than deterministic system

- Ensemble mean may replace deterministic reanalysis

- Costs could also be reduced by using deterministic analysis as ensemble guess

- On each cycle a certain member may be more accurate than the mean or det. analysis...

Observations

ECMWF

Land Synop

Sondes

Ships & buoys

Aircraft

AIRS/TOVS/SSMIS/SEVIRIclear

Precip...
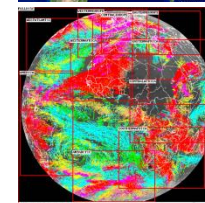
GPSRO

ECA&D

MetDB

Clouds (not avl.)

GroundGPS (not avl.)

AMVs (cant read)

Scatwinds (cant read)

© Crown copyright   Met Office

Satellite BC

- *VarBC*

$$bias = c^{scan} + \sum_{i=1}^{n} ci^{air} f(x_b)$$

Airmass-dependent bias correction of satellite radiances (based on Harris and Kelly, 2001)

VarBC will give smooth and automatic updating
(DingMin Li, Andrew Lorenc , Dale Barker)

- *Thanks to Richard Renshaw*