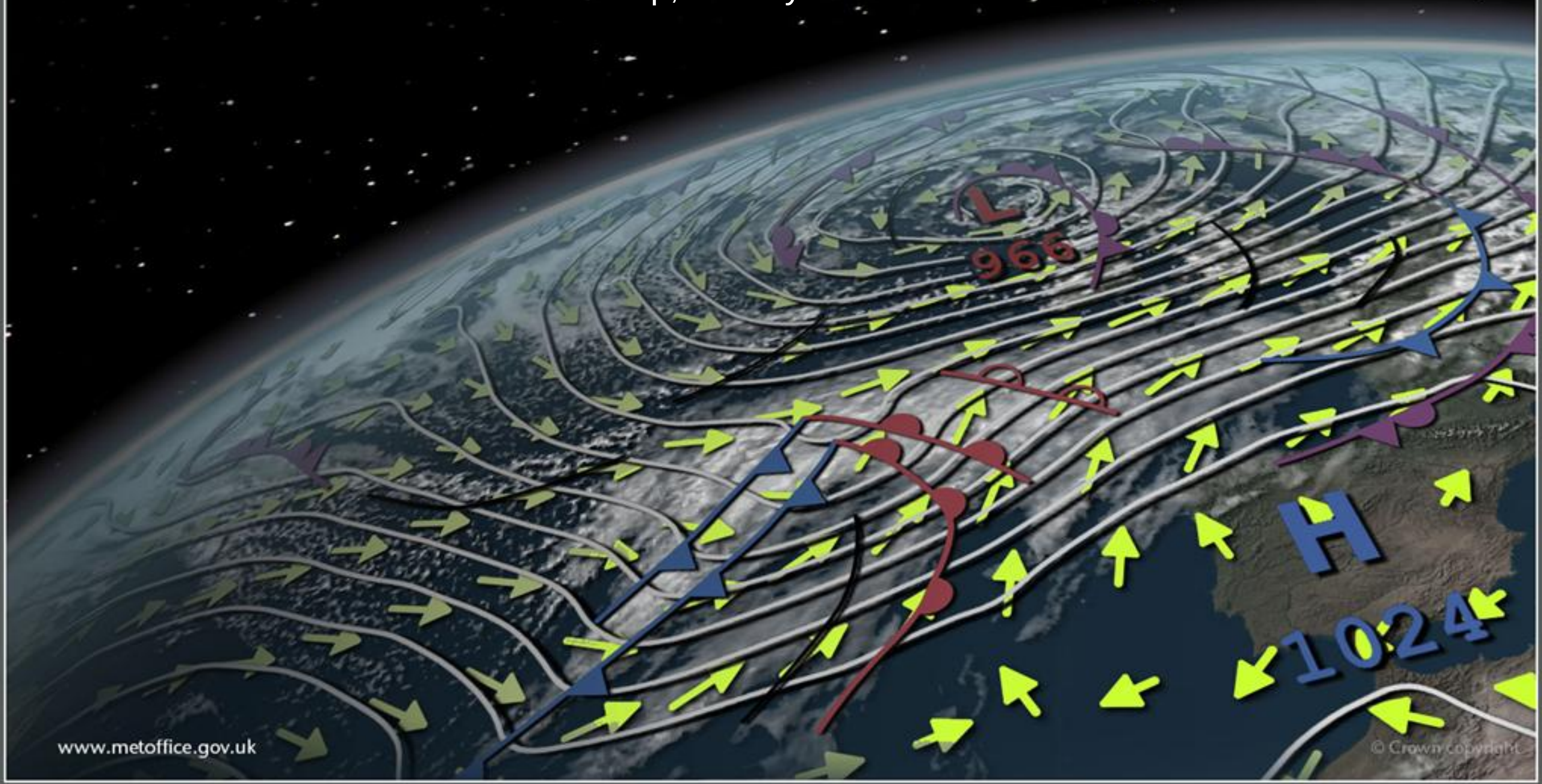
Met Office

Initial trials of convective-scale data assimilation with a cheaply tunable ensemble filter

Jonathan Flowerdew

7th EnKF Workshop, 26 May 2016

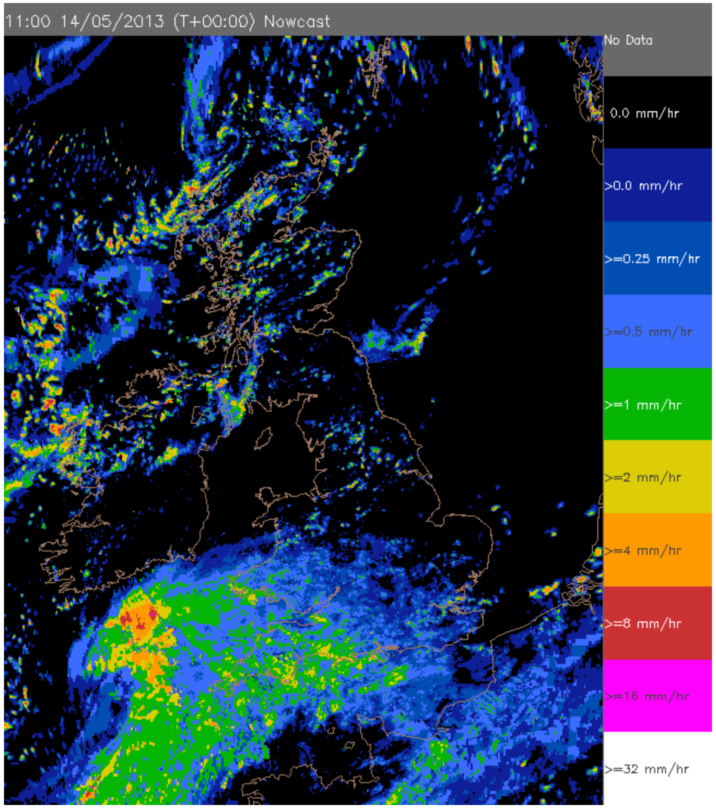


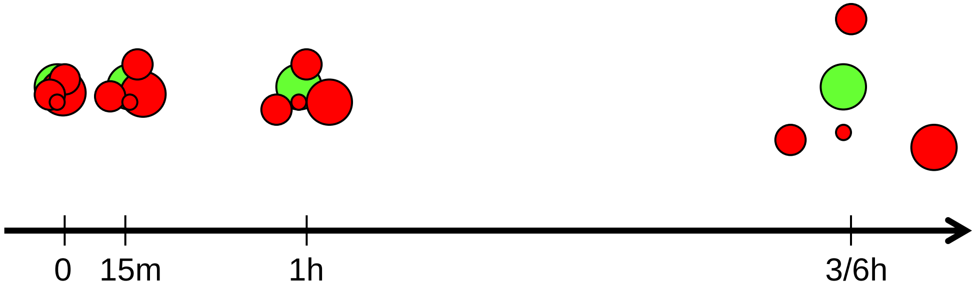
Overall strategy

- Exploring synergy between ensembles and data assimilation at convective scales:
  - Flow-dependent covariances, convective-scale balances, interaction with orography, cross-system relationships, ...
  - Can ensemble DA improve UK deterministic/ensemble forecasts?
  - Long-term research, beyond immediate operational plans such as UK hourly 4DVar
- Explore issues such as localisation, inflation, filter type, radar/cloud observations, lateral boundary handling, updating hydrometeors, …
  - Initial experiments use a serial ensemble filter
- Results will inform ultimate operational ensemble DA system (be that EnKF or EnVar)

Update frequency

- The gold standard would be to synchronise the model with individual convective cells
- This requires perturbations around the actual cell locations, which may be unlikely at T+3/6h for a small ensemble
- But might it be possible at T+1h, 15m, 5m?
- More linear, more Gaussian
- Many nudges within convective life-cycle, whilst preserving large scales
- Reduces time localisation issues
- Many EnKF studies use 5m updates
- May be limited by initialisation shock, or frequent small increments drawn from the same model may reduce such shocks





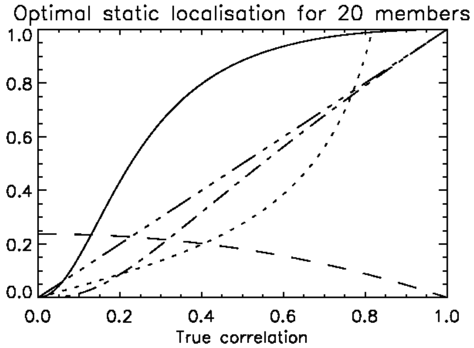
Optimal static localisation

[Flowerdew (2015)](), Tellus

- Minimise RMS error, like the Kalman Filter
- For a single observation: $\boxed{x_2^a - x_2^f = \alpha \hat{b}(y - x_1^f)}$ $\hat{b} = \dfrac{\hat{\sigma}_1 \hat{\sigma}_2 \hat{r}}{\hat{\sigma}_1^2 + \sigma_o^2}$

Minimise $\boxed{J(\alpha) = \int (\alpha \hat{b} - b)^2 P(\hat{b})\, dz}$

$\boxed{\Rightarrow \quad \alpha = b^2 / (\sigma_{\hat{b},N}^2 + b^2)}$

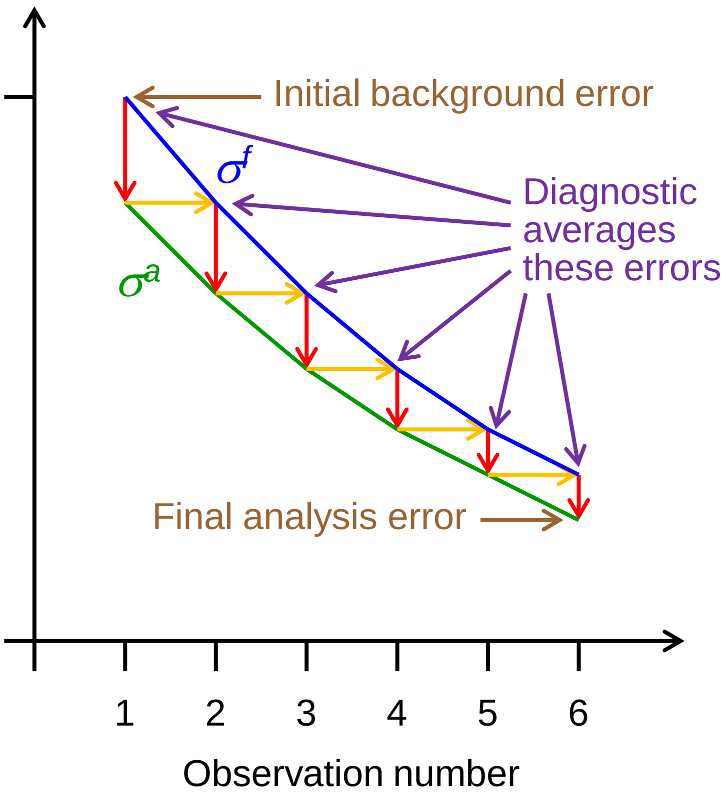


- This also applies between variables, between model and observations, etc
- Extra terms arise for dense observations, eg:

$$J(\mathbf{L}) = Tr < (\tilde{\mathbf{K}} - \mathbf{K})(\mathbf{H}\mathbf{P}^{\mathbf{f}}\mathbf{H}^{\mathbf{T}} + \mathbf{R})(\tilde{\mathbf{K}} - \mathbf{K})^{\mathbf{T}} > \qquad \tilde{\mathbf{K}} = \tilde{\mathbf{P}}^{\mathbf{f}}\mathbf{H}^{\mathbf{T}}(\mathbf{H}\tilde{\mathbf{P}}^{\mathbf{f}}\mathbf{H}^{\mathbf{T}} + \mathbf{R})^{-1} \qquad \tilde{\mathbf{P}}^{\mathbf{f}} = \mathbf{L} \circ \hat{\mathbf{P}}^{\mathbf{f}}$$

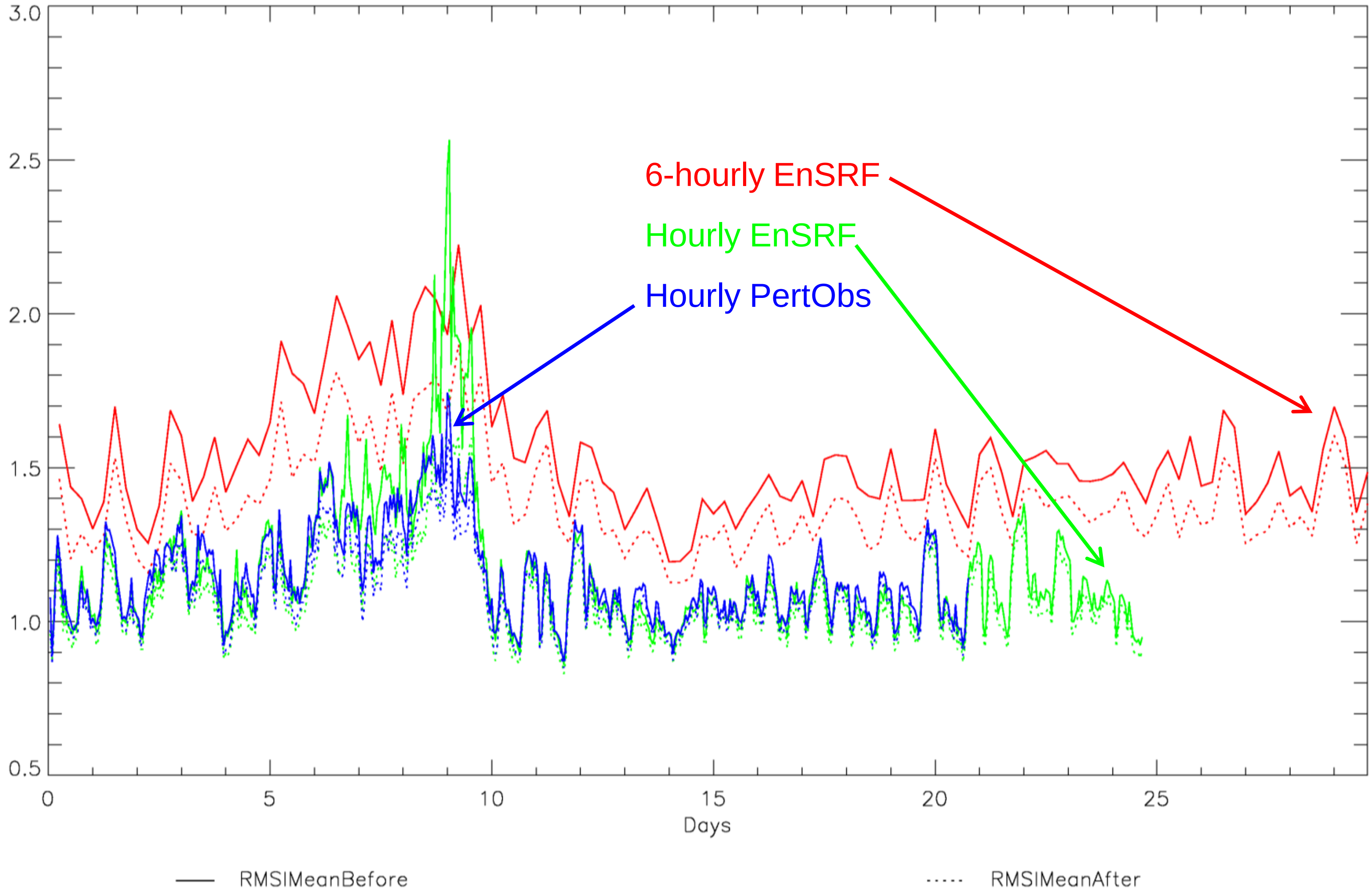
Updating $H(\mathbf{x}^f)$ – and a measure of analysis error

- Serial filters normally update $\mathbf{x}^f$ for observation *j* before calculating $H_{j+1}(\mathbf{x}^f)$
- Our separate OPS calculates all $H(\mathbf{x}^f)$ at the start using the original background state
- We can work around this by updating the observation priors as additional elements of the state vector (Anderson, 2003)
- As a bonus, this naturally gives the innovation variance for each observation after assimilating all prior observations – an independent measure of analysis error



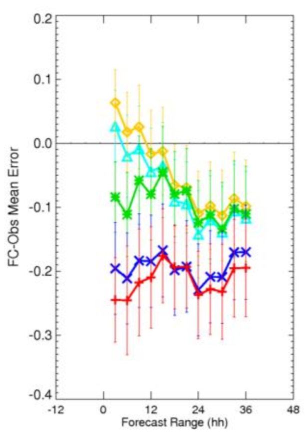
First trial diagnostics

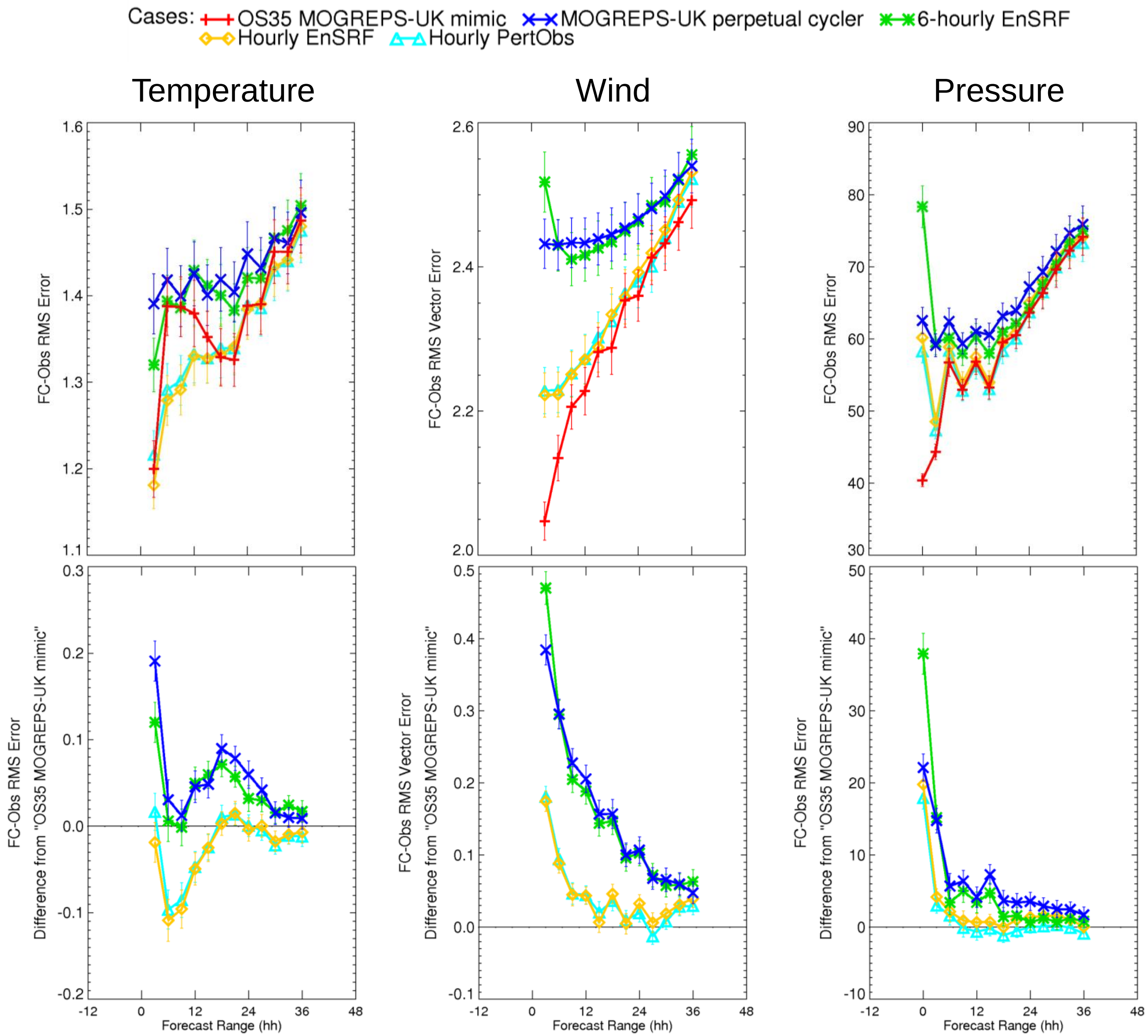
June 2014; sonde, surface and aircraft normalised innovations; 44+1 members



Forecast performance

- Cycler worse than downscaler (loss of interior global DA?)
- 6h EnDA generally improves upon cycler
- 1h EnDA much better than 6h
- Beats downscaler for temperature, not there yet for wind
- EnSRF and PertObs similar
- Also improves temperature bias





Cheaply tuning the EnKF

- The RMS innovation after assimilation of prior observations provides a way to cheaply tune many EnKF parameters
- To test the principle, run EnKF from archived input every 1.25d with all permutations of:
- Check by running full trial of suggested configuration
- The signal may be clear enough to build tuning into cycle-by-cycle EnKF

MaxRadius (km)	1500	1000	666	333	167
Equiv Gauss (km)	275	183	122	61	30
Obs/10k	592	263	117	29	7

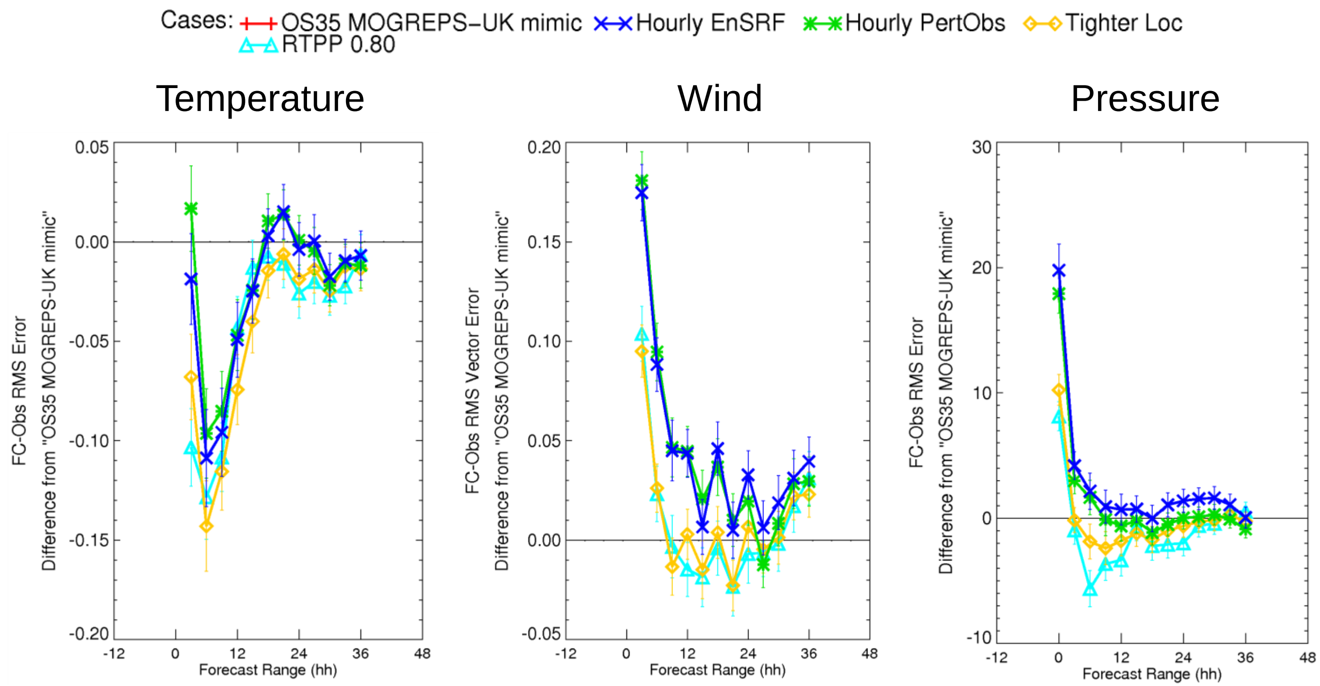
VertLocScale	“0”	2.93	1.26	1.00	0.406	0.214
Gauss P factor	Inf	5	2	1.73	1.25	1.125

V \ H	1000	666	333	167
1.0	1.1342	1.1339	1.1362	1.1398
0.406	1.1323	1.1323	1.1353	1.1393
0.214	1.1305	1.1310	1.1347	1.1391

(data from 2014061921 Hourly PertObs ensemble mean)

Further trial results

- Tighter localisation improves performance, as suggested by tuner
- RTPP plus stochastic physics brings some further improvement
- Now often beating downscaler, despite limited observation set
- Some improvement to categorical scores

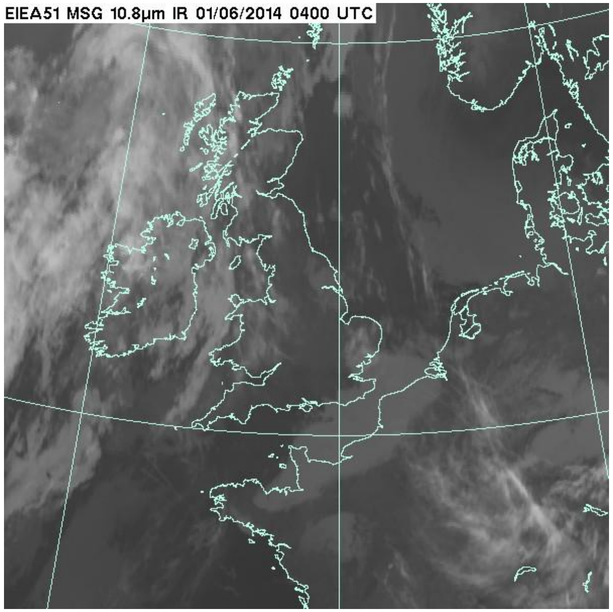


Parameter 0/6/12/18 UTC	Hourly EnSRF	Tighter Loc	RTPP	Downscaler
Surface visibility	0.056	0.064	0.067	0.071
6h Precip Accum	0.206	0.211	0.208	0.210
Total Cloud Amount	0.253	0.254	0.259	0.261
Cloud Base Height (3/8 cover)	0.247	0.248	0.249	0.255

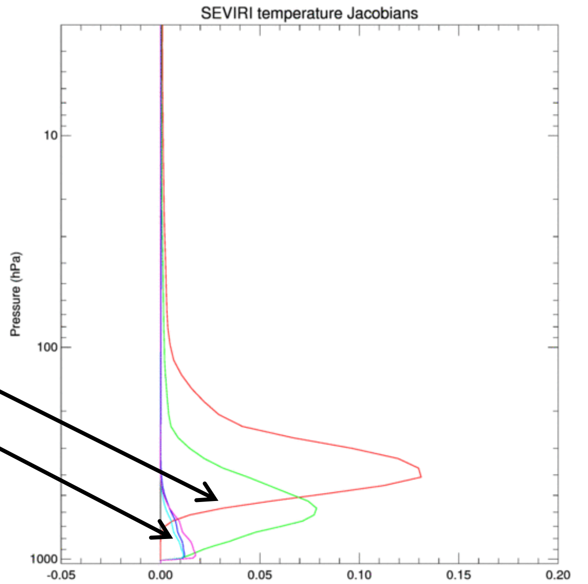
Equitable Threat Scores (higher is better)

SEVIRI cloud-affected radiances

Pete Weston

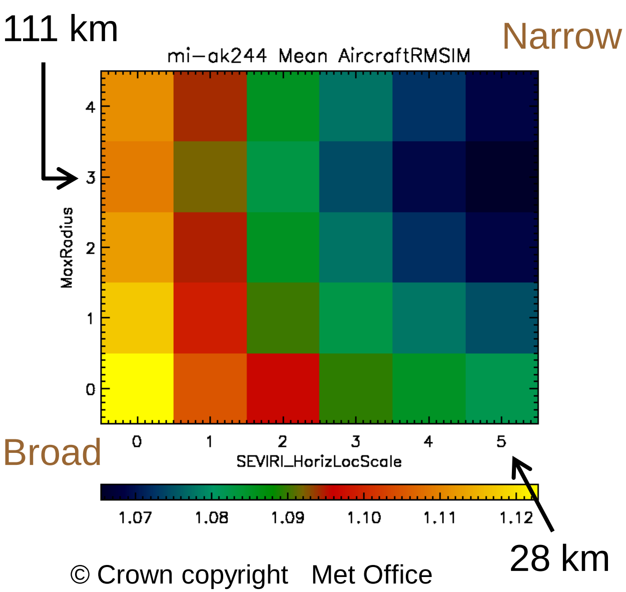


- Useful observation type to test:
  - Satellite assimilation normally ignores areas affected by cloud
  - Cloud is a key forecast variable, tied to convective-scale features
  - Probes dense observations, ‘awkward’ variable, non-trivial observation operator
- Satellite Applications independently chose to test ensemble filter
  - Natural synergy between plans
- Previous CsEnDA system, plus:
  - channel 5 (upper tropospheric humidity, red)
  - channel 9 (cloud top, or surface if clear, cyan)
  - SEVIRI localisation can differ from conventional obs
  - Increment cloud water and ice



SEVIRI trial results

- Strongly draws towards subsequent observations
- Incrementing cloud water/ice is slightly detrimental (need inter-variable localisation?)
- Adding just channel 5 is beneficial
- Adding channel 9 harms overall performance (vertical localisation needs to move with the covariances?)
- Tuner suggests SEVIRI horizontal localisation narrower than conventional observations
- Tuner suggests vertical localisation broadening similar to typical Jacobian widths



Suite ID	Description	Ctrl ID	UKV index	Cloud amount	Cloud base height
mi-af844	Conventional obs	N/A	N/A	N/A	N/A
mi-ag759	Increment cloud water/ice	af844	-0.44%	-0.056	-0.015
mi-ah038	Only channel 5	ag759	+0.97%	+0.004	+0.029
mi-ag842	Channels 5 & 9	ag759	-2.19%	-0.045	-0.045
mi-ah361	Broad localisation	ag759	-6.45%	-0.049	-0.101

Conclusions

- The serial filter is a promising technique for convective-scale data assimilation
- A complete CsEnDA suite has been developed, with flexible cycle length
- Hourly assimilation performs much better than 6-hourly
- Both the deterministic and perturbed observation filters are worth considering
- Tighter horizontal and vertical localisation is beneficial
- SEVIRI channel 5 is beneficial, channel 9 more challenging
- The RMS innovation after assimilating prior observations (IAPO) is a useful diagnostic and allows cheap tuning of parameters such as localisation radii
    - It may also provide a way to calculate the relaxation factor without having to know the observation error

Met Office

Future work

- Further SEVIRI work:
  - Further trials, tuning, diagnostics
  - Vertical localisation moving up/down for channel 9?
  - Apply tuner to observation errors?
  - Shorter (15/30 minute) cycles?
  - Inter-variable localisation?
- Complete wider EnKF experiments:
  - Ensemble verification
  - Localisation, inflation/relaxation, LBCs, ...
  - Radar assimilation
  - Diagnostics/case studies
  - Comparison to 3/4DVAR, perhaps LETKF
- Proposed PhD project extending theoretical/idealised work on localisation, inflation/relaxation and serial/parallel filters
- Then decide what to build operationally

Met Office

Any questions?

Thanks to: Neill Bowler, Gordon Inverarity, DA@Reading, SA Cloud Analysis Review Group, Susanna Hagelin, Kelvyn Robertson, Gareth Dow, Anne McCabe, Jorge Bornemann, Adam Maycock, David Davies, Rob Darvell, ...