



Hybrid variational-ensemble data assimilation at NCEP

Daryl T. Kleist NOAA/NCEP/EMC & University of Maryland-College Park

Dave Parrish¹, Jeff Whitaker², & Xuguang Wang³

With thanks to EMC DA team, John Derber, Steve Lord & Kayo Ide

¹NOAA/NCEP/EMC, ²NOAA/ESRL, ³University of Oklahoma

4th EnKF Workshop, 7-9 April 2010



Motivation



- Current background error covariance (applied operationally at NCEP) in VAR sub-optimal
 - Isotropic recursive filters
 - Poor handle on cross-variable covariance
 - Minimal flow-dependence added
 - Implicit flow-dependence through linearization in normal mode constraint (Kleist et al. 2009)
 - Flow-dependent variances (only for wind, temperature, and pressure) based on background tendencies
 - Tuned NMC-based estimate



Current flow-dependence





- Although flow-dependent *variances* are used, confined to be a rescaling of fixed estimate based on time tendencies
 - No cross-variable or length scale information used
 - Does not necessarily capture 'errors of the day'
- Plots valid 00 UTC 12 September 2008

Level 1 "Unbal" Tv Variance (Reweighted)

Level 1 Tv Ensemble Spread





NOAA EnKF*



- Serial EnSRF (Whitaker and Hamill 2002; Anderson 2003)
 - Options exist for perturbed observations and LETKF formulations
- NCEP GSI for forward operator
 - Ability to assimilate full suite of operationally available observations
- Gaspari-Cohn type localization
- Adaptive radiance bias correction (Miyoshi)
- Can apply multiplicative and/or additive ('NMC' perturbations) inflation



EnKF Experiments



- T190 L64 with GFS
 - 60 ensemble members
 - Multiplicative and additive inflation
 - Assimilate full suite of observations, including in-situ, GPSRO, radiances (AMSU, AIRS, etc.), TC minimum SLP (NHC/JTWC estimate)
- Evaluate
 - Deterministic GFS forecasts from EnKF ensemble mean analysis (two options)
 - Interpolate low resolution mean directly to T382L64
 - Results that follow used this method
 - Maintain high-resolution "control" for replacement of ensemble mean (so-called dual-resolution)
 - Ensemble forecasts (GEFS) from 20 members of 60 member ensemble (interpolated to T190L28)*



EnKF-GFS Verification





AC: P500 HGT SH 00Z, 200B0810-20080922





EnKF-GFS Verification



RMSE: 20080810-20080922 Mean for WIND G2/TRO 00Z 50 EnKF-G81 0.45 GSI 0.3 -0.15--0.156 -0.3 0:45 7.2 $100 \cdot$ -8.4 -0.6-9.6 10.8 -0.75200 -0.6300 -0. 0.15 0.3 400 500 700 -0.15850 4.80.15 ₹.₹ 23 1000 48 96 48 96 144144 Â. 0 Forecast Hour -0.75 -0.6 -0.45 -0.3 -0.150 0.15 0.30.450.60.75



AC: U250 TROPICS 00Z, 20080810-20080922



EnKF Summary



- EnKF-GFS (deterministic)
 - Competitive with GSI despite lower resolution, but...
 - Best performance obtained when utilizing offline satellite bias correction coefficients from operational 3DVAR
 - Investigation on-going, but perhaps irrelevant as we pursue hybrid var/ens DA
 - Experiment utilized data QC'ed by operational (3DVAR based) 06hr forecasts
 - No tropical cyclone relocation
 - Improve TC track forecasts despite this!
- EnKF-GEFS (ensemble)
 - Evaluation on-going, but preliminary results suggest EnKF competitive with ETR based GEFS
- NCEP pursuing hybrid approach in effort to capture benefits of ensemble within existing VAR framework



Hybrid var-ens



• Incorporate ensemble perturbations directly into variational cost function through extended control variable

- Lorenc (2003), Buehner (2005), Wang et. al. (2007), etc.

$$J(\mathbf{x}_{f}, \boldsymbol{\alpha}) = \beta_{f} \frac{1}{2} (\mathbf{x}_{f})^{T} \mathbf{B}^{-1} (\mathbf{x}_{f}) + \beta_{e} \frac{1}{2} (\boldsymbol{\alpha})^{T} \mathbf{A}^{-1} (\boldsymbol{\alpha}) + \frac{1}{2} (\mathbf{y}_{o} - \mathbf{H} \mathbf{x}_{t})^{T} \mathbf{R}^{-1} (\mathbf{y}_{o} - \mathbf{H} \mathbf{x}_{t})$$
$$\mathbf{x}_{t} = \mathbf{x}_{f} + \sum_{k=1}^{K} (\boldsymbol{\alpha}_{k} \circ \mathbf{x}_{k}^{e})$$
$$\frac{1}{\beta_{f}} + \frac{1}{\beta_{e}} = 1$$

 $\beta_{\rm f} \& \beta_{\rm e}$: weighting coefficients for fixed and ensemble covariance respectively $\mathbf{x}_{\rm t}$: (total increment) sum of increment from fixed/static $\mathbf{B}(\mathbf{x}_{\rm f})$ and ensemble \mathbf{B} α_k : extended control variable; $\mathbf{x}_k^{\rm e}$:ensemble perturbation A: correlation matrix [localization on ensemble perturbations]

*Following notation similar to Wang et al. (2007, 2008), Wang (2010)



Hybrid with GSI



- Control variable has been implemented into GSI 3DVAR*
 - Full **B** preconditioning
 - Working on extensions to $\mathbf{B}^{1/2}$ preconditioned minimization options
 - Spectral filter for horizontal part of A
 - Eventually replace with (anisotropic) recursive filters
 - Recursive filter used for vertical
 - Dual resolution capability
 - Ensemble can be from different horizontal resolution than background/ analysis (but same vertical levels)
 - Can use GFS-based or internally generated ensemble
 - Working on building I/O for other models, regional, etc.
 - Option to apply TLNMC (Kleist et al. 2009) to analysis increment

$$\mathbf{x}' = \mathbf{C}\left[\mathbf{x}_{f}' + \sum_{k=1}^{K} \left(\boldsymbol{\alpha}_{k} \circ \mathbf{x}_{k}^{e} \right) \right]$$

*Acknowledgement: Thanks to Dave Parrish who implemented the extended control variable



Single Observation





Single ps observation (-2mb O-F, 1mb error) near center of Hurricane Ike



Single Observation





Single 850mb zonal wind observation (3 m/s O-F, 1m/s error) in Hurricane Ike circulation



Single Observation





Single 850mb Tv observation (1K O-F, 1K error)



Hybrid Cycling Experiments



- Full resolution cycling experiment
 - T382L64 deterministic analysis and forecasts
- Same period as EnKF runs (2008 Hurricane Season)
 - Re-ran control run utilizing latest versions of GSI/GFS
- "One-way coupled"
 - T190L64 perturbations from offline EnKF run used for **B** estimate [no feedback to ensemble system]
- Localization scales for alpha set similar to EnKF run
 - Not exact since current localization implementation for hybrid differs from EnKF
- $(\beta_1)^{-1}=0.5$; TLNMC utilized
 - conservative parameter settings
 - Provide baseline, test mechanics, and find potential issues







Fits of 06h forecasts to radiosondes (O-F, wind) for 20080815-20080915

Hybrid fits are better than control 3DVAR and EnKF.

EnKF fits are for high-resolution deterministic forecast (not ensemble mean). Suffering from spin-up/resolution issues?







Fits of 06h forecasts to radiosondes (O-F, temperature) for 20080815-20080915

Hybrid fits are slightly better in lower troposphere, but worse in upper troposphere & stratosphere.

EnKF fits are for high-resolution deterministic forecast (not ensemble mean).







AC: P500 HGT SH 00Z, 200B0810-20080922









Differences outside of the hollow bars are 95% significant based on 10000 Monte Carlo Tests









Hybrid & EnKF improve track errors from f12-f72, slightly degrade at day 4/5.

Hybrid substantially improves intensity errors. EnKF intensity forecasts suffer from degraded analysis resolution.





- EnKF / Hybrid show substantial promise
- Large collaborative effort underway to sort out best path forward for NCEP
 - EMC (myself, ensemble team, DA team), NOAA/ESRL (Jeff Whitaker), OU (Xuguang Wang)
- Many open questions
 - Ensemble Forecasting
 - EnKF/ETR
 - Hybrid DA
 - Localization (adaptive, flow-dependent, anisotropic), balance, weighting term for **B** (adaptive), many more....
 - Coupling data assimilation and ensemble forecasting systems