Stochastic tendency perturbations for NWP ensembles

Martin Leutbecher

European Centre for Medium-Range Weather Forecasts

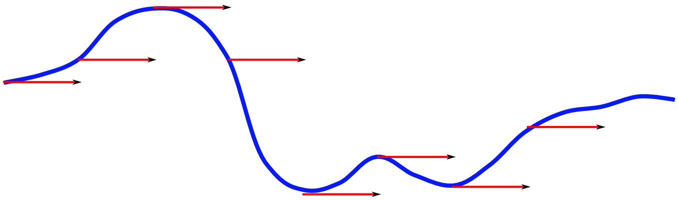
Workshop on Representing Model Uncertainty in numerical weather prediction (NWP) models and in climate models

Acknowledgements: Glenn Shutts, Martin Steinheimer & Peter Bechtold
Lars Isaksen, Massimo Bonavita & Roberto Buizza
Frederic Vitart, Tim Stockdale, Thomas Jung and Tim Palmer

Outline

1. Introduction
2. Tendency perturbations used in ECMWF ensembles
   - Stochastically Perturbed Parameterization Tendencies (SPPT)
   - Stochastic Kinetic Energy Backscatter (SKEB)
3. Impact of tendency perturbations on the EPS
4. Model uncertainty and analysis uncertainty
   - Kalman filter
   - Ensemble of 4D-Vars (EDA)
5. Summary

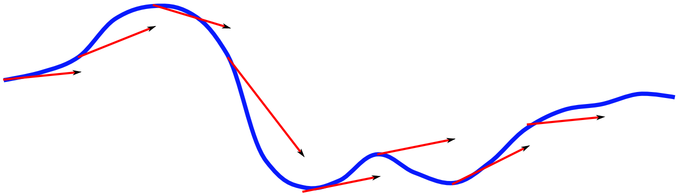
Estimating model error statistics



truth versus (unperturbed) model mismatches over interval Δt

- mismatches $\mathbf{x}_f - \mathbf{x}_t$ are state vectors
- spatial, multi-variate and temporal correlations matter
- error will be a function of the initial state

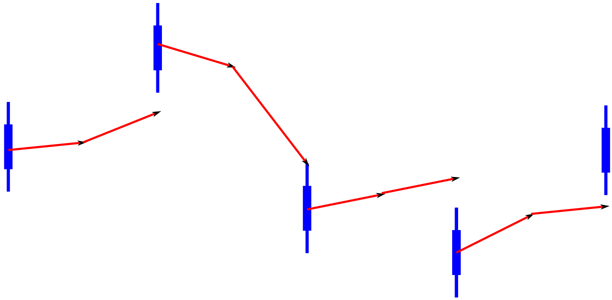
Estimating model error statistics



truth versus (unperturbed) model mismatches over interval Δt

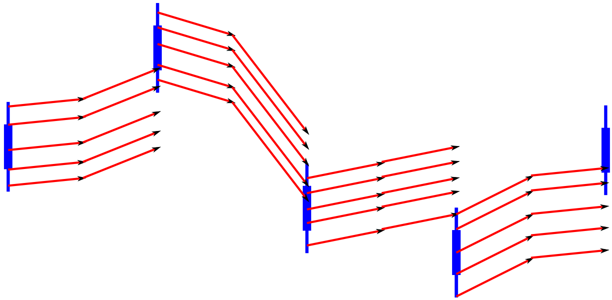
- mismatches $\mathbf{x}_f - \mathbf{x}_t$ are state vectors
- spatial, multi-variate and temporal correlations matter
- error will be a function of the initial state

Estimating model error statistics (II)



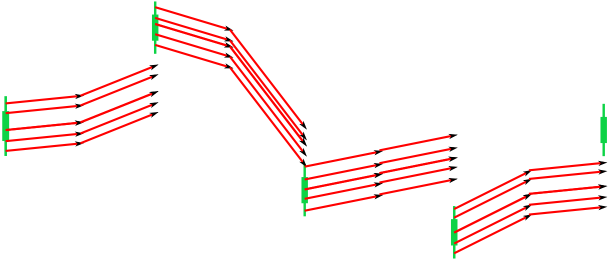
estimate of truth versus model mismatches over interval Δt

Estimating model error statistics (II)



estimate of truth versus model mismatches over interval Δt

Estimating model error statistics (II)



other estimate of truth versus model mismatches over interval Δt

Estimating model error covariances

observable are $\mathbf{G}_a = \langle(\mathbf{x}_f - \mathbf{x}_a)(\mathbf{x}_f - \mathbf{x}_a)^{\mathrm{T}}\rangle$ and
$\mathbf{G}_o = \langle(\mathbf{H}\mathbf{x}_f - \mathbf{y})(\mathbf{H}\mathbf{x}_f - \mathbf{y})^{\mathrm{T}}\rangle$
under some simplifying assumptions (linearity, temporally uncorrelated errors) we expect

$$\mathbf{G}_o = \mathbf{HMAM}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \mathbf{HQH}^{\mathrm{T}} + \mathbf{R} \qquad \text{and} \qquad \mathbf{G}_a = \mathbf{MAM}^{\mathrm{T}} + \mathbf{Q} + \mathbf{A}$$

Estimating model error covariances

observable are $\mathbf{G}_a = \langle(\mathbf{x}_f - \mathbf{x}_a)(\mathbf{x}_f - \mathbf{x}_a)^{\mathrm{T}}\rangle$ and $\mathbf{G}_o = \langle(\mathbf{H}\mathbf{x}_f - \mathbf{y})(\mathbf{H}\mathbf{x}_f - \mathbf{y})^{\mathrm{T}}\rangle$
under some simplifying assumptions (linearity, temporally uncorrelated errors) we expect

$$\mathbf{G}_o = \mathbf{HMAM}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \mathbf{HQH}^{\mathrm{T}} + \mathbf{R} \quad \text{and} \quad \mathbf{G}_a = \mathbf{MAM}^{\mathrm{T}} + \mathbf{Q} + \mathbf{A}$$

IF, initial uncertainty **A** (and **R**) precisely known, then

$$\mathbf{Q} = \mathbf{G}_a - \mathbf{MAM}^{\mathrm{T}} - \mathbf{A} \quad \text{and} \quad \mathbf{HQH}^{\mathrm{T}} = \ldots$$

yields the model error covariance **Q**. Vice versa, errors in **A** (and **R**) will alias into errors of our estimate of **Q** ($\mathbf{HQH}^{\mathrm{T}}$).

Estimating model error covariances

observable are $\mathbf{G}_a = \langle (\mathbf{x}_f - \mathbf{x}_a)(\mathbf{x}_f - \mathbf{x}_a)^{\mathrm{T}} \rangle$ and
$\mathbf{G}_o = \langle (\mathbf{H}\mathbf{x}_f - \mathbf{y})(\mathbf{H}\mathbf{x}_f - \mathbf{y})^{\mathrm{T}} \rangle$
under some simplifying assumptions (linearity, temporally uncorrelated errors) we expect

$$\mathbf{G}_o = \mathbf{HMAM}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}} + \mathbf{HQH}^{\mathrm{T}} + \mathbf{R} \qquad \text{and} \qquad \mathbf{G}_a = \mathbf{MAM}^{\mathrm{T}} + \mathbf{Q} + \mathbf{A}$$

IF, initial uncertainty $\mathbf{A}$ (and $\mathbf{R}$) precisely known, then

$$\mathbf{Q} = \mathbf{G}_a - \mathbf{MAM}^{\mathrm{T}} - \mathbf{A} \quad \text{and} \quad \mathbf{HQH}^{\mathrm{T}} = \dots$$

yields the model error covariance $\mathbf{Q}$. Vice versa, errors in $\mathbf{A}$ (and $\mathbf{R}$) will alias into errors of our estimate of $\mathbf{Q}$ ($\mathbf{HQH}^{\mathrm{T}}$).
The analyis error covariance $\mathbf{A}$ depends on the assimilation technique α, $\mathbf{H}$, $\mathbf{R}$ *and* $\mathbf{Q}$. Thus, we have

$$\mathbf{G}_a = \mathbf{M}\,\mathbf{A}(\alpha, \mathbf{H}, \mathbf{R}, \mathbf{Q})\,\mathbf{M}^{\mathrm{T}} + \mathbf{Q} + \mathbf{A}(\alpha, \mathbf{H}, \mathbf{R}, \mathbf{Q}) \tag{1}$$

$\longrightarrow$ a nontrivial inverse problem!

Ambiguity between initial uncertainty and model uncertainty

- Without constraining the estimate of **A** (completely) by data assimilation, both the representation of initial uncertainties (**A**) and tendency perturbations (**Q**) need to be set for an ensemble forecasting system.

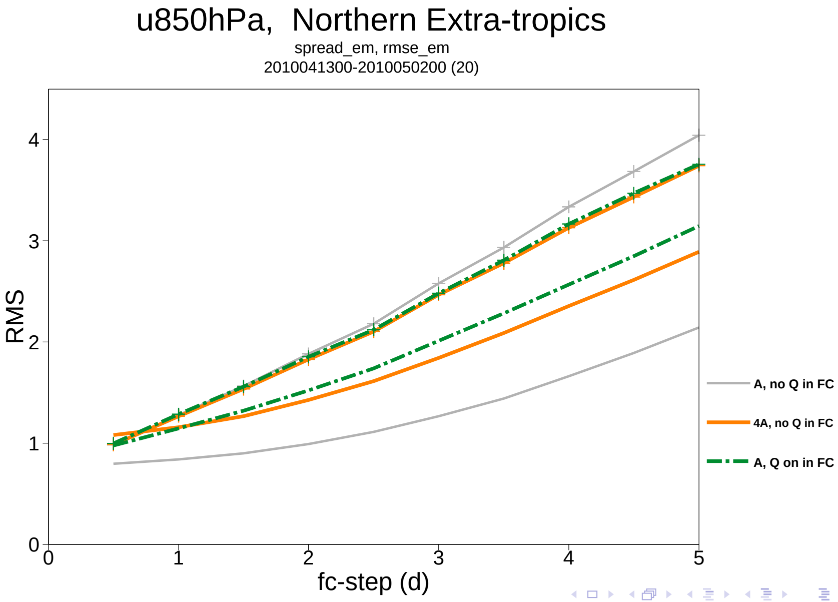
Ambiguity between initial uncertainty and model uncertainty

- Without constraining the estimate of $\mathbf{A}$ (completely) by data assimilation, both the representation of initial uncertainties ($\mathbf{A}$) and tendency perturbations ($\mathbf{Q}$) need to be set for an ensemble forecasting system.
- However, setting $\mathbf{A}$ *and* $\mathbf{Q}$ for ensemble forecasts may be an under-determined problem!
- If the estimate of $\mathbf{A}$ is "too small" (ie. the ensemble variance due to initial uncertainty represented by $\mathbf{A}$ is lower than the error variance), "larger" $\mathbf{Q}$ can compensate.

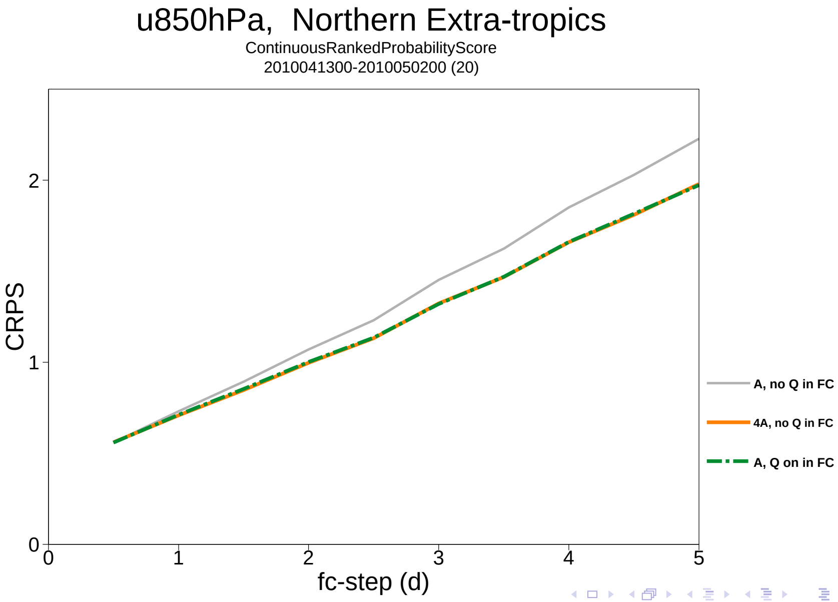
Ambiguity between initial uncertainty and model uncertainty

- Without constraining the estimate of $\mathbf{A}$ (completely) by data assimilation, both the representation of initial uncertainties ($\mathbf{A}$) and tendency perturbations ($\mathbf{Q}$) need to be set for an ensemble forecasting system.
- However, setting $\mathbf{A}$ *and* $\mathbf{Q}$ for ensemble forecasts may be an under-determined problem!
- If the estimate of $\mathbf{A}$ is "too small" (ie. the ensemble variance due to initial uncertainty represented by $\mathbf{A}$ is lower than the error variance), "larger" $\mathbf{Q}$ can compensate.
- Consider for instance $\alpha\tilde{\mathbf{A}}$ and $\beta\tilde{\mathbf{Q}}$ for two estimates of analysis error covariance and model error covariance.
  Can we determine *unambiguously* (α, β) for a NWP ensemble?

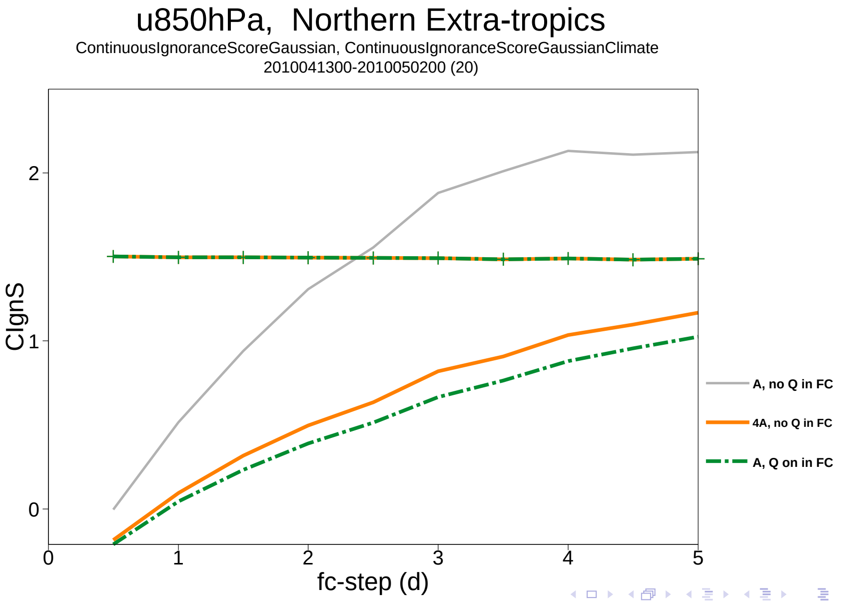
A-Q-Ambiguity (an example with the EPS)



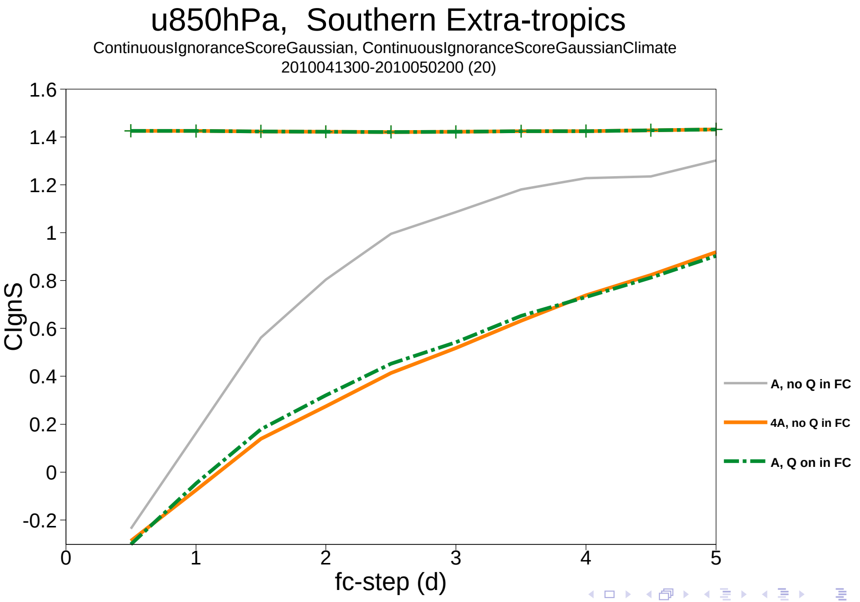
A-Q-Ambiguity (an example with the EPS)



A-Q-Ambiguity (an example with the EPS)



A-Q-Ambiguity (an example with the EPS)



Model uncertainty representation at ECMWF

Status quo

The EPS uses

- Stochastically Perturbed Parameterization Tendencies (SPPT)
  a.k.a. stochastic physics
- Stochastic Kinetic Energy Backscatter (SKEB)

Model uncertainty representation at ECMWF

Status quo

The EPS uses

- Stochastically Perturbed Parameterization Tendencies (SPPT)
  a.k.a. stochastic physics
- Stochastic Kinetic Energy Backscatter (SKEB)

The trajectory and the nonlinear forecast of the perturbed members of the EDA (Ensemble of 4D-Vars) use SPPT only.

- Work is in progress to make the representation of model uncertainties in the nonlinear forecasts in EPS and EDA consistent
- Full consistency requires more → weak-constraint 4D-Var

Stochastically Perturbed Parameterization Tendencies
SPPT

- Physics tendencies P perturbed by $\Delta P = rP$, with r a random pattern
- Improved version of the original SPPT scheme (stochastic physics, Buizza, Miller & Palmer (1999)

Stochastically Perturbed Parameterization Tendencies
SPPT

- Physics tendencies P perturbed by $\Delta P = rP$, with r a random pattern
- Improved version of the original SPPT scheme (stochastic physics, Buizza, Miller & Palmer (1999)
- 2D Random pattern r uses AR-1 processes in spectral space and is smooth in space and time (instead of $10^\circ \times 10^\circ$ tiles changing every 6 time steps)
- Three components with different correlation scales:
  6 h, 3 d, 30 d and 500 km, 1000 km, 2000 km with standard deviations of 0.52, 0.18, 0.06, respectively

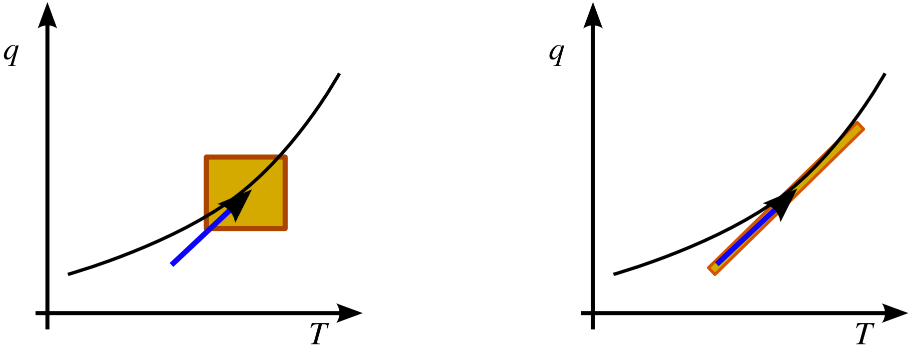
Stochastically Perturbed Parameterization Tendencies

SPPT

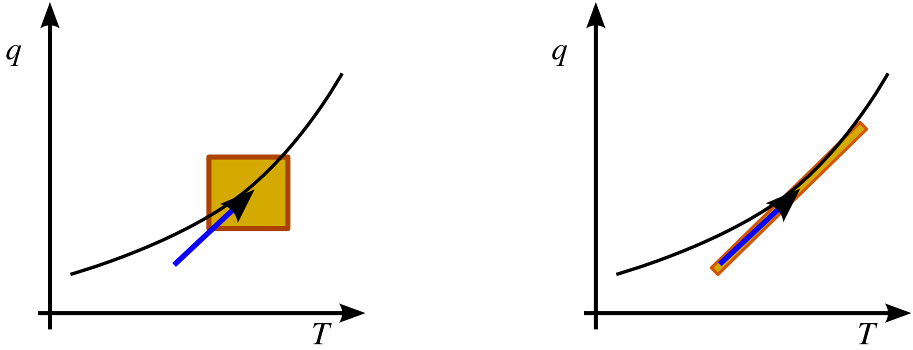
- Physics tendencies P perturbed by $\Delta P = rP$, with r a random pattern
- Improved version of the original SPPT scheme (stochastic physics, Buizza, Miller & Palmer (1999)
- 2D Random pattern r uses AR-1 processes in spectral space and is smooth in space and time (instead of $10^\circ \times 10^\circ$ tiles changing every 6 time steps)
- Three components with different correlation scales:
  6 h, 3 d, 30 d and 500 km, 1000 km, 2000 km with standard deviations of 0.52, 0.18, 0.06, respectively
- Gaussian distribution, truncated at $\pm 2\sigma$ (instead of uniform distr.)
- Same pattern r for T, q, u, v (instead of an independent patterns for each variable)

see Tech Memo 598, Palmer et al. (2009) for more details

Multi-variate uniform versus univariate Gaussian



Multi-variate uniform versus univariate Gaussian

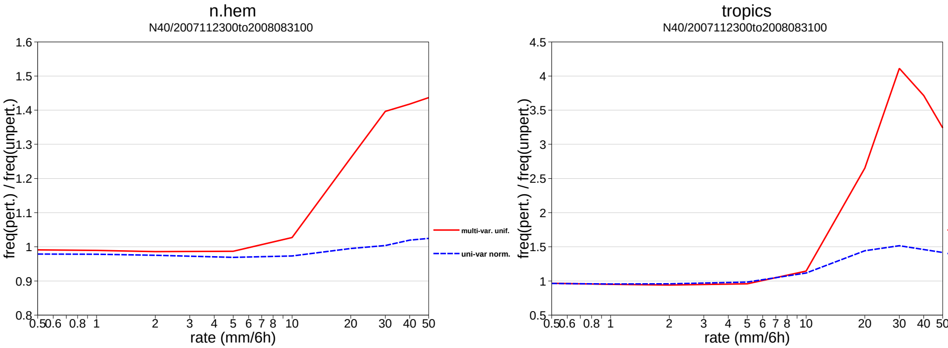


multi-variate uniform in 4 dimensions:

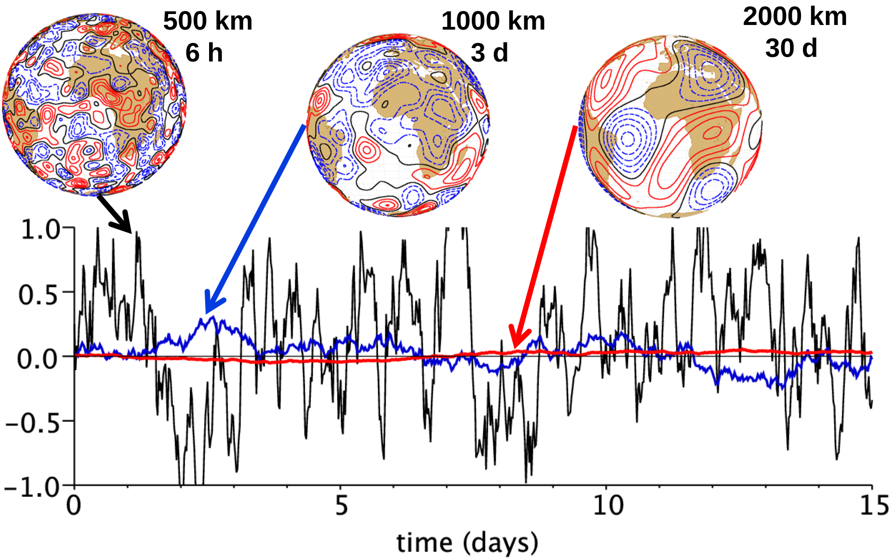
- probability to be within interquartile range for all four variables is $1/16$
- probability to perturb at least one of the four variables in excess of 0.92 of the maximum perturbation amplitude is $0.5 = (1 - 2 \times 0.08)^4$.

Tendency pert^{ns} and the frequency of heavy precipitation

multi-variate uniform distribution of (u, v, T, q) ten. perturbations
uni-variate Gaussian tendency perturbations



SPPT pattern



Stochastic Kinetic Energy Backscatter
SKEB

- Rationale: A fraction of the dissipated energy is backscattered upscale and acts as streamfunction forcing for the resolved-scale flow (Shutts and Palmer 2004, Shutts 2005, Berner et al. 2009)
- Streamfunction forcing $= [bD]^{1/2}\ F(\mathbf{x}, t)$,
  where b, D, F denote the backscatter ratio, the (smoothed) total dissipation rate and the 3-dim evolving pattern, respectively

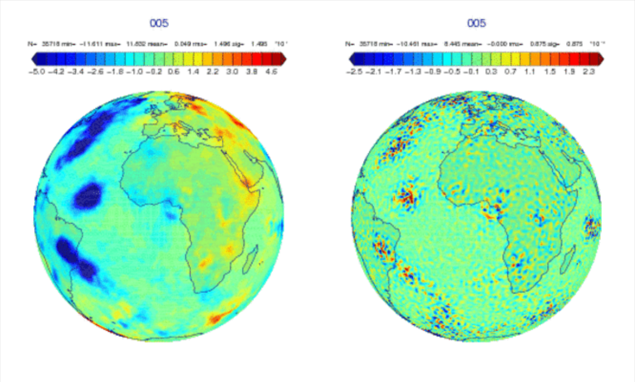
Stochastic Kinetic Energy Backscatter
SKEB

- Rationale: A fraction of the dissipated energy is backscattered upscale and acts as streamfunction forcing for the resolved-scale flow (Shutts and Palmer 2004, Shutts 2005, Berner et al. 2009)
- Streamfunction forcing $= [bD]^{1/2}\ F(\mathbf{x}, t)$,
  where b, D, F denote the backscatter ratio, the (smoothed) total dissipation rate and the 3-dim evolving pattern, respectively
- Total dissipation rate: sum of
  - "numerical" KE dissipation by numerical diffusion + interpolation in semi-Lagrangian advection
  - dissipation from orographic gravity wave drag parameterization
  - an estimate of the deep convective KE production
- Boundary layer dissipation is omitted

see also Tech Memo 598, Palmer et al. (2009) for further details

SKEB forcing

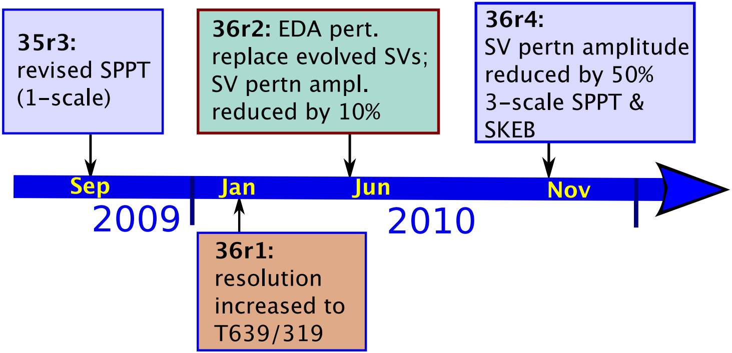
streamfunction | vorticity



- F uses AR-1 processes in spectral space with random vertical phase shifts
- decorrelation time of pattern F is set to 7 h
- structure of pattern constrained by results from coarse-graining studies with T1279 IFS and CRM

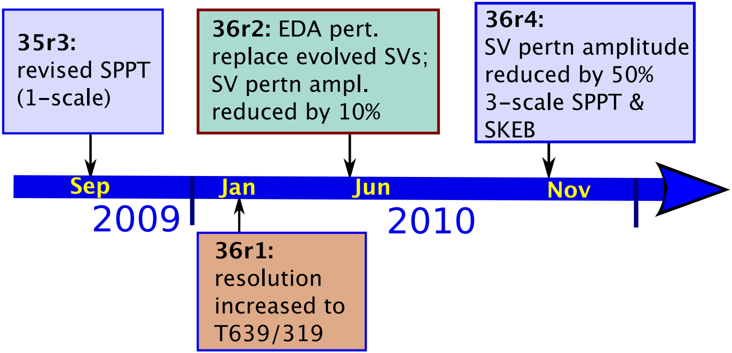
Recent operational implementations

affecting the EPS



Recent operational implementations

affecting the EPS



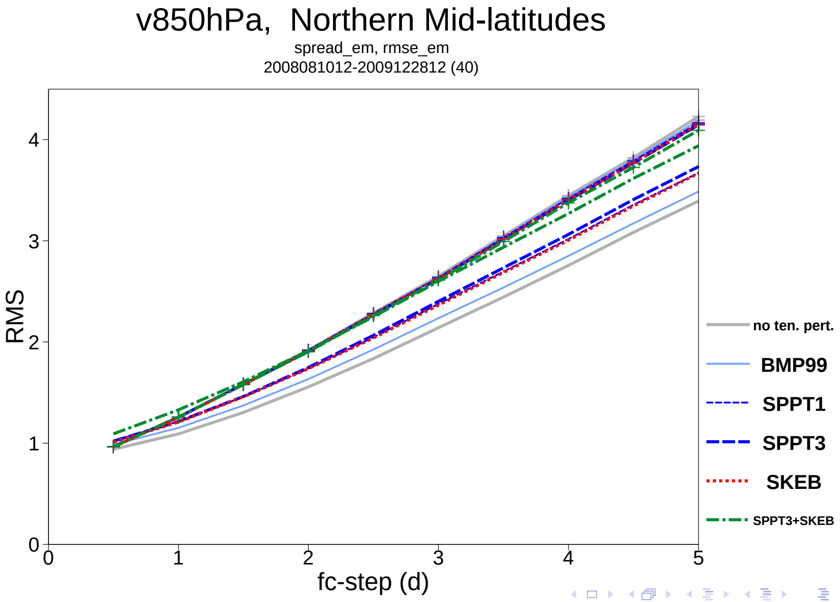
Note, EDA uses the 1-scale version of SPPT (as implemented in 35r3 in the EPS)

Impact of tendency perturbations on the EPS

for a fixed representation of initial uncertainties

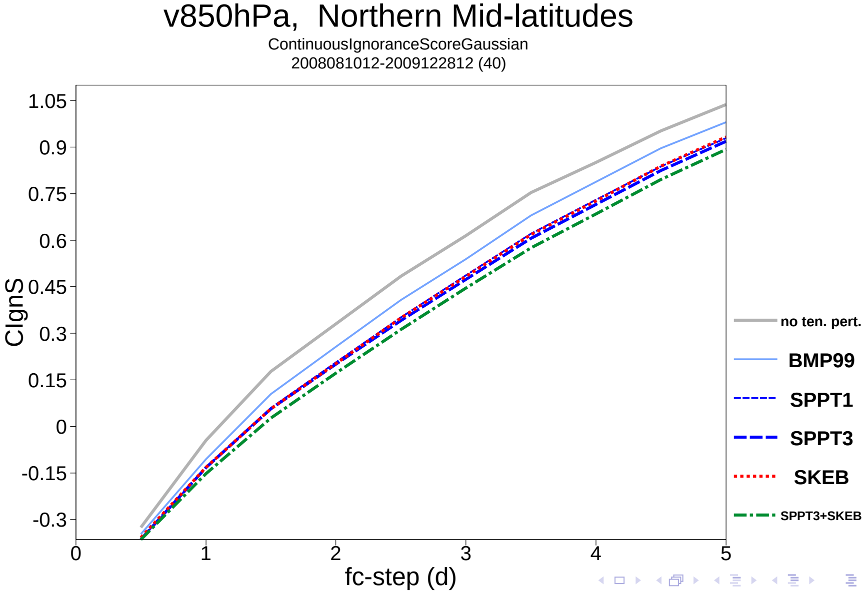
- initial perturbations as used since 36r4
  - EDA perturbations instead of evolved SV perturbations
  - 50 % reduced amplitude of initial SV perturbations
- 40 cases: Aug/Sep 2008 and Oct–Dec 2009
- T639, 50 member
- cycle 36r2
- 6 different tendency perturbations
  - no ten. perturbations
  - original SPPT, **BMP99** (Buizza, Miller & Palmer, 1999)
  - single-scale SPPT (**SPPT1** as implemented in 35r3)
  - three-scale SPPT (**SPPT3** as implemented in 36r4)
  - stochastic kinetic energy backscatter (**SKEB**)
  - **SPPT3+SKEB**

Ensemble standard deviation (no symbols), EM RMSE (+)



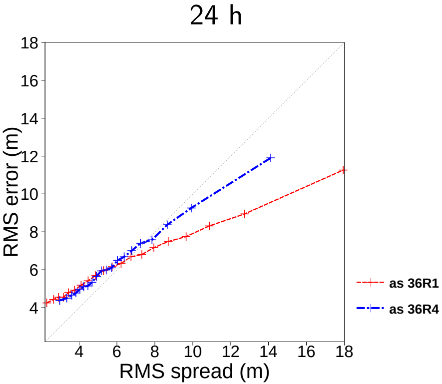
Ignorance score (=Logarithmic score)

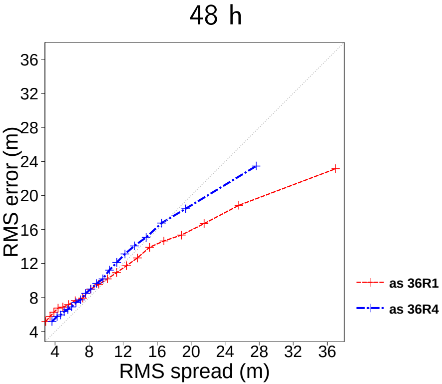
$\mathrm{CIgnS} = -\log(p_{\mathrm{fc}}(y))$; the smaller the better



Spread-reliability of 500 hPa height — 20°–90°N

Jan 2010 configuration versus Nov 2010 configuration



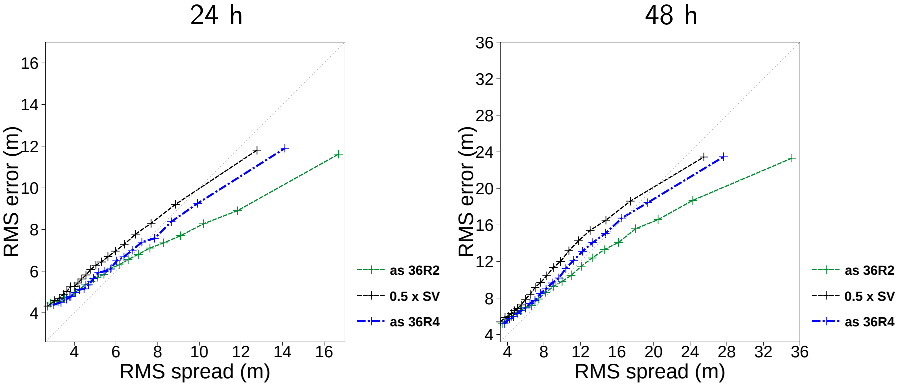


- 40 cases: Aug/Sep 2008 and Oct–Dec 2009
- T639, 50 member
- cycle 36r2

- Major improvement of a long-standing deficiency
- Which change of the EPS configuration is responsible?

Spread-reliability of 500 hPa height — 20°–90°N

Impact of halved SV perturbation amplitude



- Main improvement from reduced SV perturbation amplitude
- Probabilistic skill of 0.5 × SV is inferior to 36R2 configuration
- smaller contribution from 36R1 → 36R2
- upgraded tendency perturbations prevent underdispersion

Kalman filter and model uncertainty

see Daley & Menard (1993)

- variance evolution in the Kalman filter:

$$\text{forecast step} \qquad \mathbf{P}^f = \mathbf{M}\,\mathbf{P}^a\,\mathbf{M}^T + \mathbf{Q} \tag{2}$$

$$\text{analysis step} \qquad \mathbf{P}^a = (\mathbf{I} - \mathbf{K}\,\mathbf{H})\,\mathbf{P}^f, \qquad \text{where} \tag{3}$$

$$\text{gain matrix} \qquad \mathbf{K} = \mathbf{P}^f\,\mathbf{H}^T(\mathbf{H}\,\mathbf{P}^f\,\mathbf{H}^T + \mathbf{R})^{-1} \tag{4}$$

Kalman filter and model uncertainty

see Daley & Menard (1993)

- variance evolution in the Kalman filter:

$$\text{forecast step} \qquad \mathbf{P}^f = \mathbf{M}\,\mathbf{P}^a\,\mathbf{M}^T + \mathbf{Q} \tag{2}$$

$$\text{analysis step} \qquad \mathbf{P}^a = (\mathbf{I} - \mathbf{K}\,\mathbf{H})\,\mathbf{P}^f, \qquad \text{where} \tag{3}$$

$$\text{gain matrix} \qquad \mathbf{K} = \mathbf{P}^f\,\mathbf{H}^T(\mathbf{H}\,\mathbf{P}^f\,\mathbf{H}^T + \mathbf{R})^{-1} \tag{4}$$

- Equivalence between 4D-Var and a Kalman smoother
- Many ensemble assimilation techniques aim at approximating the Extended Kalman filter

Kalman filter and model uncertainty

see Daley & Menard (1993)

- variance evolution in the Kalman filter:

$$\text{forecast step} \qquad \mathbf{P}^f = \mathbf{M}\,\mathbf{P}^a\,\mathbf{M}^T + \mathbf{Q} \tag{2}$$

$$\text{analysis step} \qquad \mathbf{P}^a = (\mathbf{I} - \mathbf{K}\,\mathbf{H})\,\mathbf{P}^f, \qquad \text{where} \tag{3}$$

$$\text{gain matrix} \qquad \mathbf{K} = \mathbf{P}^f\,\mathbf{H}^T(\mathbf{H}\,\mathbf{P}^f\,\mathbf{H}^T + \mathbf{R})^{-1} \tag{4}$$

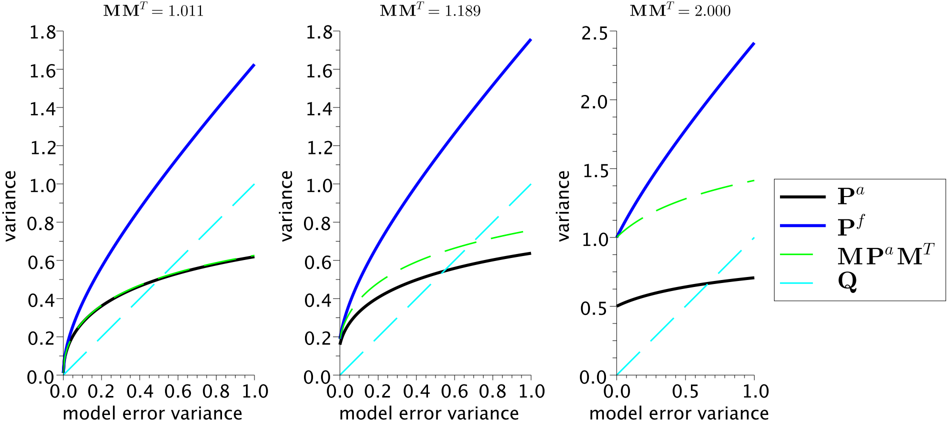
- Equivalence between 4D-Var and a Kalman smoother
- Many ensemble assimilation techniques aim at approximating the Extended Kalman filter
- What is the impact of model uncertainty in the simplest possible KF?
- DM93 studied properties of the KF with stationary **R M Q H** for the case where all matrices can be diagonalized simultaneously
  $\Rightarrow$ independent KF's, each provides the analysis for one scalar variable

Sensitivity to model error variance

Stationary Kalman filter for a scalar variable

slow pert^n growth
frequent obs^{ns}

fast pert^n growth
infrequent obs^{ns}



- all variances normalized by $\mathbf{R} = \sigma_o^2$

Impact of representing model uncertainties

in EDA and EPS on ensemble forecasts

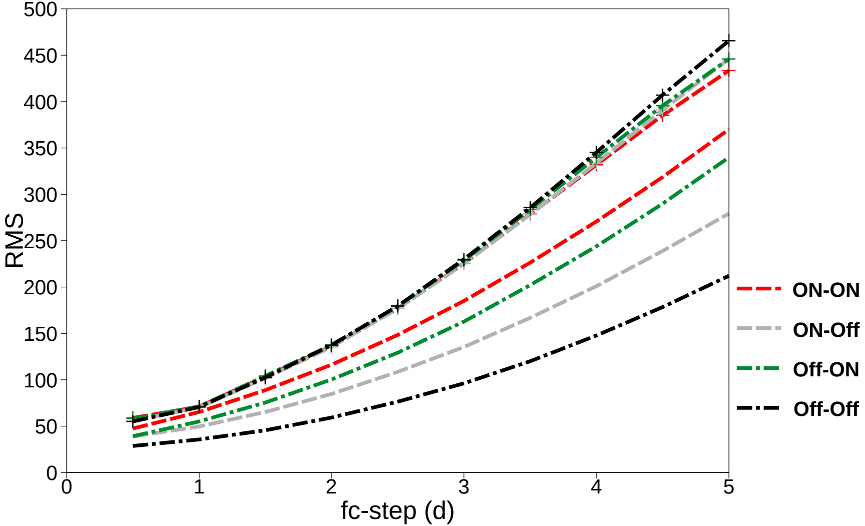
- 3 EDA experiments; (10 member, T399):
  - no tendency perturbations
  - SPPT
  - SPPT+SKEB
- 5 EPS experiments (20 member, T639):

pertn. in EDA	perturbation in EPS	
	None	SPPT+SKEB
None	Off-Off	Off-ON
SPPT+SKEB	ON-Off	ON-ON
SPPT		SPPT-ON

  - no SV perturbations
  - EDA perturbations defined with respect to EDA mean
  - analysis uncertainties accounted for in verification
- 20 cases in April/May 2010
- cycle 36r4
- see also earlier results in Sec. 3 of Tech Memo 598, Palmer et al. (2009)

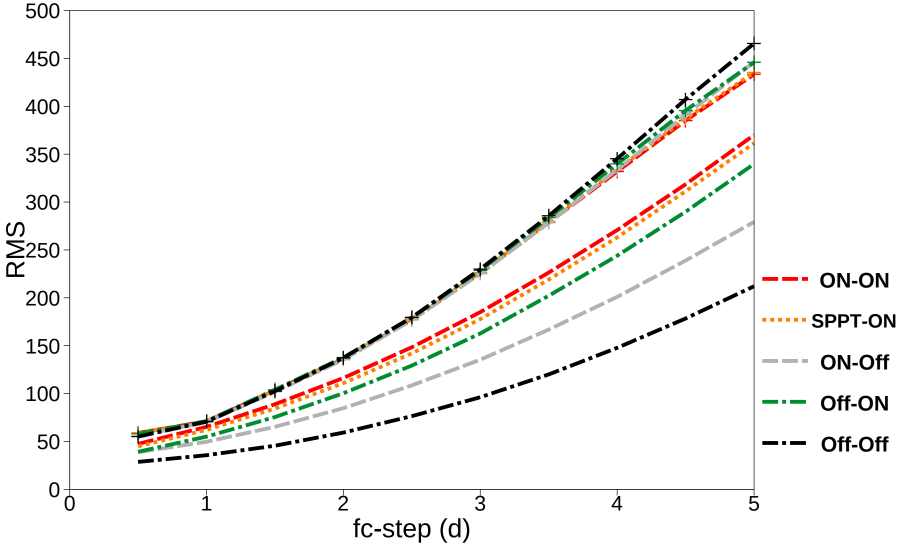
Ensemble standard deviation (no symbols), EM RMSE (+)

500 hPa geopotential — 35°N–65°N



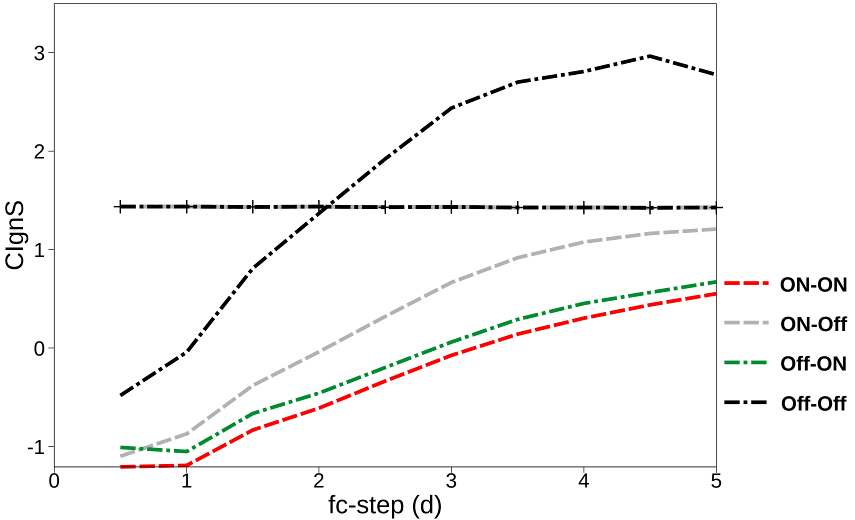
Ensemble standard deviation (no symbols), EM RMSE (+)

500 hPa geopotential — 35°N–65°N



Continuous Ignorance Score

500 hPa geopotential — 35°N–65°N

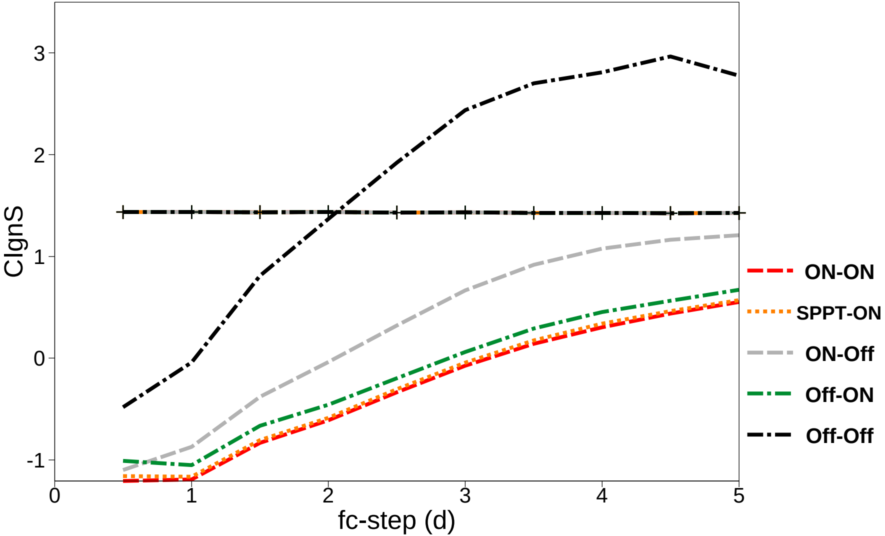


$\text{CIgnS} = -\log\left(\text{p}(\text{y})\right)$

the smaller the better

Continuous Ignorance Score

500 hPa geopotential — 35°N–65°N

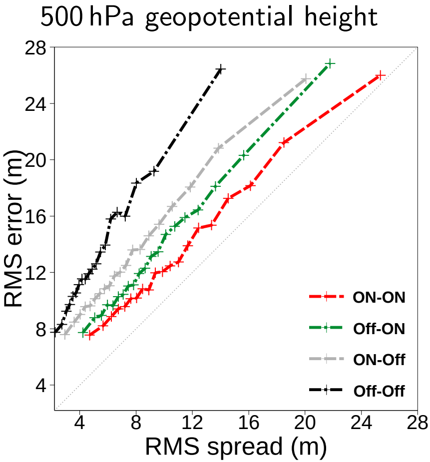


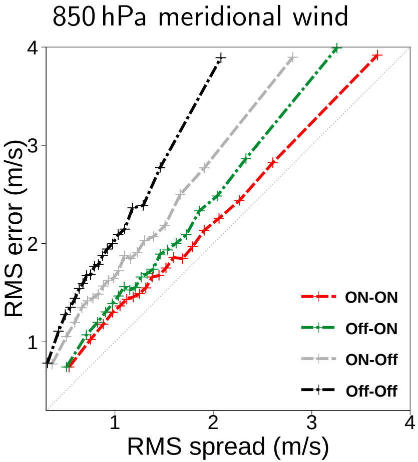
$\mathrm{CIgnS} = -\log(\mathrm{p}(\mathrm{y}))$

the smaller the better

Spread-reliability: $t = 48$ h

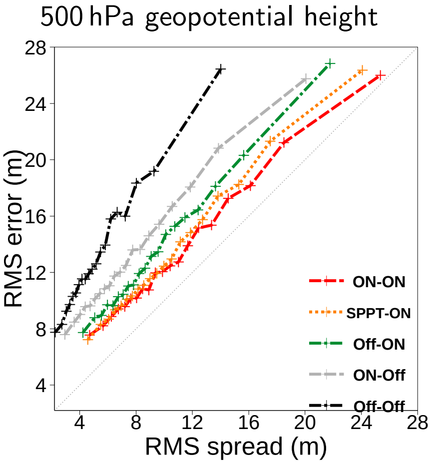
Northern mid-latitudes 35°N–65°N

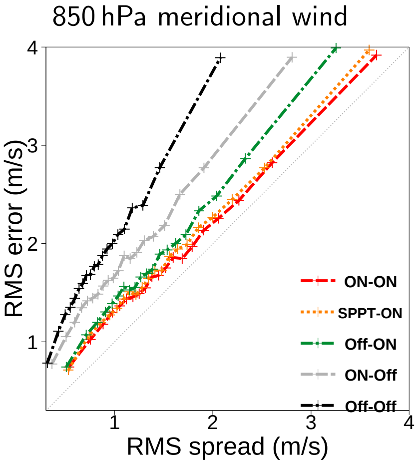




Spread-reliability: $t = 48\,\text{h}$

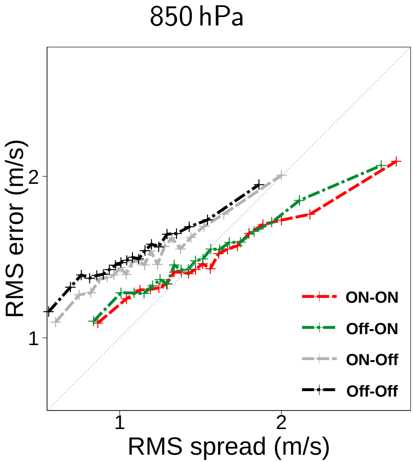
Northern mid-latitudes 35°N–65°N

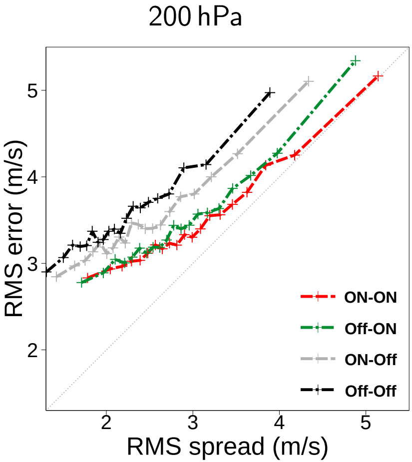




Spread-reliability: meridional wind $t = 48\,\text{h}$

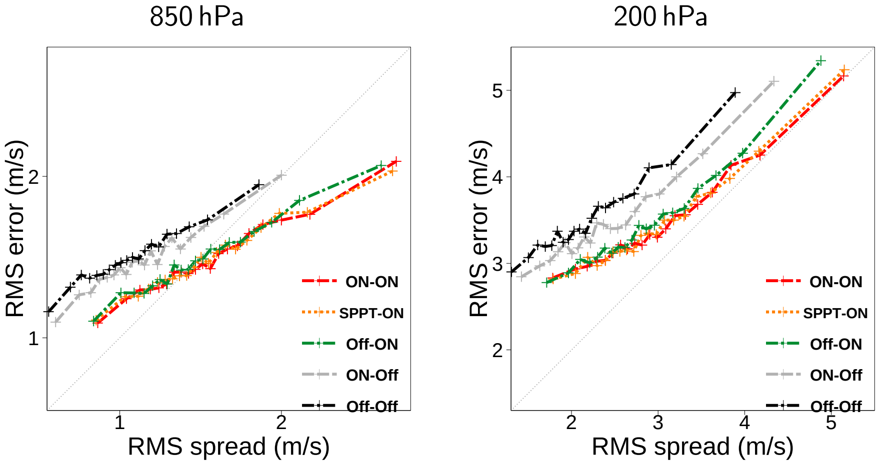
Tropics 20°S–20°N





Spread-reliability: meridional wind $t = 48$ h

Tropics 20°S–20°N



Conclusions

- Stochastic tendency perturbations used in the operational ECMWF ensembles contribute significantly to ensemble spread and improve probabilistic skill.
- Improved ensemble forecast variances and improved probabilistic skill through a combination of
  - introduction of EDA perturbations
  - reduced amplitude for SV perturbations
  - more active representation of model uncertainties

Conclusions

- Stochastic tendency perturbations used in the operational ECMWF ensembles contribute significantly to ensemble spread and improve probabilistic skill.
- Improved ensemble forecast variances and improved probabilistic skill through a combination of
  - introduction of EDA perturbations
  - reduced amplitude for SV perturbations
  - more active representation of model uncertainties
- Not having precise estimates of initial error covariances hampers diagnostic of the characteristics of (random) model tendency errors
- Diagnostics may need to be improved to distinguish well different representations of model uncertainty.
- Model uncertainty contributes to initial uncertainty whereever a short-range forecast is used as prior information. A consistent representation of model uncertainty in data assimilation and forecast can help to better constrain the formulation of model uncertainty.

Plans

- Compare operational schemes with more basic tendency perturbations. For instance, additive noise, e.g. from scaled tendencies constructed from a tendency archive, (e.g. YOTC data)
- Diagnose tendency differences from different models started from the same initial conditions (resolution, different parameters, different parameterization schemes, . . .). What is the nature of the *random component* of the differences?
- Develop improved diagnostics that permit to evaluate better the realism of different tendency perturbations.