Numerical discretization causing variance loss and the need for inflation

by Richard Ménard, Sergey Skachko
Environment and Climate Change Canada
Dorval, Quebec, CANADA

and Olivier Pannekoucke
INPT-ENM Toulouse, CNRM Météo-France, CERFACS Toulouse

2021 EnKF workshop
(virtual meeting), June 7, 2021

Motivation

- The transport of information from one observation time to the next is important for an optimal data assimilation system
- **To what extent the numerical discretization of the model influences the forecast uncertainty ?**

Transport of passive tracer

Eulerian coordinates (x,t)

$$\frac{\partial q}{\partial t} + \mathbf{V} \cdot \nabla q = 0$$

q passive tracer, e.g. mass mixing ratio of an air constituent

Lagrangian coordinates

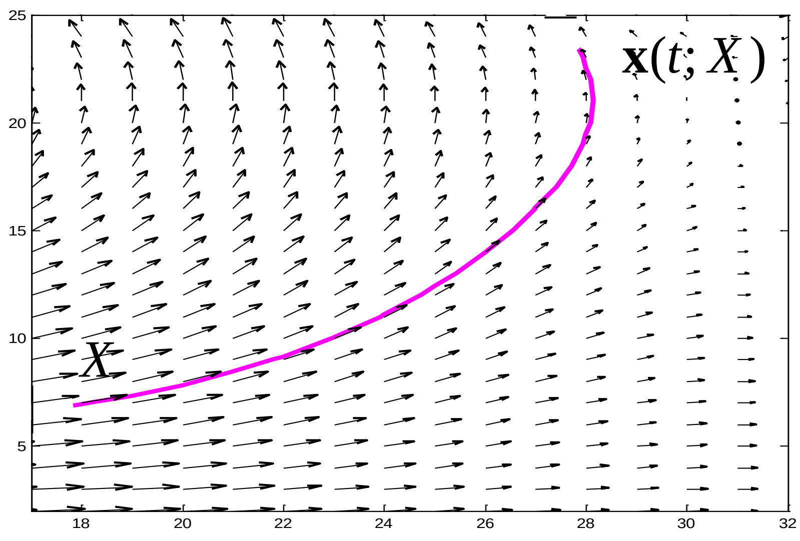
Let X represent the position of fluid particles at $t=0$, and moving with the flow

$$\mathbf{x}(t; X) = \Phi_t(X)$$

The flow Φ_t is given by the winds $\mathbf{V}$

$$\frac{d\,\mathbf{x}}{dt} = \mathbf{V}(\mathbf{x},t)$$

$$\mathbf{x}(t) = \int_0^t \mathbf{V}(\mathbf{x}(\tau),\, \tau)\; d\tau$$



Conservative properties in Lagrangian coordinates

$$\frac{Dq}{Dt} = 0 \quad \Rightarrow \quad q(\mathbf{x}(t;X),t) = q(X,0)$$

Assuming that the flow is known (i.e. deterministic) makes the advection a linear problem.

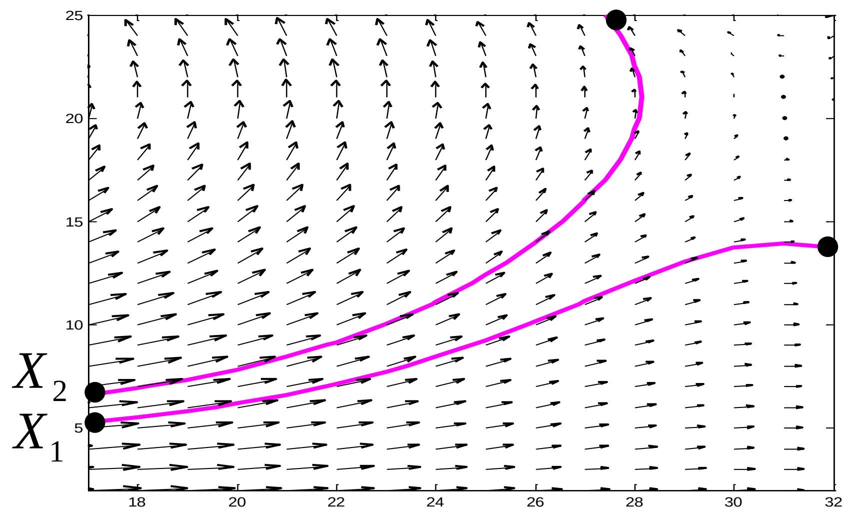
Consider an ensemble of N_e initial Tracer values for each fluid particles

$$\{\, q_i(X,0) \mid i = 1,\ldots,N_e \,\}$$

The ensemble mean is conserved,

$$\overline{q}(\mathbf{x}(t;X),t) = \overline{q}(X,0) \ ,$$

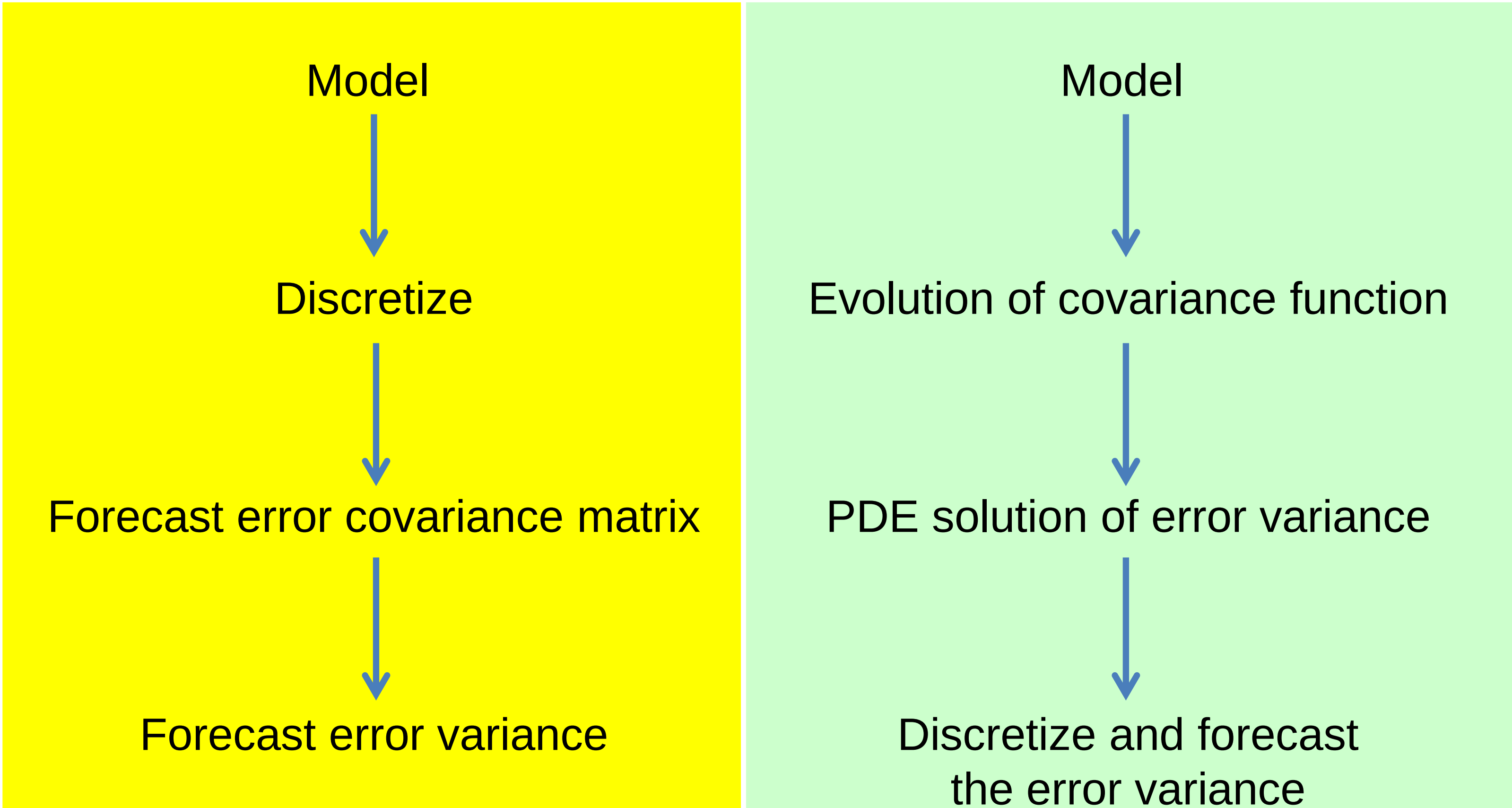
covariance between a pair particles is also conserved,



$$\mathrm{cov}\big(\, q(X_1,t),\ q(X_2,t)\,\big) = \mathrm{cov}\big(\, q(X_1,0),\ q(X_2,0)\,\big)\,,$$

in fact the full pdf , p , is conserved in the Lagrangian coordinates. A nice property it we were to apply it to particle filters.

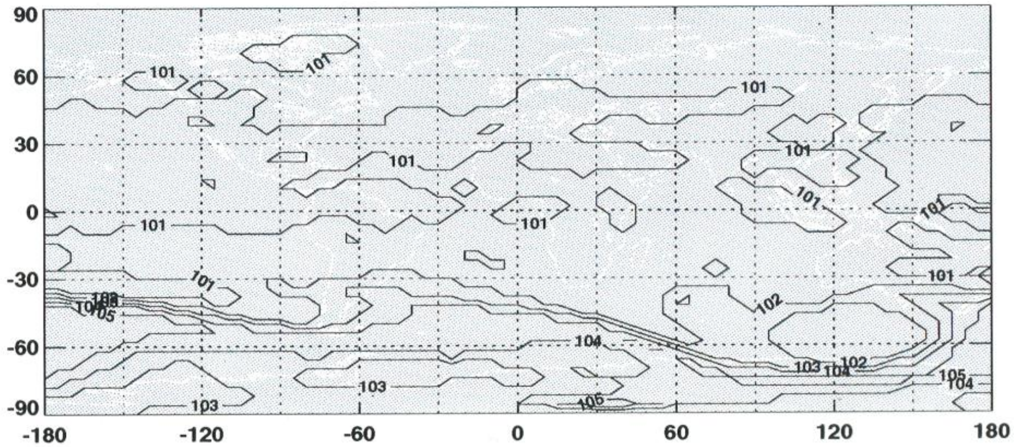
$$p_{\mathbf{q}}\big(\, q(X_1,t),\ \ldots,\ q(X_n,t)\big) = p_{\mathbf{q}}\big(\, q(X_1,0),\ \ldots,\ q(X_n,0)\big)$$


Model
Discretize
Forecast error covariance matrix
Forecast error variance
Model
Evolution of covariance function
PDE solution of error variance
Discretize and forecast
the error variance

Same numerical model to evolve V and calculate P_{ii}

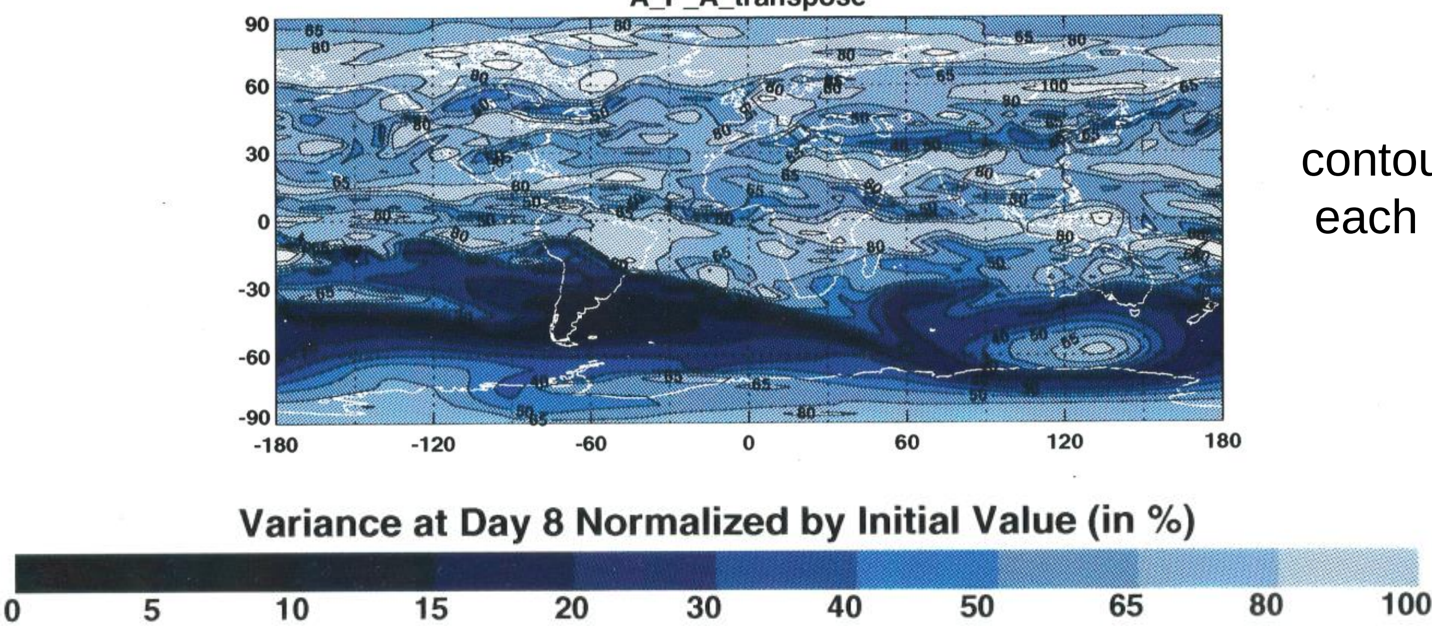
- After 8 day of transport

evolving V (Lagrangian solution)



contours are for each 1%

A_P_A_transpose



contours are for each 10-20 %

Consider 1D advection uniform wind. Donor cell upstream or semi-lagrangian

$$q_i^{n+1} = (1-\alpha)\, q_i^n + \alpha\, q_{i-1}^n$$

where $\alpha = u\,\Delta t / \Delta x$ is the Courant number.

Time stepping of the error covariance $P^n\big(x_1(i),\, x_2(j)\big) = \left\langle \tilde{q}_i^n\, \tilde{q}_j^n \right\rangle$ is done by *in two steps*

1) updating with respect to the x_1 coordinate

$$P^*(x_1(i), x_2(j)) = \text{transport of } P^n(x_1, x_2) \text{ with respect to } x_1$$

2) updating with respect to the x_2 coordinate

$$P^{n+1}(x_1(i), x_2(j)) = \text{transport of } P^*(x_1, x_2) \text{ with respect to } x_2$$

Remark: In matrix form, the result of step 1 and 2 is equivalent to $\mathbf{P}^{n+1} = \mathbf{M}\mathbf{P}^n\mathbf{M}^T$

Specifically for the 1D advection transport scheme above

$$P^*\big(x_1(i), x_2(j)\big) = P^n\big(x_1(i), x_2(j)\big) + \alpha\left\{ P^n\big(x_1(i-1), x_2(j)\big) - P^n\big(x_1(i), x_2(j)\big)\right\}$$

$$P^{n+1}\big(x_1(i), x_2(j)\big) = P^*\big(x_1(i), x_2(j)\big) + \alpha\left\{ P^*\big(x_1(i), x_2(j-1)\big) - P^*\big(x_1(i), x_2(j)\big)\right\}$$

Combining those two equations and writing the result for $i = j$ (the same point) i.e. the *error variance*

$$P^{n+1}\left(x_1(i), x_2(i)\right) \stackrel{\Delta}{=} P_{ii}^{n+1} = (1-\alpha)^2 P_{ii}^n + \alpha(\alpha-1) P_{i-1,\,i}^n + \alpha(\alpha-1) P_{i,\,i-1}^n + \alpha^2 P_{i-1,i-1}^n \quad (1)$$

Whereas from the Lagrangian description, the error variance should simply be advected

$$V_i^{n+1} = (1-\alpha) V_i^n + \alpha V_{i-1}^n \quad (2)$$

The difference, equation (2) minus equation (1), is

$$-\alpha(\alpha-1)\left\{P_{ii}^n - P_{i,i-1}^n - P_{i-1,i}^n + P_{i-1,i-1}^n\right\} \quad (3)$$

a *net loss of error variance*

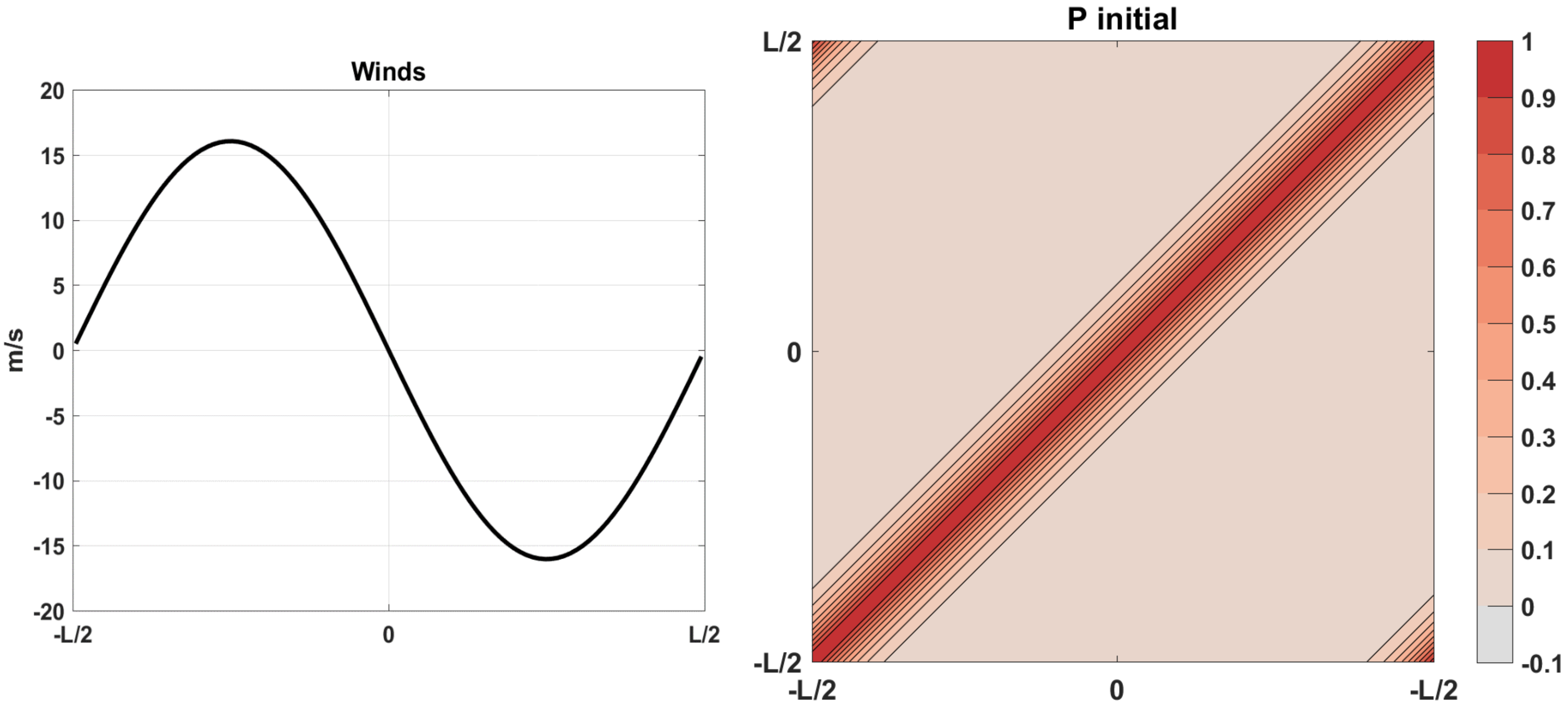
To recover the variance predicted by a Lagrangian description, the error variance obtained from an Eulerian model need to be increased by the deficit (3), i.e.

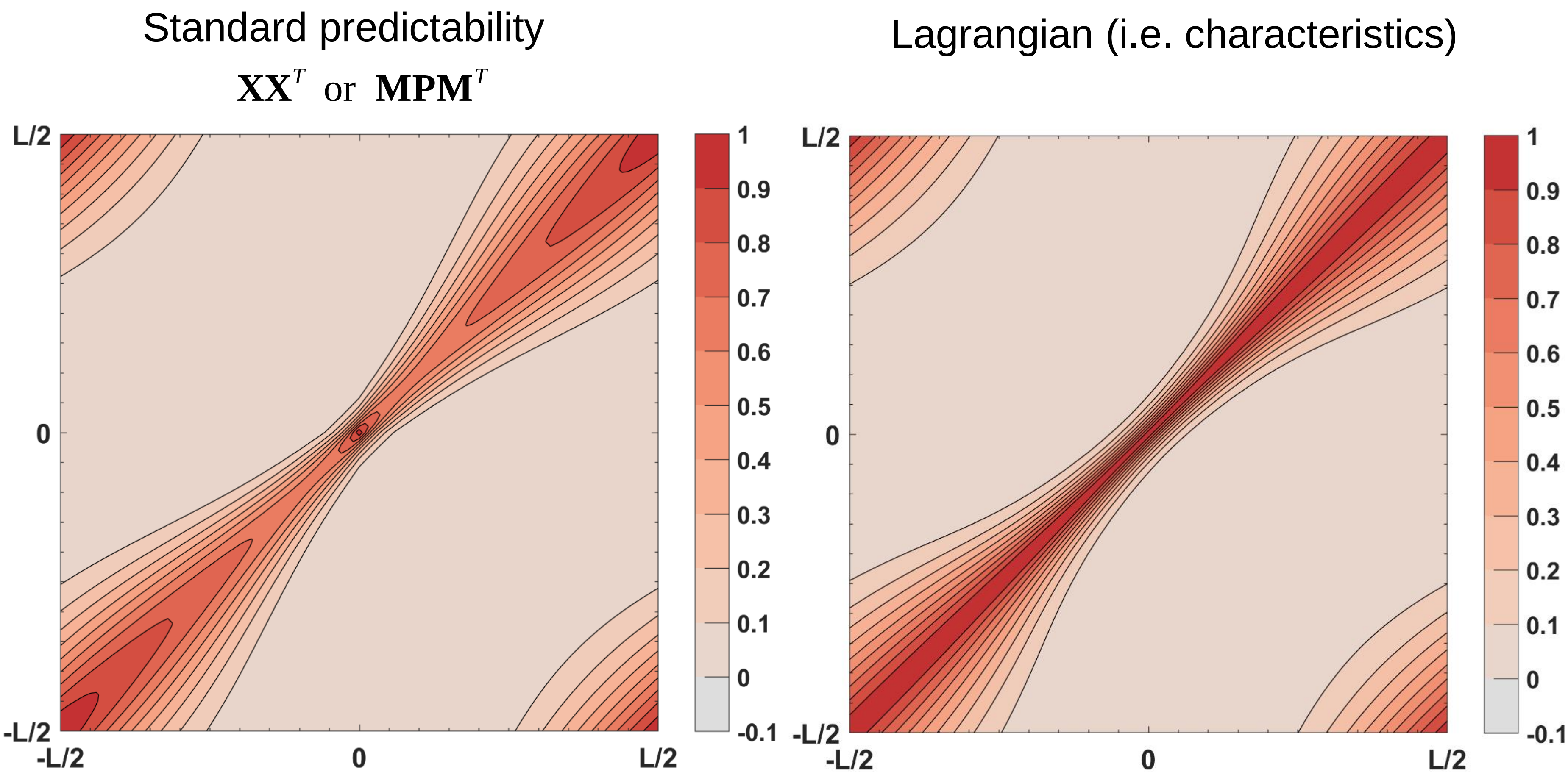
$$\alpha(\alpha-1)\left\{P_{ii}^n - P_{i,i-1}^n - P_{i-1,i}^n + P_{i-1,i-1}^n\right\}$$

This is an explicit form for inflation.

here we can see that inflation (correction) also depends on the courant number and the correlation length

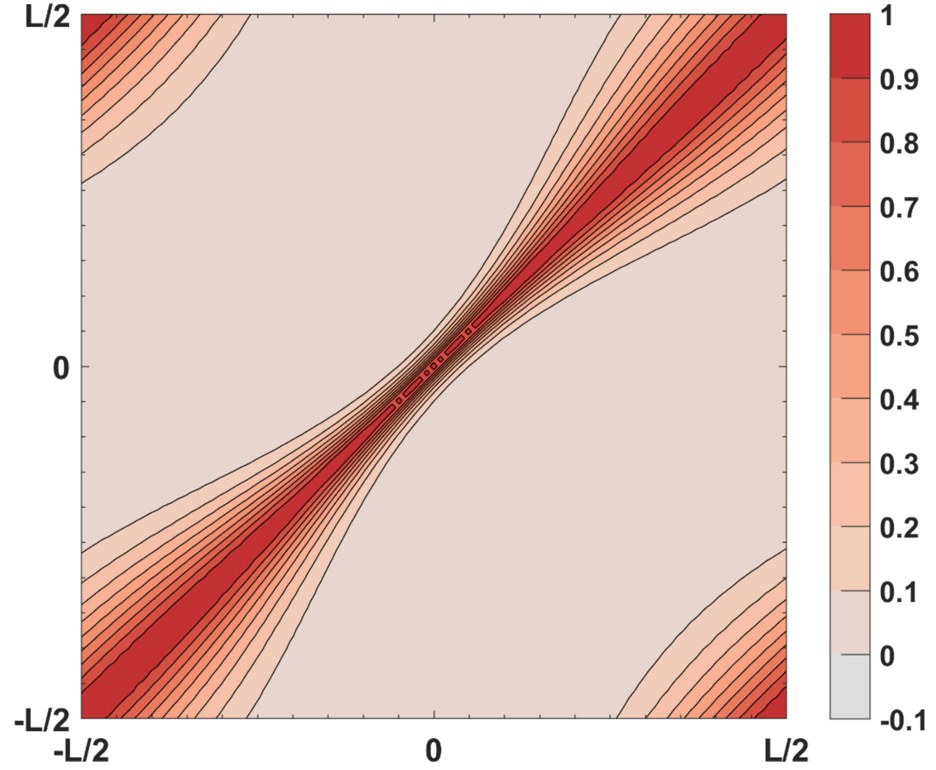
Advection by a non-uniform (stationary) wind
Initial correlation : SOAR with Lc = 400 km





Standard predictability
XX^T or MPM^T
Lagrangian (i.e. characteristics)
L/2
0
-L/2
-L/2
0
L/2
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-0.1

Using local covariance coordinates



Propagation in spatial coordinates

$$\frac{\partial P}{\partial t} + u(x_1)\,\frac{\partial P}{\partial x_1} = 0$$

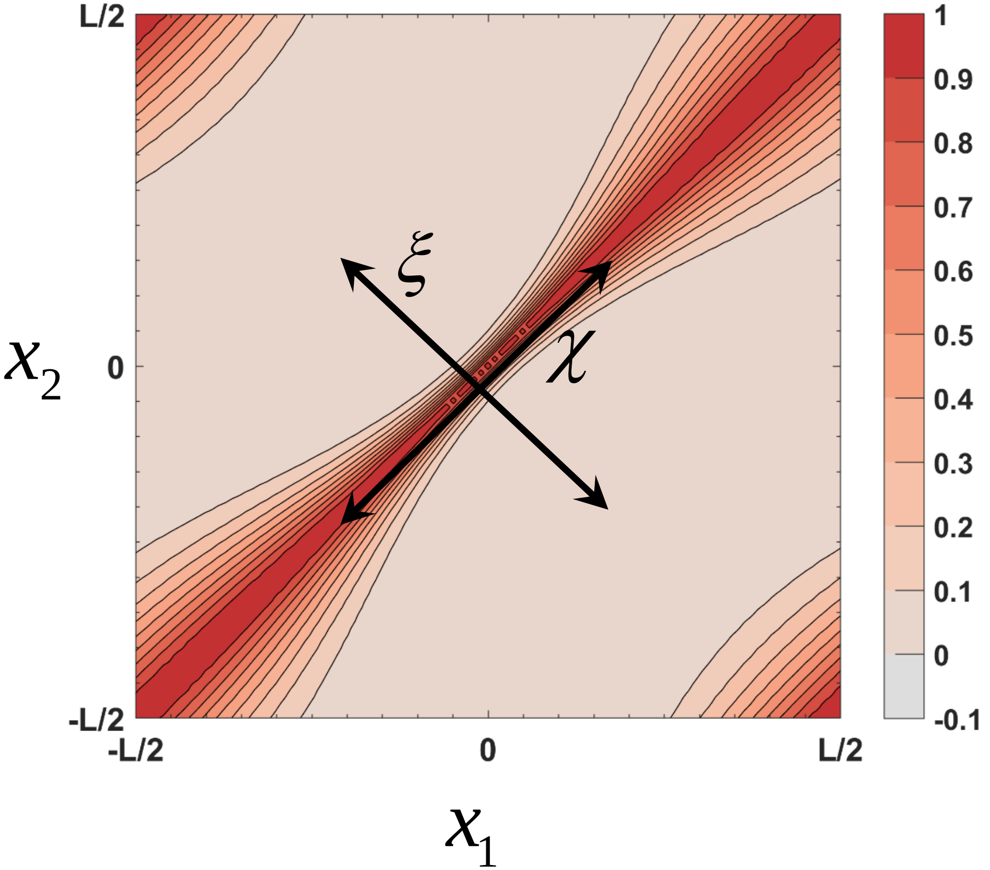
$$\frac{\partial P}{\partial t} + u(x_2)\frac{\partial P}{\partial x_2} = 0$$

Propagation in local coordinates

$$\frac{\partial \tilde{P}}{\partial t} + \frac{1}{2}\left[u(\chi+\xi)+u(\chi-\xi)\right]\frac{\partial \tilde{P}}{\partial \chi} = 0$$

$$\frac{\partial \tilde{P}}{\partial t} + \frac{1}{2}\left[u(\chi+\xi)-u(\chi-\xi)\right]\frac{\partial \tilde{P}}{\partial \xi} = 0$$

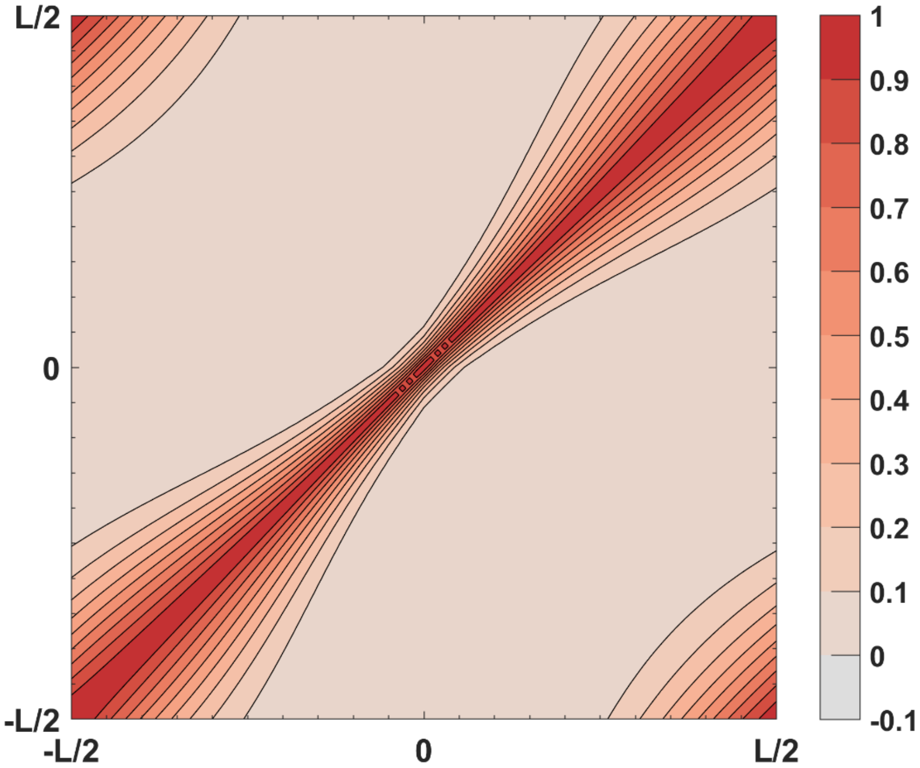
Using local covariance coordinates



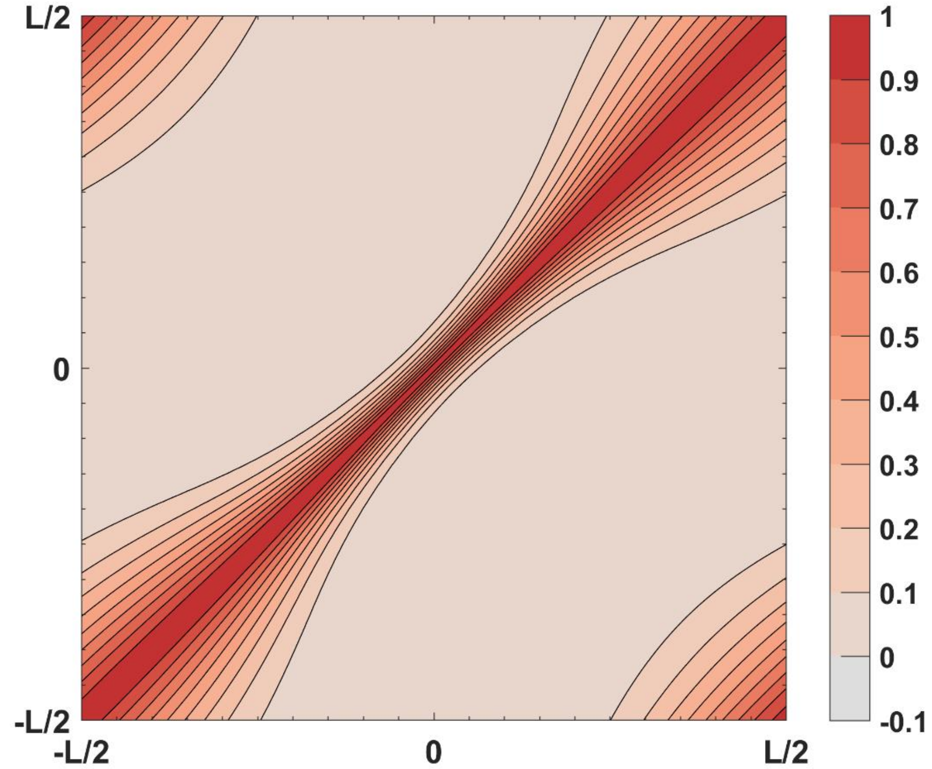
Standard predictability $\mathbf{XX}^T$ or $\mathbf{MPM}^T$

+ inflation of variance

$$\alpha(\alpha-1)\left\{P_{ii}^{n}-P_{i,i-1}^{n}-P_{i-1,i}^{n}+P_{i-1,i-1}^{n}\right\}$$

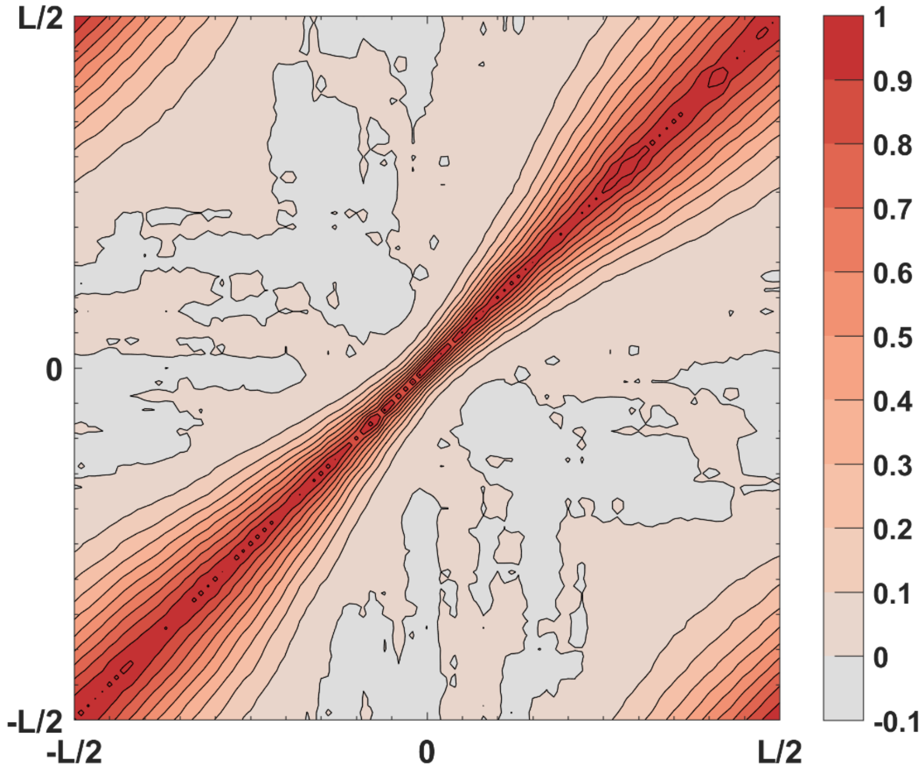


Lagrangian (i.e. characteristics)

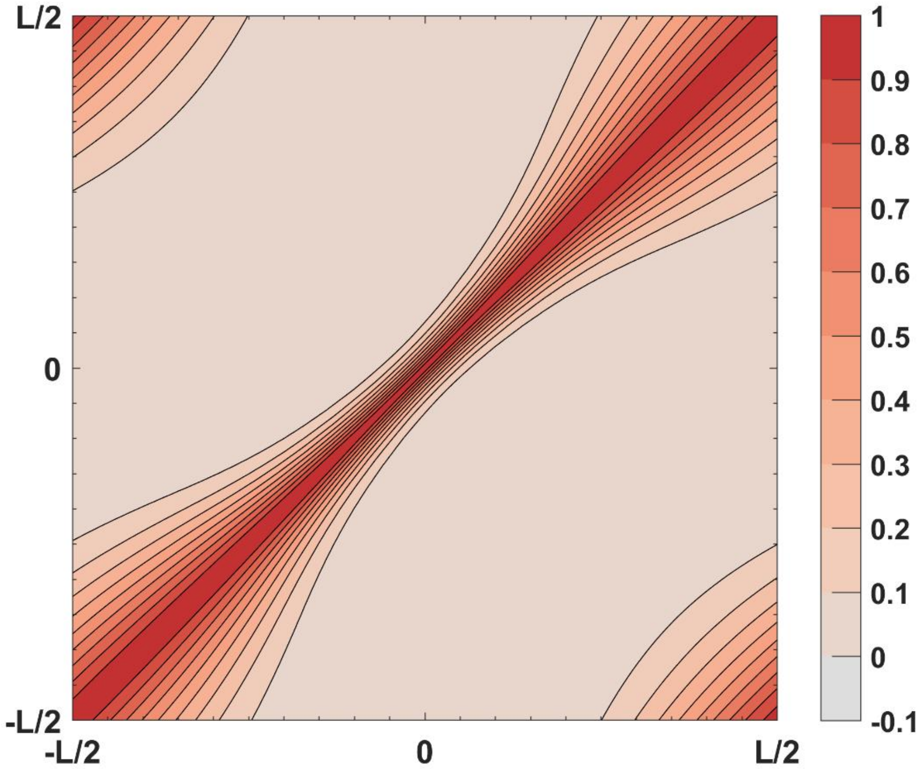


Standard predictability $\mathbf{XX}^T$ or $\mathbf{MPM}^T$
+ stochastic inflation of state

$$+\sqrt{|\alpha_i^n|(1-|\alpha_i^n|)}\,\left(q_i^n - q_{i-1}^n\right)\mathrm{U}_i \quad \text{for } \alpha_i^n > 0$$

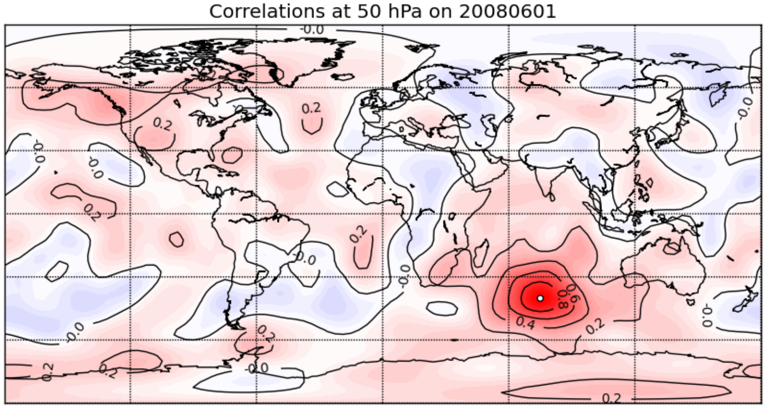


Lagrangian (i.e. characteristics)

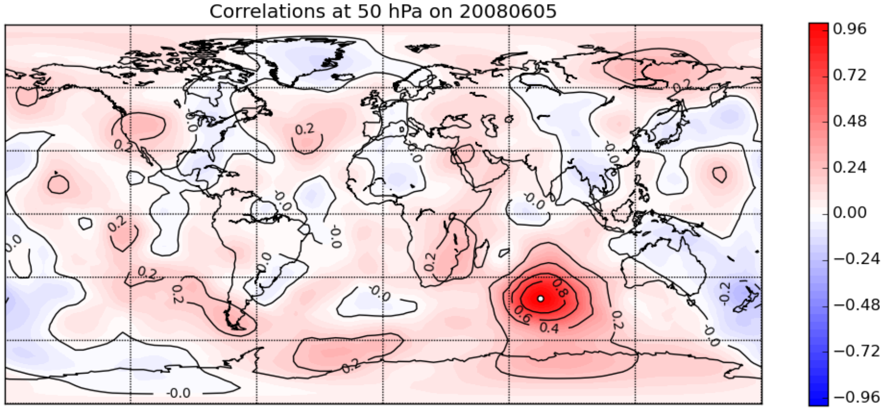


3D CTM - Solid body rotation winds

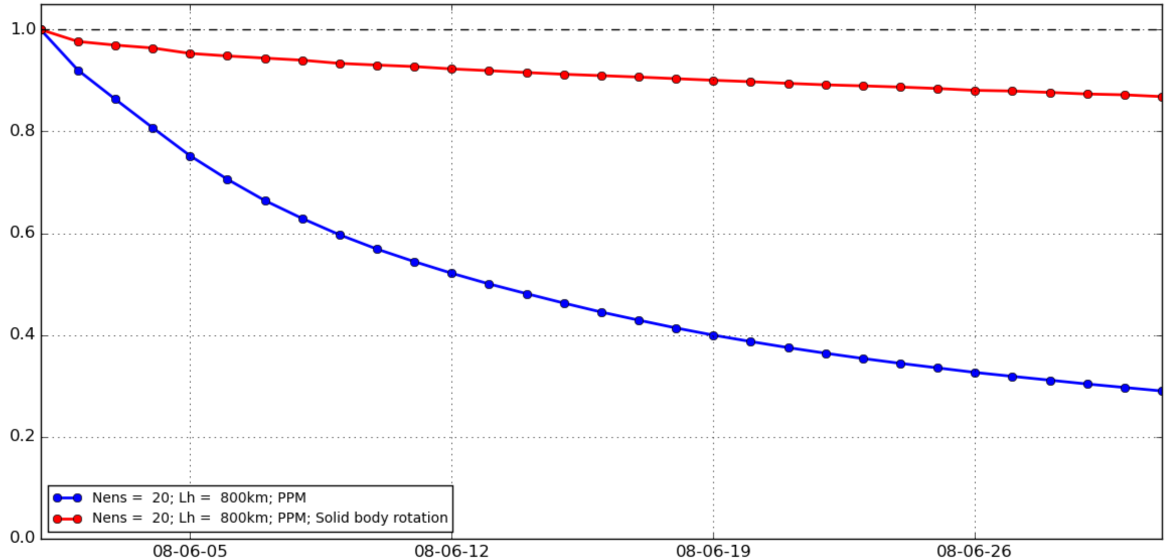
Initial error correlation



Error correlation after 4 days

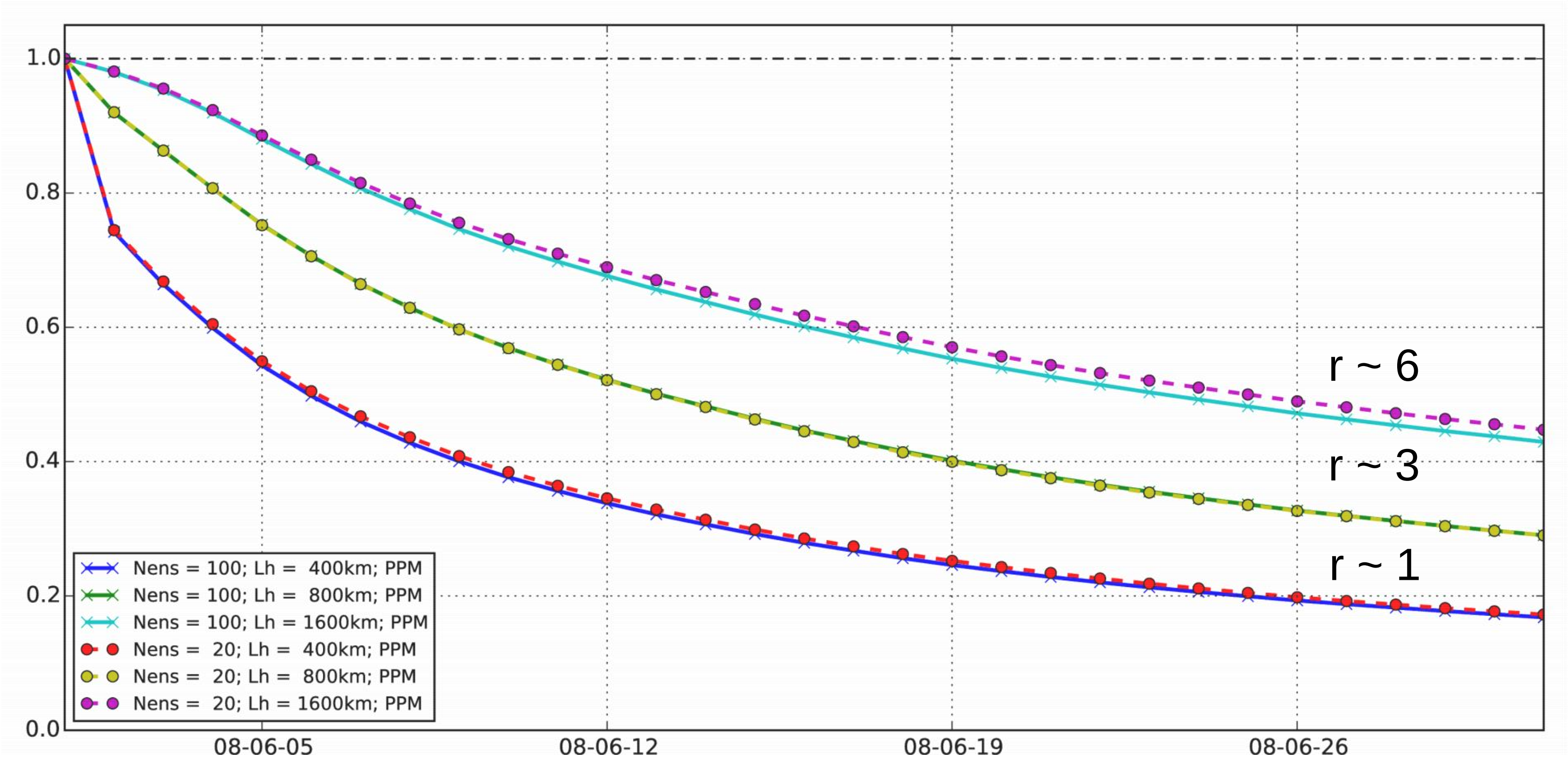


Variance loss: **solid body rotation winds**, **ERA interim meteorology**



3D CTM - ERA Interim meteorology

B/A : Ratio of the EnKF error variance with the advection of error variance

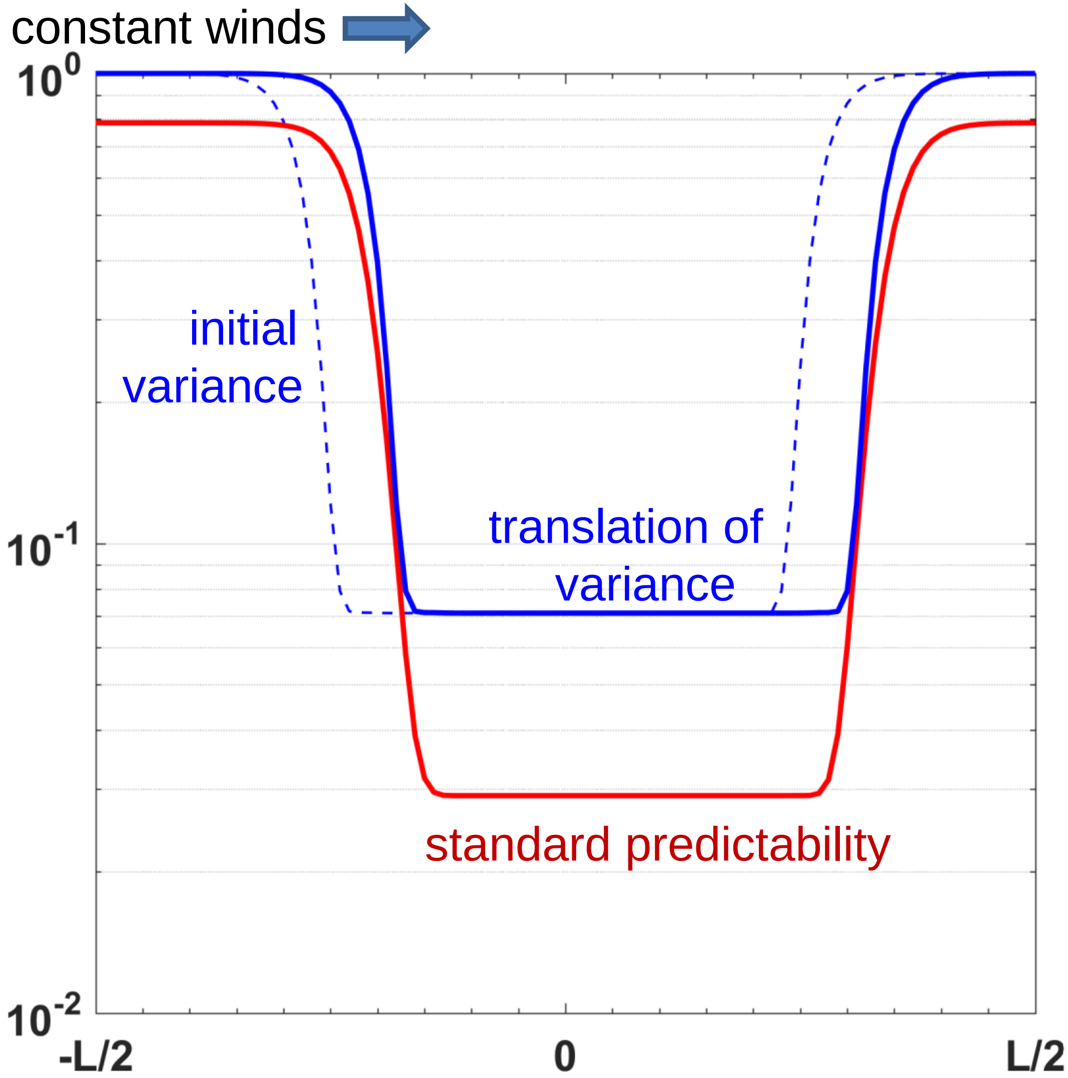


- The variance loss depends on the ratio of **r = correlation length / model resolution**
- Differences with small sample size (Nens = 20) is due to sampling errors

In which circumstances the variance loss is the largest ?

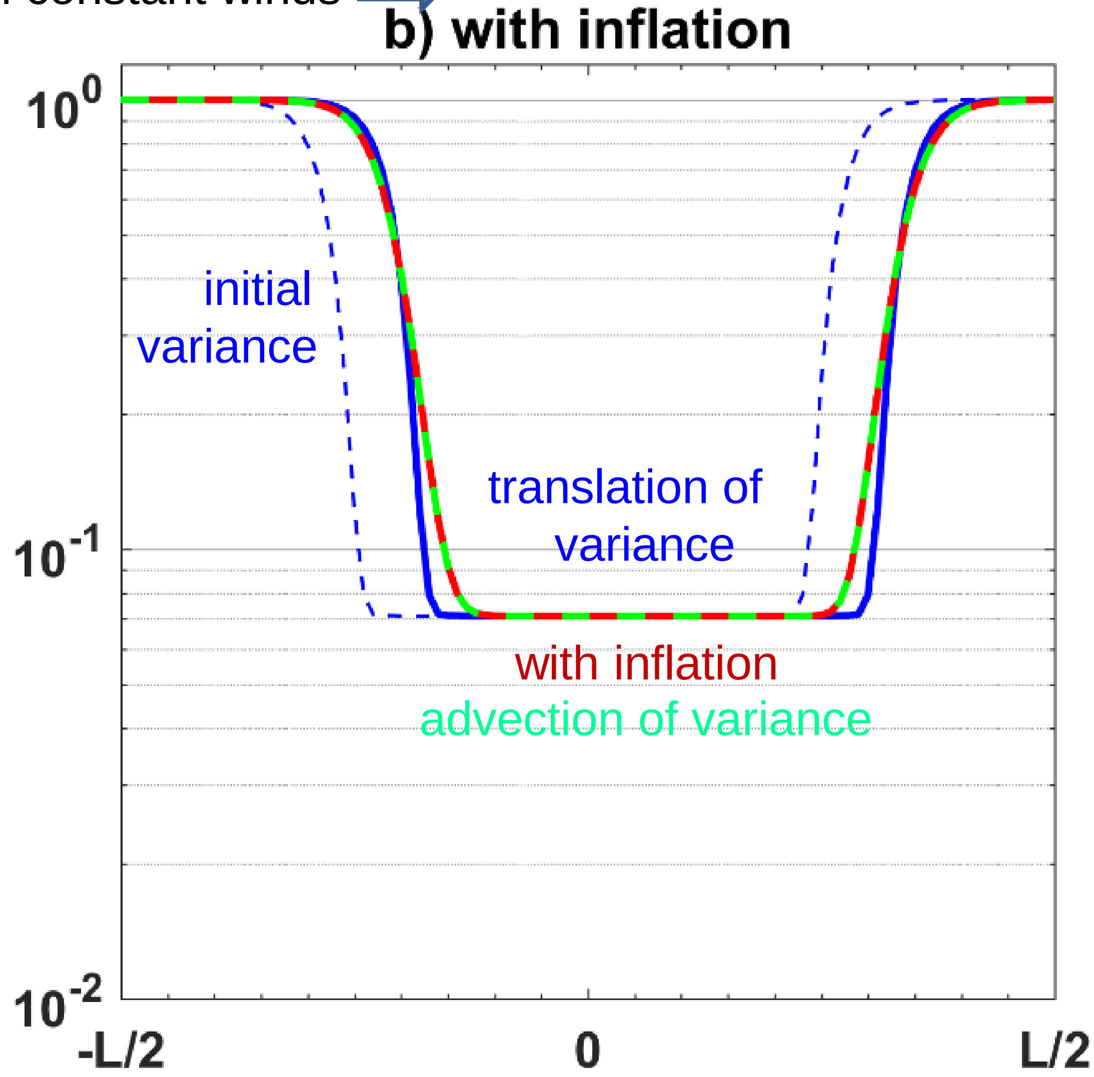
Over a region with dense observations

Case: uniform constant winds



And if we use inflation (due to numerical discretization error)

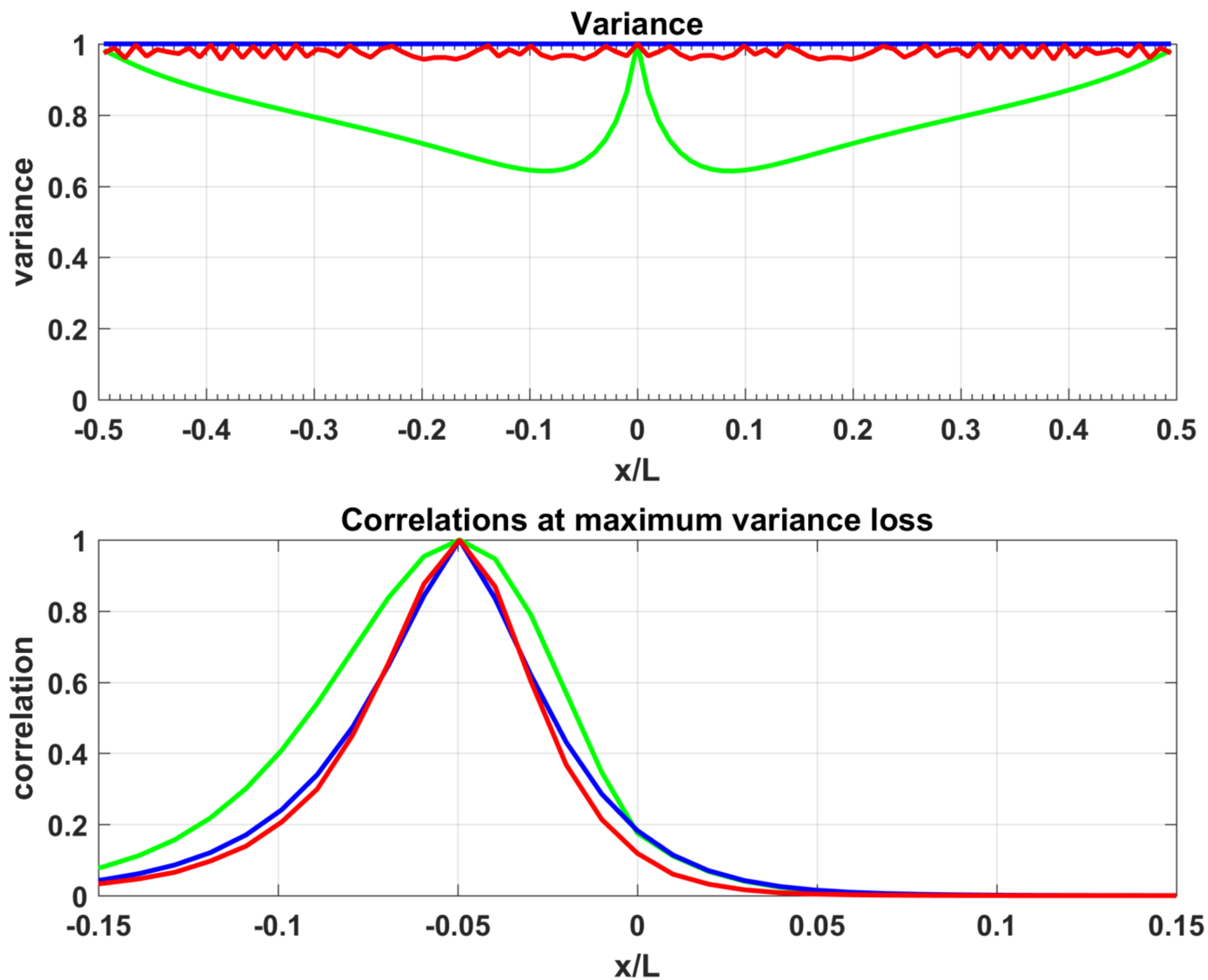
Case: uniform constant winds



Conclusions

- In addition to the variance loss we observe in ensemble DA, there exist a new loss of variance due to the model discretization error that appears in both KF and ensemble DA
- The variance loss depends on the way we obtain the error covariance, e.g. standard, local coordinate, Lagrangian
- We derived an inflation based on the variance loss of a first order discretization (either to rescale the variance or as a stochastic term)
- The variance loss is largest when the correlation length is smallest, e.g. over a densely observed region (with spatially uncorrelated obs errors)

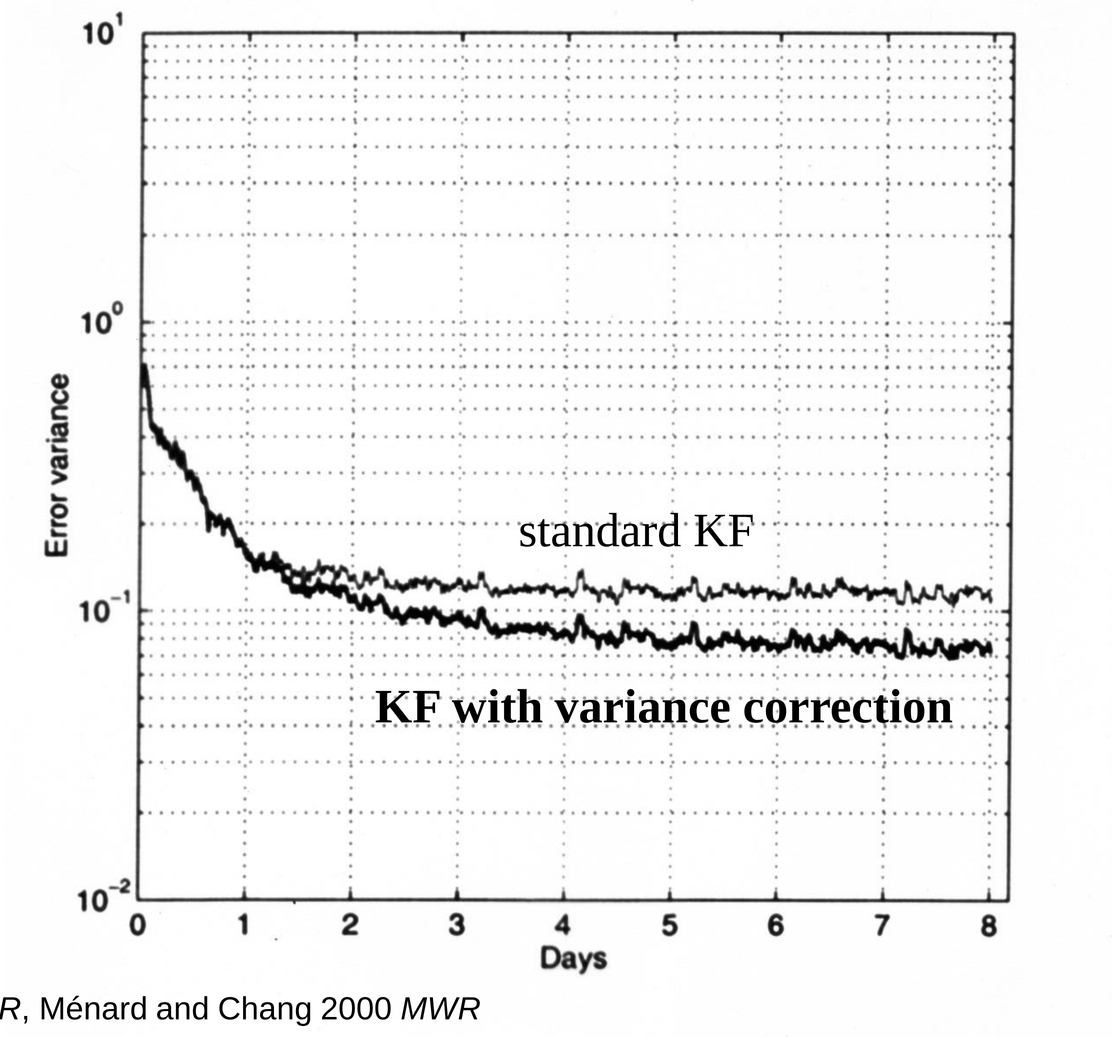
Thanks for your attention



Variance
variance
1
0.8
0.6
0.4
0.2
0
-0.5
-0.4
-0.3
-0.2
-0.1
0
0.1
0.2
0.3
0.4
0.5
x/L
Correlations at maximum variance loss
correlation
1
0.8
0.6
0.4
0.2
0
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
x/L

Motivation

- When using different numerical schemes the model error variance t can differ by a factor 2 (with a modestly dense observations)
- Analysis error variance can also be different by a factor 2



Ménard et al 2000 *MWR*, Ménard and Chang 2000 *MWR*