A methodology to obtain model-error covariances due to the discretization scheme from the parametric Kalman filter perspective

O. Pannekoucke[1,2,3], R. Menard[4], M. El Aabaribaoune[2,3], M. Plu[2]

[1]CNRM INPT-ENM, France. [2]CNRM INPT-ENM, France. [3]CERFACS, France. [4]Environment and Climate Change Canada, Canada.

2021EnKF workshop | Norway | 7-11 June 2021

METEO FRANCE CERFACS cnrs lefe Environment Canada

Model-error for the advection

As a simple example, consider that the nature is given by the advection

$$\partial_t c + \mathbf{u}\partial_x c = 0, \tag{1}$$

where $\mathbf{u}(t,x) > 0$ is an heterogeneous wind field and $c(t,x)$ a passive scalar field.

Model-error for the advection

As a simple example, consider that the nature is given by the advection

$$\partial_t c + \mathbf{u}\partial_x c = 0, \tag{1}$$

where $\mathbf{u}(t,x) > 0$ is an heterogeneous wind field and $c(t,x)$ a passive scalar field.
Note that c is conserved along the characteristic curves (see *e.g.* [Boyd, 2001, chap. 14]), that is

$$\frac{dx}{dt} = u, \tag{2a}$$

$$\frac{dc}{dt} = 0. \tag{2b}$$

Model-error for the advection

As a simple example, consider that the nature is given by the advection

$$\partial_t c + \mathbf{u}\partial_x c = 0, \tag{1}$$

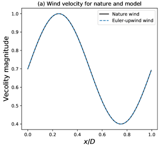
where $\mathbf{u}(t, x) > 0$ is an heterogeneous wind field and $c(t, x)$ a passive scalar field.
Note that *c* is conserved along the characteristic curves (see *e.g.* [Boyd, 2001, chap. 14]), that is

$$\frac{dx}{dt} = u, \tag{2a}$$

$$\frac{dc}{dt} = 0. \tag{2b}$$

Now suppose the dynamics is numerically solved by the Euler-upwind scheme,

$$\frac{c_i^{q+1} - c_i^q}{\delta t} = -u_i \frac{c_i^q - c_{i-1}^q}{\delta x}, \tag{3}$$

$\mathbf{u}(x) \rightarrow$ 

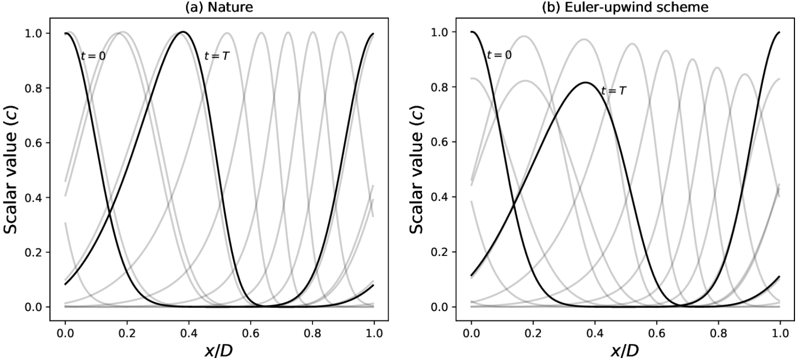


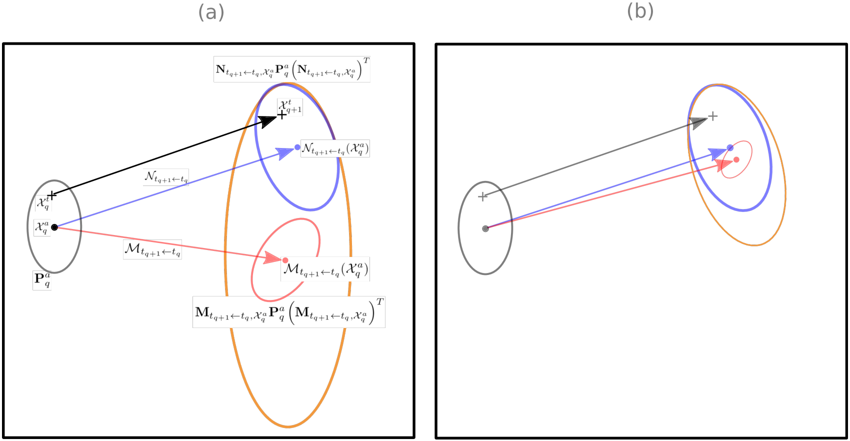
Figure: Nature versus numerical dynamics

Transport with conservation for the nature
but heterogeneous damping for the num. model == model error.

Table of contents

Dynamics of the model-error covariance

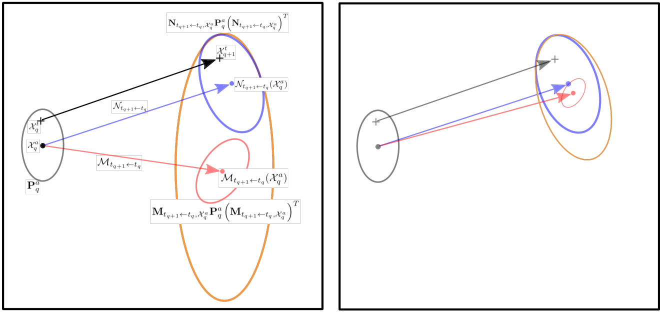
Sketch of the uncertainty evolution in presence of model-error



$\mathcal{N}$ (resp. $\mathcal{M}$) denotes the nature (the numerical model). The orange ellipse represents the forecast-error covariance matrix $\mathbf{P}_{q+1}^f$.

Modelling of the model-error covariance

(a) (b)



In the diffusive case (b) the forecast-error covariance matrix writes [Pannekoucke et al., 2021]

$$\mathbf{P}^f{}_{q+1} \approx \mathbf{P}^p_{q+1} + \Pi^m_{q+1} + \mathbf{Q}_{q+1}, \tag{4}$$

with $\mathbf{P}^p_{q+1} = \mathbf{M}\mathbf{P}^a{}_q\mathbf{M}^T$ the predictability-error covariance matrix and

$$\Pi^m_{q+1} = \mathbf{N}\mathbf{P}^a{}_q\mathbf{N}^T - \mathbf{M}\mathbf{P}^a{}_q\mathbf{M}^T. \tag{5}$$

Dynamics of the model-error covariance

With $\mathbf{P}^f{}_{q+1} \approx \mathbf{P}^p_{q+1} + \mathbf{P}^m_{q+1}$, it results that the model-error covariance matrix

$$\mathbf{P}^m_{q+1} = \Pi^m_{q+1} + \mathbf{Q}_{q+1}, \tag{6}$$

expands as a

- flow-dependent part $\Pi^m_{q+1} = \mathbf{N}\mathbf{P}^a{}_q\mathbf{N}^T - \mathbf{M}\mathbf{P}^a{}_q\mathbf{M}^T$,
- climatological part $\mathbf{Q}_{q+1}$.

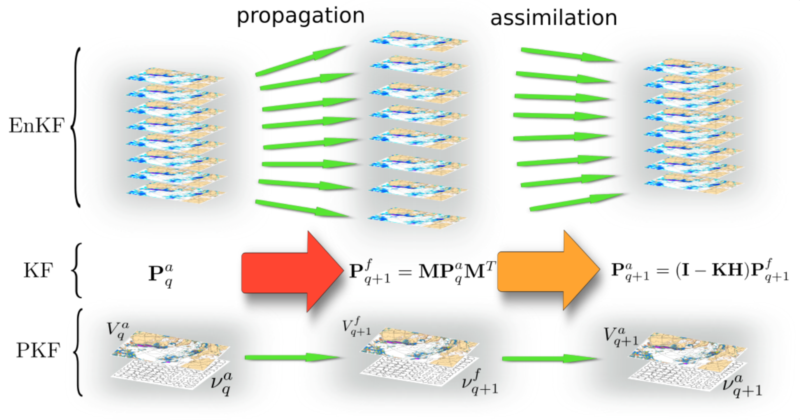
So to focuse on the flow-dependent part, we need to compute the predictability-error covariance matrices $\mathbf{N}\mathbf{P}^a{}_q\mathbf{N}^T$ and $\mathbf{M}\mathbf{P}^a{}_q\mathbf{M}^T$.

From $\mathbf{M}$ we can compute $\mathbf{M}\mathbf{P}^a{}_q\mathbf{M}^T$ from an ensemble. But in practice since we don't have $\mathbf{N}$, so we cannot compute $\mathbf{N}\mathbf{P}^a{}_q\mathbf{N}^T$.

We propose to performed the computation by using the parametric approach.

Table of contents

Parametric Kalman Filter



What are the PKF equations for the forecast steps ?

VLAT covariance matrices

In this talk we consider covariance models parameterized by the variance and the local anistropy tensor fields – the VLATcov model [Pannekoucke, 2021]. For an error field $\varepsilon(t,x)$,

- the variance is defined as $V(t,x) = \mathbb{E}\left[\varepsilon^2\right]$

VLAT covariance matrices

In this talk we consider covariance models parameterized by the variance and the local anistropy tensor fields – the VLATcov model [Pannekoucke, 2021]. For an error field $\varepsilon(t,x)$,

- the variance is defined as $V(t,x) = \mathbb{E}\left[\varepsilon^2\right]$
- the local anisotropy tensor is given either by the metric tensor, $\boldsymbol{g}(t,x)$, which measures the anisotropy of the correlation function

$$\rho(t,x,x+\delta x) = \frac{\mathbb{E}\left[\varepsilon(t,x)\varepsilon(t,x+\delta x)\right]}{\sqrt{V_x V_{x+\delta x}}} \underset{\delta x \to 0}{=} 1 - \frac{1}{2}||\delta x||^2_{\boldsymbol{g}_x} + \mathcal{O}(\delta x^2),$$

VLAT covariance matrices

In this talk we consider covariance models parameterized by the variance and the local anistropy tensor fields – the VLATcov model [Pannekoucke, 2021]. For an error field $\varepsilon(t,x)$,

- the variance is defined as $V(t,x) = \mathbb{E}\left[\varepsilon^2\right]$
- the local anisotropy tensor is given either by the metric tensor, $\boldsymbol{g}(t,x)$, which measures the anisotropy of the correlation function

$$\rho(t,x,x+\delta x) = \frac{\mathbb{E}\left[\varepsilon(t,x)\varepsilon(t,x+\delta x)\right]}{\sqrt{V_x V_{x+\delta x}}} \underset{\delta x \to 0}{=} 1 - \frac{1}{2}||\delta x||^2_{\boldsymbol{g}_x} + \mathcal{O}(\delta x^2),$$

or the the aspect tensor, $\boldsymbol{s}(t,x)$, which is the matrix inverse of the metric tensor

$$\boldsymbol{s}_x = \boldsymbol{g}_x^{-1}.$$

VLAT covariance matrices

In this talk we consider covariance models parameterized by the variance and the local anistropy tensor fields – the VLATcov model [Pannekoucke, 2021]. For an error field $\varepsilon(t,x)$,

- the variance is defined as $V(t,x) = \mathbb{E}\left[\varepsilon^2\right]$
- the local anisotropy tensor is given either by the metric tensor, $\boldsymbol{g}(t,x)$, which measures the anisotropy of the correlation function

$$\rho(t,x,x+\delta x) = \frac{\mathbb{E}\left[\varepsilon(t,x)\varepsilon(t,x+\delta x)\right]}{\sqrt{V_x V_{x+\delta x}}} \underset{\delta x \to 0}{=} 1 - \frac{1}{2}||\delta x||^2_{\boldsymbol{g}_x} + \mathcal{O}(\delta x^2),$$

or the the aspect tensor, $\boldsymbol{s}(t,x)$, which is the matrix inverse of the metric tensor

$$\boldsymbol{s}_x = \boldsymbol{g}_x^{-1}.$$

Note that $(\boldsymbol{g}_x)_{ij} = \mathbb{E}\left[\partial_i\left(\frac{\varepsilon}{\sqrt{V}}\right)\partial_j\left(\frac{\varepsilon}{\sqrt{V}}\right)\right].$

PKF for VLATcov models

The parametric Kalman filter dynamics for the variance and the local anisotropy writes as

$$\partial_t V = 2\mathbb{E}\left[\varepsilon \partial_t \varepsilon\right], \tag{7}$$

$$\partial_t \boldsymbol{g} = \partial_t \mathbb{E}\left[\partial_i \left(\frac{\varepsilon}{\sqrt{V}}\right) \partial_j \left(\frac{\varepsilon}{\sqrt{V}}\right)\right], \tag{8}$$

which can be computed by using a computer algebra system.

PKF for VLATcov models

The parametric Kalman filter dynamics for the variance and the local anisotropy writes as

$$\partial_t V = 2\mathbb{E}\left[\varepsilon \partial_t \varepsilon\right], \tag{7}$$

$$\partial_t \boldsymbol{g} = \partial_t \mathbb{E}\left[\partial_i \left(\frac{\varepsilon}{\sqrt{V}}\right) \partial_j \left(\frac{\varepsilon}{\sqrt{V}}\right)\right], \tag{8}$$

which can be computed by using a computer algebra system.

SymPKF performs the symbolic computation of the PKF for VLATcov model and can also automatically generate a finite difference implementation for the numerical exploration [Pannekoucke and Arbogast, 2021].

see https://github.com/opannekoucke/sympkf

Predictability-error covariance dynamics: the nature

The PKF dynamics for the transport equation, over a 1D domain, computed with SymPKF leads to the predictabilty error dynamics for the nature:

$$\partial_t \tilde{c} = -u\partial_x \tilde{c}, \tag{9}$$

$$\partial_t \tilde{V}^p = -u\partial_x \tilde{V}^p, \tag{10}$$

$$\partial_t \tilde{s}^p = -u\partial_x \tilde{s}^p + 2\tilde{s}^p \partial_x u, \tag{11}$$

stands for $\mathbf{N}\mathbf{P}^a{}_q\mathbf{N}^T$, where $\tilde{\cdot}$ denotes the statistics for the nature.

Predictability-error covariance dynamics: the nature

The PKF dynamics for the transport equation, over a 1D domain, computed with SymPKF leads to the predictabilty error dynamics for the nature:

$$\partial_t \tilde{c} = -u\partial_x \tilde{c}, \tag{9}$$

$$\partial_t \tilde{V}^p = -u\partial_x \tilde{V}^p, \tag{10}$$

$$\partial_t \tilde{s}^p = -u\partial_x \tilde{s}^p + 2\tilde{s}^p \partial_x u, \tag{11}$$

stands for $\mathbf{NP}^a{}_q\mathbf{N}^T$, where $\tilde{\cdot}$ denotes the statistics for the nature.

The variance is conserved, while the anisotropy is stretched by the shear (the term $2s\partial_x u$).

Predictability-error covariance dynamics: the model

The modified equation associated with the Euler-upwind scheme

$$\frac{c_i^{q+1}-c_i^q}{\delta t} = -u_i\frac{c_i^q - c_{i-1}^q}{\delta x}, \tag{12}$$

reads as

$$\partial_t C + U\partial_x C = \kappa\partial_x^2 C, \tag{13}$$

where

$$\begin{cases} U(t,x) = u - \frac{\delta t}{2}\partial_t u + \frac{\delta t}{2}u\partial_x u, \\ \kappa(t,x) = \frac{u}{2}\left(\delta x - u\delta t\right). \end{cases} \tag{14}$$

which shows that the num. model is suffering from dispersion and dissipation.

Note that similar expressions are obtained for semi-Lagrangian discretization as used in NWP and air quality.

Predictability-error covariance dynamics: the model

The PKF dynamics for the num. model, computed with SymPKF, reads as

$$
\begin{aligned}
\partial_t c &= -U\partial_x c + \kappa\partial_x^2 c,\\
\partial_t V^p &= U\partial_x V^p - \frac{2V^p\kappa}{s^p} + \kappa\partial_x^2 V^p - \frac{\kappa\left(\partial_x V^p\right)^2}{2V^p}\\
\partial_t s^p &= -U\partial_x s^p + 2\kappa {s^p}^2\mathbb{E}\left(\tilde{\varepsilon^p}\partial_x^4\tilde{\varepsilon^p}\right) - 3\kappa\partial_x^2 s^p - 2\kappa + \frac{6\kappa\left(\partial_x s^p\right)^2}{s^p}\\
&\quad - \frac{2\kappa s^p\partial_x^2 V^p}{V^p} + \frac{\kappa\partial_x V^p\partial_x s^p}{V^p} +\\
&\quad \frac{2\kappa s^p\left(\partial_x V^p\right)^2}{{V^p}^2} + 2s^p\partial_x U + \partial_x\kappa\partial_x s^p - \frac{2s^p\partial_x\kappa\partial_x V^p}{V^p},
\end{aligned}
$$

where $\tilde{\varepsilon^p} = \varepsilon^p/\sqrt{V^p}$ is the normalized error. This stands for $\mathbf{M}{\mathbf{P}^a}_q\mathbf{M}^T$ and appears as a coupled system due to the numerical diffusion which needs a closure

Predictability-error covariance dynamics: the model

With the local Gaussian closure [Pannekoucke et al., 2018] $\mathbb{E}\left(\tilde{\varepsilon^p}\partial_x^4\tilde{\varepsilon^p}\right) = \frac{2\partial_x^2 s^p}{s^{p2}} + \frac{3}{s^{p2}} - \frac{4(\partial_x s^p)^2}{s^{p3}}$, the predictability-error covariance dynamics reads as

$$\partial_t c = -U\partial_x c + \kappa\partial_x^2 c,$$

$$\partial_t V^p = U\partial_x V^p - \frac{2V^p\kappa}{s^p} + \kappa\partial_x^2 V^p - \frac{\kappa\left(\partial_x V^p\right)^2}{2V^p}$$

$$\partial_t s^p = -U\partial_x s^p + (2\partial_x U)s^p +$$

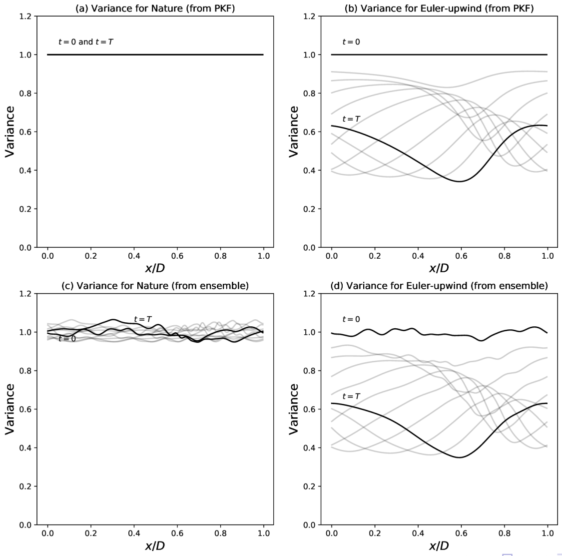
$$\kappa\partial_x^2 s^p + 4\kappa - \frac{2\left(\partial_x s^p\right)^2}{s^p}\kappa + \partial_x\kappa\partial_x s^p - \frac{2\partial_x^2 V^p}{V^p}\kappa s^p +$$

$$\frac{\partial_x V^p}{V}\kappa\partial_x s^p - \frac{2\partial_x V^p}{V^p}s^p\partial_x\kappa + \frac{2\left(\partial_x V^p\right)^2}{V^{p2}}\kappa s^p,$$

Note that closures can be found from data-driven physics considering an hybridization of physics and AI *e.g.* PDE-NetGen [Pannekoucke and Fablet, 2020] see https://github.com/opannekoucke/pdenetgen

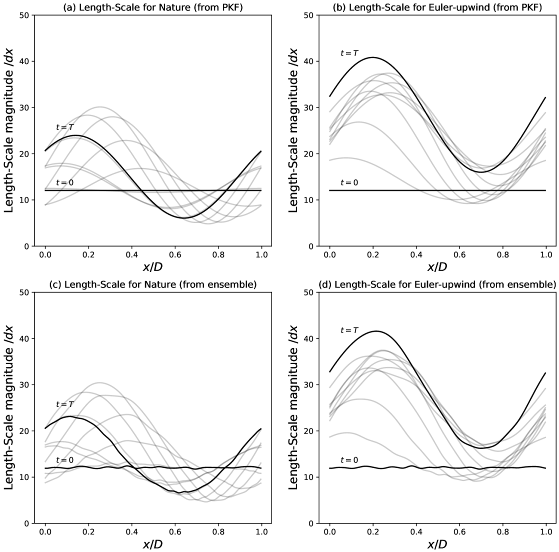
Predictability-error covariance dynamics PKF

(validated by an ensemble estimation with 6400 members, [Evensen, 2009])



Predictability-error covariance dynamics PKF

(with the correlation length-scale defined as $\sqrt{s}$)



Predictability-error covariance dynamics: the model

Note that, using the spatial averaged over the domain, $\langle\cdot\rangle = \frac{1}{D}(\int)\cdot dx$, the predictability-error dynamics for $\mathcal{M}$ approximately reads as

$$\partial_t \langle V^p \rangle = -\frac{\langle 2\kappa \rangle}{\langle s^p \rangle}\langle V^p \rangle, \tag{15}$$

$$\partial_t \langle s^p \rangle = 4\langle \kappa \rangle, \tag{16}$$

of solution

$$\langle V^p \rangle(t) = \langle V^p \rangle(0)\left(\frac{\langle s^p \rangle(0)}{\langle s^p \rangle(0) + 4\langle \kappa \rangle t}\right)^{1/2}, \tag{17}$$

$$\langle s^p \rangle(t) = \langle s^p \rangle(0) + 4\langle \kappa \rangle t. \tag{18}$$

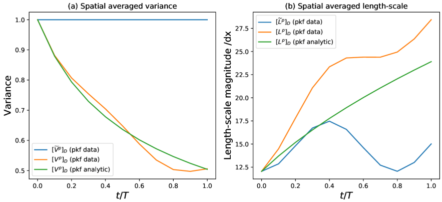
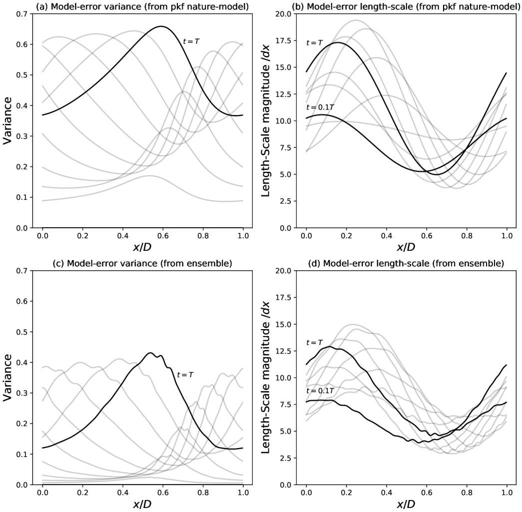


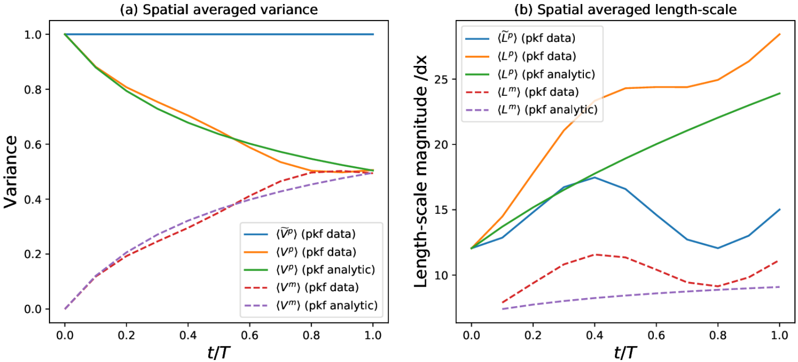
Table of contents

Time evolution of the low-dependent part of $\mathbf{P}^m$



Evolution of the flow-dependent part of the model-error covariance
$\Pi^m_{q+1} = \mathbf{N}\mathbf{P}^a{}_q\mathbf{N}^T - \mathbf{M}\mathbf{P}^a{}_q\mathbf{M}^T$

Spatial averaged of the model-error variance



For the first moments of the experiment, $\langle V^m \rangle(t) \sim t \frac{\langle \kappa \rangle}{\langle \nu^p \rangle(0)} \langle V^p \rangle(0)$,

then $\langle V^m \rangle(t) \sim 1 - \left(\frac{l_h^2}{l_h^2 + 4\langle \kappa \rangle t} \right)^{1/2}$, where $\langle s^p \rangle(0) = l_h^2$.

Table of contents

Conclusion

For the parametric Kalman filte (PKF):

- In the PKF error-covariance matrix are approximated by some covariance model
- The dynamics of the parameters is an approximation of the dynamics of the real error-covariance matrix, it often needs a closures but gives access to the physics of uncertainty.
- Some applications are under investigation – see the next presentation of Martin Sabathier with the assimilation of the Earth radiation belts

Concerning its application for the model-error covariance properties:

- We can compute the predicatability-error dynamics from the PKF
- This gives a proxy of the flow-dependent part of the model-error covariance that could be interesting in the diffusive case
- We illustrate this on a simple transport over a 1D domain
- Give some clues for inflation in 2D/3D domain applications – see the next presentation of Richard Ménard

Boyd, J. (2001).
Chebyshev and Fourier Spectral Methods.
Dover Publications, second edition.

Evensen, G. (2009).
Data Assimilation: The Ensemble Kalman Filter.
Springer-Verlag Berlin Heidelberg.

Pannekoucke, O. (2021).
An anisotropic formulation of the parametric kalman filter assimilation.
to appear in Tellus A: Dynamic Meteorology and Oceanography.

Pannekoucke, O. and Arbogast, P. (2021).
SymPKF (v1.0): a symbolic and computational toolbox for the design of parametric kalman filter dynamics.
Geoscientific Model Development.

Pannekoucke, O., Bocquet, M., and Ménard, R. (2018).
Parametric covariance dynamics for the nonlinear diffusive burgers' equation.
Nonlinear Processes in Geophysics, 2018:1–21.

Pannekoucke, O. and Fablet, R. (2020).
PDE-NetGen 1.0: from symbolic partial differential equation (PDE) representations of physical processes to trainable neural network representations.
Geoscientific Model Development, 13(7):3373–3382.

Pannekoucke, O., Ménard, R., El Aabaribaoune, M., and Plu, M. (2021).
A methodology to obtain model-error covariances due to the discretization scheme from the parametric kalman filter perspective.
Nonlinear Processes in Geophysics, 28(1):1–22.