Spatial covariance modeling for stochastic sub-grid scales parameterizations

Federica Gugole

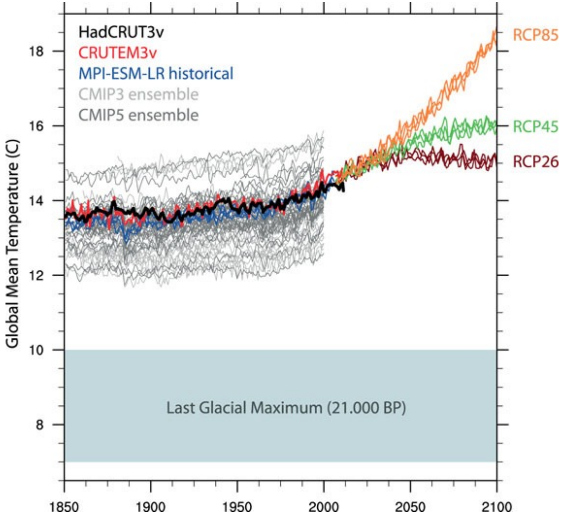
Joint work with Christian L. E. Franzke

University of Hamburg, Meteorological Institute

Center for Earth System Research and Sustainability (CEN)

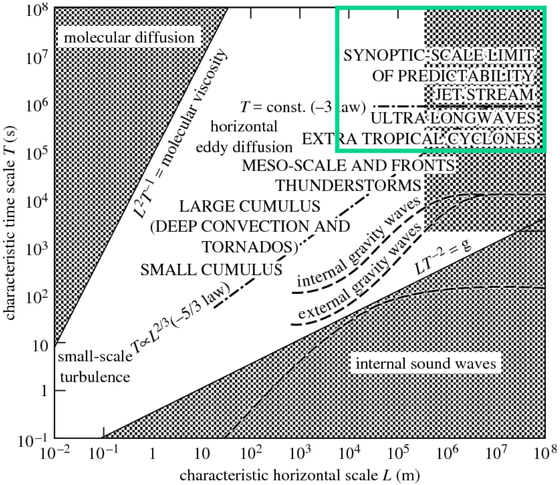
Current affiliation: Centrum Wiskunde & Informatica

Climate models agree on something, but not on everything



[T. Mauritsen et al. (2012), *JAMES*]

The climate system includes multiscales dynamics, and climate models resolve only a tiny part of it



[Cullen and Brown (2009), *Phil. Trans. R. Soc. A*]

Climate models are the discrete approximation of the Navier-Stokes equations (with some extras)

In case of a dry atmosphere or a sweet ocean:

$$
\begin{aligned}
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) &= 0 \\
\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} + 2\Omega \times \mathbf{u} &= -\frac{\nabla p}{\rho} + \nu \nabla^2 \mathbf{u} + \mathbf{F} \\
\rho &= h(\theta, p) \\
\frac{\mathrm{D}\theta}{\mathrm{D}t} - \frac{1}{c_p}\left(\frac{\theta}{T}\dot{Q}\right) &= 0
\end{aligned}
$$

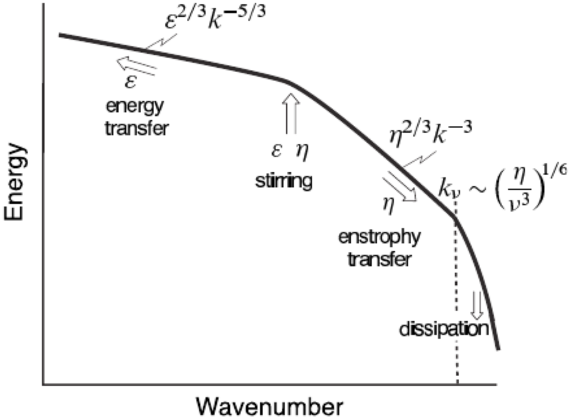
The limitation on the time step does not allow to represent the whole range of scales

The CFL condition (1928) for 1-dimensional advection reads

$$C = \frac{u\Delta t}{\Delta x} \leq C_{max}$$

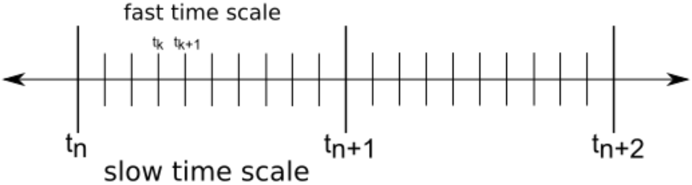
Scale	Typical size (km)	Life span
Global scale	5000	Days to weeks or more
Synoptic scale	2000	Days to weeks
Mesoscale	20	Minutes to days
Microscale	0, 002	Seconds to minutes

The kinetic energy of the system is sensitive to the grid resolution



[G. K. Vallis (2006), *Cambridge University Press*]

Stochastic parameterizations are meant to represent the unresolved scales in an indirect way



Advantages

- Gain in computational time
- Reduction of model errors
- Representation of uncertainties and model errors

[T. Palmer at al. (2009), *ECMWF Technical Memorandum*]

Many stochastic parameterizations are available and they all have different characteristics

E. Mémin (2014), Fluid flow dynamics under location uncertainty

D. D. Holm (2015), Variational principles for stochastic fluid dynamics

M. F. Jansen and I. M. Held (2014), Parameterizing subgrid-scale eddy effects using energetically consistent backscatter

D. Crommelin and W. Edeling (2020), Resampling with neural networks for stochastic parameterization in multiscale system

J. E. Frank and G. A. Gottwald (2013), Stochastic homogenization for an energy conserving multi-scale toy model of the atmosphere

and many others...

Spatial covariance modeling

1. The stochastic 2-layer Quasi-Geostrophic (QG) model
   - Overview
   - Derivation of the stochastic equations
2. Dimension reduction techniques
3. Results

The 2-layer QG model is quite an idealized setting but still dynamically challenging

Performing a scale analysis on the shallow water equations,
we obtain the so-called geostrophic balance, i.e. $-fv = -g\frac{\partial \eta}{\partial x}$

The QG approximation includes advenction and the time derivative,
and allows small departures from the geostrophic balance

It represents synoptic-scale mid-latitude atmospheric dynamics and
has features, such as a jet-like zonal flow, similar to the real atmosphere

Inviscid 2-layer QG model: starting point

Non-dimensional formulation on a β plane
$\rightarrow f = f_0 + \beta y$

No external forcing nor damping

q_B, ψ_B (q_T, ψ_T) barotropic (baroclinic) potential vorticity and streamfunction

The noise acts directly on the fast baroclinic mode and indirectly on the slow barotropic mode

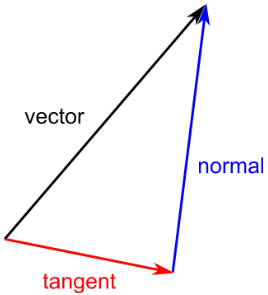
$$
\begin{aligned}
dq_B &= \left(-J(\psi_B, q_B) - J(\psi_T, q_T)\right) dt \\
dq_T &= \left(-J(\psi_T, q_B) - J(\psi_B, q_T)\right) dt \\
q_B &= \nabla^2 \psi_B + \beta y \\
q_T &= \nabla^2 \psi_T - k_d^2 \psi_T^2 \\
H(q_B, q_T) &= \frac{1}{2} \iint_{\mathbf{x}} \left[(\nabla \psi_B)^2 + (\nabla \psi_T)^2 + k_d^2 {\psi_T}^2 \right] d\mathbf{x}
\end{aligned}
$$

To guarantee conservation of energy, model the subgrid scales with a stochastic process plus a correction

$$
\begin{aligned}
dq_B &= (-J(\psi_B, q_B) - J(\psi_T, q_T))\,dt \\
dq_T &= (-J(\psi_T, q_B) - J(\psi_B, q_T))\,dt \underbrace{-\Gamma q_T dt + \Sigma dW_t}_{\texttt{subgrid processes}} \; \underbrace{+dY_t}_{\texttt{correction}} \\
dY_t &= S_t dW_t + B_t dt \\
H(q_B, q_T) &= \frac{1}{2} \iint_{\mathbf{x}} \left[(\nabla \psi_B)^2 + (\nabla \psi_T)^2 + k_d^2 {\psi_T}^2 \right] d\mathbf{x}
\end{aligned}
$$

[J. Frank, and G. A. Gottwald (2013), *Physica D*]

The projection operator approach removes the normal component of the stochastic perturbation



Projection operator

$$\mathbb{P} = \mathsf{I} - \frac{1}{|\nabla_{q_T} H|^2} \nabla_{q_T} H (\nabla_{q_T} H)^T$$
$$= \mathsf{I} - \frac{1}{|\psi_T|^2} \psi_T \psi_T^T$$

The correction should have only components orthogonal to the manifold of constant energy, i.e.

$$\mathbb{P} S_t = 0 \quad \mathbb{P} B_t = 0$$

Imposing the conservation of total energy determines B_t and S_t

By Ito's theorem

$$
\begin{aligned}
dH &= \frac{\partial H}{\partial q_B} \cdot dq_B + \frac{\partial H}{\partial q_T} \cdot dq_T + \frac{1}{2} \frac{\partial^2 H}{\partial q_T \partial q_T} : dq_T dq_T^T \\
&= \mu_H dt + \sigma_H dW_t = 0
\end{aligned}
$$

We obtain the structure of the correction

$$
\begin{aligned}
S_t &= -\left(\mathsf{I} - \mathbb{P}\right) \Sigma \\
B_t &= \left(\mathsf{I} - \mathbb{P}\right) \Gamma q_T + \frac{1}{2\left|\psi_T\right|^2} \psi_T \frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P} \Sigma \Sigma^T \mathbb{P}
\end{aligned}
$$

The stochastic inviscid 2-layer QG model contains two unknowns

$$
\begin{aligned}
dq_B &= \left(-J(\psi_B, q_B) - J(\psi_T, q_T)\right) dt \\
dq_T &= \left(-J(\psi_T, q_B) - J(\psi_B, q_T)\right) dt + \mathbb{P}\Sigma dW_t \\
&\quad - \left(\mathbb{P}\Gamma q_T - \frac{1}{2\left|\psi_T\right|^2}\left(\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}\right)\psi_T\right) dt \\
q_B &= \nabla^2\psi_B + \beta y \\
q_T &= \nabla^2\psi_T - k_d^2\psi_T^2
\end{aligned}
$$

[F. Gugole and C. L. E. Franzke (2019), *MCWF*]

The stochastic forced and damped 2-layer QG model contains one unknown

$$
\begin{aligned}
dq_B = & -\left(J(\psi_B - \frac{1}{2}Uy, q_B) + J(\psi_T - \frac{1}{2}Uy, q_T)\right) dt - \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right) dt \\
& - \frac{C_{Leith}\Delta^6}{2}\nabla^2\left(\left|\nabla^4(\psi_B + \psi_T)\right| \nabla^4(\psi_B + \psi_T) + \left|\nabla^4(\psi_B - \psi_T)\right| \nabla^4(\psi_B - \psi_T)\right) dt \ , \\
dq_T = & -\left(J(\psi_T - \frac{1}{2}Uy, q_B) + J(\psi_B - \frac{1}{2}Uy, q_T)\right) dt + \frac{1}{2}\tau_f^{-1}\left(\nabla^2\psi_B - \nabla^2\psi_T\right) dt \\
& - \frac{C_{Leith}\Delta^6}{2}\nabla^2\left(\left|\nabla^4(\psi_B + \psi_T)\right| \nabla^4(\psi_B + \psi_T) - \left|\nabla^4(\psi_B - \psi_T)\right| \nabla^4(\psi_B - \psi_T)\right) dt \\
& + \mathbb{P}\Sigma dW_t + \frac{1}{2\left|\psi_T\right|^2}\left(\frac{\partial^2 H}{\partial q_T \partial q_T} : \mathbb{P}\Sigma\Sigma^T\mathbb{P}\right)\psi_T dt \\
q_B = & \nabla^2\psi_B + \beta y \\
q_T = & \nabla^2\psi_T - k_d^2\psi_T^2
\end{aligned}
$$

[F. Gugole and C. L. E. Franzke (2020), *JAMES*]

Spatial covariance modeling

1. The stochastic 2-layer Quasi-Geostrophic (QG) model
2. Dimension reduction techniques
   - Empirical Orthogonal Functions (EOF), also known as PCA or POD
   - Dynamic Mode Decomposition (DMD)
3. Results

EOF is a statistical analysis technique that gives climatic patterns

EOF considers long time series and
derives the dominant patterns of variability

It is widely used in climate science thanks to its robustness
and ease of computation

Its patterns have global in time validity and
do not need to be recomputed during the simulation

DMD relates to the Koopman operator, which looks at the dynamics locally in time

Consider a dynamical system

$$\dot{x} = f(x) \quad x(0) = x_0 \quad x \in \mathbb{R}^d$$

We can introduce the flow map φ_t such that $x(t) = \varphi_t(x_0)$.
Let $\psi(x)$ be an observable, then

$$\psi(x(t)) = \psi(\varphi_t(x_0)) = \mathcal{K}_t \psi(x_0)$$

where $\mathcal{K}_t$ is the Koopman operator.

[I. Mezić (2013), *Annu. Rev. Fluid Mech.*]

The length of the time series can be seen as scale selection

Consider two time series such that

$$\mathbf{X} = \begin{pmatrix} | & | & & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_m \\ | & | & & | \end{pmatrix} \qquad \mathbf{X}' = \begin{pmatrix} | & | & & | \\ \mathbf{x}'_1 & \mathbf{x}'_2 & \cdots & \mathbf{x}'_m \\ | & | & & | \end{pmatrix}$$

$$\mathbf{x}_k = \mathbf{x}(t_k) \in \mathbb{R}^n \quad \mathbf{x}_k = \mathcal{K}_{\Delta t}\mathbf{x}_{k-1} \quad \mathbf{x}'_k = \mathcal{K}_{\Delta t}\mathbf{x}'_{k-1} \quad \mathbf{x}'_k = \mathcal{K}_{\delta t}\mathbf{x}_k$$

and $\delta t \leq \Delta t$. $\mathcal{K}_{\delta t}$ is approximated as

$$K_{\delta t} = X' X^{\dagger}$$

where $X^{\dagger}$ is the pseudo-inverse of X, and a low rank $r \leq m$ singular value decomposition of $X = U \Sigma V^{\star}$ have been used.

[J. N. Kutz et al. (2016), *SIAM*]

[M. Gavish and D. L. Donoho (2014), *IEEE Transactions on Information Theory*]

Let us see first if we actually need a spatial structure for Σ

Let us consider the inviscid scenario and assume that each grid point is independent from each other, i.e.

$$\Sigma = \sigma \mathsf{I} \quad \text{and} \quad \Gamma = \gamma \mathsf{I}$$

with I identity matrix and $\sigma, \gamma \in \mathbb{R}$

Numerical setting

Double-periodic boundary conditions on $2\pi \times 2\pi$ domain

Grid-point based discretization on a Cartesian grid

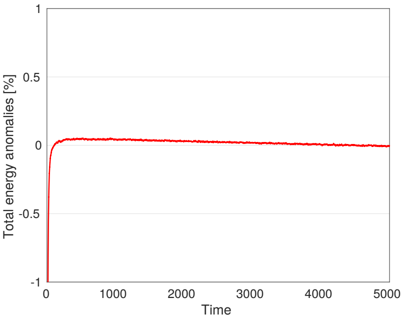
Numerical methods used:

- Time stepping: explicit Runge Kutta 4^{th} order
- Jacobian and Laplacian: Arakawa scheme, Fast Fourier Transform
- Stochastic integration: Euler-Maruyama

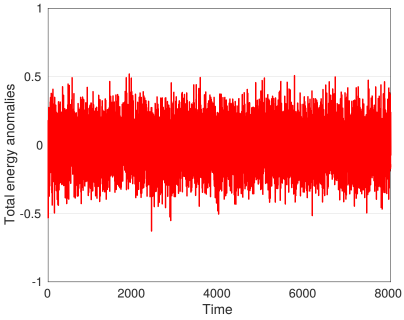
Low resolution 128×128; high resolution 512×512

iid noise: energy is conserved

Stochastic simulation



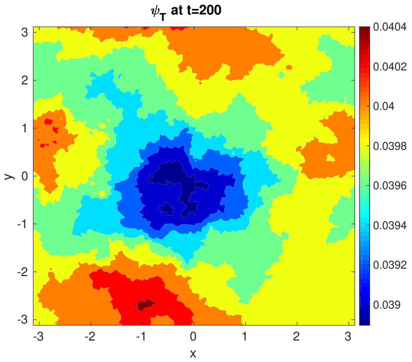
Low resolution deterministic model



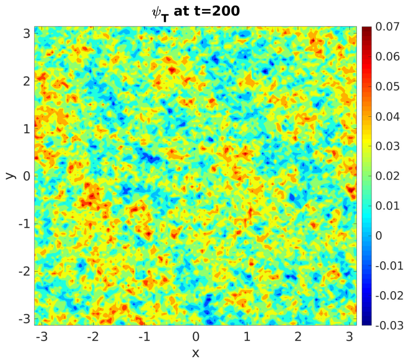
[F. Gugole and C. L. E. Franzke (2019), *MCWF*]

iid noise: the dynamics does not look promising

Stochastic simulation



Reference solution



[F. Gugole and C. L. E. Franzke (2019), *MCWF*]

Let us try now with a spatial structure for Σ

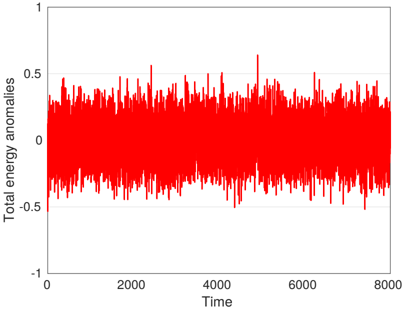
Let us consider the inviscid case and assume that there is correlation between the grid points, i.e.

$$\Sigma = \sum_{i=1}^{2} \omega_i \phi_i \quad \sum_i \omega_i = 1 \quad \omega_i \sim \mathcal{U}\{0, 1\} \qquad \Gamma = \mathbf{0}$$

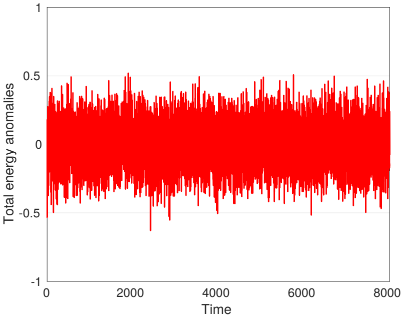
where the eigenvectors ϕ_i are computed with EOF

Spatially correlated noise: energy is conserved

Stochastic simulation



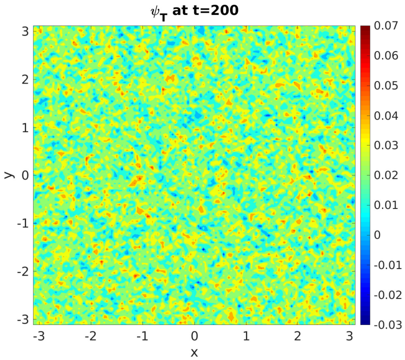
Low resolution deterministic model



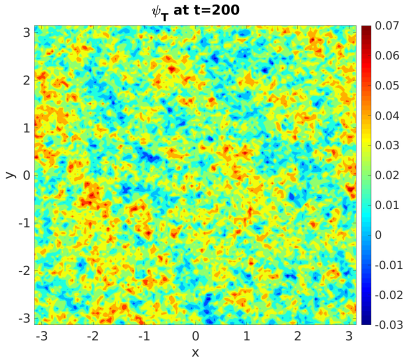
[F. Gugole and C. L. E. Franzke (2019), *MCWF*]

Spatially correlated noise: the dynamics looks much better

Stochastic simulation



Reference solution



[F. Gugole and C. L. E. Franzke (2019), *MCWF*]

Spatially correlated noise: eddy length in the barotropic mode is improved

Variable	512^2	128^2	128^2 Stochastic
ψ_B zonal	0.71428	0.52422	[0.51124, 0.56958]
ψ_B merid.	0.71632	0.52406	[0.51151, 0.56854]
ψ_T zonal	0.14918	0.07460	[0.07452, 0.07459]
ψ_T merid.	0.14918	0.07458	[0.07451, 0.07460]

[F. Gugole and C. L. E. Franzke (2019), *MCWF*]

What happens if I use a different technique?

Let consider the forced and damped system and
use two different data analysis techniques to define Σ

EOF

$$\Sigma(\mathbf{x}) = \sum_{i=1}^{5} \sqrt{\lambda_i^{EOF}} \phi_i^{EOF}$$

DMD

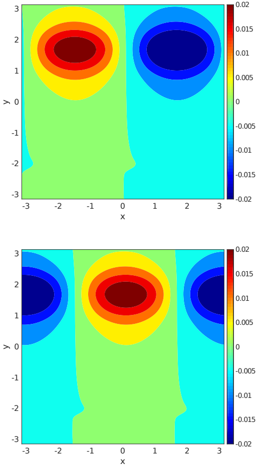
$$\Sigma(t, \mathbf{x}) = \frac{1}{2} \sum_{i=1}^{2} \left(\left(\mathrm{Re}\left(\lambda_i^{DMD}\right) + \imath \mathrm{Im}\left(\lambda_i^{DMD}\right) \right) \left(\mathrm{Re}\left(\phi_i^{DMD}\right) + \imath \mathrm{Im}\left(\phi_i^{DMD}\right) \right) + \mathrm{c.c.} \right)$$

$$= \sum_{i=1}^{2} \left(\mathrm{Re}\left(\lambda_i^{DMD}\right) \mathrm{Re}\left(\phi_i^{DMD}\right) - \mathrm{Im}\left(\lambda_i^{DMD}\right) \mathrm{Im}\left(\phi_i^{DMD}\right) \right)$$

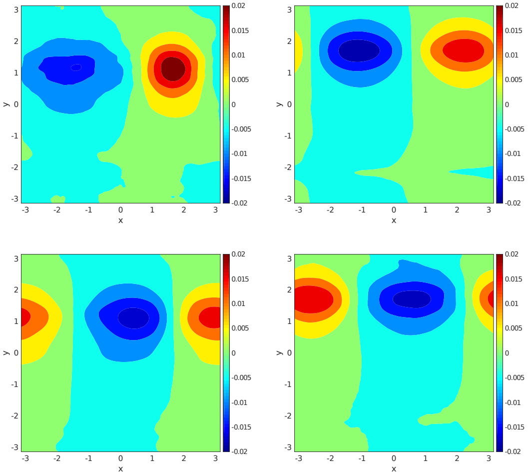
DMD parameters: $m = 16$, $r = 7$, $\delta t = 0.1$ $\Delta t = 3\delta t$,
and normalize Σ such that $\Lambda = \mathrm{Tr}(\Sigma\Sigma^T)$ is as with EOFs.
DMDs are recomputed every $m\Delta t$ time units.

EOFs 1-2 look like DMD mode 1 but the eddies in the DMD mode move also in *y*-direction

EOF



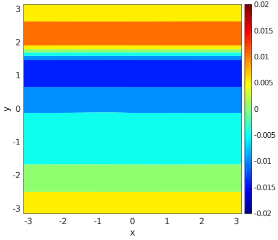
DMD



[F. Gugole and C. L. E. Franzke (2020), *JAMES*]

EOF 3 does not have a direct equivalent among DMD modes

EOF

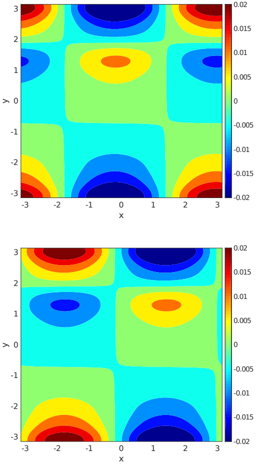


DMD

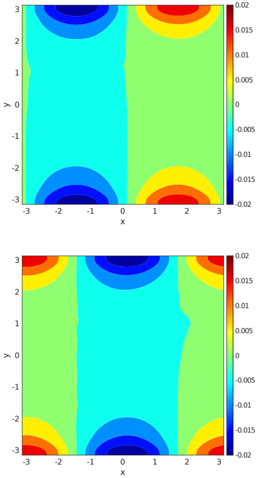
[F. Gugole and C. L. E. Franzke (2020), *JAMES*]

EOFs 4-5 are similar to DMD mode 2 but EOFs 4-5 have some extra structure

EOF



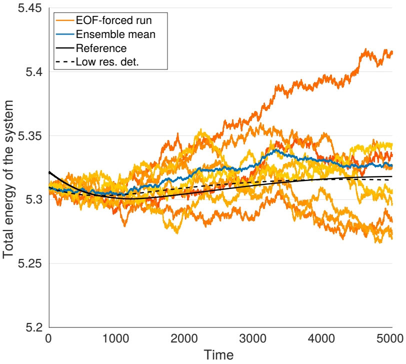
DMD



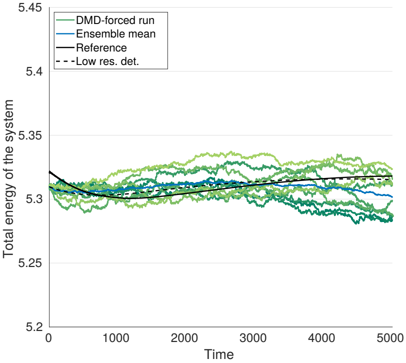
[F. Gugole and C. L. E. Franzke (2020), *JAMES*]

The parameterization is energy consistent in both cases

EOF



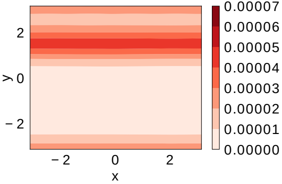
DMD



[F. Gugole and C. L. E. Franzke (2020), *JAMES*]

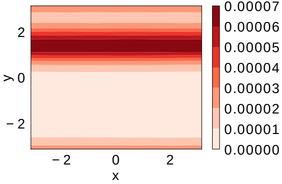
The DMD scheme has better performances with respect to the total variance

Low resolution model



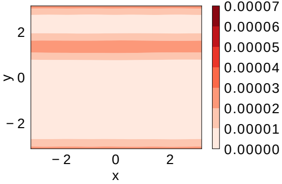
L^2 norm $= 0.0026$

EOF



L^2 norm $\in [0.0024,\ 0.0070]$

DMD

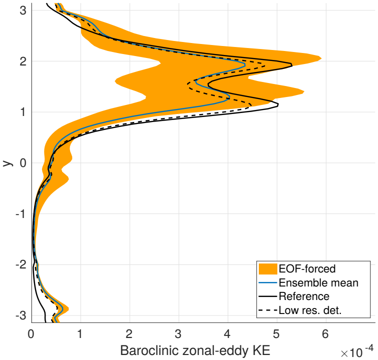


L^2 norm $\in [0.0014,\ 0.0042]$

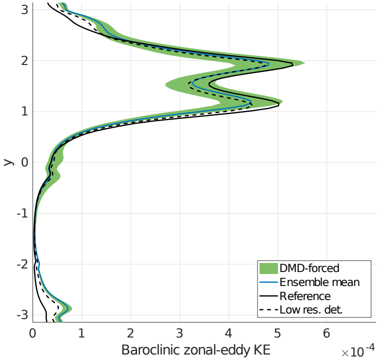
[F. Gugole and C. L. E. Franzke (2020), *JAMES*]

In the long run the DMD-forced ensemble resolves better the meridional movement of the eddies

EOF



DMD



[F. Gugole and C. L. E. Franzke (2020), *JAMES*]

Spatial covariance modeling

1. The stochastic 2-layer Quasi-Geostrophic (QG) model

2. Noise covariance structure
   - no structure (i.i.d. noise)
   - Empirical Orthogonal Functions (EOF)
   - Dynamic Mode Decomposition (DMD)

3. Results
   - dynamically consistent structure
   - constant in time
   - dynamically adaptive

What happens if we consider ... ?

- ... another parameterization?
- ... a more complex model?
- ... other data-driven techniques?
- ... a drift and/or a memory term?
- ...

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG
TRR181
CWI Centrum Wiskunde & Informatica
ECMA