Projected Data Assimilation using Dynamic Mode Decomposition

Aishah Albarakati (Clarkson University)

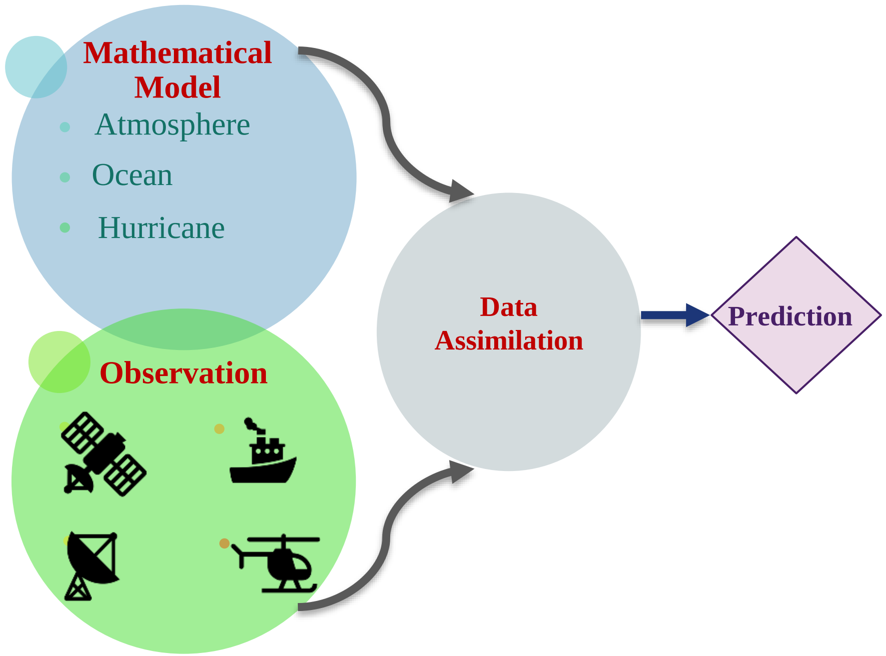
Marko Budišić (Clarkson University), Rose Crocker (Adelaide), Juniper Glass-Klaiber (Mount Holyoke), Sarah Iams (Harvard), John Maclean (Adelaide), Noah Marshall (McGill), Colin Roberts (Colorado State), and Erik Van Vleck (University of Kansas)

Graduate School Mathematics of Random Systems
September 09, 2021

This is the outgrowth of the 2019 Mathematics and Climate Research Network (MCRN)
summer school and academic year engagement program funded by NSF and AIM.

A. Albarakati, M. Budisic, R. Crocker, J. Glass-Klaiber, S. Iams, J. Maclean, N.Marshall, C. Roberts, and E. S. Van Vleck, "Model and Data Reduction for Data Assimilation: Particle Filters Employing Projected Forecasts and Data with Application to a Shallow Water Model," (2021) accepted in Model Computers and Mathematics with Applications (https://arxiv.org/abs/2101.09252)

What is Data Assimilation (DA)?



Data Assimilation (DA): combines computational models and observational data with their uncertainties to produce an estimate of the state of the physical system.

The challenges to make accurate predictions:

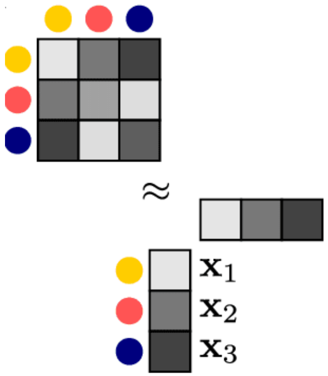
Nonlinearity of the model, Noise (Gaussian and non-Gaussian), High dimensionality

Common Data Assimilation techniques	Order	Model	Noise
Kalman Filter (KF)	High	Linear	Gaussian
Ensemble Kalman Filter (EnKF)	High	Nonlinear	Gaussian
Particle Filter (PF)	Low	Nonlinear	Non-Gaussian

Order Reduction in DA:

Order Reduction of Models

- Project model equations onto a subspace of lower dimension
- Preserve basic dynamic character of the model



Types

- Model-based
  Assimilation in the Unstable Subspace (AUS), (Trevisan and Uboldi 2004, Carrassi et al., 2007)
- Data-driven
  Proper Orthogonal Decomposition (POD), (Berkooz, Holmes, Lumley, 1993) or
  Dynamic Mode Decomposition (DMD), (Rowley et al., 2009, Schmid 2010)

Order Reduction and Data Assimilation

- Kalman filter + AUS (Bocquet and Carrassi, 2017)
- Kalman filter + DMD (Iungo et al., 2015)
- EnKF+ POD (Popov et al., 2021)
- PF + AUS (Maclean and Van Vleck, 2021)

Summary

- Find a dynamically relevant reduced basis using order reduction techniques:
  - Assimilation in the Unstable Subspace (AUS).
  - Proper Orthogonal Decomposition (POD).
  - Dynamic Mode Decomposition (DMD).
- Use this basis to project the physical and data models into a reduced dimension subspace.
- Develop projected DA techniques:
  - Projected Particle Filter (PROJ-PF)
  - Projected Optimal Proposal Particle Filter (PROJ-OP-PF).
- Perform data assimilation in the reduced dimension subspace.

Maclean, John, and Erik S. Van Vleck. “Particle Filters for Data Assimilation Based on Reduced-Order Data Models.” Quarterly Journal of the Royal Meteorological Society 147, no. 736 (2021): 1892–1907. https://doi.org/10.1002/qj.4001.

A. Albarakati, M. Budisic, R. Crocker, J. Glass-Klaiber, S. Iams, J. Maclean, N.Marshall, C. Roberts, and E. S. Van Vleck, “Model and Data Reduction for Data Assimilation: Particle Filters Employing Projected Forecasts and Data with Application to a Shallow Water Model,” (2021) accepted in Model Computers and Mathematics with Applications (https://arxiv.org/abs/2101.09252)

DA: Mathematical Formulation

- Consider a discrete time stochastic model:

Nonlinear state transition

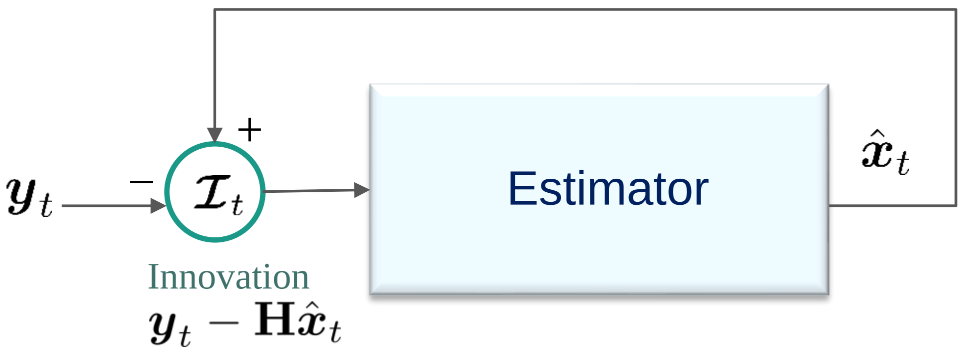
State noise

$$\textbf{STATE:} \qquad \boldsymbol{x}_t = \mathbf{F}\left(\boldsymbol{x}_{t-1}\right) + \boldsymbol{\omega}_t, \qquad \boldsymbol{\omega}_t \sim \mathcal{N}(0,\, \mathbf{Q}_t), \qquad \boldsymbol{x}_t \in \mathbb{R}^M$$

$$\textbf{OBSERVATION:} \qquad \boldsymbol{y}_t = \mathbf{H}\boldsymbol{x}_t + \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim \mathcal{N}(0,\, \mathbf{R}_t), \qquad \boldsymbol{y}_t \in \mathbb{R}^D$$

Observation noise

Linear observation operator



Particle Filters

- Particle filters run N parallel estimators $\{\hat{\boldsymbol{x}}_t^1, \hat{\boldsymbol{x}}_t^2, \ldots, \hat{\boldsymbol{x}}_t^N\}$ ("particles"), associated W_t^n
- The estimate of the true state is the weighted average of particle states.

$$\hat{\boldsymbol{x}}_t = \sum_{n=1}^{N} W_t^n \boldsymbol{x}_t^n$$

- Start with $\hat{\boldsymbol{x}}_{t-1}^n,\ n = 1, \ldots, N$ particles with equal weight initially $W_{t-1}^n = \frac{1}{N}$.

Particle Update:

$$\hat{\boldsymbol{x}}_t^n = \mathbf{F}\left(\hat{\boldsymbol{x}}_{t-1}^n\right) + \boldsymbol{\omega}_t^n,\ \boldsymbol{\omega}_t \sim \mathcal{N}(0, \mathbf{Q})$$

Weights updated:

- The weights update based on the Bayes' Theorem: $P(\boldsymbol{x}_t|\boldsymbol{y}_t) = \dfrac{P(\boldsymbol{y}_t|\boldsymbol{x}_t)\,P(\boldsymbol{x}_t)}{P(\boldsymbol{y}_t)}$

$$W_t^n \propto \exp\left[-\frac{1}{2}\left(\boldsymbol{\mathcal{I}}_t^n\right)^\top \mathbf{R}^{-1}\left(\boldsymbol{\mathcal{I}}_t^n\right)\right] W_{t-1}^n. \qquad \boldsymbol{\mathcal{I}}_t := (\boldsymbol{y}_t - \mathbf{H}\hat{\boldsymbol{x}}_t)$$

Optimal Proposal Particle Filter(OP-PF)

- **Optimal** = variance of the weights is minimized
- PF: innovation updates weight $\mathcal{I}_t^n := \left(\boldsymbol{y}_t - \mathbf{H}\hat{\boldsymbol{x}}_{t-1}^n)\right)$

  OP-PF: innovation updates weight **and state**

Particle update:

$$\begin{aligned}\hat{\boldsymbol{x}}_t^n &= \mathbf{F}\left(\hat{\boldsymbol{x}}_{t-1}^n\right) + \boldsymbol{\omega}_t + \boldsymbol{k}\mathcal{I}_t^n, \quad \boldsymbol{\omega}_t \sim \mathcal{N}(0, \mathbf{Q}_p) \\ \boldsymbol{k} &= \mathbf{Q}_p\mathbf{H}^\top\mathbf{R}^{-1} \\ \mathbf{Q}_p^{-1} &= \mathbf{Q}^{-1} + \mathbf{H}^\top\mathbf{R}^{-1}\mathbf{H}\end{aligned}$$

Weight Update:

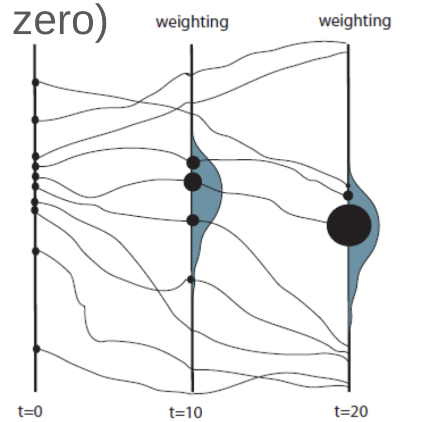
$$P(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{y}_t) = \frac{P(\boldsymbol{y}_t|\boldsymbol{x}_{t-1})\, P(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})}{P(\boldsymbol{y}_t|\boldsymbol{x}_{t-1})}$$

$$W_t^n \propto \exp\left[-\frac{1}{2}(\mathcal{I}_t^n)^\top(\mathbf{HQH}^\top + \mathbf{R})^{-1}(\mathcal{I}_t^n)\right] W_{t-1}^n$$

Particle Filter Degeneracy

- In high dimensional spaces the importance weights are more likely to be degenerate (one particle gets weight one, and all others get weight zero)

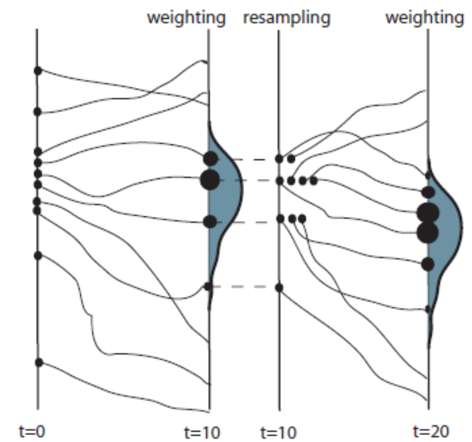
$$\underbrace{\log(N)}_{\text{Particles}} \propto (\underbrace{M}_{\text{State dim.}} \times \underbrace{D}_{\text{Observation dim.}})$$



(van Leeuwen et al., 2017)

- Resampling: abandon particles with very small weights and make multiple copies of particles with large weights.

$$\mathrm{ESS} = \frac{1}{\sum_{n=1}^{N} (W_t^n)^2} < \frac{1}{2} N$$



(van Leeuwen et al., 2017)

- Lowering either the state model dimension M, observation model dimension D or both helps in mitigating this problem.

Model Reduction by Orthogonal Projection

Full model

Unprojected

$$\boldsymbol{x}_{t+1} = \mathbf{F}\left(\boldsymbol{x}_t\right) + \boldsymbol{\omega}_t$$
$$\boldsymbol{y}_t = \mathbf{H}\boldsymbol{x}_t + \boldsymbol{\eta}_t$$

Reduction

$$\boldsymbol{v}_t = \mathbf{V}_t^\top \boldsymbol{x}_t$$

Orthogonal Proj. matrix

Reconstruction

$$\boldsymbol{x}_t = \mathbf{V}_t \boldsymbol{v}_t$$

Projected

$$\boldsymbol{v}_{t+1} = \mathbf{F}^q(\boldsymbol{v}_t) + \boldsymbol{\omega}_t^q, \quad \boldsymbol{\omega}_t^q \sim \mathcal{N}\left(0, \mathbf{Q}_t^q\right)$$
$$\boldsymbol{u}_t = \mathbf{H}_t^q \boldsymbol{v}_t + \boldsymbol{\eta}_t^q, \quad \boldsymbol{\eta}_t^q \sim \mathcal{N}\left(0, \mathbf{R}_t^q\right)$$

$$\mathbf{F}^q(\boldsymbol{v}_t) = \mathbf{V}_{t+1}^\top \mathbf{F}_t(\mathbf{V}_t \boldsymbol{v}_t)$$
$$\mathbf{H}_t^q = \mathbf{U}_t^\top \mathbf{H} \mathbf{V}_t$$

$$\mathbf{Q}_t^q = \mathbf{V}_{t+1}^\top \mathbf{Q}_t \mathbf{V}_{t+1}$$
$$\mathbf{R}_t^q = \mathbf{U}_t^\top \mathbf{R} \mathbf{U}_t$$

Reduced-order Model (ROM)

Projected Optimal Proposal (Proj-OP-PF)

Unprojected OP-PF

Particle update:

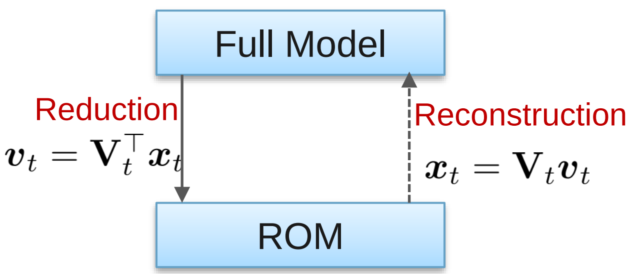
$$\hat{\boldsymbol{x}}_t^n = \mathbf{F}\left(\hat{\boldsymbol{x}}_{t-1}^n\right) + \boldsymbol{\omega}_t + K\left(\boldsymbol{y}_t - \mathbf{H}\mathbf{F}(\boldsymbol{x}_{t-1}^n)\right)$$

$$K = \mathbf{Q}_p \mathbf{H}^\top \mathbf{R}^{-1}$$

$$\mathbf{Q}_p^{-1} = \mathbf{Q}^{-1} + \mathbf{H}^\top \mathbf{R}^{-1} \mathbf{H}$$

Weight Update:

$$W_t^n \propto \exp\left[-\frac{1}{2}(\boldsymbol{\mathcal{I}}_t^n)^\top (\mathbf{H}\mathbf{Q}\mathbf{H}^\top + \mathbf{R})^{-1} (\boldsymbol{\mathcal{I}}_t^n)\right] W_{t-1}^n$$



Proj-OP-PF

Particle update:

$$\boldsymbol{v}_t^n = \mathbf{F}^q(\boldsymbol{v}_{t-1}^n) + \boldsymbol{\omega}_t + K_p\left(\boldsymbol{y}_t - \mathbf{H}\mathbf{V}_t \mathbf{F}_{t-1}^q(\boldsymbol{v}_{t-1}^n)\right)$$

$$K_p = \mathbf{Q}_p (\mathbf{H}\mathbf{V}_t)^\top \mathbf{R}^{-1}$$

$$\mathbf{Q}_p^{-1} = (\mathbf{Q}_t^q)^{-1} + (\mathbf{H}\mathbf{V}_t)^\top \mathbf{R}^{-1} (\mathbf{H}\mathbf{V}_t)$$

Weight Update:

$$W_t^n \propto \exp\left[-\frac{1}{2}(\boldsymbol{\mathcal{I}}_t^n)^\top (\mathbf{Z}_t^q)^{-1} (\boldsymbol{\mathcal{I}}_t^n)\right] W_{t-1}^n$$

$$\mathbf{Z}_t^q := (\mathbf{H}_t^q)\mathbf{Q}_t^q(\mathbf{H}_t^q)^\top + \mathbf{R}_t^q$$

$$\boldsymbol{\mathcal{I}}_t^n := \boldsymbol{y}_t^q - \mathbf{H}_t^q \boldsymbol{v}_t^n$$

A. Albarakati, M. Budisic, R. Crocker, J. Glass-Klaiber, S. Iams, J. Maclean, N.Marshall, C. Roberts, and E. S. Van Vleck, "Model and Data Reduction for Data Assimilation: Particle Filters Employing Projected Forecasts and Data with Application to a Shallow Water Model," (2021) accepted in Model Computers and Mathematics with Applications (https://arxiv.org/abs/2101.09252)

AUS: Assimilation in the Unstable Subspace

Inputs:

$$\mathbf{U}_{t+1}T_t = \mathbf{F}'_t(\boldsymbol{x}_t)\mathbf{U}_t \approx \frac{1}{\epsilon}[\mathbf{F}_t(\boldsymbol{x}_t + \epsilon\mathbf{U}_t) - \mathbf{F}_t(\boldsymbol{x}_t)], \quad t = 0, 1, \ldots$$

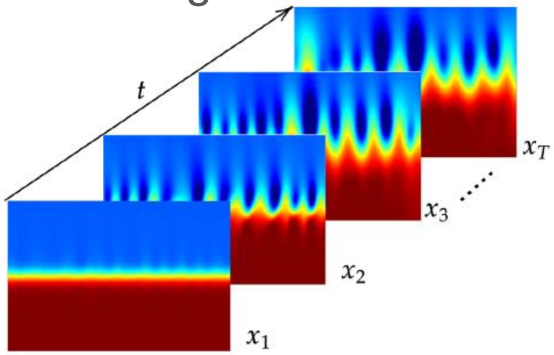
where U is orthogonal and $\boldsymbol{T}$ is upper triangular with positive diagonal elements

Outputs: Lyapunov vectors spanning expanding/neutral subspaces.

Proper Orthogonal Decomposition (**POD**) and
Dynamic Mode Decomposition (**DMD**):

Data-driven (model-free) techniques that do not require any knowledge of the underlying equations.

- Inputs: the evolution of state vectors $x_t \in \mathbb{R}^M$.
- Outputs: spatial-temporal coherent structure (modes) that dominate the observed data.

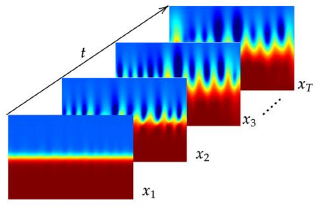


Proper Orthogonal Decomposition (POD)

Given a recording of evolution of state vectors stored as a snapshot matrix

Inputs:

$$\mathbf{X} := \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_T \end{bmatrix}$$



- Compute the singular value decomposition

$$\mathbf{X} = \mathbf{\Phi\Sigma\Psi}^\top$$

$$\mathbf{X} = \underbrace{\begin{bmatrix} \phi_1 & \phi_2 & \dots \end{bmatrix}}_{\text{Mut. orthogonal spatial profiles}} \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix}}_{\text{Singular values}} \underbrace{\begin{bmatrix} \psi_1 & \psi_2 & \dots \end{bmatrix}^\top}_{\text{Mut. orthogonal timeseries}}$$

Outputs:

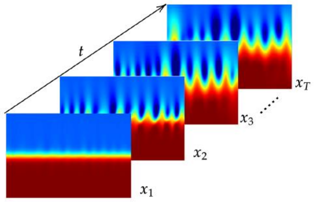
Reduce the dimension of X to a lower-dimensional matrix V by keeping M^q dominant spatial profiles.

$$\mathbf{V}_{\mathrm{POD}} = \begin{bmatrix} \phi_1 & \cdots & \phi_{M^q} \end{bmatrix}$$

DMD Projection:

Inputs: the evolution of snapshots x_t is approximated by

$$\boldsymbol{x}_t \approx \sum_{m=1}^{M} \Phi_m \exp(t\omega_m) b_m,$$



- Φ_m are DMD modes corresponding to a spatial profile.
- $\omega_m \in \mathbb{C}$ are complex-valued DMD frequencies governing growth, decay, and oscillation of time evolution.
- $b_m \in \mathbb{C}$ are linear combination coefficients.

Outputs: dynamically significant DMD reduced modes ordered by L^2 norms

$$\Phi = \begin{bmatrix} \Phi_1 & \Phi_2 & \dots & \Phi_{M^q} \end{bmatrix}$$

- Apply Gram-Schmidt to orthogonalize the DMD modes.

Experiment Set Up:

- Offline Projection Computation

Simulation Snapshot $\xrightarrow[\text{DMD}]{\text{POD}}$ Mode Matrix $\xrightarrow[\text{QR}]{}$ Orthogonal Basis

- Identical twin experiment:

Truth

$$\boldsymbol{x}_t = \mathbf{F}\left(\boldsymbol{x}_{t-1}\right)$$

$$\boldsymbol{y}_t = \mathbf{H}\boldsymbol{x}_t + \boldsymbol{\eta}_t$$

Estimates

$$\hat{\boldsymbol{x}}_t = \mathbf{F}\left(\hat{\boldsymbol{x}}_{t-1}\right) + \boldsymbol{\omega}_t$$

$$\hat{\boldsymbol{y}}_t = \mathbf{H}\hat{\boldsymbol{x}}_t$$

$\boldsymbol{x}^{\text{truth}}$

$\boldsymbol{x}^{\text{est}} = \sum_{n=1}^{N} W_t^n \hat{\boldsymbol{x}}_t^n$

$$\text{RMSE}(\boldsymbol{x}^{\text{truth}}, \boldsymbol{x}^{\text{est}}) := \|\boldsymbol{x}^{\text{truth}} - \boldsymbol{x}^{\text{est}}\|_2/\sqrt{M}$$

- Performance indicator (lower is better):

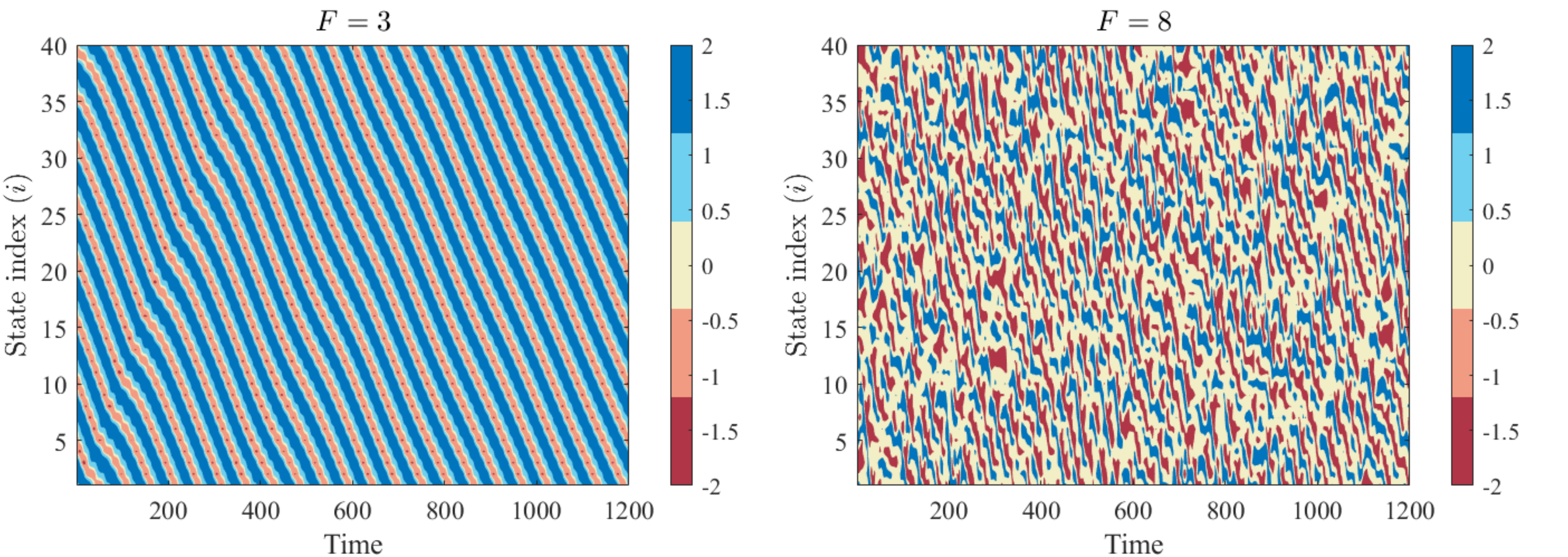
1. RMSE (compared to the standard deviation of the observation error)
2. Resampling Percentage

Case study: Lorenz '96

- The model is presented as a system of ODEs:

$$\dot{\boldsymbol{x}}_{t,i} = (\boldsymbol{x}_{t,i+1} - \boldsymbol{x}_{t,i-2})\,\boldsymbol{x}_{t,i-1} - \boldsymbol{x}_{t,i} + \boldsymbol{F}, \quad i = 1, \ldots, M$$

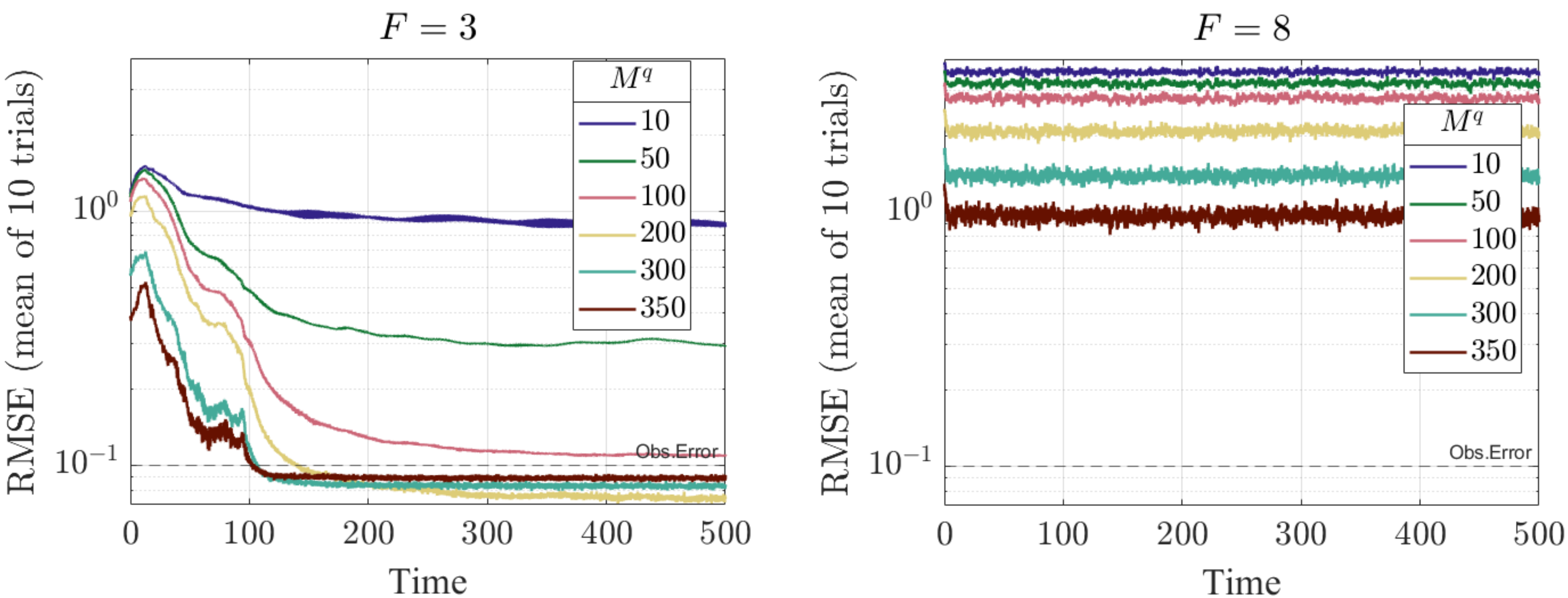
- F determines whether the evolution will be regular or chaotic.
- M is the state dimension



- All model variables are observed ($\mathrm{H} = \mathrm{I}$)

Projected Models and Data: L96 (POD)

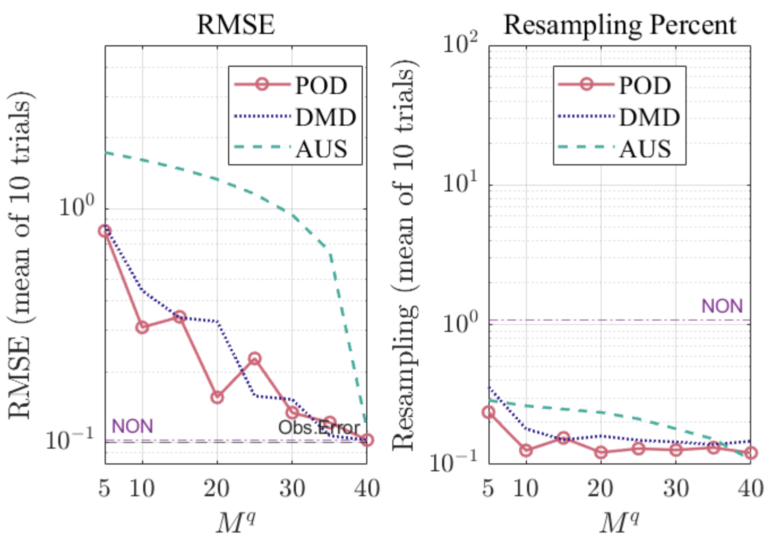
Low RMSE when the time evolution is structured



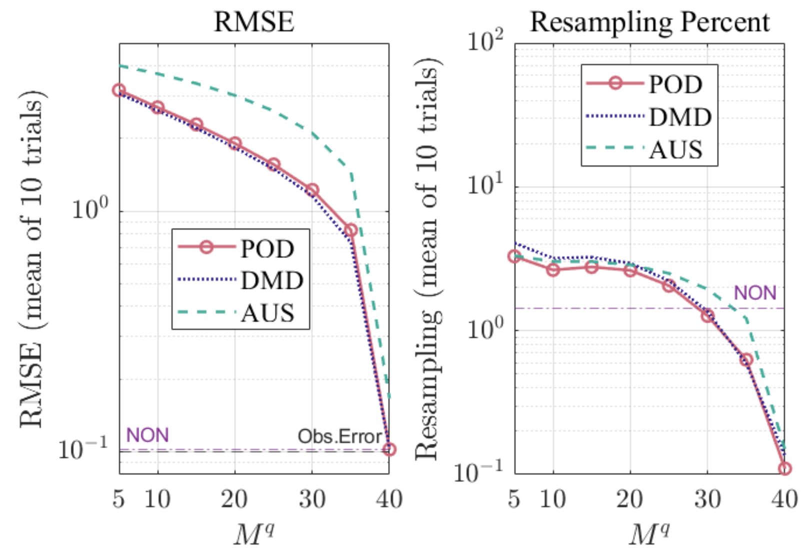
M = 400, $M^q = 400$ $D^q = 5$, N = 20 particles

L96: AUS, POD, and DMD

$F = 3,\ M = 40,\ M^q = 5 - 40,\ D^q = 5,\ N = 20$

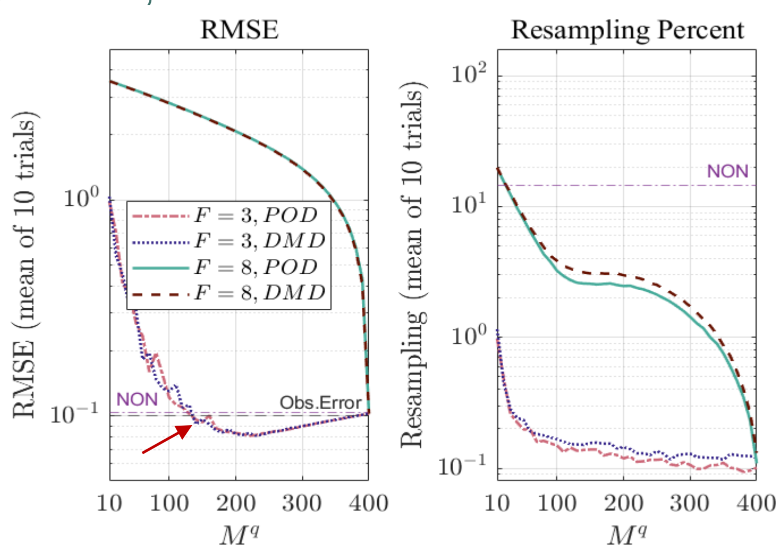


$F = 8,\ M = 40,\ M^q = 5 - 40,\ D^q = 5,\ N = 20$



$M = 400,\ M^q = 5 - 400,\ D^q = 5,\ N = 20$

Low RMSE and Resampling when the time evolution is structured (F=3)



RMSE remains an order of magnitude larger than the observation error (F=8)

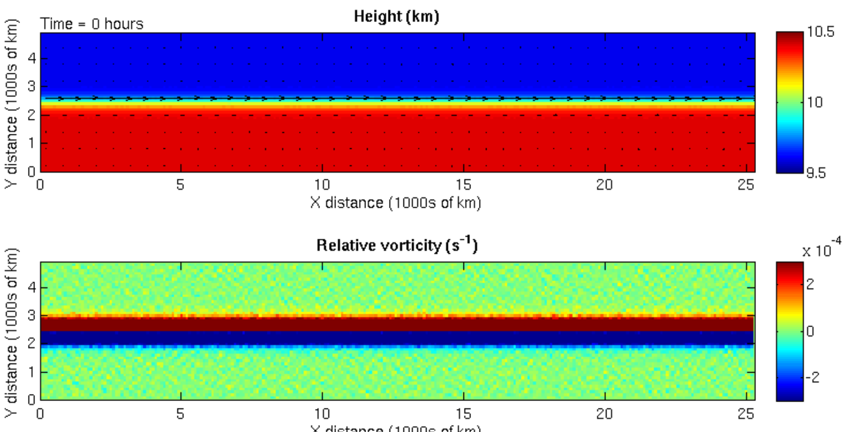
Case study: Shallow water equations (SWE)

$$\frac{\partial u}{\partial t} = \left(-\frac{\partial u}{\partial y} + f\right) v - \frac{\partial}{\partial x}\left(\frac{1}{2}u^2 + gh\right) + \nu\Delta u - c_b u$$

$$\frac{\partial v}{\partial t} = -\left(\frac{\partial v}{\partial x} + f\right) u - \frac{\partial}{\partial y}\left(\frac{1}{2}v^2 + gh\right) + \nu\Delta v - c_b v$$

$$\frac{\partial h}{\partial t} = -\frac{\partial}{\partial x}((h + \underline{h})u) - \frac{\partial}{\partial y}((h + \underline{h})v).$$

- $u(x, y, t)$ and $v(x, y, t)$ are velocity components and $h(x, y, t)$ is the height of the column of water at time t
- The three fields are evaluated at a grid of 254×50 points, resulting in a very high state dimension ($M = 38100$)

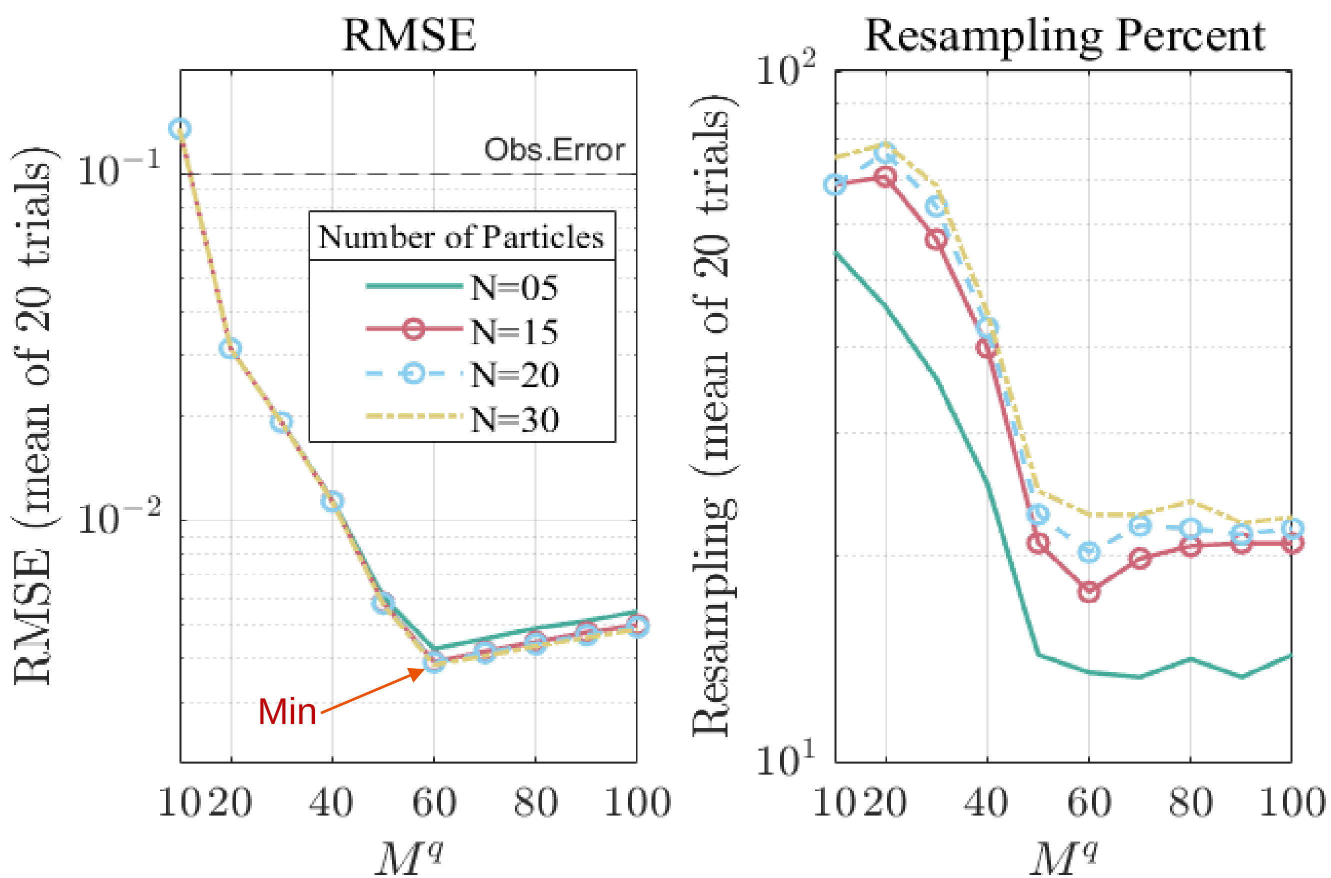


(Hogan, 2014), Barotropic instability

A. Albarakati, M. Budisic, R. Crocker, J. Glass-Klaiber, S. Iams, J. Maclean, N.Marshall, C. Roberts, and E. S. Van Vleck, "Model and Data Reduction for Data Assimilation: Particle Filters Employing Projected Forecasts and Data with Application to a Shallow Water Model," (2021) accepted in Elsevier (https://arxiv.org/abs/2101.09252)

Projected Models and Data: SWE

Assimilation is successful with relatively small # of particles



$\mathrm{M} = 38100, \mathrm{M}^q = 10 - 100 \;\; \mathrm{D}^q = 10$

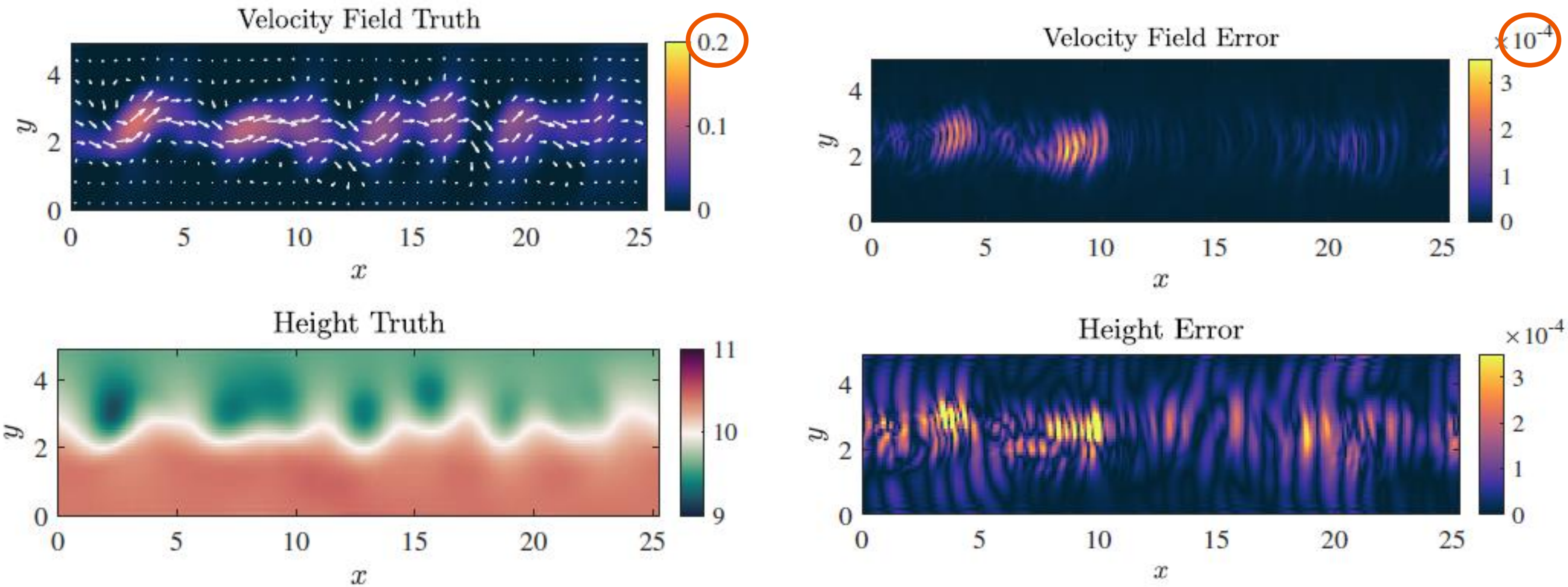
SWE+DMD: Spatial distribution of error

Full model space: $M = 38100$,
Reduced model space: $M^q = 20$

$$\frac{M^q}{M} \sim 2000$$

Full obs. space $D = 381$
Reduced obs. space: $D^q = 10$,

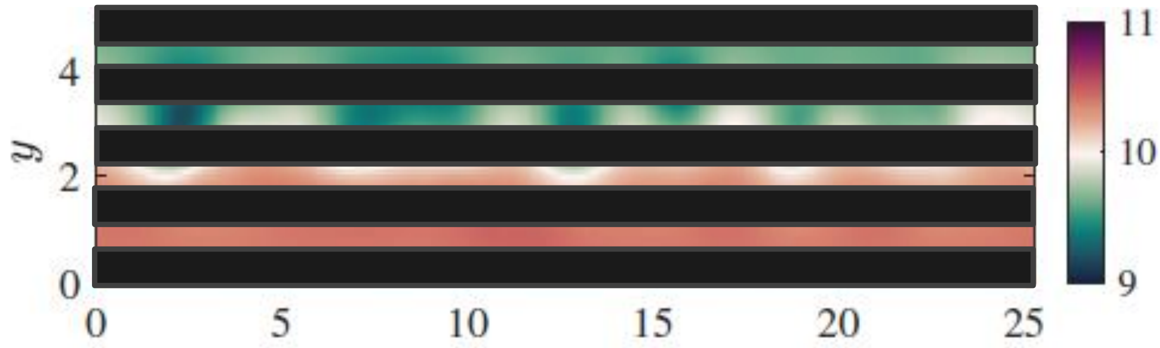
$N = 5$ particles

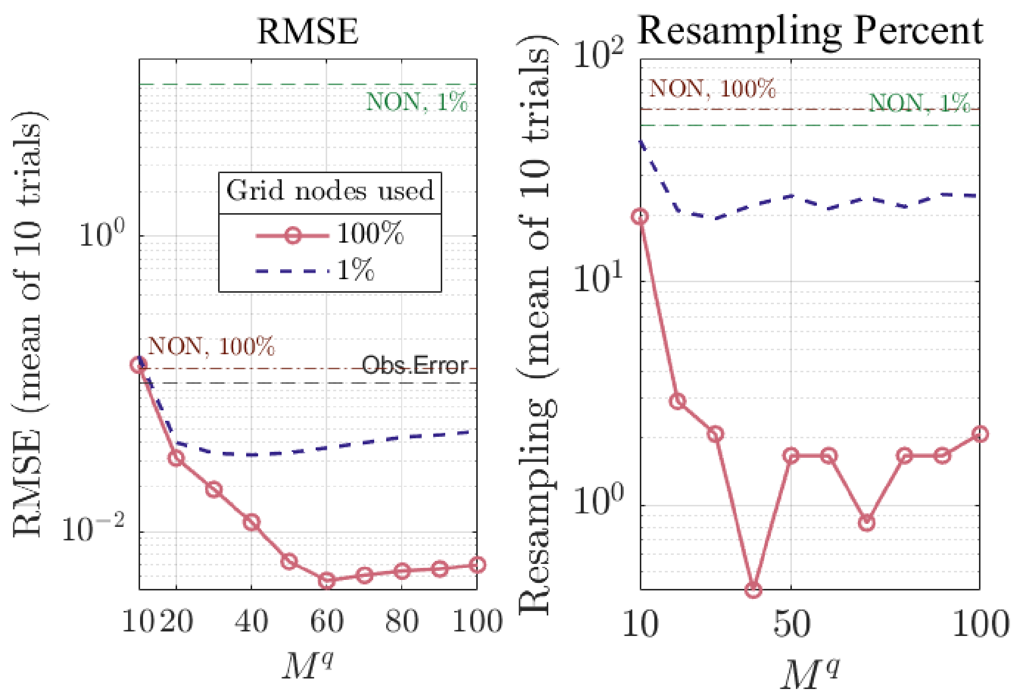


Clarkson

SWE+DMD: Low-rank observation operator

Assimilation is successful even when measurements are severely restricted





$M = 38100$, $M^q = 10 - 100$ $D^q = 10$, $N = 5$ particles

Summary and Related Publications:

- Derived a projected data assimilation framework based on the reduced order model, AUS, POD and DMD.
- Reduce the dimension of state and observation models to lower dimensions (e.g., SWE, 38100 to 10)
- Stable RMSE for L96 and low RMSE for SWE and resampling percentage.
- Promising results for the SWE, where Proj-OP-PF with minimal tuning provides good results.

Related Publications:

- Maclean, John, and Erik S. Van Vleck. “Particle Filters for Data Assimilation Based on Reduced-Order Data Models.” Quarterly Journal of the Royal Meteorological Society 147, no. 736 (2021): 1892–1907. https://doi.org/10.1002/qj.4001.
- A. Albarakati, M. Budisic, R. Crocker, J. Glass-Klaiber, S. Iams, J. Maclean, N.Marshall, C. Roberts, and E. S. Van Vleck, “Model and Data Reduction for Data Assimilation: Particle Filters Employing Projected Forecasts and Data with Application to a Shallow Water Model,” (2021) submitted (https://arxiv.org/abs/2101.09252)