Study of Variational Ensemble Methods for Image Assimilation

Etude de méthodes d'ensemble variationnelles pour l'assimilation d'images

Yin YANG

Fluminance Group, Inria Centre Rennes-Atlantique
Université de Rennes 1

16/Dec/2014

Outline

1. Introduction
   - The data assimilation
   - The image data
   - Variational methods & Ensemble Kalman filter
2. Hybrid methods
   - State of the art
   - Ensemble-based 4DVar scheme
   - Highlights of Ensemble-based 4DVar
3. Applications
   - 2D shallow water model
4. Stochastic approach
   - Shallow water equations under location uncertainty
   - Ensemble-based parameter estimation
5. Conclusion & Future work

Data assimilation problem

Dynamic model	$\partial_t \mathbf{x}(t,\mathrm{x}) + \mathbb{M}(\mathbf{x}(t,\mathrm{x}), u) = \mathbf{q}(t,\mathrm{x}),$	(1)
Prior knowledge model	$\mathbf{x}_0(\mathrm{x}) = \mathbf{x}_0^b(\mathrm{x}) + \boldsymbol{\eta}(\mathrm{x}),$	(2)
Observation model	$\mathcal{Y}(t,\mathrm{x}) = \mathbb{H}(\mathbf{x}(t,\mathrm{x})) + \boldsymbol{\epsilon}(t,\mathrm{x}).$	(3)

Variational and sequential methods

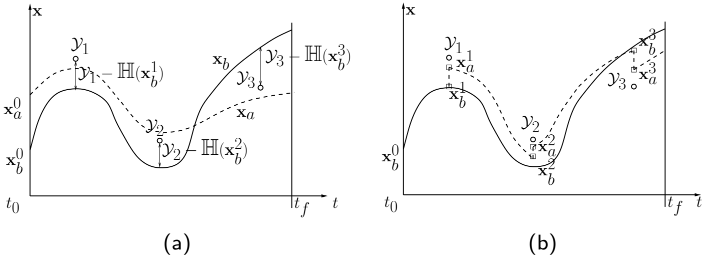


Figure: Scheme comparison between variational and sequential methods

Introduction: The image data

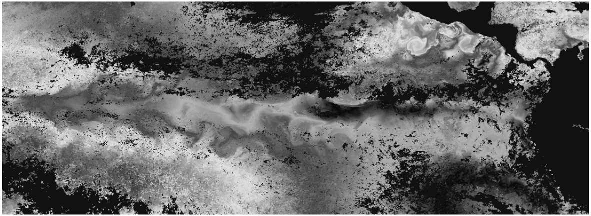


Figure: Sea surface temperature image data taken at 10/01/2008

Summary of attributes of image data:

- High resolution, especially higher than the dynamic grids
- The image is only indirectly related to the state variables
- Large areas of missing information and incomplete observations
- Potential balance issues
- Hard to evaluate the quality of analysis

Variational methods

The cost functional:

$$J \equiv \int_{t_0}^{t_f} \big(Y_s(t) - \mathcal{Y}(t)\big)^T \mathbf{R}^{-1} \big(Y_s(t) - \mathcal{Y}(t)\big)\, dt. \tag{4}$$

$$Y_s(t) \equiv \mathbb{H}(\mathbf{x}_t) \quad (\text{or } \mathbf{H}\mathbf{x}_t).$$

The necessary condition for J to be an extremum is:

$$\delta J(Y_s(t)) = 2\int_{t_0}^{t_f} \{\frac{\partial J}{\partial Y_s}, \delta Y_s(t)\} dt = 0.$$

Adjoint techniques: Expressing the constraint by transforming the cost function to one which only depends on the initial condition.

$$\{\mathbf{R}^{-1}\big(Y_s(t) - \mathcal{Y}(t)\big), \partial_{\mathbf{x}}\mathbb{H}(\mathbf{x}_t)\partial_{\mathbf{x}}\varphi_t(\mathbf{x})\delta\mathbf{x}_0\} = \{\partial_{\mathbf{x}}\varphi_t^* \partial_{\mathbf{x}}\mathbb{H}^* \mathbf{R}^{-1}\big(Y_s(t) - \mathcal{Y}(t)\big), \delta\mathbf{x}_0\}.$$

$$\boldsymbol{\nabla}_{\mathbf{x}_0} J = 2\int_{t_0}^{t_f} \partial_{\mathbf{x}}\varphi_t^* \partial_{\mathbf{x}}\mathbb{H}^* \mathbf{R}^{-1}\big(Y_s(t) - \mathcal{Y}(t)\big)\, dt.$$

Incremental 4DVar methods

A standard four dimensional variational data assimilation problem:

$$J(\mathbf{x}_0) = \frac{1}{2}\|\mathbf{x}(t_0, \mathrm{x}) - \mathbf{x}_0^b(\mathrm{x})\|_{\mathbf{B}}^2 + \frac{1}{2}\int_{t_0}^{t_f} \|\mathbb{H}(\mathbf{x}(t, \mathrm{x})) - \mathcal{Y}(t, \mathrm{x})\|_{\mathbf{R}}^2 dt. \quad (5)$$

Why increment?

Using two interleaved loops to

- incorporate the evolution of the background solution through the nonlinear dynamics,
- keeping the simplicity of the internal loop to recover an optimal increment driven by the tangent linear dynamics.

The cost functional in terms of the increment $\delta\mathbf{x}_0$:

$$J(\delta\mathbf{x}_0) = \frac{1}{2}\|\delta\mathbf{x}_0\|_{\mathbf{B}}^2 + \frac{1}{2}\int_{t_0}^{t_f} \|\partial_{\mathbf{x}}\mathbb{H}\ \delta\mathbf{x}(t, \mathrm{x}) - \mathbf{D}(t, \mathrm{x})\|_{\mathbf{R}}^2 dt, \quad (6)$$

where

$$\delta\mathbf{x}_t = \partial_{\mathbf{x}}\varphi_t(\mathbf{x}_b)\delta\mathbf{x}_0, \text{ and } \mathbf{D}(\mathrm{x}, t) = \mathcal{Y}(\mathrm{x}, t) - \mathbb{H}(\varphi_t(\mathbf{x}_b)).$$

Preconditioning and Conditioning

Conditioning

The Hessian conditional number:

$$\kappa(\mathcal{H}) \leq \kappa(\mathbf{B})(1 + \sigma^{-1}\lambda_{\min}(\mathbf{B}) \max_{t\in[t_0 t_f]}(\lambda_{\max}(C_t))(t_f - t_0)).$$

- Directly related to the conditional number of **B**.
- Perfect observation & High resolution observation lead to bad conditioning.

Preconditioned incremental form

Control variable transform (CVT): $\delta\mathbf{x}_t = \mathbf{U}\delta\mathbf{z}_t = \mathbf{B}^{\frac{1}{2}}\mathbf{z}_t,$

$$J(\delta\mathbf{z}_0) = \frac{1}{2}\|\delta\mathbf{z}_0\|^2 + \frac{1}{2}\int_{t_0}^{t_f} \|\partial_{\mathbf{x}}\mathbb{H}\ \partial_{\mathbf{x}}\varphi\ _t(\mathbf{x}_0)\mathbf{B}^{\frac{1}{2}}\delta\mathbf{z}_0 - \mathbf{D}(t,x)\|^2_{\mathbf{R}}dt.$$

$$\widetilde{\mathcal{H}} = \mathbb{I} + \int_{t_0}^{t_f} \mathbf{B}^{\frac{1}{2}T}\partial_X\varphi_t^*\partial_X\mathbb{H}^*\mathbf{R}^{-1}\partial_X\mathbb{H}\partial_{\mathbf{x}}\varphi_t\mathbf{B}^{\frac{1}{2}}dt. \tag{7}$$

Ensemble Kalman filter

Motivation

- Kalman filter is incapable of dealing with nonlinear operators.
- Extended Kalman filter only works well for slightly nonlinear models.

EnKF (Evensen 94) works as a Monte Carlo implementation of Kalman Filter.

- Forecast phase: A cloud of possible states is propagated by the nonlinear dynamics.
- Analysis phase: Update the ensemble state based on the observation.
  - Perturbed observation approach: Monte Carlo implementation.
  - Direct ensemble transformation approach.

4DVar VS EnKF

	4DVar	EnKF
Covariance Statistical Model	Not needed, Non-Gaussian errors allowed	Prior pdf of background error and observation error needed usually Gaussian
Background Errors	Fix at the beginning, Implicit evolution	Derived from ensemble, Flow dependent, can be quite noisy
Ability to Update the Background Error Covariance	No	Yes, the posterior background error covariance can be estimated directly
Localization	Not needed	Needed
Tangent Linear Adjoint Models	Needed	Not needed
Analysis Methods	Simultaneously assimilate all observations	Sequentially assimilate batches of observations
Estimation Methods	Iteratively minimizing a quadratic cost function: quasi-Newton methods	Explicitly calculate the analysis by Kalman gain matrix
Balance Constraints	Yes	Slightly lost in localization

State of the art

Variational method using flow-dependent background error covariance

EnKF-3DVar (Hamill and Snyder, 2000, Lorenc, 2003 , Buehner, 2005), ETKF-3DVar (Wang et al,2008), En4DVar (Buehner et al, 2010, Clayton et al, 2012, Zhang et al, 2009, Zhang and Zhang, 2012), 4DEnVar (Liu et al, 2008, Fairbairn et al, 2013, Decsroziers et al, 2014), Explicit 4DVar (Qiu and Zhou, 2005, Tian et al, 2008).

EnKF method adjusting to assimilate asynchronous data

4DEnKF (Hunt et al, 2004), 4DLETKF (Hunt et al, 2007, Fertig et al,2007), AEnKF (Sakov et al, 2010).

EnKF system incorporating an iterative procedures

MLEF(Zupanski, 2005), IEnKF (Sakov et al, 2012), VEnKF (Solonen et al, 2012).

4DEnVar: important aspects

Low rank approximation of the background error covariance matrix

$$\mathbf{B}^{\frac{1}{2}} \approx \mathbf{A}_b' := \frac{1}{\sqrt{N-1}}(\mathbf{x}_0^{(1)b} - \langle \mathbf{x}_0^b \rangle, \ldots, \mathbf{x}_0^{(N)b} - \langle \mathbf{x}_0^b \rangle), \tag{8}$$

The cost functional preconditioned by the empirical perturbation matrix

$$J(\delta \mathbf{z}_0) = \frac{1}{2}\|\delta \mathbf{z}_0\|^2 + \frac{1}{2}\int_{t_0}^{t_f} \|\partial_{\mathbf{x}}\mathbb{H}\, \partial_{\mathbf{x}}\varphi_t(\mathbf{x}_0)\mathbf{A}_b'\delta \mathbf{z}_0 - \mathbf{D}(t,x)\|_{\mathbf{R}}^2 dt. \tag{9}$$

The propagation of ensemble perturbation matrix projected into observation space

$$\partial_{\mathbf{x}}\mathbb{H}\, \partial_{\mathbf{x}}\varphi\,_t(\mathbf{x}_0)\mathbf{A}_b' \approx \frac{1}{\sqrt{N-1}}(\mathbb{H}(\varphi_t(\mathbf{x}_0^{(1)b})) - \mathbb{H}(\varphi_t(\langle \mathbf{x}_0^b \rangle)), \ldots, \mathbb{H}(\varphi_t(\mathbf{x}_0^{(N)b})) - \mathbb{H}(\varphi_t(\langle \mathbf{x}_0^b \rangle)) \tag{10}$$

General Algorithm

- Define the background state $\mathbf{x}_0^b$ and Initialize the background ensemble $\mathbf{X}_0^b$.
- Calculate the ensemble forecast $\mathbf{X}_t^b$ over the assimilation window.
- Empirically calculate the ensemble based background perturbation matrix $\mathbf{A}'_b$ and $\partial_{\mathbf{x}}\mathbb{H}\ \partial_{\mathbf{x}}\varphi\ _t(\mathbf{x}_0)\mathbf{A}'_b$.
- Do CVT $\delta\mathbf{x}_0 = \mathbf{A}'_b\delta\mathbf{z}_0$ and transform the cost function in terms of the control vector $\delta\mathbf{z}_0$.
- Solve for the initial increment $\delta\hat{\mathbf{z}}_0$ which minimize the problem:

$$J(\delta\mathbf{z}_0) = \frac{1}{2}\|\delta\mathbf{z}_0\|^2 + \frac{1}{2}\int_{t_0}^{t_f} \|\partial_{\mathbf{x}}\mathbb{H}\ \partial_{\mathbf{x}}\varphi_t(\mathbf{x}_0)\mathbf{A}'_b\delta\mathbf{z}_0 - \mathbf{D}_t\|_{\mathbf{R}}^2 dt.$$

- Update the analysis state $\mathbf{x}_0^a = \mathbf{x}_0^b + \mathbf{A}'_b\delta\hat{\mathbf{z}}_0$.

Ensemble generation

The ensemble can be generated by

- perturbing initial conditions:

$$\mathbf{x}_0^{f,(i)} = \mathbf{x}_0^f + \delta_0^{(i)}, \quad i = 1, \ldots, N. \tag{11}$$

  in which δ_0 is usually sampled from a peusdo-Gaussian field.

- perturbing parameters:

$$u^i = p(u). \tag{12}$$

Each member of the ensemble is independently integrated through the nonlinear model:

$$\mathbf{x}_t^{f,(i)} = \varphi_t(\mathbf{x}_0^{f,(i)}, u^i), \quad i = 1, \ldots, N. \tag{13}$$

Nested loops schemes

Different variants of the incremental technique:

- Courtier et al. 94

$$\mathbf{x}_0^k = \mathbf{x}_0^{k-1} + \delta\mathbf{x}_0^k,$$
$$J_b(\delta\mathbf{x}_0^k) = \frac{1}{2}\|\delta\mathbf{x}_0^k + \mathbf{x}_0^{k-1} - \mathbf{x}_0^b\|_{\mathbf{B}}^2,$$

- Weaver el al. 03

$$\mathbf{x}_0^k = \mathbf{x}_0^b + \delta\mathbf{x}_0^k,$$
$$\mathbf{D}_t^k = \mathcal{Y}_t - \mathbb{H}(\varphi_t(\mathbf{x}_0^{k-1})) + \partial_{\mathbf{x}}\mathbb{H}\;\delta\mathbf{x}_t^{k-1},$$

- Ensemble updating scheme

$$\mathbf{x}_0^k = \mathbf{x}_0^{k-1} + \delta\mathbf{x}_0^k,$$
$$\mathbf{B}^k = \overline{(\mathbf{x}_0^{k-1} - \mathbf{x}^{ref})(\mathbf{x}_0^{k-1} - \mathbf{x}^{ref})^T},$$
$$J_b(\delta\mathbf{x}_0^k) = \frac{1}{2}\|\delta\mathbf{x}_0^k\|_{\mathbf{B}^k}^2.$$

Ensemble Update schemes: Perturbed observation approach

Perturbed observation approach: Monte Carlo implementation

- Observation ensemble:

$$\mathcal{Y}^{i,j} = \mathcal{Y}^{i} + \epsilon^{i,j}, \quad i = 0, \dots, T, \quad j = 1, \dots, N.$$

- For each member $j = 1, \dots, N$:

$$J(\delta \mathbf{z}_0^{(j),k}) = \frac{1}{2}\|\delta \mathbf{z}_0^{(j),k}\|^2 + \frac{1}{2}\int_{t_0}^{t_f} \|\partial_{\mathbf{x}}\mathbb{H}\, \partial_{\mathbf{x}}\varphi_t(\mathbf{x}_0)\mathbf{A}_b'^{k}\delta \mathbf{z}_0^{(j),k} - \mathbf{D}^{(j),k}(t,x)\|_{\mathbf{R}}^2 dt.$$

- Background Ensemble members update:

$$\mathbf{x}_0^{(j)b,k+1} = \mathbf{x}_0^{(j)b,k} + \mathbf{A}_b'^{k}\widehat{\delta \mathbf{z}_0}^{(j),k}, \quad j = 1, \dots, N.$$

- Ensemble perturbation matrix update:

$$\mathbf{A}_b'^{k+1} = \frac{1}{\sqrt{N-1}}(\mathbf{x}_0^{(1)b,k+1} - \langle \mathbf{x}_0^{b,k+1}\rangle, \dots, \mathbf{x}_0^{(N)b,k+1} - \langle \mathbf{x}_0^{b,k+1}\rangle).$$

Ensemble Update schemes: Direct ensemble transformation approach

Direct ensemble transformation approach

- By approximating the link between the analysis error covariance matrix and the inverse Hessian matrix.

$$(\mathbf{P}^a)^{-1} \approx \mathcal{H}_{\mathbf{B}} = \mathbf{B}^{-1} + \int_{t_0}^{t_f} \partial_{\mathbf{x}}\varphi_t^* \partial_{\mathbf{x}}\mathbb{H}^* \mathbf{R}^{-1} \partial_{\mathbf{x}}\mathbb{H} \partial_{\mathbf{x}}\varphi_t dt,$$

$$\mathbf{A}'_a \approx \mathbf{A}'_b \mathcal{H}_{\mathbb{I}}^{-\frac{1}{2}}.$$

- Ensemble perturbation matrix update:

$$\mathbf{A}_b'^{k+1} = \mathbf{A}_b'^{k} \mathcal{H}_{\mathbb{I}}^{-\frac{1}{2},k} \mathbf{V}.$$

- Updated background state (Ensemble mean):

$$\widehat{\mathbf{x}}_0^{b,k} = \mathbf{x}_0^{b,k} + \mathbf{A}_b'^{k} \widehat{\delta \mathbf{z}_0}^{k}.$$

- Background Ensemble members update:

$$\mathbf{x}_0^{(j)b,k+1} = \widehat{\mathbf{x}}_0^{b,k} + \sqrt{N-1}\mathbf{A}_b'^{(j)k+1}, \quad j = 1, \ldots, N.$$

Localization technique

Localization aims at filtering the pseudo correlation arisen from the ensemble fields.

Localized covariance

- Define Schur product: $\mathbf{P}^b = \mathcal{C} \odot \mathbf{P}^e$.
  - Polynomial approximation of a Gaussian function with compact support is often employed.
- Spectral decomposition: $\mathcal{C} = E\lambda E^T$.
- Localized ensemble perturbation matrix: $\mathbf{P}'_b = (diag(\mathbf{A}_b'^{(1)})\mathcal{C}', \ldots, diag(\mathbf{A}_b'^{(N)})\mathcal{C}')$.

Local ensemble (analysis)

- The ensemble (re)propagation are dealt with full dynamic model in global space.
- The minimization w.r.t each grid point is conducted parallelly in a local region.

Outline

Conservative form:

$$\partial_t \eta + \partial_x(hu) + \partial_y(hv) = 0, \tag{14a}$$

$$\partial_t(hu) + \partial_x(hu^2) + \partial_y(huv) = -gh\partial_x\eta - ghS_bx + hfv + \nu\big(\partial_x(hu_x) + \partial_y(hu_y)\big), \tag{14b}$$

$$\partial_t(hv) + \partial_x(huv) + \partial_y(hv^2) = -gh\partial_y\eta - ghS_by - hfu + \nu\big(\partial_x(hv_x) + \partial_y(hv_y)\big). \tag{14c}$$

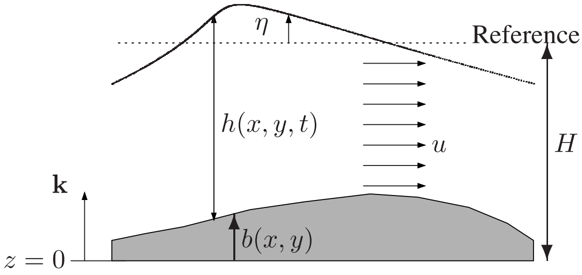


Figure: Schematic diagram of unsteady flow over an irregular bottom (Cushman-Rosin and Becker, 11).

Twin synthetic and experimental cases' setups:

	Synthetic case A	Synthetic case B	Experimental case
Domain size	100cm*25cm		
Dynamic Resolution	26*11	101*41	124*49
Assimilation time	0.2s		0.15s
Background error type	Gaussian	Non-Gaussian	Unknown
Cut-off distance		Optimally tuned	

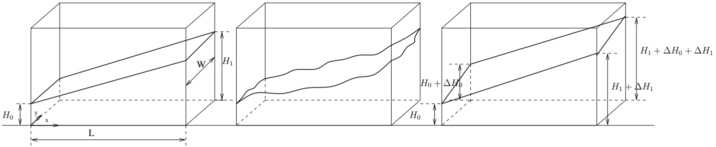


Figure: A priori initial experimental configuration (left) and the true synthetic initial conditions with a Gaussian noise–Case A–(middle) and with a 10% slope along the y-axis–Case B–(right).

Comparison tools

$$\text{RMSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(\mathbf{x}^f - \mathbf{x}^{obs})^2}. \quad (15)$$

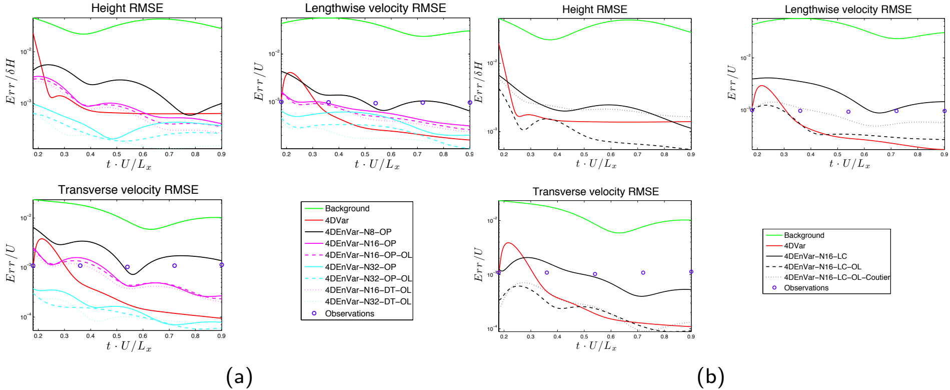


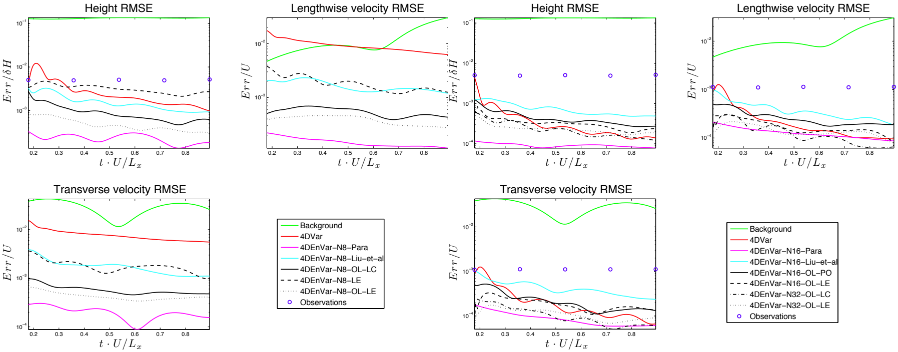
Figure: RMSE comparison between (a) group of methods with perfect background ensemble and no localization (OP, observation perturbation; DT: direct transformation; OL, extra outer loop), (b) outer loop schemes: No outer loop (black line), Algorithm of ensemble updating (black dashed line), Algorithm of Courtier et al, 94 (black dotted line).

Synthetic case A results, Imperfect background ensemble

Type	N	COD/L_x	OL Iter	IL Iter	RMSE(t_f)e^{-4}	RMSE($\bar{t}$)
Observation	-	-	-	-	-	$(9.758, 9.804)$
Background	-	-	-	-	$(275.5, 92.32)$	$(362.5, 115.2)$
4DVar	-	-	3	100	$(1.450, 0.8495)$	$(6.645, 5.693)$
4DEnVar-OP-LC	32	60%	2	100,100	$(2.768, 0.634)$	$(5.779, 3.992)$
4DEnVar-DT-LE	32	30%	2	100,100	$(4.268, 1.664)$	$(12.84, 4.268)$

Table: RMSE comparison table. Type: Group of methods with imperfect background ensemble and no localization (OP, observation perturbation; DT: direct transformation; LC, Localized covariance; LE, Local ensemble); N: ensemble members; COD/L_x: ratio of cut-off distance divided by characteristic length; OL Iter: Outer loop iteration; IL Iter: Inner loop iterations; RMSE(t_f): final RMSE; RMSE($\bar{t}$): mean RMSE.

Synthetic case B results: RMSE comparison



(a) (b)

Figure: RMSE comparison between an incremental 4DVar and 4DEnVar assimilation approaches: (a) partial observed through noisy free surface height, (b) fully observed system

Synthetic case B results: Comparison of computational power

Type	CPU Time	Memory demands
4DVar	3200s	Small
4DEnVar-PP(No Localization, N=8)	120s	Small
4DEnVar-LC(N=32)	2400s	Huge
4DEnVar-LE(N=32)	600s	Small

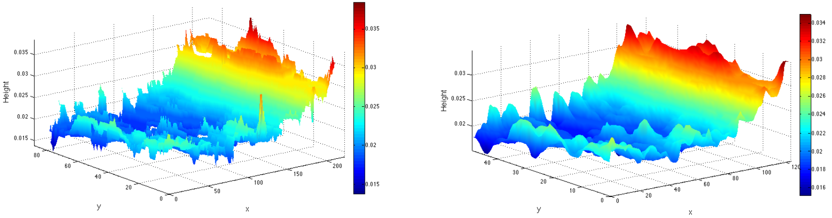
Table: Comparison of the CPU time (seconds) (2×2.66 GHz Quad-Core Intel Xeon) and memory demands (16 GB in total) with 10^5 level of state size between different methods.

Remarks

Our ensemble-based methods

- can yield comparable results as standard 4DVar.
- are especially pertinent and efficient compared to standard 4DVar when dealing with incomplete observations.
- work best if the statistics of ensemble fields correspond to the statistics of the background error (Parameter perturbation if possible).
- embrace two ensemble update schemes: The OP approach and the DT approach.
- embrace two localization schemes: The LC approach is more robust when dealing with incomplete observations whilst the LE approach needs less computational resources.
- employ the background ensemble repropagation scheme between two consecutive outer loops which is proved to be more effective.

Real image case: configuration



(a) (b)

Figure: (a) The height field observed from the kinect camera, (b) The corresponding height background state at $t = 0$.

Real image case: results

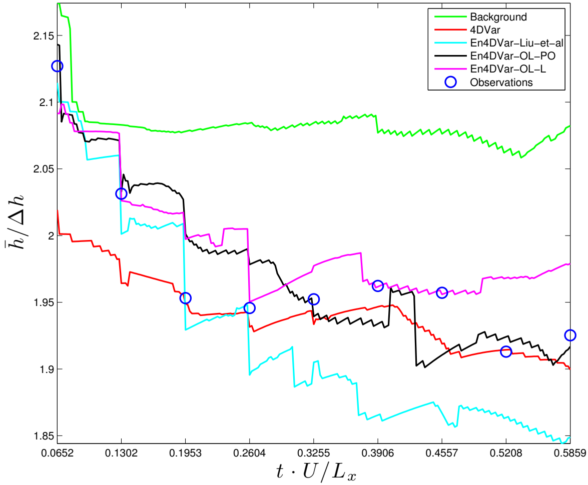


Figure: Mean surface height of the wave crest region as a function of time - comparison of different variational data assimilation approaches results

Remarks

In the case of the real image assimilation, we have shown that the ensemble-based methods are:

- easy to deal with complex, nonlinear observation model.
- to a large extent model independent.

Summary of image data assimilation

Compared to 4DVar:

- Ensemble repropagation redefines the subspaces to which the analysis increments belongs
- More reliable and accurate background error covariance magnitude and structure (tunable)
- Better reconstruction of unobserved component.
- More computationally efficient.

Compared to filtering method:

- provides directly a smoothing concerning all observations.

Outline

Shallow water equations under location uncertainty

Motivation

Explore high-resolution image data.

Common approaches

The Direct Numerical Simulation (DNS), the Large-Eddies Simulations (LES) and the Reynolds Average Numerical Simulations (RANS).

- DNS: Simulating the smallest scales of the flow, but requires extensive computational power.
- LES: Filtered governing equations are able to catch unsteady effects and requires moderate computational power.
- RANS: Time averaged and Least computational demanding,

SGS (Subgrid-scale) modeling

Additive random forcing terms.

Flow displacement field decomposition

The Lagrangian fluid particle displacement can be separated in two parts: a smooth differentiable part and a non-differentiable stochastic part:

$$d\boldsymbol{X}(x,t) = \mathbf{w}(\boldsymbol{X}(x,t),t)dt + \boldsymbol{\sigma}(\boldsymbol{X}(x,t),t)d\boldsymbol{B_t}, \tag{16}$$

where $d\boldsymbol{B_t}$ is a vector of d-dimensional Wiener increment (d=2,3).

Derivation process in deterministic approach

- Material derivative $\frac{d\phi}{dt} = \frac{\partial\phi}{\partial t} + w \cdot \boldsymbol{\nabla}\phi$.
- Differentiation of the integral over a material volume $\mathcal{V}$ (Reynolds transport theorem)

$$\frac{d}{dt}\int_{\mathcal{V}} q dx = \int_{\mathcal{V}} [\frac{\partial q}{\partial t} + \boldsymbol{\nabla} \cdot (q\mathbf{w})] dx. \tag{17}$$

Ito-Wentzell formula

In our case, $\phi(x, t)$ is necessarily a random function, consequently it must satisfy the stochastic differential of Ito-Wentzell formula:

$$d\phi(\boldsymbol{X}, t) = d_t\phi + \boldsymbol{\nabla}\phi \cdot d\boldsymbol{X} + \frac{1}{2}\sum_{i,j} d\langle X_i, X_j\rangle \frac{\partial^2\phi}{\partial x_i \partial x_j} dt + \sum_i d\langle \frac{\partial \phi}{\partial x_i}, \int_0^t (\boldsymbol{\sigma} \boldsymbol{B_t})^i \rangle dt$$
$$= 0. \quad (18)$$

This formula functions as the material derivative in the sense of deterministic derivation.

Stochastic Reynolds transport theorem (Mémin 2014)

$$d\int_{\mathcal{V}} q dx = \int_{\mathcal{V}} [d_t q + \boldsymbol{\nabla}\cdot(q\mathbf{w})dt - \frac{1}{2}\sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}q)dt + \boldsymbol{\nabla} q \cdot \boldsymbol{\sigma}\, d\boldsymbol{B_t})]dx. \quad (19)$$

Shallow water equations under location uncertainty

2D stochastic shallow water equation: Non-viscous and incompressible flow

Continuity equation:

$$d_t h + \Big(\frac{\partial (h\bar{u})}{\partial x} + \frac{\partial (h\bar{v})}{\partial y} - \frac{1}{2} \sum \partial_{x_i} \partial_{x_j} (a_{ij} h) \Big) dt + \frac{\partial (h\boldsymbol{\sigma} d\boldsymbol{B_t})_x}{\partial x} + \frac{\partial (h\boldsymbol{\sigma} d\boldsymbol{B_t})_y}{\partial y} = 0,$$

Momentum conservation equation:

$$d_t(\rho w_i) + \boldsymbol{\nabla} \cdot (\rho w_i \mathbf{w}) dt - \frac{1}{2} \sum_{ij} \frac{\partial^2}{\partial x_i \partial x_j} (a_{ij} \rho w_i) dt + \boldsymbol{\nabla} \cdot (\rho w_i \boldsymbol{\sigma} d\boldsymbol{B_t}) = (\rho \mathbf{g} - \boldsymbol{\nabla} p) dt.$$

1D stochastic shallow water equation in terms of horizontal mean field

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} \boxed{- \frac{1}{2}\frac{\partial^2 (a_{xx} h)}{\partial x^2}} = 0, \tag{20a}$$

$$\frac{\partial hu}{\partial t} + \frac{\partial (hu^2 + \frac{1}{2} g h^2)}{\partial x} \boxed{- \frac{1}{2}\frac{\partial^2 (a_{xx} hu)}{\partial x^2} - a_{xx}\frac{\partial h}{\partial x}\frac{\partial u}{\partial x}} = 0, \tag{20b}$$

Smagorinsky eddy viscosity model

$$\tau' = C \Delta_x \Delta_y \sqrt{\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \frac{1}{2}\left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)^2} S_{ij}. \tag{21}$$

Estimation of the variance tensor from data

Estimator from realized temporal variance

Suppose we have a sequence of small-scale velocity data ranging form 0 to T: $w_0, w_1, \ldots, w_M$, the second moment of random variable $\boldsymbol{X}$ can be approximated by its observed values:

$$\sum_{i=0}^{T} \mathbf{a} = \Delta t \sum_{i=0}^{T} \big(w(t_{i+1}) - w(t_i)\big)\big(w(t_{i+1}) - w(t_i)\big)^T.$$

Estimator from realized spatial variance

$$\mathbf{a}(t_i, x_k) = C \frac{\delta t}{n^2 - 1} \sum_{l \in \mathcal{C}} \left[(w(t_i, x_l) - \bar{w}(t_i, x_{\mathcal{C}}))(w(t_i, x_l) - \bar{w}(t_i, x_{\mathcal{C}}))^T\right].$$

Estimator from realized ensemble variance

$$\mathbf{a} = C \frac{(\Delta t_i)^2}{T} \sum_{i=0}^{T} \frac{1}{N-1} \sum_{j=1}^{N} [w^j(t_i) - \bar{w}(t_i)][w^j(t_i) - \bar{w}(t_i)]^T.$$

Estimation of the variance tensor from DA

Cost function

$$J(\delta \mathbf{x_0}, \delta \mathbf{a}) = \frac{1}{2}\|\delta \mathbf{x_0}(x)\|^2_{\mathbf{B_x}} + \frac{1}{2}\|\delta \mathbf{a}\|^2_{\mathbf{B_a}} + \frac{1}{2}\int_{t_0}^{t_f} \|\partial_{\mathbf{x}}\mathbb{H}\ \delta \mathbf{x}(t, x, \mathbf{a}) - \mathbf{D}(t, x, \mathbf{a})\|^2_{\mathbf{R}} dt, \tag{22}$$

where the innovation vector $\mathbf{D}(t, x, \mathbf{a})$ is defined as:

$$\mathbf{D}(t, x, \mathbf{a}) = \mathcal{Y}(t, x) - \mathbb{H}(\varphi_t(\mathbf{x_0}(x), \mathbf{a})), \tag{23}$$

Possible approaches

- Parameter estimation scheme in variational system: adjoint based gradient calculation
- Parameter estimation scheme in ensemble-based system: empirical based gradient calculation

Ensemble-based parameter estimation

State of the art

Simultaneous approach: EnKF (Zupanski and Zupanski, 2006; Yang and Delsole, 2009), EnSRF (Tong and Xue, 2008), LETKF (Kang et al, 2011), IEnKS (Bocquet and Sakov, 2013).
Separately approach: (Koyama and Watanabe, 2010).

Simultaneous approach

Estimating the initial condition and the parameters at the same time

Parameter structures

In our case, the variance tensor **a** is a function of both time and space. It is necessary to introduce cycling procedure.

Parameter evolution model

The parameter is propagated to the next cycle $\mathbf{a}^f_{j+1} = \mathbf{a}^f_j + \beta \widehat{\delta \mathbf{a}_j}$ up to a relaxing coefficient $\beta \in (0.2, 0.5)$.

- Define the augmented state vector $\mathbf{s}_0 = [\mathbf{x}_0, \mathbf{a}_0] \in \mathbb{R}^{n+p}$
- Initialize and integrate the augmented ensemble $\mathbf{S}_0$ which can be a function of both initial conditions and parameters
- Calculate the augmented ensemble perturbation matrix: $\mathbf{A}'_s := \frac{1}{\sqrt{N-1}}(\mathbf{S}_0^1 - \langle \mathbf{S}_0 \rangle, \dots, \mathbf{S}_0^N - \langle \mathbf{S}_0 \rangle)$.
- Minimize the cost function in terms of the augmented state vector:

$$J(\delta \mathbf{s}_0) = \frac{1}{2}\|\delta \mathbf{s}_0\|^2_{\mathbf{B}^f_\mathbf{s}} + \frac{1}{2}\int_{t_0}^{t_f} \|\widetilde{\partial_{\mathbf{x}}\mathbb{H}}\, \partial_{\mathbf{x}}\varphi'_t(\mathbf{s}_0)\delta \mathbf{s}_0 - \mathbf{D}(t, x, \mathbf{a})\|^2_{\mathbf{R}} dt,$$

with

$$\mathbf{B}^f_\mathbf{s} = \mathbf{A}'_s {\mathbf{A}'_s}^T = \begin{pmatrix} \mathbf{B}_{xx} & \mathbf{B}^T_{ax} \\ \mathbf{B}_{ax} & \mathbf{B}_{aa} \end{pmatrix}.$$

Ensemble-based parameter estimation

Parameter estimation effect

- $\mathbf{a}$ explicitly manifests itself on the model integration in the form of the estimated parameter $\widehat{\mathbf{a}}$.
- $\mathbf{a}$ implicitly affects the analysis state through the ensemble spread from which the propagation of the ensemble perturbation matrix is calculated.

Localization technique

The localized covariance approach has to be extended to the spatial parameter covariance and the cross-covariance between the state variable and the parameter. That is to say, we use

$$\mathbf{P}_s^f = \mathcal{C}_s \odot \mathbf{B}_s^f = \begin{pmatrix} \mathcal{C}_{xx} & \mathcal{C}_{ax}^T \\ \mathcal{C}_{ax} & \mathcal{C}_{aa} \end{pmatrix} \odot \begin{pmatrix} \mathbf{B}_{xx} & \mathbf{B}_{ax}^T \\ \mathbf{B}_{ax} & \mathbf{B}_{aa} \end{pmatrix}, \qquad (24)$$

Parameter identifiability

Definition

A criteria to decide whether or not the parameter of interest can be inferred from data

Possible approaches

- An output response function:

$$J_y(\mathbf{a}) = \int_0^T \|\mathcal{Y}(t,x) - \mathbb{H}(\varphi_t(\mathbf{x_0}(x),\mathbf{a}))\|^2 dt.$$

- The correlation coefficient:

$$cor(\mathbf{x},\mathbf{a}) = \frac{cov(\mathbf{x},\mathbf{a})}{\sqrt{var(\mathbf{x})var(\mathbf{a})}}.$$

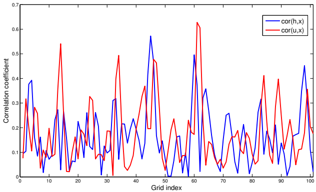


Figure: The spatial distribution of the correlation coefficients between the state variable h, u and a: $cor(h,a)$ (blue line) and $cor(u,a)$ (red line).

Experimental settings

Initial state

$h(x,0) = H_0 - \frac{fU_0x}{g} + A\xi$, with $L = 6000\text{km}$, $H_0 = 5000\text{m}$, $f = 1.03 \times 10^{-4}\text{s}^{-1}$, $U_0 = 40\text{m/s}$ and g is the gravity acceleration. And ξ is a random Gaussian covariance field (with de-correlation length equals to 20%L).

1D test

- Simulated state resolution 101 grid points.
- True state resolution 401 grid points.
- 3 cycling windows.

2D test

- Simulated state resolution 21×51 grid points.
- True state resolution 81×401 grid points.
- 5 cycling windows.

Results and Discussions: 1D synthetic

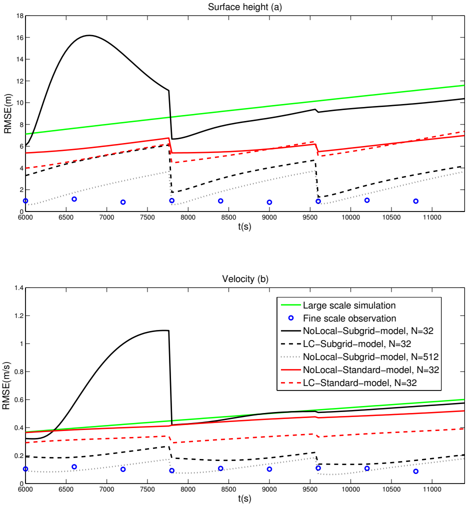


Figure: RMSE comparison in terms of free surface height (a) and velocity (b) between various configurations of 4DEnVar.

Results and Discussions: 1D synthetic, Subgrid model term effects:

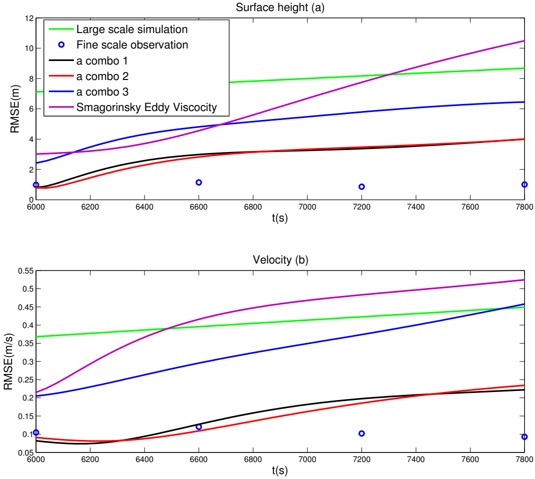


Figure: RMSE comparison in terms of free surface height (a) and velocity (b) between various subgrid model configurations of stochastic shallow water model.

Results and Discussions: 2D synthetic, RMSE results

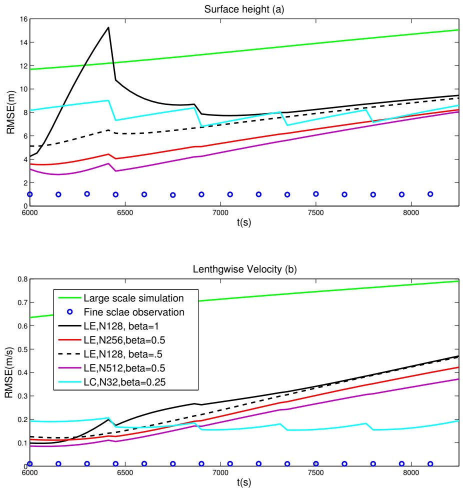
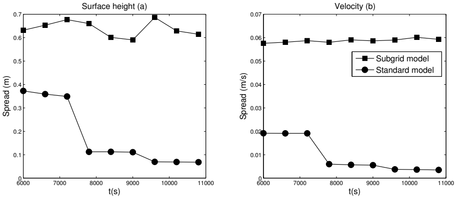
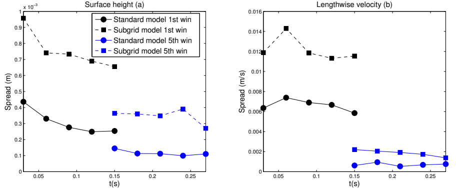


Figure: RMSE comparison in terms of free surface height (a) and lengthwise velocity (b) between various configurations of 4DEnVar.

Results and Discussions: Ensemble spread



(a)



(b)

Figure: Ensemble Spread in terms of free surface height (a) and lengthwise velocity (b): standard model (circle) with ensemble in function of initial states; subgrid model (square) with ensemble in function of initial parameters.

Remarks

- Compared to standard model, our model under location uncertainty can better reflect the interaction between resolved parts and unresolved parts of the flow.
- Compared to standard model, ensemble initialized by variance tensor perturbation can better maintain the ensemble spread.
- High-resolution image data can be used to its full extent without the need to run dynamical model on fine grids.
- Compared to Smagorinsky eddy viscosity model, the term emerged in the mass conservation equation is proved to be useful in some cases.

Outline

Conclusion & Future work

Conclusion

In general, the variational ensemble methods proposed can better cope with image data in terms of bad background (error) quality, unobserved components, high resolution

Future work

- SST image application
- Complex model application
- Uncertainty subgrid estimation: non Gaussian parameter distribution and constraints.