Ocean Modeling: Bias correction through stochastic forcing.

Canter Martin [1] Barth Alexander [1]

[1]GeoHydrodynamic and Environmental Research (GHER), University of Liège, Liège, Belgium

8 May 2015

Overview

1. Motivation
2. Our Approach: why is it innovative?
3. Lorenz '96 model
    - Model behaviour
    - Twin Experiment
4. NEMO-LIM Model:
    - Stochastic Forcing
    - Twin Experiment
5. Conclusion

Motivation: Large scale models

Most **low resolution** models suffer important errors due to poorly represented processes. This leads to a **systematic error** with a non-zero mean: bias.

Bias is considered to be the **main source** of errors in climatic model. It allows only to study the variation of a model, not its absolute results. (Zunz, 2012)

Both off-line and on-line methods aim at correcting the bias during the assimilation procedure. The **source** of bias remains.

Bias correction: Innovative approach

How to correct bias with data assimilation?

- Estimate the model's bias and its source in the model's equations.
- Create an ensemble of stochastic forcing directly added into the model's equations.
- Run the model for each forcing field separately.
- Consider this stochastic forcing as a control variable for data assimilation.
- Adjust the forcing field with data assimilation to correct the bias.
- Rerun the model with the new forcing field.
- Validate the bias correction with external and independent data.

$\Longrightarrow$ This method allows a **continuous correction** for bias during the model run, since the source of bias itself is corrected.

Lorenz '96 model: Mean Model State

Modified Lorenz '96 model equation:

$$\frac{dX_k}{dt} = -X_{k-2}X_{k-1} + X_{k-1}X_{k+1} - X_k + \mathbf{F_k} \tag{1}$$

Model is run with classic configuration:

- $0 < \mathbf{F_k} < 10$, where $k = 1, .., 40$
- 1000 time step of $t = 0.05$

However, we do not look at the variables at a specific point in time, but rather at the model's spatial and temporal mean.

- 15 different initial condition $\mathbf{X_k}$ for each $0 < \mathbf{F_k} < 10$.
- Temporal average from $200th$ time step, average over the $i = 15$ initial conditions
- **Model Mean State**: $X_k = \overline{X}_{k,i,t}$

Lorenz '96 model: Mean Model State

Model Mean State: 30 evenly distributed different $0 < \mathbf{F_k} < 10$, with 450 different initial conditions for each mean $\mathbf{F}$:

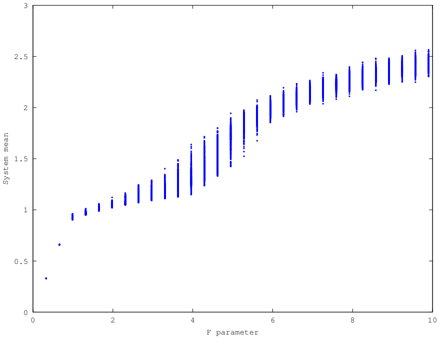

Figure : Lorenz '96 model mean state compared to a constant forcing parameter **F** .

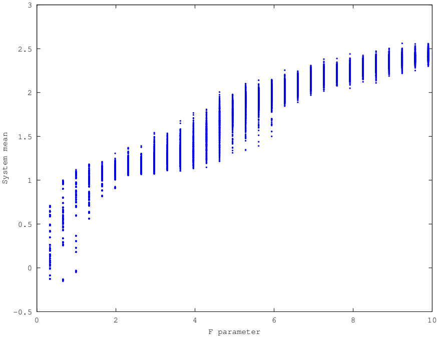

Figure : Lorenz '96 model mean state compared to a mean spatially variable forcing parameter $\mathbf{F_k}$, with a mean **F** and standard deviation of $\mathbf{F}_{\mathrm{std}} = 1$.

Lorenz '96 model: Twin Experiment

We tested our innovative approach with a Twin Experiment with the Lorenz '96 model:

- $\overline{\mathbf{F}}_{\text{tru}} = 4$, with spatially correlated perturbation

$$F_k = F_{k,\text{tru}} + z \qquad (2) \qquad\qquad z \sim N(0, P) \qquad (3)$$

- One run is considered as reality.
- Observations are extracted: spatial model mean $\overline{\mathbf{X}}_{\mathbf{k},\text{tru}}$
- Ensemble is built with 100 runs with different $\mathbf{F}_{\mathbf{k},\text{ens}}$.

Perturbation is considered as bias, we intend to find it and correct the model:

- State vector consists of $\mathbf{F}_{\mathbf{k},\text{ens}}$ and $\overline{\mathbf{X}}_{\mathbf{k},\text{ens}}$
- We assimilate $\overline{\mathbf{X}}_{\mathbf{k},\text{tru}}$, and correct our ensemble

Lorenz ’96 model: Twin Experiment

Lorenz ’96 Twin Experiment Results

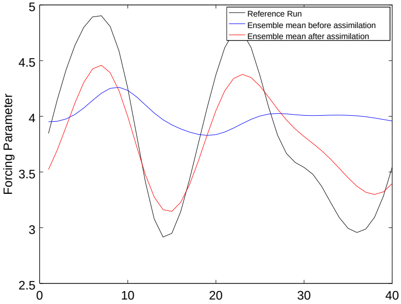

Figure : Lorenz ’96 model $\mathbf{F}_\mathbf{k}$ parameter of the truth, ensemble mean and assimilated ensemble mean runs.

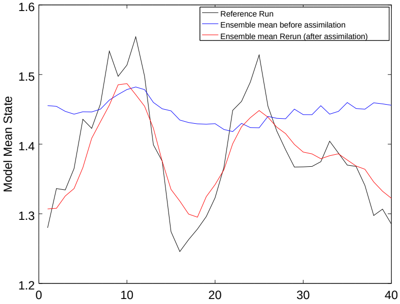

Figure : Lorenz ’96 model $\mathbf{X}_\mathbf{k}$ temporal mean of the truth, ensemble mean and assimilated ensemble mean runs.

NEMO-LIM Model

This method is currently being applied to the NEMO-LIM model.

- Global and **low resolution** (2°) coupled model with long time steps allowing simulations over several decades.
- Used in the PredAntar project (Belspo), which aims at understanding and predicting the Antartic sea ice variability at the decadal timescale.

Because of this low resolution, **ocean currents** are poorly represented and have been identified as a possible source of bias. They have a direct impact on heat transportation in the ocean, thus also on the sea surface temperature bias.

Forcing the model

Current investigation: poorly located currents in NEMO-LIM model. We apply the forcing terms diretly into the **momentum equations** of ocean dynamics in NEMO.

$$\frac{du}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + fv + \frac{1}{\rho}\frac{\partial \tau_x}{\partial z} + F_u \tag{4}$$

$$\frac{dv}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial y} - fu + \frac{1}{\rho}\frac{\partial \tau_y}{\partial z} + F_v \tag{5}$$

Constructing the forcing term: Diva-ND

Random field is created with Diva-ND (Barth et al., 2013). It allows to apply **constraints** by using a **cost function**:

1. Penalizes abrupt variations
2. Uses a given correlation length
3. Decouples disconnected areas based on topography

$$J(\Psi) = \int_{\Omega} \alpha_2 (\nabla^2 \Psi)^2 + \alpha_1 (\nabla \Psi)^2 + \alpha_0 \Psi dx \qquad (6)$$

Constructing the forcing term: Diva-ND

However, in order not to create currents **perpendicular** to the coasts, we tried different additional constraints on the stream function:

1. Constraining the stream function to be zero at the coasts. Too restrictive.
2. Constraining the gradient of the stream function along the coast. We use a strong constraint: forced to be zero.

$$J'(\Psi) = J(\Psi) + \lim_{\sigma \to 0} \frac{1}{\sigma} \int_{\Delta\Omega} (\nabla\Psi \bullet \vec{t})^2 ds \qquad (7)$$

Additionnal filtering to smoothen the first derivative, for a higher model stability.

Constructing the forcing term: Stream Function

We use this field as a stream function to construct zonal and meridional currents. The condition of zero divergence resulting from flow incompressibility gives us:

$$u = -\frac{\partial \Psi'}{\partial y} \quad (8) \qquad v = \frac{\partial \Psi'}{\partial x} \quad (9)$$

The zonal and meridional currents are then dampened towards depths depending on the yearly mean **turbocline** ($T(x,y)$), in order to keep surface currents, defined as the ocean mixed layer thickness.

$$F_u(x,y,z) = \frac{u(x,y)}{1 + \exp \frac{z - T(x,y)}{L}} \quad (10) \qquad F_v(x,y,z) = \frac{v(x,y)}{1 + \exp \frac{z - T(x,y)}{L}} \quad (11)$$

Model Forcing: Twin Experiment

With forcings on the meridional and zonal currents, we are able to obtain an internal variability of the **Sea Surface Height** (SSH) of about $28cm$, which can be compared with the RMS between a NEMO free run and the CNES Mean Dynamic Topography, of $20cm$.

We build a Twin Experiment similarly to the Lorenz '96 case:

- Ensemble of forcings on zonal and meridional currents
- The yearly mean SSH is considered as observation and control variable
- We build and ensemble of forcings, and assimilate observations from a truth run

Model Forcing: Twin Experiment Local Assimilation

Local assimilation correlation lenght: 2000km.

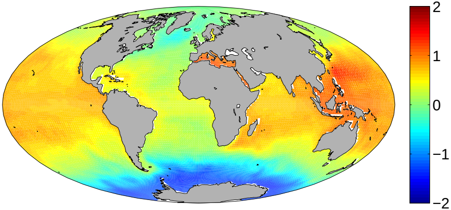

Figure : SSH NEMO Ensemble Mean before assimilation (in m)

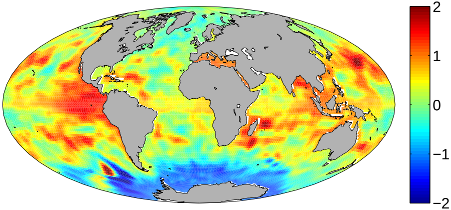

Figure : SSH NEMO Twin True run (in m)

Model Forcing: Twin Experiment Local Assimilation

Local assimilation correlation lenght: 2000km.

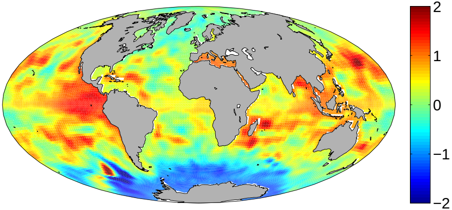

Figure : SSH NEMO Ensemble Mean after assimilation (in m)

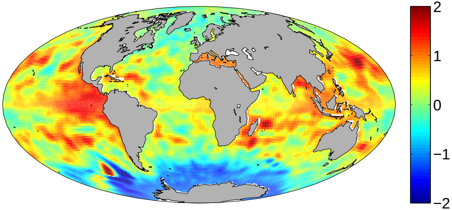

Figure : SSH NEMO Twin True run (in m)

Model Forcing: Twin Experiment Local Assimilation

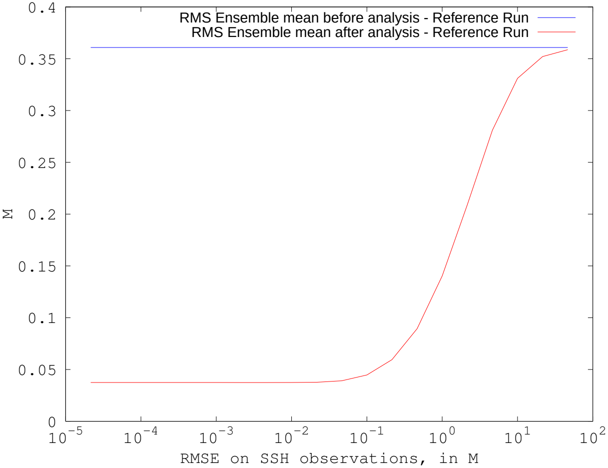

Figure : RMS on SSH from Ensemble Mean before and after analysis, with True Run

Model Forcing: Twin Experiment Local Assimilation

Background estimate of $F_u \approx 0$.

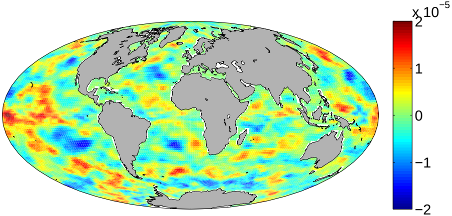

Figure : NEMO Ensemble mean after analysis, Zonal Forcing

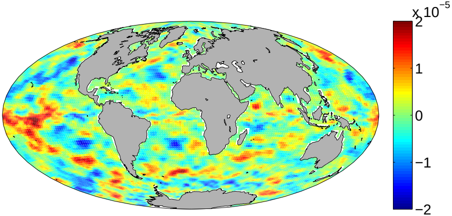

Figure : NEMO True Run, Zonal Forcing

Model Forcing: Twin Experiment Local Assimilation

Background estimate of $F_v \approx 0$.

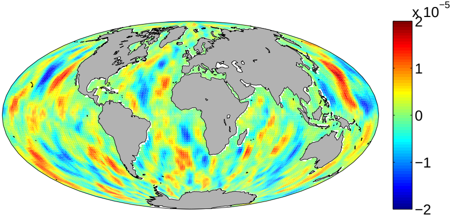

Figure : NEMO Ensemble mean after analysis, Meridional Forcing

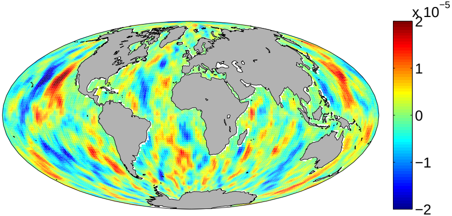

Figure : NEMO True Run Meridional Forcing

Model Forcing: Twin Experiment Local Assimilation

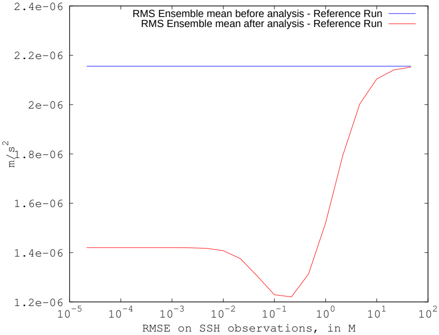

Figure : RMS on Zonal Forcing from Ensemble mean before and after Analysis, with True Run

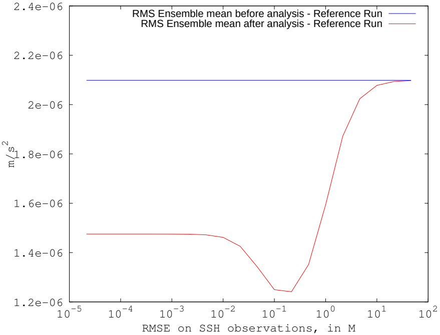

Figure : RMS on Meridional Forcing from Ensemble mean before and after Analysis, with True Run

Model Forcing: Twin Experiment Local Assimilation

NEMO Rerun, with the assimilated U and V forcings.

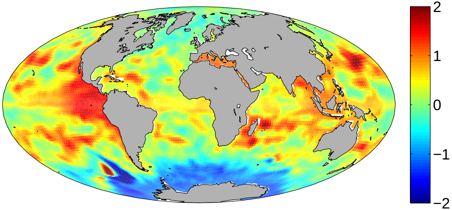

Figure : SSH NEMO Rerun (in m)

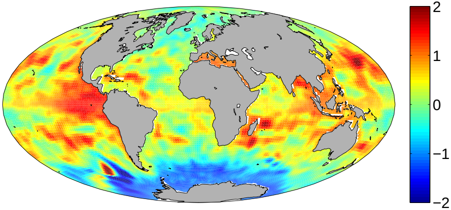

Figure : SSH NEMO Twin True run (in m)

Model Forcing: Twin Experiment Local Assimilation

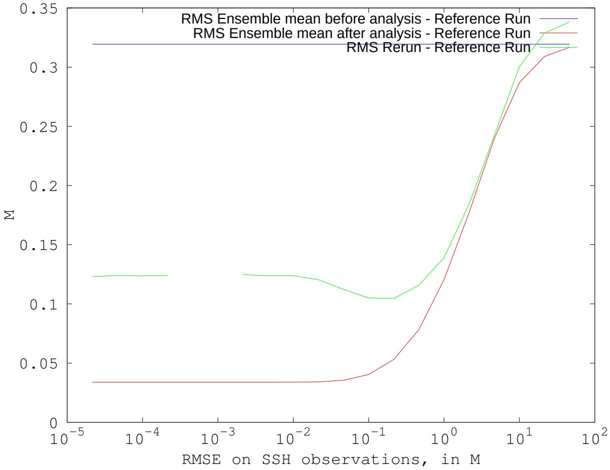

Figure : RMS on SSH from Ensemble Mean before and after analysis, and from Rerun, with True Run

Conclusion and perspectives

- Results with Lorenz '96 model were encouraging enough to start testing this method on NEMO model.
- Model stability problems with forcing have been handled
- Variability on SSH from forcing is large enough to correct the estimated bias from the model
- Local assimilation procedure gives encouraging results
- Rerun of the model with analysed forcings ongoing
- Real data and observations (CNES MDT) will be used to correct the model after the Twin Experiment

Thank you !