Application of Analog Data Assimilation to Simulated/Real Sea-Surface Height Interpolation

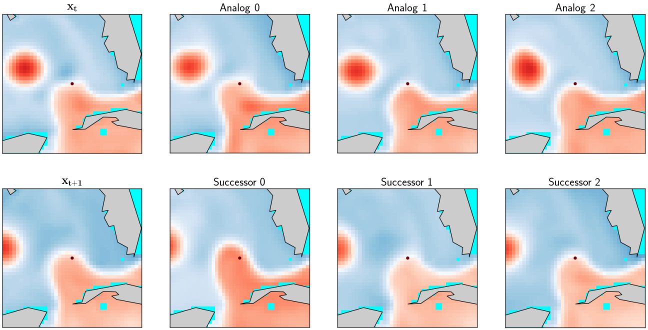
Pierre Tandeo and Yicun Zhen

IMT Atlantique, Lab-STICC, UBL, Brest, France

February 10, 2020

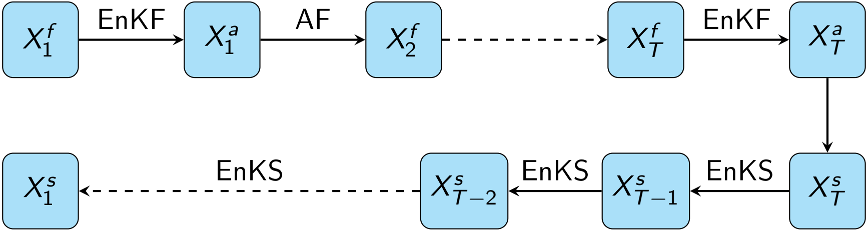
The analog forecast method

- Have a huge amount of historical data (the catalog);
- For a given initial state $\mathbf{x}_t$, find the similar states (analogs) in the historical database;
- Calculate $\mathbf{x}_{t+1}$ basesd on the analogs and the corresponding successors.



Analog Data Assimilation (AnDA) for Reanalysis

- Ensemble Kalman filter(EnKF) to calculate the state analysis;
- Analog forecast (AF) for state forecast;
- Ensemble Kalman smoother (EnKS) for calculating the state reanalysis.



The elements of AnDA

To implement AnDA, you need:

- 1, choose your catalog (dimension reduction, distance function, etc.);
- 2, choose your data assimilation scheme (EnKF, EnKS, covariance localization/inflation, etc.)
- 3, choose the analog forecast scheme (locally linear model, number of analogs, etc.)

In the experiments of this presentation,

- the catalog is always the time series of the first 100 EOF coefficients of the dataset.
- EnKS (1000 ensemble members) is alwasys used. No covariance localization. Only the natural covariance inflation in analog forecast.
- locally linear model is always chosen for the local model.

Caution of using AnDA

- 1, Analog forecast can also have huge model error! The catalog represents the hidden model you are using.
- 2, Catalog must be large enough. How large? Is my dataset large enough? Try!

Objective interpolation (OI)–a widely used model-free method for calculating the reanalysis

Thanks Maxime Beauchamp for the detailed introduction of OI!

Reanalysis of simulated SSH at Gulf of Mexico

- Dataset: OCCIPUT simulated SSH of 50 members and 20 years;
- Catalog: the time series of the first 100 principal components of OCCIPUT dataset;
- Obs: simulated along-track obs (without error) of SSH from altimeters in 2004.

⇒ Task: Compare the reanalysis results of AnDA and OI with the known truth.

Reanalysis of simulated SSH at Gulf of Mexico

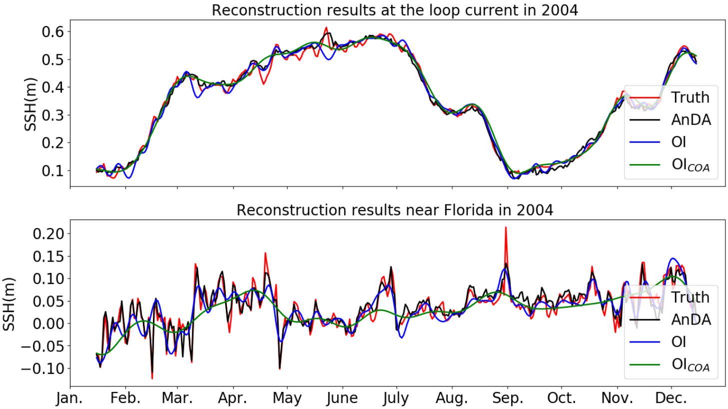


Figure 1: (OI = well tuned OI.) (OI$_{COA}$ = OI with large temporal correlation.)

Reanalysis of simulated SSH at Gulf of Mexico

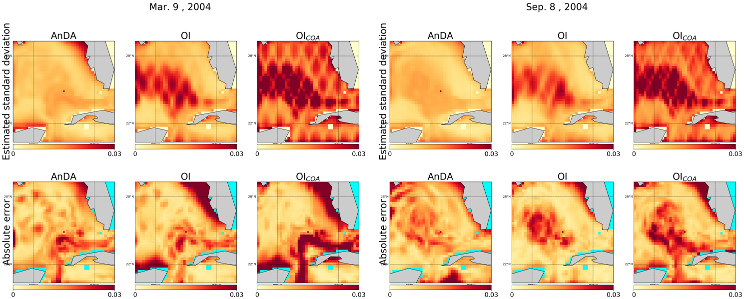


Figure 2: Error v.s. estimated standard deviation

Reanalysis of simulated SSH at Gulf of Mexico

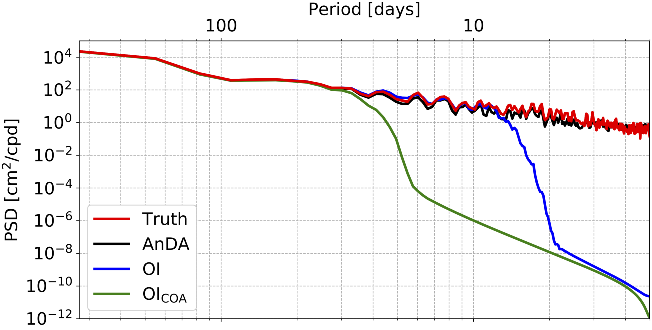


Figure 3: Power spectral density

Reanalysis of simulated SSH at Gulf of Mexico

Summary of the comparison results of AnDA and OI using simulated data:

- 1, Similar mean state estimate;
- 2, AnDA has flow-dependent standard deviation;
- 3, AnDA produces more complete power spectral density.

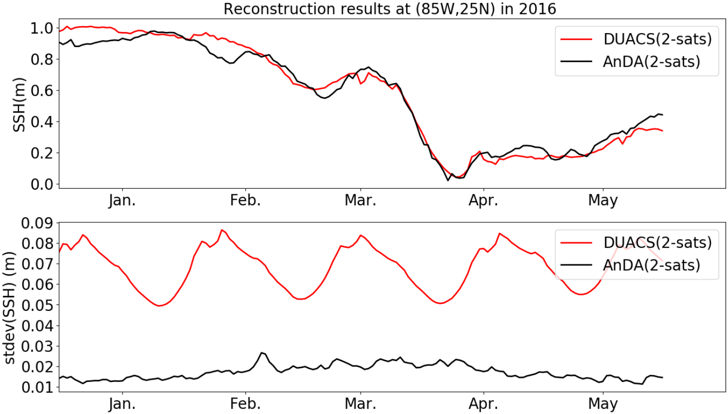
Why? Good and large enough catalog (19 years X 49 members)!

Reanalysis of Real SSH at Gulf of Mexico

- Catalog: the 20-year-long time series (from 1998.1.1 to 2018.12.31) of the first 100 principal components of (SSH,SST), where SSH is from the DUACS reanalysis, SST is from REMSS reanalysis.
- Obs: real satellite altimetry obs and REMSS's daily SST maps from 2015.12.1 to 2016.5.31.
- OI results: the DUACS reanalysis using obs from only two satellites.

$\Rightarrow$ Task: Compare the reanalysis results of AnDA and OI with the unused satellite obs.

Reanalysis of Real SSH at Gulf of Mexico



$\text{RMSE(AnDA)}/\sqrt{R} = 1.23$
$\text{RMSE(DUACS)}/\sqrt{R} = 1.38.$

Reanalysis of Real SSH at Gulf of Mexico

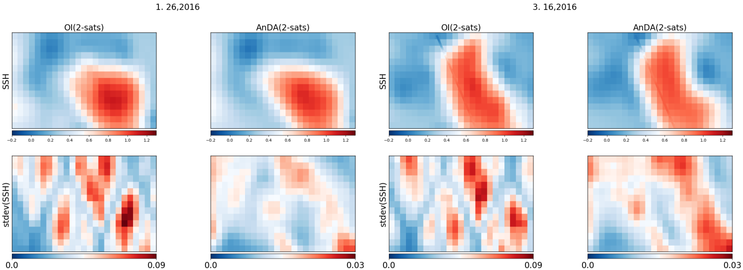


Figure 4: The stdev estimated by AnDA is still flow-dependent.

Reanalysis of Real SSH at Gulf of Mexico

Summary of the comparison results of AnDA and DUACS (OI):

- 1, AnDA produces better mean state estimate;
- 2, AnDA has flow-dependent standard deviation.

Remarks:

- Here we use 100 EOFs for AnDA. The result will be worse than DUACS if we use only 50 EOFs.

Future plan

Start with DUACS and REMSS time series, use AnDA and OI interchangably and iteratively to rebuild the 20-year-long reanalysis:

$$SSH_{DUACS} + SST_{REMSS} \Rightarrow SSH^{(1)}_{AnDA} + SST^{(1)}_{AnDA} + P^{s,(1)}_{AnDA} \Rightarrow$$
$$SSH^{(1)}_{OI} + SST^{(1)}_{OI} \Rightarrow SSH^{(2)}_{AnDA} + SST^{(2)}_{AnDA} + P^{s,(2)}_{AnDA} \Rightarrow \ldots$$

Thank you!

Lguensat, R., Tandeo, P., Ailliot, P., Pulido, M., and Fablet, R. (2017).
The Analog Data Assimilation.
Monthly Weather Review, 145(10):4093–4107.

Tandeo, P., Ailliot, P., Ruiz, J. J., Hannart, A., Chapron, B., Easton, R., and Fablet, R. (2015).
Combining analog method and ensemble data assimilation: application to the Lorenz-63 chaotic system.
In Machine Learning and Data Mining Approaches to Climate Science, pages 3–12.