Autolocalization techniques for ocean modelling using OpenDA

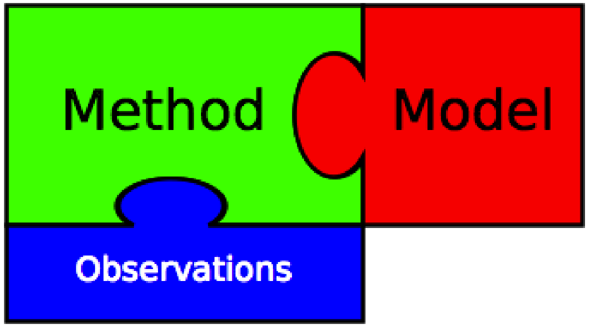
Umer Altaf, Nils van Velzen, Martin Verlaan, Arnold Heemink

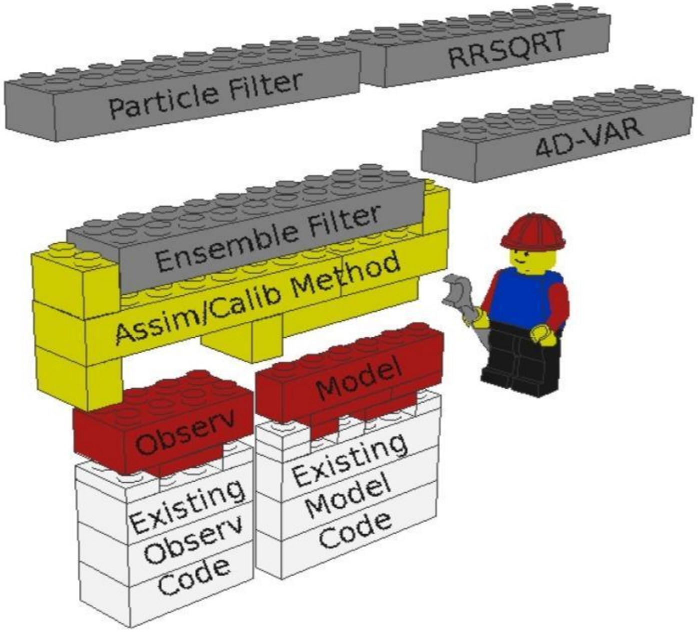
Outline

- OpenDA
- Ensemble methods and localization
- Automatic localization techniques
- Experiments with NEMO model

OpenDA: framework for Data Assimiation

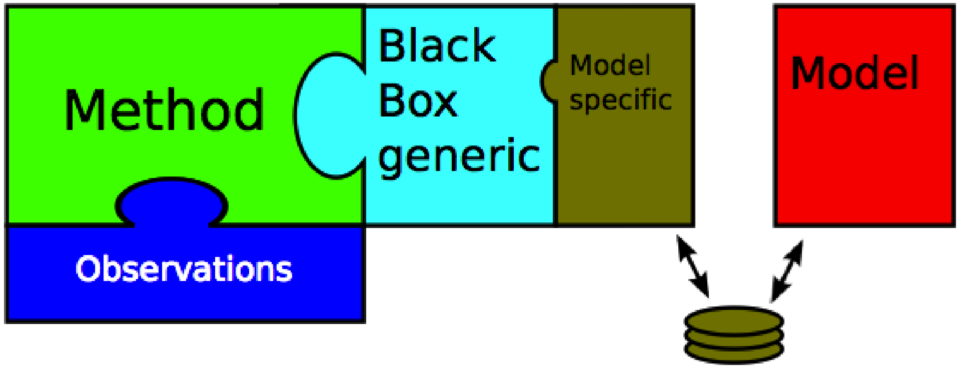
- Content:
  - Set of interfaces that define interactions between components
  - Library of data-assimilation algorithms
  - DA philosophy
  - Building blocks only need to be implemented once





OpenDA: framework for Data Assimiation

- Black box coupling
- Model needs proper restart functionality
- + Easy to implement
- + No change to model code
- - Restart/file overhead
- - localization might be difficult to implement



Localization

- Ensemble Kalman methods:

$$\hat{\mathbf{P}}_k^f = \frac{1}{n-1}\sum_{i=1}^{n}\left(\mathbf{x}_{k,i}^f - \hat{\mathbf{x}}_k^f\right)\left(\mathbf{x}_{k,i}^f - \hat{\mathbf{x}}_k^f\right)^T$$

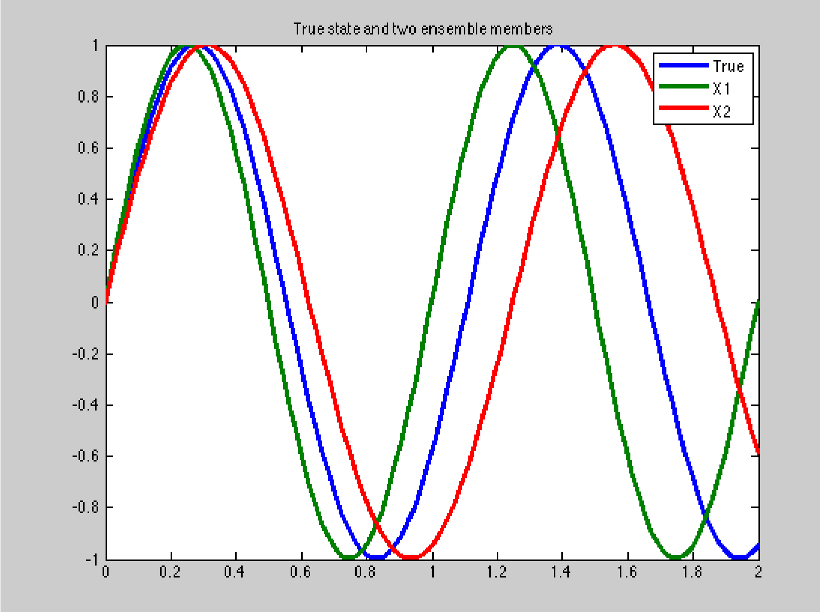
$$\mathbf{K}_k = \hat{\mathbf{P}}_k^{xh}\left(\hat{\mathbf{P}}_k^{hh} + \mathbf{R}_k\right)^{-1}$$

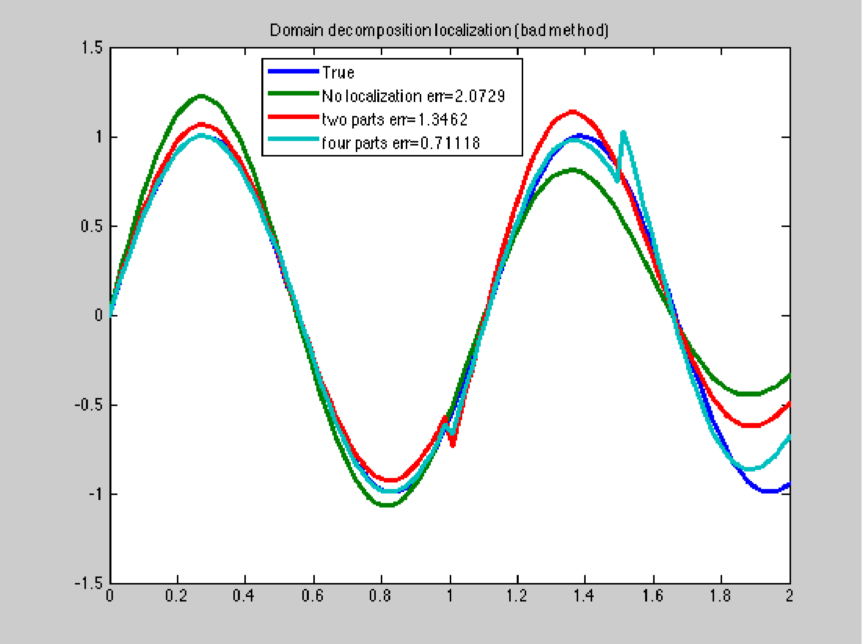
- Analyzed ensemble is given by:

$$\mathbf{x}_{k,i}^a = \mathbf{x}_{k,i}^f + \mathbf{K}_k\left(\mathbf{y}_{k,i}^s - \mathcal{H}_k(\mathbf{x}_{k,i}^f)\right), i = 1, 2, ..., n,$$

Localization

- Ensemble is set of “basis” vectors for representing errors and update





Localization

- Covariance: low rank approximation compared to model state:

$$\mathcal{O}(10-100)$$

Localization

- Covariance: low rank approximation compared to model state:

$$\mathcal{O}(10-100)$$

- Structural underestimation of errors
- Spurious correlations

Localization

- Localization of gain matrix

$$\mathbf{K}^{Loc} = \mathbf{K} \circ \beta_{xy}$$

- Increases "dimension" of the update

Localization

- Localization of gain matrix

$$\mathbf{K}^{Loc} = \mathbf{K} \circ \beta_{xy}$$

- Increases "dimension" of the update
- How to determine β_{xy} ?

Localization

- Localization of gain matrix

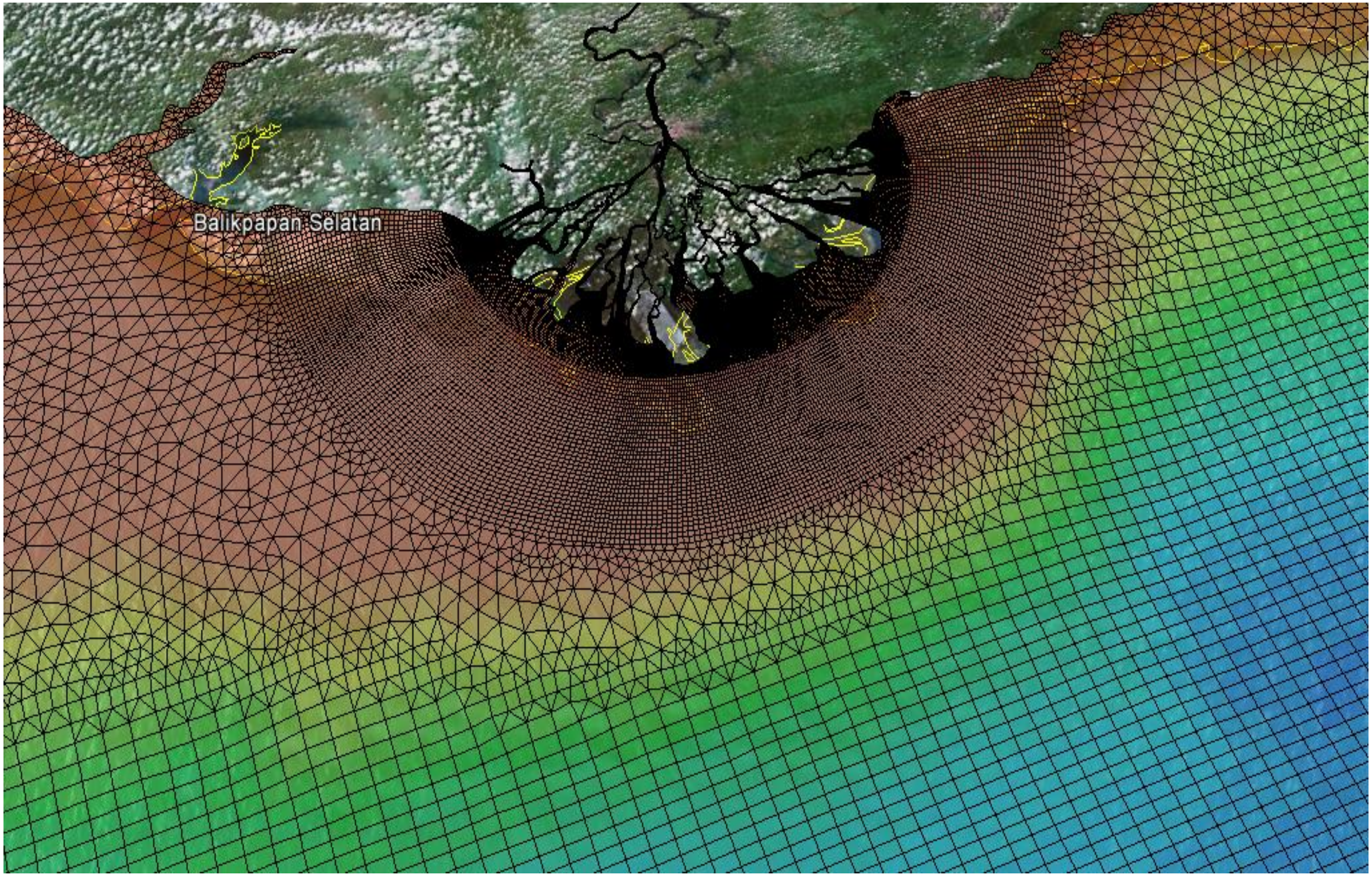
$$\mathbf{K}^{Loc} = \mathbf{K} \circ \beta_{xy}$$

- Increases "dimension" of the update
- How to determine β_{xy} ?
- Is distance a good measure?
- What is the location of state variables

Localization

Source:wikipedia

Localization



Source:deltares

Auto Localization

- Anderson 2004:
  - Define N_g groups of N_e ensembles
  - Each group has its own gain matrix $\mathbf{K}_j, j = 1, .., N_g$

Auto Localization

- Anderson 2004:
  - Define N_g groups of N_e ensembles
  - Each group has its own gain matrix $\mathbf{K}_j, j = 1, .., N_g$

  - Assume elements from gain matrices are drawn from distribution containing the "true" gain matrix.
  - Minimize:

$$\sqrt{\sum_{j=1}^{N_g} \sum_{k=1, k\neq j}^{N_g} (\alpha_i k_{i,k} - k_{i,j})}.$$

  - Too much computational work for real time application
  - Good for investigating good selections of weights

Auto Localization

- Zhang and Oliver 2011:
  - Create ensemble (N_B) of gain matrices based on the initial ensemble using bootstrapping
  - Estimation of variance of each element of the gain matrix

$$\hat{\sigma}^2_{k_{i,j}} = \frac{\sum_{m=1}^{N_B} \left(\hat{k}^*_{i,j,m} - \bar{k}_{i,j}\right)^2}{N_B}$$

  - Ratio between mean and variance

$$\hat{C}^2_{v_{i,j}} = \frac{\hat{\sigma}^2_{\theta_{i,j}}}{\bar{k}^2_{i,j}}$$

  - Localization weights

$$\beta_{i,j} = \frac{1}{1 + \left(1 + 1/\sigma^2_\alpha \hat{C}^2_{v_{i,j}}\right)}$$

  - Balance parameter (Zhang and Oliver 2010) $\sigma^2_\alpha = 0.36$

Coupling NEMO into OpenDA

- Namelist File:
  - NemoNamelistFileWrapper
  - Edit namelist file after each analysis step.
- Restart.nc
  - NemoRestartFileWrapper
  - Modifying data at analysis step
- nemo_exchangeitem
  - Input and Output exchange items interface

NEMO-OpenDA

```
<?xml version="1.0" encoding="UTF-8"?>
<openDaApplication xmlns="http://www.openda.org">

<stochObserver className="org.openda.observers.NoosTimeSeriesStochObserver">
     <workingDirectory>./stochObserver</workingDirectory>
     <configFile>noosObservations.xml</configFile>
</stochObserver>
```

NEMO-OpenDA

```xml
<?xml version="1.0" encoding="UTF-8"?>
<openDaApplication xmlns="http://www.openda.org">

<stochObserver className="org.openda.observers.NoosTimeSeriesStochObserver">
     <workingDirectory>./stochObserver</workingDirectory>
     <configFile>noosObservations.xml</configFile>
</stochObserver>

<stochModelFactory className="org.openda.blackbox.wrapper.BBStochModelFactory">
     <workingDirectory>./stochModel</workingDirectory>
     <configFile>nemoStochModel.xml</configFile>
</stochModelFactory>
```

NEMO-OpenDA

```
<?xml version="1.0" encoding="UTF-8"?>
<openDaApplication xmlns="http://www.openda.org">

<stochObserver className="org.openda.observers.NoosTimeSeriesStochObserver">
     <workingDirectory>./stochObserver</workingDirectory>
     <configFile>noosObservations.xml</configFile>
</stochObserver>

<stochModelFactory className="org.openda.blackbox.wrapper.BBStochModelFactory">
     <workingDirectory>./stochModel</workingDirectory>
     <configFile>nemoStochModel.xml</configFile>
</stochModelFactory>

<algorithm className="org.openda.algorithms.kalmanFilter.EnKF">
     <workingDirectory>algorithm</workingDirectory>
     <configString>enkf.xml</configString>
</algorithm>

<resultWriter className="org.openda.resultwriters.NetcdfResultWriter">
     <workingDirectory>.</workingDirectory>
     <configFile>enkf_.nc</configFile>
</resultWriter>

</openDaApplication>
```

NEMO-OpenDA

```
<?xml version="1.0" encoding="UTF-8"?>
<blackBoxStochModel xmlns="http://www.openda.org".....>
<modelConfig>
     <file>./nemoModel.xml</file>
</modelConfig>
<vectorSpecification>
                <state>
               <vector id="ub"        />
          <vector id="vb"        />
          <vector id="tb"        />
          <vector id="sb"        />
          <vector id="sshb"      />
          <vector id="un"        />
          <vector id="vn"        />
          <vector id="tn"        />
          <vector id="sn"        />
          <vector id="sshn"      />
          </state>
     <predictor>
                         <subVector id ="pointssh0101.none"
sourceVectorId="sshn"><selection index1="0101" /> </subVector>
        </predictor>
</vectorSpecification>
```

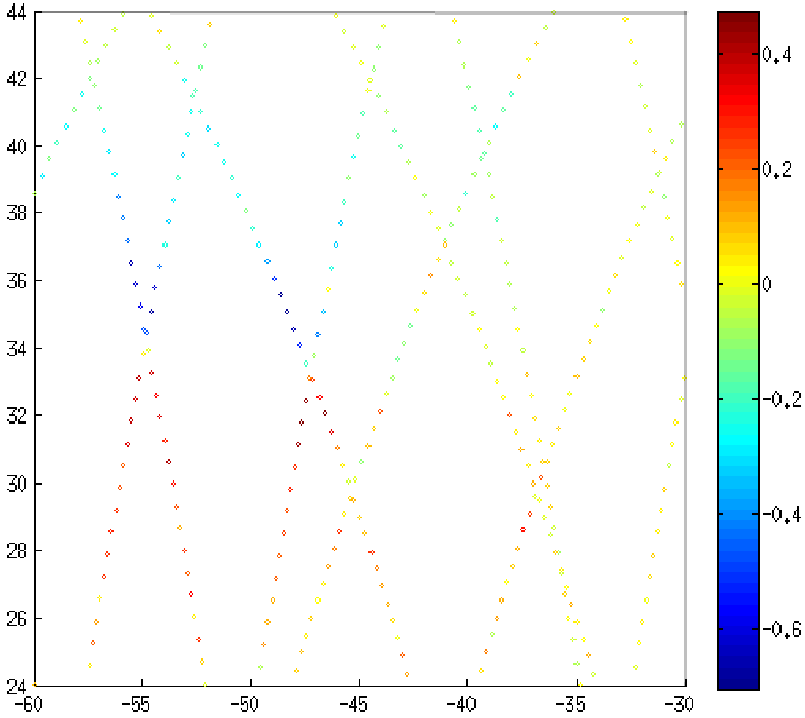
NEMO-OpenDA

```
<?xml version="1.0" encoding="UTF-8"?>
<noosObserver xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">

        <timeSeries id="pointssh0101" status="use" standardDeviation="0.06">
                pointssh0101.noos
        </timeSeries>
        <timeSeries id="pointssh0199" status="use" standardDeviation="0.06">
                pointssh0199.noos
        </timeSeries>
</noosObserver>
```

```
     #=======================================================
# Timeseries
#=======================================================
# Location    : pointssh0101
# Position    : (-2.045543,57.361939)
# Source      : observed
# Timezone    : GMT
#=======================================================
200801010000   -0.8300
200801010010   -0.8800
200801010020   -0.9100
```

Experiment Setup



Medium Case Benchmark

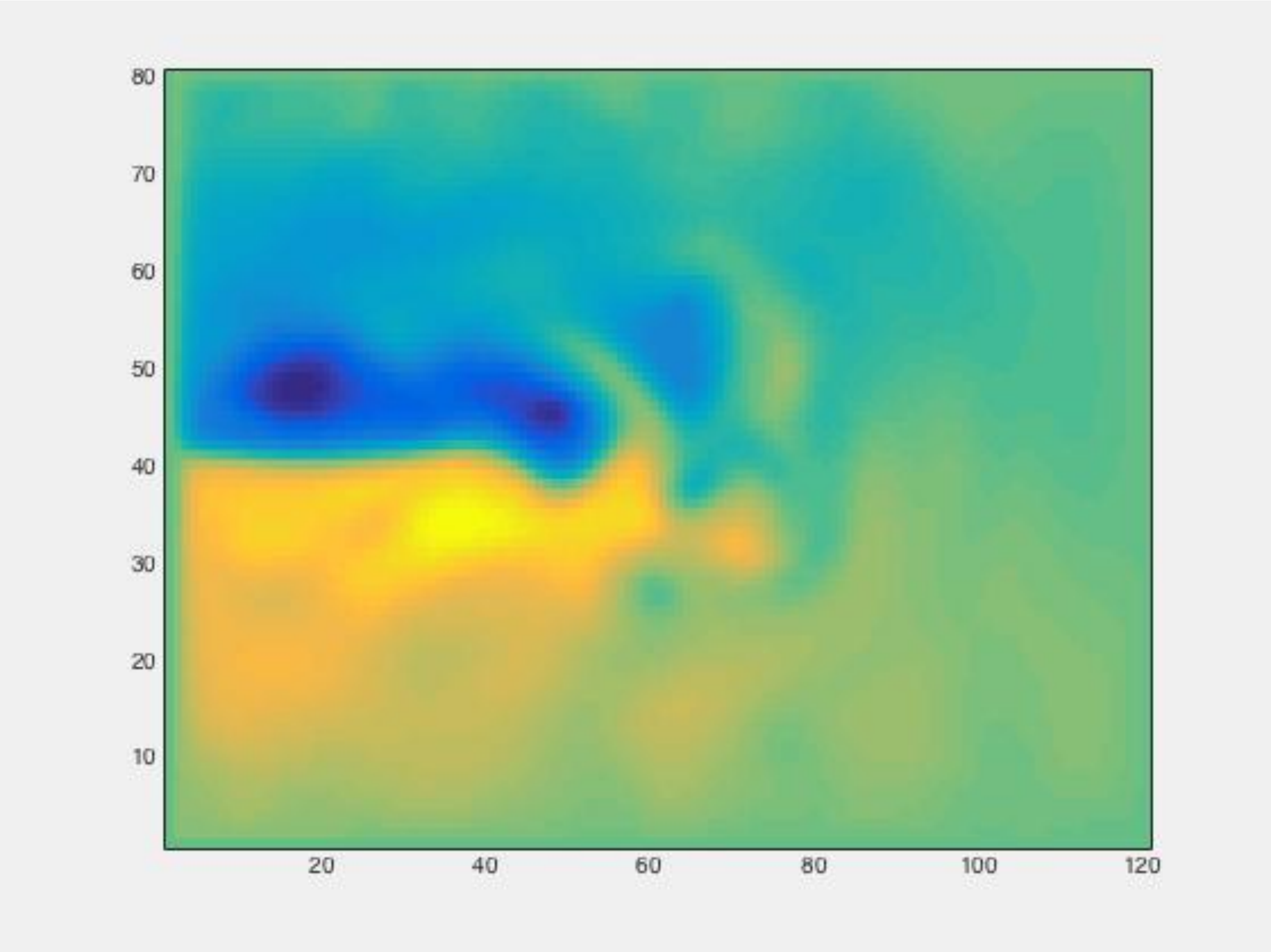
- Free run 40 years.

- State vector includes:
  (ub,vb,tb,sshb,un,vn,tn,sshn)

Ensemble Kalman filter

- 30/100 Ensemble members
- SSH observations. ENVISAT, Jason-1
- Analysis: 2 days

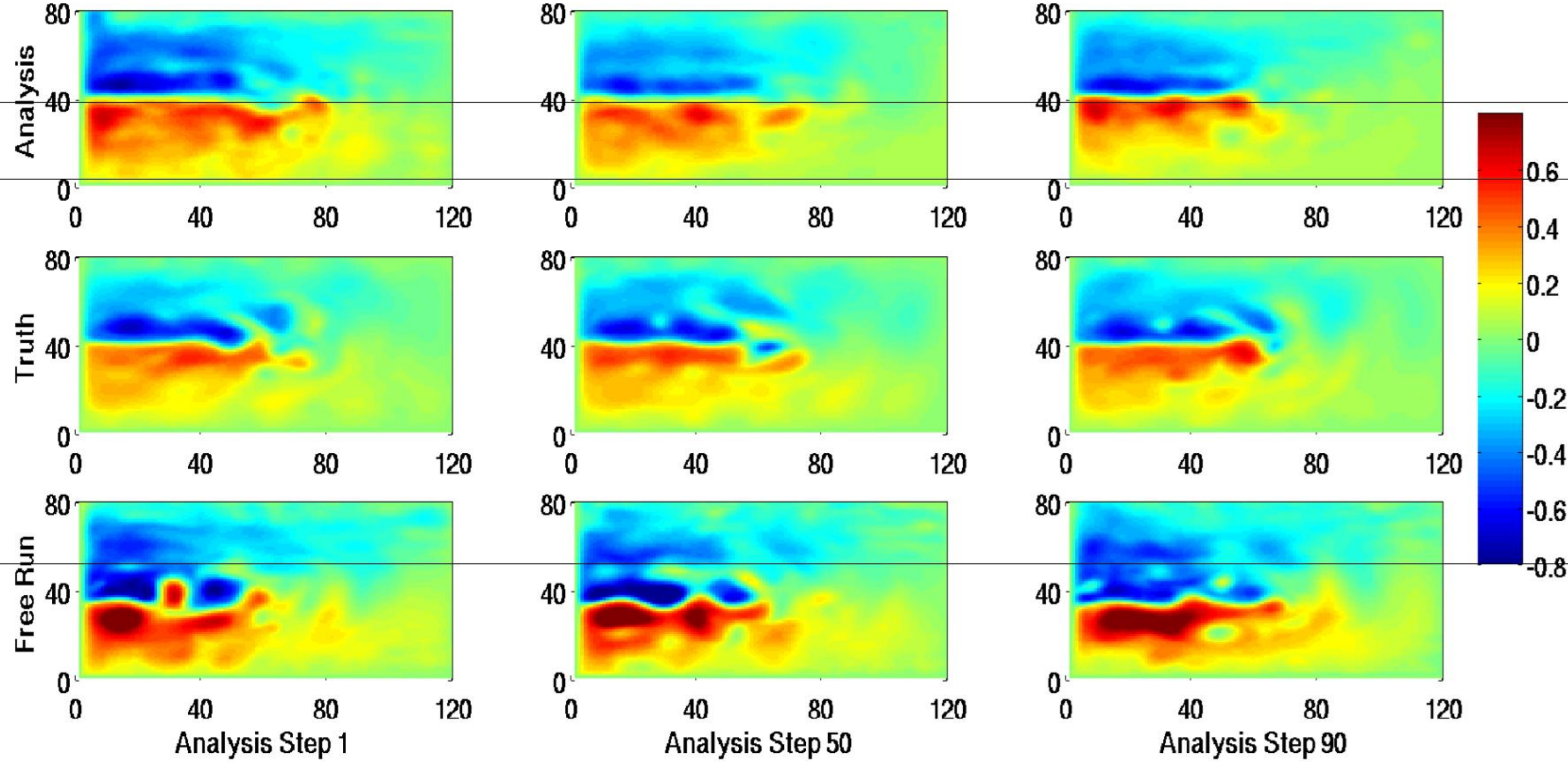
Experiment Setup

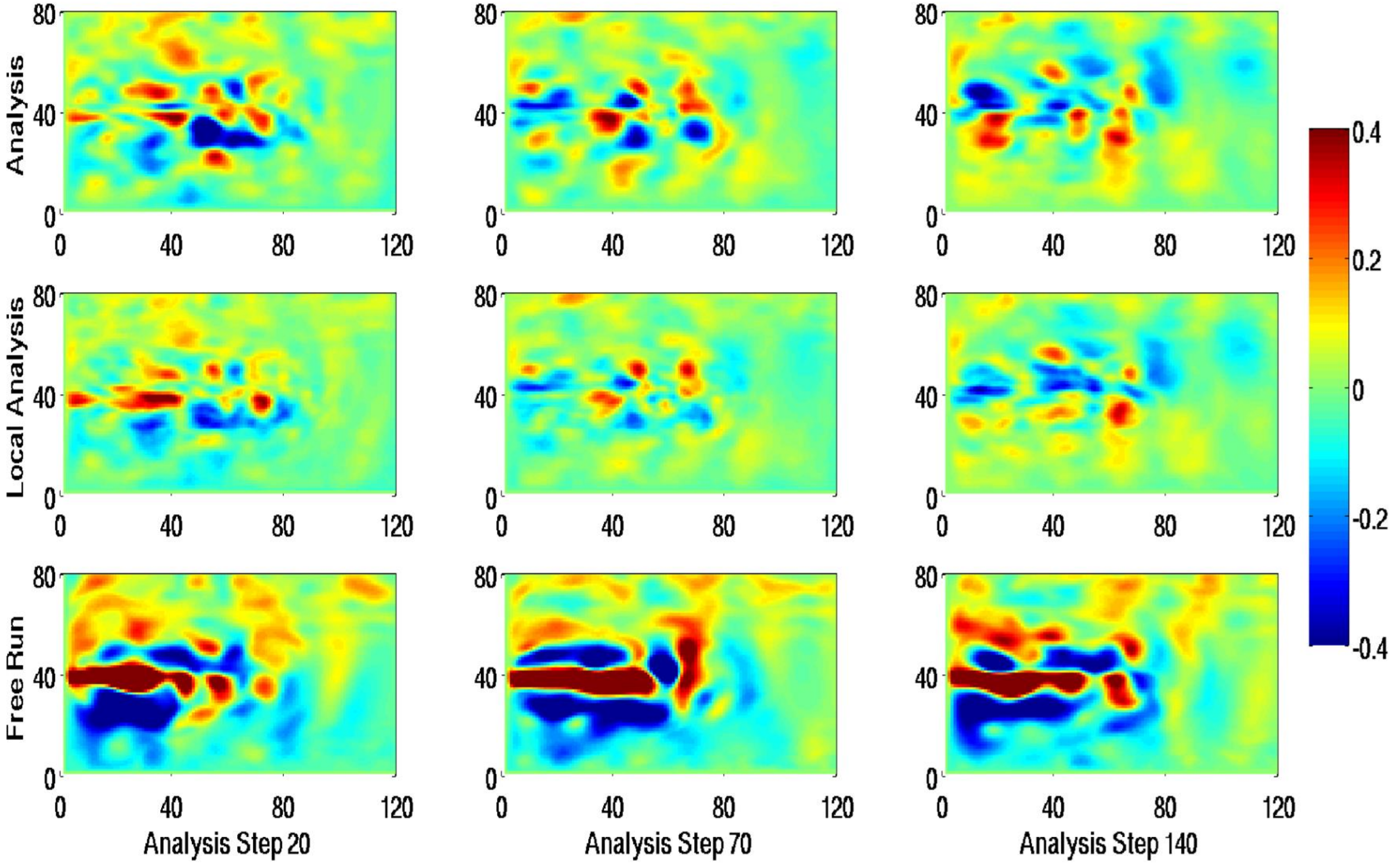


Results

Medium Case Benchmark

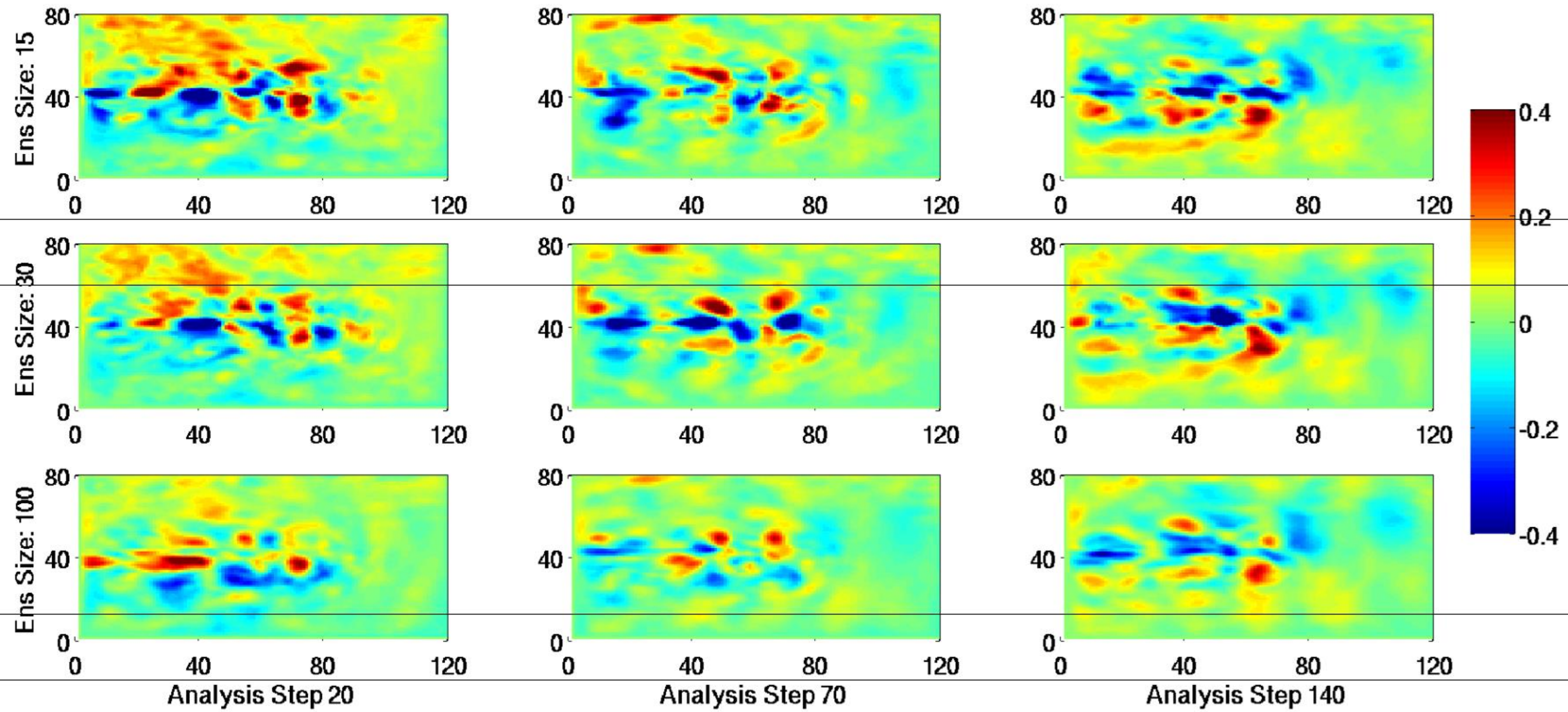
- Assimilating SSH every 2 days
- Assimilation period: 1 year




Analysis
Local Analysis
Free Run
80
40
0
0
40
80
120
Analysis Step 20
Analysis Step 70
Analysis Step 140
0.4
0.2
0
-0.2
-0.4

Medium Case Benchmark Results

- Localization improves results
- Limitation: No. of observation



Summary

- NEMO-OpenDA framework is established.
- Localization is not trivial to implement
- Auto localization seems promising
- Next steps:
    - More experimenting
    - Combine auto localization with normal localization

Acknowledgements:

This work is supported by SANGOMA a European FP7-SPACE-2011 project, Grant 283580.