Estimation of tidal boundary conditions and surface winds by assimilation of high-frequency radar surface currents in the German Bight

Alexander Barth, Aida Alvera-Azcárate, Joanna Staneva,
Alex Port, Klaus-Werner Gurgel,
Jean-Marie Beckers, Emil V. Stanev

GHER (University of Liege), University of Hamburg, GKSS, University of Oldenburg

JONSMOD, 2010

Outline

- Weakly constrained ensemble perturbations
- Example 1: Estimation of tidal boundary conditions using HF radar observations
- Example 2: Estimation of wind forcing using HF radar observations

Weakly constrained ensemble perturbations

- For ensemble schemes, unknown initial and boundary conditions, parameters, ... have to be perturbed within their range of uncertainty.
- By validation of the model with observations one can obtain an estimate of the magnitude of the perturbation.
- But which spatial structure?
- Method to create ensemble perturbation that satisfy *a priori* linear constraints
- Example of constraints:
  - geostrophic equilibrium
  - zero horizontal divergence of surface winds
  - stationary solution to the advection-diffusion equation
  - the linear shallow water equations
  - perturbations should be close to a subspace defined by *e.g.* empirical orthogonal functions (EOFs).
  - ...

Probability of a perturbation

- To describe our *a priori* knowledge of what a realistic perturbation is, we introduce a cost function J, similar to the cost function used in variational analysis techniques:

$$J(\mathbf{x}) = \text{“linear balance”}^2 + \text{“smooth”}^2 + \text{“limited amplitude”}^2$$

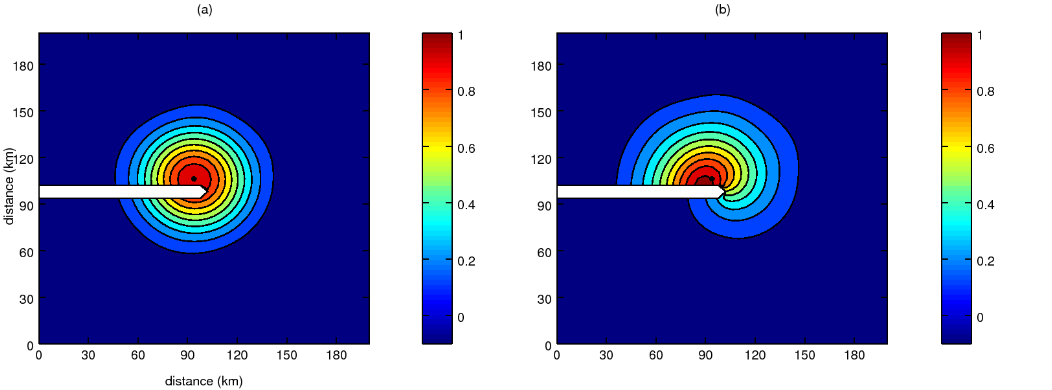
- The cost function can be used to define the probability of a perturbation $\mathbf{x}$ (*e.g.* **?**):

$$p(\mathbf{x}) = \alpha \exp\left(-J(\mathbf{x})\right) \tag{1}$$

- Perturbations are derived from the Hessian matrix of J.
- Article and source code (for MATLAB and GNU Octave) is available at http://modb.oce.ulg.ac.be/mediawiki/index.php/WCE
- “Dynamically constrained ensemble perturbations - application to tides on the West Florida Shelf”. Ocean Science, 5(3):259–270, 2009. http://www.ocean-sci.net/5/259/2009.

Impact of barriers

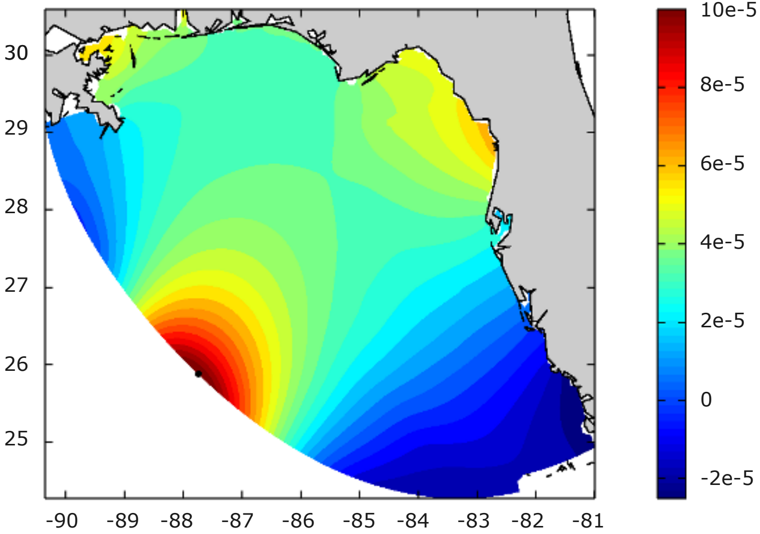
- The "smoothness" constraint is implemented through a diffusion operator (laplacian), it takes thus the land-sea mask into account



- Ensemble covariance using "classical" Fourier modes (a) and constrained perturbations based on the land-sea mask (b).

Harmonic shallow water equations

- For tidal models, perturbations should be approximately a harmonic solution to the shallow water equations

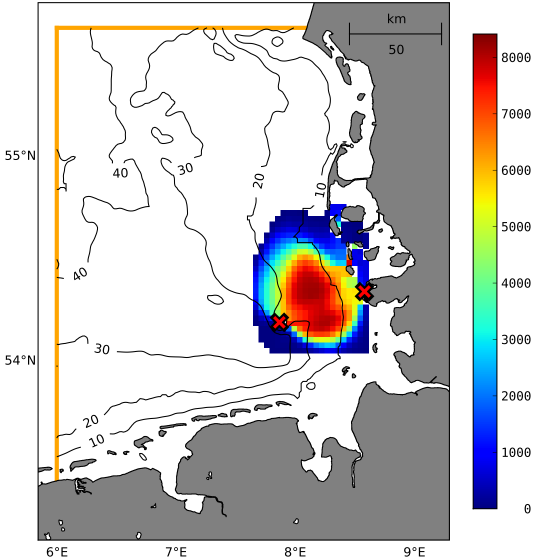


- Horizontal covariance of the constrained perturbations between the point near the open boundary marked by a black dot and all other grid points.

German Bight model

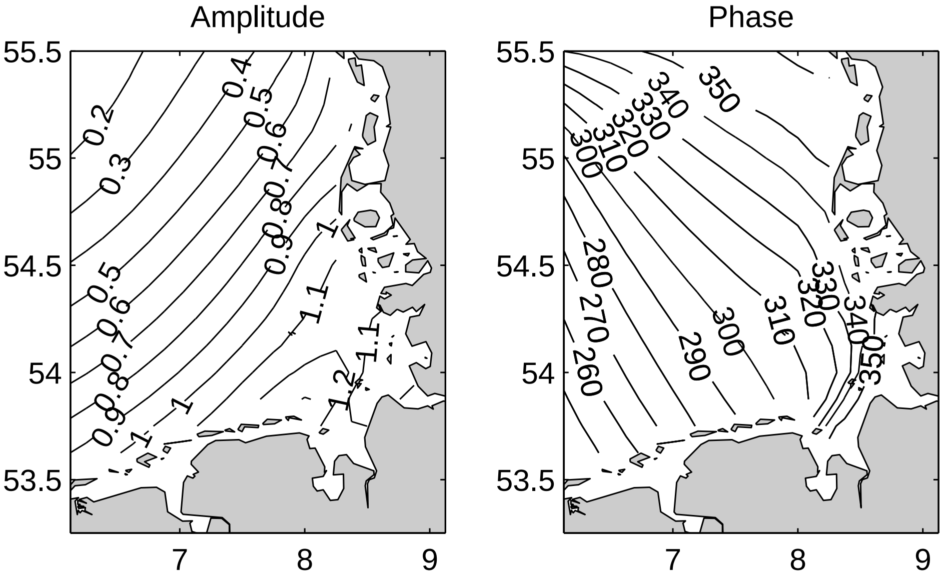
- General Estuarine Ocean Model (GETM **?**)
- 3-D primitive equations with a free-surface
- 21 σ levels, resolution of about 0.9 km.
- nested in a 5-km resolution North Sea-Baltic Sea model
- ETOPO-1 topography with observations from BSH
- Atmospheric fluxes are estimated by the bulk formulation using 6-hourly ECMWF re-analysis
- Implementation by GKSS (**?**).

HF radar observations



- Spatial coverage of the HF radar zonal and meridional surface velocity observations
- The number of samples available at each observation grid point is color–coded according to the color-bar.
- The crosses show the location of HF radar antennas.
- The operating frequency: 29.85 MHz (coupling to 5.02 m long ocean waves).
- HF Radar measurements from University of Hamburg (PRISMA project)

Empirical Ocean Tides (EOT08a)

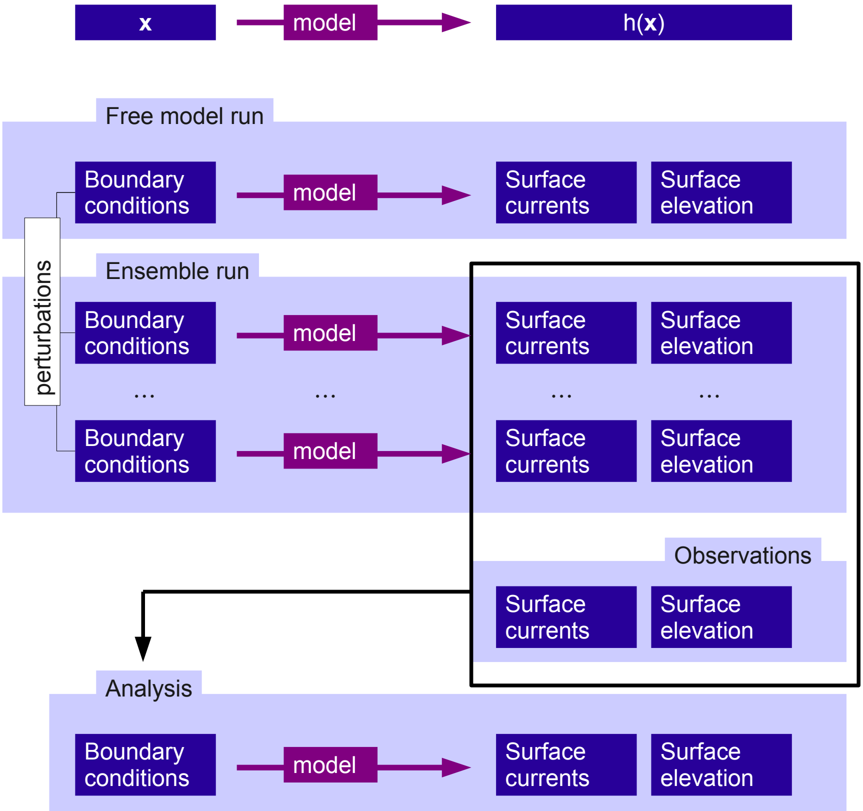


- M2 amplitude (in m) and phase (in degrees) of EOT08a for the German Bight based on altimetry.
- complex tidal parameters are assimilated

Smoother scheme

- M2 tidal boundary conditions are perturbed within the range of their uncertainty to create a ensemble with 51 members. Perturbations are constrained by the linear shallow water equations.
- The GETM model is run for 40 days with each of those perturbed boundary values.
- All HF radar observations at any time instance within the integration period and the EOT parameters are grouped in the observation vector (vector $\mathbf{y}^o$) with their corresponding error covariance (matrix $\mathbf{R}$) estimated by cross-validation.
- Observations are extracted from perturbed model solution (vector $h(\mathbf{x}^{(k)})$).
- Schematically, the non-linear operator $h(\cdot)$ performs the following operations:

$$h(\cdot) = \text{Interpolation to obs. location} \circ \text{Model integration with perturbed forcing} \tag{2}$$


x
model
h(x)
Free model run
Boundary conditions
model
Surface currents
Surface elevation
perturbations
Ensemble run
Boundary conditions
model
Surface currents
Surface elevation
...
...
...
...
Boundary conditions
model
Surface currents
Surface elevation
Observations
Surface currents
Surface elevation
Analysis
Boundary conditions
model
Surface currents
Surface elevation

Smoother scheme

- The optimal perturbation is given the Kalman analysis (using non-linear observation operators as in **?**):

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{A}\left(\mathbf{B}+\mathbf{R}\right)^{-1}\left(\mathbf{y}^o - h(\mathbf{x}^b)\right) \quad (3)$$

- where the matrices $\mathbf{A}$ and $\mathbf{B}$ are covariances estimated from the ensemble.

$$\mathbf{A} = \mathrm{cov}(\mathbf{x}^b, h(\mathbf{x}^b)) = \left\langle \left(\mathbf{x} - \langle \mathbf{x}\rangle\right)\left(h(\mathbf{x}) - \langle h(\mathbf{x})\rangle\right)^T \right\rangle \quad (4)$$

$$\mathbf{B} = \mathrm{cov}(h(\mathbf{x}^b), h(\mathbf{x}^b)) = \left\langle \left(h(\mathbf{x}) - \langle h(\mathbf{x})\rangle\right)\left(h(\mathbf{x}) - \langle h(\mathbf{x})\rangle\right)^T \right\rangle \quad (5)$$

where $\langle\cdot\rangle$ is the ensemble average.

- But covariance matrices do not need to be formed explicitly. Analysis is performed in the subspace defined by the ensemble members.

Smoother scheme

- For a linear model and an infinite large ensemble, equation (3) minimizes,

$$J(x) = (\mathbf{x} - \mathbf{x}^b)^T {\mathbf{P}^b}^{-1} (\mathbf{x} - \mathbf{x}^b) + (\mathbf{y}^o - h(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y}^o - h(\mathbf{x})) \quad (6)$$

or

$$J(x) = (\mathbf{x} - \mathbf{x}^b)^T {\mathbf{P}^b}^{-1} (\mathbf{x} - \mathbf{x}^b) + \sum_n (\mathbf{y}_n^o - (h(\mathbf{x})_n))^T {\mathbf{R}_n}^{-1} (\mathbf{y}_n^o - (h(\mathbf{x})_n)) \quad (7)$$

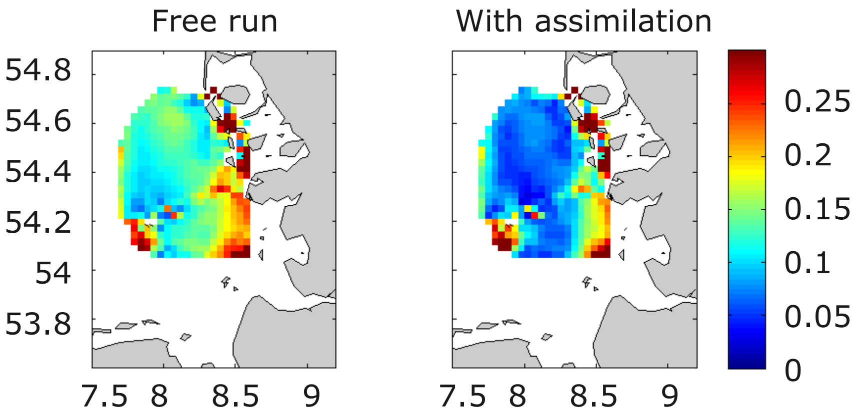
where n references to the indexed quantifies at time n. This is the cost function from which 4D-Var and Kalman Smoother can be derived.

- Approach is closely related to Ensemble Smoother (**?**), 4D-EnKF (**?**) and AEnKF (**?**) where model trajectories instead of model states are optimized and to the Green's method with stochastic "search directions"

- The model is rerun with the optimized boundary values for 60 days.

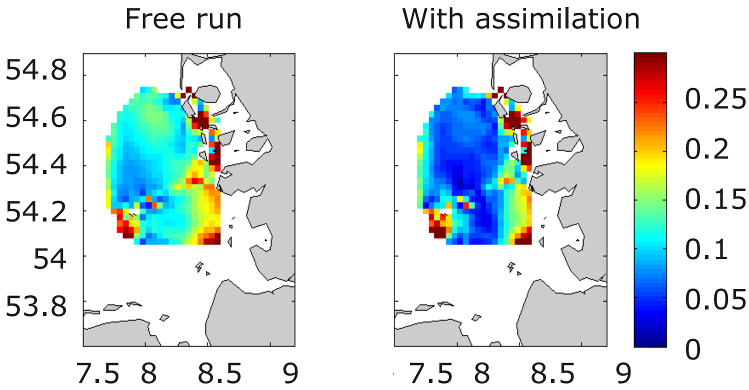
RMS difference

$$
\begin{aligned}
\mathrm{RMS}^2 &= \lim_{T\to\infty} \frac{1}{T} \int_0^T (A\cos(\omega t - \phi) - A' \cos(\omega t - \phi'))^2 dt \qquad (8) \\
&= \frac{A^2 + A'^2}{2} - AA' \cos(\phi - \phi') \qquad (9)
\end{aligned}
$$



RMS difference between surface current observations due to the M2 tides and the corresponding model results without (left panel) and with assimilation (right panel).

Comparison with un-assimilated observations (M2)



- RMS difference between surface current observations (not used in the assimilation) due to the M2 tides and the corresponding model results without (left panel) and with assimilation (right panel).
- Analysis RMS compared to unassimilated data is only 0.002 m/s larger than compared to assimilated data

Tide gage observations

	Helgoland			Cuxhaven		
	amplitude	phase	RMS	amplitude	phase	RMS
Observations	1.13	304		1.36	334	
Free	0.81	318	0.28	0.95	15	0.63
Assimilation	0.97	302	0.12	1.08	2	0.46

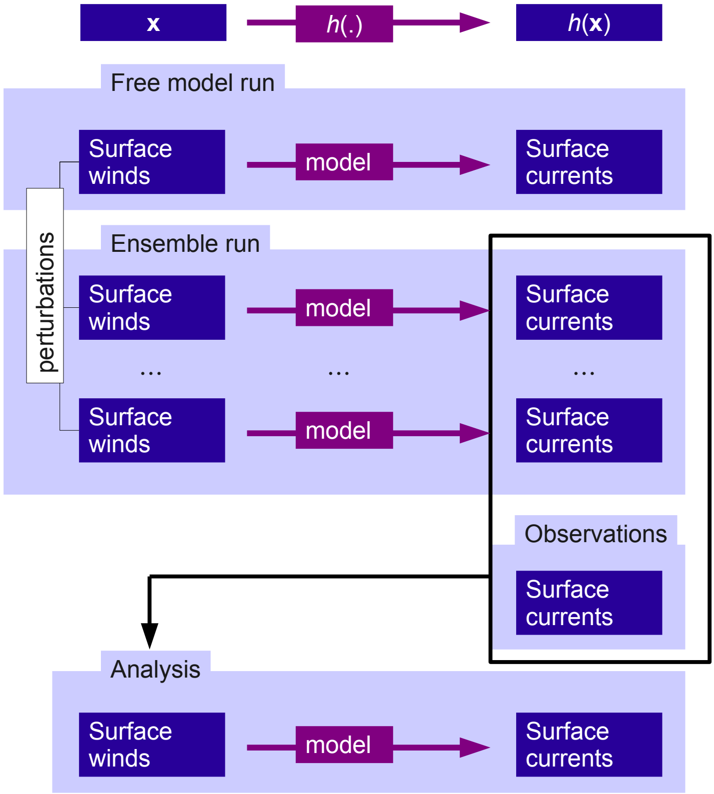
Table 1: Comparison with tide gage observations. Amplitude is in m and phase in degrees.

- Tide gage observations from different time period $\rightarrow$ only comparison of tidal parameters
- Helgoland within the area covered by radar, but not Cuxhaven
- The assimilation reduces the RMS error by a factor of 2 for Helgoland and by a factor of 1.4 for Cuxhaven.
- Ocean Science, 6, 161–178, 2010 http://www.ocean-sci.net/6/161/2010/os-6-161-2010.pdf.

Wind estimation from HF radar observations

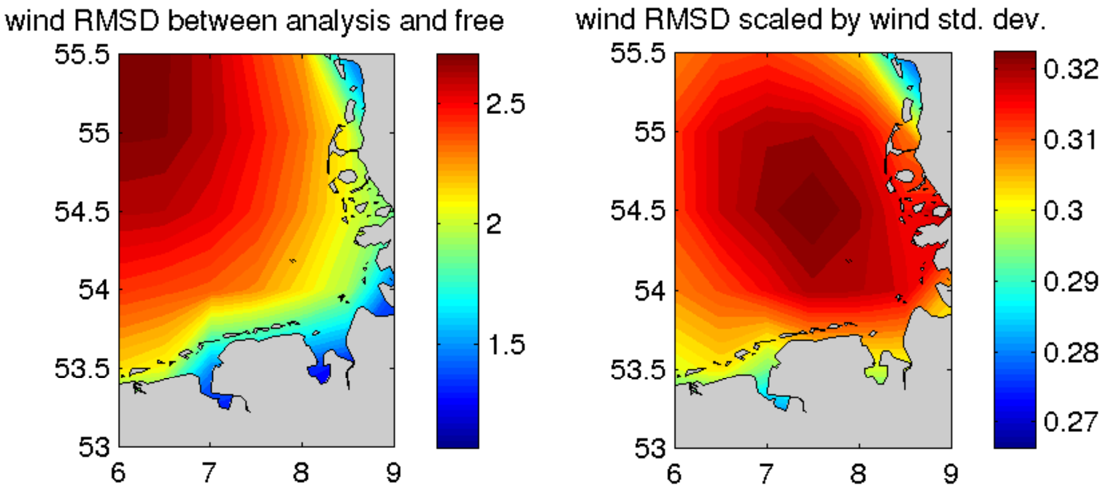
- Ensemble of 100 wind forcings are created (by using a Fourier decomposition)
- estimation vector $\mathbf{x}$: u- and v- component of wind forcing
- observations: $\mathbf{y}^o$: surface currents
- "observation operator" $h(\cdot)$:

$$h(\cdot) = \text{Interpolation to obs. location} \circ \text{Model integration with perturbed wind} \tag{10}$$



x
h(.)
h(x)
Free model run
Surface winds
model
Surface currents
perturbations
Ensemble run
Surface winds
model
Surface currents
...
...
...
Surface winds
model
Surface currents
Observations
Surface currents
Analysis
Surface winds
model
Surface currents

Time-averaged wind correction statistics



- RMS difference between analyzed winds and ECMWF winds (averaged over time)
- RMS difference scaled by wind standard deviation

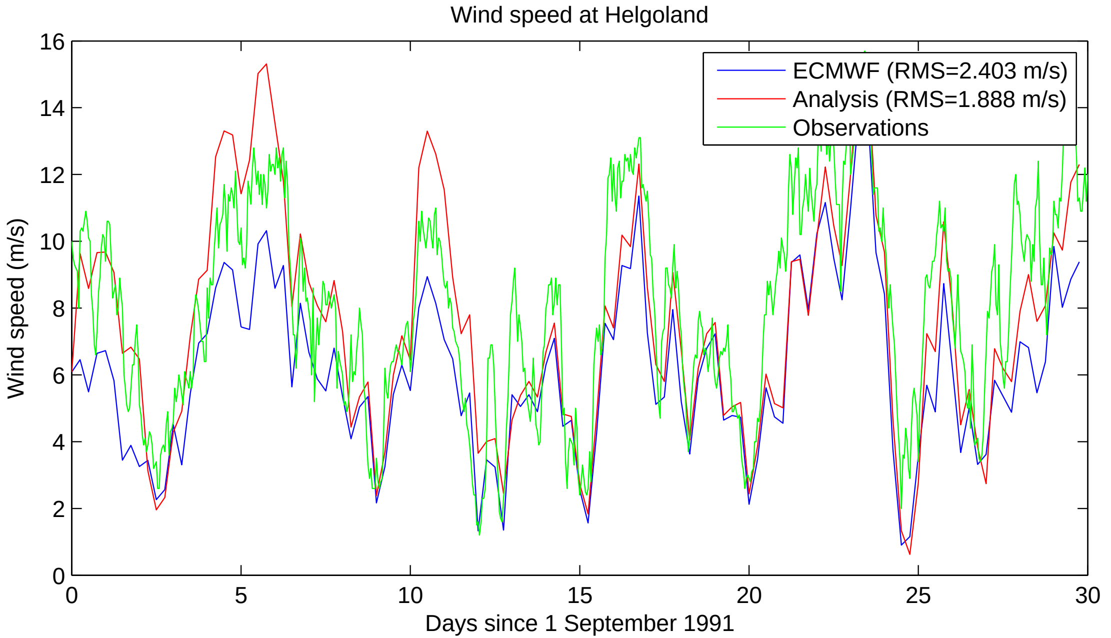


Figure 1: Measured wind speed, wind speed from ECMWF and analyzed wind speed at Helgoland. Units are m/s.

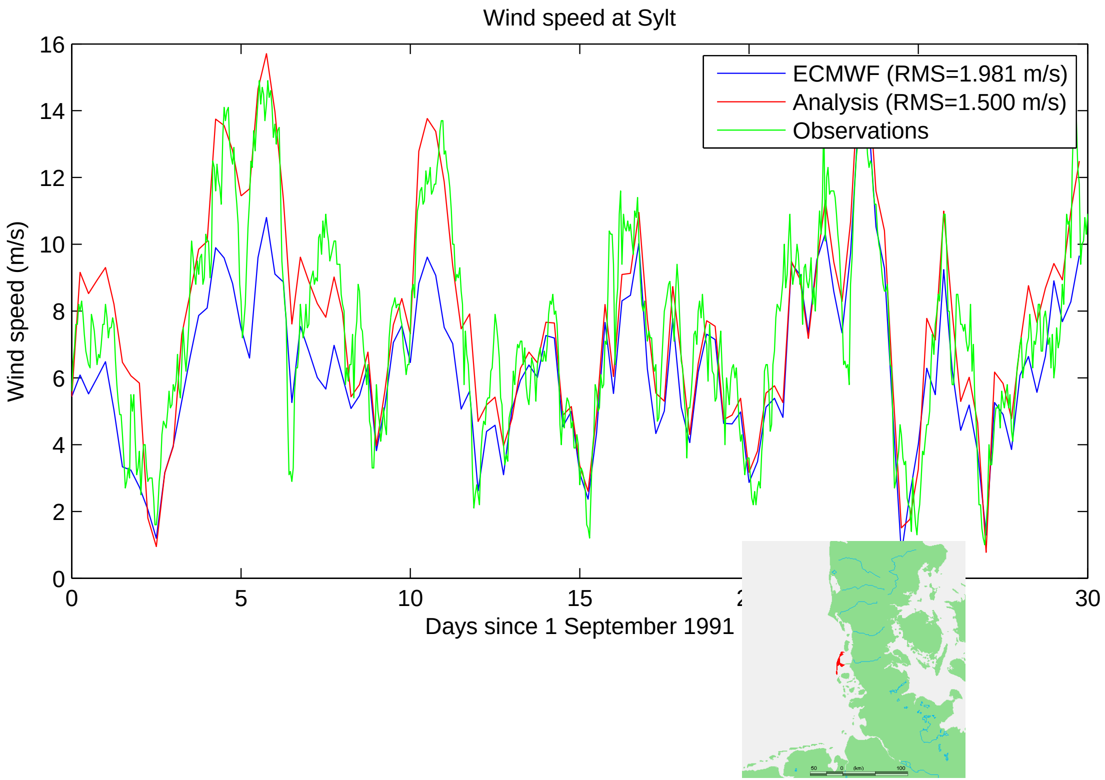


Figure 2: Measured wind speed, wind speed from ECMWF and analyzed wind speed at Sylt. Units are m/s.

Summary

- Ensemble assimilation methods require realistic perturbation schemes (error co-variances)
- Use of dynamical relationships (similar to Variational analysis)
- Optimizing tidal boundary conditions and wind forcing with a smoother scheme
- HF radar observation is a very valuable data set for constraining regional and coastal models

Skill score

- How strong can we force the model towards the observations (S_{HF})?
- Taking into account that:
  - redundancy between the individual time instances is high (data every 30 min)
  - the difference between observed currents and model currents is not only due to errors in the wind field
  - If S_{HF} would be too small, errors in *e.g.* the density structure could be compensated in a unrealistic way by modifying the wind fields, only to match observations closely.

		Helgoland		Sylt	
	S_{HF}	RMS	skill score	RMS	skill score
Free	–	2.40	0.00	1.98	0.00
Analysis	0.5	2.14	0.21	1.96	0.03
	1.0	2.02	0.29	1.82	0.15
	1.5	1.93	0.35	1.64	0.31
	2.0	1.89	0.38	1.50	0.43
	2.5	1.88	0.39	1.43	0.48
	5.0	2.05	0.27	1.54	0.39

- Comparison with *in situ* wind measurements. RMS and S_{HF} are expressed in m/s.

Comparison with satellite SST

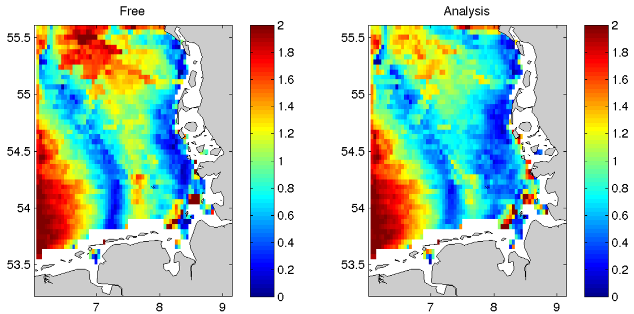


Figure 3: RMS difference between AVHRR SST and model SST without assimilation (left panel) and with assimilation (right panel)

	S_{HF}	RMS	skill score
Free	–	1.21	0.00
Analysis	0.5	1.09	0.19
	1.0	1.09	0.19
	1.5	1.10	0.18
	2.0	1.11	0.16
	2.5	1.12	0.14
	5.0	1.16	0.08

RMS is expressed in °C and S_{HF} in m/s.

Perturbations scheme

The cost function is a quadratic function in $\mathbf{x}$ and can thus be written as:

$$\begin{aligned} 2J(\mathbf{x}) &= \mathbf{x}^T(\mathbf{M}^T\mathbf{W}_M\mathbf{M} + \mathbf{D}^T\mathbf{W}_D\mathbf{D} + \mathbf{W}_E)\mathbf{x} && (11) \\ &= \mathbf{x}^T\mathbf{B}^{-1}\mathbf{x} && (12) \end{aligned}$$

where the matrix $\mathbf{B}$ (covariance matrix, not formed explicitly) is defined as:

$$\mathbf{B} = (\mathbf{M}^T\mathbf{W}_M\mathbf{M} + \mathbf{D}^T\mathbf{W}_D\mathbf{D} + \mathbf{W}_E)^{-1} \tag{13}$$

To generate an ensemble of perturbations that follows the previous pdf, the matrix $\mathbf{B}$ is decomposed in eigenvectors (rows of $\mathbf{U}$) and eigenvalues (diagonal elements of $\mathbf{\Lambda}$) :

$$\mathbf{B} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T \tag{14}$$

The smaller an eigenvalue is, the stronger the corresponding eigenvector violates the dynamical and smoothness constraint.

An ensemble of vectors $\mathbf{z}^{(k)}$ where the subscript k is the ensemble member, is created following a normal distribution.

$$\mathbf{z} \sim N(0, \mathbf{I}_n) \tag{15}$$

An ensemble of perturbations $\mathbf{x}^{(k)}$ following (1) can be obtained by:

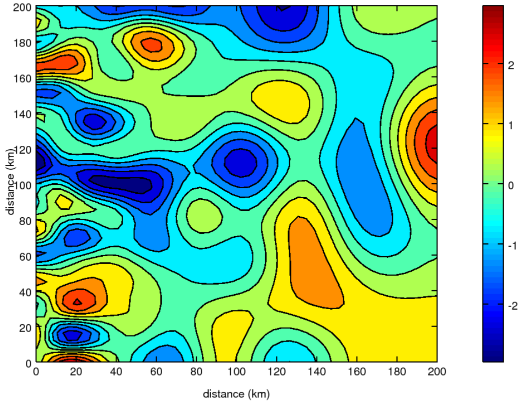
$$\mathbf{x}^{(k)} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{z}^{(k)} \tag{16}$$

Alternatively, one can use the 2nd order exact re-sampling method (SEIK):

$$\mathbf{x}^{(k)} = \mathbf{U}\mathbf{\Lambda}^{1/2}\mathbf{H}_w(\Omega)_k \tag{17}$$

where columns of $\mathbf{H}_w$ are all perpendicular to the vector $\mathbf{1}_{N\times 1}$ and $(\Omega)_k$ is the k-column of a random orthogonal matrix Ω.

- Also perturbations with a spatially varying correlation length can be created.
- Scale of mesoscale variability $\longrightarrow$ internal radius of deformation which varies in space:



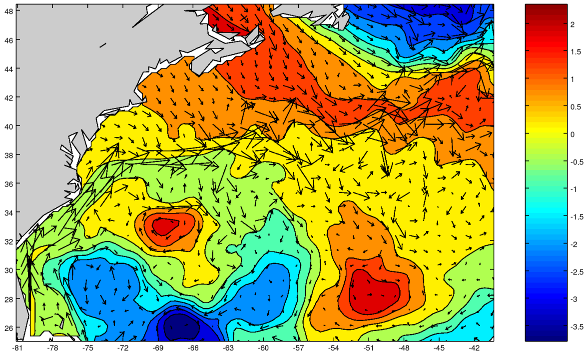
- Illustration of a random field with a variable correlation length.

Examples for linear constraints

Advection constraint

- For large-scale models, perturbations should be approximately stationary solutions to the advection equation

$$\mathbf{v} \cdot \nabla \phi = 0 \tag{18}$$



- Example of ensemble perturbations using the advection constraint

Application to HF Radar assimilation in the German Bight (tidal BC)

- Only M2 tidal boundary conditions are perturbed:

$$\zeta^{(k)} = \zeta^{(b)} + \Re\left(\zeta'(x,y)\,\exp(i\omega t)\right) \tag{19}$$

where ω is the M2 angular frequency and $\zeta'(x,y)$ is a random field satisfying approximately the harmonic shallow water equations:

$$i\omega\zeta' + \frac{\partial(hu')}{\partial x} + \frac{\partial(hv')}{\partial y} = 0 \tag{20}$$

$$i\omega u' - fv' + g\frac{\partial\zeta'}{\partial x} = 0 \tag{21}$$

$$i\omega v' + fu' + g\frac{\partial\zeta'}{\partial y} = 0 \tag{22}$$

- The 50 eigenvector with the largest eigenvalues of the matrix $\mathbf{B}$ from (14) are calculated (providing the spatial structure of the perturbation).

- From those 50 eigenvector/eigenvalues an ensemble of 51 members is created with zero mean (2nd order exact re-sampling).
- The GETM model is run for 40 days with each of those perturbed boundary values.
- Observations are assimilated with an expected RMS error of 0.3 m/s (including representativity error and error that cannot be corrected modifying only the boundary conditions) providing an optimal increment of the boundary values.
- The model is rerun with the optimized boundary values for 60 days.