D9.6
OSSE FINAL REPORT

Table of Contents

Document description

Introduction

In WP9 OSSE experiments are applied in the Adriatic Sea, the North Sea, The Baltic Sea and the Bay of Biscay and the Irish Sea. The methodologies range from estimating the observation impact in twin experiments to methods based on the evaluation of properties of the matrix of analysis errors.

OSSE experiments are very important for planning the deployment of future observational platforms. They can be performed either by estimating the impact of historical observational experiments that lasted for a limited time, or by producing synthetic observations that can later be assimilated. Some methods are based on the evaluation of the change in the properties of analysis error covariances. The methods applied in the WP9 use different methodologies in different experiments covering different coastal areas. This report documents OSSE are feasible to optimize the design of different observation systems. Even though the methods to evaluate the effects of desgined observation systems are different, the improvements are certain nevertheless to different extents.

1. OSSE in the Baltic Sea (Z. Wan, DMI)

1.1 Geographical setup

The Baltic Sea is a semi-closet coast sea. Its average depth is about 60 m, while typical basin depths vary 100 to 250 m. It connects to the North Sea through Danish Straits – the Great Belt and the Little Belt. Its surface salinity ranges from 8-9 PSU in the south to 2-3 PSU in the northern Bothenian Bay. Salty and oxic water can intrude into the bottom of Gotland Deep in the central basin due to weather events. In the surface, the Baltic Sea recieves over 70 rivers' discharges. The semi-resdence time in the Baltic Sea is assessed about 30 years. The coastal regions are highly populated. The Baltic Sea was believed one of the best observed seas in the world, as massive moorings and research vessels and also radar and satellite remote sensors are providing information.

1.2 Model description

In the Baltic Sea, DMI is running a two-way nested, free surface, hydrostatic three-dimensional (3D) circulation model called HIROBM-BOOS (HBM). The model code forms the basis of a common Baltic Sea model for providing GMES Marine Core Service since 2009. The finite difference method is adopted for its spatial discretization in which a staggered Arakawa C grid is applied on a horizontally spherical and vertically z-coordinate.

The model has a horizontal resolution of about 6 nautical miles (nm) and 50 vertical layers. The top layer thickness is selected at 8 m in order to avoid tidal drying of the first layer in the English Strait. The rest of the layers in the upper 80 m have 2 m vertical resolution. In the Danish Strait, the horizontal resolution is increased to 1 nm to better resolve the complex bathymetry. A detailed description of the model can be found in Berg and Poulsen (2011).

The meteorological forcing is based on a reanalysis using the regional climate model HIRHAM through a dynamic downscaling (including a daily re-initialization) from ERA-Interim Global reanalysis. HIRHAM is a regional atmospheric climate model (RCM) based on a subset of the HIRLAM and ECHAM models, combining the dynamics of the former model with the physical parameterization schemes of the latter. The original HIRHAM model was a collaboration between DMI, the Royal Netherlands Meteorological Institute (KNMI) and

MPI. A detailed description of HIRHAM Version 5 can be found in Christensen et al. (2006).

1.3 Data assimilation system description

A 3DVAR method has been applied to assimilate the satellite SST, in situ temperature and salinity profiles into a coupled physical-biogeochemical model in the Baltic Sea. In general, the basic scheme of 3DVAR is to find the optimal solution of the model state x which minimizes the following cost function:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x_b})^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \frac{1}{2}(H(\mathbf{x}) - \mathbf{y}_o)^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}_o) \qquad (1)$$

x is the model state to be estimated. It usually refers to analysis state vector. $\mathbf{x}_b$ is the background state vector, $\mathbf{y}_o$ is the observation state vector. H is the non-linear observational operator with which the analysis equivalent of observation $\mathbf{y} = H(\mathbf{x})$ can be obtained to compare with the observation measurements. The superscript T denotes matrix transpose. In the cost function, the misfit between analysis and background is weighted by the background error covariance B, and the misfit between analysis and observation is weighted by the observational error covariance R. Usually the optimal solution is found by minimizing the cost function $J(\mathbf{x})$ with respect to $\mathbf{x}$, in which its gradient is also needed for determining the search direction and iteration steps in the minimizing algorithm:

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) + \nabla_{\mathbf{x}} H(\mathbf{x})^T \mathbf{R}^{-1}(H(\mathbf{x}) - \mathbf{y}_o) \qquad (2)$$

Following an incremental method (Courtie, etc. 1994), Equation (1) is linearized around the background state into the following form:

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^T \mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d})^T \mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}) \qquad (3)$$

where $\mathbf{d} = \mathbf{y}_o - H(\mathbf{x}_b)$ is the innovation vector, $\mathbf{H}$ is the linearized observation operator evaluated at $\mathbf{x} = \mathbf{x}_b$ and $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$ is the analysis incremental vector. In this way, the original problem converts into finding an incremental analysis $\delta\mathbf{x}$. Equation (2) becomes :

$$\nabla J(\delta\mathbf{x}) = \mathbf{B}^{-1}\delta\mathbf{x} + \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}) \qquad (4)$$

In our current scheme, the state vector contains only temperature and salinity model state variables:

$$\mathbf{x} = \begin{bmatrix} T & S \end{bmatrix}^T \qquad (5)$$

1.4 Sampling Strategy

With Observing System Simulation Experiments (OSSEs), we are planning to examine the impacts on the forecasts of using observations from gliders operating along two selected routes (Route 1 and Route 2) in the Baltic Sea (Fig. 1.1). The glider routes are designed to test along which route the glider observation system can improve prediction products more and which area each glider can improve. The gliders are navigated to

move back and forth along the designed route 1 km/h and to release data once per day. The gliders are assumed to take a snapshot of profiles for temperature and salinity with vertical resolution of 2m.
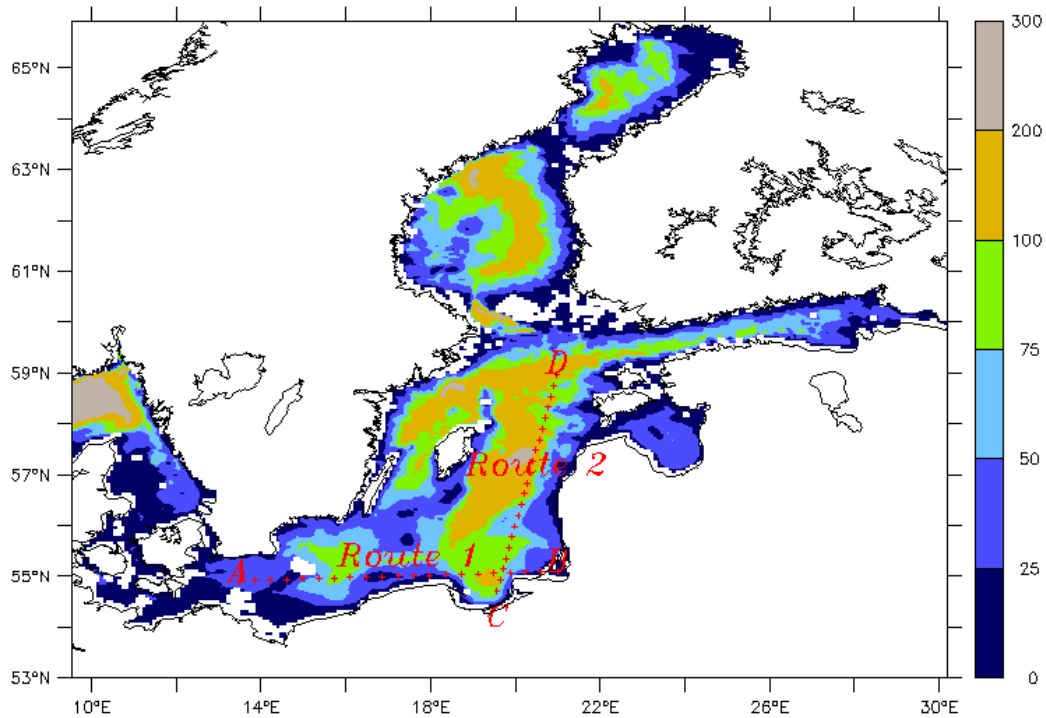
Fig. 1.1 Model domain and glider routes: Route 1 – AB; Route 2 - CD

OSSE approach has been widely used as proof-of-concept tool for assessing the impacts of proposed observing systems. It relies on the assumption that there exists one model configuration which provides a complete knowledge of the true state of the ocean. The OSSE experiments in the Baltic Sea are described as follows. The 'true' ocean state is generated by a model run starting on January 1, 2009, while the model is assimilating T/S profiles from research vessels and SST from satellite remote sensing. We consider the model boundary and flux conditions as error-free. In addition to the 'true' ocean state, a reference experiment is conducted by starting the model from a false initial condition (taken from another year). The temperature and salinity profiles will be assimilated with the 3DVAR scheme to reconstruct the 'true' state. Four scenarios are selected:

- Exp. 0: deviating from the 'true' ocean state with perturbation of a year-to-year variation of initial fields;
- Exp. 1: identical to Exp. 0, but assimilating data from one glider operating along Route 1;
- Exp. 2: identical to Exp. 0, but assimilating data from one glider operating along Route 2;
- Exp. 3: identical to Exp. 0, but assimilating data from two gliders operating along Route 1 and Route 2 respectively.

All the four experiments run from March 1, 2009 to April 29, 2009.

1.5 Results

a) Temporal evolution of mean deviations of temperature and salinity from scenario state to 'true' state

Mean deviations of temperature and salinity integrated over entire model domain are compared among the four experiments in Fig. 1.2. As we can see, mean deviations of experiments with data assimilation (DA) are smaller than that without DA, and the mean deviation of experiment assimilating data from two gliders is smaller than that of experiment assimilating data from individual glider.

Fig. 1.2 Comparison of mean deviations among four experiments (Exp. 0 – black, Exp. 1 – red, Exp. 2 – green, Exp. 3 – blue). Panel a – temperature, panel b – salinity.

b) Profiles of mean deviations of temperature and salinity from scenario state to 'true' state

Mean deviations of temperature and salinity integrated over model planes are compared among four experiments in Fig. 1.3. As we can see, mean deviations of experiments with DA are smaller than that without DA, and the mean deviation of experiment assimilating data from two gliders is smaller than that of experiment assimilating data from individual glider.
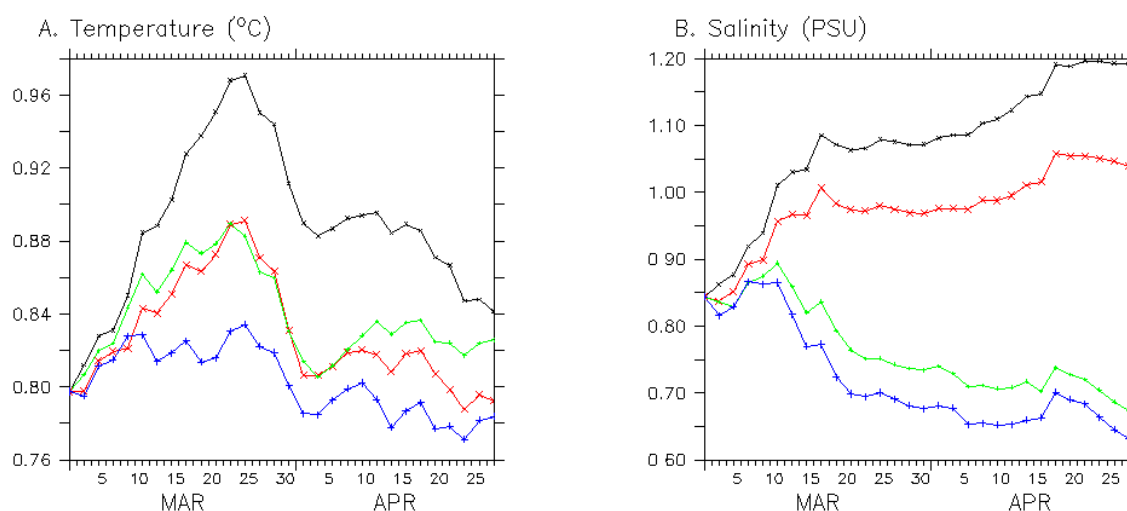
A. Temperature (°C)   B. Salinity (PSU)
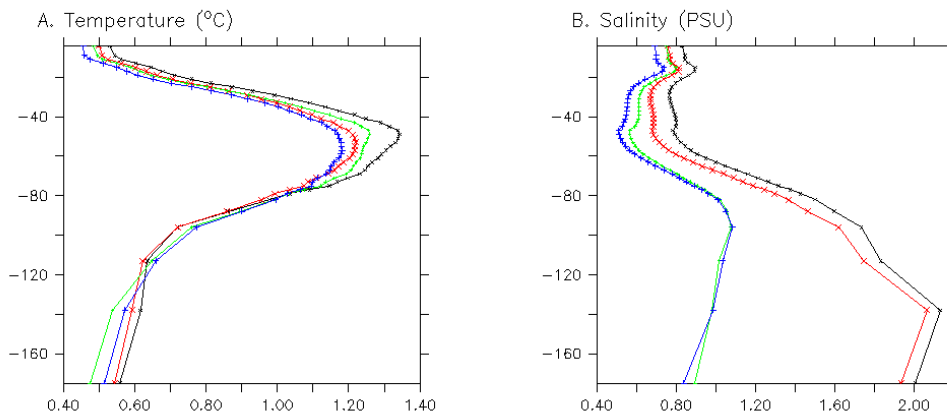
Fig. 1.3 Comparison of profiles of mean deviations among four experiments (Exp. 0 – black, Exp. 1 – red, Exp. 2 – green, Exp. 3 – blue). Panel a – temperature, panel b – salinity.

c) Regional distribution of mean deviations and improvements

Regional distribution of mean deviations from Exp. 0 to 'true' state is depicted in Fig. 1.4a for temperature. Percentage improvements of Exp. 1-3 relative to Exp. 0 are depicted in Fig. 4b-d. It is clear that improvements appear mainly nearby the route, but also in areas circulation can immediately impact. Improvement in Exp. 2 is larger than in Exp. 1. Improvement in Exp. 3 looks like a sum of those in Exp. 1 and Exp. 2. The result for salinity (Fig. 1.5) is rather similar to temperature.

A. OSSE Ref, Absoulte Bias (°C)   C. OSSE R2, Relative Improvement

B. OSSE R1, Relative Improvement   D. OSSE R1+R2, Relative Improvement

Fig. 1.4 Mean deviations of temperature from the reference state Exp. 0 to 'true' state (a) and percentage improvements of Exp. 1 (b), 2 (c) and 3 (d) relative to Exp. 0.

Fig. 1.5 Mean deviations of salinity from the reference state Exp. 0 to 'true' state (a) and percentage improvements of Exp. 1 (b), 2 (c) and 3 (d) relative to Exp. 0.

d) Results in statistics

Statistics based on daily means of temperature at all model grids show that mean deviation from the reference state to the 'true' state is 2.8% due to the introduced perturbation in initial fields. The glider observation system can reduce mean deviations for the entire Baltic Sea up to 6.6%, 2.3%, 13% in c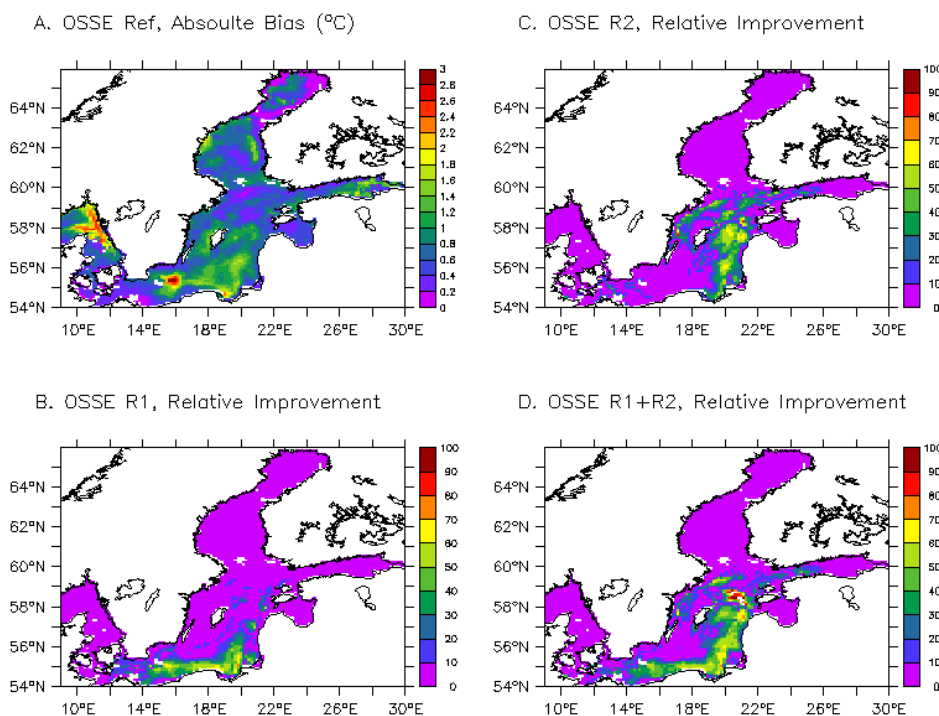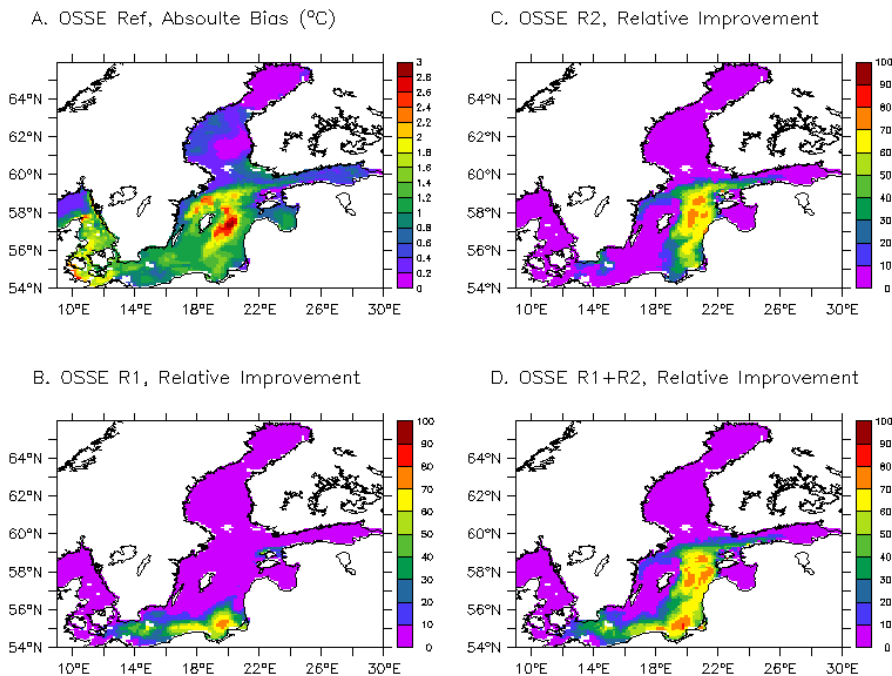ircumstance with one glider operating along Route 1, Route 2 and two gliders along Route 1 and Route 2 respectively, comparing to the reference run (without DA). For salinity, mean deviation from the reference state to the 'true' state is 1.2%. The glider observation system can reduce mean deviations for the entire Baltic Sea up to 3.8%, 27%, 30%.

e) Discussion and lessons learnt

It is found that assimilating data from gliders operating along both Route 1 and Route 2 can reduce the model 24h forecast error by 13% for temperature and 30% for the salinity. Assimilating data from glider operating along Route 2 can improve more than that along Route 1. The combination of two gliders can improve more than individuals. This investigation documents that OSSE approach can facilitate the design of glider observation system in order to optimize the system function. This practice to employing OSSE is carried out through introducing perturbation to initial fields. As we know, data assimilation can reduce prediction bias generated from initial fields, forcing conditions and parameter values. Model bias from different resources may differ in propagation. In principle, effects from assimilating glider data might have minor difference when model bias is introduced from different resources.

The 'true' state is presented by a reanalysis model state. Both the 'true' state and the reference state are solutions to model equations. How much difference data assimilation will make using data from the 'true' state versus real ocean is worthy of further investigation.

2. OSSE in the Adriatic Sea (A.Aydogdu, N.Pinardi, CMCC)

2.1 Geographical setup

The Adriatic Sea is located in the northern part of the Central Mediterranean, between the Italian peninsula and the Balkans. It has a highly variable depth varying from about 30 m in the Northern Adriatic to 1200 m in the Southern Adriatic (Fig. 2.1). It is connected to the Mediterranean Sea through Otranto Strait. The Adriatic Sea is characterized by a large seasonal and spatial variability of the atmospheric forcing and the river discharge. In particular the large river discharge exceeds the evaporation and determines the estuarine like exchange with the Mediterranean Sea. The near surface circulation is mainly cyclonic with a permanent South Adriatic Cyclonic Gyre and the Middle Adriatic Cyclonic Gyre. The Mediterranean water inflow along the eastern coast, and the permanent Western Adriatic Current along the western coast is mainly determined by the salinity gradient due to the large river run-off at the northern coasts. The dense water generated during the winter in the Adriatic Sea outflows through the lower layer of the Otranto Strait and represents a significant source of dense water in the Eastern Mediterranean.
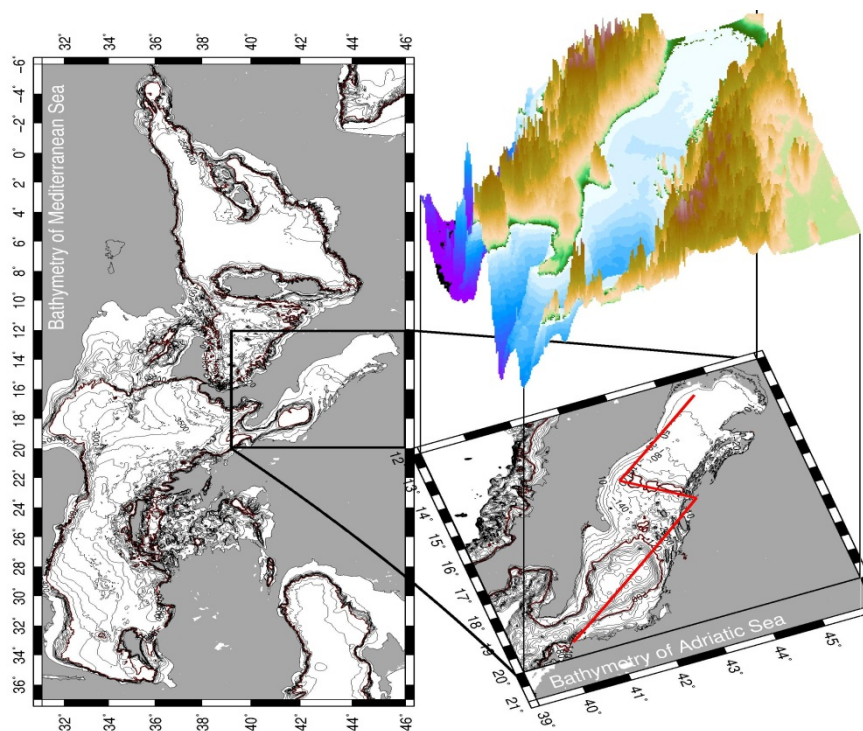
2.2 Model description

Fig. 2.1: Adriatic Sea Geometry and Bathymetry. The red line indicates the section for vertical field visualization.

The numerical model set-up in the Adriatic Sea has a constant grid resolution of 1/48° along the longitudinal and latitudinal directions that corresponds to 1.8 and 2.3 km respectively and it uses the NEMO (Nucleus for

European Modeling of the Ocean, Madec, 2008) code in its free-surface formulation. The model grid has 432 points in the zonal, and 331 in the meridional direction. In the vertical direction, it is conFig.d with 120 unevenly spaced horizontal z-levels. The bottom topography is represented by the partial cell method. Vertical grid spacing is 1 m in the top 60 m, then increases to 9 m at 100 m depth and further to 50 m at the deepest point in the Adriatic Sea. The largest spacing of 70 m is in the Ionian Sea at the deepest point (2800 m). This configuration of the model has been described in a recent paper (Gunduz et al., 2013) and in this work we have used the restart from the published 10 year simulation in January 2007.

Atmospheric forcing fields, except for precipitation, were obtained from the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim data set. The precipitation was obtained from the Merged Analysis of Precipitation (CMAP) observational data set (Xie and Arkin 1997). The ERA-Interim atmospheric forcing fields are available at the frequency of 6 hour and the horizontal resolution of 0.25°. The monthly mean CMAP data set has the horizontal resolution of 2.5°.

The model set-up has one open boundary communicating with the Mediterranean Sea positioned south of the Otranto Strait (see Fig. 2.1). The boundary conditions for temperature, salinity, sea surface height, zonal and meridional current are provided daily from outer basin scale MFS Mediterranean Model. The model uses scale selective open boundary conditions.

2.3 OceanVar Data Assimilation System

The system consists of the Adriatic set-up of the NEMO ocean model described above and the OceanVar data assimilation scheme (Dobricic and Pinardi 2008). In the OceanVar set-up the slowly evolving vertical part of temperature and salinity background error covariances is represented by monthly varying Empirical Orthogonal Functions (EOFs). They are calculated in each model point separately by the method described in Dobricic et al. (2006).

The horizontal part of background error covariances is assumed to be Gaussian isotropic depending only on distance. It is modeled by the successive application of the recursive filter in longitudinal and latitudinal directions, that provides a high computational efficiency in each iteration of the algorithm. The rapidly evolving part of the background error covariances, consisting of the sea level and the barotropic velocity components is modeled in each step of the minimization algorithm by applying a barotropic model forced by the vertically integrated buoyancy force resulting from temperature and salinity variations. The velocity is then estimated by applying the geostrophic relationship, modified along the coast by the application of the divergence dumping filter in order to eliminate the horizontal divergence. In this way OceanVar combines long term three dimensional variational scheme for the slow processes with a scheme that fully dynamically evolves the covariances by model equations for the fast processes. In particular the dynamical model used for the simulation of covariances between sea level errors and errors in temperature and salinity fields allows their very accurate estimate over areas with highly variable or shallow bottom topography typical for coastal areas.

2.4 Fishery Observing System
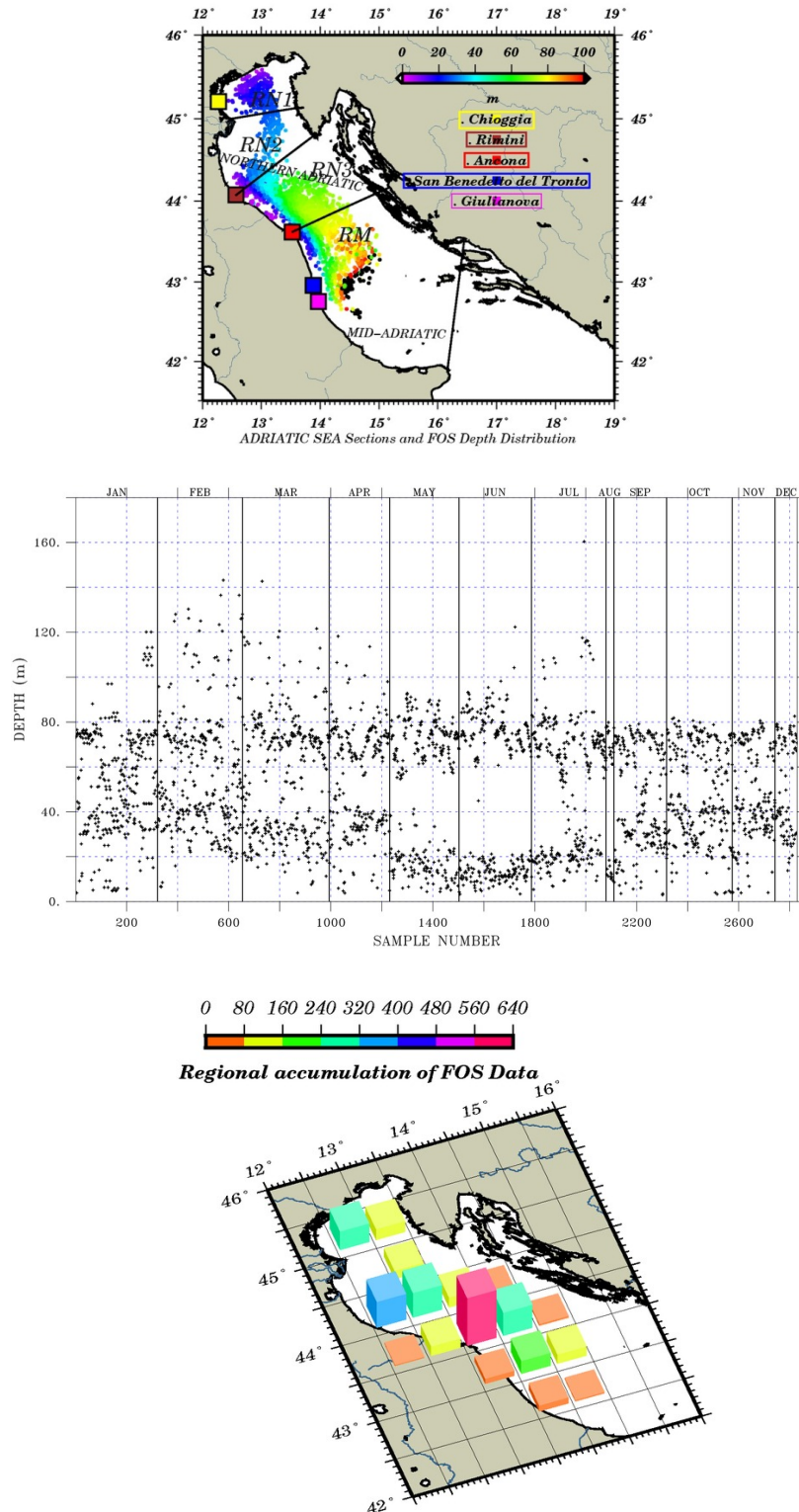
Fig. 2.2: The depth distribution of the FOS data and regional subdivision of the Adriatic Sea (top panel). Depth of the measurement caluclated as an average of the net fishing depth (middle panel). Horizontal monthly cumulative distribution of the measurements (bottom panel).

In JERICO, the Observing System Experiments (OSE) assimilate observations, provided by CNR partner, from the voluntary fishing vessels (Fishing Observing System, FOS). The FOS data used in this study consists of

seven different vessels from five different fleets in 2007. Fleets are located in Chioggia, Rimini, Ancona, San Benedetto del Trento and Giulianova from north-west to mid-west Adriatic Sea, respectively. StarOddi sensors are installed to the nets of the pelagic pair trawlers and purse seine fishing vessels. Sensor measures temperature with an accuracy of ± 0.1°C. Depth is calculated from pressure which has an accuracy ± 0.4% with the selected range for 50m-270m range. Profiles taken during the releasing and hauling of the net is excluded from the assimilation experiments due to the stabilization problem of the sensor. Moreover, the horizontal profile taken during the net drift is averaged along the track. As a result, the dataset used in the assimilation consists of point data well-distributed along the Italian side of the northern and middle Adriatic Sea (Fig. 2.2).

The measurement points reach a maximum depth of 160m but most of them stay within the first 100 m. Largest amount of data is collected by the Ancona and Rimini fleets. There were two vessels in those fleets whereas only one vessel was available in each of the other fleets. Least amount of data is collected during August due to the restrictions on fishing activities (Fig. 2.2).

2.3 Data assimilation system description

The classical OSSE methodology is explained in detail in various studies (e.g. Arnold Jr and Dey, 1986; Halliwell Jr et al., 2014; Lahoz et al., 2010). Here we use identical twin experiments in which the truth (nature run) is chosen to be a simulation for the 2007 year without any data assimilation. The model run is instead realized by perturbing the nature run mimiquing the effect of unknown physics and model inaccuracies.

The perturbed simulation is produced by using the thermocline intensified random perturbation (TIRP) method described by Pinardi et al. (2008, 2011). An initial condition perturbation is applied to the temperature and salinity fields as follows:

$$T_p(x, y, z) = T_0(x, y, z) + p(x, y) \sum e_i f_i(z) \quad (1)$$

$$S_p(x, y, z) = S_0(x, y, z) + p(x, y) \sum e_i g_i(z) \quad (2)$$

where   are the unperturbed initial conditions; $p(x,y)$ is a pseudo-random field with values in the interval [0,1.8] for temperature and [0,0.4] for salinity;   are 20 vertical empirical orthogonal functions computed from model statistics and   are their eigenvalues. The technique leads to bigger perturbations where the error covariances are larger, i.e. around the thermocline and the continental shelf. The perturbation have the possibility of grow in time due to baroclinic instabilities of the flow field. The perturbation is applied to June 01, 2006. A spin-up is performed until January 1, 2007 and the experiments are reinitialized with the perturbed initial conditions (Fig. 2.3).

PERTURBATION–TRUTH TEMPERATURE 2007/01/01 degC

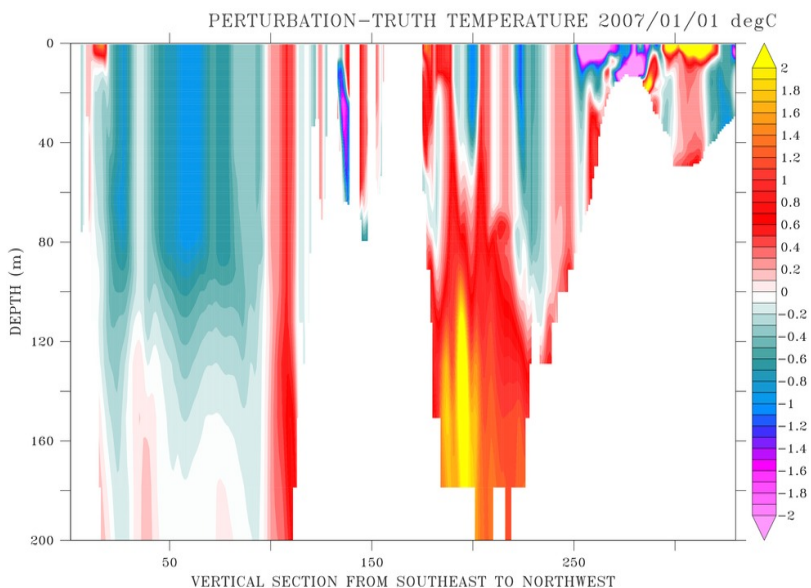VERTICAL SECTION FROM SOUTHEAST TO NORTHWEST

Fig. 2.3: Temperature difference between the perturbed run and the truth at the initial time. The field is displayed along the section of Fig. 2.1.

The Adriatic Sea circulation consists of two important permanent gyres which dominate the circulation, especially during the mixing seasons. The initial condition perturbation amplitude does not grow for more than three months and therefore it was chosen to select a wrong wind field to force the perturbed simulation: the year 2006 instead of 2007 is the used to achieve a change in these wind-driven permanent large scale structures.

The growth of the perturbation is demonstrated by the spread which is calculated by normalizing the RMS difference between perturbation run and truth by the standard deviation of the truth. According to spread, the difference between the truth and the perturbation run is below 20% in April, May and June. The perturbation grows significantly after June (Fig. 2.4). This perturbation run is used to see the impact of the synthetic data in the OSSE.

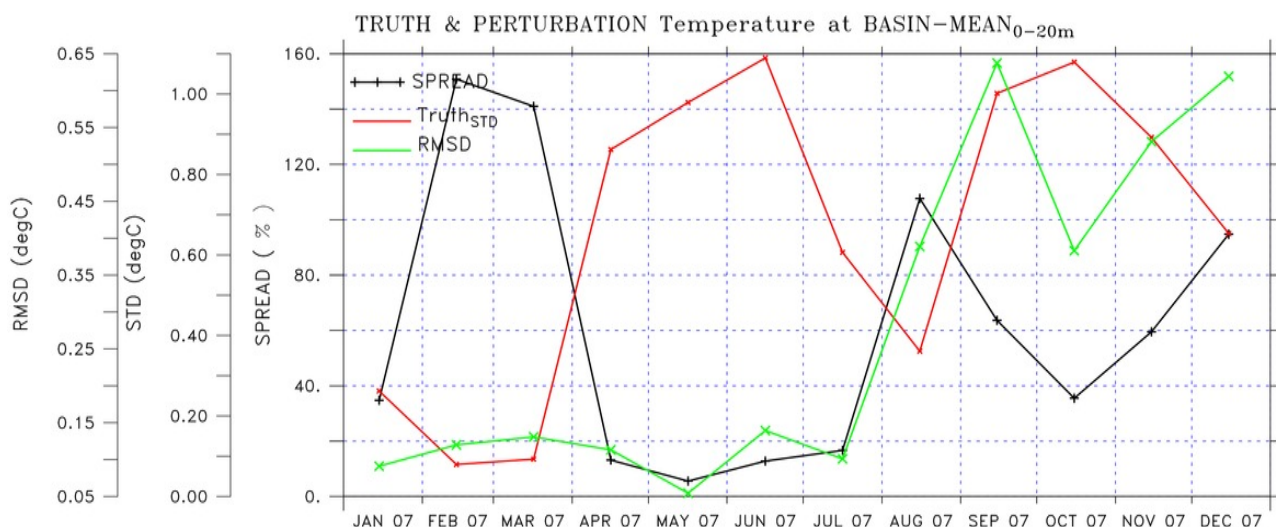TRUTH & PERTURBATION Temperature at BASIN-MEAN$_{0-20m}$

Fig. 2.4: Growth of the temperature perturbation in the 0-20 m layer in terms of the RMSD (green) and the spread (black). Perturbation significantly grows after July.

2.4 Sampling Strategy

The synthetic observations are extracted from the truth (Fig. 2.5). The first set of synthetic observations (FOS#1) is sampled exactly from the same positions of the real observations. In FOS#2 instead the measurements are extracted always at 10 m depth. An observational error is randomly added in both datasets. The error is assumed to be in the interval ±0.1°C for temperature considering the sensor accuracy. In both synthetic datasets, there are also synthetic salinity observations produced by using the same method as temperature with an observational error of 0.02 psu.
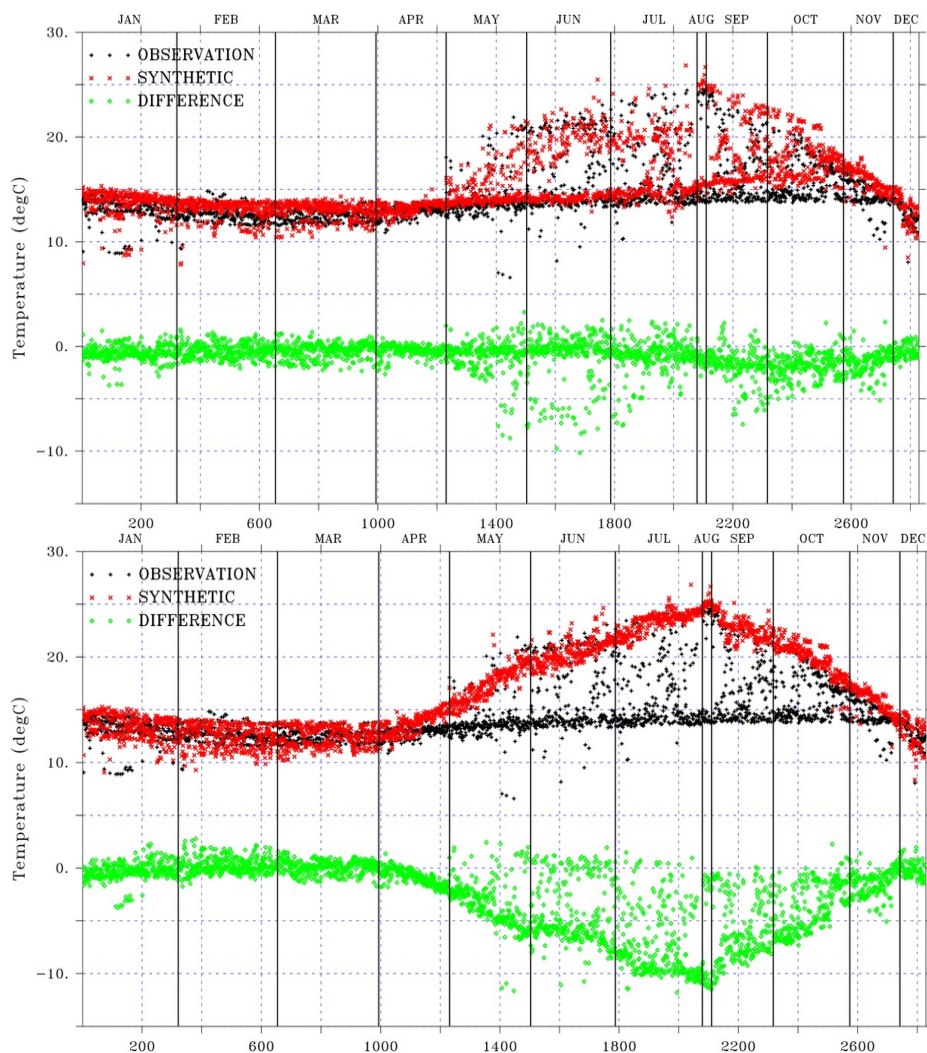
Fig. 2.5: Top panel: FOS#1 synthetic temperature data (black dots) vs real observations (red dots) and difference in temperature between the two measurements (green dots). Bottom panel: FOS#2 synthetic temperature data (black dots) vs real observations (red dots) and difference in temperature between the two measurements (green dots). FOS#1 is assimilated in OSSE01, OSSE02, OSSE03 and OSSE04. Thermocline intensified synthetic data FOS#02 is used in OSSE05.

Five OSSEs are performed to test the different FOS design (Table 1). The design of the OSSEs should be realistic and applicable in practice. Here, a possible implementation of salinity sensors in the FOS network is

tested. Two OSSEs are performed by assimilating synthetic salinity observations to see the impact. Moreover, the positions of the sensors on the net during catchments is also an important issue for the efficiency of the FOS design. Majority of the real observations are at the top 100 m. An OSSE is done to assess the impact of the observing system with seasonal thermocline focused data collection (10 m depth).

First two OSSE are performed to see the impact of the synthetic temperature observations FOS#1 which are at the same positions with real observations. The results are compared with the corresponding OSEs to see whether the OSSEs are realistic. In OSSE01, all the FOS#1 synthetic temperature data are assimilated. Then, the observations provided by the Ancona fleet are excluded in OSSE02. The OSSE03 and OSSE04 are the same with OSSE01 and OSSE02, respectively with the assimilation of synthetic salinity observations in FOS#1. Finally, in OSSE05, the idea of seasonal thermocline focused FOS design is tested. For this purpose, the second set FOS#2 of synthetic temperature data is assimilated to assess the impact.

NAME	TYPE	ASSIMILATION	VARIABLE
TRUTH	*NATURERUN*	NO	-
PERTURBATION	*SIMULATION**	NO	-
OSSE01	*ASSIMILATION**	ALL Synthetic FOS#1	TEMP
OSSE02	*ASSIMILATION**	SELECTIVE Synthetic FOS#1	TEMP
OSSE03	*ASSIMILATION**	ALL Synthetic FOS#1	TEMP + SALT
OSSE04	*ASSIMILATION**	SELECTIVE Synthetic FOS#1	TEMP + SALT
OSSE05	*ASSIMILATION**	ALL Synthetic FOS#2	TEMP

Table 1: description of OSSE experiments

The misfits, defined as the difference between the observations and the model background, are used to calculate the RMS error during the post-processing to compare the experiments. The lower the RMS error the better the solution since the synthetic observations are extracted from the truth.

2.5 Results

The OSSEs are compared with the truth run at different water column depths, namely 0-20 m, 20-50 m, 50-100 m, to have with a more robust statistics. The weekly temperature and salinity RMS errors of the perturbation run, OSSE01 and OSSE03 are shown in Fig. 2.6. The synthetic temperature observations FOS#1 improves the solution generally. Below 20 m depth, the FOS#1 reduces the RMS error along the year with some exceptions. The correction between 20-50 m, where the data is abundant, is significant. However, in the first 20 m, the results are more complicated. In this layer it is not possible to say that the assimilation is increasing the analysis quality. It improves the solution significantly during months like June and July whereas it degrades it in January and May. There is no persistent pattern along the year. The question why the assimilation of synthetic observations doesn't correct surface and sub-surface layers is still open and discussed further below.
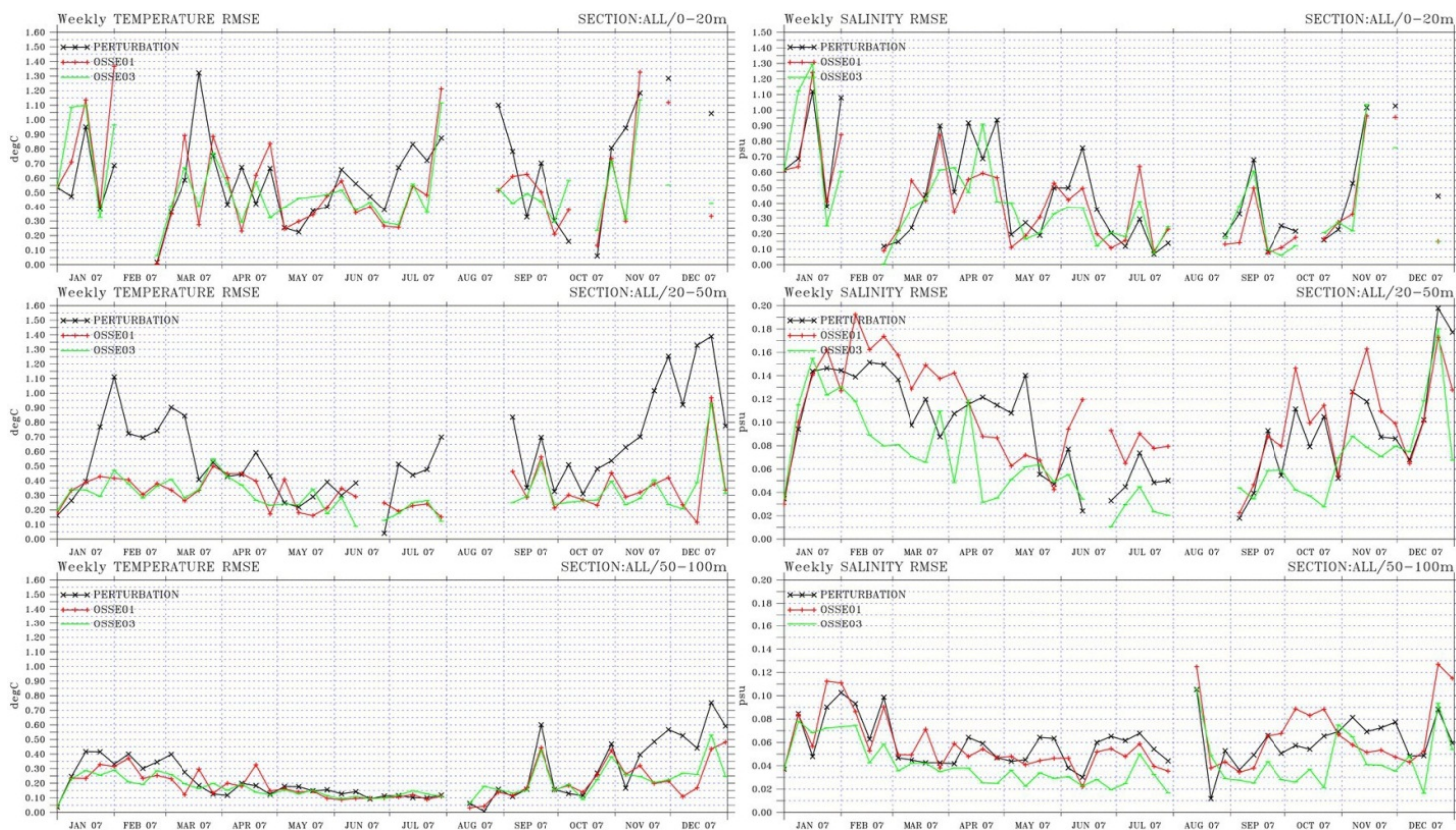
Fig. 2.6: Weekly temperature (left) and salinity (right) RMS error of perturbation run (black), OSSE01 (red) and OSSE03 (green).

Another important aspect is that the salinity and temperature corrects each other considering the error covariances provided to the 3dvar. When we assimilate the salinity in OSSE03 (Fig. 6, green), the solution for temperature is getting better below 20 m. On the other hand, the opposite is not true. Assimilation of only temperature in OSSE01 actually degrades salinity almost everywhere (Fig. 6, red). Therefore, the salinity assimilation becomes important to improve the skills. Actually, in OSSE03, the impact of salinity assimilation is significant along the year below 20 m (Fig. 2.7, green). Very similar results are obtained in OSSE02 and OSSE04 (not shown). Although the Ancona fleet provides significant amount of data, the impact of the synthetic observations by the rest of the vessels is high. Therefore, it may not be able to improve the solution furthermore.
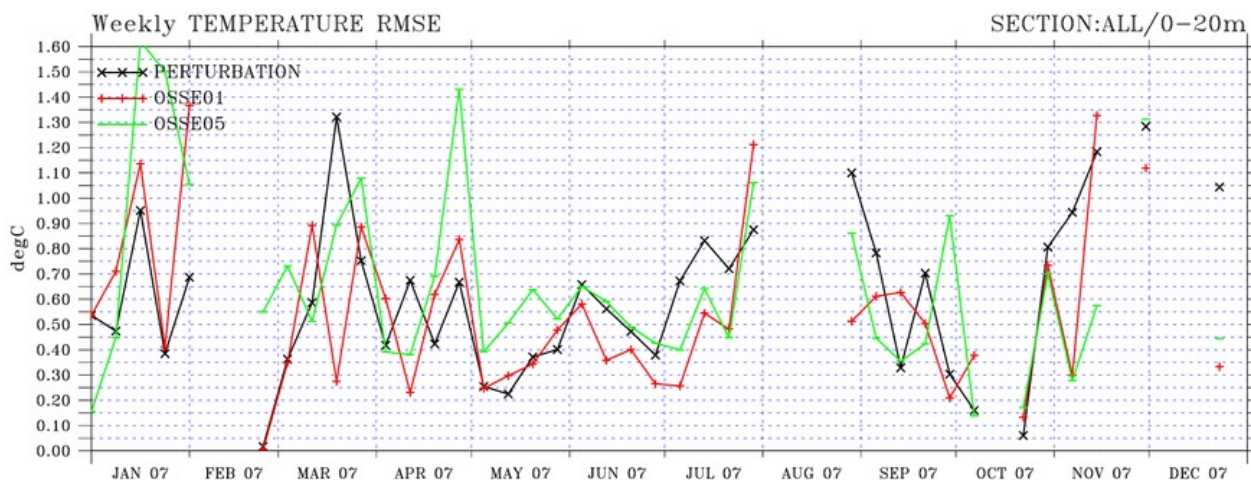
Fig. 2.7: Weekly temperature RMS error of perturbation run (black), OSSE01 (red) and OSSE05 (green).

The correction problem in the first 20 m suggested to perform another OSSE with a seasonal thermocline focused FOS. All the synthetic temperature data extracted from the same geographical coordinates of real observations but from a depth of 10 m are assimilated (Fig. 2.5). Assimilating more data around seasonal thermocline degrades the skill of prediction at that layer (Fig. 2.7). However, the data improves the prediction after October 2007. In order to understand this, we show the snapshot at 31st of December 2007 (Fig. 2.8): shows that the difference of temperature between the truth and the OSSE05 is small along the western Adriatic jet where the data is mostly accumulated (Fig. 2.8). More analysis is needed to understand the dynamics which results in a degradation of the skill of the model before October at that layer.
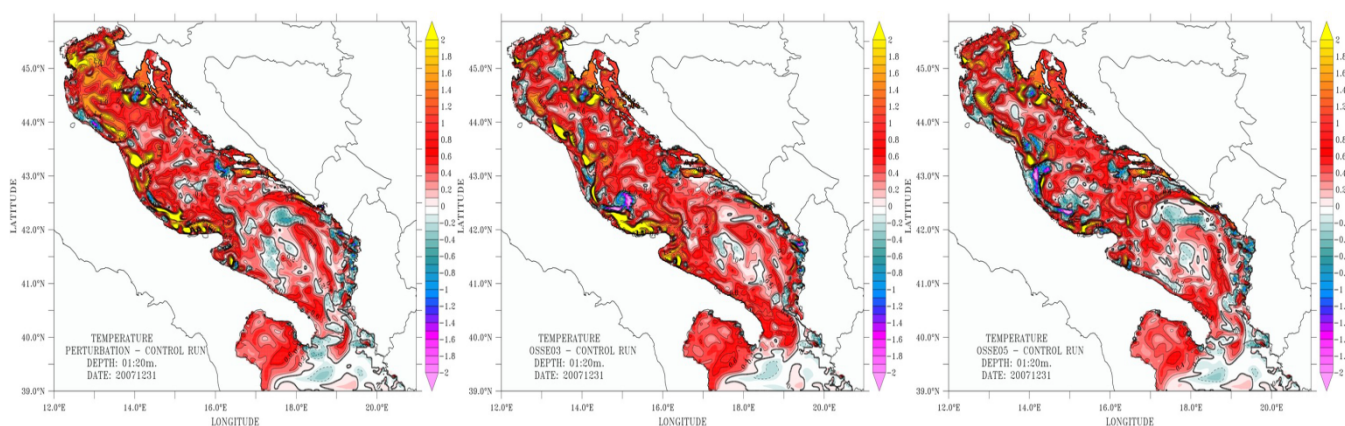
Fig. 2.8: Temperature difference between the truth and the perturbation run, OSSE03 and OSSE05 from left to right respectively.

In summary, five OSSEs were performed to assess the impact of the FOS sampling scheme and the introduction of a salinity sensor in Adriatic FOS. The results reveal that salinity observations ameliorate the analysis along the year. Moreover, the coastal processes should be investigated more to improve the impact of the data assimilation especially along the northwestern Adriatic.

References

Adcroft, A., Hill, C., & Marshall, J. (1997). Representation of topography by shaved cells in a height coordinate ocean model. *Monthly Weather Review, 125(9),* 2293-2315.

Dobricic, S., Pinardi, N., Adani, M., Tonani, M., Fratianni, C., Bonazzi, A., & Fernandez, V. (2007). Daily oceanographic analyses by Mediterranean Forecasting System at the basin scale. *Ocean Science, 3(1),* 149-157.

Dobricic, S., & Pinardi, N. (2008). An oceanographic three-dimensional variational data assimilation scheme. *Ocean Modelling, 22(3),* 89-105.

Falco, P., Belardinelli, A., Santojanni, A., Cingolani, N., Russo, A., & Arneri, E. (2007). An observing system for the collection of fishery and oceanographic data. *Ocean Science, 3(2),* 189-203.

Gunduz, M., Dobricic, S., Oddo, P., Pinardi, N., & Guarnieri, A. (2013). Impact of Levantine Intermediate Water on the interannual variability of the Adriatic Sea based on simulations with a fine resolution ocean model. *Ocean Modelling, 72,* 253-263.

Madec, G. (2008). NEMO ocean engine.

Mellor, G. L., Häkkinen, S. M., Ezer, T., & Patchen, R. C. (2002). A generalization of a sigma coordinate ocean model and an intercomparison of model vertical grids. In *Ocean Forecasting* (pp. 55-72). Springer Berlin Heidelberg.

Pasaric, M., & Orlic, M. (2004). Meteorological forcing of the Adriatic: present vs. projected climate conditions. *Geofizika, 21,* 69-87.

Raicich, F. (1994). *Note on the flow rates of the Adriatic rivers* (Vol. 94, No. 8). Tech. Rep. RF 02.

Xie, P., & Arkin, P. A. (1997). Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society, 78(11),* 2539-2558.

3. OSSE in the Bay of Biscay-English Channel

3.1 Geographical setup

3.2 Model description

3.3 Data assimilation system description

3.4 Sampling Strategy

3.5 Results

4. OSSE in the North Sea (S. Ponsar, RBINS-OD)

4.1 Geographical setup

The North Sea domain under consideration is located between 4°W to 10°E in longitude and 48.5°N to 60°N in latitude. There are three open sea boundaries: a narrow connection to the English Channel through the Dover Strait, a connection to the Baltic Sea through the Skagerrak, and a wide northern boundary. Its bathymetry varies widely, with large areas that are less than 40 meters deep (Southern and German Bights as well as the Dogger Bank) while there are deeper regions east and west of the Dogger Bank where the depths exceed 90 meters. Along the Norwegian Trench, the depth is up to 700 meters. The most important forcing mechanisms are the tides and the wind. Semi-diurnal tides are predominant at the latitude under consideration. The dominant factor governing the temperature field is the surface seasonal heating and cooling which, in the central part of the North Sea, leads to a thermal stratification of the water column in summer.

4.2 Model description

The COHERENS (Coupled Hydrodynamical-Ecological Model for Regional and Shelf Seas) model (Luyten, 2011) is a finite difference model. Simulations are performed with a horizontal resolution of 4 nautical miles in the horizontal and 20 σ-sigma levels in the vertical.

Meteorological data are supplied by the Danish Meteorological Institute (DMI) from the HIRLAM model with a temporal resolution of one hour. Tidal harmonics and daily profiles of currents, temperature, salinity and inflow/outflow conditions at the boundaries of the domain are derived from simulations with the POLCOMS (Proudman Oceanographic Laboratory) model covering a larger area. River runoffs from the Elbe, Scheldt, Rhine/Meuse, Thames, Humber, Tyne/Tees are taken into account. Baroclinic inflow/outflow conditions are imposed at the eastern boundary to include the exchange of water masses with the Baltic Sea.

4.3 Data assimilation system description

The ensemble Kalman filter developed by Evensen (1994) combines the traditional Kalman filter with Monte-Carlo methods to generate an ensemble of states representing the model error. A square root algorithm is applied at the analysis step.

Simulations are carried out for September 2001, a month during which the two dynamical regimes of the North Sea coexist (well mixed and summer stratified). An initial ensemble of states is generated from the 1st of September and is integrated without data assimilation till the 11th of September. The model error is sampled once a day using 50 ensemble members. Eight temperature profiles are assimilated once a day at midnight from the 12th of September till the 28th of September.

Synthetic temperature profiles are assimilated, they are derived using the above mentioned model setup with a horizontal resolution of one nautical mile. They show significant differences in comparison with the temperature modelled with a horizontal resolution of four nautical miles [She et al., 2006]:

- Eddy structures with scales of a few kilometers, visible along the thermal fronts in the one nautical mile resolution simulations, not resolved by the coarser grid,
- Large vertical displacements of the thermocline: this feature of the North Sea dynamics is induced by winds and tides, the amplitudes are much larger in the high resolution run.

4.4 Sampling Strategy

It has been chosen to focus on a network of buoy data for two reasons. First, satellites provide surface data with a very high spatial coverage and a temporal resolution of around one day. However, even in cloud free conditions their errors remain larger than that from CTD in situ data. Furthermore, no vertical information is available from satellite data while profile measurements can be provided either by CTD data or buoys. CTD or buoy data have a very high temporal resolution but are limited to a fixed location and exist only at a few locations; such data can be transferred in near real time and are less expensive to operate than satellites.

The data set consists of 20 temperature profiles. They are extracted at the assimilation time step from model runs generated with the same set-up but with a higher horizontal resolution of one nautical mile. Their impact on the neighboring temperature field is limited by means of an assimilation cutoff radius.

Four data sets of eight synthetic temperature profiles are assimilated. They represent four observational networks whose impact on the modelled temperature will be assessed and compared:

- one existing network,
- one existing network + 1 station,
- one existing network in which 3 stations are moved,
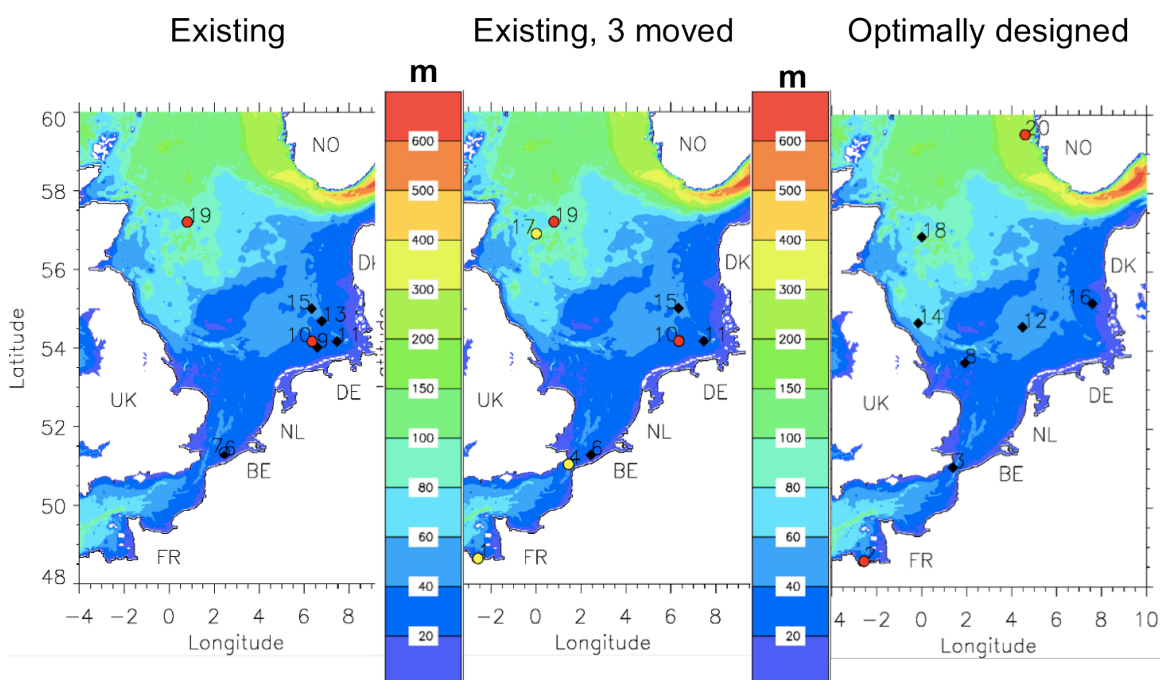- one optimally designed network.

Fig. 4.1: North Sea observational networks – Left: Existing, Center: Existing 3 stations moved, Right: Optimally designed.

Fig. 4.1 presents the observational networks that are assessed in the framework of the OSSE experiments. The existing network is made of eight buoys extracted from existing observation locations along the Belgian coast, in the German Bight and in the United Kingdom central part of the North Sea. This network is characterized by an overlap between stations as they are essentially located in coastal areas. The center of the panel represents the existing network in which three stations have been moved to reduce their overlap, particularly in the German Bight. The optimally designed network is optimal in the sense that it maximizes the number of neighboring points to which a given station is correlated. The correlation scales were deduced from SST satellite data [She et al., 2006].

The impact of the assimilation of data from these networks on model forecasts is assessed in terms of two criteria:

- the reduction of the ensemble spread on the whole North Sea domain [Mourre et al., 2006],
- the root mean square error between the model results obtained with data assimilation and the assimilated data [Wei and Malanotte-Rizzoli, 2010].

4.5 Results

Temperature profiles

Fig. 4.2 presents the root mean square error and model bias between the assimilated data and the model results without and with data assimilation for the eight stations of the optimally designed network.

At all stations, the assimilation process clearly reduces the error between the assimilated data and the model over the whole water column. This indicates that the ensemble Kalman filter is adequate for data assimilation in a North Sea model. However, at station 18, the root mean square error and bias are slightly larger at the thermocline than at the surface and bottom of the water column.
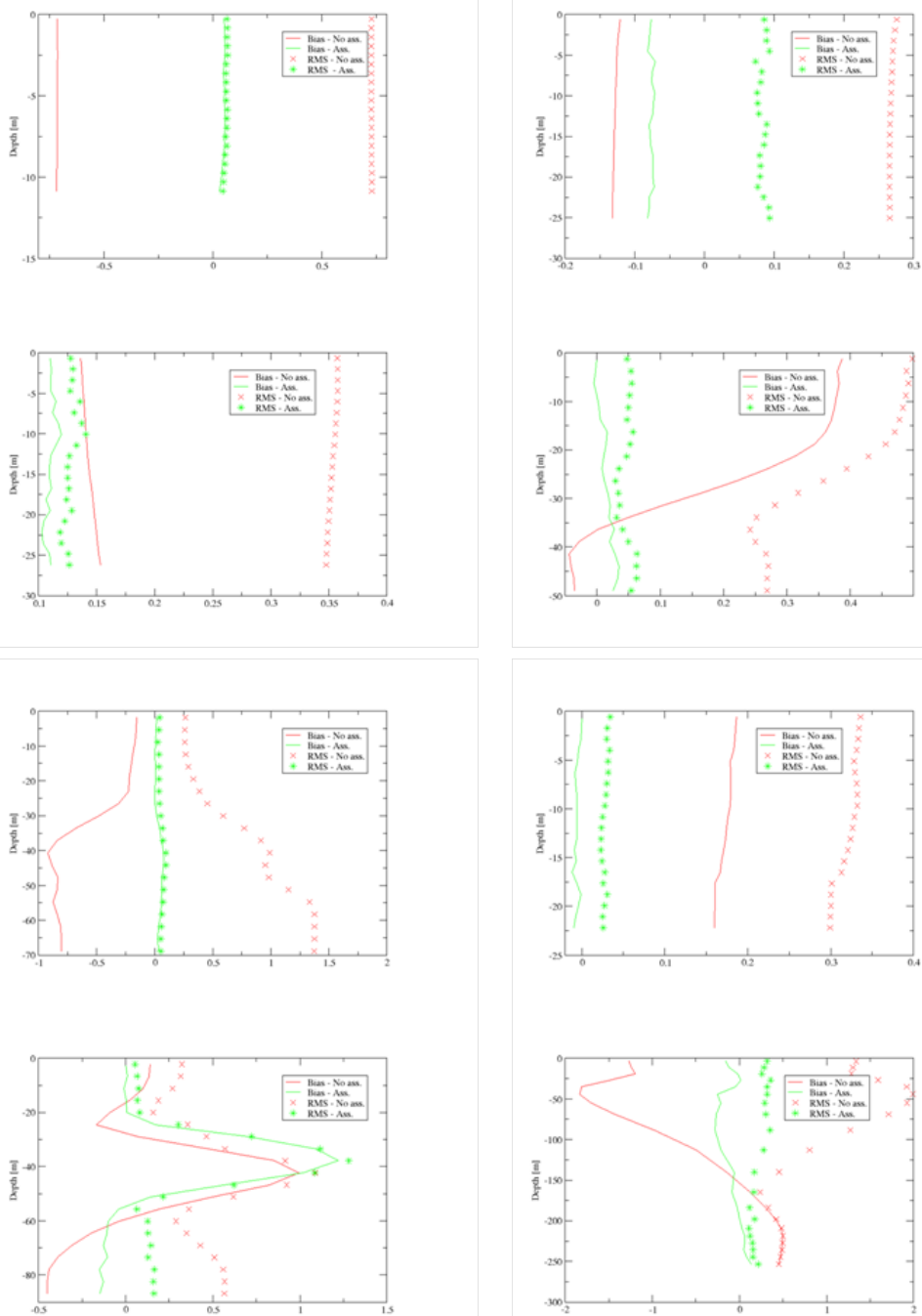
Fig. 4.2: Root mean square error (symbols) and model bias (straight lines) between the assimilated data and the model results without (red) and with (green) data assimilation – Stations of the optimally designed network.

Sea surface temperature

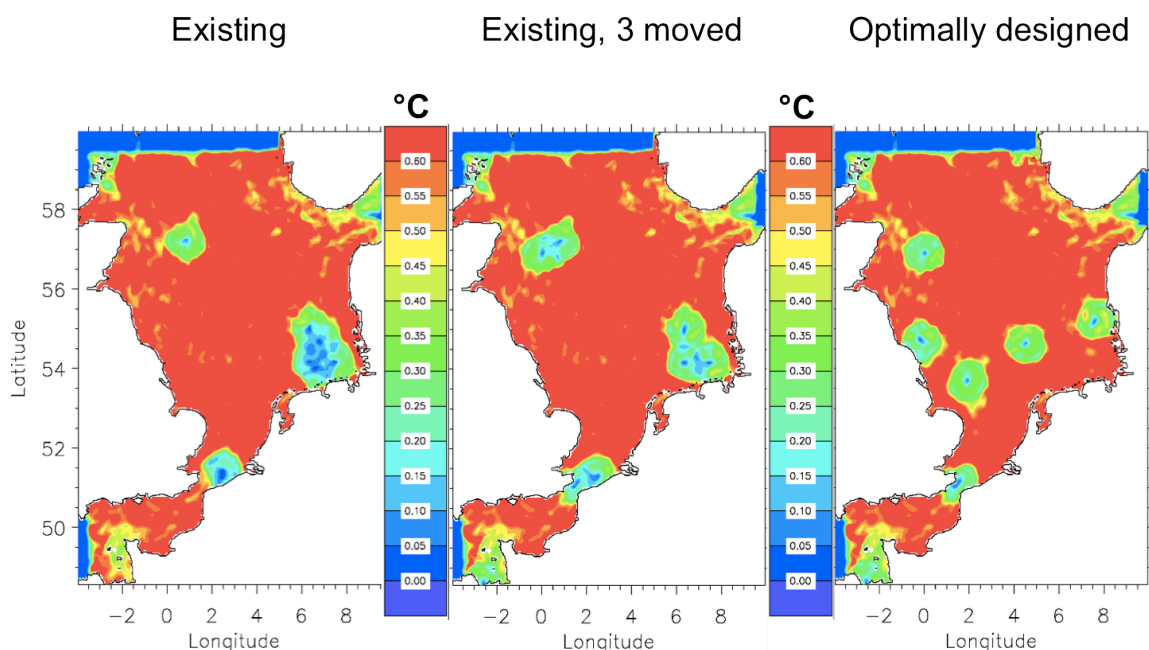Existing   Existing, 3 moved   Optimally designed

Fig. 4.3: Standard deviation of the ensemble for the sea surface temperature - Left: Existing network, Center: Existing 3 stations moved, Right: Optimally designed.

One way to assess the impact of an observational network on the model forecasts is the reduction of the ensemble spread induced by that network, with the aim of maximizing the number of points where the ensemble spread is reduced. The ensemble spread for the sea surface temperature is presented on Fig. 4.3. A comparison of the results for the existing network to those of the existing network in which 3 stations are moved and the optimally designed network indicates that the removal of the overlap between stations improves the observational network efficiency.

Network comparison

In order to further assess the network efficiency, advantage is taken of the overlap between the networks. The data of some stations are located in the assimilation radius of a station of another network in which they are not assimilated. The root mean square error between the model and the data at the assimilation location is computed.
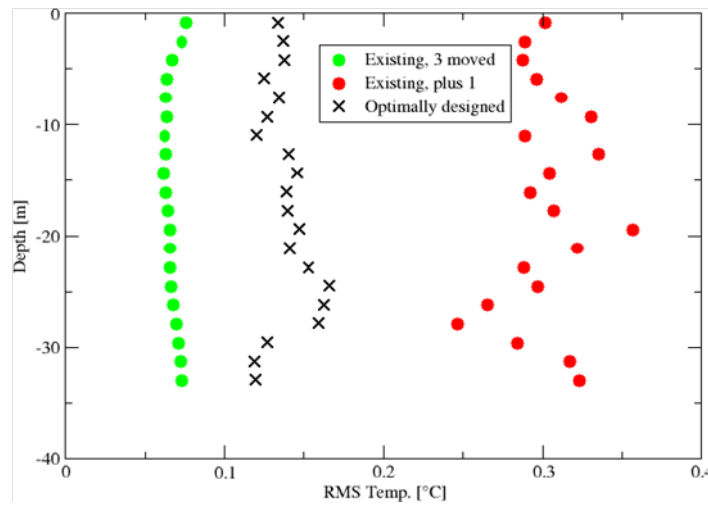
Fig. 4.4: Root mean square error between the model and data at station 4.

Fig. 4.4 presents results at station 4. Data at that station are assimilated in the "existing, three moved" network and can be used for comparison in the "optimally designed" network corrected by station 3 and in the "existing, plus one" network corrected by station 5.

As expected, the root mean square error between the model and data in the network in which these data are assimilated are the smallest (green dots). For networks in which they are not assimilated (black crosses and red dots), the "optimally designed" network (black crosses) performs better than the "existing, plus one" network.

The same kind of assessment is performed at station 13 (see Fig. 4.5). Data at that station are assimilated in the "existing" network and can be used for comparison in the "optimally designed" network corrected by station 16 and in the "existing, three moved" network corrected by stations 11 and 15.
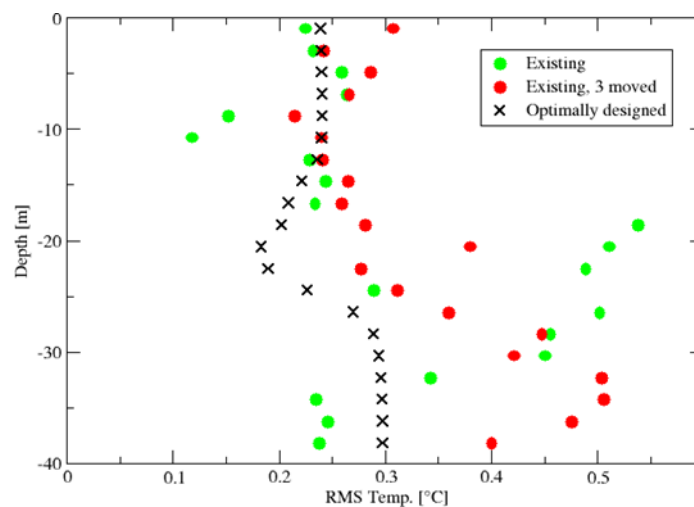

Fig. 4.5: Root mean square error between the model and data at station 13.

Results at station 13 indicate that the smallest root mean square error between the model and data is obtained in the "optimally designed" network (black crosses). In the network where the data are assimilated (green dots), this error is quite large. This is attributed to the overlap between stations of the German Bight in the "existing network".

The optimally designed network seems to perform better than existing ones mainly for two reasons. A first reason stems from the fact that it is designed to maximize the correlations with the neighboring points i.e. to minimize the variance; a second reason is that the overlap between stations is absent from the "optimally designed" network while it creates spurious features in the temperature modelled with data assimilation.

References

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, Journal of Geophysical Research, 99, 10 143- 10 162, 1994.

Luyten, P.: COHERENS – A coupled Hydrodynamical-Ecological Model for Regional and Shelf Seas: User Documentation. Version 2.1.2, RBINS-MUMM Report, Royal Belgian Institute of Natural Sciences, 2011.

Mourre, B., P. De Mey, Y. Ménard, F. Lyard and C. Le Provost: Relative performance of future altimeter systems and tide gauges in constraining a model of North Sea high-frequency barotropic dynamics, *Ocean Dynamics*, 2006 – DOI 10.1007/s10236-006-0081-2

She, J., B. Armstrup, K. Borenas, E. Buch, L. Funkquist, P. Luyten and R. Proctor: ODON: Optimal Design of Observational Networks, *Scientific report*, 2006.

Wei, J. and P. Malanotte-Rizzoli: Validation and application of an ensemble Kalman filter in the Selat Pauh of Singapore, *Ocean Dynamics*, **60**, 395-401, 2010 – DOI 10.1007/sl0236-009-0253-y

5. OSSE in the German Bight and North Sea (J. Schulz-Stellenfleth, HZG)

5.1 Geographical setup

The OSSE experiments for FerryBox and HF radar data were performed for the North Sea and some experiments also specifically for the German Bight. The North Sea is a marginal sea of the Atlantic Ocean. As a shelf sea on the European Continental Shelf, it connects to the ocean through the English Channel in the south and the Norwegian Sea in the north. It is more than 970 kilometres long and 580 kilometres wide, with an area of around 750,000 square kilometres. The dynamics in the North Sea is very much dominated by tides. Due to the small water depth also the bathymetry and bottom roughness play an important role. The German Bight is the southeastern bight of the North Sea bounded by the Netherlands and Germany to the south, and Denmark and Germany to the east. To the north and west it is limited by the Dogger Bank. The Bight contains the Frisian and Danish Islands. The German Bight has a typical tidal range of 2–4 m and a dominant period of 12.4 h. The largest non-tidal variations are caused by atmospheric low pressure systems, either as external surges from the North Atlantic or internally generated surges. During strong storm events water levels can exceed 4 m above mean sea level. The German Bight is furthermore characterised by very shallow water with Wadden Sea areas falling dry during low tide. The region is very busy regarding offshore operations (e.g. offshore wind farms) and ship traffic (e.g., to the harbor of Hamburg). Apart from the complicated tidal dynamics as e.g. described in Stanev et al. (2014), sediment transport and current/ocean wave interaction processes play an important role. Another important component is the fresh water input from the rivers Elbe, Weser and Ems, which has an impact on the stratification and coastal circulation (e.g., Staneva et al., 2009).
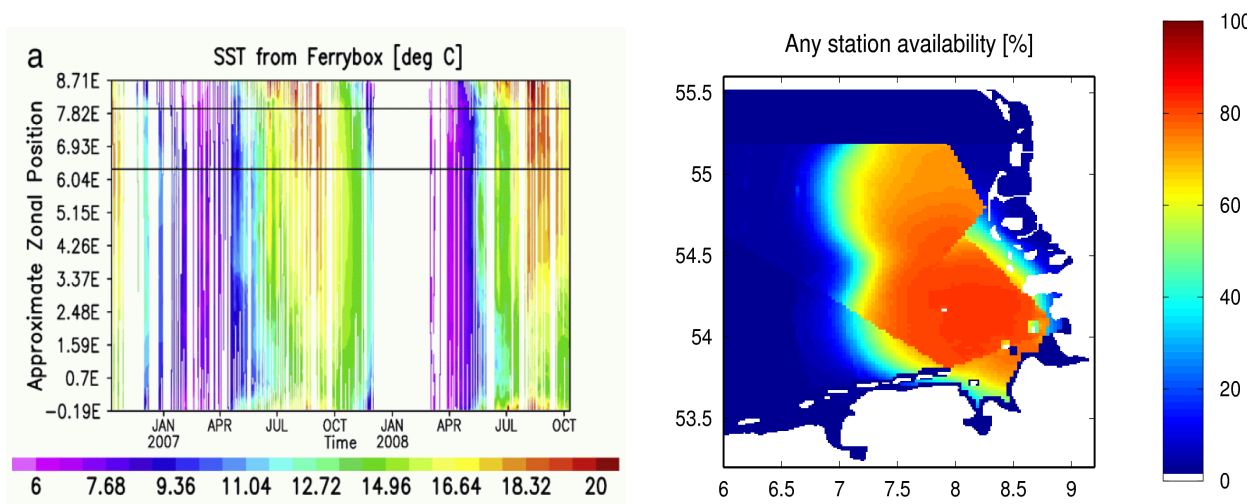

Fig. 5.1: (left) Example of SST data acquired by the FerryBox system operating between Cuxhaven and Immingham. (right) Example of HF radar coverage obtained with three radar stations located at Wangerooge, Büsum, and Sylt.

5.2 Model Description

For the OSSE experiments in the North Sea and the German Bight data from two models were used:

I) GETM

GETM is a primitive equation model, in which the equations for the three velocity components and sea surface height, as well as the equations for turbulent kinetic energy and the eddy dissipation rate are solved. The model is run on a spherical grid with 1 km resolution. Terrain following equidistant coordinates (σ-coordinates) are used in the vertical. OSSE studies for the North Sea were conducted using a setup with 5 km horizontal resolution. The water column is discretised into 21 nonintersecting layers. The model is forced by 1) atmospheric fluxes estimated by the bulk aerodynamic formula using 6-hourly ECMWF re-analysis data (wind, atmospheric temperature, relative humidity and cloud cover) and simulated by the model SST, 2) hourly river run-off data provided by the Bundesamt für Seeschifffahrt und Hydrographie (BSH), and time varying lateral boundary conditions of sea surface elevations and salinity. More details on the model setup can be found in Staneva et al. (2009).

II) MYOCEAN NEMO

In parallel data from the MYOCEAN North West Shelf model were used. The Forecasting Ocean Assimilation Model 7km Atlantic Margin model (FOAM AMM7) is a coupled hydrodynamic-ecosystem model, nested in a series of one-way nests to the Met Office global ocean model. The hydrodynamics are supplied by the Nucleus for European Modelling of the Ocean (NEMO) with an Analysis Correction (AC) data assimilation scheme for sea surface temperature. As part of the MYOCEAN system the model provides among others hourly data of surface elevation as well as ocean currents, salinity and temperature at 24 vertical layers.

5.3 Data assimilation system description

OSSE experiments were performaned using the approaches described in Grayek et al. (2011) and Schulz-Stellenfleth and Stanev (2010). The first approach is a modified optimal interpolation technique with dayly model restarts described in more detail in Grayek et al. (2011). The second technique uses a statistical method to estimate error bars for state estimates based on model and observation errors.

The basic concept explained in more detail in Schulz-Stellenfleth and Stanev (2010) is as follows: Lets denote by P the background covariance matrix derived from the numerical model data. In practice a reduced rank approximation of P was used according to

$$P = V\,U\,V^{T} \tag{1}$$

with an orthogonal matrix V containing the EOFs and a diagonal matrix U containing the respective eigenvalues. If we furthermore denote the observation error covariance matrix by G and the observation operator by H, the reconstruction error ρ, which tells us how well we can estimate the ocean state from a combination of model and observation information is given by

$$\rho = tr\!\left(P - P\,H^{T}B^{-1}H\,P\right)^{1/2} \tag{2}$$

where "tr" denotes the trace of a matrix and B is defined as

$$B = H P H^T + G \quad .$$
(3)

This technique was applied for different observation data sets, where the selection of measured variables and their location go into the definition of H and the respective observation errors are defined in G. As commonly done in literature it was assumed that the measurement errors are independent, i.e, G is diagonal.

5.4 Sampling Strategy

For the OSSE experiments two different data sets were used:

I) FerryBox data

A FerryBox is an autonomous measurement, data logging and transmission system, which operates continuously while the carrying ship is on its way (Petersen et al., 2011). Measurements are made using devices, which are either in direct contact with or sample from a continuous flow of seawater taken from a water depth of 4-6 m. The vessel position is tracked by Global Positioning System (GPS). It is connected to a station on shore via Global System for Mobile Communications (GSM) or satellite for remote control and data transfer. The basic sensors used in this study measure turbidity, temperature, salinity, and chlorophyll a fluorescence. The North Sea routes so far equipped with FerryBox systems are the ones between Buesum and Helgoland, Cuxhaven and Harwich, Cuxhaven and Immingham (compare Fig. 5.1) and recently
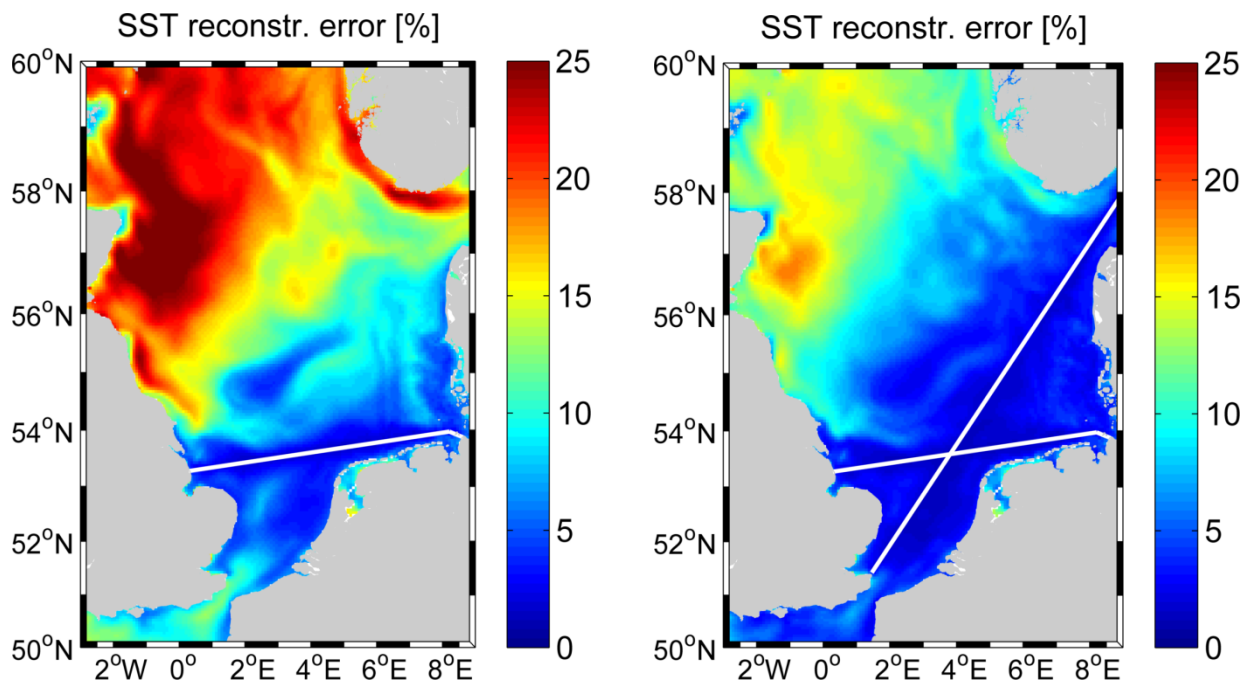
Fig. 5.2: Relative re-construction errors for sea surface temperature using one (left) and two (right) FerryBox lines.

between Hamburg, Cuxhaven, Chatham, Moss and Halden. The typical cruising speed is 15 knts. The sampling rate is 10 seconds. Depending on the travel distance, the routes provide the following revisit times: Buesum-Helgoland, daily, Cuxhaven-Immingham, less than 36h, Hamburg-Cuxhaven-Chatham-Moss-Halden about 8 days. The Ferry routes do not change substantially. However, individual tracks show small deviations one from another. Therefore, to simplify the analysis we relate the data to an averaged track. The maximum deviations from the averaged track can reach 10km.

I) <u>HF radar data</u>

An ocean surface current transporting the Bragg resonant ocean waves causes a Doppler shift (Barrick,1978; Stewart and Joy, 1974). This shift can be converted to the underlying current speed towards or away from the radar, which is the radial component u_r of the 2-dimensional (2D) surface current. Based on this principle, three HF radars have been installed on the island of Wangerooge, at Büsum, and on the island of Sylt to monitor ocean currents and waves in the German Bight. These systems cover the eastern part of the German Bight and are WERA type radars (Gurgel et al., 1999) operated in the 10.8 MHz (Büsum and Sylt) and 12.1 MHz (Wangerooge) frequency range. The spatial resolution is 1.5 km in range and about 3 degrees in azimuth. Measurements are taken every 20 min and represent 10 min averages. Due to the working frequency, the radar couples to 12.5 m (12.1 MHz) and 13.9 m (10.8 MHz) long ocean waves by Bragg scattering and the radar echoes provide information on ocean currents within a surface layer of about 1 m (Stewart and Joy, 1974). The working range of the WERAs mainly depends on salinity, sea state, working frequency, and electromagnetic noise (Radio Frequency Interference (RFI), background noise, and ionospheric reflections). Typically the radar reaches out to 120 km off the coast.
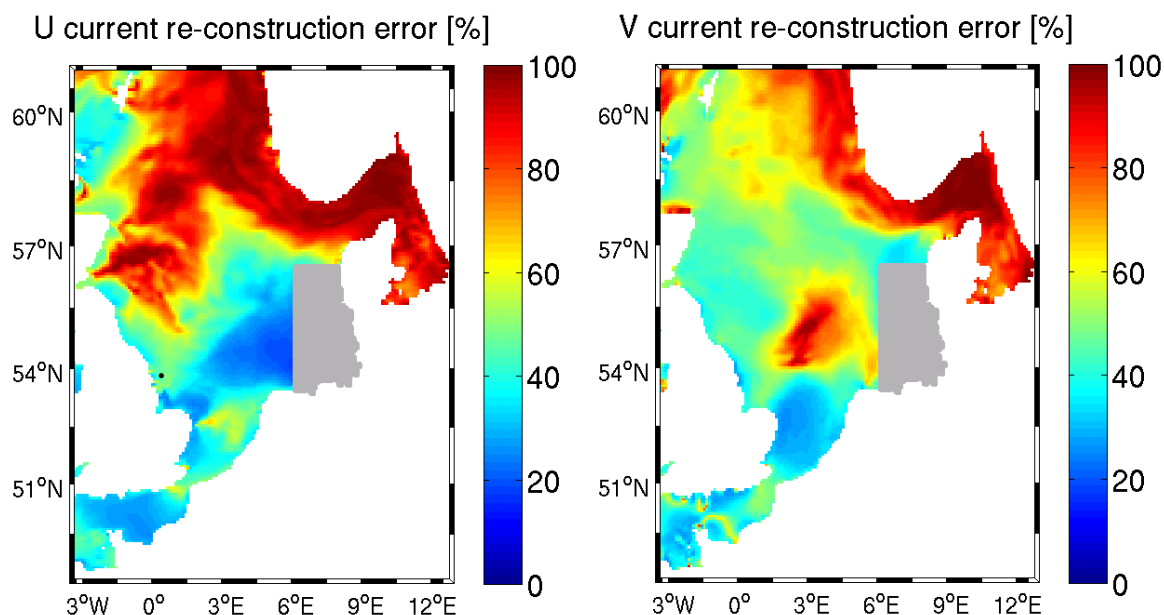
Fig. 5.3 Reconstruction errors for the zonal (left) and meridional (right) surface current component assuming that the entire German Bight (grey box) is covered with surface current measurement.

The range and coverage achieved by the antenna stations is illustrated in Fig.2. Colors indicate the availabiliy of at least one station. In the OSSE experiments comparable surface current measurments were siumulated at different locations around the North Sea.

5.5 RESULTS

As an example Fig. 5.2 shows the relative reconstruction errors to be expected from combined use of FerryBox and numerical model SST data. The errors are normalised, i.e., given in percent with respect to the background variance.   The situation with one FerryBox line (left) is compared to a configuration with one additional line further North. One can see that the additional line has a significant impact on the SST errors. In particular the SST estimation fore the Norwegian trench is improved significantly. Similar experiments were also performed for surface salinity. The main results resulting from these experiments can be summarized as follows:

- As expected the area for which useful estimates for SST can be obtained from the FerryBox data is more or less centred around the ship track with correlation length of the order of 100 km.

- The correlation lengths obtained for surface salinity are in general shorter than those for SST

- The surface salinity in the estuaries is of particular complexity and the extrapolation of FerryBox measurements in these areas is difficult

- The information propagation for SST and SSS shows a seasonal cycle. SST has longer correlation length in winter and SSS particularly short correlation lengths in spring

One has to emphasize that these studies are based on quite simplified models for the observation errors. In particular no bias errors, which can have a significant impact in practice, were considered. Another important aspect to stress is that alternative SST measurements are also provided by satellite radiometers and an efficient approach seems to be a combination of both data sources. Regarding SSS measurements there are no reliable measurements from space available for coastal areas and FerryBox systems can provide valuable information on spatial salinity gradients.

For the HF radar different locations and HF radar configurations were analysed using the described statistical approach as well. Fig. 5.3 shows an example of reconstruction errors maps obtained for the zonal (left) and meridional (right) surface current component assuming that both current components are measured within the entire German Bight (grey box). One can see that there are quite strong far field effects in particular with the English, Belgian and Dutch coast. For the Norwegian coast we see a stronger decorrelation, which is probably due to the dynamics associated with the Norwegian trench. In the meridional component we furthermore see an area of higher re-construction errors in the vicinity of the amphidromic point. This is most likely due to the weak currents in this area, which lead to a low signal to noise ratio in the re-construction. As a general conclusion one can say the following:

- It was demonstrated that HF radar data have upscaling capabilities. The specific information propagation depends on the type of model errors. For example the observations can impact large areas in case of simple timing and amplitude errors.

- From the statistical point of view there seems to be a decoupling between the currents along the Norwegian trench and the remaining North Sea, which means that HF radars along the Norwegian coast would be beneficial to get a more complete picture of the North Sea

- In the near coastal areas, e.g., the Elbe estuaries, currents are strongly steered by bathymetric features, which are often not well known. In these areas HF radar data have a big potential to improve systematic errors in the model.

As before one has to stress that the used errors models in this analysis are very simple, i.e., no bias errors were considered. Such errors can have a significant impact and will be investigated in more detail in future studies. It is also necessary to investigate more closely the potential of a combined use of HF radar data with traditional tide gauge measurements.

Literature

Barrick, D.E. "HF Radio Oceanography - a Review." *Boundary Layer Metereology* 13 (1978): 23–43.

Burchard, H., and K. Bolding. *GETM - a General Estuarine Transport Model*. European Comission, 2002.

Grayek, S., J. Staneva, J. Schulz-Stellenfleth, W. Petersen, and E.V. Stanev. "Use of FerryBox Surface Temperature and Salinity Measurements to Improve Model Based State Estimates for the German Bight." *J. Mar. Syst.* 88 (2011): 45–59. doi:10.1016/j.jmarsys.2011.02.020.

Gurgel, K.-W., G. Antonischki, H.-H. Essen, and T. Schlick. "Wellen Radar (WERA): A New Ground Wave HF Radar for Ocean Remote Sensing." *Coastal Engineering* 37 (1999): 219–134.

Petersen, W., F. Schroeder, and F.-D. Bockelmann. "FerryBox: Application of Continuous Water Quality Observations along Transects in the North Sea." *Ocean Dynamics*, 2011.

Schulz-Stellenfleth, J., and E.V. Stanev. "Statistical Assessment of Ocean Observing Networks - A Study of Water Level Measurements in the German Bight." *Ocean Modelling* 33 (2010): 270–82. doi:10.1016/j.ocemod.2010.03.001.

Staneva, J., E. Stanev, J.-O. Wolff, T. H. Badewien, R. Reuter, B. Flemming, A. Bartholomae, and K. Bolding. "Hydrodynamics and Sediment Dynamics in the German Bight. A Focus on Observations and Numerical Modeling in the East Frisian Wadden Sea." *Cont. Shelf Res.* 29 (2009): 302–19. doi:10.1016/j.csr.2008.01.006.

Stanev, Emil V, Rahma Al-Nadhairi, Joanna Staneva, Johannes Schulz-Stellenfleth, and Arnoldo Valle-Levinson. "Tidal Wave Transformations in the German Bight." *Ocean Dynamics* 64 (2014): 951–68. doi:10.1007/s10236-014-0733-6.

Stanev, E.V., F. Ziemer, J. Schulz-Stellenfleth, J. Seemann, J. Staneva, and K.W. Gurgel. "Blending Surface Currents from HF Radar Observations and Numerical Modelling: Tidal Hindcasts and Forecasts," Accepted by J. Atm. Ocean Techn., 2014.

Stewart, R.H., and J.W. Joy. "HF Radio Measurements of Surface Currents." *Deep-Sea Research* 21 (1974): 1039–49.