Assimilating High-resolution Sea Surface Temperature Data Improves the Ocean Forecast Potential in the Baltic Sea

Ye Liu[1], Weiwei Fu[2]

1. Swedish Meteorological and Hydrological Institute, Norrköping 60176, Sweden.

2. Department of Earth System Science, University of California, Irvine, California, 92697, USA.

Correspondence to: Ye Liu (ye.liu@smhi.se)

Abstract. We assess the impact of assimilating the satellite sea surface temperature (SST) data on the Baltic forecast, practically on the forecast of ocean variables related to SST. For this purpose, a multivariable DA system has been developed based on a Nordic version of the Nucleus for European Modelling of the Ocean (NEMO-Nordic). We use a local Singular Evolutive Interpolated Kalman (LSEIK) filter to characterize correlation scales in the coastal regions. High resolution SST from OSISAF is assimilated to verify the performance of DA system. The assimilation run shows very stable improvements of the model simulation as compared with both independent and dependent observations. The SST prediction of NEMO-Nordic is significantly enhanced by the DA system. Temperatures are also closer to observation in the DA system than the model results in the water above 100 m in the Baltic Sea. In the deeper layers, salinity is also slightly improved. Besides, we find that Sea level anomaly (SLA) is improved with the SST assimilation. Comparison with independent tide gauge data show that overall root mean square error (RMSE) is reduced by 1.8% and overall correlation coefficient is slightly increased. Moreover, the sea ice concentration forecast is improved considerably in the Baltic proper, the Gulf of Finland and the Bothnian Sea during the sea ice formation period, respectively.

1. Introduction

Monitoring the marine status of the Baltic Sea with relevant resolution and accuracy is a key requirement to serve the marine policy for detecting the influence of human activities on the environment and better understanding the response of ocean to accelerating global climate change. The Baltic Sea is one of the largest brackish seas in the world. It is a semi-enclosed basin, whose hydrography is highly variable and influenced by large-scale atmospheric processes and significant influx of freshwater from rivers runoff and precipitation (Leppäranta and Myrberg, 2009). In addition, the water exchange between the North Sea and Baltic Sea through the Danish straits is hindered by shallow topographic restrictions in the transition zone (Fig. 1).

A characteristic feature of numerical forecast in the Baltic Sea is in itself a major challenge because of complex topography and rich dynamics. A number of ocean forecasting systems for the Baltic Sea have been developed using hydrological model by operational agencies around this region. Traditionally, these models have a horizontal resolution of 1–5 km and approximately 20–100 layers in vertical structure (Omstedt et al. 2014). Due to the geographic location and conditions of the Baltic Sea, even higher resolutions are often needed to better understand the circulation dynamics. However, even ocean circulation models with a particularly high spatial resolution (e.g. 1 km) cannot resolve all dynamically important physical processes in the ocean (Malanotte-Rizzoli and Tziperman, 1996). In general, the forecast quality for a numerical model depends on initial conditions, boundary conditions (lateral, open boundaries as well as meteorological forcing and bathymetry) and a robust numerical model itself. As an operational forecasting agency, the Swedish Meteorological and Hydrological Institute's (SMHI) needs to issue well-informed forecasts and warnings for decision making by other authorities during e.g. severe weather events, but also to the public. To improve the forecast quality, the core three-dimensional dynamic model of the SMHI operational forecast system has recently migrated to the Nordic version of the Nucleus for European Modelling of the Ocean (NEMO-Nordic).

In additional to model development, an extended observational network has been established by the joint efforts of the countries surrounding the Baltic Sea. The observation platforms include vessels, buoys, coastal stations, satellite, etc. Specially, the observations from satellite have dominated the coverage of SST observational networks in the Baltic Sea (She et al. 2007). Among satellite products, the SST is most popularly and widely used for the operational forecast, reanalysis or validation of the model because of both its coverage and properties. SST acts as a medium between atmospheric and oceanic variations through activation of coupling mechanisms. SST is also a key ocean variable to link many processes that occur in the upper ocean, for example, air-sea exchange of energy, primary productivity, and formation of water masses (Tranchant et al., 2008).

A realistic forecast of SST is essential to an ocean forecasting system. SST is especially important for the Baltic Sea that the average water depth is only 56 m and its surface water is directly related to the bottom water by the mixing in the shallow sub-basins. Recently, the applications of SST for forecasting and analyzing the status of the North Sea and Baltic Sea have received particular attention. In the short-term forecast, Losa et al. (2012, 2014) investigated the systematic model uncertainties for forecasting the North and Baltic Seas by assimilating the Advanced Very High Resolution Radiometer (AVHRR) SST data. Nowicki et al. (2015) applied SST observed from Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) into 3D coupled ecosystem model of the Baltic Sea with the Cressman analysis scheme. O'Dea et al. (2016) enhanced the SST prediction skill of the operational system by assimilating both in-situ data and level 2 SST data provided by the Global Ocean Data Assimilation Experiment High-Resolution SST (GHRSST) into a European North-West shelf operational model. Moreover, SST has been used in the long-term analysis in this region. For instance, Stramska and Bialogrodzka (2015) analyzed spatial and temporal variability of SST in the Baltic sea based on 32-years of satellite data, which indicate that there is a statistically significant trend of increasing SST in the entire Baltic sea. However, these long-term SST data haven't been used to verify the application of sophisticated DA methods for hydrography model in the Baltic profiles simulation, especially at the Baltic deep water regions. Another important question is: what amount of satellite SST can improve long-term forecast of ocean variables related to SST in the Baltic Sea.

The objective of this study is to address the impact of assimilating a high resolution SST product on the forecast of the Baltic Sea, particularly the forecast of SST related variables like sea level and sea ice. It is also the first time that satellite SST from the Ocean and Sea Ice Satellite Application Facility (OSISAF) was assimilated into NEMO-Nordic model (NEMO variant for the North Sea and Baltic Sea). For operational forecast, the SST from OSISAF is the most important dataset in the Baltic Sea because it differs from hindcast analyzed product like OSTIA (Operational SST and Sea Ice Analysis) data. As a level 2 product, the OSISAF SST has both good temporal and spatial coverage in the Baltic Sea. As there is no hindcast information included in the OSISAF SST, we are able to assess direct impacts of assimilating SST observations. Therefore, exploring the potential of this product is critically important to further improving the new operational forecast system. In addition, our study will enrich the reanalysis database of the Baltic Sea. In this study, we use the Singular Evolutive Interpolated Kalman (SEIK) filter (Pham, 2001) to account for the model uncertainties arising from a wide range of spatial and temporal scales (Haines, 2010). One of our focuses is the impact of SST on the modeled sea level and the sea ice in the Baltic Sea. For the whole Baltic Sea, how the SST assimilation influences the temperature and salinity (T/S) on the different depth is another focus of this study.

The outline of the paper is as follows: the model configuration and SEIK scheme are described in Section 2. An overview of the observations used in this study is presented in Section 3. The implementation of DA experiment is given in section 4 together with the sampling of ensemble and localization. Results are compared with observations for temperature, salinity, sea level anomaly and sea ice in Section 5. In this section, the impact of data assimilation on the forecasts is also investigated. Conclusions and discussions are given in section 6.

2. Methodology

2.1 NEMO-Nordic

NEMO (Nucleus for European Modelling of the Ocean; Madec, 2008) has been set up at SMHI for the North Sea and the Baltic Sea, a configuration called NEMO-Nordic (Hordoir et al., 2015) (Fig. 1). Open boundaries are implemented in northern North Sea between Scotland and Norway and in the English Channel between Brittany and Cornwall, respectively (Hordoir et al., 2013). In this study, NEMO-Nordic employs a horizontal resolution of 2 nautical miles (3.7 km) and 56 vertical levels, and with a vertical resolution of 3 m close to the surface, decreasing to 22 m at the bottom of the deepest part of the Norwegian trench. NEMO-Nordic uses a fully nonlinear explicit free surface (Adcroft and Campin, 2004). A bulk formulation is used for the surface boundary condition (Large and Yeager, 2004). The ocean model is coupled to the Louvain-la-Neuve Sea Ice Model (LIM3) sea ice model (Vancoppenolle et al., 2008) with a constant value of 10^{-3} PSU for the sea-ice salinity. A time-splitting approach is used to compute a barotropic and a baroclinic mode, as well as the interaction between them. A Tidal Inversion Model is used to define the barotropic mode at the open boundary conditions (Egbert and Erofeeva, 2002). 11 tidal harmonics are defined for sea level and barotropic tidal velocities. In addition, a coarse resolution barotropic storm surge model covering a large area of the Northern Atlantic basin provides wind-driven sea level that is added to the tidal contribution. The T/S data at the open boundary are provided by the Levitus climatology (Levitus and Boyer, 1994). Radiation conditions are applied to calculate baroclinic velocities at these boundaries. A quadratic friction is applied with a constant bottom roughness of 3 cm, and the drag coefficient is computed for each bottom grid cell. NEMO-Nordic uses a TVD advection scheme with a modified leapfrog approach that ensures a very high degree of tracer conservation (Leclair and Madec, 2009). Unresolved vertical turbulence is parameterized with κ-ε scheme (Umlauf and Burchard, 2003). In addition, Galperin parameterization is used to obtain a stable long-term stratification for the Baltic Sea (Galperin et al., 1988).

A Laplacian isopycnal diffusion is used for both momentum and tracers with a diffusion parameter that is constant in time, but varies in space. Additional strong isopycnal diffusion is used close to the Neva river inflow (Gulf of St. Petersburg) in order to avoid negative salinities. The bottom boundary layer is parameterized to ease the propagation of saltwater inflows between the Danish Straits and the deepest layers of the Baltic Sea (Beckmann and Doscher, 1997). A free-slip option is used for lateral boundaries.

The model is forced by meteorological forcing derived from a downscaled run of Euro4M reanalysis (Dahlgren et al., 2014). The downscaling is based on the regional atmospheric model RCA4 (Samuelsson et al., 2011) which uses the reanalysis data as boundary conditions. A runoff database provides the river flow to NEMO-Nordic (Donnelly et al. 2016); it includes inter-annual variability for the Baltic Sea basin and is based on climatological values for the North Sea basin. The salinity of the river runoff is set to a constant value of 10^{-3} PSU, which is the same value used for the sea-ice to avoid any negative salinity.

2.2 Local Sigular Evolutive Interpolated Kalman (LSEIK) filter

The method used to assimilate SST into NEMO-Nordic is the Local Singular Evolutive Interpolated Kalman (LSEIK) filter (Pham et al., 2001, Nerger et al. 2006). This is a sequential data assimilation scheme, which is an error subspace extend Kalman filter that uses a minimum number of ensemble members to reduce the prohibitive computation burden (Pham, 2001). The LSEIK filter proceeds in correction and forecast step:

1. Forecast: the analysis state $\mathbf{X}^a$ at time t_{i-1} is integrated forward to the time of the next available observations t_i to compute the forecast state $\mathbf{X}^f$,

$$\mathbf{X}^f(t_i) = \mathbf{M}(t_{i-1}, t_i)\mathbf{X}^a(t_{i-1}) \qquad (1),$$

where $\mathbf{M}$ denotes the nonlinear dynamic model operator that integrates a model state from time t_{i-1} to time t_i . The superscript 'f' and 'a' denote the forecast and analysis. The corresponding error covariance matrix can be expressed as:

$$\mathbf{P}^f(t_i) = \mathbf{L}_i[(r+1)\mathbf{T}^T\mathbf{T}]^{-1}\mathbf{L}_i^T + \mathbf{Q}_i \qquad (2),$$

$$\mathbf{L}_i = \mathbf{X}^f(t_i)\mathbf{T} \qquad (3),$$

with $\mathbf{Q}_i$ being the covariance matrix of model uncertainties and $r+1$ is the minimum number of sample ensemble members for error covariance matrix. The superscript 'T' denotes the transpose of matrix. The full rank matrix $\mathbf{T}$ has a dimension of $(r+1)\times r$ with zero column sums and $\mathbf{L}$ is a full rank $(r+1)\times r$ matrix which implicitly represents the model variability.

2. Correction: when the observation is available at time t_i, the LSEIK filter merged the information from model and observation to produce the analysis state with the formula:

$$\mathbf{X}^a(t_i) = \mathbf{X}^f(t_i) + \mathbf{K}_i[\mathbf{Y}^o(t_i) - \mathrm{H}_i\mathbf{X}^f(t_i)] \qquad (4).$$

Here $\mathbf{Y}^o$ is a vector of observations. The gain matrix $\mathbf{K}$, which linearly interpolates between the observations and the forecast, is given by

$$\mathbf{K}_i = \mathbf{P}_i^f \mathrm{H}_i^T\left(\mathrm{H}_i\mathbf{P}_i^f\mathrm{H}_i^T + \mathbf{R}_i\right)^{-1} = \mathbf{L}_i\mathbf{U}_i(\mathrm{H}_i\mathbf{L}_i)^T\mathbf{R}_i^{-1} \qquad (5),$$

where $\mathbf{H}_i$ denotes the linearization of observation operator, which mapping the model space to the observation space. $\mathbf{R}$ is the observation error covariance matrix. The matrix $\mathbf{U}_i$ is updated according to

$$\mathbf{U}_i^{-1} = [\mathbf{U}_{i-1} + (\mathbf{L}_i^T\mathbf{L}_i)^{-1}\mathbf{L}_i^T\mathbf{Q}_i\mathbf{L}_i(\mathbf{L}_i^T\mathbf{L}_i)^{-1}]^{-1} + \mathbf{L}_i^T\mathrm{H}_i^T\mathbf{R}_i^{-1}\mathrm{H}_i\mathbf{L}_i \qquad (6).$$

Localization was used to remove the unrealistic long-range correlation with a quasi-Gaussian function and a uniform horizontal correlation scale (Liu et al. 2013). It was performed by neglecting observations that were beyond correlation distance from an analyzed grid point. In other words, only data located in the "neighborhood" of an analyzed grid point should contribute to the analysis at this point(Liu et al. 2009; Janjić et al. 2011).

A second-order exact sampling is used to initialize the LSEIK filter. At time t_{i-1}, a analysis state $\mathbf{X}^a(t_{i-1})$ and its corresponding error covariance matrix $\mathbf{P}^a(t_{i-1})$, in the factorized form $\mathbf{L}_{i-1}\mathbf{U}_{i-1}\mathbf{L}_{i-1}^T$, are available. The samples can be given by the following formular:

$$\mathbf{X}_k^a(t_{i-1}) = \overline{\mathbf{X}^a}(t_{i-1}) + \sqrt{r+1}\,\mathbf{L}_{i-1}\left(\mathbf{\Omega}_{k,i-1}\mathbf{C}_{i-1}\right)^T \qquad (7).$$

For $1 \leq k \leq r+1$, the $\mathbf{C}_{i-1}$ is the Cholesky decomposition of $\mathbf{U}_{i-1}^{-1}$ and Ω_{i-1} is a $(r+1)\times r$ matrix with orthonormal columns and zero column sums, where $\mathbf{\Omega}_{k,i-1}$ denotes the k^{th} row of $\mathbf{\Omega}_{i-1}$. $\overline{\mathbf{X}^a}$ is the average of the analysis state.

3. Observations

3.1 Satellite observations

The satellite SST used in DA was provided by OSISAF (http://osisaf.met.no/p/sst/index.html). OSISAF ~~products are using in priority the European Meteorological satellites METEOSAT and MetOp and also several American satellites operated by NOAA, DMSP and NASA. Its~~ aim is to produce, control and distribute operationally in near real-time products using available satellite data. The satellite datasets product used here includes the observations from polar orbiting satellites (the EUMETSAT MetOp-A and NOAA-18, -19) with the AVHRR instrument. The SST product has a resolution of 5 km and is produced twice daily at 00 UTC and 12 UTC. It covers the Atlantic Ocean from 50°N to 90°N. The SST observations are thermal infrared observations from the AVHRR instrument and are therefore limited by cloud cover (Kilpatrick et al. 2001). The cloud mask in use is based on a multi-spectral thresholding algorithm by SMHI. The products were retrieved using a nonlinear split window algorithm (Walton et al. 1998). The coefficients in the retrieval algorithm are determined through regression toward in situ observations, and the dataset thus represents the subskin temperature of the oceans. Further, subskin observations are subject to diurnal warming effects, which can be significant in the Baltic Sea. Here only the subskin SST at night (00 UTC), which is comparable to in situ (buoy) measurement, is used to minimum this effect. The SST is controlled with the climatology check. A quality level from 0 to 5 is associated with every pixel. The higher level value, the better the quality of the observations (Brisson et al., 2001). Observations with quality level 4 (good) or 5 (excellent) are collected for the analysis and low quality observations were removed. By applying the above quality control processes, only a subset of the original OSISAF products is kept in this study. Based on the former validation, a bias value of 0.5^{o}C is given for this product.

Further, the IceMap from a sea ice concentration dataset with a high spatial resolution of 5 km (http://www.smhi.se/oceanografi/iceservice/is_prod_en.php) is used to validate the DA results. It is produced by SMHI and originates from digitized ice charts. An advantage of this data is that the ice charts are quality checked manually. However, the drawback is that they include some subjective steps. The temporal resolution of the IceMap SST is twice a week in the experiment period. Sea ice occurs most frequently in the Bay of Bothnia, with up to 100 ice covered days per year. However, sea ice can occur in all parts of the Baltic Sea and Danish straits, demonstrating the need for careful treatment of sea ice in the SST analysis.

3.2 In situ data

The observations from the German Maritime and Hydrographic Agency (BSH) moored buoy stations were collected as independent dataset to validate the assimilation results. The observations have high temporal resolution and long continuous record. The second dataset was downloaded from the Swedish Oceanographic Data Centre -SHARK database (http://sharkweb.smhi.se). SHARK mainly contains low-resolution CTD data from a list of predefined standard stations in the Baltic Sea, as well as in Kattegat and Skagerrak. Only observations that have passed gross quality control procedures are collected into the SHARK database. This procedure includes, for example, location checks and local stability checks. In addition, validating data records from tide gauges are also used. The sea level anomaly measurements from tide gauges (sea level stations) are measured in a local height system and values are presented relative to theoretical mean sea level, a level calculated from many years of annual means, which takes into account the effect of land uplift and sea level rise. The values are averaged over one hour period.

Not all the available observations from satellite, moored buoys, CTDs, tide gauges were included in this study. To obtain the high assimilation quality results, another quality control was applied for these data before they were used into assimilation and validation. These controls include examination of forecast observation differences by excluding those observations for which the difference between the forecast and the measurement exceeded given standard maximum deviations. The criteria were set up empirically based on past validation results of the model (Liu et al. 2013). Furthermore, stations located on land, according to the NEMO-Nordic grid, were excluded. We also removed the duplicate records of these data.

The accuracy of observation error is difficult to be defined for all water points. ~~The observation error mainly comes from the observation instrument itself, the observation representativeness, the temporary reading error and imperfect retrieval algorithm.~~ The observation is commonly assumed to be spatially irrelevant, which results in an error covariance matrix that is time-invariant diagonal and its diagonal elements equal the variance of observation error. In this study, the observation error was estimated to one value as the sum of all observation uncertainties used in the analysis. Besides, the uncertainties of satellite SST varies from coast to the open sea, i.e. higher uncertainties in the coast region relative to the open sea. We used a constant standard deviation value of 0.4^{o}C based on the standard deviation of satellite SST, which ranged from the ~0.1^{o}C to ~0.5 ^{o}C in the Baltic Sea (She et al. 2007, Høyer et al. 2016).

4. Configuration of LSEIK in the experiment

As above mentioned, the initialization of the filter requires an initial analyzed state and a low rank approximation of the corresponding estimation of error covariance matrix. The data assimilation process was initialized by a free model simulation. First the model was spinning up 20 years to reach a statistically steady state. Then a further (free-run) integration covered the period 2006-2009 was carried out to generate a historical sequence of model state. To reduce the calculation cost, we took a snapshot in every 6 days and saved 183 state vectors, which includes sea level, temperature and salinity, in total to describe the model variability because successive states are quite similar. The initial ensemble provided an estimate of the initial model state and its uncertainty before the assimilation of SST observations. The quantity of the model variability was expected to be reasonably comparable with the forecast error, which was dominated by misplacement of mesoscale features and varies in location and intensity seasonally. Further, the very high frequencies of model variability were also unfavourable in an ensemble of state vectors for SST data assimilation (Oke et al., 2005). Therefore, a band-pass filter was used to remove the unwanted frequency of model variability. To initial low rank error covariance matrix, a multivariable Empirical Orthogonal Functions (EOF) analysis was applied on the 183 state vectors of model variables (sea level, temperature and salinity). In the North Sea and Baltic Sea, error covariances of different variables are not uniform and strongly dependent on whether the variable resides in the open sea or coastal zone. Each state variable was then normalized by the inverse of its spatially averaged variance at every model level. At last, 34 leading EOF modes were kept and they explained 85% overall variability. Then the initial error covariance matrix was estimated by $P^a(t_0) \approx L_0 U_0 L_0^T$, where the L_0 is composited by the leading EOF modes and U_0 is diagonal matrix with the corresponding eigenvalues on its diagonal. We used a time-invariant sample ensemble to approximate the background error covariance during the experimental period (Korres et al, 2004; Liu et al. 2017). This stationary ensemble affords a good approximation of the ocean's background error covariance. Meanwhile, it is computationally efficient for our objective.

A forgetting factor ρ was introduced to parameterize the imperfect model by amplifying the already existing modes of the background error (Nerger et al, 2006). The matrix U_i was calculated by

$$\mathbf{U}_i^{-1} = \rho(r+1)\mathbf{T}^T\mathbf{T} + \mathbf{L}_i^T\mathbf{H}_i^T\mathbf{R}_i^{-1}\mathbf{H}_i\mathbf{L}_i \qquad (8).$$

The localization scale is another import factor to the assimilation system, especially at the coastal region. Large correlation scale may transfer artificial increments to the positions far away from the analysis observation during the DA process. However, small correlation scale is prone to cause the singularity of ocean state around analyzed observation and break the continuity of the ocean state. Hence, an unreasonable scale causes the instability of the model integration or degrades the assimilation quality. Unfortunately, the accuracy length for the correlation is unknown for the North Sea and Baltic Sea. The correlation length scale is to some extent dependent on the Rossby radius of deformation (Losa et al., 2012), which varies from ~ 200 km in the barotropic mode to ~ 10 km or even less in the baroclinic mode (Fennel et al., 1991; Alenius et al, 2003). According to the former researches like Liu et al. (2013, 2017), a length scale of 70 km was specified for both the North Sea and Baltic Sea in this study. Not that this value may be not perfect and more accurate correlation length needs to be tested for LSEIK. For example, spatially variable length scales are the next step for the regional DA simulations.

To define the forgetting factor, a one-month simulation experiment with varying the factor ρ was done in January 2010. At last, a factor $\rho = 0.3$ resulted in the best assimilation performance. Further, we define a two-day assimilation window in assimilation experiment. As a result, the observations in the two days before the assimilation time were used to calculate the innovation with observation operator. When we calculated the innovation we also changed the observation error according to the observation time by

$$\varepsilon = 0.4 \times \exp(-0.15\Delta t) \qquad (9),$$

here Δt is the absolute time difference between observation time and DA time.

5. Results

In the following sub-sections, we conducted two runs with and without assimilation of the SST observations from the OSISAF database, both runs with the above setup of the analysis system.

Accordingly, the runs with and without assimilation are called ASSIM and FREE, respectively. We considered the evolution of SST based on 48-hourly local analysis from 1 January 2010 to 31 December 2010. The 48-hourly forecast SST from two runs was assessed with observations from different dataset. Then we analyzed the impact of the data assimilation on the profile simulation of T/S. At last, we evaluated the system performance with respect to sea surface anormaly and sea ice, respectively.

5.1 Comparison with satellite data

First, we presented two cases to show the ocean state before and after the assimilation of the OSISAF SST data in Fig. 2. The first case was given at 11 January 2010, a date with clear weather and many observations available. The model has obvious difficulties in reproducing the observed SST. The cold biases in the forecast were found in the Skagerrak, west coast of the Baltic proper and the Bothnian Bay, respectively. However, the warm biases appeared in the interior of the Baltic Sea and the Kattegat. The largest deviation in the FREE reached 2.2 ^{o}C at the Skagerrak. Apparently, temperature by assimilation analysis agreed with the satellite-derived data much better. This correction at the analysis step has allowed us to reduce the deviation of the SST forecast from the observations. ~~The SST bias of model forecast possibility has seasonal variability because of the errors in the forcing and/or heat flux parameterization used in the ocean model (Fu et al. 2012).~~ The DA system simulation was also verified at 2 June 2010, which has also many available OSISAF observations. The biases on 2 June 2010 were obviously different from that on 11 January 2010. Moreover, it was found they had a roughly opposite bias signal. For example, relative to the OSISAF SST at the Baltic proper, Bothnian Sea and Bothnian Bay, FREE produced relatively warmer water at January 11 and colder water at 2 June (Fig. 2), respectively. After data assimilation, the analysis increments were appropriately added to the model field. In general, the SST DA has improved the simulated SST in both cases (Fig. 2).

Maps of annual averaged RMSE of SST from two runs relative to the IceMap observation are shown in Fig. 3. Obviously, the RMSE in FREE and ASSIM had different distribution in the Baltic Sea. In general, FREE had smaller error in the Skagerrak, eastern the Kattegat and the interior of the Bothnian Sea relative to other subbasin of the Baltic Sea. The largest RMSE was found at the connection region between the Baltic proper and the Bothnian Sea. This could be caused by the shallow water, complicated bathymetry and large observation biases in this area. It was also noted that the RMSE was larger in the coast region compared to its interior in the Baltic proper and Bothnian Sea. After the assimilation, the SST has been significantly improved. The RMSE of SST from ASSIM was generally smaller than 1.0 °C. However, there were still some regions where the improvements were relatively small and the RMSE of SST was greater than 1.0 °C. These large errors were predominantly located at the edge of the Baltic Sea and the Danish straits. For instance, the RMSE of SST was greater than 1.2 °C at both the entrance of the Gulf of Finland and the west coast of the Bothnian Sea. The relatively small improvements were regularly caused by the rare observations or the less accurate observations near the coast water.

The overall daily averaged SST errors against the IceMap observations have been estimated (Fig. 4). The observations had better coverage in summer and autumn than in winter and spring. The variability of the number of observation directly affected the assessment of DA results. The model biases had pronounced seasonal variability, which had small values in spring and winter. In general, the assimilation provided better SST estimations. The free run had a RMSE of 1.47 °C. After the assimilation, the RMSE was reduced to 1.03 °C, whereas the bias was reduced by 0.73 °C. An interesting feature was that the SST error reduction due to the assimilation was almost consistent with the variability of the number of IceMap observations. For example, the improvement became large with increasing the number of IceMap observations from March to June 2010. However, the number of observations was kept constant during the period June-November 2010 and the improvement shown in both the bias and RMSE of SST did not exhibit large variability, which meant reliable performance of the DA system.

5.2 Comparison with independent in-situ data

The time series of T/S were compared with independent observations located at Arkona station (13.87°E, 54.88°N) in the Arkona Basin and at BY15 (20.05 °E, 57.33 °N) in the Eastern Gotland Basin, respectively. These two stations were selected to verify the experiment results because of their relatively completed observation records for the experiment period. In the Arkona Basin, the water depth was shallow and the water column can be well mixed between surface and bottom water. Thus, the bottom T/S was largely affected by the surface dynamic (Liu et al. 2014). Relative to observations, the model had warm biases at this station (Fig. 5). ~~The temperatures differ by about 15–22 °C between summer and winter.~~ At a depth of 25m, the observed temperature showed the largest variability, which was a good representation of the bottom characteristics of the mixed layer. In mid-August, the temperature was abruptly increased by 10°C at a depth of 25m and slightly decreased at surface, respectively. The reason is that the surface water suddenly sinks to deeper layers, which warm the deep water. However, this dynamic process hasn't reached to Arkona bottom and it didn't cause the obvious bottom temperature variability (Fig. 9). Both FREE and ASSIM had reproduced this process, whereas FREE showed larger temperature biases. To the salinity at the Arkona station, the surface observations were missing, the comparison at 7 m depth verified the subsurface simulations. The observations showed larger salinity variability in winter relative to summer. This pronounced seasonal variation is associated with the variation of fresh river runoff and net E–P (Evaporation–Precipitation) flux (Fu et al, 2012). At a depth of 7 m, salinity was obviously underestimated from April to September and overestimated after November although the ASSIM had slightly better results compared to FREE. The DA also provided better simulation of salinity at 25 m depth. For example, the salinity bias in the October was reduced by 3 psu by DA. At a depth of 40 m, the saltwater inflows were observed, resulting in sudden increases of salinity. For instance, the salinity was increased by 3.5 psu in February followed by a decreasing trend. The variations were reproduced in both FREE and ASSIM, whereas the intensity of the decreased process is weakly simulated with a difference of 3 psu and the inflow in March was not strong enough relative to the observed one. Observations also showed a large salinity variability amounts to 4–8 psu in the autumn. Although FREE and ASSIM had shown these changes, their magnitude was obvious weaker than observations. The possible reason was that the model's resolution was inadequate to well resolve the topography and eddies in this area. Both the large runoff and the complicated bathymetry posed challenges for the model to tackle the small-scale dynamic process in such a shallow basin. A higher resolution model perhaps was more preferable to study this dynamic process.

The Eastern Gotland Basin has deeper water depth compared to the Arkoan Basin, in which the water column is permanently stratified and the halocline lies at about 60–80 m (Fu et al, 2012). The mixing and sinking of T/S are hindered by the strong stratification. Unlike observations in the Arkona Basin (Fig. 5), the CTD observations at BY15 had lower temporal resolution with almost one observation per month. In the mixing layer, it can be seen model had overestimated the temperature (Fig. 6). At a depth of 10 m, ASSIM has remarkably improved the simulation of temperature relative to FREE. The bias has been reduced by 3^{o}C in the spring of 2010. At 175 m depth, observed temperature showed very small variation. The reason was that the main source for deep water ventilation is the saltwater inflows which are suppressed by runoff within a depth range of 75–135 m in the Eastern Gotland Basin (Vali et al. 2013). As a result, updating the bottom water is very slow. Both FREE and ASSIM overestimated the temperature in the spring and the beginning of summer of 2010. Further, ASSIM has increased the temperature bias after mid-summer relative to FREE. This result might be explained by that the strong correlation isn't expected between surface and layers bellow the halocline because of the strong stratification in this basin, which perhaps yield the artificial correction. Therefore, the improvement of the surface temperature cannot guarantee its positive influence on the bottom temperature. To the salinity, the model had less accurate simulation with generally low salinity biases at 10 m depth. ASSIM provided better salinity simulation compared to FREE. At 70 m depth, the small variation of salinity was found after DA. Moreover, at 175 m depth, the observation had very small variability about 0.1 psu. In general, both experiments have reproduced these variations. However, FREE increased salinity by 0.2 psu from March to April relative to the observation, which caused the overall salinity overestimated amount to 0.2 psu. This increasing process wasn't shown in observations and the reason remained unclear. The DA has shown slight improvement, but it still salter than the observations.

The mixed layer depth (MLD) was calculated at the Arkona and BY15 station and compared with the SHARK observation in Fig. 7. We used the temperature criterion to define the MLD, i.e., the depth at which the temperature deviated from the surface value by 0.5 ^{o}C (Fu et al., 2012). Figure 7 shows that the MLD at Arkona had larger variability relative to the MLD at BY15. The reason contributed to this feature is that the deeper water at Arkona is easy affected by wind forcing because of the shallow bathymetry and well mixing, whereas the temperature variation in upper water at BY15 difficulty influences the deeper water because of the strong stratification. Both runs had reproduced the MLD variability feature similar as the observations. For example, the minimum MLD appeared in summer, which was about several meters. The assimilation of satellite SST caused strong changes in the MLD at both stations, especially in winter. One explanation was that the Baltic Sea was largely affected by wind forcing and the winter wind was much stronger than the summer wind. Further, strong heating in summer promoted stratification in summer and shoaled the MLD.

Further, the temporal and spatial distribution of the SHARK observations is shown in Fig.8. These observations were unevenly distributed in the Baltic Sea. In the Skagerrak, the observations appeared at the Danish and Swedish coast. However, in the Bornholmn Basin, Kattegat, and Baltic proper, the observations mainly were found in the central and the Swedish coast side. There were also many observations in the Bothnian Sea and rare observations in the central of the Bothnian Bay. It must be noticed that there aren't SHARK observations in both the Gulf of Finland and Gulf of Riga during the experiment period. Moreover, these SHARK profiles in the first four months were mainly located from the Skagerrak to the Baltic proper, which are relatively rare in the northern Baltic Sea. In the Bothnian Bay, the observations are mainly in the winter period.

Figure 9 shows the change of overall bias and RMSE of T/S with depth against the SHARK dataset. In the Baltic Sea, DA had large impact on the temperature forecast in the water above 100 m. The RMSE showed that the forecast of temperature was obviously improved from surface to thermocline in the ASSIM and the improvements generally decreased with depth. Above 100 m, the overall RMSE of temperature in ASSIM was decreased by 21.38% (from 1.59 to 1.25 ^{o}C). It was also found the temperature error had similar variability as the warm biases in two runs. In the transition zone, the RMSE in the ASSIM was reduced by 5.59% and -20.31% above and below 100 m relative to the FREE, respectively. Below 90 m, the temperature was also over-adjusted, which changed the warm bias to cold bias. It is worth noting that the number of the deeper water observation in the transition zone is substantially less than that in the Baltic Sea. For the salinity, both RMSE and bias of the ASSIM showed very minor changes relative to the FREE inside the Baltic Sea. For the water above 100 m, the total RMSE of salinity was increased by 3.48% (from 1.15 psu in the FREE to 1.19 psu in the ASSIM) in the transition zone and 1.04% (from 0.96 psu in the FREE to 0.97 psu in the ASSIM) in the Baltic Sea.

5.3 Sea Level Anomaly

SLA represents a vertically integrated effect of the T/S variations over the whole water column. The accurate simulation of SLA is thus a good indicator of the model performance. Therefore, validating the impact of SST assimilation on the simulation of SLA is very important to the Baltic Sea forecast. The observations from the 24 tide gauge stations were used. These gauge stations are mainly located at the Swedish coast (see Fig.8b). Since only the SST is assimilated in this study, the SLA observations are completely independent.

We calculated the RMSE and correlation coefficients for both the FREE and ASSIM against the observations from tide gauges (Fig. 10). The overall RMSE was reduced by 1.8% and the correlation coefficients were slightly increased. Among these stations, RMSE at the Oskarshamn was decreased by 5.6%, which is larger than that in other station. The minimum RMSE change of SLA was seen at the Klagshamn. For the correlation coefficient, improvement on the SLA by the DA is very small. Simrishamn station showed the biggest change of correlation coefficient, which is 1.1%. The RMSE and correlation comparison demonstrated that the SST DA has generally positive effects on the forecast of the SLA.

In addition, the time series of the SLA error discrepancy (ASSIM minus FREE) in two runs at four stations were selected to evaluate the simulation results (Fig. 11). These four stations were selected to represent the model performance at different positions of the Swedish coast. Two runs showed evidently different performance in these four stations. The variability of the SLA difference between two experiments at the Smogen station had higher frequency compared to other stations. The reason was that the Smogen station was located at the transition zone where the water had higher frequency variations caused by the brackish Baltic in/outflowing relative to other three stations. At these four stations, the improvements were mainly in later spring and summer, whilst the degraded simulations were mostly happened after Mid-September, respectively. The SST assimilation had less impact in late winter and early spring compared to other seasons. Besides, the impact of SST assimilation on SLA simulation was not same in the four positions. For instance, during the period from Mid-November to Mid-December, the SLA in ASSIM was improved at Simrishamn and degraded at both Ratan and

LandsortNorra stations, respectively. This phenomenon was possibly caused by the imperfect correlation between SST and SLA in the stationary samples. Further, these steric small changes of SLA by DA were what we expected because only SST was assimilated into Nemo-Nordic.

5.4 Sea ice

Sea ice in the Baltic Sea occurs primarily in its north region and influences the Baltic climate. Accurate detecting the sea ice is very useful to the northern Baltic living because too much or too little sea ice can be a problem for wildlife and people. Sea ice concentration (SIC) and Sea ice extent (SIE) are two important and common indicator to modeling sea ice environment. We assessed the SIC and SIE from simulations against the IceMap observations in Fig. 12-13. Differ from the daily evaluation in Losa et al. (2014), the monthly mean SIC was used to represent the general status of sea ice in the Baltic Sea. Besides, SIC in January, February and December showed the variation of the sea ice in winter.

In January 2010, the observations showed large ice coverage in the Bothnian Bay and the Gulf of Finland and small SIC in the Gulf of Riga, respectively. Model generally reproduced this distribution of sea ice. However, FREE simulated too much sea ice in the Gulf of Finland and the eastern coast of the Baltic proper relative to observations. For example, SIC from FREE almost to 30% higher than observations along the Estonia coastline. It could be seen that the SST DA reduced these biases. The reason is the SST DA modified the thermal expansion by providing the well temperature fields above the thermocline. The temperature in February became colder relative to January in the Baltic Sea. As a result, the sea ice in February extended to the Bothnian Sea and the whole Gulf of Riga. Observation also showed small SIC in Kattegat and Skagerrak. Model simulated higher SIC in the Bothnian Sea with largest biases along the Swedish and Finnish coast. As an example, the observed ice in the Bothnian Sea was characterized by concentrations mainly smaller than 0.5, whereas modeled ice in FREE had concentration greater than 0.9 in the shallow region of the Bothnian Sea. FREE also had smaller ice coverage with lower SIC in the transition zone between the North Sea and the Baltic Sea relative to IceMap. After the SST assimilation, ASSIM reduced SIC in the Bothnian Bay and the west coast of the Baltic Sea, which was closer to the observations. The ice in ASSIM didn't have obvious variation in Kattegat and Skagerrak yet. ASSIM also reduced too much ice at the southern of the Bothhomn Basin. The reason is that the satellite SST observations had limited accuracy near the coast and they could bring artificial information into the modeling. In March, compared to observation, the FREE produced low SIC in the western coast of the Bothnian Sea, Gulf of Finland, Gulf of Riga and the connect zone between the Bothnian Sea and Gulf of Finland. However, the model SIC in the FREE was higher than IceMap in the interior the Bothnian Bay. For instance, the SIC from FREE in the western Bothnian Sea was 40% higher than observation. In the south coast of the Arkona basin and Baltic proper, the FREE failed to reproduce the sea ice as in observation. After the DA, the high SIC was decreased in western Bothnian Sea and closer to that in IceMap in Bothnian Sea. In the Gulf of Finland and Gulf of Riga, the SIC error was increased in the ASSIM. In April, the large SIC error in the FREE was shown in the Bothnian Sea, the Bothnian Bay, Gulf of Rig and Gulf of Finland, where no clear improvements were seen in the ASSIM. In December, sea ice coverage was smaller because of relatively warm temperature compared to that in other winter month. Most of the sea ice with high concentration was observed at the edge of the Bothnian bay. Nevertheless, high concentration ice in FREE also happened at the transition zone between the Bothnian Sea and Bothnian bay. Relatively, ASSIM reduced the high concentration biases of sea ice. By contrast, both ASSIM and FREE had lower concentration ice than observation in the eastern coast of the Bothnian Sea. The SIC from ASSIM was relatively lower than that from FREE in the northern Finish coast, whereas the observations had high concentration ice there.

The daily SIE from FREE and ASSIM was compared with observations in Fig.13. The observed SIE was generally increased from January to February and reached the maximum in mid-February. During the period of March-May, SIE was decreased as temperature was increasing. SIEs in both the FREE and ASSIM experiments were generally underestimated by comparison with the observation in 2010, especially in the period from Mid-March to early April. The SIE bias in both runs was roughly increased from January to early April. In early April, the maximum negative bias of SIE was found to be 105000 km^2 for ASSIM and 10000 km^2 for FREE. The impact of SST assimilation on the SIE was positive during the phase of sea ice formation. For example, the SIE bias was reduced 25000 km^2 at end of February and in the Mid-December. However, during the phase of sea ice melting (March to

April), the SIE error was increased in ASSIM even with the error of SST decreased. For example, the SIE bias in ASSIM was increased by 42000 km^2 relative to FREE in the early March. These increased SIE error in March mainly happened in the Gulf of Riga and Gulf of Finland (Fig.11).

6. Conclusion and discussions

A DA system based on a LSEIK filter has been coupled to the NEMO circulation model of the North and Baltic Seas. The method was successfully applied for assimilating high resolution satellite SST data. We demonstrated that, over the period of 2010, the agreement of the SST forecast with the independent satellite observation was improved by ~ 29.93% in comparison with the regular forecast without DA. The assimilation quality is directly related to the number of observation.

Compared with independent in-situ data from SHARK, the RMSE of temperature was reduced by 21.38% and 5.59% for the water above 100 m inside and outside of the Baltic Sea, respectively. However, in the deeper layers, the temperature was slightly degraded in the Baltic Sea. This is partially caused by the artificial correlation between surface layer and deeper layers. The improvement of temperature by SST DA can't guarantee corresponding improvement of the salinity. The statistics displays the salinity RMSE was increased by 1.04% and 3.48% in the transition zone and the Baltic Sea, respectively. Both ASSIM and FREE have captured the main dynamic process in the Baltic Sea, for example, the inflow and the sink. However, ASSIM is closer to the observed one relative to FREE.

The forecast results were further validated with the independent SLA observations. The result shows that all RMSEs and correlations for all 21 stations are smaller than 0.12 m and greater than 0.86, respectively. After DA, the SLAs at these stations have been slightly improved. In general, the RMSE was reduced by 1.8% and correlation coefficients were slightly increased, respectively. Further, the model-observation comparison at selected four stations indicates that these improvements are mainly in later of spring and summer. The comparisons also denote the SST assimilation has less impact in the late winter and early spring relative to other seasons.

When compared with monthly mean observations of SIC, both assimilation run and free run reproduced main spatial distributions of sea ice in the Baltic Sea. During the sea ice formation period, the SST assimilation has improved the results of SIC from FREE in the Gulf of Finland, the Bothnian Sea and eastern coast of the Baltic proper. However, minor improvements were found in Kattegat and Skagerrak. Besides, over the sea ice melting period, the SIE comparison showed the SST assimilation increased the SIE error, especially in the Gulf of Finland and Gulf of Riga.

The daily MLD from two runs has been compared with the observations at Arkona and BY15 stations. Model could capture the variability features of the MLD. Similar as Fu et al.(2012), it was found that SST assimilation had less impact on the MLD in summer than that in winter. In general, the SST DA produced less influences on the MLD in the deeper region (BY15) relative to that in the shallow region (Arkona).

Further, the reliability of the DA system is worth being assessed. In Rodwell et al.(2006), a perfect reliable system error variance for ensemble assimilation was calculated by the sum of the variance of the sample ensemble, the square of innovation(misfit between observation and model) and the variance of observation at assimilation time. In this study, we used a constant observation error similar to Rodwell et al. (2016) because our DA design is different from that paper. The major difference between these two studies is that we estimate the background error covariance from stationary ensemble and avoid the perturbation of observation error. Therefore, the variance of the sample ensemble and observation is univariate and the diagnostic of the assimilation stability can be directly obtained from the forecast error like the RMSE in Fig.4.

The results of the SST assimilation are encouraging and the assimilation helps to ameliorate some model deficiencies such as the simulation of sea ice in the Gulf of Finland. However, some problems need to be further addressed in the SST DA in the future: firstly, the SST assimilation has worse influence on the simulation of salinity in the upper layers and temperature in the deeper layers. Losa et al.(2012) denoted that the salinity simulation quality crucially depends on the assumptions about the model and data error statistics. Here a stationary ensemble sample was used to represent the correlation between T/S and between surface and deep water. These relationships could be changed with the varying dynamics and forcing conditions. More sophisticated assumption should be used in the DA of Baltic Sea. Secondly, the SHARK observations in this study are absent at the Gulf of Finland and Gulf of Riga. This denotes the validation results with SHARK observation didn't include the evaluation of the simulation of T/S in deep water of these two basins. Thirdly, the univariate localization scale used in this study could be another problem. The spreading of observation information strongly depended on the correlation scale~~ around the observation position~~. The large localization scale can introduce the artificial information, which could degrade the assimilation quality. A flow-dependent background error covariance with varying correlation scale may be more appropriate for the Baltic Sea with complex bathymetry and rich dynamics. Fourthly, the remote sensing observations near the coast could have large bias because of the limit of the instrument itself. More strict quality controlling method needed to be used for the satellite coastal observations before their assimilation.

Acknowledgment

The research presented in this study was funded by the Swedish Space Board within the project 'Assimilating SLA and SST in an operational ocean forecasting mode for the North Sea and Baltic Sea using satellite observations and different methodologies' (grant no.172/13).

References

Adcroft, A., and Campin, J. M.: Re-scaled height coordinates for accurate representation of free-surface flows in ocean circulation model, Ocean Modell., 7, 269–284, 2004.

Alenius, P. A., Nekrasov, A., and Myrberg, K.: Variability of the baroclinic Rossby radius in the Gulf of Finland, Cont. Shelf Res., 23 (6), 563–573, 2003.

Beckmann, A., and Döscher, R.: A method for improved representation of dense water spreading over topography in geopotential-coordinate models, J. Phys. Oceanogr., 27, 581–591, 1997.

Brisson, A., Le Borgne, P., and Marsouin, A.: Results of one year of preoperational production of sea surface temperatures from GOES-8, J. Atmos. Oceanic Technol., 19(10), 1638–1652, 2002.

Dahlgren, P., Kållberg, P., Landelius, T. and Undén, P.,: EURO4M Project Report, D 2.9 Comparison of the Regional Reanalyses Products with Newly Developed and Existing State-of-the Art Systems. Technical Report, Online at: http://www.euro4m.eu/Deliverables.htm, 2014.

Donnelly, C., Andersson, J. C., and Arheimer, B.: Using flow signatures and catchment similarities to evaluate the E-HYPE multi-basin model across Europe, Hydrological Sciences Journal, 61, 255–273, 2016.

Egbert, G. D., and Erofeeva, S. Y.: Efficient inverse modeling of barotropic ocean tides, J. Atmos. Oceanic Technol., 19(2), 183–204, doi: 10.1175/1520-0426, 2002.

Fennel, W., Seifert, T., and Kayser, B.: Rossby radii and phase speeds in the Baltic Sea. Cont. Shelf Res., 11(1), 23–26, 1991.

Fu, W.W., She, J., and Dobrynin, M.: A 20-year reanalysis experiment in the Baltic Sea using three-dimensional variational (3DVAR) method. Ocean Sci., 8, 827–844, 2012.

Galperin, B., Kantha, L. H., Hassid, S., and Rosati A.: A quasi-equilibrium turbulent energy model for geophysical flows, J. Atmos. Sci., 45, 55–62, 1988.

Haines, K.: Ocean data assimilation. In: Data Assimilation: Making Sense of Observations. . Springer-Verlag, Berlin Heidelberg, pp. 517-548. ISBN 9783540747024, 2010.

Hordoir, R., Axell, L., Löptien, U., Dietze, H., and Kuznetsov, I.: Influence of sea level rise on the dynamics of salt inflows in the Baltic Sea, J. Geophys. Res. Oceans, 120, doi:10.1002/2014JC010642, 2015.

Hordoir, R., Dieterich, C., Basu, C., Dietze, H., and Meier M.: Freshwater outflow of the Baltic Sea and transport in the Norwegian current: A statistical correlation analysis based on a numerical experiment, Cont. Shelf Res., 64, 1–9, doi:10.1016/j.csr.2013.05.006, 2013.

Høyer J.L., and Karagali, I.: Sea Surface Temperature Climate Data Record for the North Sea and Baltic Sea. JOURNAL OF CLIMATE. 29, 2529–2541, 2016.

Janjić, T., Nerger, L., Albertella, A., Schröter, J., Skachko, S.,. On domain localization in ensemble based Kalman filter algorithms. Monthly Weather Review, 136 (7), 2046–2060, 2011.

Kilpatrick, K. A., Podesta, G. P., and Evans, R.: Overview of the NOAA/NASA Advanced Very High Resolution Radiometer Pathfinder algorithm for sea surface temperature and associated matchup database, J. Geophys. Res., 106(C5), 9179–9197, doi:10.1029/1999JC000065, 2001.

Large, W. G., and Yeager, S.: Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies, NCAR Tech. Note, NCAR/TN-4601STR, CGD Div. of the Natl. Cent. for Atmos. Res., 2004.

Leclair, M., and Madec, G.: A conservative leapfrog time stepping method, Ocean Modell., 30, 88–94, doi:10.1016/j.ocemod.2009.06.006, 2009.

Leppäranta, M., and Myrberg, K.: The Physical Oceanography of the Baltic Sea, pp. 378, Springer-Verlag, Berlin-Heidelberg, New York, 2009.

Levitus, S., and Boyer, T. P.: Salinity, in World Ocean Atlas 1994, NOAA Atlas NESDIS, vol. 3, 99 pp., U.S. Gov. Print. Off., Washington, D. C., 1994.

Liu, Y., Zhu, J., She, J., Zhuang, S. Y., Fu, W.W., and Gao, J.D.: Assimilating temperature and salinity profile observations using an anisotropic recursive filter in a coastal ocean model. Ocean Model. 30, 75–87, 2009.

Liu, Y., Meier, H. E. M., and Axell, L.: Reanalyzing temperature and salinity on decadal time scales using the ensemble optimal interpolation data assimilation method and a 3-D ocean circulation model of the Baltic Sea. J. Geophys. Res.Oceans., 118, 5536–5554, 2013.

Liu, Y., Meier, H. E. M., and Eilola, K.: Improving the multiannual, high-resolution modelling of biogeochemical cycles in the Baltic Sea by using data assimilation, Tellus A, 66, 24908, doi:10.3402/tellusa.v66.24908, 2014.

Liu, Y., Meier, H. E. M., and Eilola, K.: Nutrient transports in the Baltic Sea – results from a 30-year physical–biogeochemical reanalysis. Biogeosciences, 14, 2113–2131, 2017.

Losa S.N., Danilov, S., Schröter, J., Nerger, L., Maβmann, S., and Janssen, F.: Assimilating NOAA SST data into the BSH operational circulation model for the North and Baltic Seas: Inference about the data. Journal of Marine Systems, 105–108,152–162, 2012.

Losa S.N., Danilov, S., Schröter, J., Janjic, J., Nerger, L., and Janssen, F.: Assimilating NOAA SST data into the BSH operational circulation model for the North and Baltic Seas: Part 2. Sensitivity of the forecast's skill to the prior model error statistics. Journal of Marine Systems, 259–270, 2014.

Madec, G.: NEMO ocean engine, version 3.3, Note du Pôole de modélisation de l'Inst. Pierre-Simon Laplace 27, Inst. Pierre-Simon Laplace, Paris. (Available at http://www.nemo-ocean.eu/.), 2010.

Malanotte-Rizzoli, P, and Tziperman, E.: The oceanographic data assimilation problem: overview, motivation and purposes. In Modern Approaches to Data Assimilation in Ocean Modeling, Amsterdam: Elsevier, 3–17, 1996.

Nerger, L., Danilov, S., Hiller, W., and Schröter, J.: Using sea level data to constrain a finite-element primitive-equation ocean model with a local SEIK filter. Ocean Dynamics 56, 634–649, 2006.

Nowicki, A., Dzierzbicka-Głowacka, L., Janecki, M., and Kałas, M.: Assimilation of the satellite SST data in the 3D CEMBS model. Oceanologia, 57, 17–24, 2015.

O'Dea E. J., Arnold, A. K., Edwards, K. P., Furner, R., Hyder, P., Martin, M. J., Siddorn, J. R., Storkey, D., While, J., Holt, J. T., and Liu H.: An operational ocean forecast system incorporating NEMO and SST data assimilation for the tidally driven European North-West shelf, Journal of Operational oceanography, 5(1), 3-17, 2012.

Oke, P. R., Schiller, A., Griffin, D. A., and Brassington, G. B.: Ensemble data assimilation for an eddy-resolving ocean model of the Australian Region. Q. J. Roy. Meteorol. Soc. 131, 3301–3311, 2005.

Omstedt, A., Elken, J., Lehmann,A., Leppäranta, M., Meier, H.E.M., Myrberg, K., and Rutgersson, A.: Progress in physical oceanography of the Baltic Sea during the 2003–2014 period. Progress in Oceanography, 128, 139-171, 2014.

Pham, D.T.: Stochastic methods for sequential data assimilation in strongly nonlinear systems. Mon. Weather Rev. 129, 1194–1207, 2001.

Rodwell, M. J., Lang, S. T. K., Ingleby, N. B., Bormann, N., Hólm, E., Rabier, F., Richardson, D. S., and Yamaguchi, M.: Reliability in ensemble data assimilation, Q. J. Roy. Meteor. Soc., 142, 443–454, doi:10.1002/qj.2663, 2016.

Samuelsson, P., Jones, C., Willen, U., Ullerstig, A., and co-authors.: The Rossby Centre Regional Climate model RCAS3: model description and performance, TellusA, 63, 4–23, 2011.

She, J, Høyer, J. L., and Larsen, J.: Assessment of sea surface temperature observational networks in the Baltic Sea and North Sea. Journal of Marine Systems 65, 314–335, 2007.

Stramska, M., and Białogrodzka, J.: Spatial and temporal variability of sea surface temperature in the Baltic Sea based on 32-years (1982—2013) of satellite data. Oceanologia, 57, 223–235, 2015.

Tranchant B., Reffray, G., Greiner, E., Nugroho, D., Koch-Larrouy, A., and Gaspar, P.: Evaluation of an operational ocean model configuration at 1/12◦ spatial resolution for the Indonesian seas (NEMO2.3/INDO12) –Part 1: Ocean physics. Geosci. Model Dev., 9, 1037–1064, 2016.

Umlauf, L., and Burchard, H.: A generic length-scale equation for geophysical turbulence models, J. Mar. Syst., 61, 235–265, 2003.

Vancoppenolle, M., Fichefet, T., Goosse, H., Bouillon, S., Madec, G., and Maqueda, M. A. M.: Simulating the mass balance and salinity of arctic and Antarctic sea ice, Ocean Modell., 27(1–2), 33–53, doi:10.1016/j.ocemod.2008.10.005, 2008.

Väli, G., Meier, H. E. M., and Elken, J.: Simulated halocline variability in the Baltic Sea and its impact on hypoxia during 1961–2007, J. Geophys. Res.-Ocean., 118, 6982–7000, doi:10.1002/2013JC009192, 2013.

Walton, C. C., Pichel, W. G., Sapper, F. J., and May, D. A.: The development and operational application of nonlinear algorithms for the measurement of sea surface temperatures with NOAA polar-orbiting environmental satellites, J. Geophys. Res., 103(C12), 27,999–28,012, doi:10.1029/98JC02370, 1998.

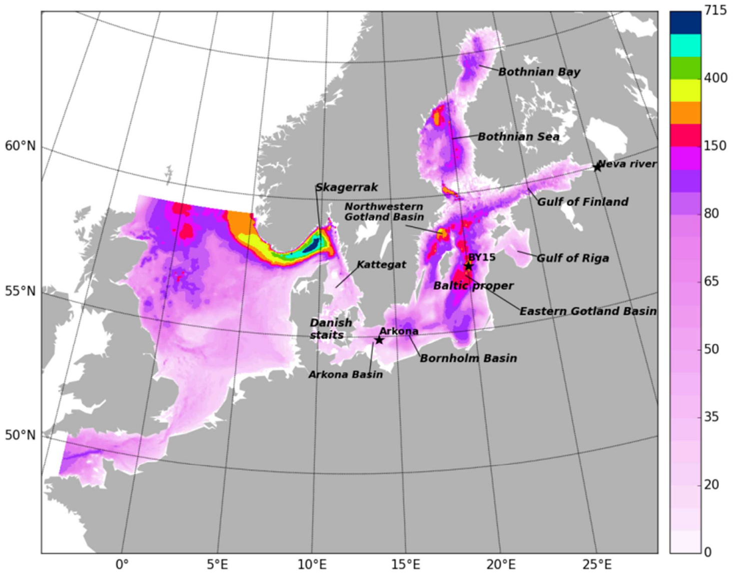

Figure 1. Geographical domain and bathymetry (in m) of the NEMO-Nordic configuration.

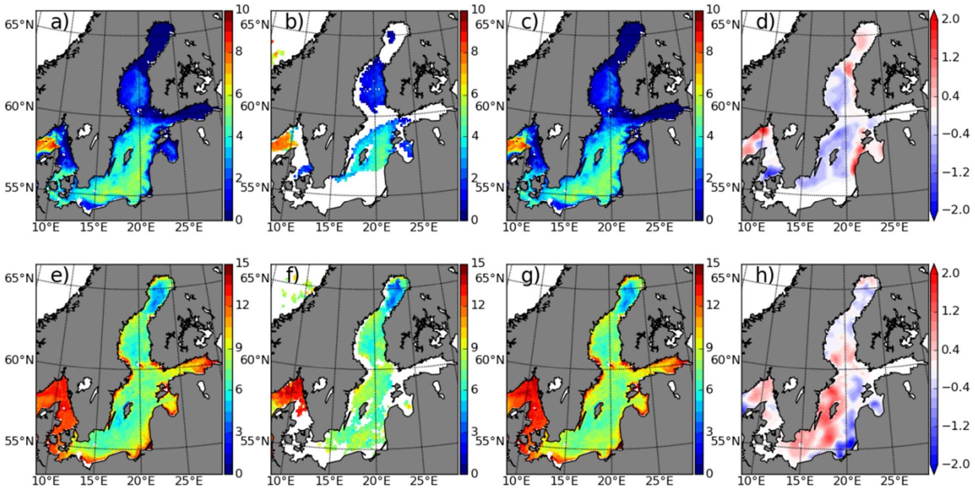

Figure 2. Map of SST from FREE (a,e), OSISAF (b, f), ASSIM (c, g) and the assimilation increments (d, h) on 11 January 2010 (first row) and 2 June 2010 (second row), respectively.

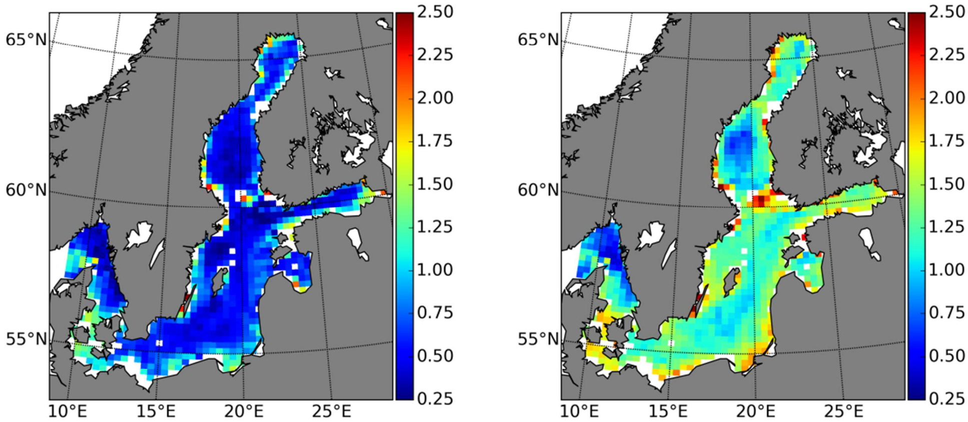

Figure 3. Map of the RMSE of SST from ASSIM (left panel) and FREE (right panel) calculated against IceMap SST in 2010, respectively.

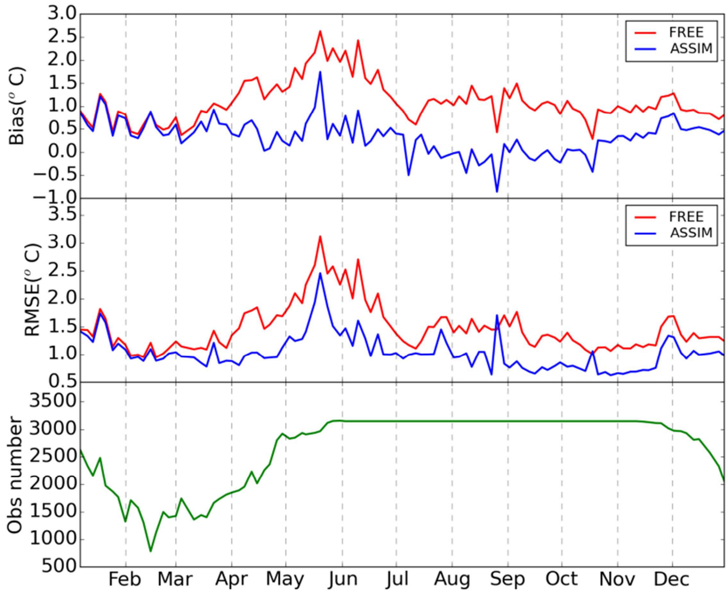

Figure 4. The evolution of basin-averaged bias and RMSE of SST from FREE and ASSIM relative to IceMap SST and the number of IceMap observation in 2010.

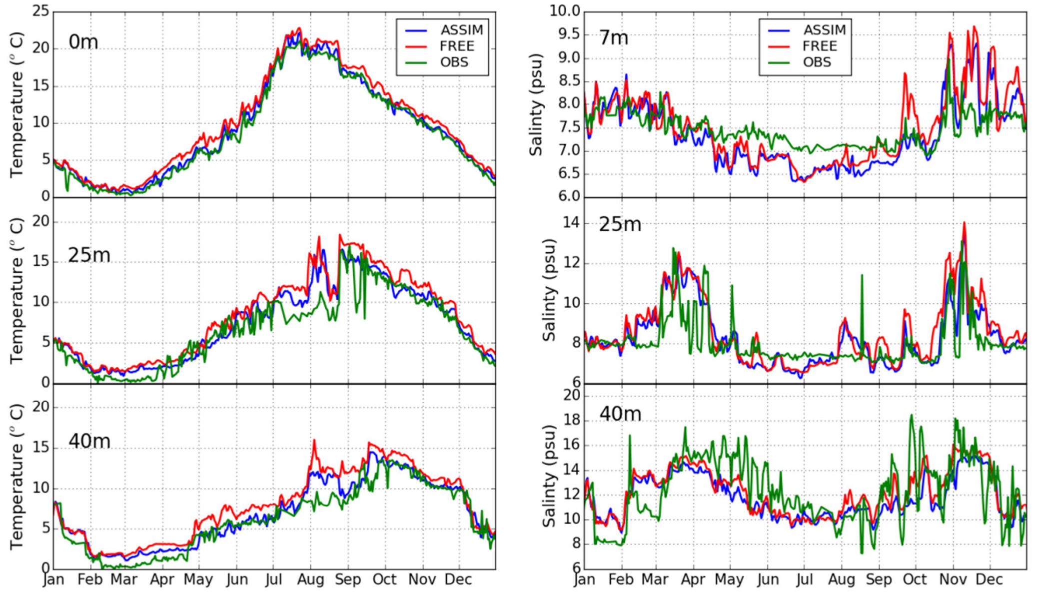

Figure 5. The time series of temperature (left panel) at a depth of 0, 25 and 40 m and salinity (right panel) at a depth of 7, 25 and 40 m at the Arkona station (13.87^{o}E, 54.88^{o}N), respectively.

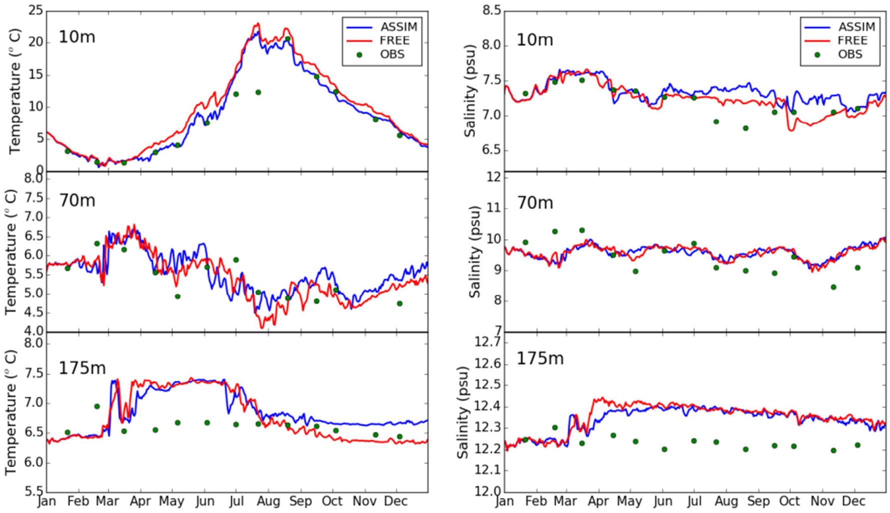

Figure 6. The time series of temperature (left panel) and salinity (right panel) at the BY15 station (20.05°E, 57.33°N) at a depth of 10, 70 and 175 m, respectively.

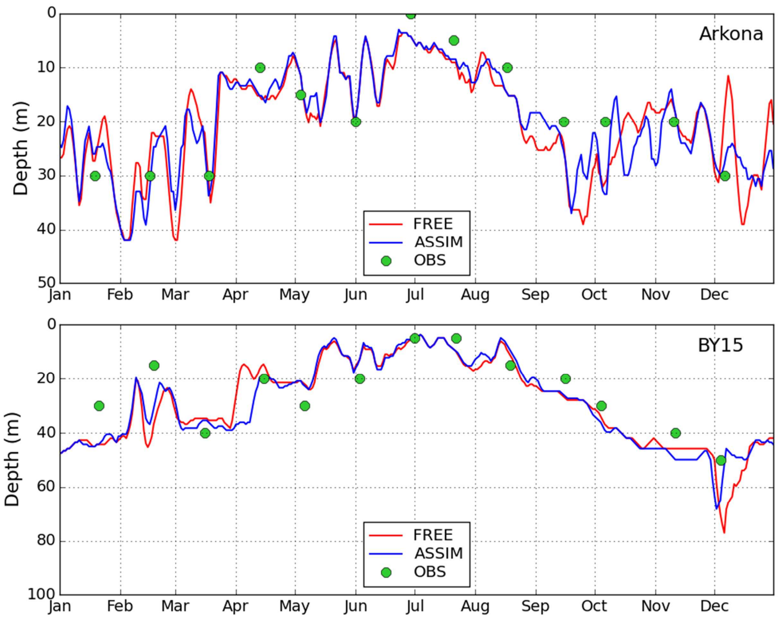

Figure 7. The time series of mixed layer depth at Arkona and BY15 station.

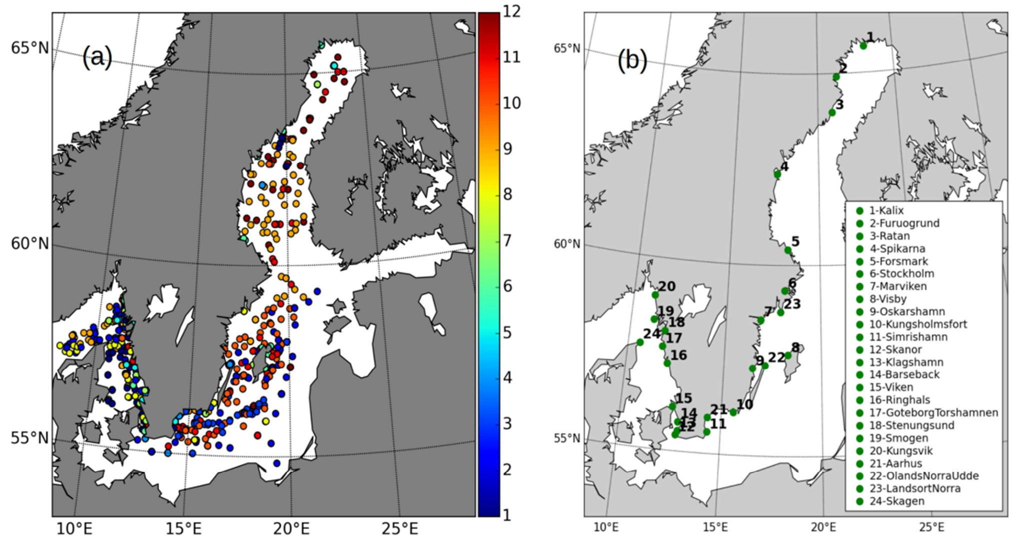

Figure 8. (a) Map of the temperature and salinity profiles from SHARK database in 2010. The colors show the observations months.(b) The tide gauges station along the Swedish coast.

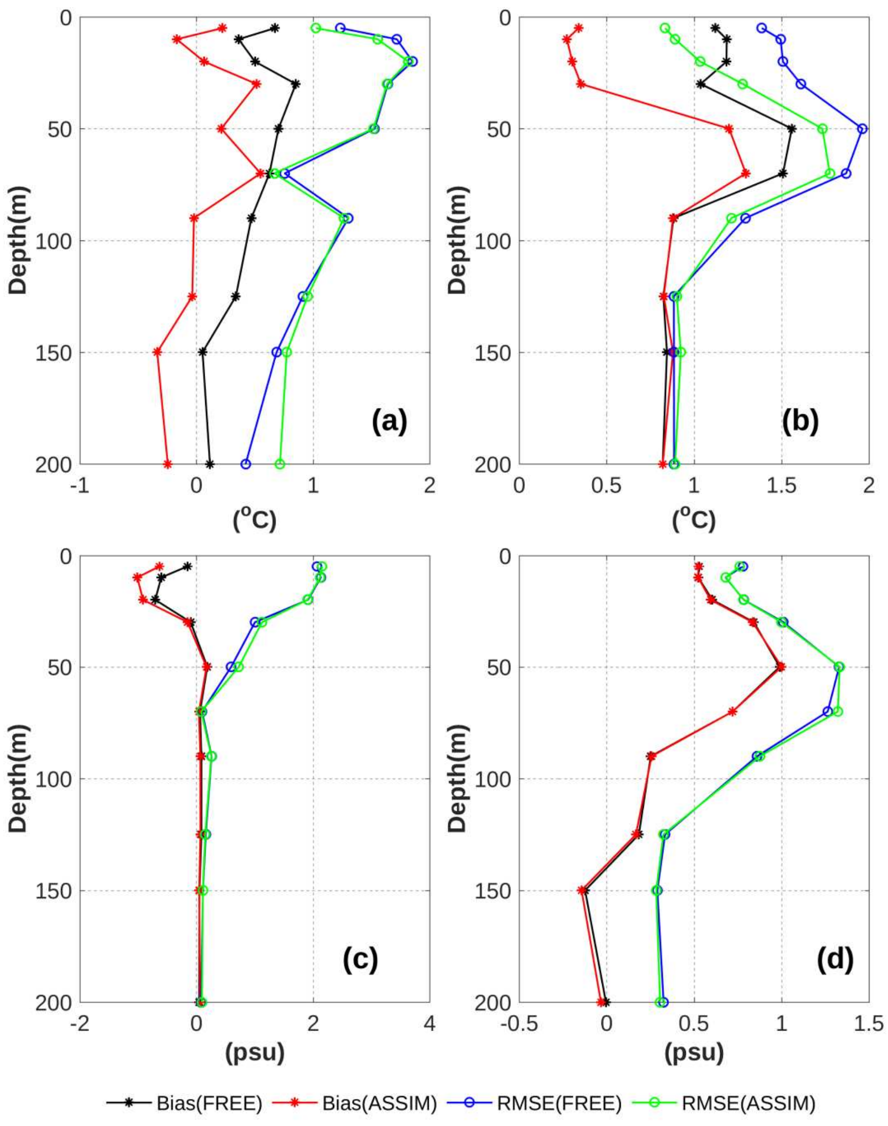

Figure 9. The overall RMSE and bias of temperature (up panel) and salinity (down panel) from FREE and ASSIM relative to observations as a function of water depth inside (b,d) and outside (a,c) of the Baltic Sea.

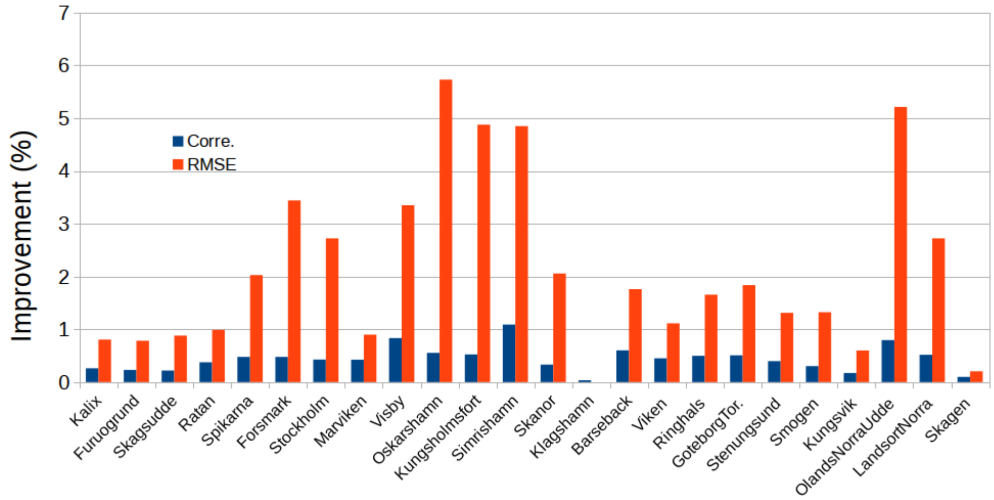

Figure 10. The improvement (%) of correlation and RMSE for the SLA at the tide gauges stations. The station position is in the Figure 8b.

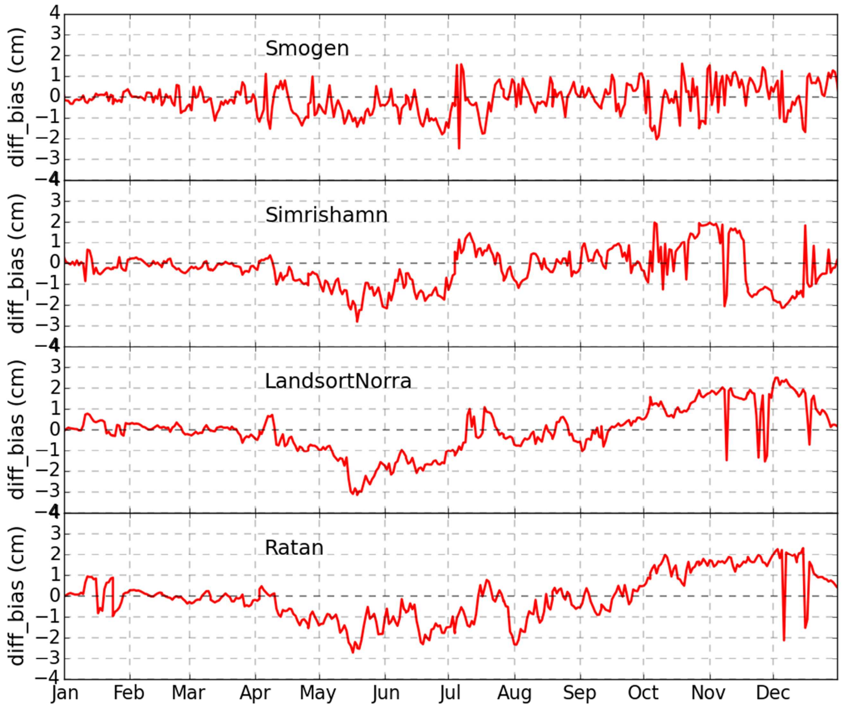

Figure 11. The variation of SLA biases in ASSIM relative to FREE against observations as a function of time. The station position is shown in the Figure 8b.

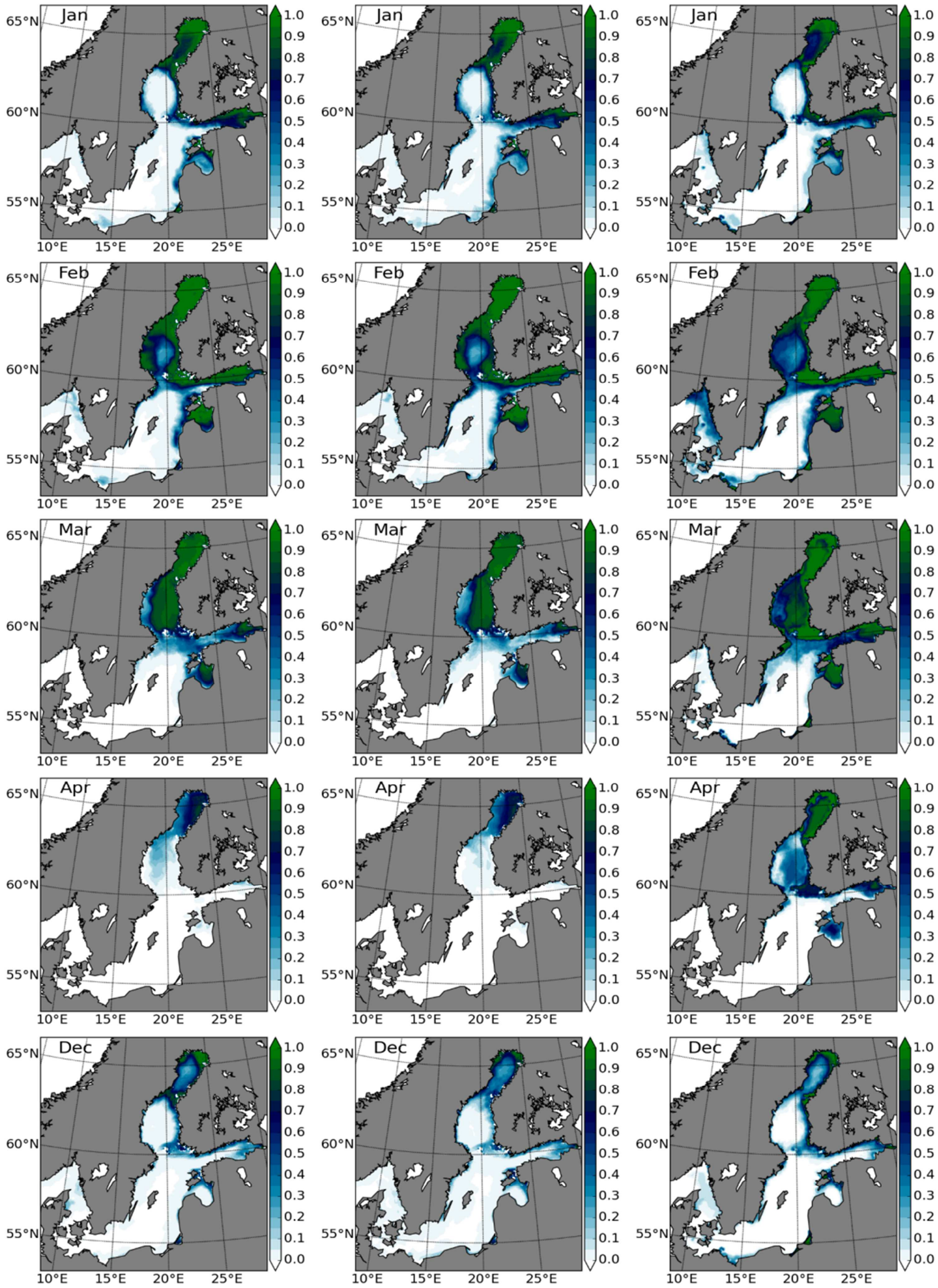

Figure 12. The monthly mean sea ice concentrations in FREE (left panel), ASSIM (middle panel) and IceMap (right panel), respectively.

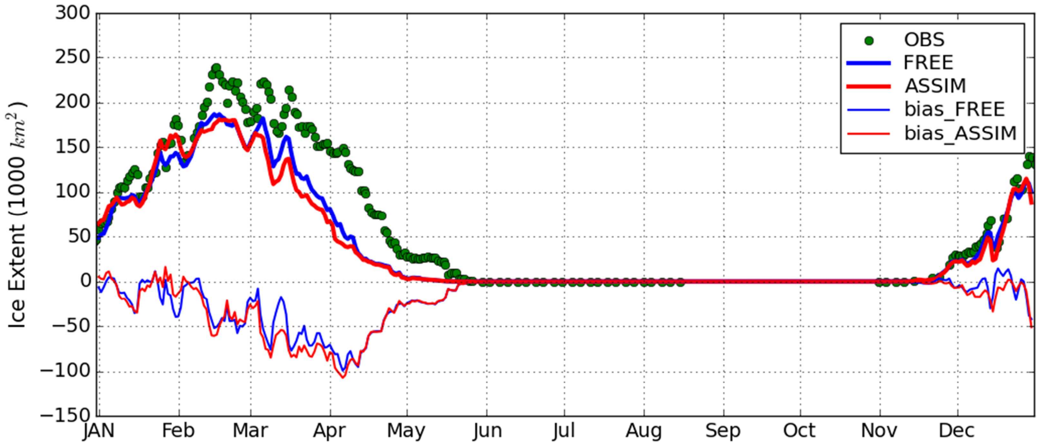

Figure 13. The daily sea ice extent from FREE, ASSIM and IceMap and the sea ice extent bias (modelled minus observed field), respectively.