

**DOCTORAL THESIS**

# Development of Data Assimilation for Numerical Forecasts in Estonian Marine Areas

Mihhail Zujev

TALLINN UNIVERSITY OF TECHNOLOGY  
DOCTORAL THESIS  
35/2021

# **Development of Data Assimilation for Numerical Forecasts in Estonian Marine Areas**

MIHHAIL ZUJEV





TALLINN UNIVERSITY OF TECHNOLOGY

School of Science

Department of Marine Systems

This dissertation was accepted for the defence of the degree 04/06/2021

**Supervisor:**

Prof Jüri Elken  
School of Science  
Tallinn University of Technology  
Tallinn, Estonia

**Co-supervisor:**

Dr Priidik Lagemaa  
School of Science  
Tallinn University of Technology  
Tallinn, Estonia

**Opponents:**

Dr Laura Tuomi  
Finnish Meteorological Institute

Dr Uldis Bethers  
University of Latvia

**Defence of the thesis:** 09/07/2021, Tallinn

**Declaration:**

I hereby declare that this doctoral thesis, my original investigation, and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been submitted for a doctoral or equivalent academic degree.

Mihhail Zujev

-----  
signature



European Union  
European Regional  
Development Fund



Investing  
in your future

Copyright: Mihhail Zujev, 2021

ISSN 2585-6898 (publication)

ISBN 978-9949-83-711-3 (publication)

ISSN 2585-6901 (PDF)

ISBN 978-9949-83-712-0 (PDF)

Printed by Auratrükk

TALLINNA TEHNIKAÜLIKOOL  
DOKTORITÖÖ  
35/2021

**Eesti merealade prognoosisüsteemi  
arendamine vaatlusandmete assimileerimise  
abil**

MIHHAIL ZUJEV





# Contents

List of Papers.....	6
Author's Contribution to the Papers.....	7
Abbreviations.....	8
1 Introduction.....	9
2 Materials and methods.....	11
2.1 Description of the study area.....	11
2.2 Model description.....	12
2.3 Assimilation algorithms.....	13
2.3.1 General approach.....	13
2.3.2 Optimal interpolation.....	13
2.3.3 Successive corrections.....	14
2.3.4 DA using reconstruction from empirical orthogonal functions.....	15
2.4 Observational data.....	17
2.4.1 FerryBox observations and shipborne monitoring.....	17
2.4.2 Satellite observations.....	18
2.5 Validation of assimilation results.....	19
3 Results.....	20
3.1 Assimilation of satellite SST data from CMEMS.....	20
3.2 Assimilation of in-situ data with EOF method.....	22
3.2.1 SST and SSS reconstruction with EOF.....	22
3.2.2 Assimilation using EOF reconstruction.....	24
4 Conclusion.....	28
5 References.....	29
Abstract.....	35
Lühikokkuvõte.....	36
Appendix.....	37
Paper I.....	37
Paper II.....	53
Paper III.....	75
Curriculum Vitae.....	96
Elulookirjeldus.....	97

## List of Papers

The list of author's Papers, on the basis of which the thesis has been prepared:

- I Zujev, M. and Elken, J., 2018. Testing marine data assimilation in the northeastern Baltic using satellite SST products from Copernicus Marine Environment Monitoring Service. *Proceedings of the Estonian Academy of Sciences*, 67 (3), 217–230. 10.3176/proc.2018.3.03.
- II Elken, J., Zujev, M., She, J. and Lagemaa, P., 2019. Reconstruction of Large-Scale Sea Surface Temperature and Salinity Fields Using Sub-Regional EOF Patterns From Models. *Frontiers in Earth Science*, 7, 1–20. 10.3389/feart.2019.00232.
- III Zujev, M., Elken, J. and Lagemaa, P., 2020. Data assimilation of sea surface temperature and salinity using basin-scale EOF reconstruction: a feasibility study in the NE Baltic Sea. *Ocean Science*, 10.5194/os-2020-43.

Conference presentations:

- I Zujev, M. and Elken, J., 2017. Seasonal features of marine data assimilation using satellite SST from Copernicus Marine Service. *The 11th Baltic Sea Science Congress: 12-16 June 2017, Rostock, Germany*. 10.12754/procs-2017-bssc.
- II Zujev, M. and Elken, J., 2017. Testing marine data assimilation in the northeastern Baltic using satellite SST products from Copernicus Marine Environment Monitoring Service. *Baltic from Space Workshop 2017. 29-31 March 2017 Helsinki, Finland*.
- III Elken, J. and Zujev, M., 2018. Using model-based sub-regional EOF patterns to reconstruct temperature and salinity fields from observations. In: S. Köppen and M. Reckermann (Ed.). *2nd Baltic Earth Conference the Baltic Sea in Transition (18–19). Geesthacht, Germany: International Baltic Earth Secretariat Papers. (International Baltic Earth Secretariat Papers; 13)*.
- IV Zujev, M., Elken, J., She, J., and Lagemaa, P., 2019. Examples of reconstruction of fine-grid SST and SSS patterns using less than 0.5% of grid coverage by observations. *Gulf of Finland Science Days "Facing our Common Future". 13-14 November 2019, Helsinki, Finland*.
- V Elken, J., Zujev, M., and Lagemaa, P., 2018. Reconstructing sea surface temperature and salinity fields in the northeastern Baltic from observational data, based on sub-regional Empirical Orthogonal Function (EOF) patterns from models. *2018 IEEE/OES Baltic International Symposium (BALTIC): 12-15 June 2018, Klaipeda, Lithuania. IEEE*. 10.1109/BALTIC.2018.8634845.

## **Author's Contribution to the Papers**

Contribution to the papers in this thesis are:

- I The author updated the map of the research domain, wrote and debugged the code for DA algorithm, collected and prepared observational data for further calculations, carried out DA experiments, performed the analysis of the results, contributed to the visualization of the results, made all the figures, wrote initial text of the results, contributed to the writing of other parts, and presented the paper at the 11<sup>th</sup> Baltic Sea Science Congress (Rostock) and at the Baltic from Space Workshop (Helsinki).
- II The author collected and prepared model data, participated in developing new algorithm, wrote and debugged the code, performed the EOF calculations, performed analysis of the results, made drafts of figures, wrote initial text of the results, contributed to the writing of other parts, and presented the paper at the IEEE/OES Baltic International Symposium (Klaipeda) and at the Gulf of Finland Science Days (Helsinki).
- III The author collected and prepared observational data, wrote and debugged the code, carried out DA experiments, performed the analysis of the results, contributed to the visualization of the results, wrote initial text of the results part, including drawing the figures, and contributed to the writing of other parts.

## Abbreviations

CMEMS	Copernicus Marine Environment Monitoring Service
BOOS	Baltic Operational Oceanographic System
DA	Data assimilation
EOF	Empirical orthogonal functions
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FB	FerryBox
FerryBox	The through-flow system installed on board of a ship-of-opportunity in order to continuously measure physical, chemical and biological parameters
FR	Free model run without assimilation
HBM	HIROMB-BOOS Model
HBM-EST	HBM setup for Estonian coastal waters
HIROMB	High Resolution Operational Model for the Baltic Sea
HIRLAM	High Resolution Limited Area Model
Ifremer	Oceanographic institution in France, L'Institut Français de Recherche pour l'Exploitation de la Mer
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
OI	Optimal interpolation
RMSD	Root-mean-square difference
SC	Successive corrections
SSS	Sea surface salinity
SST	Sea surface temperature
STD	Standard deviation

# 1 Introduction

Oceans and marginal seas constitute the domains of the Earth System with increasing societal significance, as defined by a number of global and European policy documents, such as the UN Sustainable Development Goals, Paris Climate Agreement, and the EU Green Deal. Improved ocean governance necessitates that present marine information systems, based on in situ observations, remote sensing, and numerical modelling (both in operational forecasting and in climate projection modes) would undergo fast development, in order to provide timely, more detailed, and more accurate information products.

When making numerical forecasts using models, improved accuracy in relation to observations is usually achieved by data assimilation (DA), which includes interpolation or reconstruction of observations and/or model errors, both in meteorology and oceanography (Ghil and Malanotte-Rizzoli, 1991; Ide et al., 1997). Oceanographic data assimilation has specific features (Ghil, 1989), owing to the nature of governing processes (landlocked basins, shallow areas, and wind driving characterized oceans; the atmosphere is unbounded, “deep”, and self-driving by polar-tropical gradients), but also of methods and observation coverage.

In the Baltic Sea data assimilation tests started in 2000s. Studies on sea level assimilation have been performed by a number of research groups (Canizares et al., 2001; Sørensen and Madsen, 2004, Ivanov et al., 2012), based on the different variations of Kalman filter. Assimilation of scalar variables like temperature and salinity has been tested by Funkquist (2006) who used 3D optimal interpolation (3D OI) for satellite and profile data. The OI method needs prescription of correlation functions which were estimated by Høyer and She (2007), She et al. (2007) and Fu et al. (2011a). Cressman method of successive corrections (SC) for satellite-based SST data was used by Nowicki et al. (2015). Regarding operational forecasts, several experiments have been performed to test the results of DA methods: 3DVAR with isotropic (Zhuang et al., 2011) and anisotropic (Liu et al., 2009) recursive filters to estimate covariance functions, Ensemble Optimal Interpolation (Fu et al., 2011b) and Singular Evolutive Interpolated Kalman Filter (Losa et al., 2012, 2014). In long-term studies, reanalysis has been made using 3DVAR (Fu et al, 2012; Fu, 2016), SC (Axell, 2013), Ensemble Optimal Interpolation (Liu et al., 2013; 2014) and Ensemble 3DVAR (Axell and Liu, 2016). (Paper I)

In optimal interpolation (OI) (Gandin, 1963) it is usually assumed that local covariance decreases with a distance between the points by some fading functions, like Gaussian, damped cosine or exponential decay. In the open sea where observations are dense (e.g., satellite SST in cloud-free conditions), the OI is sufficiently good (Høyer and She, 2007). For cases with sparse observations or in coastal areas with complicated covariance patterns, a more comprehensive reconstruction method should be needed. (Paper II)

In relation to improvements of DA, several statistical methods have been developed for the data reconstruction, like various options of regression, optimal interpolation and Empirical Orthogonal Functions (EOFs). Reconstruction of acceptable quality (in terms of statistics) should account for the multiscale spatial and temporal covariance. There is a number of processes that cause significant covariance over large distance, for example, warm and cold weather events, occurrence of storms, differential heating or cooling of shallow coastal areas compared to the deeper offshore regions (Legrand et al., 2015), patterns of freshwater plumes from rivers (Soosaar et al., 2016). Covariance patterns



have geometrically often elliptical form, that are stretched along the coasts or isobaths (Fu et al., 2011a).

Complementary to the classical OI, methods based on the EOF have been developed and applied. The methods have produced reliable large-scale patterns, which were approximated by dominant EOFs (Kaplan et al., 1997; Kim, 1997; Menemenlis et al., 1997; Beckers and Rixen, 2003); in the regions of dense sampling multivariate approach can be applied, and the anomalies from large-scale fields are interpolated using OI or some of its variations. (Paper II)

The novelty of this research is developing and testing a new statistical data assimilation algorithm. Testing the traditional assimilation methods with new sea surface temperature data sets from satellite remote sensing (Paper I) pointed to the need for a more detailed evaluation of spatial covariance statistics of the assimilated variable. A new EOF-based method for the reconstruction of gridded data fields of sea surface temperature and salinity was developed in the course of these covariance studies (Paper II). The new reconstruction method was applied in the new data assimilation algorithm (Paper III) that revealed a significant reduction of forecast errors, compared to the control run without data assimilation.

Surface temperature and salinity were chosen as variables for DA in the present study since they are important parameters for the modelling of the hydrodynamics, and there are abundant observational datasets available. DA of subsurface data was not included in the present study because of a much smaller amount of data, and will be considered in the future.

The thesis has the following specific objectives:

- to test marine data assimilation into the model of the northeastern Baltic using satellite SST products from CMEMS and FerryBox observations (Paper I),
- to develop and test the method for large-scale EOF analysis of sub-regional time-dependent SST and SSS data, based on the covariance estimates from the model results (Paper II),
- to implement this statistical reconstruction technique into the DA of the forecast model (Paper III),
- to study the feasibility of this assimilation method (Paper III),
- to assess the performance of the model with DA (Papers I and III).

## 2 Materials and methods

### 2.1 Description of the study area

The study was conducted in the northeastern part of the Baltic Sea (Fig. 1), between 21°E and 30°E from west to east and between 57°N to 61°N from south to north. The **region** includes the Gulf of Finland, the Gulf of Riga, and the northeastern Baltic Proper.

The **topography** of the Baltic Sea is irregular with variable depths. There are narrow and shallow straits (Skagerrak, Kattegat, Belts and Öresund), deep basins (Arkona basin, Gotland deep) and several large gulfs (the Gulf of Finland, the Gulf of Riga, the Bothnian sea). Some sub-basins have sills (Bothnian Bay, the Gulf of Riga) restricting the inflow of bottom water from the Baltic Proper. The entire sea is regarded as a brackish estuarine-type multi-basin water body (Elken and Matthäus, 2008; Leppäranta and Myrberg, 2009).

The **coastline** of the region is very diverse. There is the Åland archipelago in the NW part and many small islands along the southern shore of Finland. Neva Bay in the NE has a triangular shape and smooth forms. The Estonian coast of the Gulf of Finland has interchanging sections of rocky and sandy beaches and a few islands. In the W there are two large islands (Saaremaa and Hiiumaa) with narrow straits connecting enclosed Väinameri with Baltic Proper, the Gulf of Riga, and the Gulf of Finland. There is a smooth coastline along the Gulf of Riga that has a partly elliptical shape.

The region lies in the temperate **climatic zone**. In the summer, SST maximum usually exceeds 15 °C in July or August (Alenius et al., 1998), whereas temperatures up to 25 °C can occur in the shallow coastal zones (Stramska and Białogrodzka, 2015). Almost every winter, sea ice forms with variable extent and thickness; during severe winters (Vihma and Haapala, 2009), the Gulf of Finland and the Gulf of Riga are fully ice-covered (Jevrejeva et al., 2004).

The **wind regime** is anisotropic, with dominant SW winds (Soomere, 2003; Soomere and Keevallik, 2003). During a period covered in this study, five storms were observed in the Baltic Proper at Utö station (Arra, 2018).

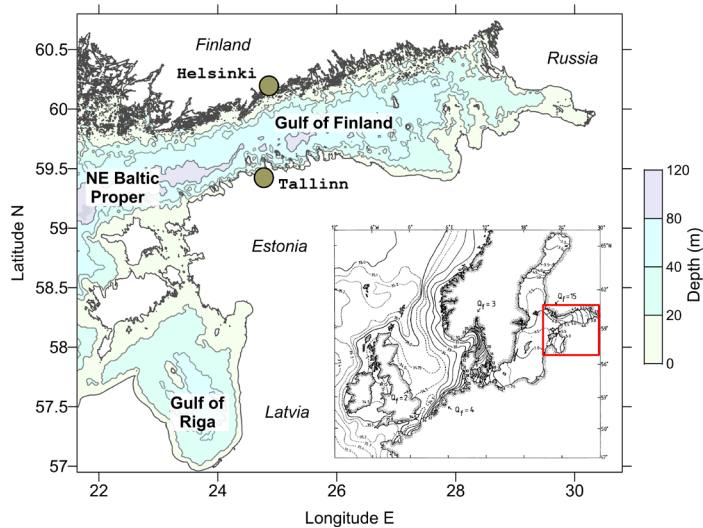
The **river discharge** impacts the basin in different ways. The Gulf of Finland and the Gulf of Riga together receive 34% of the total freshwater discharge to the Baltic Sea as can be calculated from the data by Bergström (1994) and Johansson (2017). This increases SSS values from east to west, characteristic to estuarine areas (Yurkovskis et al., 1993; Alenius et al., 1998), reaching 7–8 g kg<sup>-1</sup> in the Baltic Proper (Köuts and Omstedt, 1993). SSS has numerous high-gradient frontal regions influenced by fragmented coastline and multiple rivers entering the basin. Large-scale SSS patterns are guided by unsteady circulation that depends on the climatic variations of atmospheric forcing (Alenius et al., 1998).

The **thermal regime** is dominated by a seasonal heat cycle and modified by differential heating and cooling at variable depths in coastal and offshore areas. The warm upper layer of 10–20 m thickness is well mixed down to the thermocline or to the bottom, depending on which of them is shallower. Occasionally, wind-driven coastal upwelling processes disrupt this warm layer (e.g., Uiboupin and Laanemets, 2009). Upwelling and downwelling patterns induced by the transient wind fields have strong impact on SST (e.g., Laanemets et al., 2011).

The **salinity distribution** varies in different sub-basins of the region. Water exchange of the Gulf of Finland (Maljutenko and Raudsepp, 2019) and the Gulf of Riga (Lips et al,

2016; Maljutenko, 2019) is governed by an inflow of more saline waters in the deep layers and outflow of less saline waters in the surface layers. During stronger winds with easterly directions, this estuarine circulation may be temporarily reversed (Elken et al. 2003; Liblik et al., 2013). The Gulf of Finland waters have lower salinity as of those in the Baltic Proper, and there is no barrier between them, so saline waters can freely enter the Gulf of Finland and create an estuarine halocline there (Elken et al., 2003; Liblik et al., 2013). In the Gulf of Riga situation is different: only surface waters from the Baltic Proper can overpass the sill between two sub-basins (Lilover et al., 1998).

The **surface circulation** in the Baltic Proper is cyclonic (Lehmann and Hinrichsen, 2000), unlike the Gulf of Riga where clockwise circulation may also occur (Maljutenko, 2019). Depending on the time scale, season, period, method of calculation, and depth there are various estimates of gyres in the Gulf of Finland according to different studies. Surface circulation in the Gulf of Finland near the coast of Finland has a weak eastward component and strong westward flow in Neva Bay and along the Estonian coast (Soomere et al., 2011). While horizontal circulation in both of the gulfs has been historically considered cyclonic, anticyclonic gyres occur as well. Mesoscale variability has rather short spatial scales; the Rossby deformation radius ( $R_d$ ) values are from a few km to about seven km (Alenius et al., 2003).



**Figure 1.** Map of the study area in the northeastern Baltic with depth contours. Shown are the sea areas of the Gulf of Finland, Gulf of Riga and part of the northeastern Baltic Proper. Insert presents the map of surface salinity of the Baltic and North seas by Rodhe (1998). The arrows present the mean basin-wide river discharges in  $1000 \text{ m}^3 \text{ s}^{-1}$ . The location of our study area is given on the insert by a red box. (Paper III)

## 2.2 Model description

The HBM model (Berg and Poulsen, 2012) originates from the BSHcmod model (Kleine, 1994) initially developed by the Federal Maritime and Hydrographic Agency at the beginning of the 1990s. Later it was modified and used by a variety of institutions within the HIROMB cooperation.

The model was chosen for the present research since it was routinely used for operational forecast in Estonia.

HBM uses the Arakawa C-grid, and produces a forecast for 16 ocean variables including temperature, salinity (which were studied in this research), current speed, ice concentration and others. More information is given in the report by Berg and Poulsen (2012) which has full description of the model and its validation.

The larger domain of the model covers the entire Baltic Sea including Danish straits with a portion of the southern part of the North Sea. For smaller areas, there are setups with finer resolution; one of them covers the northeastern Baltic, it has been used in the present research.

The HBM-EST setup has a 0.5' N x 1' E resolution containing the Gulf of Finland, the Gulf of Riga and the northeastern portion of the Baltic Proper (Lagemaa, 2012). The model fields are three-dimensional having 455 x 529 x 39 points (by latitude, longitude and depth correspondingly) with 750 088 wet-points, and 71 986 of them on the surface. Data for the western open boundary originate from the Baltic-wide HBM model, provided by the Copernicus Marine Environment Monitoring Service (CMEMS, <https://marine.copernicus.eu/>, last access: 2 May 2020). Atmospheric forcing is provided by the Estonian implementation of HIRLAM (Männik and Merilain, 2007). Forecasts can be made up to 48 hours with a time step of one hour. In this research 24 hour forecast was used.

For DA experiments, the model was run with data from the 1<sup>st</sup> of May until the 31<sup>st</sup> of December 2015. This short period was chosen due to computational reasons, since the aim was to keep the high resolution needed for the operational forecasts, which are made routinely by the Estonian Weather Service. Model data from the control run (free run) without DA were available in the archive for all the DA experiments. For the calculation of EOF modes, model data for the longer period were used, as will be described in Sect. 2.3.3.

## 2.3 Assimilation algorithms

### 2.3.1 General approach

All assimilation methods were performed in the two-dimensional surface layer for computational efficiency; there were no observations on deeper model levels. In the model results, vertical profiles did not have distinctive graphical signatures of assimilation (abrupt steps or jumps) since vertical mixing dominated over introduced innovations, as the observations were assimilated every day.

### 2.3.2 Optimal interpolation

OI was developed by Gandin (1963), and in this research it was used for assimilation (Paper I). The algorithm uses the least-square minimization of analysis errors to calculate the weight coefficients. The preparation of data and the algorithm itself are described below.

**Preparation** includes the filtering and averaging of observational data. Irregular satellite observations were averaged in each wet-point of the 455x529 grid.

Existing values were then averaged again for the coarse grid 45x53, leaving 744 wet-points.

The **algorithm** has three steps.

1) The first step is to derive the equations, based on the minimization of function

$Q = \left[ f_0 - \sum_{j=1}^n w_j (f_j + \varepsilon_j) \right]^2 \Rightarrow \min$ . It is minimized with respect to interpolation weights  $w_j$ , where  $f_0 = \tilde{x}_a - x_b$ , which is the difference between the unknown “true” state ( $\tilde{x}_a$ ) and background ( $x_b$ ) in the model point  $i$ ;  $f_j = y_j - \hat{y}_j$  is the difference between observed values ( $y_j$ ) and modelled values ( $\hat{y}_j$ ) taken at the observation points;  $\varepsilon_j$  – random errors of observations.

Minimization is done by setting the  $n$  constraints for the derivatives  $\frac{\partial Q}{\partial w_j} = 0$  using the conditions  $\bar{f}_j = 0, \bar{f}_0 = 0, \bar{\varepsilon}_j = 0, \bar{\varepsilon}_j f_j = 0, \bar{\varepsilon}_0 f_0 = 0$ . As a result we obtain for the  $i$ -th model point the system of  $n$  linear equations regarding  $w_j$

$$\sum_{j=1}^n \bar{f}_k \bar{f}_j w_j + \sigma_{\varepsilon_k}^2 w_j = \bar{f}_k \bar{f}_0, \quad (1)$$

where  $k = 1 \dots n$  is a number of equation.

By dividing Eq. (1) with the variance  $\sigma_f^2 = \bar{f}_k^2$ , we obtain correlation instead of spatial covariance.

2) The second step is calculating the vector of weights  $\mathbf{w} = \mathbf{b}(\mathbf{B} + \eta^2 \mathbf{I})^{-1}$ , based on Eq. (1) divided by variance. Here  $\mathbf{b} = \frac{\bar{f}_k \bar{f}_0}{\sigma_f^2}$  is correlation vector between the observation point and the  $i$ -th model point,  $\mathbf{B} = \left\{ \frac{\bar{f}_k \bar{f}_j}{\sigma_f^2} \right\}$  – correlation matrix between the individual observation points,  $\mathbf{I}$  – a unit matrix,  $\eta^2$  – relative noise variance.

3) The third step is a calculation of the assimilated value

$$x_a = x_b + \mathbf{w}(\mathbf{y} - \hat{\mathbf{y}}) = x_b + \sum_{j=1}^n w_j (y_j - \hat{y}_j). \quad (2)$$

Correlations  $\mathbf{B}$  and  $\mathbf{b}$  were approximated by the Gaussian function from the distance  $r$  between the correlated points. Anisotropic correlation features were taken into account by the directional distribution of the correlation scale from the angle  $\theta$  in the form of ellipse dependence  $D = a \sin(\theta - \theta_0) + b \cos(\theta - \theta_0)$  relative to the reference angle  $\theta_0$ . Ellipse semi-axes are designated as  $a$  and  $b$ . This way the correlation has been adopted in the form  $B(r, \theta) = \exp\left(\frac{-r^2}{D^2(\theta)}\right)$ , where  $D = D(\theta)$  was pre-calculated in the each model grid point according to the coastline and topography. According to the results by Høyer and She (2007), longer correlation scales were taken along the coasts and the isobaths and shorter scales in the perpendicular direction. The typical horizontal impact scale along the coast or isobath was chosen at 15 km. Standard deviations for the entire run were taken  $\sigma_0^2 = 0.5$  and  $\sigma_m^2 = 1.0$ .

### 2.3.3 Successive corrections

The successive correction method or Cressman method was introduced by Cressman (1959). The algorithm is based on the following assumptions: a) between the state variables are univariate relations; b) weights of the individual observations  $w_j$  in Eq. 3. decrease with the distance  $d_j$  between the observation point  $j$  and the model grid point  $i$ . The weights are positive within the influence radius and zero outside of it.

**Assimilation** is done in two steps.

1) The first step is a calculation of weights:

$$W_j = \frac{\max\left(0, \frac{R^2 - d_j^2}{R^2 + d_j^2}\right)}{\sum_{j=1}^n \max\left(0, \frac{R^2 - d_j^2}{R^2 + d_j^2}\right) + \eta^2}, \quad (3)$$

where  $R$  is influence radius (37 km, see details below) around the model point  $i$ ;  $j$  – observation point running index;  $k$  – number of observations out of total  $n$  observations are located;  $\eta^2$  – relative noise variance; introduced for reduction of the assimilation weights in realistic noisy conditions. It is estimated from the variances of observation errors  $\sigma_o^2$  and background errors  $\sigma_b^2$ ,  $\eta^2 = \frac{\sigma_o^2}{\sigma_b^2}$ . In the noiseless case ( $\eta^2 = 0$ ) the sum of the weights is equal to unity (Paper I).

2) The second step is a calculation of analysis (Eq. 2).

Data assimilation for SST (Paper I) was made with influence radius value of 37 km (20 nautical miles, 40 grid points). This length is about ten times larger than Rossby deformation radius. Therefore, the impact of individual mesoscale eddies is suppressed, but basin-scale SST features are kept. Weight function has a greater impact within the nearest 5 km, then it decreases to zero for 37 km.

The **preparation** for assimilation includes averaging the observations over each grid cell in order to avoid oversampling problems. During the testing of the scheme, the values of  $\sigma_o^2$  and  $\sigma_b^2$  were not known in advance. For the chosen dataset acceptable results were obtained with  $R = 37$  km and  $\eta^2 = \frac{\sigma_o^2}{\sigma_b^2} = 2$ . These values were used throughout the entire model run.

### 2.3.4 DA using reconstruction from empirical orthogonal functions

EOF is a statistical method, developed for meteorological applications (Lorenz, 1956), and is mainly used for the decomposition of continuous space-time field into the sum of basic functions of space and expansion functions of time. This approach allows for the reduction of the number of variables without significant loss of signal. A detailed description of the classical EOF technique with examples can be found e.g. in von Storch ja Zwiers (1999).

A new method was developed (Paper II) that allows for the making of a gridded reconstruction of irregular point observations, using EOF modes calculated from the model results.

The **preparation** is done in several steps:

1) The first step is the calculation of eigenvectors based on the covariance matrix:

$$\mathbf{B} = \frac{1}{N-1} \mathbf{X}^T \mathbf{X},$$

where  $\mathbf{X}$  is matrix of deviations from mean model results,  $N$  – number of wet-points,  $\mathbf{B}$  – covariance matrix between all possible pairs of wet-points.

2) The second step is the calculation of eigenvalues and space-dependent eigenvectors:

$$\mathbf{B}\mathbf{E} = \mathbf{\Lambda}\mathbf{E},$$

where  $\mathbf{E}$  is matrix of eigenvectors,  $\mathbf{\Lambda}$  – matrix of eigenvalues.

As far as modes are orthogonal, covariance is additive with respect to the EOF modes, *i.e.*, the full covariance is the sum of the covariance of the component data sets. Calculated covariance is not homogeneous, which is usually assumed in the implementation of OI.

3) The third step is a reconstruction of observations to the entire grid at time  $i$ :

$$\hat{\mathbf{x}}_i = \mathbf{E}\hat{\mathbf{a}}_i, \quad (4)$$

where  $\hat{\mathbf{x}}_i$  is vector of reconstructed observations,  $\mathbf{E}$  – matrix of eigenvectors,  $\hat{\mathbf{a}}_i$  – vector of observational amplitudes.

The expression to find observational amplitudes is

$$\hat{\mathbf{a}}_i = (\mathbf{E}^T\mathbf{H}_i^T\mathbf{H}_i\mathbf{E})^{-1}\mathbf{E}^T\mathbf{H}_i^T\mathbf{y}_i, \quad (5)$$

where  $\mathbf{H}_i$  is observation operator,  $\mathbf{y}_i$  – observation values.

Usually only  $L$  most energetic modes are taken into account.

The  $\hat{\mathbf{a}}_i$  values should follow the least-square minimization of reconstruction error in relation to observations  $\|\mathbf{y}_i - \mathbf{H}_i\mathbf{E}\hat{\mathbf{a}}_i\|^2 \Rightarrow \min$ . Note, that during standard EOF decomposition, there are values of “observations” available at all space points ( $\mathbf{y}_i = \mathbf{x}_i$ ), the observation operator  $\mathbf{H}_i$  is unity, and therefore Eq(5) reduces to the standard EOF formula  $\hat{\mathbf{a}}_i = \mathbf{E}^T\mathbf{x}_i$ .

While (4)-(5) present the time-fixed reconstruction when all observations are taken exactly at time  $i$ , then extended time-dependent reconstruction assumes that within the short time span the amplitudes depend linearly on time:  $\hat{\mathbf{b}}_p = \hat{\mathbf{a}}_i + \mathbf{d}_i \cdot \delta t_p$ , where  $\hat{\mathbf{a}}_i$  is the time-fixed amplitude,  $\mathbf{d}_i$  is the rate of change vector and  $\delta t_p = t_p - t_i$  is the difference between the observation and reference times. The amplitudes  $\hat{\mathbf{b}}_i$  are found again by least-square minimization and they are used in (4) instead of  $\hat{\mathbf{a}}_i$ .

**DA algorithm** uses the transformation of the reconstructed one-dimensional vector of observations  $\hat{\mathbf{x}}_i$  back into the two-dimensional gridded field  $\psi^o$ . The dynamic equation with Newtonian relaxation of the field  $\psi$  to observations is written

$$\partial\psi/\partial t = F(\psi) - \frac{1}{\tau}(\psi - \psi^o), \quad (6)$$

where  $F$  is model operator,  $\tau$  – adjustable relaxation time (5 or 10 days in this research).

Discrete DA has two steps:

1) The first step is the calculation of raw forecast from the analysis of the previous DA step:

$$\psi^f = \psi^{a-1} + \Delta t F(\psi^{a-1}),$$

where  $\psi^f$  is raw forecast field without DA,  $\psi^{a-1}$  – analysis field from previous DA step,  $\Delta t$  – DA time step,  $F(\psi^{a-1})$  – cycle of model time-stepping operators within  $\Delta t$ .

2) The second step is the calculation of analysis:

$$\psi^a = (1 - \alpha)\psi^f + \alpha\psi^o,$$

where  $\psi^a$  is analysis field for the new DA step,  $\psi^o$  – reconstructed observations on the model grid,  $\alpha = \Delta t / \tau$  – inverse non-dimensional relaxation time.

In time-dependent reconstruction (see details in Papers II and III), it is necessary to choose both the reference time and duration of the time interval. As with the time-fixed reconstruction, the highest mode suitable for assimilation is determined by comparing the amplitude values against statistical limits. The method allows for estimation of EOF amplitudes and reconstruction of gridded fields solely by backward observational data. This is the case for operational forecasts, where only past observations can be assimilated for producing the present nowcast maps.

## 2.4 Observational data

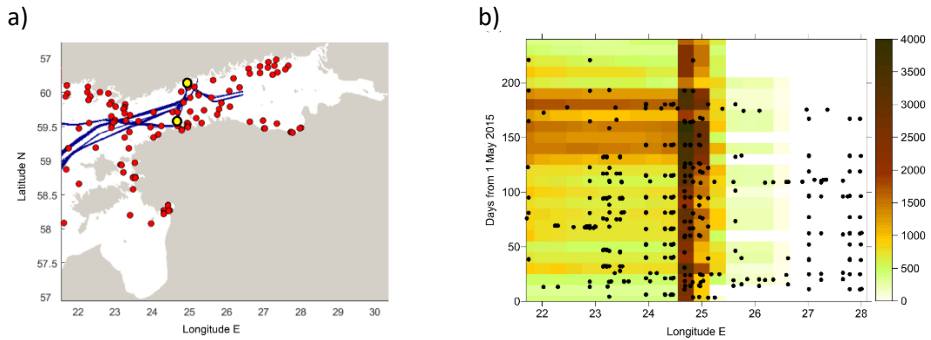
### 2.4.1 FerryBox observations and shipborne monitoring

Automatic observations made from ships crossing the sea areas were used both as independent data for validation and quality assessment (Paper I) and as input for assimilation (Paper III). FerryBox is a measurement system installed on board commercial ferries, collecting temperature, salinity, chlorophyll *a* fluorescence and turbidity data. This technology is used to study basin-scale and mesoscale processes, upwellings and to measure horizontal profiles of oceanographic parameters e.g., temperature and salinity (Kikas and Lips 2016). The water is sampled at about 4 m below the surface at different rates, but every 20 s measurement is recorded, thus covering roughly 160–200 m in a horizontal direction. There are quality check procedures to eliminate unexpected and physically unrealistic values and cross-checking with the same data from the return trip is performed as well (Kikas and Lips, 2016). A comprehensive description of technical parameters of the FerryBox system can be found in Lips et al. (2008).

Observations are available on the routes starting from the ports of Tallinn, Helsinki, Travemünde, and Stockholm (Fig. 2a) on the forth-and-back tracks twice a day (Tallinn – Helsinki) or less frequently (other routes). For the particular grid cell of the model, one mean SST value (Papers I and III) and one SSS value (Paper III) for each day were calculated regardless of the time within the day and the number of observations. Within these intervals, mean coordinates and observation time were collected for the use in EOF analysis. Observations from the shipborne monitoring were also included in the data set of EOF assimilation. Compared with the FerryBox data, shipborne monitoring enlarged the area covered by observations (locations of the observations are shown in Fig. 2b by dots), but the amount of data was very small and time intervals were large.

The data were taken as they are within the CMEMS depository for 2015. The data had passed an advanced quality check.





**Figure 2.** Distribution of observations. (a) Map of FerryBox observation points along ship tracks (blue) and shipborne monitoring observations (red) over the study period. Shown are also the locations near Tallinn (south) and Helsinki (north) depicting the route of FB line (yellow circles with black contour); (b) Observation amount over longitude and time. FerryBox data are shown by colour image; each image cell presents the number of initial observations over intervals of 10 days and 18' E longitude. Shipborne observations are shown by black dots. (Paper III)

## 2.4.2 Satellite observations

Sea surface temperature data, observed from satellites, were used as input observational data within data assimilation with OI and SC algorithms (Paper I). Gridded observation maps were obtained from the CMEMS multi-sensor product, which is built from bias-corrected mono-sensor products at a horizontal resolution of 0.02 by 0.02 degrees.

SST product from satellite observations is composed by merging of various satellite SST level 2 data (Bonekamp et al., 2016). The data have passed a significant number of quality controls. The raw data have been calibrated through an inter-sensor bias correction procedure. Nighttime SST maps were based on original SST observations without any smoothing or interpolation. Details of the product are described on the CMEMS web resource <http://cmems-resources.cls.fr/documents/QUID/CMEMS-OSI-QUID-010-009-a.pdf> (last access: 8 May 2020). Data were acquired from NASA, NOAA, IFREMER, EUMETSAT OSI-SAF and ESA using sensors like METOP\_B, SEVIRI, VIIRS\_NPP, MODIS and others. (Paper I)

Depending on cloud cover, there were from 200 up to 21000 observations per day. Some obviously erroneous SST values were filtered out (which differed more than 10 °C from model ones). All of them were used for assimilation with the Cressman method. A data thinning algorithm was used for optimal interpolation in order to reduce oversampling, leaving one value for the area of 2.5 by 5 nautical miles.

A single SST value was used for each day; it was reduced to midnight based on several available observations at different times (near-real-time).

In order to make sure that the EOF algorithm will perform in the best possible way, experiments with pseudo-observations were conducted. It was found that ca 5000 observations were needed to produce reliable results. As input data for assimilation, model values at specified locations were taken instead of real observations (Paper II).

## 2.5 Validation of assimilation results

Following the approach by Taylor (2001), for each comparison of the two variables  $f_n$  and  $g_n$  a common data set is defined where missing values of one or both data sets are ignored. If the standard deviations of the data sets are  $\sigma_f$ ,  $\sigma_g$  and the coefficient of their mutual correlation is  $r_{f,g}$ , then the centred (with bias removed) root-mean-squared difference (RMSD) of the data sets  $Q'$  reads

$$Q'^2 = \sigma_f^2 + \sigma_g^2 - 2\sigma_f\sigma_g r_{f,g} . \quad (7)$$

Maps or transect plots can be visually compared and described (Crosnier and Le Provost, 2007). This method is appropriate for situations when no rigorous criteria are set or they are unknown. Some features can be identified, inconsistencies removed and algorithms altered.

From an assimilation perspective, time-series of good analysis lay between the free run and the observations. Maps of good analysis should not have artificial and physically impossible properties, like bull's eye formation around a single point observation, rectangular shapes etc.; this should be valid also in the regions of missing observations where quantitative validation is not possible.

The model performance with respect to observations was evaluated over the grid cells – time span pairs, where and when the observations were available. Since there were low number of observations, DA results were also checked against control run without DA, though in this case it is not possible to conclude on improvement, but only to analyse the changes due to DA. Standard statistical characteristics were calculated for the individual fields: mean, standard deviation, in case of differences (for example, relative to observations): bias, RMSD (root-mean-square difference, equals to the standard deviation of difference field), and the Pearson correlation coefficient.

A comparison was made against independent data for SST (Paper I), and partially against the data which was used in assimilation (Paper III) as there were no salinity observations available from satellites.

When SST and SSS data were assimilated simultaneously using EOF reconstruction (Paper III), data from the DA experiments with a relaxation time of 5 and 10 days were compared to the same averaged observational FB data as the data from the model run without assimilation (FR). The problems concerning performance evaluations of operational ocean models were described by Hernandez et al. (2015). In this research different portion of the observations were withheld in order to see the impact of the assimilation using 50% of the available data (Gregg et al., 2009). The implementation of EOF DA described in the thesis involved about 13 k observational averages over coarse  $5' \text{ N} \times 10' \text{ E}$  grid. The reconstruction procedure by Eqs. (4)–(5) has no direct connection to the ongoing modelling (although it includes statistical results from longer model runs) and the fields of  $\psi^o$  in Eq. (6) are the only link where observations enter the DA process.

### 3 Results

The most important result of the entire work is the development of a method that allows the assimilation of fragmented observations into a model and gets better results than with free-run or simple algorithms. Additionally, it is shown that SST and SSS fields can be realistically reconstructed using the EOF technique with model-generated values. CMEMS data can be used as input for assimilation in the NE Baltic.

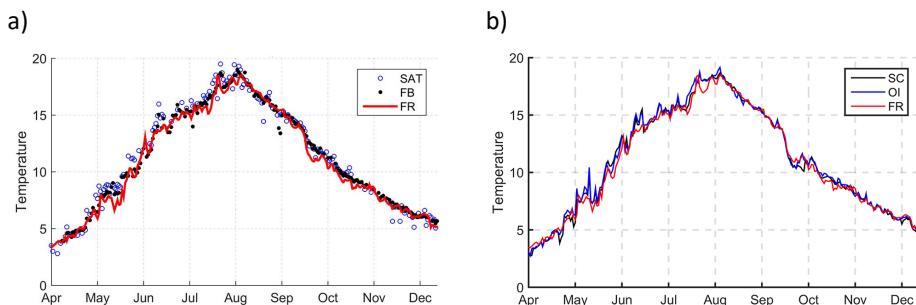
Several experiments were conducted in this study. The first experiment was conducted to make sure CMEMS data are acceptable for data assimilation. The second experiment revealed possibilities for two simple algorithms (SC and OI) to improve forecast comparing with free run (without assimilation). The third experiment showed that EOF can be used for the reconstruction of SST fields. The fourth experiment confirmed the ability of the EOF technique to assimilate sparse observations to the large fine-grid.

#### 3.1 Assimilation of satellite SST data from CMEMS

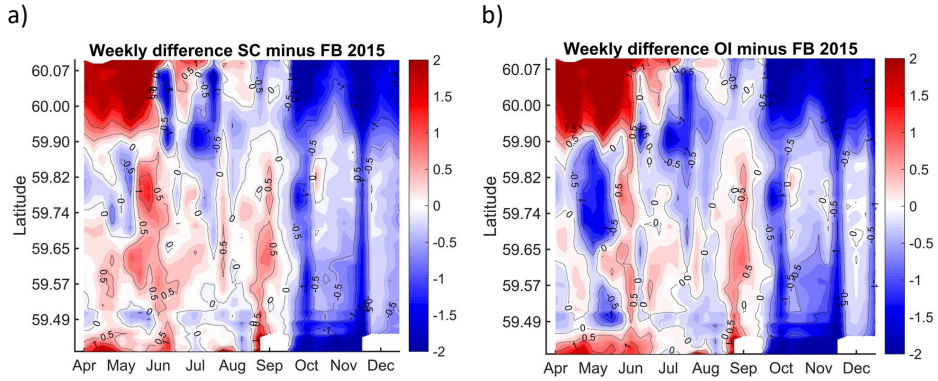
In the adopted data assimilation approach, SST satellite observations were used to correct the model forecast. FerryBox data were taken from the Tallinn-Helsinki route (Fig. 2a), which has the most frequent observations. Therefore, comparisons are presented for this route.

A statistical comparison of weekly mean values of free run (FR) forecast and assimilated SC and OI forecasts with FB data revealed that assimilation provided better correspondence to the independent observations (Table 1). Improvements were noted in degrees of bias, root-mean-squared difference RMSD, and the overall correlation (Table 1). The main performance estimator – RMSD – was  $< 1\text{ }^{\circ}\text{C}$  in all the cases.

Time series of SST from daily SAT data revealed during the warming period in the open part of the Gulf of Finland (Fig. 3), generally higher temperatures than FB. The SAT data were spiky compared with the FB data: warmer spikes occurred during the warming period and colder spikes during the cooling. The free run (FR) forecast provided in the offshore waters slightly smaller SST than observed. Data assimilation using SC and OI “dragged” the model results towards SAT observations (Fig. 3b), still, the SST spikes did not appear in the assimilated model results.



**Figure 3.** Daily SST time series in 2015 on the Tallinn-Helsinki FerryBox transect in the central part of the Gulf of Finland: FB during observation time and nightly values for FR and SAT. See Abbreviations for explanations of the legend. (Paper I)



**Figure 4.** Time-latitude map of weekly mean sea nightly surface temperature difference of assimilated with SC (a) and OI (b) in reference to Ferrybox data, between Tallinn and Helsinki. (Paper I)

The difference between the two assimilated SST datasets from the FB data is shown in Fig. 4. In the central part of the Gulf of Finland (latitudes 59.5–60 N) there is a strong seasonal signal. Thin layer temperature registered by satellite was 0.3–0.7 °C larger than the bulk temperature of the upper layer (observed at a 4 m depth) during spring and summer until August, and insignificantly (less than 0.5 °C) smaller in autumn and early winter. Larger SAT minus FB differences emerged occasionally in areas immediate to the coasts. In December the thin surface layer cooled down by 0.5–1.5 °C more than the deeper surface layer along the whole transect, also including the coastal waters.

Total RMSD between analysis and validated data was less than 1 °C for SST. Within the selected parameters of assimilation algorithms, the computationally effective SC algorithm gave slightly better results than OI in relation to independent FerryBox data.

A statistical comparison of weekly mean values of FR forecast and assimilated SC and OI forecasts with FB data (Table 1) revealed that assimilation provides better correspondence to the independent observations. Presented deviations contain the seasonal cycle, since all the statistics were calculated with respect to the constant mean value over the whole period from April to December. Standard deviations of SST were in the range of 4.2 to 4.6 °C. Calculated correlations were quite high – more than 0.93.

*Table 1. Statistics of FR, SC, OI and SAT with reference to FB data (weekly data).*

	FR	SC	OI	SAT	FB
bias [°C]	-0.45	-0.34	-0.42	-0.31	0.00
RMSD [°C]	0.97	0.84	0.96	0.66	0.00
correlation	0.931	0.937	0.934	0.936	1.000
mean [°C]	10.99	11.11	11.03	11.13	11.45
standard deviation [°C]	4.35	4.45	4.48	4.57	4.19

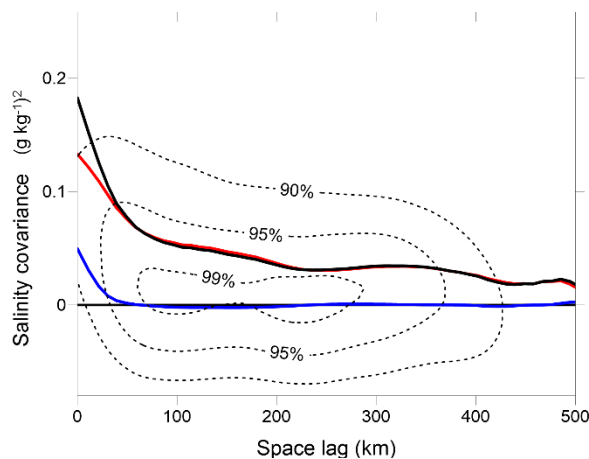
With reference to the SAT data, FR had RMSD = 0.96. Data assimilation reduced this value to 0.82 (SC) and 0.93 (OI).

## 3.2 Assimilation of in-situ data with EOF method

### 3.2.1 SST and SSS reconstruction with EOF

The main idea behind EOF is temporal covariance between distant locations.

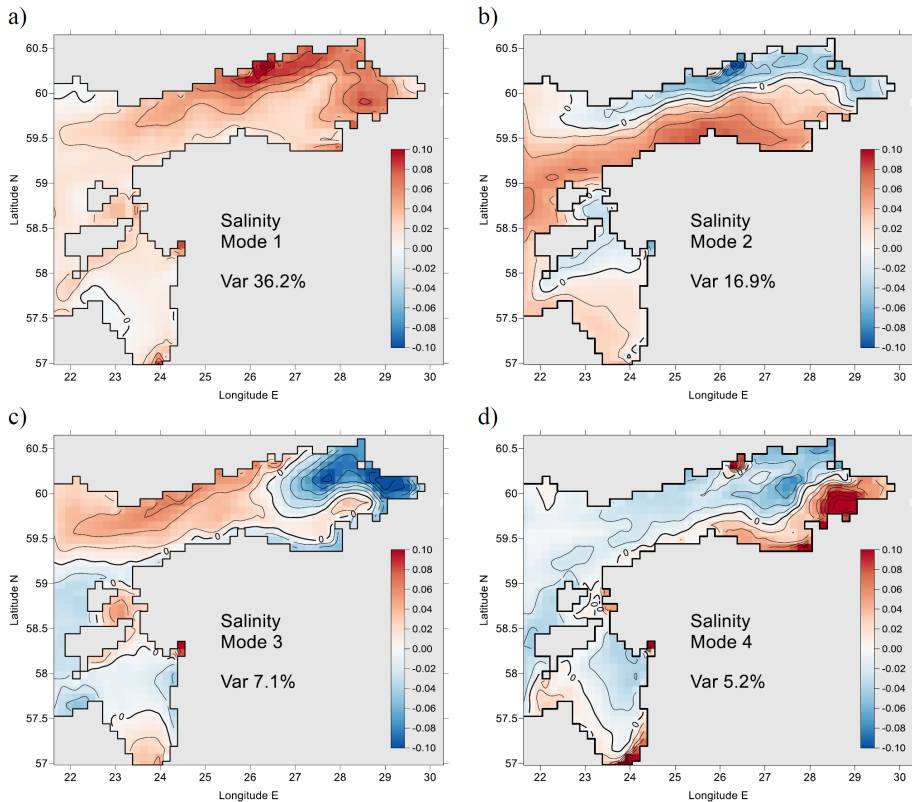
Covariance as a function of the space lag of original data (Fig. 5) usually did not follow the normal distribution neither did it exponentially decrease. Large covariance values, especially for SST (not shown), for remote wet points, are due to the similarity of their thermohaline regime (shallow coastal areas, zones influenced by river discharge, local basins with low-intensity circulation). The covariance of the most energetic modes followed closely at larger scales the covariance of original data. Covariance of residual fields (sum of the remaining higher EOF modes) had a strong normal distribution and it decayed fast with increasing space lag. Correlation (not shown) dipped 0.2 at a distance of 30 km for both SST and SSS, justifying the use of OI for this part of the variability.



**Figure 5.** Covariance of SSS as a function of space lag between the model points. Shown are heavily smoothed two-dimensional relative histograms of the original data (dotted lines, percentiles 90, 95, and 99 %) and mean covariance of original data (black line). The SSS covariance of the sum of the six most energetic EOF modes is also shown (red line) and higher EOF modes (blue line). (Paper II)

Remapping one-dimensional vectors  $e_k$  back into the two-dimensional geographical framework gives insight into underlying physical processes. First, most energetic modes (Fig. 6 shows the modes for SSS), had nearly “flat” patterns without sign change; their amplitudes were dominated by a seasonal signal. Higher modes were considered random due to mesoscale eddies and other fine-scale processes, therefore their correlation decayed rapidly with increasing distance. In the SST patterns, the first mode dominated heavily (97.64% of variance explained) due to the seasonal cycle. Regarding the SSS patterns, the share of different modes was more distributed and the first six modes explained 72.88% of the total variance. The details of the calculated modes and their interpretation can be found in Paper II (Table 1 and Table 2).

In order to make sure that the EOF algorithm will perform in the best possible way, experiments with pseudo-observations were conducted. It was found that ca 5000 observations were needed for producing reliable results. As input data for assimilation, model values at specified locations were taken instead of real observations (Paper II).



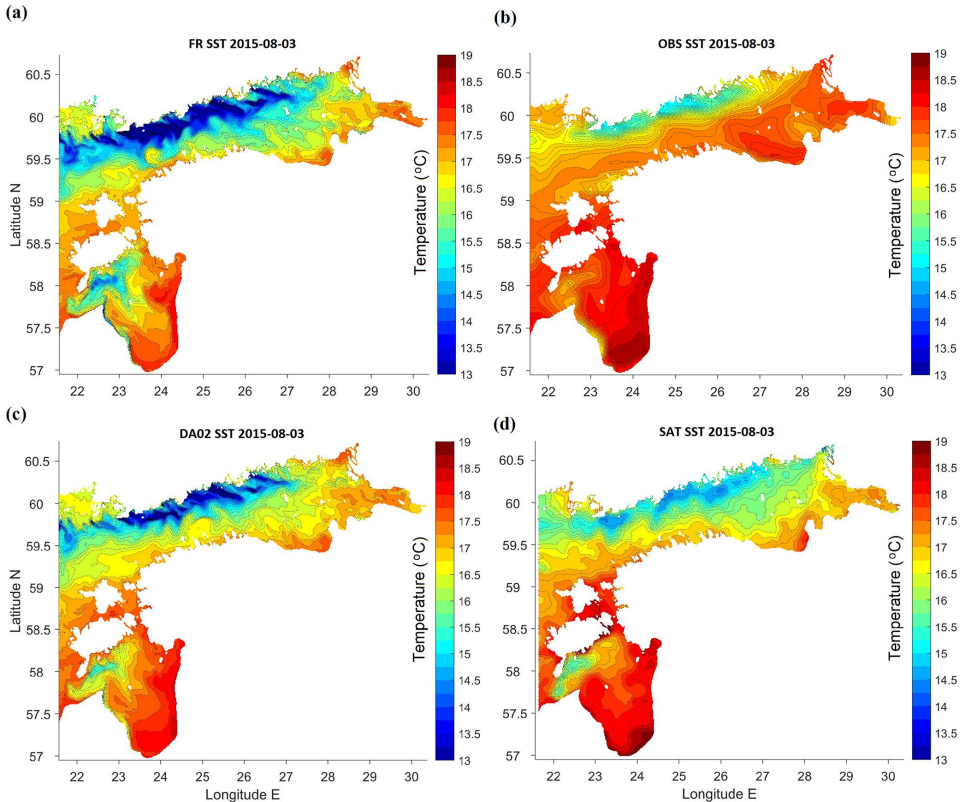
**Figure 6.** EOF patterns for the four first modes of SSS. Shown are the explained variance percentage of each mode. The contour interval for non-dimensional normalized modes (744 points) is 0.02. (Paper II)

With a decreasing number of observations  $K$ , errors slightly increased when  $K > L$ , where  $L$  is the number of modes. For example, SSS absolute error was  $< 0.3 \text{ g kg}^{-1}$  for 88% of cases with  $K = 51$  and 80% of cases with  $K = 10$ . Regarding SST, the errors were  $< 0.6 \text{ }^\circ\text{C}$  in 90% and 82% of cases, respectively. Regression of all the values of both SST and SSS yields tangent between initial and reconstructed data 0.99, their correlations follow  $r > 0.95$ . Relative errors of all the SST data, compared with the horizontal standard deviation of each time instance, were from 6.7% (observation grid step 37 km) to 8.6% (93 km). Relative errors of SSS were somewhat larger – 18% and 25%, respectively. For  $K < L$  the errors increased abruptly and singularity errors occurred in Eqs. (5)-(6).

In one of the experiments (Paper II), the whole region was split into three sub-regions: the Gulf of Finland, Gulf of Riga, and northeastern Baltic Proper (Fig. 1), and individual EOF modes were calculated for each of the sub-areas. Except for the northeastern Baltic Proper, the first two SST modes for the Gulf of Finland and Gulf of Riga were similar to the patterns obtained for the whole area. Pairwise correlations of the SST amplitudes were  $> 0.95$  between the Gulf of Finland, the Gulf of Riga and the whole region. There was no evidence that any particular season or sub-region was modelled with greater accuracy than this.

### 3.2.2 Assimilation using EOF reconstruction

When waters with an SST above 17 °C dominated the area, all the maps indicated moderate upwelling near the northern coasts of the basins (Fig. 7). However, the spatial extent of the colder waters and the minimum temperatures were different. Warmest “cold” waters were observed on satellite images.

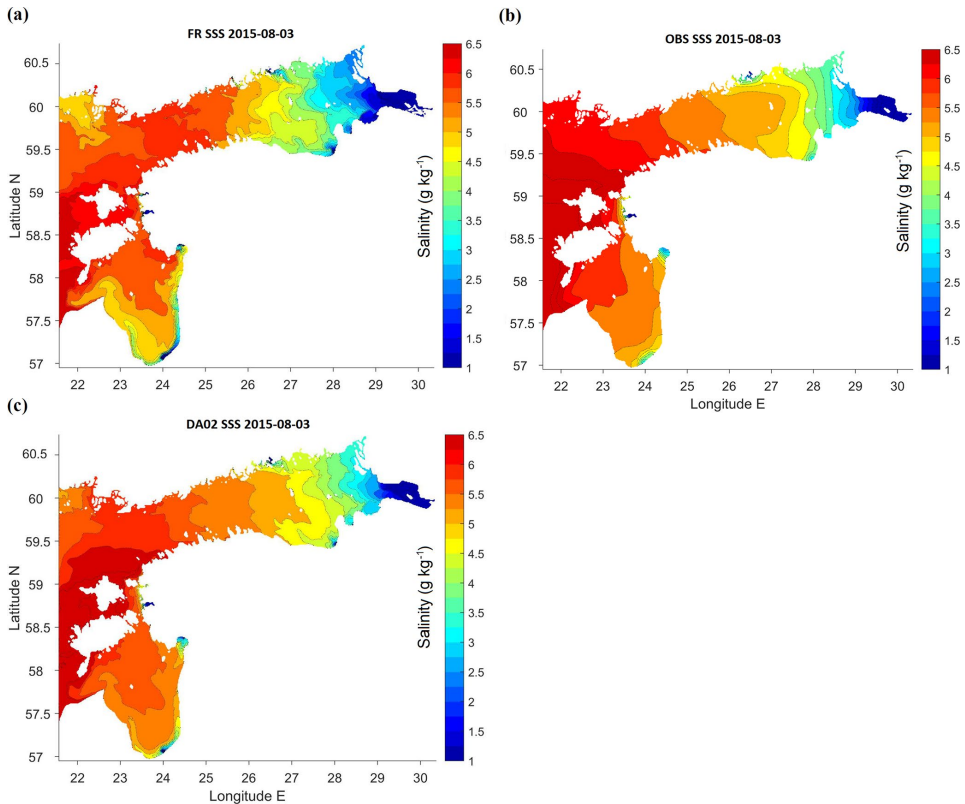


**Figure 7.** Maps (longitude E, latitude N) of the SST in the study area on 3 August 2015: (a) free model run without DA, (b) in situ observations reconstructed using EOF method, (c) DA with a relaxation time of 5 days (weight 0.2), (d) satellite observations. (Paper III)

Various mesoscale features emerged on SST (Fig. 7) and SSS (Fig. 8) maps, like colder upwelling filaments along the northern coasts of the Gulf of Finland and the Gulf of Riga, and decaying anticyclonic warm-core eddies near the southern coast of the Gulf of Finland.

The model forecasted spreading of the Daugava river waters by narrow coastal strips of lower salinity in the NE and NW directions (Fig. 8). Locations with dense observations allowed us to validate the model and visually evaluate assimilation quality. While SST followed the seasonal cycle (Fig. 9), with weather-dependent deviations, then SSS behavior was more irregular. All the compared SST data sources showed less variability than that of SSS in the given variation scales of SST and SSS (16 °C and 2 g·kg<sup>-1</sup> respectively). Still, for most of the time the assimilation curve (blue line, Fig. 9b) was closer to the FerryBox observations than the control run, for both SST and SSS (Fig. 9).





**Figure 8.** Maps (longitude E, latitude N) of SSS in the study area on 3 August 2015: (a) free model run without DA, (b) in situ observations reconstructed using EOF method, (c) DA with relaxation time 5 days (weight 0.2). (Paper III)

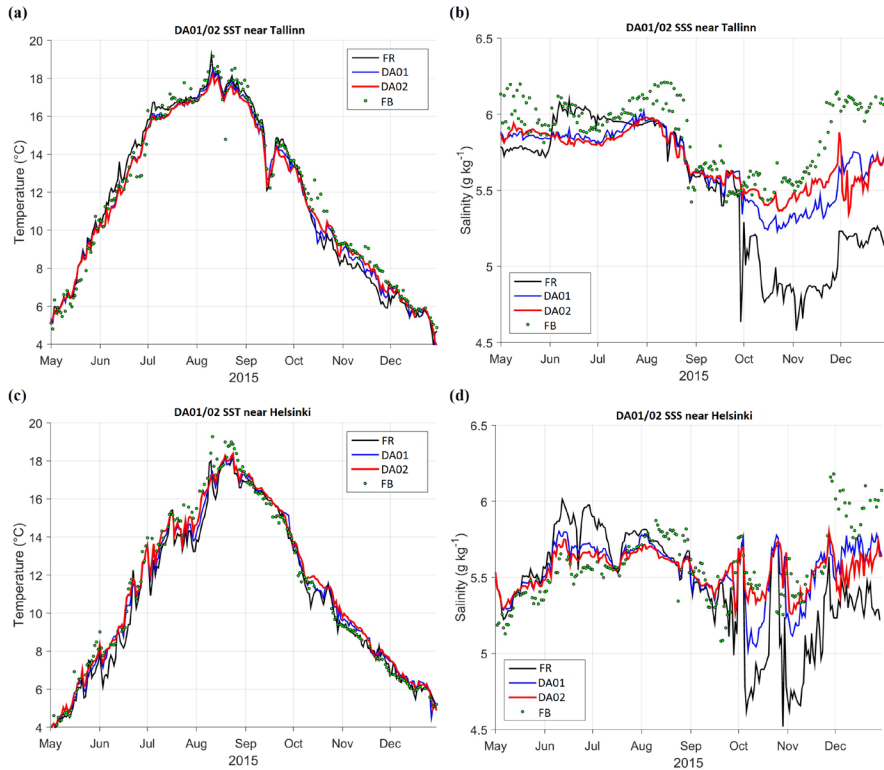
DA improved the model performance significantly: centered RMSD of SST was reduced by 22% and SSS by 34%, compared to the control run (Table 2). From DA01 to DA02, slight improvement of DA performance was observed, therefore DA02 was adopted. Spatial pattern of RMSD changes between the DA and FR (Fig. 10) indicates that most significant improvement (RMSD reduction up to 50%), both for SST and SSS, took place in the observation-covered areas in the Gulf of Finland. Too cold waters produced by FR near the northern coast of the Gulf of Finland were effectively corrected by DA (see also Fig. 7), therefore highest improvement percentage scores were detected in this region. Near the western open boundary, non-assimilated SST and SSS values of the larger model were advected into the area, therefore RMSD reduction was small, or even negative for SSS.

Daily maps of EOF DA data were converted also to weekly averages, as it was done for OI and SC methods in Table 1 and Fig. 4, see details in Paper I. Weekly RMSD is 13% lower for SST and 9% for SSS, compared to the daily data (Table 2). Weekly statistics suppressed the mesoscale variability, not resolved by EOF reconstruction, and revealed better match between the DA and the observations. DA decreased the bias, especially for SSS, and increased correlation of SSS between analysis and observations.

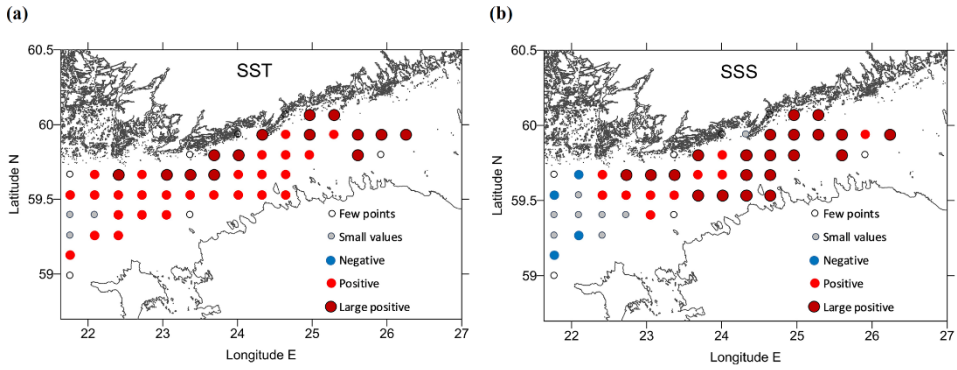


Table 2: Statistics of daily data in  $0.5' N \times 1' E$  grid cells with FerryBox (FB) observations: free model run without data assimilation (FR), data assimilation DA01 (observation weight 0.1), DA02 (weight 0.2) and FB. Bias, RMSD and correlation are taken with reference to FB. (Paper III)

	FR	DA01	DA02	FB
<b>SST [°C]</b>				
Mean	12.03	12.15	12.25	12.48
Standard deviation	3.98	3.92	3.93	3.97
Bias	-0.45	-0.33	-0.23	0
RMSD	0.72	0.59	0.56	0
Correlation	0.98	0.99	0.99	1.00
<b>SSS [g kg<sup>-1</sup>]</b>				
Mean	5.61	5.79	5.85	5.93
Standard deviation	0.35	0.29	0.31	0.37
Bias	-0.31	-0.14	-0.08	0
RMSD	0.35	0.24	0.23	0
Correlation	0.52	0.76	0.78	1.00



**Figure 9.** Time series of SST (a, c) and SSS (b, d) near Tallinn (a, b, 59.4833° N, 24.7667° E) and Helsinki (c, d, 59.9500° N, 24.8833° E), locations shown in Fig. 2a. FerryBox data are shown by dots, black lines represent control run without DA, red lines correspond to DA with relaxation time 5 days (weight of observations 0.2), blue lines for 10 days (weight 0.1). (Paper III)



**Figure 10.** Improvement of RMSD of DA compared to that of FR, both taken in reference to 110 k FerryBox observations. Comparison is made for 20 x 20 grid cells (10' N x 20' E) for SST (a) and SSS (b) over the whole study period. Legend codes: few points – less than 100 observations in a box, small values - absolute percentage change less than 10%, negative – DA RMSD growth more than 10%, positive – DA improvement (RMSD reduction) from 10% to 30%, large positive – improvement more than 30%. (Paper III)

## 4 Conclusion

Three different assimilation methods were used with HBM-EST simulations – OI, SC and EOF. Calculation was performed on model data for year 2015 and satellite observations provided by CMEMS, validation was made with in situ data. The findings can be summarized as follows:

- The tests with marine data assimilation into the model of the northeastern Baltic showed that satellite SST products from CMEMS can be well used for data assimilation in the sub-regional marine forecasts (Paper I).
- SST from the DA, produced using satellite data, was compared with FB observations; it was confirmed that for areas with dense observations OI and SC algorithms are appropriate and can be used to produce good quality results (Paper I).
- Estimation of the model performance for both cases with or without DA gave for the SC and OI methods similar pattern of differences relative to FB data. Comparing the results by SC and OI methods, SC produced slightly more accurate results than OI with the given set of parameters. Namely, SC provided smaller bias and RMSD and larger correlation as compared to the FB data. (Paper I)
- Statistically justified EOF reconstruction method is developed that allows to handle large-scale patterns of observed fields in the sub-regions. The entire region can be covered with interpolated and extrapolated observations using model-based EOF patterns. Summation of all modes yields initial field variance (Paper II).
- Study of the EOF reconstruction method revealed that in the smaller sea regions, which are affected by the same large-scale forcing patterns, the dominating EOF patterns have obvious physical interpretations. Their shape does not depend very much on the selection of boundaries as shown by experiments with split regions (Paper II).
- Implementation of the EOF statistical reconstruction technique into the DA of the forecast model yielded reduction of RMSD and interpolation errors. Mesoscale deviations from large-scale EOF patterns follow well-defined covariance decay with space lag; therefore, they could be treated by optimal interpolation or similar method (Paper III).
- EOF DA method has small computational effort compared to the localized methods like OI etc (Paper III).
- Intermediate results are in the form of maps that are easily understandable and can be checked visually or taught to be analysed by artificial intelligence (Paper III).
- Since the quality of DA and forecast are primarily determined by the quality of EOF reconstruction (when extensive mesoscale observations are not available), then it would be possible to perform faster calculations with orthogonal EOF basis vectors (Paper III).
- DA made major improvement in modelling of SSS. Further improvements can be made as RMSD to the observations makes 62% of observed standard deviations (Paper III).

## 5 References

- Alenius, P., Myrberg, K. and Nekrasov, A., 1998. The physical oceanography of the Gulf of Finland: a review. *Boreal Environ. Res*, 3(2), pp. 97-125.
- Alenius, P., Nekrasov, A. and Myrberg, K., 2003. Variability of the baroclinic Rossby radius in the Gulf of Finland. *Continental Shelf Research*, 23(6), pp. 563-573.
- Arra, V., 2018. Storm Frequency in the Northern Baltic Sea Region and its Association to the North Atlantic Oscillation.
- Axell, L., 2013. BSRA-15: A Baltic Sea Reanalysis 1990–2004. SMHI.
- Axell, L. and Liu, Y., 2016. Application of 3-D ensemble variational data assimilation to a Baltic Sea reanalysis 1989–2013. *Tellus A: Dynamic Meteorology and Oceanography*, 68(1), p. 24220.
- Beckers, J.M. and Rixen, M., 2003. EOF calculations and data filling from incomplete oceanographic datasets. *Journal of Atmospheric and oceanic technology*, 20(12), pp. 1839-1856.
- Berg, P. and Poulsen, J.W., 2012. Technical Report 12-11 Implementation details for HBM. *Danish Meteorological Institute, Ministry for Climate and Energy*.
- Bergström, S. and Carlsson, B., 1994. River runoff to the Baltic Sea-1950-1990. *Ambio*, 23(4-5), pp. 280-287
- Bonekamp, H., Montagner, F., Santacesaria, V., Loddo, C.N., Wannop, S., Tomazic, I., O'Carroll, A., Kwiatkowska, E., Scharroo, R. and Wilson, H., 2016. Core operational Sentinel-3 marine data product services as part of the Copernicus Space Component. *Ocean Science*, 12(3) 787-795.
- Canizares, R., Madsen, H., Jensen, H.R. and Vested, H.J., 2001. Developments in operational shelf sea modelling in Danish waters. *Estuarine, Coastal and Shelf Science*, 53(4), pp. 595-605.
- Cressman, G.P., 1959. An operational objective analysis system. *Mon. Wea. Rev*, 87(10), pp. 367-374.
- Crosnier, L. and Le Provost, C., 2007. Inter-comparing five forecast operational systems in the North Atlantic and Mediterranean basins: *The MERSEA-strand1 Methodology*. *Journal of Marine Systems*, 65(1-4), pp. 354-375.
- Elken, J., and W. Matthäus. "Baltic Sea oceanography, [in:] *Assessment of climate change for the Baltic Sea basin*, Annex A. 1.1, H. von Storch & A. Omstedt (eds.), BALTEX Publ." (2008).
- Elken, J., Raudsepp, U. and Lips, U., 2003. On the estuarine transport reversal in deep layers of the Gulf of Finland. *Journal of Sea Research*, 49(4), pp. 267-274.
- Fu, W., Høyer, J.L. and She, J., 2011a. Assessment of the three dimensional temperature and salinity observational networks in the Baltic Sea and North Sea. *Ocean Science*, 7(1), p. 75.
- Fu, W., She, J. and Zhuang, S., 2011b. Application of an Ensemble Optimal Interpolation in a North/Baltic Sea model: assimilating temperature and salinity profiles. *Ocean Modelling*, 40(3), pp. 227-245.
- Fu, W., She, J. and Dobrynin, M., 2012. A 20-year reanalysis experiment in the Baltic Sea using three-dimensional variational (3DVAR) method. *Ocean Science*, 8(5).
- Fu, W., 2016. On the Role of Temperature and Salinity Data Assimilation to Constrain a Coupled Physical–Biogeochemical Model in the Baltic Sea. *Journal of Physical Oceanography*, 46(3), pp. 713-729.

- Funkquist, L., 2006. An operational data assimilation system for the Baltic Sea. *European Operational Oceanography: Present and Future*, p. 656-660.
- Gandin, L. S. 1963. Objective analysis of meteorological fields. *Israel Program for Scientific Translations*.
- Ghil, M., 1989. Meteorological data assimilation for oceanographers. Part I: Description and theoretical framework. *Dynamics of Atmospheres and Oceans*, 13(3-4), pp. 171-218.
- Ghil, M., and Malanotte-Rizzoli, P. 1991. Data assimilation in meteorology and oceanography. *Advances in geophysics Vol. 33*. 141-266.
- Gregg, W.W., Friedrichs, M.A., Robinson, A.R., Rose, K.A., Schlitzer, R., Thompson, K.R. and Doney, S.C., 2009. Skill assessment in ocean biological data assimilation. *Journal of Marine Systems*, 76(1-2), pp. 16-33.
- Hernandez, F., Blockley, E., Brassington, G.B., Davidson, F., Divakaran, P., Drévilion, M., Ishizaki, S., Garcia-Sotillo, M., Hogan, P.J., Lagemaa, P. and Levier, B., 2015. Recent progress in performance evaluations and near real-time assessment of operational ocean products. *Journal of Operational Oceanography*, 8(sup2), pp. s221-s238.
- Holland, W.R. and Malanotte-Rizzoli, P., 1989. Assimilation of altimeter data into an ocean circulation model: Space versus time resolution studies. *Journal of physical oceanography*, 19(10), pp. 1507-1534.
- Høyer, J.L. and She, J., 2007. Optimal interpolation of sea surface temperature for the North Sea and Baltic Sea. *Journal of Marine Systems*, 65(1-4), pp. 176-189.
- Ide, K., Courtier, P., Ghil, M. and Lorenc, A.C., 1997. Unified notation for data assimilation: Operational, sequential and variational (gtspecial issue\data assimilation in meteorology and oceanography: Theory and practice). *Journal of the Meteorological Society of Japan. Ser. II*, 75(1B), pp. 181-189.
- Ivanov, S.V., Kosukhin, S.S., Kaluzhnaya, A.V. and Boukhanovsky, A.V., 2012. Simulation-based collaborative decision support for surge floods prevention in St. Petersburg. *Journal of Computational Science*, 3(6), pp. 450-455.
- Jevrejeva, S., Drabkin, V.V., Kostjukov, J., Lebedev, A.A., Leppäranta, M., Mironov, Y.U., Schmelzer, N. and Sztobryn, M., 2004. Baltic Sea ice seasons in the twentieth century. *Climate Research*, 25(3), pp. 217-227.
- Johansson, J., 2017. Total and regional runoff to the Baltic Sea. *Baltic Sea environment fact sheet*, available at: <http://www.helcom.fi/baltic-sea-trends/environment-fact-sheets/>, last access: April 2020.
- Kaplan, A., Kushnir, Y., Cane, M.A. and Blumenthal, M.B., 1997. Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *Journal of Geophysical Research: Oceans*, 102(C13), pp. 27835-27860.
- Kikas, V. and Lips, U., 2016. Upwelling characteristics in the Gulf of Finland (Baltic Sea) as revealed by Ferrybox measurements in 2007–2013. *Ocean Science*, 12(3), pp. 843-859.
- Kim, K.Y., 1997. Statistical interpolation using cyclostationary EOFs. *Journal of climate*, 10(11), pp. 2931-2942.
- Kleine, E. 1994. *Das Operationelle Modell des BSH für Nordsee und Ostsee. Konzeption und Übersicht*, Bundesamt für Seeschifffahrt und Hydrographie (manuscript report).
- Köuts, T. and Omstedt, A., 1993. Deep water exchange in the Baltic Proper. *Tellus A: Dynamic Meteorology and Oceanography*, 45(4), pp. 311-324.

- Laanemets, J., Väli, G., Zhurbas, V., Elken, J., Lips, I. and Lips, U., 2011. Simulation of mesoscale structures and nutrient transport during summer upwelling events in the Gulf of Finland in 2006. *Boreal environment research*, 16(Suppl. A), pp. 15-26.
- Lagemaa, P., 2012. *Operational forecasting in Estonian marine waters*. Tallinn: TUT Press.
- Legrand, C., Fridolfsson, E., Bertos-Fortis, M., Lindehoff, E., Larsson, P., Pinhassi, J. and Andersson, A., 2015. Interannual variability of phyto-bacterioplankton biomass and production in coastal and offshore waters of the Baltic Sea. *Ambio*, 44(3), pp. 427-438.
- Lehmann, A. and Hinrichsen, H.H., 2000. On the wind driven and thermohaline circulation of the Baltic Sea. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 25(2), pp. 183-189.
- Leppäranta, M. and Myrberg, K., 2009. *Physical oceanography of the Baltic Sea*. Springer Science & Business Media.
- Liblik, T., Laanemets, J., Raudsepp, U., Elken, J. and Suhhova, I., 2013. Estuarine circulation reversals and related rapid changes in winter near-bottom oxygen conditions in the Gulf of Finland, Baltic Sea. *Ocean Science*, 9(5).
- Lilover, M.J., Lips, U., Laanearu, J. and Liljebldh, B., 1998. Flow regime in the Irbe Strait. *Aquatic Sciences*, 60(3), pp. 253-265.
- Lips, U., Lips, I., Kikas, V. and Kuvaldina, N., 2008, May. Ferrybox measurements: a tool to study meso-scale processes in the Gulf of Finland (Baltic Sea). In *2008 IEEE/OES US/EU-Baltic International Symposium* (pp. 1-6). IEEE.
- Lips, U., Zhurbas, V., Skudra, M. and Väli, G., 2016. A numerical study of circulation in the Gulf of Riga, Baltic Sea. Part I: Whole-basin gyres and mean currents. *Continental Shelf Research*, 112, pp. 1-13.
- Liu, Y., Zhu, J., She, J., Zhuang, S., Fu, W. and Gao, J., 2009. Assimilating temperature and salinity profile observations using an anisotropic recursive filter in a coastal ocean model. *Ocean Modelling*, 30(2), pp. 75-87.
- Liu, Y., Meier, H.E. and Axell, L., 2013. Reanalyzing temperature and salinity on decadal time scales using the Ensemble Optimal Interpolation data assimilation method and a 3D ocean circulation model of the Baltic Sea. *Journal of Geophysical Research: Oceans*, 118(10), pp. 5536-5554.
- Liu, Y., Meier, H.M. and Eilola, K., 2014. Improving the multiannual, high-resolution modelling of biogeochemical cycles in the Baltic Sea by using data assimilation. *Tellus A*, 66.
- Lorenc, A.C., 1986. Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 112(474), pp. 1177-1194.
- Lorenz, E.N., 1956. *Empirical orthogonal functions and statistical weather prediction*. Cambridge, Massachusetts: MIT.
- Losa, S.N., Danilov, S., Schröter, J., Nerger, L., Maßmann, S. and Janssen, F., 2012. 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, pp. 152-162.
- Losa, S.N., Danilov, S., Schröter, J., Janjić, T., Nerger, L. and Janssen, F., 2014. Assimilating NOAA SST data into 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*, 129, pp. 259-270.

- Maljutenko, I., 2019. *Water Circulation in Gulf Type Regions of Freshwater Influence - the Gulf of Finland and Gulf of Riga*. Tallinn: TUT Press.
- Maljutenko, I. and Raudsepp, U., 2019. Long-term mean, interannual and seasonal circulation in the Gulf of Finland —the wide salt wedge estuary or gulf type ROFI. *Journal of Marine Systems*, 195, pp. 1-19.
- Männik, A. and Merilain, M., 2007. Verification of different precipitation forecasts during extended winter-season in Estonia. *HIRLAM Newsletter*, 52, pp. 65-70.
- Menemenlis, D., Fieguth, P., Wunsch, C. and Willsky, A., 1997. Adaptation of a fast optimal interpolation algorithm to the mapping of oceanographic data. *Journal of Geophysical Research: Oceans*, 102(C5), pp. 10573-10584.
- Moore, A.M. and Reason, C.J., 1993. The response of a global ocean general circulation model to climatological surface boundary conditions for temperature and salinity. *Journal of physical oceanography*, 23(2), pp. 300-328.
- Nowicki, A., Dzierzbicka-Głowacka, L., Janecki, M. and Kałas, M., 2015. Assimilation of the satellite SST data in the 3D CEMBS model. *Oceanologia*, 57(1), pp. 17-24.
- Raudsepp, U. and Elken, J., 1999. Application of the Bryan-Cox-Type Ocean Model to reproduce synoptic and mesoscale variability of the Irbe Strait salinity front. *Deutsche Hydrografische Zeitschrift*, 51(4), pp. 477-488.
- Ravichandran, M., Behringer, D., Sivareddy, S., Girishkumar, M.S., Chacko, N. and Harikumar, R., 2013. Evaluation of the global ocean data assimilation system at INCOIS: the tropical Indian Ocean. *Ocean Modelling*, 69, pp. 123-135.
- Rodhe, J. 1998. The Baltic and North Seas: a process-oriented review of the physical oceanography. *The sea*, 11 699-732.
- She, J., Høyer, J.L. and Larsen, J., 2007. Assessment of sea surface temperature observational networks in the Baltic Sea and North Sea. *Journal of Marine Systems*, 65(1), pp. 314-335.
- Soomere, T., 2003. Anisotropy of wind and wave regimes in the Baltic Proper. *Journal of Sea Research*, 49(4), pp. 305-316.
- Soomere, T. and Keevallik, S., 2003. Directional and extreme wind properties in the Gulf of Finland. *Proc. Estonian Acad. Sci. Eng*, 9(2), pp. 73-90.
- Soomere, T., Delpeche, N., Viikmäe, B., Quak, E., Meier HEM, Döös K. 2011. Patterns of current-induced transport in the surface layer of the Gulf of Finland. *Boreal Environmental Research* 16 (SUPPL. A), 49-63.
- Soosaar, E., Maljutenko, I., Uiboupin, R., Skudra, M. and Raudsepp, U., 2016. River bulge evolution and dynamics in a non-tidal sea-Daugava River plume in the Gulf of Riga, Baltic Sea. *Ocean Science*, 12(2).
- Sørensen, J.V.T., Madsen, H. and Madsen, H., 2004. Efficient Kalman filter techniques for the assimilation of tide gauge data in three-dimensional modeling of the North Sea and Baltic Sea system. *Journal of Geophysical Research: Oceans*, 109(C3).
- Stramska, M. and Białogrodzka, J., 2015. Spatial and temporal variability of sea surface temperature in the Baltic Sea based on 32-years (1982–2013) of satellite data. *Oceanologia*, 57(3), pp. 223-235.
- Taylor, K.E., 2001. Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research: Atmospheres*, 106(D7), pp. 7183-7192.
- Uiboupin, R. and Laanemets, J., 2009. Upwelling characteristics derived from satellite sea surface temperature data in the Gulf of Finland, Baltic Sea. *Boreal Environment Research*, 14(2).

- Vihma, T. and Haapala, J., 2009. Geophysics of sea ice in the Baltic Sea: A review. *Progress in Oceanography*, 80(3-4), pp. 129-148.
- von Storch, H. and Zwiers, F.W. 1999. *Statistical Analysis in Climate Research*. Cambridge: Cambridge University Press.
- Yurkovskis, A., Wulff, F., Rahm, L., Andruzaitis, A. and Rodriguez-Medina, M., 1993. A nutrient budget of the Gulf of Riga; Baltic Sea. *Estuarine, Coastal and Shelf Science*, 37(2), pp. 113-127.
- Zhuang, S.Y., Fu, W.W. and She, J., 2011. A pre-operational three Dimensional variational data assimilation system in the North/Baltic Sea. *Ocean Science*, 7(6), pp. 771-781.



## **Acknowledgements**

I would like to thank both my supervisors Jüri Elken and Priidik Lagemaa for their help and guidance throughout these difficult years.

This study was supported by the Ph.D. programme and the institutional research funding IUT 19-6 of the Estonian Ministry of Education and Research.

A larger BAL MFC team within the EU projects MyOcean, MyOcean2, and MyOcean-FO did the development of the HBM model. This cooperation is highly acknowledged.

Sincere thanks to Lars Axell, who shared programme code and supplementary documentation for OI, and to Germo Väli for reviewing the thesis.

I also express my gratitude to the Department of Marine Systems.

## Abstract

### Development of data assimilation for forecasts in Estonian marine areas

This research was performed in order to solve practical problems concerning quality of operational marine forecast for Estonian coastal waters. This was done through data assimilation using available observation data from satellites and FerryBox.

The main objective of the work was to find assimilation algorithm which allows to produce forecast with smaller errors compared to model run without assimilation. Observation coverage and spatiotemporal characteristics can vary significantly, but this should not impact quality of the assimilation.

HBM model was used for experiments as it currently produces operational forecast for Estonian marine areas. The region was limited to northeastern part of Baltic Sea, including Gulf of Finland, Gulf of Riga and northeastern part of Baltic Proper. Temporal resolution of the model was taken 24 hours, spatial resolution is 0.5' N x 1' E.

Assimilation was performed only for the first water layer (depth 3m) since observations were taken from satellite or FerryBox. Experiments were performed with three algorithms.

Two of them are well-known straightforward methods – successive corrections (SC) and optimal interpolation (OI). These algorithms are computationally robust and can be used for areas with abundant observation coverage. Assimilation reduces RMSD of e.g. SST by 0.01–0.13 °C.

The novelty of the research is development of a new method, which allows to assimilate small amount of scattered observations and improve forecast over entire grid, including areas without observations. The algorithm is based on calculation of EOF modes of model fields and reconstruction of observations using small portions of data. This method uses grid transposition (from fine to coarse and then back to fine grid through bilinear interpolation) in order to reduce computational load.

New method can be used for assimilation for regions that are under similar forcing resulting in high covariance over large distances. Therefore variability of assimilated fields can be presented by limited number of dominating EOF modes. Calculation of EOF modes should be done with model that describes marine physical processes very well.

During all the experiments the same time period was used (01.05.2015–31.12.2015) in order to maintain comparability between different algorithms.

Feasibility study of the EOF assimilation method showed that the EOF patterns have obvious physical interpretations and their shape does not depend very much on the selection of boundaries of sub-regions. The output is presented as two-dimensional maps which can be interpreted in terms of underlying physics.

EOF DA method reduces RMSD for SST by 0.14–0.16 °C and SSS by 0.11–0.12 g·kg<sup>-1</sup>. Calculations of EOF modes were done with 5-year model data. First four modes can be used for assimilation, they explain 99.3% and 65.4% of variance for SST and SSS respectively.

EOF DA method has small computational effort compared to the localized methods and can be used in operational forecast at Department of Marine Systems.

## Lühikokkuvõte

### Eesti merealade prognoosisüsteemi arendamine vaatlusandmete assimileerimise abil

Käesolev doktoritöö kirjeldab uuringuid, mis viidi läbi eesmärgiga pakkuda täpsemat lahendust Eesti merealade operatiivsele prognoosisüsteemile. Arendamine hõlmas satelliitandmete ja kruisilaevade vaatlusandmete assimileerimist.

Peamine eesmärk oli leida arvutialgoritm, mis korrigeeriks prognoosi ning vähendaks vigu, mis on iga mudeli puhul paratamatud, võrreldes mudeli tulemustega ilma assimileerimiseta. Läbi selle saavutatakse prognoosi kõrgem täpsus, kuna arvutusi järgmiseks päevaks alustatakse korrigeeritud väärtusest. Vaatluste tihedus võib kõikuda üle kogu basseini, samuti võivad olla väga erinevad ajalised parameetrid, kuid see ei tohiks oluliselt mõjutada mudeli prognoosi kvaliteeti peale assimileerimist.

Eksperimentide jaoks kasutati HBM mudelit, mis töötab operatiivsel režiimil Eesti merealade jaoks. Uurimispiirkonnaks on Läänemere kirdeosa, mis hõlmab Soome ja Liivi lahte, Väinamerd ning Läänemere avaosa, mis jääb 21°E meridiaanist ida poole. Mudeli assimileerimise ajasamm oli võetud 24 tundi ning ruumiline võrgupesa suurus oli 0.5' N x 1' E.

Assimileerimine oli rakendatud ainult ülemise kihi jaoks (1. kihi paksus 3 m) kuna satelliitandmed olid pinnakihi kohta ning FerryBox andmed pärinesid sügavuselt 3–4 m. Eksperimentide käigus katsetati kolme erinevat algoritmi.

Nendest kaks on tuntud ja lihtsad algoritmid – järjestikulised muudatused (SC) ja optimaalinterpolatsioon (OI). Neile meetoditele on omane madal arvutusvõime, neid saab kasutada assimileerimisel piirkondades, mille kohta on rohkelt vaatlusi üle kogu basseini. Algoritmid vähendavad nt temperatuuri ruutkeskmist viga 0.01–0.13 °C võrra.

Töö uudsuseks on uue meetodi arendamine, mis lubab taastada vaatluste väärtusi üle kogu mudelivõrgu väheste vaatluste korral, mis ei tarvitse olla samas alambasseinis. Algoritm põhineb EOF moodide arvutamisel mudeliandmete alusel ning vaatluste rekonstrueerimisel, kasutades väikest andmekogust. Selle meetodi arvutusvõime tõhustamiseks kasutatakse jämevõrku, mis on peenvõrgust, kus toimub assimileerimine, 100 korda väiksema ruumilise resolutsiooniga, ning teisendus tagasi peenvõrgule on tehtud bilineaarse interpolatsiooniga.

Uus meetod sobib kasutamiseks assimileerimisel piirkondades, mis asuvad homogeense mõjuvälja alal, ja mille merevee parameetreid iseloomustab kõrge kovariatsioon suurtel kaugustel. Seega terve välja muutlikkust saab kirjeldada domineerivate moodide vähese arvuga. EOF moodide arvutus eeldab, et kasutatav mudel kirjeldab meres toimuvaid füüsikalisi protsesse küllaltki hästi.

Kõikides kirjeldatud eksperimentides oli kasutatud mudeliandmeid samast ajalisest perioodist (01.05.2015–31.12.2015), et tagada võrreldavus erinevate meetodite vahel.

EOF assimileerimismeetodi võimekuste uuring näitas, et tugevamatel EOF muustritel on füüsikaline tagapõhi ning nende kuju ei sõltu sellest, millises alambasseinis pärinevad lähteandmed. Assimileerimise väljundiks on kahedimensioonilised kaardid, mida saab tõlgendada füüsikaliste seaduspärasustega.

EOF assimileerimismeetod vähendab pinnatemperatuuri keskmist ruutviga 0.14–0.16 °C võrra ning pinnasoolsuse oma 0.11–0.12 g·kg<sup>-1</sup> võrra. EOF moodid olid arvutatud 5 aasta mudeliandmete põhjal. Esimesed neli moodi kirjeldavad üle 99,3% pinnatemperatuuri ja 65,4% pinnasoolsuse muutlikkusest.

Arendatud meetod on arvutuslikult tõhus ning sobib kasutamiseks operatiivses režiimis Tallinna Tehnikaülikooli Meresüsteemide instituudis.

# Appendix

## Paper I

Zujev, M. and Elken, J., 2018. Testing marine data assimilation in the northeastern Baltic using satellite SST products from Copernicus Marine Environment Monitoring Service. *Proceedings of the Estonian Academy of Sciences*, 67 (3), 217–230. [10.3176/proc.2018.3.03](https://doi.org/10.3176/proc.2018.3.03).





## Testing marine data assimilation in the northeastern Baltic using satellite SST products from the Copernicus Marine Environment Monitoring Service

Mihhail Zujev\* and Jüri Elken

Department of Marine Systems, Tallinn University of Technology, Akadeemia tee 15A, 12618 Tallinn, Estonia

Received 14 August 2017, accepted 2 January 2018, available online 5 June 2018

© 2018 Authors. This is an Open Access article distributed under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>).

**Abstract.** Satellite SST products from the Copernicus Marine Environment Service were tested for data assimilation in the sub-regional marine forecasts. The sub-regional setup of the HBM model was used in the northeastern Baltic, covering also the Gulf of Finland and the Gulf of Riga. Two assimilation methods – successive corrections and optimal interpolation – were implemented on the daily forecasts from April to December 2015. Independent daily FerryBox data from the ship track between Tallinn and Helsinki were used for validation. Higher SST forecast errors of the reference model were found near the shallower northwestern coasts. During the calm heating period in spring and early summer, the reference model produced in these regions too warm waters compared with the satellite and FerryBox observations. Too cold waters, compared to the observations, were modelled during the cooling period from late summer to winter. Although data assimilation reduced these errors, improving the treatment of coastal–offshore exchange in the core forecast model would be useful.

**Key words:** remote sensing, data assimilation, successive corrections, optimal interpolation, short-term forecast, HBM model, SST assimilation.

### 1. INTRODUCTION

Assimilation of observational results into oceanographic forecast models has a history of several decades, following with some delay developments of data assimilation in meteorology. In parallel to statistical forecast correction methods based on linear filtering and prediction theories (e.g. Kalman and Bucy, 1961), Cressman (1959) proposed a robust ‘manually tunable’ method directly applicable for correcting weather forecasts. Meteorology reached the state of working operational assimilation and forecast systems already in the 1970s (McPherson et al., 1979). In oceanography, only a few offshore regular time series observations have

been made. Shipborne observations, which provide most of the water column data, are non-synoptic and usually separated by a distance larger than the scale (size) of mesoscale motions. About a decade later than in meteorology, the first global oceanic data assimilation system (Derber and Rosati, 1989) was proposed based on sea surface temperature (SST) observations from merchant ships; quite sparse profile data from XBT, CTD, and Nansen bottles were incorporated as well.

Acquisition and assimilation of remote sensing data have been a common procedure for both meteorology and oceanography. However, the data coverage and accuracy are different for the atmosphere and the ocean. Reliable spaceborne thermal emissivity observations started in the 1960s on terrestrial (Buettnner and Kern, 1965) and ocean (Anding and Kauth, 1970) surfaces.

\* Corresponding author, [mihhail.zujev@ttu.ee](mailto:mihhail.zujev@ttu.ee)

Further, atmospheric infrared and microwave sounding allowed estimation of temperature and humidity profiles with height, also in the cloudy areas. Operational assimilation of remote sensing data into the weather forecast models was introduced in the 1970s, based on the adding of interpolated difference between observed and forecast values to the original forecast in order to obtain a corrected model state for the next forecast interval. The tests showed (e.g. Ghil et al., 1979) that the impact of data assimilation is highly sensitive to the quantity of data available; the choice of the assimilation method to determine the interpolation weights is also of importance.

Ocean sea surface temperature (SST) can be determined by most satellite sensors only in the cloud-free areas. Again, the amount of ocean data on the surface is irregular both in time and space as these are temperature–depth profile data; this causes significant problems in ocean data assimilation, compared with the more regular atmospheric observational data. Skin-layer corrected accurate (Donlon et al., 2002) and operationally available SST data that have high resolution both in space (<10 km) and time (6–12 h) have been synthesized and used for data assimilation since the beginning of the 2000s (Tang et al., 2004; Brasseur et al., 2005).

New remote sensing SST products were developed and made publicly available during the MyOcean project (Nardelli et al., 2013). This procedure is further developed and continued as the Copernicus Marine Environment Monitoring Service (CMEMS), <http://marine.copernicus.eu/>, providing Level 3 (or L3, supercollated or merged multisensor data, different options are available) and Level 4 (or L4, gap-free, interpolated from L3) SST products (Martin et al., 2012) in numerical format. These data open up new possibilities for operational data assimilation.

Statistical methods for data assimilation have many variants (e.g. Ide et al, 1997), which are all based on the estimated spatial–temporal correlation functions and variance fields. While optimal interpolation, 3DVAR, and 4DVAR methods assume in most applications prescribed statistical fields, then Kalman filters estimate and predict their variations depending on the evolution of oceanographic state variables. Good estimations are found when true ensemble forecasts can be made, with parallel forecasts starting from slightly modified initial conditions. Another option, demanding much less computing power, is to generate pseudo-ensembles from a single forecast. Sea level innovations, introduced by the assimilation procedure, propagate fast as barotropic long gravitational surface waves, therefore continuous assimilation with small increments during an assimilation cycle is advisable. Assimilation of temperature and/or salinity modifies the forecast density field, corres-

ponding perturbations propagate much slower as baroclinic internal waves or advective plumes; hence larger innovations are acceptable. When observations of different state variables are combined into the same forecast, multivariate optimal interpolation provides reliable results (Cummings, 2005).

In the Baltic Sea, probably the first data assimilation system was made by Sokolov et al. (1997) for ‘smart’ interpolation of temperature, salinity, and chemical profile data from monitoring stations, using a hydrodynamic model. Some tests have been devoted to assimilation of sea level data (Canizares et al., 2001; Sørensen and Madsen, 2004; Ivanov et al., 2012), based on the various options of the Kalman filter.

Assimilation of scalar variables such as temperature and salinity into the Baltic Sea models has a quite rich history. The present study has some specific features. Firstly, assimilation is designed into the operational forecast system and is prepared for the routine use; therefore, the methods must be robust and computationally effective. Secondly, the study is based on the downstream forecasts from the Baltic-wide core service system and both the model results and SST observations are of high spatial resolution. However, in the present study we deal with the validation of averaged data sets and do not consider high-resolution details.

Earlier, the operational assimilation system presented by Funkquist (2006) used 3D optimal interpolation (3D OI) for satellite and profile data. Correlation functions, needed for OI, were also estimated by Høyer and She (2007), She et al. (2007), and Fu et al. (2011a). Nowicki et al. (2015) used the Cressman method of successive corrections (SC) for satellite-based SST data. A number of pre-operational experiments have been conducted to test the performance of new assimilation methods: 3DVAR with isotropic (Zhuang et al., 2011) and anisotropic (Liu et al., 2009) recursive filters to estimate covariance functions, Ensemble Optimal Interpolation (Fu et al., 2011b), Singular Evolutive Interpolated Kalman Filter (Losa et al., 2012, 2014). Long-term reanalysis studies have used SC (Axell, 2013), 3DVAR (Fu et al., 2012; Fu, 2016), Ensemble Optimal Interpolation (Liu et al., 2013, 2014), and Ensemble 3DVAR (Axell and Liu, 2016).

This study is aimed at testing marine data assimilation into the operational high-resolution sub-regional forecast model of the northeastern Baltic, using new, routinely available satellite SST products from the CMEMS. The paper starts with the presentation of the model, data, and methods. The results section considers satellite data comparison with FerryBox, spatial, and seasonal features of assimilated data. Model skill estimates are given for the free run without assimilation, and for the model run with different options of

assimilation. We discuss possible ways to use less costly assimilation methods, yielding the results of nearly the same quality as with more sophisticated methods. Finally, conclusions are presented.

## 2. MODEL, DATA, AND METHODS

### 2.1. Sub-regional marine forecast model HBM

For assimilation tests we used the HBM-EST model (Lagemaa, 2012), which is an Estonian implementation of the HBM model (the abbreviation comes from HIROMB-BOOS Model). The model was originally constructed by the Bundesamt für Seeschifffahrt und Hydrographie, Hamburg, Germany, named as BSHCmod. Further developments have been made within Baltic-wide cooperation in operational forecasting; different options were merged to a new thoroughly tested HBM code within the EU MyOcean project. Details of the HBM model and its implementation are given by Berg and Poulsen (2012).

The HBM is a free-surface baroclinic 3D ocean model written in geographical latitude/longitude spherical coordinates that uses a horizontal staggered Arakawa C-grid and a fixed vertical grid with variable spacing and time-varying top layer thickness. There is also an option for dynamical vertical coordinates where the grid spacing changes in time. Vertical turbulence is treated by the  $\kappa$ - $\omega$  turbulence model and horizontal turbulence is treated within the well-known Smagorinsky formulation. The sea ice module is an integrated part of the HBM model, including both dynamics and thermodynamics. The model can take into account the results from independent wave models.

The forcing of the model is done by externally prescribed surface fields, point sources, and open boundary conditions. Surface fields are adopted from the numerical weather prediction model: 10-m wind components, mean sea level atmospheric pressure, surface air temperature, surface air humidity, and cloud cover. The HBM model calculates surface energy fluxes (mechanical, radiative, thermodynamic) using bulk parameterization formulae. Point source data are freshwater fluxes from rivers. If actual discharge forecasts are missing, the climatology for each calendar day will be used.

On the ocean side, the HBM model is forced by the tidal sea surface elevation, sea level forecasts from the barotropic storm surge model of the Northern Atlantic, and monthly climatological hydrography. The Baltic Sea implementation of the HBM uses a nested approach: the largest area of the 3D forecast covers the North and Baltic seas with a grid step of 12 nautical miles, the intermediate resolution with a grid step of 3 nautical

miles is further refined into a 1-mile grid covering the whole Baltic Sea. In the Danish Straits, two-way nesting with a finer grid model with the resolution of 0.5 nautical miles is used.

The Estonian implementation of the HBM (Lagemaa, 2012) covers the Baltic Sea sub-area east from 21.55°E (Fig. 1), including the Gulf of Finland and the Gulf of Riga, with the resolution of 0.5 nautical miles. The horizontal grid of the HBM-EST model consists of 425 by 529 grid points. The grid cell length by longitude is 1', by latitude it is 30". In the vertical grid, 39 depth layers are used, with a 3-m grid step near the surface and larger grid steps in the deeper layers. Forcing at the western open boundary is taken from the Baltic-wide HBM model, which operates routinely within the CMEMS with the resolution of 1 nautical mile. Forcing on the sea surface is obtained from the Estonian version of the HIRLAM model that is run by the national weather service for operational forecasts on a 11-km grid. For analysis of observation and forecast errors in relation to wind conditions, time series of wind speed components were extracted in the central part of the Gulf of Finland.

We chose the year 2015 for the forecasting and assimilation experiment. The forecasts were updated daily by introducing a new weather forecast at midnight.

### 2.2. Satellite SST data

Sea surface temperature data, observed from satellites, were used as input observational data within data assimilation. Gridded observation maps were obtained from the CMEMS multi-sensor product, which is built from bias-corrected L3 mono-sensor products at the horizontal resolution of 0.02 by 0.02 degrees. For each day a single SST value was used, which was reduced to midnight based on available observations at different times (near-real-time).

This product (Bonekamp et al., 2016) contains results from the merging of various satellite SST level 2 data, which have passed a significant number of quality controls and which have been calibrated through an inter-sensor bias correction procedure to provide an estimate of the night time SST based on original SST observations without any smoothing or interpolation. Details of the product are described on CMEMS web resource <http://cmems-resources.cls.fr/documents/QUID/CMEMS-OSI-QUID-010-009-a.pdf>.

Sensors used include METOP\_B, SEVIRI, VIIRS\_NPP, MODIS, and others. Observations were collected from different producers: NASA, NOAA, IFREMER, EUMETSAT OSI-SAF, and ESA.

Depending on the cloud cover there were from 200 up to 21 000 observations per day. Some obviously erroneous SST values (which differed more than 10 K



from model ones) were filtered out. All of them were used for the assimilation with the Cressman method. For optimal interpolation a data thinning algorithm was implemented, leaving one value for the area of 2.5 by 5 nautical miles.

An example of data extracts is shown in Fig. 2a, representing the observations on the line between Tallinn and Helsinki during the whole test period. The temperatures presented here are averaged over one week.

### 2.3. FerryBox data

Automatic observations made from ships crossing the sea areas were used as independent data for validation and quality assessment. FerryBox is a measurement system installed on board commercial ferries that collects temperature, salinity, chlorophyll *a* fluorescence, and turbidity data. This technology is used to study basin-scale temperature and salinity patterns together with mesoscale processes and upwellings (Kikas and Lips, 2016). The water is sampled at about 4 m below the surface at different rates and every 20 s the measurement is recorded, thus covering roughly 160–200 m in the horizontal direction. There are quality check procedures to eliminate unexpected and physically unrealistic values; cross-checking with the same data from the return trip is performed as well (Kikas and Lips, 2016). Detailed description of technical features of the FerryBox system is given by Lips et al. (2008).

Observations are available on the route Tallinn–Helsinki on the forth and back tracks twice a day with a time step of 20 s. Temperatures on the same latitude were averaged across multiple tracks. For each day one mean SST value was calculated regardless of the time of the day and the number of observations in the particular grid cell. Weekly averages of the temperatures observed by the FerryBox system are shown in Fig. 2b.

The data were taken as they are within the Copernicus system, in which the procedures include an advanced quality check. As our aim was to check the working of automatic systems, no additional quality control or processing was performed.

### 2.4. Data assimilation

Let us consider the model state represented by a vector  $\mathbf{x} = \{u, v, T, S, \rho, \xi, \dots\}$ , where the values in brackets are the model state variables (velocity components, temperature, salinity, water density, sea level, etc.), generally given at discrete grid points. When the model predicts the state vector  $\mathbf{x}_b$  on  $m$  grid points, it has some errors regarding the true state. Observations  $\mathbf{y}$

are taken usually at different  $n$  locations than  $\mathbf{x}_b$ . Assimilation is a procedure to create the analysis vector  $\mathbf{x}_a$  on the same set of coordinates as  $\mathbf{x}_b$  with a condition that by a given set of criteria,  $\mathbf{x}_a$  is closer to  $\mathbf{y}$  than  $\mathbf{x}_b$ . When the model state values  $\hat{\mathbf{y}}$  in the points of observations are obtained by an interpolation procedure, then the analysis is calculated from the innovation vector  $\mathbf{y} - \hat{\mathbf{y}}$  by the formula

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \hat{\mathbf{y}}), \quad (1)$$

where  $\mathbf{K}$  is the gain matrix containing  $m \times n$  weights for interpolation over  $1 \times n$  observation points. At individual model grid point  $i$  formula (1) can be written using the weight vector  $\mathbf{w}_i = (\mathbf{K})_i$ . In the following formulae we consider one state variable only and omit the index  $i$ , describing the specific model grid point. As a result we obtain

$$x_a = x_b + \mathbf{w}(\mathbf{y} - \hat{\mathbf{y}}) = x_b + \sum_{j=1}^n w_j (y_j - \hat{y}_j). \quad (2)$$

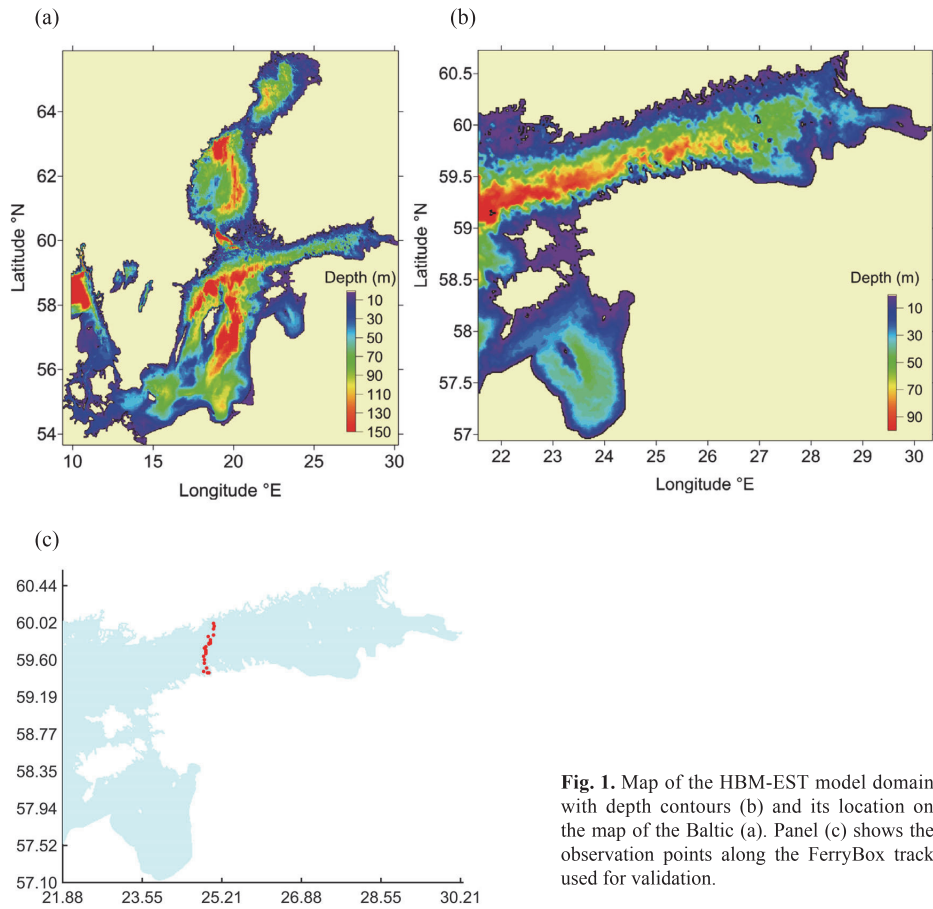
#### 2.4.1. Successive corrections

The successive correction method or the Cressman (1959) method assumes univariate relations between the state variables and that weights of the individual observations  $w_j$  in (2) decrease with the distance  $d_j$  between the observation point  $j$  and the model grid point  $i$ . Let us define the influence radius  $R$  around the model point  $i$ , where  $k$  observations out of total  $n$  observations are located. Good assimilation results are obtained with the weights given by the formula

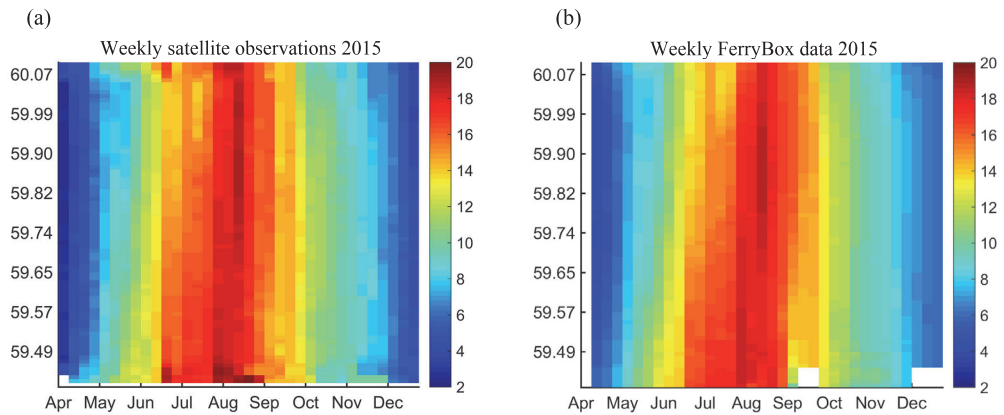
$$w_j = \frac{\max\left(0, \frac{R^2 - d_j^2}{R^2 + d_j^2}\right)}{\left[\sum_{j=1}^n \max\left(0, \frac{R^2 - d_j^2}{R^2 + d_j^2}\right) + \eta^2\right]}, \quad (3)$$

where the weights are positive within the influence radius and zero elsewhere. Reduction of the assimilation weights in real noisy conditions is done by introducing relative noise variance  $\eta^2$ , estimated from the variances of observation errors  $\sigma_o^2$  and background errors  $\sigma_b^2$ ,  $\eta^2 = \sigma_o^2 / \sigma_b^2$ . In the noiseless case ( $\eta^2 = 0$ ) the sum of weights is equal to unity.

Data assimilation for SST was made using a medium-scale value 37 km (20 nautical miles, 40 grid points) for the influence radius. This length is about ten times larger than the Rossby deformation radius. Therefore, the impact of individual mesoscale eddies is suppressed, but basin-scale SST features are kept. The weight function has the greatest impact within the nearest 5 km, then it goes almost linearly to zero for 37 km. Before



**Fig. 1.** Map of the HBM-EST model domain with depth contours (b) and its location on the map of the Baltic (a). Panel (c) shows the observation points along the FerryBox track used for validation.



**Fig. 2.** Time–latitude map of the weekly mean sea surface temperature observed from satellites (a) and FerryBox (b) between Tallinn (south) and Helsinki (north).

assimilation, the observations were averaged over each grid cell in order to avoid oversampling problems. During the testing of the scheme, the values of  $\sigma_a^2$  and  $\sigma_b^2$  were not known in advance. For the chosen data set we obtained acceptable results with  $R = 37$  km and  $\eta^2 = \sigma_a^2/\sigma_b^2 = 2$ . We used these values throughout the entire model run.

For computational efficiency, assimilation was performed in the two-dimensional domain for the surface layer only; deeper model levels were left unaffected. Since the observations were assimilated every day, the introduced innovations were moderate (compared to the vertical mixing) and not visible in the graphs of vertical profiles.

#### 2.4.2. Optimal interpolation

Optimal interpolation as developed by Gandin (1963) uses least-square minimization of analysis errors to calculate the weight coefficients  $w_j$  in expression (2). We follow the original point-wise presentation (see also Høyer and She, 2007) and denote  $f_j = y_j - \hat{y}_j$ ,  $f_0 = \tilde{x}_a - x_b$ . Here  $\tilde{x}_a$  is the ‘true’ unknown state in the model point  $i$ . The observations include random errors  $\varepsilon_j$ ; the error variance is  $\sigma_{\varepsilon_j}^2$ . The squared interpolation error, averaged over an ensemble,

$$E = \left[ f_0 - \sum_{j=1}^n w_j (f_j + \varepsilon_j) \right]^2 \Rightarrow \min \text{ is minimized with}$$

respect to  $w_j$ . This is achieved by setting  $n$  constraints for the derivatives  $\partial E / \partial w_j = 0$ , using the conditions  $\overline{f_j} = 0$ ,  $\overline{f_0} = 0$ ,  $\overline{\varepsilon_j} = 0$ ,  $\overline{\varepsilon_j f_j} = 0$ ,  $\overline{\varepsilon_j f_0} = 0$ . As a result, we obtain for the  $i$ th model point the following system of  $n$  linear equations regarding  $w_j$ :

$$\sum_{j=1}^n \overline{f_k f_j} w_j + \sigma_{\varepsilon_k}^2 w_j = \overline{f_k f_0}, \quad k = 1 \dots n, \quad (4)$$

which can be easily solved. By dividing equations (4) by the variance  $\sigma_f^2 = \overline{f_k^2}$ , we obtain correlation instead of spatial covariance. The weight coefficients of optimal interpolation are determined by the correlation matrix between the individual observation points  $\mathbf{B} = \{\overline{f_k f_j} / \sigma_f^2\}$  and the correlation vector between the observation point and the  $i$ th model point  $\mathbf{b} = \overline{f_k f_0} / \sigma_f^2$ , and by the relative variance of observation error  $\eta^2 = \sigma_{\varepsilon_k}^2 / \sigma_f^2$ . Equation (4) can be rewritten as  $\mathbf{w}(\mathbf{B} + \eta^2 \mathbf{I}) = \mathbf{b}$ , where  $\mathbf{I}$  is a unit matrix. The vector of weights is calculated in the form

$$\mathbf{w} = \mathbf{b}(\mathbf{B} + \eta^2 \mathbf{I})^{-1}. \quad (5)$$

More general matrix-vector formulations of optimal interpolation can handle also the case where observation errors may be correlated between each other and with

the background field (Lorenc, 1986; Ide et al., 1997). However, most of the practical implementations, like in our case, are limited to equations (4) with solution (5), where the spatial correlation is prescribed as a function with ‘tuned’ parameters.

Correlations  $\mathbf{B}$  and  $\mathbf{b}$  were approximated by the Gaussian function from the distance  $r$  between the correlated points. Anisotropic correlation features were taken into account by the directional distribution of the correlation scale from the angle  $\theta$  in the form of the ellipse dependence  $L = a \sin(\theta - \theta_0) + b \cos(\theta - \theta_0)$  relative to the reference angle  $\theta_0$ . So the correlation was adopted in the form  $B(r, \theta) = \exp(-r^2 / L^2(\theta))$ , where  $L = L(\theta)$  was pre-calculated in each model grid point according to the coastline and topography. Following the results by Høyer and She (2007), longer correlation scales were adopted along the coasts and the isobaths than in the perpendicular direction. The typical horizontal impact scale along the coast or isobath was chosen as 15 km. Standard deviations for the entire run were taken  $\sigma_a^2 = 0.5$  and  $\sigma_m^2 = 1.0$ .

Before performing the assimilation according to equations (2) and (5), the observations were filtered with a thinning algorithm to avoid oversampling and a huge computation amount, leaving up to  $n = 700$  points. The distance between the generated super-observations was kept at 10 km or more. For computational efficiency, wet points (located in the sea) were divided into 30 blocks to cover the entire basin. That leaves up to 81 observations per block. Observations from each block were cross-compared with all other observations in the same block plus the neighbouring observations falling into the adjacent area within a radius of 100 km, and the correlation coefficients were calculated.

#### 2.5. Methods for data quality and model skill assessments

There are a number of different SST data sets that can be compared. Here remote sensing SST products (SAT) from CMEMS are used for data assimilation. FerryBox observations (FB), carried out by the Department of Marine Systems but accessed from the Copernicus service, provide independent data for the assessment of the skill of data assimilation. Before estimating the skill of model versions (without data assimilation or with different assimilation options) in reference to one or another observational data set, the observations from different platforms are compared.

Following the approach by Taylor (2001), for a common comparison of the two variables  $f_n$  and  $g_n$  a common data set is defined where missing values of one or both data sets are ignored. If the standard deviations of the

data sets are  $\sigma_f$ ,  $\sigma_g$  and the coefficient of their mutual correlation is  $R_{f,g}$ , then the centred (with bias removed) root-mean-squared difference (RMSD) of the data sets  $E'$  will read

$$E'^2 = \sigma_f^2 + \sigma_g^2 - 2\sigma_f\sigma_g R_{f,g}. \quad (6)$$

The size of the original data sets is very different. Therefore we reduced in many cases the compared data sets to the FerryBox transect between Tallinn and Helsinki. Such data form a time–latitude matrix with a daily step in time and a 2.222 km step by latitude. For the data gap treatment we used the weekly average of all the available data. The missing values were just omitted from the averaging. Since satellite data were recalculated to the midnight values, for this comparison the forecast was treated as the weekly average of nightly (23 h advance) values. Other data set definitions are explained in the results section when necessary.

The model data sets are named as FR ('free' model run without data assimilation), SC (model run with assimilation of Copernicus SST with successive correction method), and OI (model run with assimilation of Copernicus SST with the optimal interpolation method). The model data sets have 71986 values for each time step.

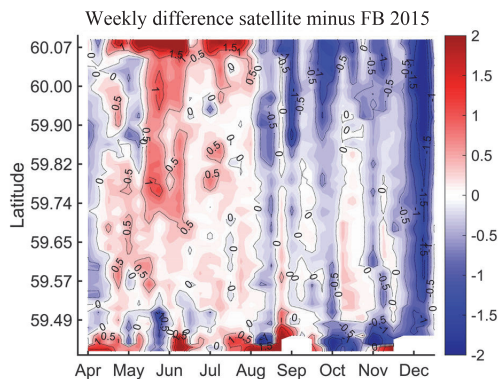
### 3. RESULTS

#### 3.1. Comparison of SST from satellite and FerryBox

There are principal differences between the temperature observed from satellites and from in situ sensors located a few metres below the surface. These observations give closer results in well-mixed conditions (Siegel et al., 2006; Uiboupin and Laanemets, 2015), which occur during stronger winds.

Here we compare averaged satellite observations from the Copernicus service (SAT) during 2015 along the FerryBox track with the in situ data of the latter (FB). Temperature values were merged by latitude, leaving out the impact of ship track variations. Several values for the same bin were replaced by their mean. The SAT data were provided as of midnight while the FB observations were made at different times during a particular day. The FB data represent the average temperature in the upper mixed layer, whose depth is variable. The data sets are independent of each other and have their maxima in August.

The difference between the two SST data sets is shown in Fig. 3. In central parts of the Gulf of Finland (latitudes 59.5–60 N) there is a clear seasonal behaviour. The thin layer temperature registered by the satellite was 0.3–0.7 K larger than the bulk temperature of the



**Fig. 3.** Weekly mean sea surface temperature difference between the satellite and FerryBox observations on a transect between Tallinn (south) and Helsinki (north) as a function of time and latitude.

upper layer (observed at 4 m depth) during spring and summer until August but insignificantly (less than 0.5 K) smaller in autumn and early winter. Larger SAT minus FB differences emerged occasionally in areas close to the coasts: a range 0.7–1.0 K was observed in Tallinn Bay and a larger range, 1.0–2.5 K, was found near Helsinki. In December the thin surface layer cooled down by 0.5–1.5 K more than the deeper surface layer along the whole transect, including also the coastal waters.

In most of the cases the difference between the two data sets (SAT minus FB) was of the same sign over the whole transect. This is consistent with the results by Uiboupin and Laanemets (2015), who studied similar data from 2000–2009. They found that during wind speeds less than 5 m/s different satellite sensors give up to 3 K larger SST than it is observed by FB; the difference is largest at smaller wind speeds of 2 m/s and less during temporary stratification of the thin surface layer. In our case with CMEMS data the difference was larger near the coasts, but both the coastal areas usually appeared in the SAT data either warmer or colder than it was found from the FB data. The reasons for large SST differences between satellite and in situ observations were discussed by Siegel et al. (2006). They also noted a seasonal behaviour as it is evident from our data presented in Fig. 3.

#### 3.2. Spatial patterns of SST

Sea surface temperature patterns manifest a variety of physical processes like different heating or cooling in coastal versus offshore areas, coastal upwelling, thermal fronts between the water masses, and signatures of mesoscale eddies and filaments.



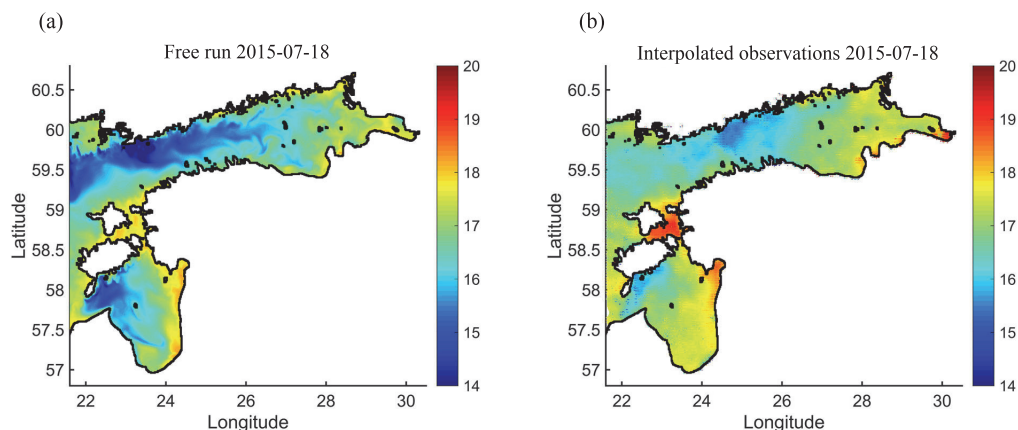


Fig. 4. Example of a HBM forecast (a) and satellite observations interpolated on the same grid (b). Both data from 2015-07-18.

We present an example of SST distributions for a summer date when there was good coverage of the sea area with satellite data. Comparison of the results of the model forecast (Fig. 4a) and remote sensing (Fig. 4b) reveals a different extent of warmer and colder areas of coastal waters. Model results show wide bands of colder water off the northern coasts, both in the Gulf of Finland and the Gulf of Riga. Satellite data registered only a small fraction of these colder water masses. In the warmer shallow areas, located between the Estonian larger islands and near the eastern coast of the Gulf of Riga, satellite observations yielded higher SST (by about 1 K) than the model. This can be partly attributed also to the wind-dependent positive bias of satellite data (Uiboupin and Laanemets, 2015). We note that the model reproduces in the above example more distinct mesoscale patterns than are evident from the satellite data.

### 3.3. Time series of SST

In the adopted data assimilation approach, SST satellite observations (example given in Fig. 4b) are used to correct the model forecast (example given in Fig. 4a). FerryBox data are independent and intended for validation. The data as defined in Section 2 are from observations (SAT and FB) and from models (FR, SC, and OI). The data are compared on the FerryBox transect (see Fig. 1c), where all the data are available. The model results are with a regular time interval (1 h), but daily observations have gaps.

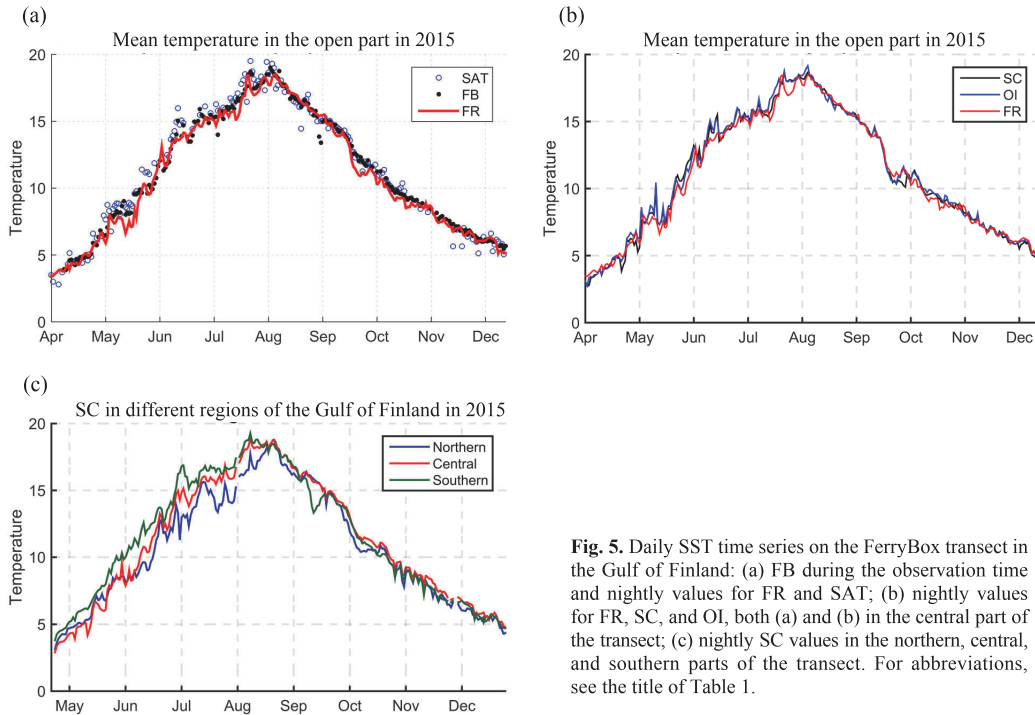
In the open part of the Gulf of Finland daily SAT data from CMEMS (Fig. 5a) revealed generally higher

temperatures than FB during the warming period (see also Fig. 3). The SAT data were spiky compared with the FB data: warmer spikes occurred during the warming period and colder spikes during the cooling. The FR forecast provided in the offshore waters slightly smaller SST than observed. Data assimilation using SC and OI ‘dragged’ the model results towards SAT observations, still the SST spikes did not appear in the assimilated model results.

As the cross-gulf SST pattern (Fig. 5b) shows, the southern part of the gulf warmed up faster than the central and northern parts, based on the results from SC-assimilated model data (see also weekly SAT and FB data in Fig. 2). This resulted in warmer by up to 5 K waters on a specific day. Cooling took place more uniformly across the research area, temperature differences were up to 1.5 K. Similar regional differences were evident in other data sets.

### 3.4. Skill assessment for non-assimilated and assimilated model results

Independent FB data obtained in a cross-section of the Gulf of Finland form the most comprehensive off-shore in situ data set within the forecast area. In the following we compare the assimilated model results obtained using the two methods with the FB data. For the non-assimilated model data (FR, reference run) we present comparisons for model validation. Since the two observations, SAT and FB, have different SST values (see Section 3.1), comparison is also made in reference to remote sensing data that were used in the assimilation process.



**Fig. 5.** Daily SST time series on the FerryBox transect in the Gulf of Finland: (a) FB during the observation time and nightly values for FR and SAT; (b) nightly values for FR, SC, and OI, both (a) and (b) in the central part of the transect; (c) nightly SC values in the northern, central, and southern parts of the transect. For abbreviations, see the title of Table 1.

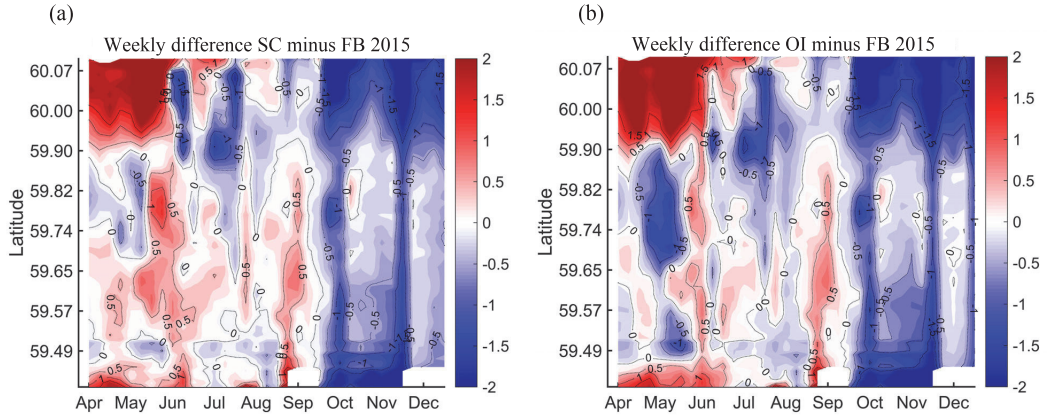
Further we consider temporal evolution of cross-gulf SST patterns in the Gulf of Finland on the basis of weekly average transects. Since the SAT data are from midnight, we used nightly model data for finding the weekly averages. The difference between the daily mean and nightly model values (not shown) due to the diurnal cycle was up to 0.5 K from April to August and slightly negative from September to December.

Both the SC and OI methods gave similar patterns of differences relative to FB data (Fig. 6). These patterns are similar to that of FR minus FB (not shown), but the variations of difference have been reduced by the data assimilation. Greater differences were found in the northern part of the Gulf of Finland, where a positive difference was evident from April to August (although the SC method gave some shorter negative differences as well) and a negative difference from September to December. In the central part of the gulf data assimilation corrected the errors of FR effectively and the forecast absolute difference from the FB data was in most cases less than 0.5 K. Exceptions were found during stronger winds at the beginning of October and November, when the assimilated SST forecast remained by about 1 K smaller than FB observations. Positive anomalies were

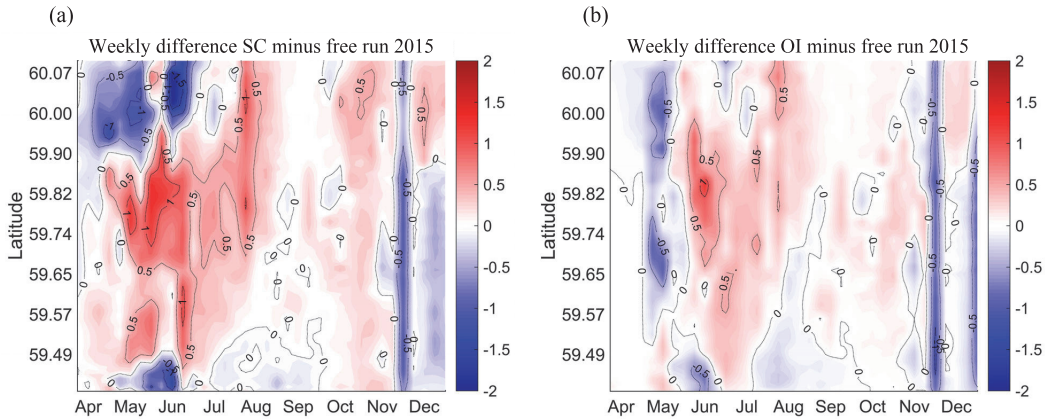
observed during the periods of calmer winds in the second half of May and August. Comparison of the results by SC and OI methods indicates that OI provided with a given set of parameters generally smaller differences from FB than SC, except in June and July near the northern coast when OI produced larger differences than SC. We note that the SC method used the interpolation weights not dependent on the direction between the points; in case of OI longer correlation scales were used along the coasts than in the cross-shore direction.

Within the data assimilation, greater changes in reference to FR (Fig. 7) were made by the OI than by the SC method. We selected the OI parameters on the basis of some trials. However, investigation of the best combination of different parameters for OI is outside the scope of this paper, reduction of sigma ratio  $\varepsilon^2 = \sigma_o^2 / \sigma_m^2$  by the factor of two did not yield plausible results.

Statistical comparison of weekly mean values of the FR forecast and the assimilated SC and OI forecasts with FB data (Table 1) revealed that assimilation provided better correspondence to the independent observations. Based on formula (6), improvements are visible in bias, RMSD, and correlation  $R$ . The main skill estimator –



**Fig. 6.** Time–latitude map of the weekly mean sea nightly surface temperature difference of assimilated with SC (a) and OI (b) in reference to FerryBox data between Tallinn and Helsinki. For abbreviations, see the title of Table 1.



**Fig. 7.** Time–latitude map of the weekly mean nightly sea surface temperature difference of assimilated with SC (a) and OI (b) in reference to non-assimilated model data between Tallinn and Helsinki. For abbreviations, see the title of Table 1.

**Table 1.** Statistics of free model run without data assimilation (FR), model run with assimilation with the successive correction method (SC), model run with assimilation of Copernicus SST with the optimal interpolation method (OI), and remote sensing SST data (SAT) with reference to FerryBox data (FB)

	FR	SC	OI	SAT	FB
Bias	−0.45	−0.34	−0.42	−0.31	0
RMSD	0.97	0.84	0.96	0.66	0
Correlation $R$	0.931	0.937	0.934	0.936	1.00
Mean	10.99	11.11	11.03	11.13	11.45
Standard deviation $\sigma$	4.35	4.45	4.48	4.57	4.19

RMSD – was less than 1 K in all the cases, which can be considered acceptable. In the given selection of assimilation parameters, SC provided slightly better results than OI, that is SC had a smaller bias and RMSD and a larger correlation. Since the statistics was calculated in relation to the constant mean value over the whole data set (period from April to December), deviations contained the seasonal cycle. Therefore standard deviations of SST were in the range of 4.2–4.6 °C. This is also the reason why the calculated correlations were quite high – more than 0.93.

In reference to the SAT data, FR had RMSE = 0.96. Data assimilation reduced this value to 0.82 (SC) and 0.93 (OI).

#### 4. DISCUSSION

Massive validation of Baltic marine forecast products was conducted earlier by the Baltic Monitoring and Forecasting Centre. Results from different Baltic-wide model setups were compared with offshore and coastal in situ observations and with satellite L3 supercollated products. From the project report (<http://marine.copernicus.eu/documents/QUID/CMEMS-BAL-QUID-003-006.pdf>) it can be found that in the year 2008 a common feature of the present HBM model versions was faster heating during spring and summer and faster cooling during autumn. For example, comparison of the monthly mean SST maps from remote sensing (SAT) to the forecast SST maps reveals a positive bias of the model forecast of up to 2 K from April to July and a negative bias amounting to –1 K from September to December. During another period, in 2013, the bias was negative throughout the whole year, with largest forecast–observation differences found in winter. Spatial differences in the bias were evident as well. In the Gulf of Finland the forecast SST tends to be smaller than that observed by SAT data. In our study with the sub-regional model of higher resolution, a seasonally and spatially varying bias appears also in our FR data (reference run without data assimilation).

Compared to the errors in the open sea, higher modelled SST monthly and quarterly scale errors were found in our study in the coastal areas. A possible reason is that the model produces too fast heating or cooling in shallow coastal areas, where stratification is usually absent. This indicates probable underestimation of modelled coastal–offshore exchange. Recent studies emphasize the importance of sub-mesoscale exchange processes (Lips et al., 2016; Väli et al., 2017), which need to be adequately captured by the models. During the spring heating period, the SAT data taken from the surface showed higher SST in the coastal areas than the

FB data observed at 4-m depth (Fig. 3). This points to the importance of accounting for the shallow stratification that may develop during the calm days.

Another important forecast aspect in coastal areas is reproduction of upwelling events (Uiboupin and Laanemets, 2009; Laanemets et al., 2011). Although upwelling patterns are modelled quite well, models tend to produce too low temperatures near the northern coast of the Gulf of Finland compared to the SAT data. On the larger estuarine systems this mismatch may be related to the interaction of surface circulation (Elken et al., 2011; Soosaar et al., 2014) and lateral salinity gradients that create thin layers of less saline water on the surface during calm weather and suppress mixing; such layers may not be resolved well by the models.

Introduction of SAT data assimilation by the SC and OI methods slightly improved the forecast. However, since SAT and FB data sets of SST have differences, the assimilated model results kept to some degree the main error features relative to FB such as too warm waters in the northern part of the Gulf of Finland during the heating period in May and June and too cold waters on the whole Gulf of Finland cross-section during the cooling period from October to December. Usually OI provides more accurate results than SC, but in our case OI did not improve the forecast as much as SC. This may be due to the inadequate description of the correlation function and noise parameters. The focus of a further study could be improvement of the description of these parameters.

We tested routine tools – data products and standard, validated model – in the data assimilation into sub-regional ocean forecast models. The results were satisfactory, and there is a need and possibility for further developments that can be divided into two categories: improvements in data and in models. SAT products of SST are often too smooth and lack mesoscale details under cloudy areas. In such areas there is deviation of heat exchange with the atmosphere compared to cloud-free conditions and sometimes also heavy precipitation, which both cause anomalies in SST. If we assimilate SST towards SAT data, it may happen that the assimilation result will be more different from the independent FB data than the model results without assimilation. There could be a need for an algorithm to incorporate FB observations into the satellite-based SST product. The spatial effect of such correction is presently not known. The SST in coastal stations depends very much on very-high-resolution local topography, and these data can be used mainly in the local forecast models.

Both the used data assimilation algorithms have several important parameters that influence the forecasting of the key variables and yield results of different quality.



For this study these values were picked as the best choice to our knowledge compared to other options. However, this does not mean the results cannot be improved.

There is a challenge to use more observations for data assimilation, but there should be some reliable independent data remaining for validation. If FerryBox data could be used in combination with CMEMS SAT products, we would be left with sparse point observations from coastal stations for checking the quality of assimilation.

## 5. CONCLUSIONS

It was found that satellite SST products from the Copernicus Marine Environment Monitoring Service can be well used for data assimilation in the sub-regional marine forecasts. Although the reference model without data assimilation – sub-regional setup of the HBM model – provided SST forecasts with root-mean-square difference to the observational data sets (satellite products and independent FerryBox data) less than 1 K, further improvements of the forecast were achieved by robust implementation of two assimilation methods: successive corrections and optimal interpolation. Within the selected parameters of assimilation algorithms, a computationally effective successive corrections algorithm gave slightly better results in reference to independent FerryBox data than optimal interpolation.

Higher SST forecast errors of the reference model were found near the shallower northwestern coasts. During the calm heating period in spring and early summer, the reference model produced in these regions too warm waters compared with the satellite and FerryBox observations. Too cold waters, compared to the observations, were modelled during the cooling period from late summer to winter. Although data assimilation reduced these errors, improving the treatment of coastal-offshore exchange in the core forecast model should be useful.

## ACKNOWLEDGEMENTS

Development of the HBM model was done by a larger BAL MFC team within the EU projects MyOcean, MyOcean2, and MyOcean-FO. Details of the model implementation were introduced by Priidik Lagemaa. Lars Axell from Swedish Meteorological and Hydrological Institute gave guidance on the data assimilation algorithms and kindly provided the FORTRAN codes. The study was supported by the PhD programme for Mihhail Zujev and institutional research funding IUT 19-6 of

the Estonian Ministry of Education and Research. The publication costs of this article were partially covered by the Estonian Academy of Sciences.

## REFERENCES

- Anding, D. and Kauth, R. 1970. Estimation of sea surface temperature from space. *Remote Sens. Environ.*, **1**(4), 217–220.
- Axell, L. 2013. *BSRA-15: A Baltic Sea Reanalysis 1990–2004*. SMHI.
- Axell, L. and Liu, Y. 2016. Application of 3-D ensemble variational data assimilation to a Baltic Sea reanalysis 1989–2013. *Tellus A*, **68**, 24220.
- Berg, P. and Poulsen, J. W. 2012. *Implementation Details for HBM*. DMI Technical Report No. 12-11. Copenhagen.
- Bonekamp, H., Montagner, F., Santacesaria, V., Nogueira Loddo, C., Wannop, S., Tomazic, I., et al. 2016. Core operational Sentinel-3 marine data product services as part of the Copernicus Space Component. *Ocean Sci.*, **12**(3), 787–795.
- Brasseur, P., Bahurel, P., Bertino, L., Birol, F., Brankart, J. M., Ferry, N., et al. 2005. Data assimilation for marine monitoring and prediction: the MERCATOR operational assimilation systems and the MERSEA developments. *Q. J. Roy. Meteor. Soc.*, **131**(613), 3561–3582.
- Buettner, K. J. and Kern, C. D. 1965. The determination of infrared emissivities of terrestrial surfaces. *J. Geophys. Res.*, **70**(6), 1329–1337.
- Canizares, R., Madsen, H., Jensen, H. R., and Vested, H. J. 2001. Developments in operational shelf sea modelling in Danish waters. *Estuar. Coast. Shelf S.*, **53**(4), 595–605.
- Cressman, G. P. 1959. An operational objective analysis system. *Mon. Weather Rev.*, **87**(10), 367–374.
- Cummings, J. A. 2005. Operational multivariate ocean data assimilation. *Q. J. Roy. Meteor. Soc.*, **131**(613), 3583–3604.
- Derber, J. and Rosati, A. 1989. A global oceanic data assimilation system. *J. Phys. Oceanogr.*, **19**(9), 1333–1347.
- Donlon, C. J., Minnett, P. J., Gentemann, C., Nightingale, T. J., Barton, I. J., Ward, B., and Murray, M. J. 2002. Toward improved validation of satellite sea surface skin temperature measurements for climate research. *J. Climate*, **15**(4), 353–369.
- Elken, J., Nömm, M., and Lagemaa, P. 2011. Circulation patterns in the Gulf of Finland derived from the EOF analysis of model results. *Boreal Environ. Res.*, **16**, 84–102.
- Fu, W. 2016. On the role of temperature and salinity data assimilation to constrain a coupled physical-biogeochemical model in the Baltic Sea. *J. Phys. Oceanogr.*, **46**(3), 713–729.
- Fu, W., Hoyer, J. L., and She, J. 2011a. Assessment of the three dimensional temperature and salinity observational networks in the Baltic Sea and North Sea. *Ocean Sci.*, **7**(1), 75.
- Fu, W., She, J., and Zhuang, S. 2011b. Application of an Ensemble Optimal Interpolation in a North/Baltic Sea

- model: assimilating temperature and salinity profiles. *Ocean Model.*, **40**(3), 227–245.
- Fu, W., She, J., and Dobrynin, M. 2012. A 20-year reanalysis experiment in the Baltic Sea using three-dimensional variational (3DVAR) method. *Ocean Sci.*, **8**(5), 827–844.
- Funkquist, L. 2006. An operational data assimilation system for the Baltic Sea. In *European Operational Oceanography: Present and Future* (Dahlin, H., Flemming, N. C., Marchand, P., and Petersson, S. E., eds). EuroGOOS Office, Norrköping, Sweden, and European Commission, Brussels, Belgium, pp. 656–660.
- Gandin, L. S. 1963. *Objective Analysis of Meteorological Fields*. Gidrometizdat, Leningrad. (English translation No. 1373 by Israel Program for Scientific Translations (1965), Jerusalem.)
- Ghil, M., Halem, M., and Atlas, R. 1979. Time-continuous assimilation of remote-sounding data and its effect on weather forecasting. *Mont. Weather Rev.*, **107**(2), 140–171.
- Høyer, J. L. and She, J. 2007. Optimal interpolation of sea surface temperature for the North Sea and Baltic Sea. *J. Marine Syst.*, **65**(1), 176–189.
- Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C. 1997. Unified notation for data assimilation: operational, sequential and variational. *J. Meteorol. Soc. Jpn. Ser. II*, **75**(1B), 181–189.
- Ivanov, S. V., Kosukhin, S. S., Kaluzhnaya, A. V., and Boukhanovsky, A. V. 2012. Simulation-based collaborative decision support for surge floods prevention in St. Petersburg. *J. Comput. Sci.*, **3**(6), 450–455.
- Kalman, R. E. and Bucy, R. S. 1961. New results in linear filtering and prediction theory. *J. Basic Eng.*, **83**(3), 95–108.
- Kikas, V. and Lips, U. 2016. Upwelling characteristics in the Gulf of Finland (Baltic Sea) as revealed by Ferrybox measurements in 2007–2013. *Ocean Sci.*, **12**(3), 843–859.
- Laanemets, J., Väli, G., Zhurbas, V., Elken, J., Lips, I., and Lips, U. 2011. Simulation of mesoscale structures and nutrient transport during summer upwelling events in the Gulf of Finland in 2006. *Boreal Environ. Res.*, **16**, 15–26.
- Lagemaa, P. 2012. *Operational Forecasting in Estonian Marine Waters*. TUT Press.
- Lips, U., Lips, I., Kikas, V., and Kuvaldina, N. 2008. Ferrybox measurements: a tool to study meso-scale processes in the Gulf of Finland (Baltic Sea). In *2008 IEEE/OES US/EU-Baltic International Symposium, 27–29 May 2008, Tallinn, Estonia*, pp. 334–339.
- Lips, U., Kikas, V., Liblik, T., and Lips, I. 2016. Multi-sensor in situ observations to resolve the sub-mesoscale features in the stratified Gulf of Finland, Baltic Sea. *Ocean Sci.*, **12**(3), 715–732.
- Liu, Y., Zhu, J., She, J., Zhuang, S., Fu, W., and Gao, J. 2009. Assimilating temperature and salinity profile observations using an anisotropic recursive filter in a coastal ocean model. *Ocean Model.*, **30**(2), 75–87.
- Liu, Y., Meier, H. E., and Axell, L. 2013. Reanalyzing temperature and salinity on decadal time scales using the Ensemble Optimal Interpolation data assimilation method and a 3D ocean circulation model of the Baltic Sea. *J. Geophys. Res. Oceans*, **118**(10), 5536–5554.
- Liu, Y., Meier, H. M., and Eilola, K. 2014. Improving the multiannual, high-resolution modelling of biogeochemical cycles in the Baltic Sea by using data assimilation. *Tellus A*, **66**, 24908.
- Lorenc, A. C. 1986. Analysis methods for numerical weather prediction. *Q. J. Roy. Meteor. Soc.*, **112**(474), 1177–1194.
- Losa, S. N., Danilov, S., Schröter, J., Nerger, L., Maßmann, S., and Janssen, F. 2012. Assimilating NOAA SST data into the BSH operational circulation model for the North and Baltic Seas: inference about the data. *J. Marine Syst.*, **105**, 152–162.
- Losa, S. N., Danilov, S., Schröter, J., Janjić, T., Nerger, L., and Janssen, F. 2014. Assimilating NOAA SST data into BSH operational circulation model for the North and Baltic Seas: Part 2. Sensitivity of the forecast’s skill to the prior model error statistics. *J. Marine Syst.*, **129**, 259–270.
- Martin, M., Dash, P., Ignatov, A., Banzon, V., Beggs, H., Brasnett, B., et al. 2012. Group for High Resolution Sea Surface temperature (GHRSSST) analysis fields inter-comparisons. Part I: A GHRSSST multi-product ensemble (GMPE). *Deep Sea Res. Part II Top. Stud. Oceanogr.*, **77**, 21–30.
- McPherson, R. D., Bergman, K. H., Kistler, R. E., Rasch, G. E., and Gordon, D. S. 1979. The NMC operational global data assimilation system. *Mon. Weather Rev.*, **107**(11), 1445–1461.
- Nardelli, B. B., Tronconi, C., Pisano, A., and Santoleri, R. 2013. High and Ultra-High resolution processing of satellite Sea Surface Temperature data over Southern European Seas in the framework of MyOcean project. *Remote Sens. Environ.*, **129**, 1–16.
- Nowicki, A., Dzierzbicka-Głowacka, L., Janecki, M., and Kałas, M. 2015. Assimilation of the satellite SST data in the 3D CEMBS model. *Oceanologia*, **57**(1), 17–24.
- She, J., Høyer, J. L., and Larsen, J. 2007. Assessment of sea surface temperature observational networks in the Baltic Sea and North Sea. *J. Marine Syst.*, **65**(1), 314–335.
- Siegel, H., Gerth, M., and Tschersich, G. 2006. Sea surface temperature development of the Baltic Sea in the period 1990–2004. *Oceanologia*, **48**(S), 119–131.
- Sokolov, A., Andrejev, O., Wulff, F., and Rodriguez Medina, M. 1997. *The Data Assimilation System for Data Analysis in the Baltic Sea*. Systems Ecology Contributions 3. Stockholm University.
- Soosaar, E., Maljutenko, I., Raudsepp, U., and Elken, J. 2014. An investigation of anticyclonic circulation in the southern Gulf of Riga during the spring period. *Cont. Shelf Res.*, **78**, 75–84.
- Sørensen, J. V. T. and Madsen, H. 2004. Efficient Kalman filter techniques for the assimilation of tide gauge data in three-dimensional modeling of the North Sea and Baltic Sea system. *J. Geophys. Res. Oceans*, **109**, CO3017.
- Tang, Y., Kleeman, R., and Moore, A. M. 2004. SST assimilation experiments in a tropical Pacific Ocean model. *J. Phys. Oceanogr.*, **34**(3), 623–642.
- Taylor, K. E. 2001. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmospheres*, **106**(D7), 7183–7192.
- Uiboupin, R. and Laanemets, J. 2009. Upwelling characteristics derived from satellite sea surface temperature

- data in the Gulf of Finland, Baltic Sea. *Boreal Environ. Res.*, **14**(2), 297–304.
- Uiboupin, R. and Laanemets, J. 2015. Upwelling parameters from bias-corrected composite satellite SST maps in the Gulf of Finland (Baltic Sea). *IEEE Geosci. Remote Sens. Lett.*, **12**(3), 592–596.
- Väli, G., Zhurbas, V., Lips, U., and Laanemets, J. 2017. Submesoscale structures related to upwelling events in the Gulf of Finland, Baltic Sea (numerical experiments). *J. Marine Syst.*, **171**, 31–42.
- Zhuang, S. Y., Fu, W. W., and She, J. 2011. A pre-operational three Dimensional variational data assimilation system in the North/Baltic Sea. *Ocean Sci.*, **7**(6), 771–781.

## Mereandmete assimileerimise testimine Läänemere kirdeosas, kasutades Copernicuse mereseire programmi satelliitproduktide veepinna temperatuuri andmeid

Mihhail Zujev ja Jüri Elken

Katsetati Copernicuse mereseire programmi satelliitproduktide veepinna temperatuuri (SST) andmete assimileerimist piirkondlikku mereprognoside mudelisse HBM, mis oli seadistatud Läänemere kirdeosa jaoks, kaasa arvatud Soome ja Liivi laht. Igapäevastele prognoosidele ajavahemikus aprillist detsembrini 2015 rakendati kaht assimileerimise algoritmi: järjestikuseid parandusi ja optimaalinterpolatsiooni. Valideerimine oli tehtud Tallinna-Helsingi liinil sõitvate laevade pardalt kogutud FerryBoxi andmetega. Suuremad SST prognoosivead (kasutades assimileerimiseta referentsmudelit) esinesid väiksemate sügavustega looderanniku lähedal. Tuulevaiksete soojenemisperioodide jooksul, kevadel ja varasuvel, tekitas mudel soojema veemassi, kui näitasid satelliidi ning FerryBoxi andmed. Vaatlustega võrreldes külmemad piirkonnad olid modelleeritud hilissuvest talveni. Kuigi assimileerimise tulemusena õnnestus vigu vähendada, on otstarbekas parendada referentsmudeli osavust prognoosida veevahetust rannaala ja avamere vahel.

## **Paper II**

Elken, J., Zujev, M., She, J., Lagemaa, P., 2019. Reconstruction of Large-Scale Sea Surface Temperature and Salinity Fields Using Sub-Regional EOF Patterns From Models. *Frontiers in Earth Science*, 7, 1–20. [10.3389/feart.2019.00232](https://doi.org/10.3389/feart.2019.00232).





# Reconstruction of Large-Scale Sea Surface Temperature and Salinity Fields Using Sub-Regional EOF Patterns From Models

Jüri Elken<sup>1\*</sup>, Mihhail Zujev<sup>1</sup>, Jun She<sup>2</sup> and Priidik Lagemaa<sup>1</sup>

<sup>1</sup> Department of Marine Systems, Tallinn University of Technology, Tallinn, Estonia, <sup>2</sup> Research and Development Department, Danish Meteorological Institute, Copenhagen, Denmark

## OPEN ACCESS

### Edited by:

Marcus Reckermann,  
Helmholtz Centre for Materials  
and Coastal Research (HZG),  
Germany

### Reviewed by:

Antonio Turiel,  
Spanish National Research Council  
(CSIC), Spain  
Fabien Roquet,  
University of Gothenburg, Sweden

### \*Correspondence:

Jüri Elken  
juri.elken@taltech.ee

### Specialty section:

This article was submitted to  
Interdisciplinary Climate Studies,  
a section of the journal  
Frontiers in Earth Science

**Received:** 14 January 2019

**Accepted:** 22 August 2019

**Published:** 11 September 2019

### Citation:

Elken J, Zujev M, She J and  
Lagemaa P (2019) Reconstruction  
of Large-Scale Sea Surface  
Temperature and Salinity Fields Using  
Sub-Regional EOF Patterns From  
Models. *Front. Earth Sci.* 7:232.  
doi: 10.3389/feart.2019.00232

A method for reconstruction of gridded fields of sea surface variables from time-dependent observations, using sub-regional EOF (Empirical Orthogonal Functions) patterns from models, is presented and tested. Covariance fields, calculated from the model results over long enough time span, are used to find EOF modes. The gravest “observational” amplitudes and their first temporal derivatives are determined from the least-square minimization of fitting errors in relation to the observed values. The field is reconstructed by superposition of continuous model-based mode patterns multiplied by observational amplitudes that meet adopted statistical limits. If the observational amplitude exceeds the limits, gridded fields for this and higher modes are not produced. We applied the method in the northeastern Baltic over the model time series 2010–2015. Daily averages of sea surface temperature (SST) and salinity (SSS) from the high-resolution (grid step 0.5 nautical miles) sub-regional HBM model were spatially averaged over bins of 5 × 5 nautical miles. Three first modes cover 99% of variance of temperature and 61.4% of salinity. As shown by experiments with pseudo-observations (model values at these points reconstructed to the model grid and then compared with the original model data), reconstruction performance depends on the configuration of the observation points in the model domain. Still, a few first modes usually produce acceptable results. When removing the SST seasonal cycle prior to EOF analysis, spatial patterns of leading modes remained practically unchanged, share of variance of the three first modes was reduced to 88.6% and reconstruction errors were reduced by about 25%. Sufficient spatial data coverage of the larger basin with ship-born observations usually takes quite long time – of the order of month; therefore, time correction of the amplitudes using the found temporal derivatives improves the accuracy of reconstruction. The method is compared with the Optimal Interpolation (OI) by using the pseudo-observations. Results show that, for SST reconstruction, the OI method is significantly worse than the EOF method. For SSS, OI is slightly better than EOF. The superiority of EOF is that the remote correlation patterns can be used in the reconstruction, which is important when the observations are sparse.

**Keywords:** sea surface observations, model-based patterns, EOF analysis, reconstruction of gap-free data, Baltic Sea

## INTRODUCTION

Many oceanographic tasks require appropriate reconstruction of gridded fields from different observational data: shipborne monitoring, coastal stations, offshore buoy stations, FerryBoxes, gliders and remote sensing. As a result, densely sampled sections may be neighbored with areas of rare or missing observational data.

Meteorology and oceanography share the same theoretical foundations of interpolation and data assimilation (Ghil and Malanotte-Rizzoli, 1991; Ide et al., 1997). Their practical implementation is, however, rather different (Ghil, 1989), owing to the nature of governing processes (landlocked basins, shallow areas and wind driving characterize oceans; atmosphere is unbounded, “deep” and self-driving by polar-tropical gradients), but also of techniques and amount/density of observations.

Many different methods have been applied for the data reconstruction, including both statistical [e.g., regression, optimal interpolation and Empirical Orthogonal Functions (EOFs)] and dynamic methods (e.g., data assimilation). Good reconstruction (in some statistical merit) should be based on the knowledge of multiscale spatial and temporal covariance. Atmospheric and ocean variability are similar (Woods, 1980; Cushman-Roisin and Beckers, 2011), if the lengths are scaled to the different values of baroclinic Rossby deformation radius ( $R_d$ ). On the shorter scales, marginal seas and/or their sub-basins which have typical lateral dimensions less than 1000 km (typical  $R_d$  in the atmosphere), are forced by the same or neighboring weather patterns. This causes for example coherent upwelling/downwelling patterns (Lehmann et al., 2012) on the left-hand/right-hand coasts from the direction of weather-generated wind. Considering also faster heating or cooling of shallow coastal areas compared to the deeper offshore regions (Legrand et al., 2015), and freshwater spreading patterns due to the dynamics of river plumes (Soosaar et al., 2016), there could be significant covariance of sea surface temperature (SST) and sea surface salinity (SSS) in marginal seas over large distances, mainly stretched along the topography isolines and/or coasts (Fu et al., 2011).

Classical optimal interpolation (OI) (Gandin, 1963) assumes that covariance is represented by Gaussian, damped cosine or exponential decay of covariance with distance between the points. In case of open sea or dense observations (e.g., satellite SST), the OI is sufficiently good for the reconstruction (Høyer and She, 2007). However, when the observations are sparse or in the coastal waters where the covariance pattern is complicated, more comprehensive reconstruction method is needed.

Spatiotemporal variability of the ocean state can be regarded as an attractor of the dynamic ocean system in a linear phase space, in which any state can be presented as a linear summation of a complete set of orthogonal base functions. The EOFs (Davis, 1976) is one kind of such base functions, which is derived as eigenfunctions of the observed spatial covariance matrix. It projects the spatiotemporal variability of the system state onto correlation patterns of different scales, which are orthogonal. The time-space matrix of the field of interest is decomposed into the sum of space-dependent mode patterns multiplied

by time-dependent amplitudes of each mode (eigenfunction). Eigenvalues present the variance of particular mode; the sum of all eigenvalues is equal to the variance of the initial field. Usually, a few most energetic modes present majority of the initial field variance. The method is not restricted to Gaussian or other similar decay over space lag.

One of the first developments of EOF interpolation in oceanography (Smith et al., 1996) considered SST on the global scale. During the period 1982–1993, when data coverage was good, SST data were gridded using traditional OI. Further, EOF modes were calculated from the gridded data. Subsequently, the EOF method was expanded to the globe in a longer period of 1950–1992. Compared to the traditional OI, the EOF produced enhanced large-scale patterns like ENSO. A number of studies (Kaplan et al., 1997; Kim, 1997; Menemenlis et al., 1997; Beckers and Rixen, 2003) have considered multivariate combined methods of interpolation: large-scale background field is approximated by the dominant EOFs; in the regions of dense sampling, the anomalies from large-scale fields are interpolated using OI or some of its modified method. There are also examples how iterated EOF method (DINEOF – Data Interpolating Empirical Orthogonal Functions) is used to reconstruct gap-free satellite images (Alvera-Azcárate et al., 2015; Jayaram et al., 2018).

The present paper has been initially motivated by the need of detailed examination of spatial covariance characteristics in a specific region – the northeastern Baltic, in relation to the data assimilation. Although using OI with Gaussian correlation function provided satisfactory results (Zujev and Elken, 2018), need for improved description of statistics deemed obvious. During different test options, we used also traditional EOF method. The covariance was determined from the model results since observational data were too irregular. The vast amount of available data was limited to the sea surface data, namely SST and SSS. Although SST is densely sampled by remote sensing, most demanding in terms of methodical aspects is using *in situ* data from a variety of platforms, e.g., research vessels, FerryBoxes and buoy stations. During the tests, we developed an easy algorithm, where “observational” amplitudes of leading base functions (EOF modes) can be evaluated by limited amount of instantaneous observational data using least-square minimization. Smith et al. (1996) have already developed this mathematics earlier, but they used the method in oceanic conditions where EOF behavior is quite different. Applying the method in the sub-region of the marginal sea, preliminary results were promising and they were presented in a recent conference paper by Elken et al. (2018).

The aim of the present paper is to develop and test the method for large-scale EOF analysis of sub-regional time-dependent SST and SSS data, based on the covariance estimates from the model results. In real oceanographic situations, spatial observations are spread over a certain time span (mapping of a sub-region by different countries/ships may take about month), therefore time correction of variables of reconstruction procedure would be useful. “Observational” EOF amplitudes and their temporal derivatives are calculated from the conditions of least-square minimization of EOF analysis error at observation points, compared to the observed values. After evaluating the covariance and EOF modes for 5-years test period, we analyze



the reconstruction accuracy using “pseudo-observations,” i.e., extracting of model data at variable “observation” locations and comparing the reconstruction result with the original model result, using the EOF reconstruction but also OI. Further tests of the method include removal of SST seasonal cycle prior to reconstruction, partitioning of the study region into smaller sub-regions, comparison of reconstruction using time correction, and calculation of long sequences of gridded data using only ship-borne observations. The paper ends with discussion and conclusions.

## METHODS AND DATA

### Notations for Empirical Orthogonal Functions (EOF)

We follow the vector-matrix notation and consider the model results as  $M \times N$  space-time matrix  $\mathbf{X}$  containing deviations from space-dependent temporal mean values  $\bar{\mathbf{x}}_m$ . The columns  $\mathbf{x}_i$  of matrix  $\mathbf{X}$  are spatial time slices consisting of  $M$  points at time  $i$ , out of  $N$  time instances. Determine then the empirical orthogonal functions as  $M \times M$  matrix  $\mathbf{E}$ , which columns are eigenvectors (normalized orthogonal spatial modes)  $\mathbf{e}_k$ . In the decomposition, the eigenvalues  $\lambda_k$  form the diagonal matrix  $\mathbf{\Lambda}$  that has zeros outside the main diagonal. The eigenvalue of the specific mode presents the variance attributed to this mode; the sum of all eigenvalues is the variance of  $\mathbf{X}$ .

Time-dependent part of the decomposition is  $M \times N$  matrix  $\mathbf{A}$ , which columns  $\mathbf{a}_i$  are the values of time-dependent amplitude vectors (one amplitude time series value for each mode) at time  $i$ . As a result, we obtain

$$\mathbf{x}_i = \mathbf{E}\mathbf{\Lambda}\mathbf{a}_i \quad \text{or} \quad \mathbf{x}_i = \mathbf{E}\tilde{\mathbf{a}}_i \tag{1}$$

where  $\tilde{\mathbf{a}}_i = \mathbf{\Lambda}\mathbf{a}_i$  is dimensional amplitude. For the whole data set  $\mathbf{X} = \mathbf{E}\mathbf{\Lambda}\mathbf{A}$ . Note the orthonormality as  $\mathbf{e}_i\mathbf{e}_j = \delta_{i,j}$  and  $\mathbf{a}_i\mathbf{a}_j = \delta_{i,j}$ , where  $\delta_{i,j}$  is the Kronecker symbol. For the amplitudes, orthonormality usually is interpreted that amplitudes of different modes are uncorrelated in time.

The eigenvalue problem is

$$\mathbf{B}\mathbf{E} = \mathbf{\Lambda}\mathbf{E} \tag{2}$$

(equivalent to  $|\mathbf{B} - \lambda\mathbf{I}| = 0$ ), where the covariance matrix averaged over  $N$  instances of time is

$$\mathbf{B} = \frac{1}{N-1}\mathbf{X}^T\mathbf{X} \tag{3}$$

Matrix  $\mathbf{E}$  can be found by a number of methods for solving linear system of equations. One favorite method is singular value decomposition (SVD). Due to the orthonormality  $\mathbf{E}^T\mathbf{E} = \mathbf{I}$ . The dimensional amplitudes are determined by the relation

$$\tilde{\mathbf{a}}_i = \mathbf{\Lambda}\mathbf{a}_i = \mathbf{E}^T\mathbf{x}_i \tag{4}$$

### Reconstruction of Observed Fields Using EOF Modes

Consider now the case where observations at a specific time instance  $i$  are represented by vector  $\mathbf{y}_i$  that has different set of

$K$  points than  $M$  points for  $\mathbf{x}_i$ . If observations include high-resolution data that contain multiple data points within the grid cell and time interval of model lattice, such oversampling has to be removed prior to further analysis, usually by averaging over the grid cell. Therefore  $K \leq M$ . Gridded data  $\mathbf{x}_i$  are transformed to the observation points by matrix  $\mathbf{H}_i$  (observation function) in a way that  $\mathbf{H}_i\mathbf{x}_i$  has the same dimension as  $\mathbf{y}_i$  and has to be directly compared with it. To be more specific, vector  $\mathbf{y}_i$  presents the observed deviations from the temporal mean value  $\mathbf{H}_i\bar{\mathbf{x}}_m$  whereas the observation function  $\mathbf{H}_i$  depends on the configuration of observation points. Eigenvalue transformation takes the form  $\mathbf{H}_i\tilde{\mathbf{x}}_i = \mathbf{H}_i\mathbf{E}\hat{\mathbf{a}}_i$ , where  $\hat{\mathbf{a}}_i$  is the “observational” amplitude, determined from observed values  $\mathbf{y}_i$  at  $K$  observation points, using the full patterns of EOF modes  $\mathbf{e}_k$  with  $M$  spatial points. For the least-squares minimization of  $\|\mathbf{y}_i - \mathbf{H}_i\mathbf{x}_i\|^2 = \|\mathbf{y}_i - \mathbf{H}_i\mathbf{E}\hat{\mathbf{a}}_i\|^2$ , the system of equations is  $\mathbf{H}^T\mathbf{E}^T\mathbf{H}\mathbf{E}\hat{\mathbf{a}}_i = \mathbf{H}^T\mathbf{E}^T\mathbf{y}_i$ , where the amplitudes as  $K \times 1$  vector and interpolated field  $\tilde{\mathbf{x}}_i$  are found from

$$\hat{\mathbf{a}}_i = (\mathbf{H}_i^T\mathbf{E}^T\mathbf{H}_i\mathbf{E})^{-1}\mathbf{H}_i^T\mathbf{E}^T\mathbf{y}_i, \quad \tilde{\mathbf{x}}_i = \mathbf{E}\hat{\mathbf{a}}_i \tag{5}$$

Note, that we cannot here anymore use the condition that the mode patterns are orthonormal.

In the matrix of eigenvectors  $\mathbf{E}$ , where different modes are presented by column vectors, we take only  $L$  first vectors and the rest of the columns are truncated to zero. When using only  $L$  modes for reconstruction, contribution of truncated modes is added in the error variance.

For the clarity of the calculations, we spell out also the element-wise summation form without presenting the time index. Minimization is done for  $Q = \sum_{k=1}^K (y_k - \sum_{l=1}^L \hat{a}_l \hat{e}_l^k)^2$ , leading to the  $L$  conditions  $\partial Q / \partial \hat{a}_l = 0$ . It results in the  $L \times L$  system of linear equations

$$\mathbf{D}\hat{\mathbf{a}} = \mathbf{h} \tag{6}$$

where the matrix and vector elements are

$$D_{mn} = \sum_{k=1}^K \hat{e}_m^k \hat{e}_n^k, \quad h_m = \sum_{k=1}^K y_k \hat{e}_m^k \tag{7}$$

Here  $\hat{e}_m^k$  is the  $m$ -th eigenvector mapped to the observation point  $k$ .

The original dimensional amplitudes  $\tilde{\mathbf{a}}$  have some statistical regularities determined over a large number of samples. Such regularities contain e.g., standard deviation  $\sigma$  or variance  $\sigma^2$ , percentiles and covariance in relation to time lag etc. The observational amplitudes  $\hat{\mathbf{a}}$  are determined from much less amount of information and are rather uncertain. There is a caution that with bad configuration of the observation points, observational EOF amplitudes of particular modes may get larger than limits determined from full statistics (details in section Covariance and EOF Characteristics). Therefore it is important to determine the maximum number of modes  $L$  by checking if determined  $\hat{\mathbf{a}}$  values lie within the statistical limits of  $\tilde{\mathbf{a}}$ ; if the limits are exceeded then this and higher modes are removed from the further analysis.



## Extension of the EOF Reconstruction Method to Time-Dependent Data

Quite often in oceanographic practice, there are not enough observational data at a specific time instance  $i$  to perform reliable construction of observations. Shipborne surveys over larger sub-regions may take several days or even weeks. Usual procedure is to consider the observations  $y_p$  made within the time window  $n_1 \dots n_2$ ,  $p \in n_1 \dots n_2$  as instantaneous, and reference the (non-processed) result to the time instance  $i$ , when  $n_1 \leq i \leq n_2$ . Such procedure can introduce apparent distortions, when the observations are conducted during the increase or decrease period within the seasonal cycle. For example, when during the spring warming period the observations are first acquired in the southern part and later in the northern part, then higher temperatures presented in the northern part of the map compared to the southern part are just an artifact, due to missing treatment of time-dependence of the data.

Take now the  $P$  observed data  $y_p$  within window  $p \in n_1 \dots n_2$  and construct modified observation function  $\hat{H}_p$  that allows pointwise comparison of  $y_p$  at different times and  $\hat{H}_p \mathbf{x}$  based on gridded data at specified time  $i$ . Time difference of observation  $p$  and reference time  $i$  is determined  $\Delta t_p = t_p - t_i$ . Eigenvalue transformation is  $\hat{H}_p \hat{\mathbf{x}}_i = \hat{H}_p \mathbf{E} \hat{\mathbf{b}}_p$ , where modified amplitude, accounting for linear time dependence given by rate of change vector  $\alpha_i$ , is  $\hat{\mathbf{b}}_p = \hat{\mathbf{a}}_i + \alpha_i \cdot \Delta t_p$ . The function to be minimized is  $Q = \|y_p - \hat{H}_p \mathbf{E} \hat{\mathbf{b}}_p\|^2 = \|y_p - \hat{H}_p \mathbf{E}(\hat{\mathbf{a}}_i + \alpha_i \cdot \Delta t_p)\|^2$  regarding  $\partial Q / \partial \hat{a}_i = 0$  and  $\partial Q / \partial \alpha_i = 0$ . Define the  $2L$  unknown coefficients  $\mathbf{z} = \{\hat{a}_1 \dots \hat{a}_L, \alpha_1 \dots \alpha_L\}$  and modified EOF mode values at observation points  $\hat{\mathbf{f}}_m^p = \{\hat{c}_1^p \dots \hat{c}_L^p, \hat{c}_1^p \Delta t_p \dots \hat{c}_L^p \Delta t_p\}$ , we obtain  $2L \times 2L$  system of linear equations

$$\mathbf{Gz} = \mathbf{w} \tag{8}$$

where the matrix and vector elements are

$$G_{mn} = \sum_{p=1}^P f_m^p f_n^p, \quad w_n = \sum_{p=1}^P y_p f_n^p \tag{9}$$

We note that when all observations have the same time stamp and  $\Delta t_p = 0$ , the system (8)-(9) is reduced to (6)-(7).

From the found vectors  $\mathbf{z}$  we extract separately observational amplitudes  $\hat{\mathbf{a}} = \{\hat{a}_1 \dots \hat{a}_L\}$  and their temporal derivatives  $\alpha = \{\alpha_1 \dots \alpha_L\}$ , where they both are checked for the statistics of full data set, in order to determine the highest acceptable mode  $L$ .

Reference time  $i$  for observational amplitudes (and corresponding reconstructions) can be modified on the condition of acceptable accuracy. Finding these bounds is a subject of separate study. In principle, it is possible to perform centered referencing, including the data from past and future times (like it is done in processing of existing time series), but also backward referencing, including only the past data (like within data assimilation for on-line forecasts).

## Estimation of Reconstruction Accuracy Using Pseudo-Observations

Accuracy of EOF reconstruction by a limited number of modes is performed by evaluating the reconstructed fields versus original

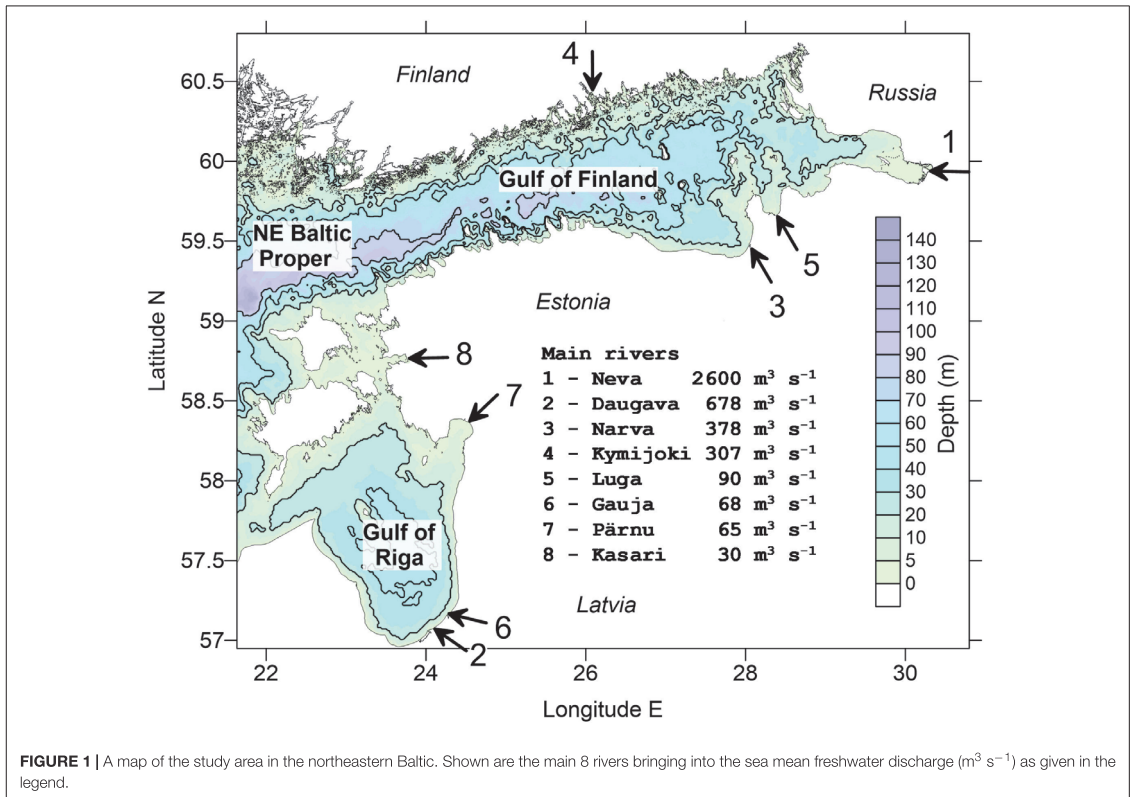
fields over a sufficiently long span of time. In case of observational data, another key factor, besides the number of modes, is configuration (including the number) of observation points. Assuming that statistical features of observations are close to that of the model results, we introduce pseudo-observations as extract of model results in the predefined locations where usually observations are taken. Accuracy of reconstruction from the pseudo-observations was checked by a series of experiments containing the following steps: (i) Configuration of observation points was selected; (ii) model values were extracted at observation points (pseudo-observations were taken); (iii) reconstructed fields  $\hat{\mathbf{x}}_i$  were calculated from pseudo-observations using (5)-(7); (iv) calculations were repeated for all time instances available, statistical characteristics like root-mean-square deviation (RMSD) between the reconstructed and original fields were evaluated.

The main experiments were made for the case of pseudo-observations on the variable grid. The factor  $N$  by which the grid step of observations were larger than the model grid was varied from 1 to 11. Additional experiments were performed with configurations typical to the FerryBox observation points and typical to the marine shipborne monitoring with reduced sampling network (Elken et al., 2018; not shown here).

In addition to the EOF reconstruction, OI was used for comparison purposes in two configurations: (i) interpolation of deviations from locally resolved mean (modeled) fields that includes high gradients in the coastal areas of river influence zones, (ii) interpolation of deviations from smooth climatological mean fields. Both configurations used Gaussian correlation function in the form  $C(r) = \exp(-r^2/R^2)$  (e.g., Zujev and Elken, 2018), where  $r$  is the space lag and  $R$  is the correlation scale. The OI configurations used for smoothing purposes prescribed noise-to-signal ratio  $\eta^2$ .

## Regional Setting of Experiments

We chose the area of our study in the northeastern Baltic (Figure 1) that contains two distinct geographical areas – Gulf of Finland and Gulf of Riga – and includes the northeastern part of the Baltic Proper. The Baltic Sea is a brackish estuarine-type multi-basin marginal sea (Elken and Matthäus, 2008; Leppäranta and Myrberg, 2009), where complex coastline and topography essentially guide the dynamics of SST and SSS. In the Gulf of Finland, EOF modes have profound structure (Elken et al., 2011). Thermal regime is dominated by the seasonal heat cycle, but it is also modified by differential heating/cooling above variable depths in the coastal and offshore areas. Ice cover occurs in the coastal areas every winter, while open parts of the sub-area are ice-covered during severe winters (Vihma and Haapala, 2009). SST is heavily modified by upwelling and downwelling patterns induced by the transient wind fields (e.g., Laanemets et al., 2011). Due to the fragmented coastline and multiple rivers entering the area, SSS has numerous high-gradient regions. Large scale SSS patterns are guided by unsteady circulation that depend on the climatic variations of atmospheric forcing; while earlier studies suggested cyclonic circulation patterns in both the Gulf of Finland and the Gulf of Riga and right-hand spreading of less saline waters from the large Neva and Daugava rivers, then recent



**FIGURE 1** | A map of the study area in the northeastern Baltic. Shown are the main 8 rivers bringing into the sea mean freshwater discharge ( $\text{m}^3 \text{s}^{-1}$ ) as given in the legend.

studies frequently reveal also anticyclonic patterns (Soosaar et al., 2016). Mesoscale variability has rather short spatial scales; the  $R_d$  values are from a few km to about 7 km (Alenius et al., 2003).

We used the HBM model (Berg and Poulsen, 2012) with sub-regional 0.5 NM (nautical mile, 926 m) setup (Lagemaa, 2012; Zujev and Elken, 2018) in the geographical bounds shown in **Figure 1** to produce the SST and SSS data. This HBM-EST model domain contains  $529 \times 455$  horizontal grid points. Forcing at the western open boundary is taken from the Baltic-wide HBM model, which operates routinely within the Copernicus Marine Environment Monitoring Service (CMEMS) with the resolution of 1 nautical mile. Forcing on the sea surface is obtained from the Estonian version of theHIRLAM model that is run by the national weather service for operational forecasts on 11-km grid.

HBM is a 3D oceanographic model for the North and Baltic Sea, which uses Arakawa C-grid and is forced by surface energy fluxes (mechanical, radiative, thermodynamic) using bulk parameterization formulae. The model includes sub-models for turbulence parameterization. A model for ice thermodynamics and ice mechanics is embedded into the model system. The HBM model has been upgraded within the CMEMS from earlier BSHCmod versions. The Baltic-wide HBM setup is extensively validated within CMEMS. The quality information document for physical variables can be found on the

web <http://cmems-resources.cls.fr/documents/QUID/CMEMS-BAL-QUID-003-006.pdf> as accessed on 10 July 2019.

For the analysis we used daily model data of free run (without data assimilation) averaged over  $10 \times 10$  grid points, resulting in 744 wet points with 5 NM (9.26 km) resolution on the coarse grid. Since the grid step of the averaged fields is larger than the Rossby deformation radius, mesoscale patterns were suppressed in the analysis results. The 5-year analysis period covered 1826 dates from July 1, 2010 to June 30, 2015.

In the observational data we focused on the *in situ* SST and SSS data and remotely sensed SST data were occasionally used for the comparison. Shipborne profile observations were acquired from HELCOM/ICES database, downloaded from <https://ocean.ices.dk/helcom/> on 12 February 2018. After extraction of surface data within the study area, 2915 data records were retrieved within 2009–2014. Prior to using the data for the reconstruction, oversampling for each particular time instance was eliminated by taking averages on the coarse grid and selected time interval. CMEMS remote sensing SST Level 4 (L4) data were downloaded from the service portfolio <http://marine.copernicus.eu/services-portfolio/access-to-products/> as the product SST\_BAL\_SST\_L4\_NRT\_OBSERVATIONS\_010\_007\_b. FerryBox data were obtained from the same portfolio as the product INSITU\_BAL\_NRT\_OBSERVATIONS\_013\_032.

Climatological monthly temperature and salinity fields were adopted from the study made by Janssen et al. (1999), covering the period 1900–1996.

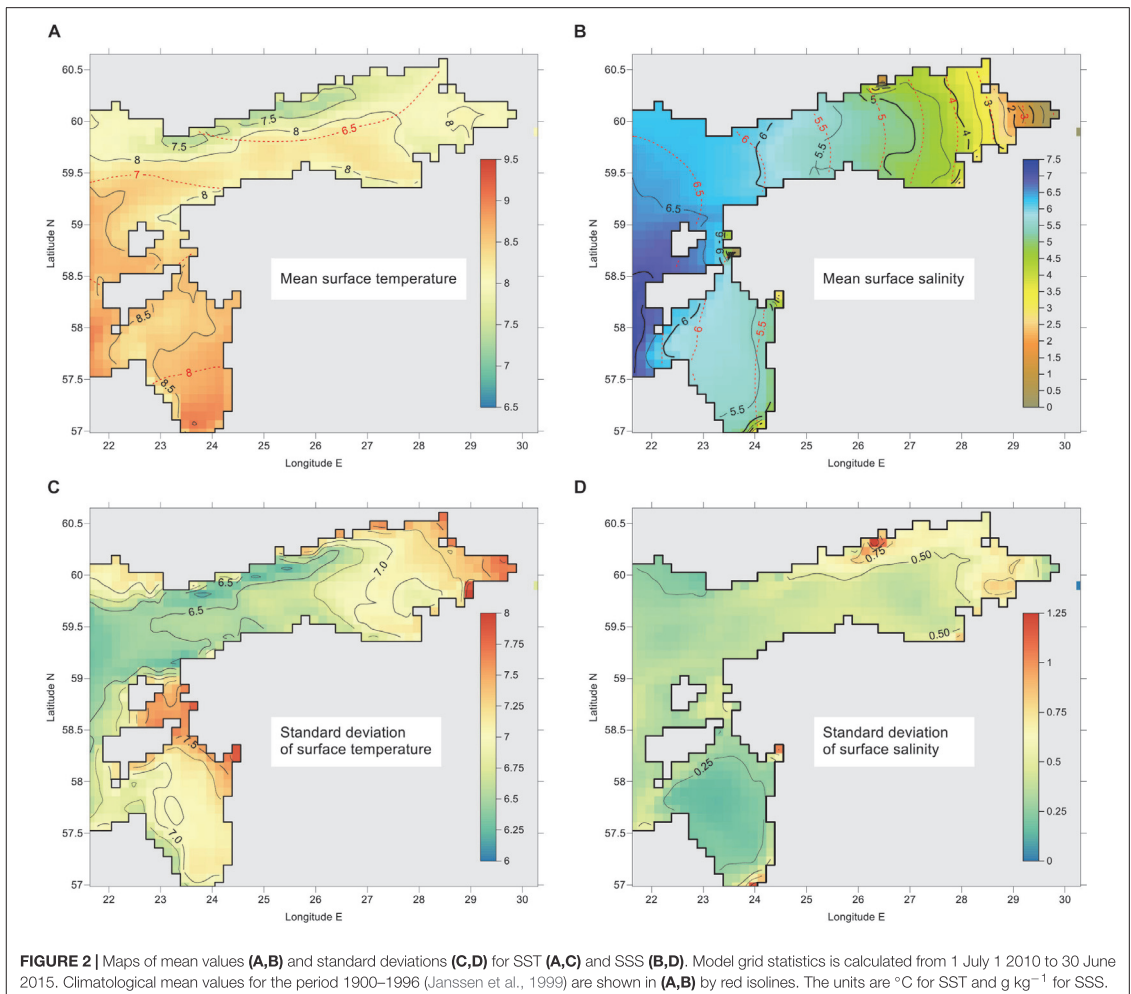
## RESULTS

### Mean and Standard Deviations of SST and SSS

The mean fields of surface temperature and salinity, shown in **Figures 2A,B**, were calculated as temporal mean values of individual grid points  $\bar{x}_m$  over the whole study period. Data assimilation was not performed; therefore, the presented maps may include some model bias. Primary purpose of the mean fields is to set the background for the variability study, i.e., investigate statistical properties of SST and SSS deviations from their fields.

The maps of mean SST and SSS are dependent on the average atmospheric conditions during the period summer 2010 – summer 2015. The period covered severe ice winter (2010/2011), and average (2011/2012, 2012/2013) and mild (2013/2014, 2014/2015) winters (FMI, 2018). The mean SST map reveals lower temperature along the Finnish coast; that occurs during dominating westerly winds favoring upwelling in that region. This is consistent with mean salinity distribution in the Gulf of Finland that exhibited pattern typical to the dominance of reversed estuarine circulation (Westerlund et al., 2019), where tongue of less saline water near the Finnish coast is not present. While our SSS map is close to the yearly climatological map (Janssen et al., 1999) then SST is in the Gulf of Finland higher by 1–1.5°C and in the Gulf of Riga by 0.5–1°C.

Based on all the model values for the period 2010–2015, we calculated total mean value and the corresponding standard



deviation  $\sigma$  for SST as 8.2 and 7.0°C, and for SSS as 5.45 and 1.3 g kg<sup>-1</sup>. While variability of SST is strongly dominated by temporal changes, SSS variability reveals dominant spatial changes. Namely, mean temporal variance  $\overline{\sigma_t^2(k)}$ , calculated on the basis of all spatial points  $k$ , comprises 99% of total variance for SST but only 11% for SSS. The remaining percentage of variance is due to variability of spatial means. Alternatively, mean spatial variance  $\overline{\sigma_x^2(i)}$ , calculated on the basis of all temporal instances  $i$ , reveals 97% of total variance of SSS and only 3% of SST. During individual time instances, spatial standard deviation  $\sigma_x(i)$  for SST was found from 0.08 (during winter) to 2.8°C; SSS values range from 0.95 to 1.5 g kg<sup>-1</sup>.

Maps of temporal standard deviations  $\sigma_t(k)$ , calculated for each spatial point  $k$ , are presented in **Figures 2C,D**. These maps include seasonal cycle but also interannual and shorter period variations. Despite the small fraction of spatial variance of SST, some distinct spatial features over the area are evident. Higher temporal standard deviations of SST (above 7.2°C) are found in the shallow areas in the eastern part of the Gulf of Finland and in the Moonsund located between the large Estonian islands and the mainland. Spatial variations  $\sigma_t(k)$  of SSS are in the range from 0.14 g kg<sup>-1</sup> to 1.5 g kg<sup>-1</sup> (**Figure 2D**), whereas higher values occur near the entrance of larger rivers. High spatial variations of standard deviation make difficult using spatial correlation functions, which calculation require normalizing covariance with variance.

## Covariance and EOF Characteristics

We calculated SST and SSS covariance according to Equation (3). After EOF decomposition of **B** using Equation (2), we also calculated covariance of the sum of six most energetic modes and of the rest higher modes. Due to orthogonality, covariance is additive regarding the EOF modes, i.e., the full covariance is the sum of covariance of the component data sets (six most energetic modes, and the rest higher modes). In **Figure 3** SST and SSS covariance are presented as a function of space lag between the model points. We see significant spreading of individual values of covariance over pairs of points and conclude that calculated covariance is not homogeneous, which is usually assumed in implementation of OI.

In the bins of space lags, distribution (histogram) of covariance of original fields and of the sum of most energetic EOF modes (not shown) usually does not follow the normal distribution. Therefore, mean covariance values can be considered only as indicative, since they differ significantly from the median values. Still it is clear that big covariance values occur over large distances, especially for SST. Covariance of residual fields (sum of higher EOF modes) has a good normal distribution and it decays fast with increasing space lag. Correlation (not shown) goes below 0.2 at a distance of 30 km for both SST and SSS, justifying the use of OI for this part of the variability.

Spatial EOF mode patterns for 4 leading modes are given in **Figures 4, 5** for SST and SSS, respectively. The one-dimensional vectors  $e_k$  of the SST and SSS modes are remapped back into the two-dimensional geographical framework.

Among the spatial patterns, large-scale physical interpretation can be easily found for four to six modes. The first, most energetic modes have nearly “flat” patterns without sign change; their amplitudes are dominated by a seasonal signal. Higher modes are considered random due to eddies and other mesoscale processes, therefore their correlation decays rapidly with increasing distance (see the earlier sub-chapter). In the SST patterns, the first mode dominates heavily (97.64% of variance explained) due to the seasonal cycle (**Table 1**). In the SSS patterns (**Table 2**), the share of different modes is more distributed and the first six modes explain 72.88% of the total variance.

Temporal variance of the mode amplitudes  $\tilde{\mathbf{a}}$  equals to the eigenvalues of covariance matrix **B**. Based on the statistical features of the amplitudes, it is possible to set the “natural” limit  $F_k$  for each of the mode  $k$ . During EOF reconstruction, we use only the modes  $k$  where the estimated individual amplitude values at time  $i$  follow the condition  $|\hat{a}_{i,k}| < F_k$ . Since the absolute values depend on the number of grid point, configuration of the sea area and other factors, we do not present the numerical values of  $F_k$ . Excluding 10% of the higher and lower values of “natural” (calculated from full set of model results) amplitudes, a reasonable limit is  $F_k = 2 \sigma(\tilde{\mathbf{a}}_k)$ .

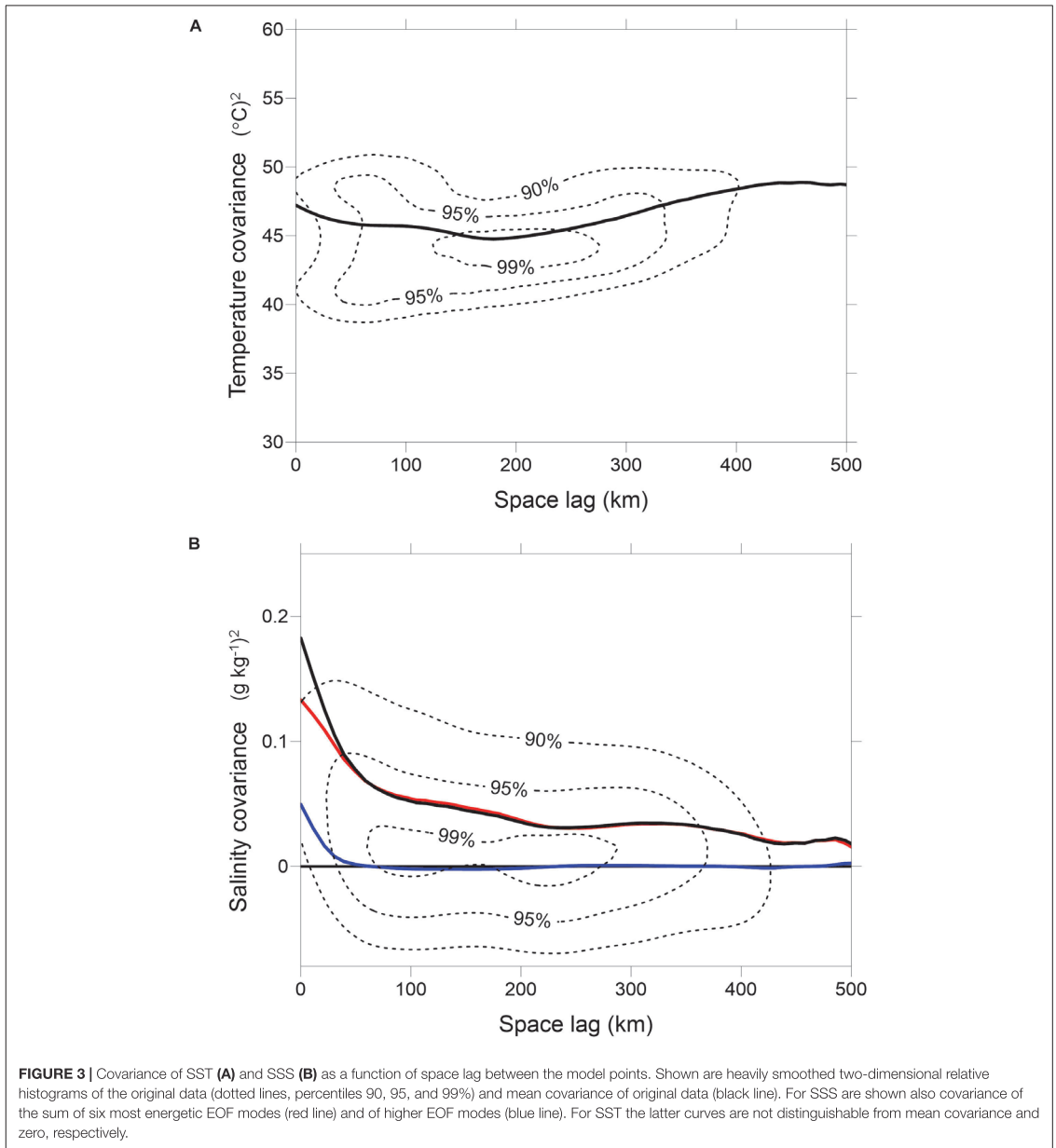
We have presented in this paper the formulae (5)-(7) how to reconstruct gridded fields from observations made during one fixed time instance. Actual spatial observations are quite often not instantaneous in time. The weights of observations from past and future times depend on the temporal covariances (or correlations). Within the EOF decomposition, amplitudes of SST and SSS modes have different temporal correlation patterns, as shown in **Figure 6**. For the SST, the first and the second modes are nearly annually periodic with correlation  $r > 0.9$  and shifted phases. Moderate semi-annual periodicity ( $r \sim 0.2-0.3$ ) appears on the fifth mode. The first SSS mode has annual harmonic with  $r \approx 0.4$ . The second SSS mode has even stronger annual harmonic with  $r \approx 0.6$ . Based on long correlation times, we consider the method of EOF reconstruction of time-dependent observations, presented by formulae (8)-(9), justified for the time window up to 1–2 months.

## Reconstruction Errors: Experiments With Pseudo-Observations

Dependence of accuracy of EOF reconstruction on the number and spacing of observation points was firstly studied by grid configuration of pseudo-observations (see section Estimation of Reconstruction Accuracy Using Pseudo-Observations) with variable grid step. At prescribed locations, model data were extracted on specific time instance; then the result of gridded reconstruction was compared with the original model data. Observational grid step  $\Delta_o$  was taken as integer  $n$  times the model grid step  $\Delta_m$ ,  $\Delta_o = n \Delta_m$ . Observation grid step factor  $n$  was cycled from 1 (observations taken at all the model points) to 11 or more (leaving 2–6 observation points).

We made pointwise comparison of all the 744 spatial points during 1826 time instances, using 6 EOF modes for the reconstruction. Frequency histograms of deviations for SST and SSS are presented in **Figure 7** for two spacing options between

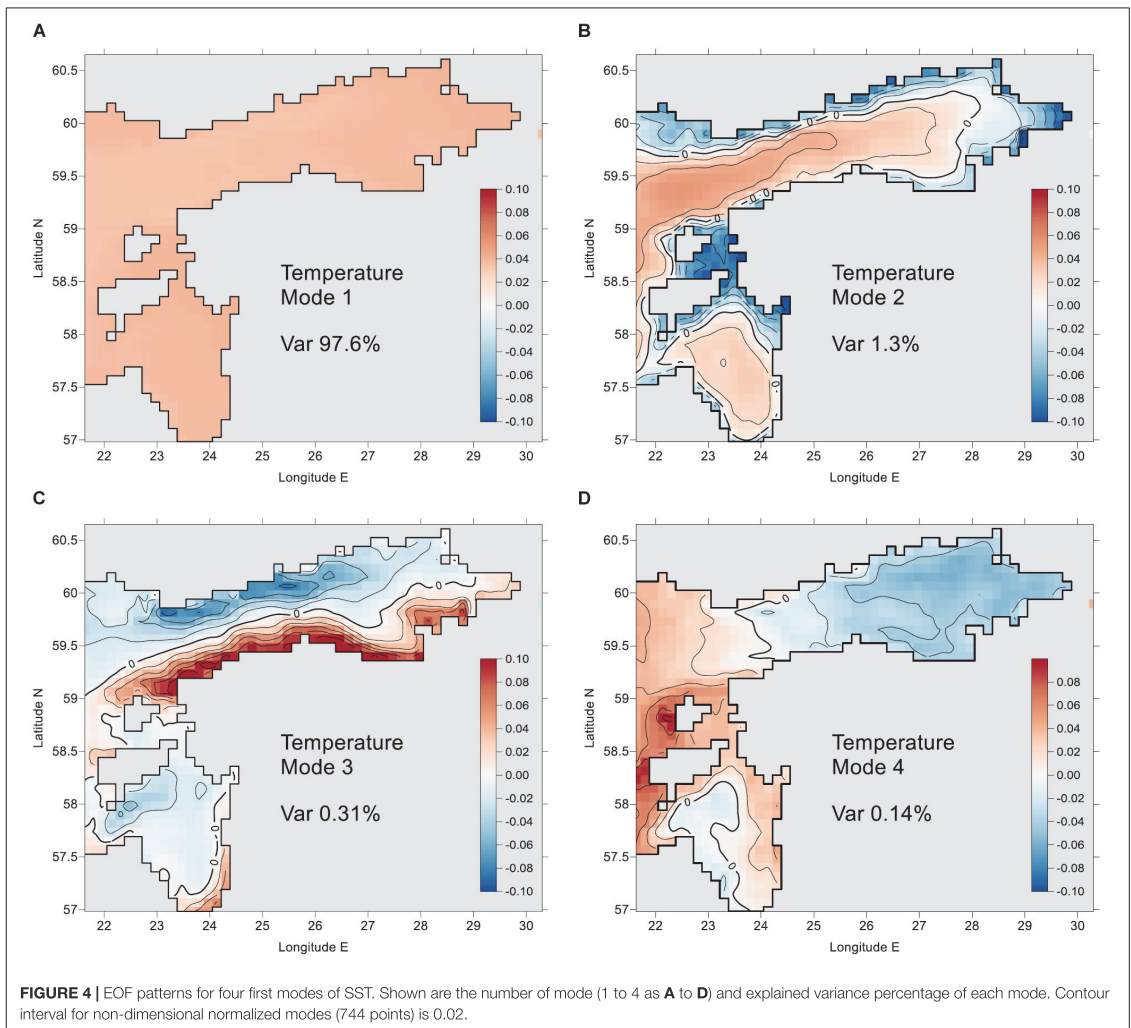




the pseudo-observations. We reveal that error histograms are quite insensitive to the number of observations  $K$ , when it is larger than the number of significant modes  $L$ . However, at small observation amounts the number of larger errors (can be considered as outliers regarding normal distribution) increases. On the background of grid points of 37 km spacing, reconstructed SST and SSS maps are shown for one arbitrary date 19 June

2015 (Figures 8C,D) together with the original model data (Figures 8A,B).

With decreasing number of observations  $K$ , errors slightly increase when still  $K > L$ . For example, SSS absolute error is less than  $0.3 \text{ g kg}^{-1}$  for 88% of cases with  $K = 51$  and 80% of cases with  $K = 10$ . Regarding SST, the errors are less than  $0.6^{\circ}\text{C}$  correspondingly for 90 and 82% cases. Regression of all



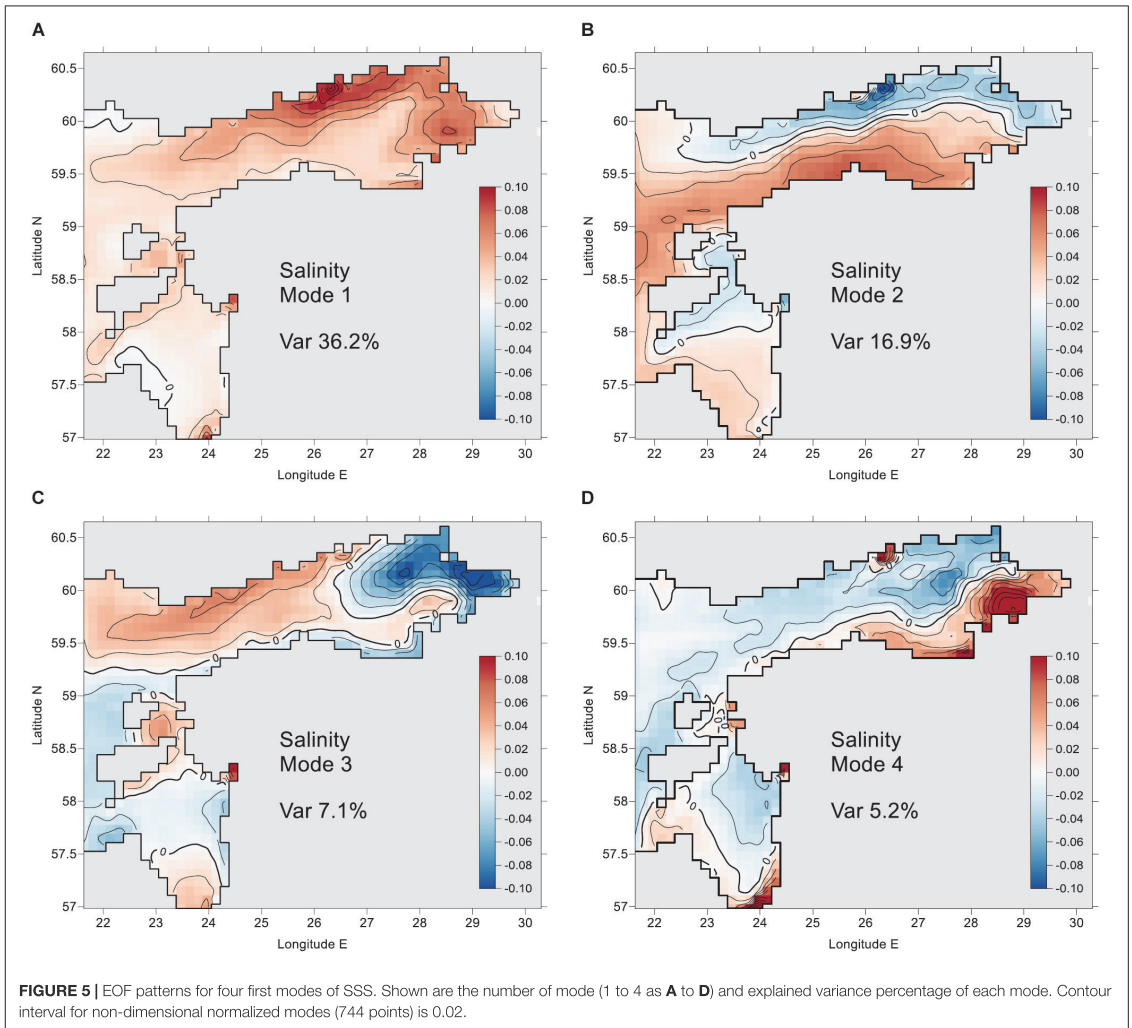
**FIGURE 4 |** EOF patterns for four first modes of SST. Shown are the number of mode (1 to 4 as **A** to **D**) and explained variance percentage of each mode. Contour interval for non-dimensional normalized modes (744 points) is 0.02.

the values of both SST and SSS yields tangent between initial and reconstructed data 0.99, their correlations follow  $r > 0.95$ . Relative errors of all the SST data, compared with horizontal standard deviation  $\sigma_x(t)$  of each time instance, are from 6.7% (observation grid step 37 km) to 8.6% (93 km). Relative errors of SSS are somewhat larger, correspondingly 18 and 25%. For  $K < L$  the errors increase abruptly and singularity errors may occur in Equations (5)–(7).

Reconstruction capability from realistic sampling schemes was evaluated for typical monitoring network (with smaller number of stations than usual) and for two routes of FerryBox along Tallinn–Helsinki and Tallinn–Stockholm (Petersen, 2014; Kikas and Lips, 2016). Pseudo-observations from the selected configurations were run through all the daily model maps. The error statistics did not differ much from that of the

above-described observation grid experiments. Inspecting the reconstructed maps (not shown), even with small number of observations the reconstructed maps generally match well to the original maps. The monitoring type of stations has observations in all the three main areas: Gulf of Finland, Gulf of Riga and adjacent Baltic Proper. The SST and SSS maps reconstructed from the pseudo-observations (not shown) match well the original maps. FerryBox data set has no data in the eastern Gulf of Finland and in the Gulf of Riga (see an example by Elken et al., 2018), therefore larger deviations of reconstructed data from initial data occur in these regions. However, main large-scale SST and SSS features, present in the initial model data, can be identified in the reconstructed maps rather well.

For comparison of EOF reconstruction with OI, we set up an experiment where EOF statistics were calculated during 4 years



from 1 July 2010 to 30 June 2014, and the remaining 1 year period from 1 July 2014 to 30 June 2015 was dedicated for comparison. For each set of  $\Delta_o = n \Delta_m$ , we included all the possible shifts of observation grid into the comparison. For example, in case of  $n = 8$  there are 64 options for grid shift. All those shift options produce different result of reconstruction, because of different resolution of topographic and coastline features and freshwater input areas. The points just neighboring the coast were excluded. The method of OI was used with correlation scale  $R = 200$  km and noise-to-signal ratio  $\eta^2 = 0.1$  (see section Estimation of Reconstruction Accuracy Using Pseudo-Observations).

Dependence of RMSD of reconstruction on the spacing for 3 compared methods – EOF, OI-M with modeled mean field, and OI-C with climatological mean fields – is presented in **Figure 9** as median values taken over all shift options. The spread

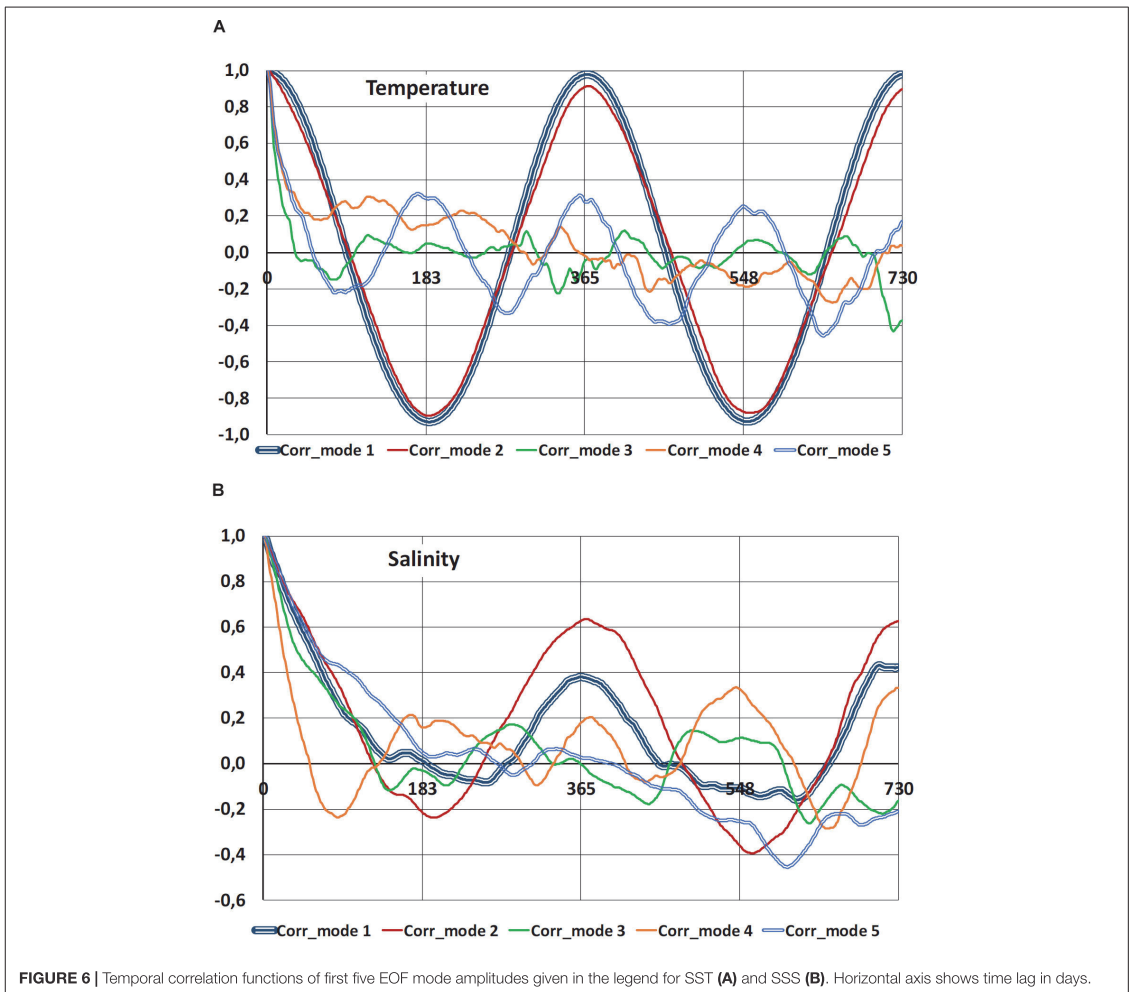
**TABLE 1** | Characteristics of SST modes.

Mode nr	% variance	Description
1	97.6%	Nearly uniform over space increase or decrease of SST, represents seasonal heating and cooling.
2	1.3%	Faster heating (in spring) or cooling (in autumn) in the shallow coastal areas, compared with deeper offshore areas.
3	0.31%	Transverse colder or warmer anomaly stripes near northern or southern coasts, like upwelling and downwelling.
4	0.14%	Longitudinal colder or warmer anomalies appearing in east-west direction.
5	0.10%	Different heating or cooling of the SW Gulf of Riga and NW-N Gulf of Finland.
6	0.07%	Physics not clear.

**TABLE 2 |** Characteristics of SSS modes.

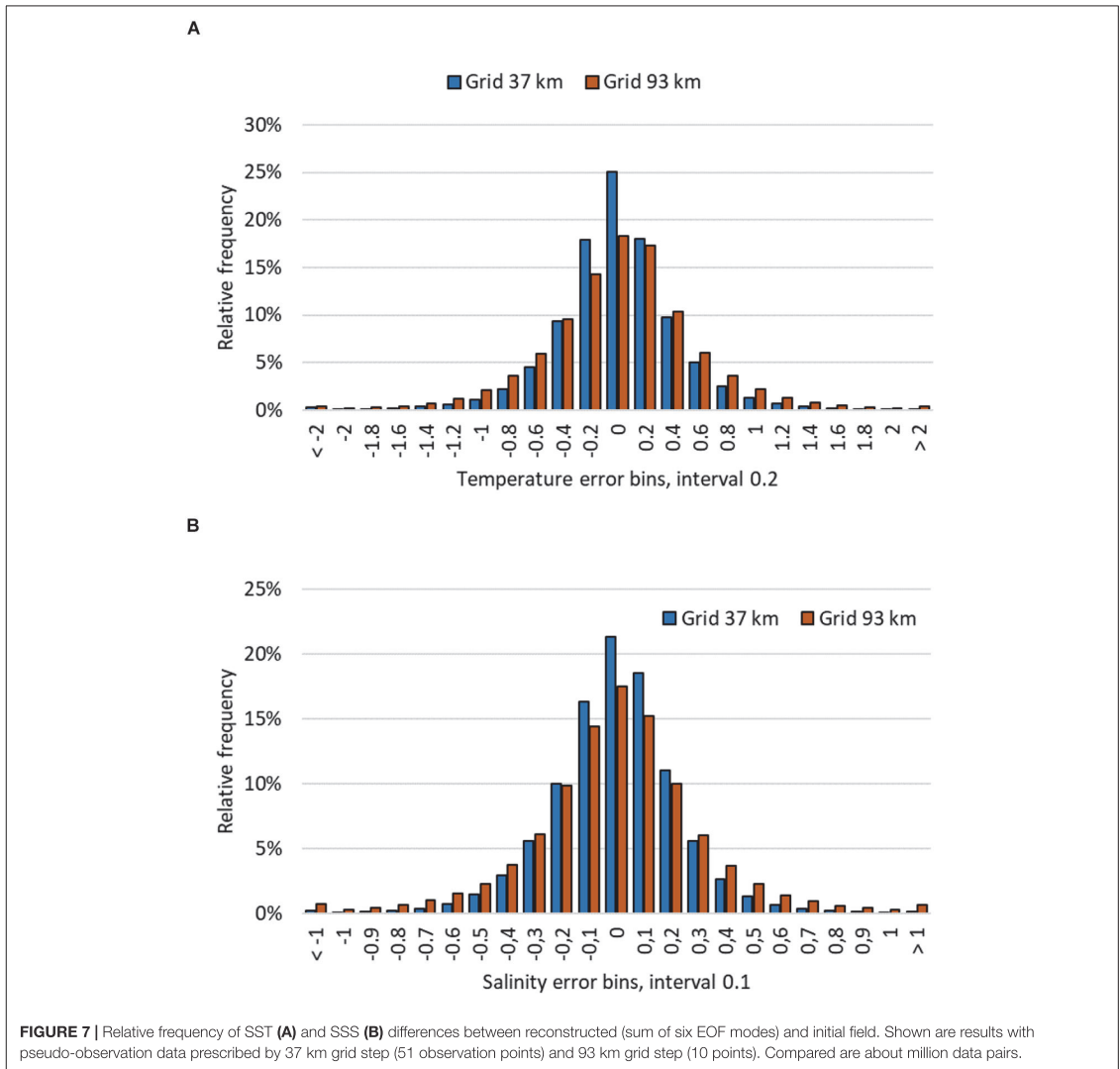
Mode nr	% variance	Description
1	36.2%	Increase or decrease of salinity over the whole study area (all changes have the same sign). Larger changes occur in the northeastern Gulf of Finland, near the discharge of the largest rivers in the region.
2	16.9%	Transverse anomalies of salinity near northern or southern coasts, like upwelling and downwelling.
3	7.1%	Salinity changes in the freshwater spreading pathway near the northern coast of the Gulf of Finland, reminds cyclonic circulation.
4	5.2%	Salinity changes near the southeastern coasts, characteristic of cyclonic and anticyclonic circulation.
5	4.1%	Physics not clear.
6	3.5%	Physics not clear.

of individual shift estimates increases from the spacing 46 km toward greater spacing and smaller number of pseudo-observations, especially for the EOF method. In case of SST, EOF methods produces on the average more accurate reconstruction than both of the OI methods using the values given above. Still, during individual time instances the reconstruction results may have similar difference pattern at large spacing of sampling (Figures 10A,C) since coastal features may remain unresolved. We had to choose large correlation scale and noise-to signal ratio in order to have reconstruction over the whole area even if the spacing of observations is large; then OI tends to make heavy smoothing that is reflected in Figure 9A by larger RMSD. SSS field is dominated by spatial variations; one control value of such domination is RMSD = 0.411 g kg<sup>-1</sup> of “no observations” (at each grid point, mean value is taken instead of observed value) that lies in the range of RMSD variation (Figure 9B).



**FIGURE 6 |** Temporal correlation functions of first five EOF mode amplitudes given in the legend for SST (A) and SSS (B). Horizontal axis shows time lag in days.



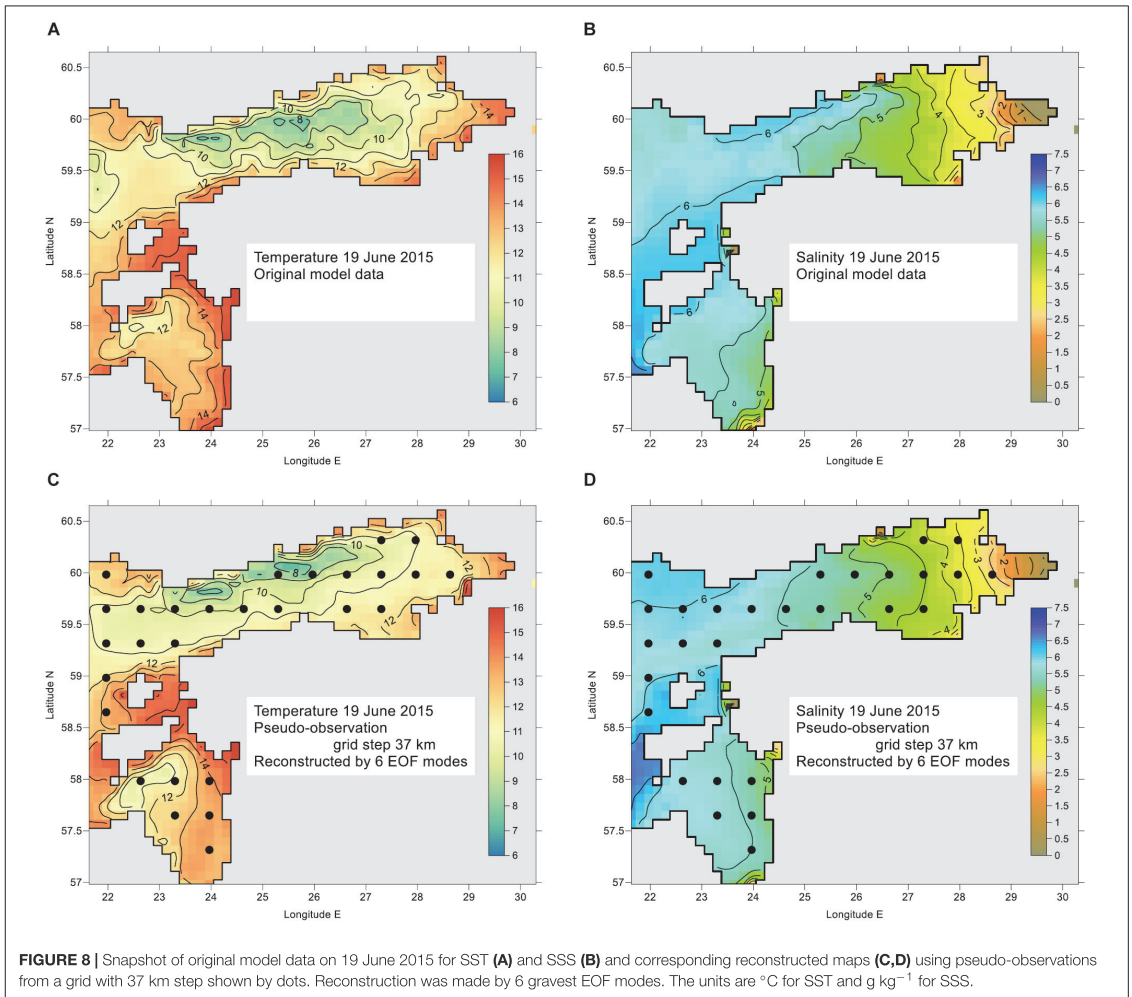


Evaluation reveals that EOF reconstruction has slightly larger error than OI relative to the modeled mean field. Commonly used OI with deviations from climatological mean reveals larger RMSD than EOF; it even exceeds the “no observations” already with spacing larger than 28 km. These features of RMSD appear due to the specific topographic and hydrographic features of the region: fragmented coastline and prevalence of low-salinity regions with surrounding higher spatial gradients and temporal standard deviations (Figure 2) near the entrances of larger rivers (Figure 1). If OI considers and interpolates the SSS deviations from highly variable spatial mean map (determined by the model that resolves local features), then such deviations follow normal distribution without significant outliers

(not shown) and local features appear in the reconstruction product without remarkable distortion. Deviations from spatially smooth climatological mean values (Figure 2B) contain a high number of outliers to the normal distribution, that cause larger distortions of OI-C reconstruction in the river influence areas than EOF reconstruction (Figures 10B,D).

### Seasonality Issues in EOF Reconstruction of SST

Among variety of physical processes, original SST data from model reveal significant seasonal variation in time. Annual cycle is evident in temporal correlation of the amplitudes of first and



**FIGURE 8** | Snapshot of original model data on 19 June 2015 for SST (A) and SSS (B) and corresponding reconstructed maps (C,D) using pseudo-observations from a grid with 37 km step shown by dots. Reconstruction was made by 6 gravest EOF modes. The units are °C for SST and g kg<sup>-1</sup> for SSS.

second EOF mode (Figure 6) that cover 98.9% of total variance. This cycle is slightly variable in space, whereas highest spatial variations (not shown) occur during the spring heating period and smallest variations take place in the winter when SST is close to or equal to freezing temperature.

It is interesting to consider what will happen to the EOF reconstruction results when seasonal signal is removed from space-time matrix  $\mathbf{X}$  prior to the procedures by Equations (1)–(7). Following Høyer and She (2007), we introduce a modified data set where time slices of spatial data with seasonality removed are defined at time index  $i$  as

$$\mathbf{x}_i^s = \mathbf{x}_i - \mathbf{s}_i, \tag{10}$$

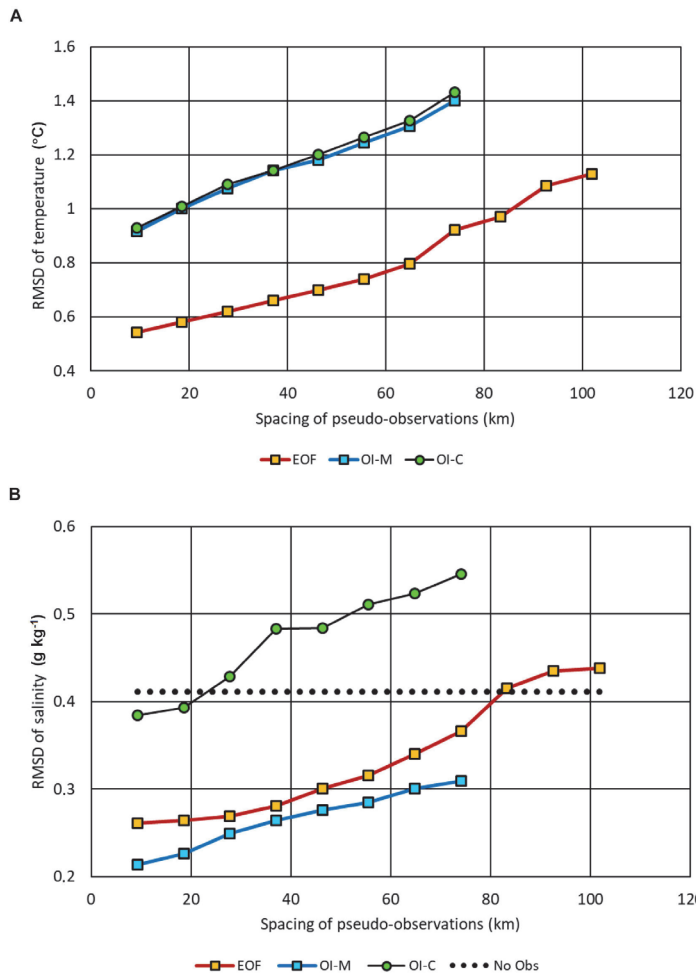
where seasonal data  $\mathbf{s}_i$  are evaluated in each model grid point  $k$ . Consider a time series vector  $\mathbf{x}_k$  which values are available on times  $t_i$  counted as fractions of decimal year. Based on the total

$M$  data of  $x_k(t_i)$ , make an approximation of seasonal cycle by a biharmonic function

$$s_k(t) = C_{1,k} \sin 2\pi t + C_{2,k} \cos 2\pi t + C_{3,k} \sin 4\pi t + C_{4,k} \cos 4\pi t + C_{5,k}. \tag{11}$$

The coefficients from  $C_{1,k}$  to  $C_{5,k}$  are found to obtain best fit of  $s_k(t_i)$  to the values  $x_k(t_i)$  in terms of minimizing their RMSD. The fitting coefficients and resulting seasonal cycle are spatially variable, whereas earlier and higher SST maxima generally appear in shallower coastal waters.

Overall variance of SST, determined in reference to the constant mean value, was 47.35 (°C)<sup>2</sup> whereas spatial variability due to temporally constant mean values in each grid point covered 0.3%. By introducing the seasonal cycle removal procedure by Equations (10)–(11), variance percentage of the

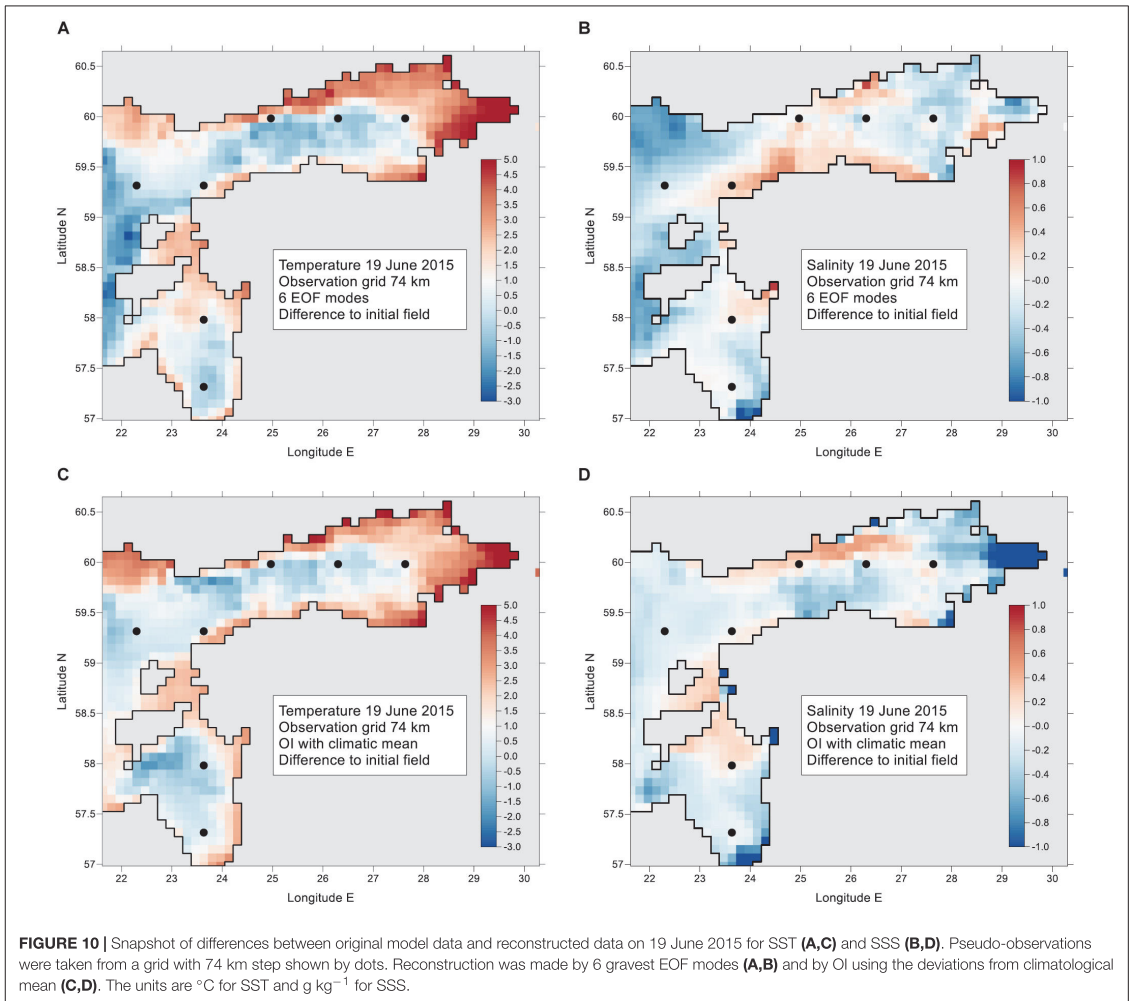


**FIGURE 9 |** Dependence of RMSD between reconstructed from pseudo-observations fields and original model output fields of SST **(A)** and SSS **(B)** for different methods: EOF, reconstruction by 6 gravest modes; OI-M, optimal interpolation with modeled mean field; OI-C, optimal interpolation with climatological mean field. The grid step of pseudo-observations cycled from 1 to 11 model grid steps, all possible shifted configurations of off-shore stations were taken into account. Median values of RMSD are presented. “No Obs” means the spatial standard deviation of temporal mean values.

input field for the EOF analysis was significantly reduced: from 99.7% for  $\mathbf{x}_i$  to 6.1% for  $\mathbf{x}_i^s$ . The fitting biharmonic seasonal cycle contained 80.0% of total variance and the remaining 13.8% appeared in the covariance between  $\mathbf{x}_i^s$  and  $\mathbf{s}_i$ .

Although variability of  $\mathbf{x}_i^s$  was reduced as compared with  $\mathbf{x}_i$ , spatially mean deviation from the seasonal cycle was typically in the range from  $-2^\circ\text{C}$  (mostly in autumn) to  $+4^\circ\text{C}$  (in summer). Spatial standard deviations of  $\mathbf{x}_i^s$  had maximum values during summer, amounting typically to  $2.5^\circ\text{C}$ . Wintertime minimum of spatial standard deviation, apparent in the initial  $\mathbf{x}_i$  data, was not anymore apparent since biharmonic  $s_k(t_i)$  had in winter problems to follow the constant level of freezing temperature.

Spatial covariance estimates of  $\mathbf{x}_i^s$  (not shown) reveal significant similarity to the estimates based on the original data  $\mathbf{x}_i$  (**Figure 3A**): covariance at distances of several hundreds of kilometers is close to the covariance at zero lag since significant part of SST variability is caused by weather events and interannual variations that occur nearly uniformly over smaller sub-regions like in our case. Based on the full covariance matrix, EOF analysis revealed highly similar patterns of leading modes of  $\mathbf{x}_i^s$  to the modes of  $\mathbf{x}_i$  which are shown in **Figure 4**. The share of “flat” first mode (**Figure 4A**) decreased from 97.6 to 80.5% after removal of biharmonic seasonal cycle. At the same time, the shares of higher modes 2–6 increased from 1.91 to 12.79%.



In addition, the second and third mode changed their order and the “upwelling” mode (Figure 4C for  $x_i$ ) got a bit higher share of variance as the “differential heating” mode (Figure 4B for  $x_i$ ) since those variations were already partly included into the spatially variable seasonal cycle.

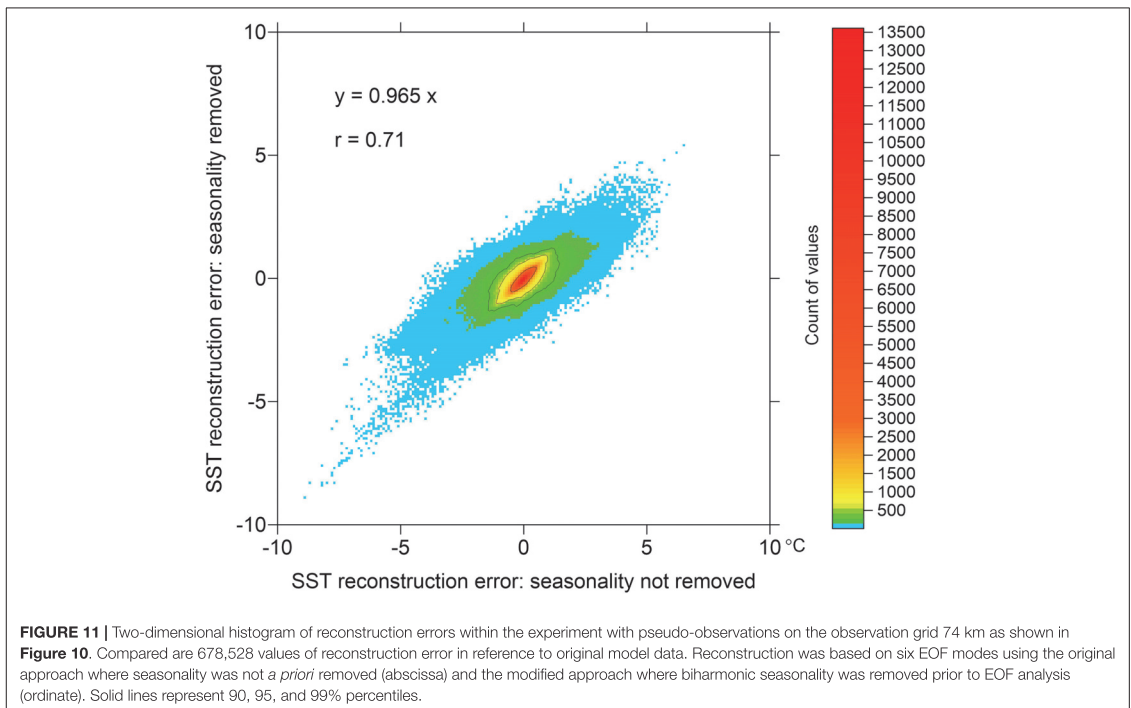
The original model field  $X$  can be approximated by leading EOF modes using equation (1), whereas the higher eigenvectors are truncated to zero. Using six leading modes both for  $x_i$  and  $x_i^s$  datasets, the decompositions seem highly similar. However, evaluation of RMSD allowed detecting of 4% reduction when seasonal cycle was removed prior to the EOF procedures.

When there are much less observations than the number of grid points, reconstruction accuracy can be estimated using the pseudo-observations method. Example of comparison of the two datasets is presented in Figure 11, based on the “observation” data that were extracted in seven locations shown in Figure 10.

The reconstruction errors were correlated with  $r = 0.71$ , whereas the regression line was (errors of  $x_i^s$ ) = 0.965 (errors of  $x_i$ ). Both reconstructions had very high correlation with the initial model data  $r > 0.99$  and the scatterplot graphs (not shown) created the impression that the data sets were nearly identical. Actually, already small variations in the correlation modify RMSD and in case of our example, there can be about 25% of RMSD reduction when seasonality is removed prior to the EOF analysis.

### Splitting the Region Into Sub-Areas

In one of the experiments, the whole region presented in Figure 1, was split in three sub-regions: Gulf of Finland (GOF), Gulf of Riga (GOR), and NE Baltic Proper (NEBP). Individual EOF modes were calculated for each of the sub-area. Except for NEBP, the first two SST modes for GOF and GOR were similar to the patterns obtained for the whole area. Pairwise correlations of the SST



amplitudes were  $r > 0.95$  between the GOF, GOR and the whole region. The whole region and GOR correlations for the first three SSS modes were  $r > 0.9$  while the first mode of GOR correlated with GOF with  $r > 0.72$  and with the whole region with  $r > 0.78$ . The first three modes of SSS of the whole region covered 61.4% of variance. The split regions had the mode coverage: GOF – 69.9%, GOR – 53.7%, NEBP – 72.8%. Although by splitting the region, mode convergence (share of variance of the lowest modes) increased slightly, we judged that EOFs of the whole region cover the regional dynamics in sufficient accuracy. Note, that region splitting may become important in other regions, where the mode convergence of the whole region might not be satisfactory.

## Examples Using Actual Observations Taking Into Account Time Dependence of Observations

It is usual practice, that spatial shipborne monitoring is carried out by different ships belonging to different institutes and countries. Covering the whole region may take quite long time. On an example, given in **Figure 12**, four ships with ICES codes 3499, 34AR, ESLV and LAVA made observations during 17 days from 18 May to 3 June 2009. During this spring heating period, SST generally increased from 7 to 15°C, but local SST variations were also evident. Over the time span, observations in the northern area were taken in the first part when water was not yet heated as much as by the end of the period. Considering the observations as instantaneous, reconstruction

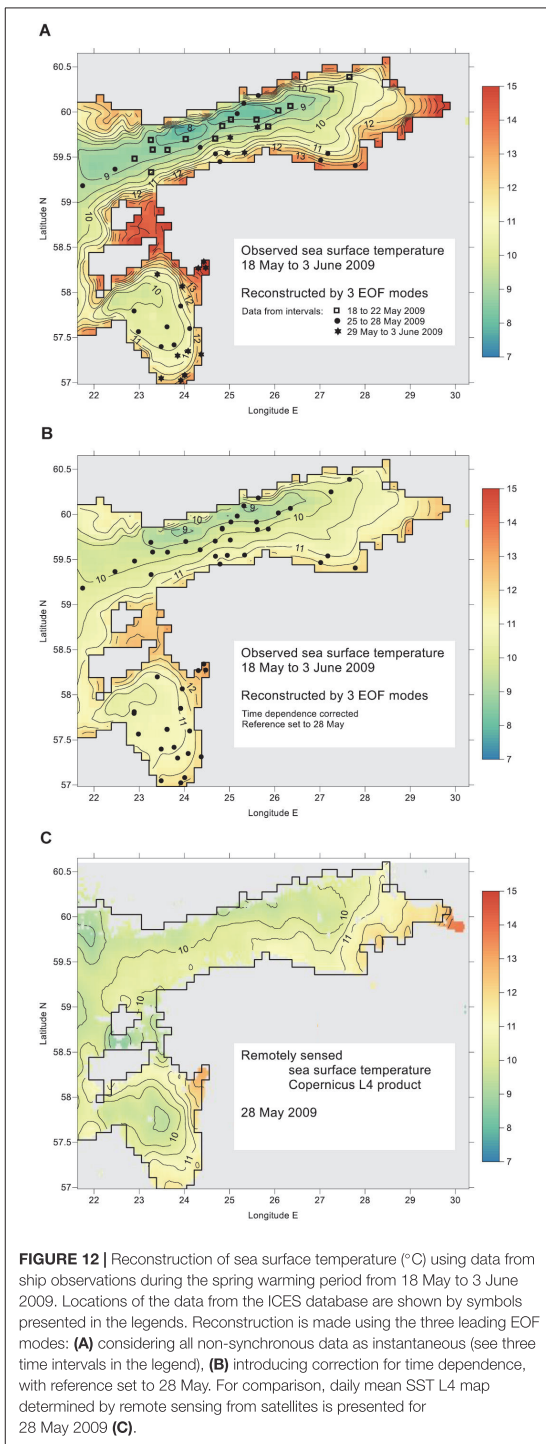
using Equations (5)–(7) provided rather cold waters there. In turn, warm waters were drawn in the coastal areas where observations were taken at the end of the period (**Figure 12A**). EOF reconstruction using the time dependence of observations based on Equations (8)–(9), setting the reference time in the middle of the period (**Figure 12B**), increased the temperature in the region of earlier observations and decreased in the region of later observations, reducing this way the artificial contrasts due to non-synchronous observations. Comparison with satellite based SST map from CMEMS L4 product (**Figure 12C**) reveals good similarity to the time-corrected map. Numerical differences are mostly less than 1°C, not exceeding the range of unresolved here diurnal oscillations (Karagali and Høyer, 2014). From the number of calculations we got the experience that reference time may be modified within the observational window without loosing the realism of reconstruction. However, extrapolation outside the window should be avoided like in the case of one-dimensional linear regression.

## Automatic Reconstruction of Time Series of Maps

It is technically easy to set up procedures for automatic reconstruction of time series of maps, using the time dependence of observations based on Equations (8)–(9). We took the shipborne profile data from ICES database (see section Regional Setting of Experiments).

During the reconstruction procedure, EOF amplitudes for each map were checked against the  $|\hat{a}_{i,k}| < F_k = 2 \sigma(\hat{a}_k)$  criteria





**FIGURE 12 |** Reconstruction of sea surface temperature (°C) using data from ship observations during the spring warming period from 18 May to 3 June 2009. Locations of the data from the ICES database are shown by symbols presented in the legends. Reconstruction is made using the three leading EOF modes: **(A)** considering all non-synchronous data as instantaneous (see three time intervals in the legend), **(B)** introducing correction for time dependence, with reference set to 28 May. For comparison, daily mean SST L4 map determined by remote sensing from satellites is presented for 28 May 2009 **(C)**.

(see section Covariance and EOF Characteristics) in order to determine the number of “good modes.” Taking the time window for one map 1/12 year (about 1 month) and the limit of at least 8 non-duplicate observations within the time window (time average was taken over 2 days), we obtained 148 maps for SST and 137 maps for SSS, out of maximum 209 maps. Average number of data points is 25 and maximum amounts to 83 both for SST and SSS. We note that large number of observation points does not necessarily mean high number of good EOF modes. For example, the observations with highest amount 83 were fragmentarily located along the coasts as short-scale repeated maps; the number of good modes was only 3. Good six modes were in other configuration obtained already from 20 to 30 observation points.

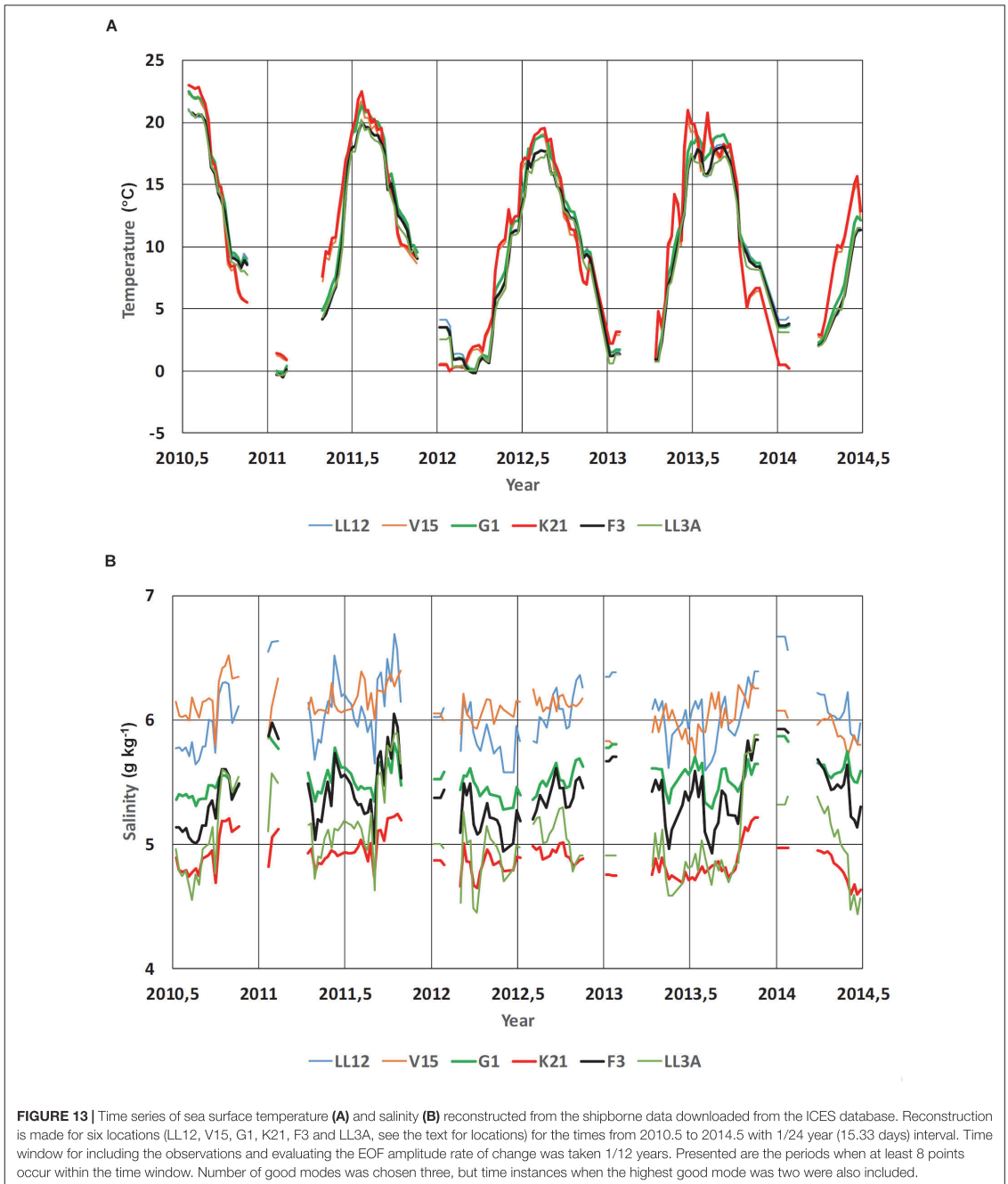
From reconstructed maps, SST and SSS time series were extracted around monitoring stations LL12 (western Gulf of Finland, 59.4835 N, 22.8968 E), V15 (Moonsund, 58.8167 N, 23.2167 E), G1 (central Gulf of Riga, 57.6167 N, 23.6167 E), K21 (Pärnu Bay, 58.2167 N, 24.3083 E), F3 (central Gulf of Finland, 59.8383 N, 24.8383 E) and LL3A (northeastern Gulf of Finland, 60.0672 N, 26.3467 E). Such reconstructed time series can be easily extended to climate studies.

Individual maps, reconstructed by the automatic procedure (example is given in **Figure 13**), follow closely the observed values but also reveal realistic patterns compared to the published knowledge on Baltic Sea climatology and monthly and instantaneous distributions (e.g., Leppäranta and Myrberg, 2009).

## DISCUSSION

There is a continued need for producing gap-free gridded oceanographic data using observations. Although new observation techniques became available, the problem of fragmentation remains in oceanographic data management. Reanalysis is a powerful, but costly method for production of gridded fields. Widespread statistically based methods like OI (optimal interpolation), DIVA (Data-Interpolating Variational Analysis, Troupin et al., 2010) etc. use mainly localized covariance patterns. Covariance estimates from models reveal large values over basin scales due to trends, seasonal signal and basin-wide dynamics (e.g., coherent upwelling-downwelling near opposite coasts). Such covariance estimates suggest using of methods that use full covariance fields. In this context, the classical EOF method is again gaining interest (e.g., Yang et al., 2017; Pilo et al., 2018), whereas the statistics of the studied field can be estimated from the model results. Amplitudes of EOF modes can be approximately estimated from the observational data set which dimension is much smaller than the number of model EOF mode grid points.

When using model data to create the EOF statistics, it is important to know how reliable the estimates of modes and amplitudes with respect to model uncertainties are. Thorough treatment of the above question cannot be found in the literature. However, on the sea surface, temperature and salinity results from different models are rather well validated by observations and the model-based covariance patterns can be considered trustful. As a common practice, modeled covariance have been



used in data assimilation. Fu et al. (2011) compared covariance patterns from modeled SST and satellite SST, and found them agreeing well. CMEMS QUID report has presented validation of

SSS against ferrybox data, showing that the SSS patterns were well simulated by the model. In deeper layers, however, there is usually larger spread between different model results.

While ocean data come from different platforms and observations are non-synchronous, taking into account the time dependence of observations is a challenge. Amplitudes of lowest EOF modes reveal distinct temporal correlation patterns, with time scale several months. This could be used to handle temporal gaps and/or non-synchronous data, if they lie within the temporal correlation scale. As a first step, we introduced linear correction to the EOF amplitude depending on the difference of observation time relative to reference time.

We have made comparison of the results from EOF reconstruction with the results from classical OI method. Our test region – NE Baltic – is characterized by fragmented coastline and highly variable topographic and hydrographic conditions. In such region the spatial changes of oceanographic variables may have quasi-permanent anomalies like low salinity near river influence areas or faster warming and cooling in shallow coastal areas, compared to the offshore SST variations. Therefore, EOF reconstruction has a potential to achieve comparable or better accuracy than OI method.

EOF modes and amplitudes depend on the selection of domain. We used a sub-region, containing two geographical regions and their transition area. It was found that some of the modes (mainly the first mode) do not change significantly when the domain is separated into smaller parts. If the convergence of leading EOF modes of large sea area is low (share of explained variance is small), refinement into smaller areas might be useful. In our example, the convergence improved only slightly when partition into three smaller areas was made. Workable criteria for the region selection is not yet established, although following geographical regions seems to be acceptable.

In case of data assimilation into the high-resolution model, it is reasonable to separate low-resolution component and make large-scale corrections into that, keeping high-resolution deviation patterns unchanged. The workable approach for correction and/or assimilation of the data on the basin scales is as follows: (1) to separate the coarse-grid part from the fine grid data by spatial averaging, (2) to perform corrections on the smaller dimension coarse grid, (3) to interpolate coarse grid data back to the fine grid and add fine grid deviations determined from the interpolation of the initial coarse grid. This is based on the assumption, that correction of basin-scale features does not influence the mesoscale patterns, apparent on the fine grid but filtered out on the coarse grid. The EOF approach allows additional assimilation of mesoscale patterns in the regions of high data coverage.

We have tested reconstruction of SST and SSS in one sub-region of the Baltic, based on the *in situ* observations. The method itself allows to be applied on different data sets (for example, including high-resolution remotely sensed SST) of different variables and their combinations (for example, joint data vector of temperature and salinity), on the condition that significant part of variability can be presented by a few leading modes. There could be obvious extensions of the approach to

cover the whole water column. This could be especially important to properly match the consequences of large salt water inflows etc.

One straightforward application of the approach could be in marine ecology, where building the gap-free patterns of nutrients and biomass variables could allow more precisely estimate the total amounts and budgets of ecosystem variables, and to evaluate the values of environmental indicators that are important for environmental management.

## CONCLUSION

We have developed statistically justified EOF reconstruction method to handle large-scale patterns of observed fields in the sub-regions. The method uses model-based EOF patterns to interpolate and extend the observational data over the full study region. In the smaller sea regions, which are affected by the same large-scale forcing patterns, the EOF patterns have obvious physical interpretations and their shape does not depend very much on the selection of boundaries. When removing the SST seasonal cycle prior to EOF analysis, spatial patterns of leading modes remained practically unchanged, share of variance of the three first modes was reduced from 99 to 88.6% and reconstruction errors were reduced by about 25%.

Since we use only the first most energetic EOF modes, we can cover with this method basin and sub-basin scales of variability. The relative interpolation errors, estimated over the full area, usually remain below 10% for SST and 20% for SSS, compared with multi-year standard deviation of all variability relative to their mean value over the basin. In comparing with OI, EOF is especially useful for reconstruction with very sparse observations. In the regions of denser sampling, EOF cannot exactly follow the observations. Mesoscale deviations from large-scale EOF patterns follow well-defined covariance decay with space lag; therefore, they could be treated by optimal interpolation or similar method.

## AUTHOR CONTRIBUTIONS

All authors contributed in analysis and writing, with JE as lead author and JS making strong co-contribution. MZ made the EOF calculations while PL was in charge with the HBM model.

## FUNDING

This study was supported by the Ph.D. program for MZ and the institutional research funding IUT 19-6 of the Estonian Ministry of Education and Research.

## ACKNOWLEDGMENTS

A larger BAL MFC team within the EU projects MyOcean, MyOcean2, and MyOcean-FO did the development of HBM model. This cooperation is highly acknowledged.



## REFERENCES

- Alenius, P., Nekrasov, A., and Myrberg, K. (2003). Variability of the baroclinic Rossby radius in the Gulf of Finland. *Cont. Shelf Res.* 23, 563–573. doi: 10.1016/S0278-4343(03)00004-9
- Alvera-Azcárate, A., Vanhellemont, Q., Ruddick, K., Barth, A., and Beckers, J. M. (2015). Analysis of high frequency geostationary ocean colour data using DINEOF. *Estuar. Coast. Shelf Sci.* 159, 28–36. doi: 10.1016/j.ecss.2015.03.026
- Beckers, J. M., and Rixen, M. (2003). EOF calculations and data filling from incomplete oceanographic datasets. *J. Atmos. Ocean. Technol.* 20, 1839–1856. doi: 10.1175/1520-0426(2003)020<1839:ecadff>2.0.co;2
- Berg, P., and Poulsen, J. W. (2012). *Implementation Details for HBM*. DMI Technical Report No. 12-11. Copenhagen.
- Cushman-Roisin, B., and Beckers, J. M. (2011). *Introduction to Geophysical Fluid Dynamics: Physical and Numerical Aspects*. Cambridge, MA: Academic Press, 875.
- Davis, R. E. (1976). Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.* 6, 249–266. doi: 10.1073/pnas.1610708114
- Elken, J., and Matthäus, W. (2008). “Baltic Sea oceanography,” in *Regional Climate Studies, Assessment of Climate Change for the Baltic Sea Basin Annex A*, eds H.-J. Bolle, M. Meneti, and I. Rasool. (Berlin: Springer), 379–385.
- Elken, J., Nömm, M., and Lagema, P. (2011). Circulation patterns in the Gulf of Finland derived from the EOF analysis of model results. *Boreal Environ. Res.* 16, 84–102.
- Elken, J., Zujev, M., and Lagema, P. (2018). “Reconstructing sea surface temperature and salinity fields in the northeastern baltic from observational data, based on sub-regional Empirical Orthogonal Function (EOF) patterns from models,” in *Proceedings of the IEEE/OES Baltic International Symposium (BALTIC)*, Klaipeda.
- FMI, (2018). *Ice Winter in The Baltic Sea*. Helsinki: Finnish Meteorological Institute.
- Fu, W., Høyer, J. L., and She, J. (2011). Assessment of the three dimensional temperature and salinity observational networks in the Baltic Sea and North Sea. *Ocean Sci.* 7, 75–90. doi: 10.5194/os-7-75-2011
- Gandin, L. S. (1963). *Objective Analysis of Meteorological Fields*. Jerusalem: Israel program for scientific translations, 242.
- Ghil, M. (1989). Meteorological data assimilation for oceanographers. Part I: description and theoretical framework. *Dyn. Atmos. Oceans* 13, 171–218. doi: 10.1016/0377-0265(89)90040-7
- Ghil, M., and Malanotte-Rizzoli, P. (1991). Data assimilation in meteorology and oceanography. *Adv. Geophys.* 33, 141–266.
- Høyer, J. L., and She, J. (2007). Optimal interpolation of sea surface temperature for the North Sea and Baltic Sea. *J. Mar. Syst.* 65, 176–189. doi: 10.1016/j.jmarsys.2005.03.008
- Ide, K., Courtier, P., Ghil, M., and Lorenc, A. C. (1997). Unified notation for data assimilation: operational, sequential and variational. *J. Meteorol. Soc. Japan Ser. II* 75, 181–189. doi: 10.2151/jmsj1965.75.1b\_181
- Janssen, F., Schrum, C., and Backhaus, J. O. (1999). A climatological data set of temperature and salinity for the Baltic Sea and the North Sea. *Deutsche Hydrografische Zeitschrift* 51, 5–245. doi: 10.1007/bf02933676
- Jayaram, C., Priyadarshi, N., Pavan Kumar, J., Udaya Bhaskar, T. V. S., Raju, D., and Kochuparampil, A. J. (2018). Analysis of gap-free chlorophyll-a data from MODIS in Arabian Sea, reconstructed using DINEOF. *Int. J. Remote Sens.* 39, 7506–7522. doi: 10.1080/01431161.2018.1471540
- Kaplan, A., Kushnir, Y., Cane, M. A., and Blumenthal, M. B. (1997). Reduced space optimum analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *J. Geophys. Res. Oceans* 102, 27835–27860. doi: 10.1029/97jc01734
- Karagali, I., and Høyer, J. L. (2014). Characterisation and quantification of regional diurnal SST cycles from SEVIRI. *Ocean Sci.* 10, 745–758. doi: 10.5194/os-10-745-2014
- Kikas, V., and Lips, U. (2016). Upwelling characteristics in the Gulf of Finland (Baltic Sea) as revealed by Ferrybox measurements in 2007–2013. *Ocean Sci.* 12, 843–859. doi: 10.5194/os-12-843-2016
- Kim, K. Y. (1997). Statistical interpolation using cyclostationary EOFs. *J. Clim.* 10, 2931–2942. doi: 10.1175/1520-0442(1997)010<2931:siuce>2.0.co;2
- Laanemets, J., Väli, G., Zhurbas, V., Elken, J., Lips, L., and Lips, U. (2011). Simulation of mesoscale structures and nutrient transport during summer upwelling events in the Gulf of Finland in 2006. *Boreal Environ. Res.* 16, 15–26.
- Lagema, P. (2012). Operational forecasting in Estonian marine waters. *TUT Press B* 128, 130.
- Legrand, C., Fridolfsson, E., Bertos-Fortis, M., Lindehoff, E., Larsson, P., Pinhassi, J., et al. (2015). Interannual variability of phyto-bacterioplankton biomass and production in coastal and offshore waters of the Baltic Sea. *AMBIO* 44, 427–438. doi: 10.1007/s13280-015-0662-8
- Lehmann, A., Myrberg, K., and Höfllich, K. (2012). A statistical approach to coastal upwelling in the Baltic Sea based on the analysis of satellite data for 1990–2009. *Oceanologia* 54, 369–393. doi: 10.5697/oc.54-3.369
- Leppäranta, M., and Myrberg, K. (2009). *Physical Oceanography of the Baltic Sea*. Berlin: Springer Science & Business Media, 378.
- Menemenlis, D., Fieguth, P., Wunsch, C., and Willsky, A. (1997). Adaptation of a fast optimal interpolation algorithm to the mapping of oceanographic data. *J. Geophys. Res. Oceans* 102, 10573–10584. doi: 10.1029/97jc00697
- Petersen, W. (2014). FerryBox systems: state-of-the-art in Europe and future development. *J. Mar. Syst.* 140, 4–12. doi: 10.1016/j.jmarsys.2014.07.003
- Pilo, G. S., Oke, P. R., Coleman, R., Rykova, T., and Ridgway, K. (2018). Impact of data assimilation on vertical velocities in an eddy resolving ocean model. *Ocean Modell.* 131, 71–85. doi: 10.1016/j.oceomod.2018.09.003
- Smith, T. M., Reynolds, R. W., Livezey, R. E., and Stokes, D. C. (1996). Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J. Clim.* 9, 1403–1420. doi: 10.1175/1520-0442(1996)009<1403:rohsst>2.0.co;2
- Soosaar, E., Maljutenko, I., Uiboupin, R., Skudra, M., and Raudsepp, U. (2016). River bulge evolution and dynamics in a non-tidal sea—Daugava River plume in the Gulf of Riga. *Baltic Sea. Ocean Sci.* 12, 417–432. doi: 10.5194/os-12-417-2016
- Troupin, C., Machin, F., Ouberdous, M., Sirjacobs, D., Barth, A., and Beckers, J. M. (2010). High-resolution climatology of the northeast atlantic using data-interpolating variational analysis (DIVA). *J. Geophys. Res. Oceans* 115: C08005.
- Vihma, T., and Haapala, J. (2009). Geophysics of sea ice in the Baltic Sea: a review. *Prog. Oceanogr.* 80, 129–148. doi: 10.1016/j.pocan.2009.02.002
- Westerlund, A., Tuomi, L., Alenius, P., Myrberg, K., Miettunen, E., Vankevich, R. E., et al. (2019). Circulation patterns in the Gulf of Finland from daily to seasonal time scales. *Tellus A Dyn. Meteorol. Oceanogr.* 1–24. doi: 10.1080/16000870.2019.1627149
- Woods, J. D. (1980). Do waves limit turbulent diffusion in the ocean? *Nature* 288, 219–224. doi: 10.1038/288219a0
- Yang, C., Masina, S., and Storto, A. (2017). Historical ocean reanalyses (1900–2010) using different data assimilation strategies. *Q. J. R. Meteorol. Soc.* 143, 479–493. doi: 10.1002/qj.2936
- Zujev, M., and Elken, J. (2018). Testing marine data assimilation in the northeastern Baltic using satellite SST products from copernicus marine environment monitoring service. *Proc. Estonian Acad. Sci.* 67, 217–230.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Elken, Zujev, She and Lagema. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

### **Paper III**

Zujev, M., Elken, J., Lagemaa, P., 2020. Data assimilation of sea surface temperature and salinity using basin-scale EOF reconstruction: a feasibility study in the NE Baltic Sea. *Ocean Science*, 10.5194/os-2020-43.





# Data assimilation of sea surface temperature and salinity using basin-scale reconstruction from empirical orthogonal functions: a feasibility study in the northeastern Baltic Sea

Mihhail Zujev, Jüri Elken, and Priidik Lagemaa

Department of Marine Systems, Tallinn University of Technology, Tallinn, EE12618, Estonia

**Correspondence:** Jüri Elken (juri.elken@taltech.ee)

Received: 11 May 2020 – Discussion started: 3 June 2020

Revised: 5 November 2020 – Accepted: 8 November 2020 – Published: 14 January 2021

**Abstract.** The tested data assimilation (DA) method based on EOF (Empirical Orthogonal Functions) reconstruction of observations decreased centred root-mean-square difference (RMSD) of surface temperature (SST) and salinity (SSS) in reference to observations in the NE Baltic Sea by 22 % and 34 %, respectively, compared to the control run without DA. The method is based on the covariance estimates from long-term model data. The amplitudes of the pre-calculated dominating EOF modes are estimated from point observations using least-squares optimization; the method builds the variables on a regular grid. The study used a large number of in situ FerryBox observations along four ship tracks from 1 May to 31 December 2015, and observations from research vessels. Within DA, observations were reconstructed as daily SST and SSS maps on the coarse grid with a resolution of  $5 \times 10$  arcmin by N and E (ca. 5 nautical miles) and subsequently were interpolated to the fine grid of the prognostic model with a resolution of  $0.5 \times 1$  arcmin by N and E (ca. 0.5 nautical miles). The fine-grid observational fields were used in the DA relaxation scheme with daily interval. DA with EOF reconstruction technique was found to be feasible for further implementation studies, since (1) the method that works on the large-scale patterns (mesoscale features are neglected by taking only the leading EOF modes) improves the high-resolution model performance by a comparable or even better degree than in the other published studies, and (2) the method is computationally effective.

## 1 Introduction

In the coastal oceans and marginal seas, basin-scale observation, modelling and forecasting of oceanographic and biogeochemical variables is a continuing challenge. As an example from the Baltic Sea, large-scale nutrient dynamics (Andersen et al., 2017; Savchuk, 2018) control the level of eutrophication and hypoxia, affected by nutrient loads and changing climate (Meier et al., 2019). Placke et al. (2018) have recently shown, by comparison of different models, that temperature is much better reproduced than salinity. A similar evaluation has been obtained earlier by Golbeck et al. (2015), based on 13 operational models used routinely in the Baltic and North seas.

Data assimilation (DA) is a key element to improve the model accuracy with respect to observations, both in the operational forecast and the reanalysis context (Martin et al., 2015; Buizza et al., 2018; Moore et al., 2019). DA methods are built upon dynamical models and they are based on some kind of minimization (minimum variance, variational cost function formulation etc.) of modelling errors (Carrassi et al., 2018), using estimated statistical characteristics of the studied variables. Most of the widespread methods (optimal interpolation, 3DVar, 4DVar, various options of the Kalman filter, including their ensemble formulations) use covariance as the basic statistical characteristic. Recent overviews on different DA applications in the Baltic Sea can be found in the papers by Liu and Fu (2018), Zujev and Elken (2018), Goodliff et al. (2019) and She et al. (2020). Whereas there are several results from Baltic Sea reanalysis studies available (Axell and Liu, 2016; Liu et al., 2017), the operational Baltic Sea forecasts within CMEMS (Copernicus Marine En-

vironment Monitoring Service) do not presently include DA (Huess, 2020) and there is ongoing work to implement an automated DA system which would be robust, reliable and well validated.

Results of DA-based forecasting depend heavily on the spatio-temporal configuration of the observing system (Le-Traon et al., 2019). Unlike the regular weather observing networks, observation systems in marginal seas are rather fragmented, where areas and periods of dense sampling can be neighbored by large observation gaps. Therefore, special OSE (observing system experiment) studies have been initiated, to find optimal observation network configurations to achieve best skill of DA (Fuji et al., 2019). However, most of the observations of the Baltic Sea surface variables, not yet detectable by remote sensing (like salinity, nutrients etc.), stem from the FerryBox systems installed on board regularly cruising commercial passenger or cargo ships (She, 2018), and planning can be done only within the existing routes. Therefore, development of improved gap-filling techniques is a challenge and it would be highly beneficial for a region with sparse observations.

Recently, a novel method for EOF reconstruction of gridded sea surface temperature (SST) and salinity (SSS) fields, using the data from (mostly) irregular and (often) sparse observations, was presented and thoroughly tested in the NE Baltic Sea (Elken et al., 2019). The method relies on the estimate of covariance matrix from the long-term model data, which is decomposed into the full set of EOF modes. The mode values at observation points, together with the observed values, enable least-squares estimation of observational amplitudes. The method is able to follow on the regular grid the pointwise observed temporal changes of the mean state and of the major basin-scale gradients. The aim of the present study is to implement this statistical reconstruction technique into the data assimilation of the forecast model, and to study the feasibility of such an assimilation method.

The paper is organized as follows. In the section on data and methods, an overview of the sub-regional oceanographic background and a short model description are presented. Observational in situ data have been compiled from three sources, and they contain shipborne monitoring and Ferry-Box platforms. The reconstruction method is presented in detail, and the section ends with the description of the data assimilation method used. The results section starts with the presentation of experiments in order to find the optimized parameters for reconstruction of gridded fields. The rest of the section is devoted to the analysis of the results of data assimilation experiments, ending with the performance evaluation. Finally, discussion and conclusions are presented.

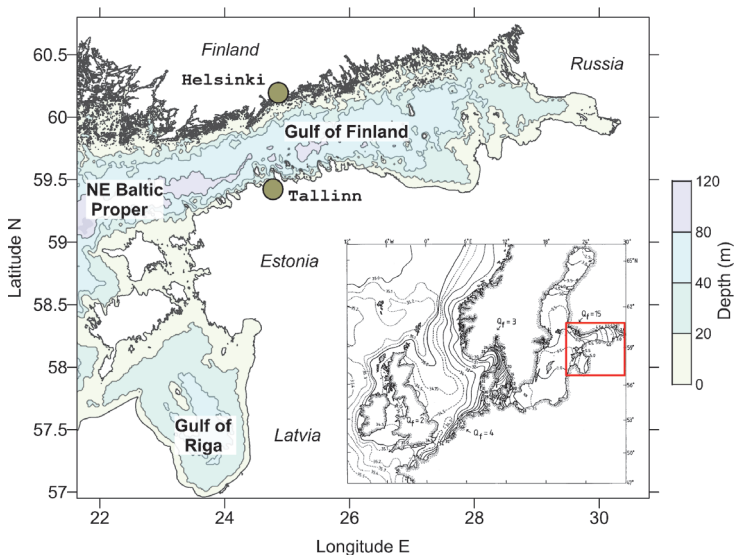
## 2 Data and methods

### 2.1 Study area and the circulation model

We have chosen the study area in the NE Baltic within 56.9417–60.725° N, 21.55–30.35° E (Fig. 1), motivated by several Estonian national interests within the operational forecast of sea state and assessments of the marine environment. The region covers the Gulf of Finland, the Gulf of Riga and part of the Baltic Proper adjacent to these gulfs. The region is rather shallow: the mean and maximum depths are 26 and 62 m in the Gulf of Riga (Yurkovskis et al., 1993) and 37 and 123 m in the Gulf of Finland (Alenius et al., 1998), respectively.

The region lies in the temperate climatic zone. During the summer, SST exceeds usually 15 °C in July or August (Alenius et al., 1998), with highest values up to 25 °C observed in some years in the shallow coastal zones (Stramska and Białogrodzka, 2015). The warm upper layer of 10–20 m thickness is well mixed down to the thermocline or bottom, whichever of them is shallower. Occasionally, wind-driven coastal upwelling processes disrupt this warm layer (Uiboupin and Laanemets, 2009). Nearly every winter, sea ice forms with variable extent and thickness; during severe winters, the Gulf of Finland and the Gulf of Riga are fully ice-covered (Jevrejeva et al., 2004). The region is impacted by large rivers: the Gulf of Finland and the Gulf of Riga together receive 34 % of the total freshwater discharge to the Baltic Sea as can be calculated from the data by Johansson (2017). As a result, there is an estuarine increase in SSS from east to west (Alenius et al., 1998; Yurkovskis et al., 1993), reaching 7–8 g kg<sup>-1</sup> in the Baltic Proper (Kõuts and Omstedt, 1993). The Gulf of Finland has a free connection to the Baltic Proper without a sill or any other topographic restriction; therefore deeper more saline waters of the Baltic Proper penetrate into the Gulf of Finland and form an estuarine halocline (Liblik et al., 2013). A shallow sill with a depth of 15 m connects the Gulf of Riga with the Baltic Proper; therefore deep layers of the Gulf of Riga can receive only surface waters of the Baltic Proper (Lilover et al., 1998). The two gulfs, located in the NE Baltic, play an essential role in the dynamics of the whole Baltic Sea (Omstedt and Axell, 2003).

For the modelling, an Estonian sub-regional set-up (Fig. 1) of the Baltic-wide HBM model was applied with a resolution of 0.5 × 1 arcmin by N and E (ca. 0.5 nautical miles) containing the entire Gulf of Finland, Gulf of Riga and NE portion of Baltic Proper (Lagemaa, 2012; Zujev and Elken, 2018). The model fields are three-dimensional having 455 × 529 × 39 grid cells (by latitude, longitude and depth correspondingly), with 750 088 wet points and 71 986 of them on the surface with a layer thickness of 3 m. At the western open boundary, the data were taken from the Baltic-wide HBM model (Huess, 2020), operated by the Copernicus Marine Environment Monitoring Service (CMEMS, <https://marine.copernicus.eu/>, last access: 2 May



**Figure 1.** Map of the study area in the NE Baltic with depth contours. Shown are the sea areas of Gulf of Finland, Gulf of Riga and part of the NE Baltic Proper. Insert presents the map of surface salinity of the Baltic and North seas by Rohde (1998). Location of our study area is shown in the insert by a red box.

2020). Atmospheric forcing was provided by the Estonian implementation of HIRLAM (Männik and Merilain, 2007). HBM uses the Arakawa C-grid and produces a forecast for 16 ocean variables including temperature, salinity, current speed and ice concentration. A detailed description of the HBM model and its validation can be found by Berg and Poulsen (2012); further analysis and evaluations are given by Golbeck et al. (2015), Hernandez et al. (2015), Tuomi et al. (2018), Huess (2020) and She et al. (2020). In particular, the CMEMS Quality Information Document (Golbeck et al., 2018) concludes that temperature forecast between the surface and about 100 m depth is one of the major strengths of the CMEMS-V4 product, below which the halocline deviations of forecast from observations increase. Regarding salinity, the values are slightly underestimated and the underestimation increases with depth.

The model set-up has been designed for operational forecasting. For computational reasons, it was decided to keep the operational 0.5 nautical mile grid resolution and to perform shorter feasibility experiments, instead of choosing larger grid steps and making longer experiments. The model is used routinely by the Estonian Weather Service (implemented by one of the authors, Priidik Lagemaa); SST is displayed on the web page <https://ilmateenistus.ee/meri/mereprognoosid/merevee-temperatuur/> (last access: 8 May 2020) and SSS is shown on the page <https://ilmateenistus.ee/meri/mereprognoosid/soolsus/> (last access: 8 May 2020). In compliance and for comparability rea-

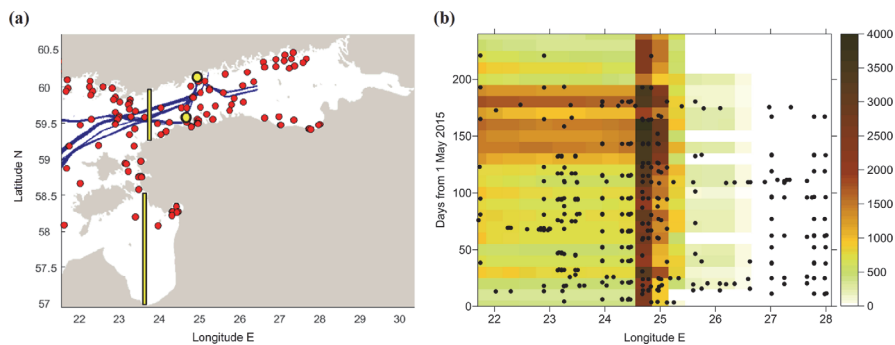
sons with the recent study by Zujev and Elken (2018), we chose the study period from 1 May to 31 December 2015, to be used for the DA experiments. The model experiments were conducted in the framework of operational forecasting, where the forcing files were updated daily. There were no gaps during the study period in meteorological data nor in open-boundary conditions nor any other input.

## 2.2 Observational data

All available SST and SSS data from three sources were compiled:

1. The Copernicus Marine Environment Monitoring Service (CMEMS, <https://marine.copernicus.eu/>, last access: 8 May 2020) contains among other data sources the quality-checked data set of Baltic in situ near-real-time multiparameter observations: [https://resources.marine.copernicus.eu/?option=com\\_csw&view=details&product\\_id=INSITU\\_BAL\\_NRT\\_OBSERVATIONS\\_013\\_032](https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=INSITU_BAL_NRT_OBSERVATIONS_013_032) (last access: 24 October 2019). This data set, accessible through free-of-charge registration, contains in our study region data from several FerryBox systems (automatic observations made from ferries and other ships crossing the sea areas on a regular basis). There are also a number of coastal stations, but they record mainly sea level and water temperature, whereas salinity observations are missing; therefore we are not using coastal stations. In our study





**Figure 2.** Distribution of observations. **(a)** Map of FerryBox observation points along ship tracks (blue) and shipborne monitoring observations (red) over the study period. Shown are also the locations for time–latitude graphs and time series (black contours with yellow background). **(b)** Observation frequency over longitude and time. FerryBox data are shown by colour image; each image cell presents the number of initial observations over intervals of 10 d and 18 arcmin of longitude (ca. 16.7 km). Shipborne observations are shown by black dots.

area and time interval, there were not any operating buoy stations, gliders or Argo floats.

- HELCOM/ICES database contains the results from the HELCOM marine monitoring programme and is hosted by ICES, together with other oceanographic data (<https://ocean.ices.dk/HydChem/HydChem.aspx?>, last access: 22 October 2019). It mainly includes the data from shipborne monitoring stations, where SST and SSS are easily extracted.
- National monitoring database KESE (<https://kese.envir.ee/kese/viewProgramNew.action?uid=473556>, last access: 11 December 2020, search for “mereseire”) contains detailed records of all variables observed under the national environmental monitoring programme. The data that were downloaded on 18 October 2019 contain different data records for every environmental variable. Except for a few cases, these data are also found in the ICES/HELCOM database. Duplicate entries were avoided from the composite data set by averaging over small time and space intervals.

The largest amount of synchronous SST and SSS data originates from the FerryBox systems, accessed through the CMEMS (Table 1). There were about 330 000 initial observation points from FerryBox, distributed over a few ship lanes (Fig. 2a) with a resolution of a few hundred metres and from daily to a few days interval. The analysed water is strongly mixed in the surface layer by the moving ship. Typical observation depth may be considered 5 m, although variations between the ships and due to the variable shipload exist (Lips et al., 2008; Karlson et al., 2016). There were also about 370 observations from shipborne monitoring stations. Distribution of the amounts of observations in selected temporal and longitude intervals (Fig. 2b) reveals a highly irregular

pattern. Most of the observations were concentrated on the Tallinn–Helsinki transect located across the Gulf of Finland between the longitudes 24.6–25° E. FerryBox observations were missing in the Gulf of Riga and in the eastern part of the Gulf of Finland, east from 26.5° E. In the southern part of the Gulf of Riga, available data were missing during the study period.

Two sets of compressed (averaged) FerryBox data were created for further data analysis, containing mean observed values, coordinates and observation times over the selected intervals. Firstly, for the model validation study, daily mean spatial averages over a fine grid with a resolution of  $0.5 \times 1$  arcmin by N and E (as in the used model) cells were created, resulting in about 110 000 values. Secondly, for the EOF pattern analysis and reconstruction of SST and SSS fields, daily mean spatial averages over the coarse grid ( $5 \times 10$  arcmin by N and E, about 5 nautical miles) were created. The main benefit of the coarse grid is to save computational costs while keeping the large-scale patterns well resolved (see Sect. 2.4 for more details on the advantages and disadvantages of using the coarse grid). In this procedure, the initial observations were compressed on the coarse grid by roughly 25 times yielding about 13 000 average values for SST and SSS. Within the temporal averaging, it was chosen not to apply any diurnal cycle correction, and all the observations at different hours were averaged to the closest midnight.

For the interpretation of model and DA results, meteorological data were taken from the model forcing fields. For the occasional comparison, CMEMS remote sensing SST Level 4 data were retrieved from the service portfolio [https://resources.marine.copernicus.eu/?option=com\\_csw&view=details&product\\_id=SST\\_BAL\\_SST\\_L4\\_NRT\\_OBSERVATIONS\\_010\\_007\\_b](https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=SST_BAL_SST_L4_NRT_OBSERVATIONS_010_007_b) (last access: 8 May 2020).

**Table 1.** FerryBox data from 1 May to 31 December 2015 in the NE Baltic used in the present study.

Ship	Main route	Operating institute	Number of initial observations
Baltic Queen	Tallinn–Helsinki	Marine Systems Institute, Tallinn University of Technology	63 368
FinnMaid	Helsinki (Vuosaari)–Travemünde	Finnish Environment Institute	142 235
Silja Serenade	Helsinki–Mariehamn–Stockholm	Finnish Environment Institute	60 228
Victoria	Tallinn–Mariehamn–Stockholm	Estonian Marine Institute, University of Tartu	65 037

### 2.3 Reconstruction of gridded data from point observations

For the purpose of DA, we chose to use EOF reconstruction of large-scale SST and SSS fields, using the orthogonal patterns from models following the detailed outline by Elken et al. (2019), and subsequent relaxation of gridded observations within the model time-stepping. In order to correct the modelled basin-scale patterns towards observations, the spatio-temporal distribution of in situ data was too irregular to use standard interpolation and filtering algorithms like the Cressman method or optimal interpolation with approximated covariance (see an example from the same region by Zujev and Elken, 2018). In this section, we summarize the well-known EOF decomposition and present general features of EOF reconstruction as a problem when the number of observations is less than the number of EOF modes (equal to the number of model grid cells).

The basic option of EOF reconstruction uses at each DA time step time-fixed amplitudes (Elken et al., 2018), encountering the observations spanning over a certain time frame (which can be longer than DA time step) that are transferred to the fixed times by some interpolation or filtering/averaging procedure. The amplitudes are estimated using time-fixed observations by minimizing the root-mean-square-difference between the observations and the EOF reconstruction. The amplitudes at adjacent time moments are not directly related, but in the case of longer temporal filters when observations overlap on different DA time steps, indirect relations between adjacent amplitudes become evident.

Elken et al. (2019) also proposed an advanced method with time-dependent amplitudes. Within this approach, the amplitudes and their time derivatives are estimated together with observations within a selected time interval, in order to find least squares between the observations and EOF reconstruction in the observational framework.

The main steps of EOF reconstruction are as follows. During the standard EOF decomposition, the orthonormal eigenvector matrix  $\mathbf{E}$  (contains the spatial eigenvectors  $e_k$ ) is found from the eigenvalue problem  $\mathbf{BE} = \mathbf{AE}$ , where  $\mathbf{B}$  is  $M \times M$  spatial covariance matrix, calculated from the

$M \times N$  spatio-temporal matrix  $\mathbf{X}$  of the “values of interest” by time averaging, and  $\mathbf{A}$  is a diagonal matrix that contains eigenvalues  $\lambda_k$ . The data set  $\mathbf{X}$  contains time slices  $\mathbf{x}_i$  that are spatial state vectors at time  $i$ . Although in the present study we use the data set  $\mathbf{X}$  selection as 2D sub-sets of individual oceanographic fields, applications towards multivariate analysis and/or extending over the 3D physical domain are straightforward. While  $\mathbf{E}$  is non-dimensional, the dimensional amplitudes (or in other words, factors) of EOF decomposition are found by  $\tilde{\mathbf{a}}_i = \mathbf{E}^T \mathbf{x}_i$ , and the decomposition is reconstructed to the “values of interest” by  $\mathbf{x}_i = \mathbf{E} \tilde{\mathbf{a}}_i$ . Here we have used the notation  $\tilde{\mathbf{a}}_i = \mathbf{A} \mathbf{a}_i$ , where  $\mathbf{a}_i$  is the non-dimensional amplitude. The eigenvalues  $\lambda_k$  present the variance (energy) of the eigenvectors  $e_k$  over the whole period, and the sum of all eigenvalues is equal to  $\sigma^2$ , the variance of  $\mathbf{X}$ . EOF decomposition offers the possibility to keep only the most energetic modes in the reconstruction and truncate the higher modes in  $\mathbf{E}$ . When  $L$  most energetic modes are taken into account in the sorted list of eigenvalues and vectors, the sum from  $\lambda_1$  to  $\lambda_L$  presents the explained variance, and the contribution of truncated modes forms the error variance. If white noise with a variance  $\varepsilon^2$  is present in the decomposed data due to sub-grid-scale processes and/or sampling errors, the noise variance appears only as additive to the diagonal elements of the covariance matrix. The eigenvalue problem becomes  $(\mathbf{B} + \varepsilon^2 \mathbf{I}) \mathbf{E} = \mathbf{A} \mathbf{E}$ , where  $\mathbf{I}$  is a unity matrix. Patterns of spatial modes remain unaffected by adding the white noise, but the eigenvalues and energy share of the modes decrease according to a factor  $(1 + \varepsilon^2/\sigma^2)^{-1}$ . When the sum of eigenvalues of the included dominating modes is less than  $\sigma^2 - \varepsilon^2$ , the contribution of noise is effectively smoothed.

During EOF reconstruction from observations  $\mathbf{y}_i$ , the number of observations  $K$  is assumedly smaller than the number of points  $M$  in the spatial eigenvectors  $e_k$  that are determined on the model grid and evaluated from the model statistics. For the comparison with observations, the model data  $\mathbf{x}_i$  are transformed to the observation points by the observation operator  $\mathbf{H}_i$  by the formula  $\mathbf{H}_i \hat{\mathbf{x}}_i = \mathbf{H}_i \mathbf{E} \hat{\mathbf{a}}_i$ , where  $\hat{\mathbf{a}}_i$  denotes the “observational” amplitudes. Further, the  $\hat{\mathbf{a}}_i$  values should follow least-square minimization of reconstruction error in relation to observations  $\|\mathbf{y}_i - \mathbf{H}_i \mathbf{E} \hat{\mathbf{a}}_i\|^2 \Rightarrow$



min. The expressions to find observational amplitudes and reconstructed fields are

$$\hat{\mathbf{a}}_i = \left( \mathbf{E}^T \mathbf{H}_i^T \mathbf{H}_i \mathbf{E} \right)^{-1} \mathbf{E}^T \mathbf{H}_i^T \mathbf{y}_i, \quad \hat{\mathbf{x}}_i = \mathbf{E} \hat{\mathbf{a}}_i. \quad (1)$$

In the reconstruction by Eq. (1), the critical point is a possibility of spurious amplitudes based on few and unfavourably spaced observation points. Experiments with pseudo-observations (Elken et al., 2019) revealed that the values of  $\hat{\mathbf{a}}_i$  of dominating  $L$  modes should match the limits derived from statistics of  $\tilde{\mathbf{a}}_i$ , whereas higher modes with outlying amplitudes should be neglected.

Most of the oceanographic observations are not made at the same time. It may take several days or even weeks to cover a larger sea area with shipborne monitoring. When  $P$  observations  $\mathbf{y}_p$  are taken at different times  $p$ , then construct an observation operator  $\hat{\mathbf{H}}_p$  that allows pointwise comparison of  $\mathbf{y}_p$  and  $\hat{\mathbf{H}}_p \mathbf{x}_i$  converted from gridded values at specified time  $i$ . Assume that within the short time span the amplitudes depend linearly on time and introduce  $\hat{\mathbf{b}}_p = \hat{\mathbf{a}}_i + \mathbf{d}_i \cdot \delta t_p$ , where  $\hat{\mathbf{a}}_i$  is the time-fixed amplitude,  $\mathbf{d}_i$  is the rate of change vector, and  $\delta t_p = t_p - t_i$  is the difference between the observation and reference times. The function to be minimized regarding reconstruction errors is  $Q = \left\| \mathbf{y}_p - \hat{\mathbf{H}}_p \mathbf{E} \hat{\mathbf{b}}_p \right\|^2 = \left\| \mathbf{y}_p - \hat{\mathbf{H}}_p \mathbf{E} (\hat{\mathbf{a}}_i + \mathbf{d}_i \cdot \delta t_p) \right\|^2$ , which for fixed time  $i$  yields a system of  $2L$  linear equations obtained from  $\partial Q / \partial \hat{\mathbf{a}}_i = 0$ ,  $\partial Q / \partial \mathbf{d}_i = 0$ ,  $i = 1 \dots L$ :

$$\mathbf{G} \mathbf{z} = \mathbf{w}, \quad \mathbf{G}_{mn} = \sum_{p=1}^P f_m^p f_n^p, \quad w_n = \sum_{p=1}^P y_p f_n^p. \quad (2)$$

Here the vector of unknowns combines the amplitudes and their rates of change  $\mathbf{z} = \{ \hat{a}_1 \dots \hat{a}_L, d_1 \dots d_L \}$ . Instead of the full set of EOF mode values, as would be used during standard decomposition, we take the modified/interpolated mode values at observation points; then  $f_n^p = \{ \hat{e}_1^p \dots \hat{e}_L^p, \hat{e}_1^p \delta t_p \dots \hat{e}_L^p \delta t_p \}$ . We note that when all observations have the same time stamp and  $\delta t_p = 0$ , Eq. (2) is reduced to Eq. (1).

Time-dependent reconstruction allows the reference time and length of time interval to be selected. As with the time-fixed reconstruction, the highest “usable” mode is determined by checking the amplitude values with statistical limits. The method also allows estimation of amplitudes and reconstruction only by backward observational data. This feature makes the method useful in operational forecasts, where only past observations can be taken into account for drawing the present nowcast maps.

## 2.4 Method for data assimilation

Many DA techniques use (irregular) point observations of a variable  $\psi$  as the input source. In our approach, gridded maps  $\psi^o$  are used; they are optimized by EOF reconstruction as described in Sect. 2.3. Therefore, in the continuous equivalent, DA is performed by Newtonian relaxation (e.g. Holland and Malanotte-Rizzoli, 1989):

$$\partial \psi / \partial t = F(\psi) - \frac{1}{\tau} (\psi - \psi^o), \quad (3)$$

a discrete form of which has been applied for DA, for example, using gridded climate data (Moore and Reason, 1993) or using optimally interpolated daily satellite-based SST data (Ravichandran et al., 2013). Equation (3) is then written for DA time step  $\Delta t$  in two stages as

$$\psi^f = \psi^{a-1} + \Delta t F(\psi^{a-1}), \quad \psi^a = (1 - \alpha) \psi^f + \alpha \psi^o, \quad (4)$$

where  $\psi^f$  is the raw forecast field calculated from the previous analysis field  $\psi^{a-1}$  using only the model operator  $F$  without DA during this time step, and  $\psi^a$  is the new analysis field. Equation (3) contains adjustable relaxation time  $\tau$  that is transformed in Eq. (4) to non-dimensional  $\alpha = \Delta t / \tau$ . This is the main DA calibration parameter, since extensive use of covariance statistics, including the effects of observation errors, has been included in the estimation of gridded reconstruction of point observations. Newtonian relaxation of gridded observations, applied during the model run at DA time steps is also named “analysis nudging” (e.g. Stauffer and Seaman, 1990), which has had recent meteorological applications (Bullock et al., 2018).

In practical calculations, SST and SSS observational data were reconstructed on the coarser grid with a resolution of  $5 \times 10$  arcmin by N and E (ca. 5 nautical miles) and interpolated or extrapolated by bilinear procedure to the finer model grid with a resolution of  $0.5 \times 1$  arcmin by N and E (ca. 0.5 nautical miles). Such a simple transition of data from a coarse to a finer grid includes smoothing, since  $\psi^o$  lacks the details that are present on the finer grid. We have tested that the effect of added smoothing is smaller than the physical diffusion. In our study area, generation of meso- and small-scale features is of high intensity; therefore relaxation to the smooth observation fields does not apparently damp the fine-grid variability. The approach of using two grids with different resolutions is justified by irregular distribution of observations; reliable estimation is possible only for large-scale patterns of SST and SSS fields. The computationally more efficient coarser grid resolves these patterns with enough details.

The above DA method is computationally efficient. The EOF modes are calculated prior to DA cycles. For each DA time step, only one system of linear equations of rank of the number of EOF modes (about 3–6) has to be solved for the

entire grid. The coefficients of the matrix are found by summation of the products of EOF mode values over the observation points (Eq. 2). For comparison, optimal interpolation requires solving the system of linear equations of rank of the number of observation points (about 100) for each grid cell (about 1000), with a single inverse matrix calculated for the time step.

The model performance with respect to observations was evaluated over those grid cells – time span pairs when observations were available. Since observations covered only a small part of the study domain, DA results were also compared with the control run without DA, but then it is possible to only analyse the changes due to DA, without evidence of possible improvement. Standard statistical characteristics were calculated for individual fields such as mean and standard deviation, and in the case of differences (for example, relative to observations) bias, RMSD (centred root-mean-square difference that equals to the standard deviation of difference field) and the Pearson correlation.

### 3 Results

#### 3.1 Experiments on EOF reconstruction

##### 3.1.1 Covariance, modes and reconstruction tests

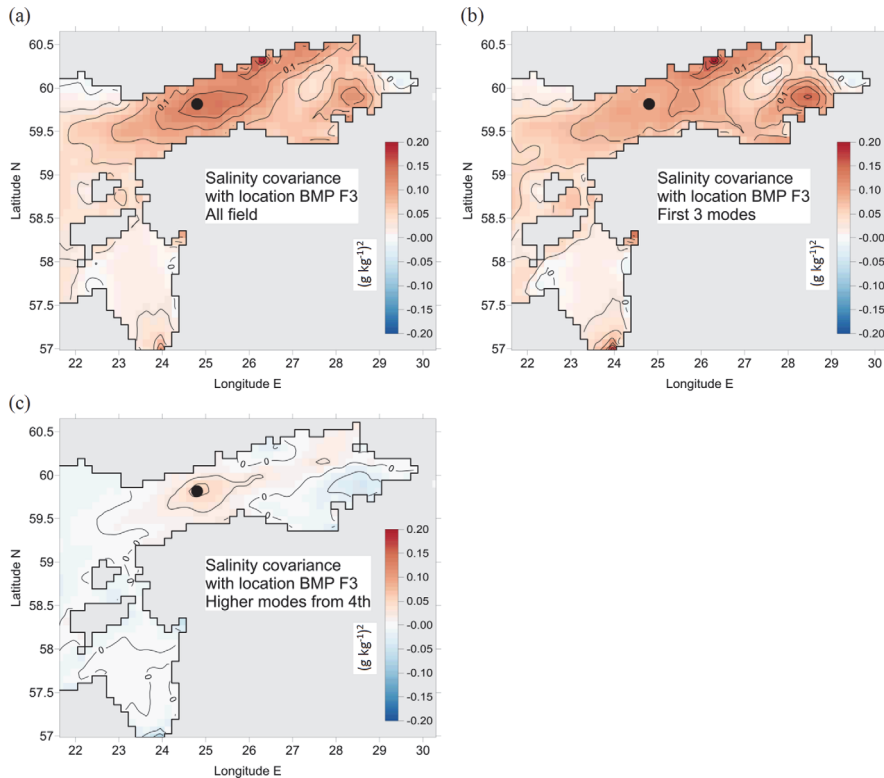
The EOF modes were calculated on the coarse grid ( $5 \times 10$  arcmin by N and E) on the basis of space-averaged results from the fine-grid ( $0.5 \times 1$  arcmin by N and E) model, running from 1 July 2010 to 30 June 2015 (Elken et al., 2019). This analysis revealed that mean distributions of modelled SST and SSS, serving as the basis for calculation of deviations in the variability studies, were close to the climatological maps calculated on the basis of observations (Janssen et al., 1999). The highest temporal variability was found in the shallow coastal areas for SST, whereas the largest SSS variations were revealed near the larger river mouths and in the NE area of the Gulf of Finland. While temporal changes strongly dominate in the variability of SST, spatial changes prevail in SSS variability.

Calculated SST and SSS covariance matrices have significant spreading of individual values over pairs of points, especially for the dominating gravest modes where big covariance values may occur over large distances. Covariance of residual fields (sum of higher EOF modes) has a decay scale of about 30 km with increasing space lag, both for SST and SSS. The first, most energetic EOF modes have nearly “flat” patterns without sign change (energy share 97.6 % for SST and 36.2 % for SSS); their amplitudes are dominated by a seasonal signal. A space-dependent mean biharmonic seasonal cycle was not removed from the model time series prior to the analysis, since special experiments revealed only a small effect of seasonality suppression on EOF mode patterns. The second EOF mode of SST (1.3 %) presents dif-

ferential heating and cooling in shallow areas, compared to the deeper offshore waters. Transverse anomaly stripes near northern or southern coasts, like those due to coherent upwelling and downwelling in the region, were evident in the second SSS mode pattern (16.9 %) and third SST mode pattern (0.31 %). There is also a pattern of SSS changes in the freshwater spreading pathway near the northern coast of the Gulf of Finland (third SSS mode, 7.1 %) and longitudinal SST changes in the east–west direction (fourth SST mode, 0.14 %).

The data set used in the present DA study (Fig. 2) is rather irregular, compared to the reconstruction experiments by Elken et al. (2019). Therefore, we revisit the covariance issues and perform additional reconstruction tests, before finding in the next subsection the best options for the automatic reconstruction procedure. Spatial interrelation of observed values at a specific point to the values in the rest of the region is found from the extract of the spatial covariance matrix, which can be shown as a map. One example of SSS covariance with a frequently sampled HELCOM monitoring station BMP F3 is shown in Fig. 3. The covariance of three dominating EOF modes (Fig. 3b) comprises most of the unfiltered data covariance (Fig. 3a) at large distances. High covariance locations have clear basin-scale geographical explanations: under the similar weather and seasonal forcing, which is spatially nearly uniform, SSS changes in distant river influence areas are closely interlinked. Correlation (not shown) may exceed 0.4 at distances greater than 500 km; therefore, assumptions of fast decay of correlation with space lag (like using the Gaussian covariance approximation), adopted in offshore areas with negligible coastal influence, are not valid. Covariance of residuals to the large-scale variations are presented by higher EOF modes (Fig. 3c). Such smaller-scale variations have nearly Gaussian structure, with elliptical anisotropy stretched along the axis of the basins similar to the results by Høyer and She (2007): spatial scales in Fig. 3c are 30 and 15 km along the main axis and perpendicular to the axis, respectively. Similar regularities – physically explained high covariance at large distances, localized covariance patterns for the higher EOF modes – were found for other points of reference, both for SSS and SST fields.

The EOF reconstruction method relies on the full covariance matrix, without any approximation. Covariance is further treated using EOF modes. For the reconstruction procedure, we keep the lowest EOF modes without any approximation, and covariance from higher modes as shown in Fig. 3c is not taken into account. The large-scale features of the EOF reconstruction and associated DA exclude the possibility of creating spurious “bullseye” patterns around observation points, which may happen for instance during unfavourable selection of optimal interpolation parameters. Subsequently, our DA method handles the large-scale features and excludes the possibility to assimilate smaller-scale features, which can be described by the higher modes.



**Figure 3.** Spatial covariance of SSS with the values in the grid cell near the HELCOM monitoring station BMP F3 (59.8383° N, 24.8383° E), extracted from the full covariance matrix calculated from the model data over 5 years. Covariance is decomposed by EOF modes: covariance of unfiltered data with all the modes included (a), the sum of covariance of the first three modes (b) and of the remaining higher modes, starting from the fourth mode (c).

A full covariance matrix can be implemented in optimal interpolation as well. While the EOF method needs to limit the number of included modes, smoothing in such a way smaller-scale variability and observational errors, optimal interpolation needs to include observational error variance (“nugget effect” in terms of Kriging method, equivalent to optimal interpolation); otherwise the system of underlying linear equations may become close to singular and the result may become unrealistically spiky. In some examples (not shown), EOF reconstruction and optimal interpolation based on full covariance produced similar results, but these relations need further studies. When observed values were close to the model-computed climatological background, visual similarity was caused mainly by the dominance of spatial gradients of mean SSS over the spatio-temporal variability. Optimal interpolation with Gaussian approximation to the covariance produced realistic results in the neighbourhood of observation points but gave unrealistic patterns and values in the distant SW extrapolation area.

### 3.1.2 Finding the parameters for reconstruction of gridded observation fields

Multiple checks performed on our data set suggested that only the three leading modes were included in the EOF reconstruction. In order to find the best options for reconstruction, experiments were made with different intervals (time window)  $t_R$  around the reference time  $t_i$ ; including the observations within time window from  $t_i - t_R/2$  to  $t_i + t_R/2$ . The results were evaluated to fulfil the following goals:

- a small RMSD between the observed values and the reconstructed fields;
- a small number of gaps in the reconstructed time series;
- a low number or missing presence of “spikes” and/or “jumps” in the time series.

Two basic options for temporal handling of the reconstruction procedures were tested:

- a. application of procedure by Eq. (1) of time-fixed amplitudes; time average of observations was taken for each grid cell, time adopted in each grid cell as constant reference time;
- b. full application of the procedure by Eq. (2) of time-dependent amplitudes; all the daily mean observations (average was taken also over coordinates and time) were kept separate for each coarse grid cell where the observations existed.

In addition, the procedure by Eq. (2) was tested with an option with a time average of observations in each grid cell, and with selection of observations closest to the reference time. These experiments provided more spikes and 70 % higher RMSD than the basic options (a) and (b) and they were neglected from further consideration.

As a first step in all the experiments with variable time windows, the EOF amplitudes of the mode  $k$  were checked for the limit  $|\hat{a}_{i,k}| < 2\sqrt{\lambda_k} = 2\sigma(\bar{a}_k)$ , where  $\sigma$  denotes standard deviation. DA data for the days with higher amplitudes were left blank since these reconstruction results most frequently became unrealistic. In addition, when the number of observations was less than six, reconstruction was not performed and the DA step using Eq. (4) was skipped.

The time windows  $t_R$  for experiments (a) and (b) were selected to be 10, 20 and 30 d. Elken et al. (2019) have found that the correlation timescales ( $e$ -folding drop, correlation value 0.368) of EOF SST amplitudes were 65 d for the seasonal first (overall heating/cooling) and second (faster heating/cooling in shallow coastal areas) modes, and 15 d for the third “upwelling” mode. Timescales of the SSS modes were 65 d for the second and third modes, representing the large-scale gradients, and 110 d for the first mode describing long-term variations of mean salinity.

Methods of time-fixed (a) and time-dependent (b) reconstructions revealed similar statistical results during the study period in 2015, whereas RMSD between observed and reconstructed values of (a) was by 5 % larger than that of (b). By increasing the time window, RMSD of reconstruction slightly increases due to the stronger smoothing. The smoothing effect can be seen from the reconstruction examples given in Fig. 4. It should be noted that the reconstruction is designed to yield the best approximation to the observations over the entire region; therefore, it does not need to present the local best fit at individual points.

A network of observations, available during the study period, appeared favourable for the reconstruction, although observations were missing in the southern part of the Gulf of Riga and eastern part of the Gulf of Finland. With a time window of 30 d, there were no reconstruction gaps identified during the study period, determined for both of the methods by the above-described amplitude limit criteria. Smaller time windows yielded some gaps in 2015. During the longer period from 2010–2018, gaps were found in most of the years (except our study period), whereas shorter time windows

result in more reconstruction gaps. Detailed comparison of the time-fixed (a) and time-dependent (b) methods revealed that time-fixed reconstruction might create spurious “jumps” when there is a gap in observations which has a length close to the time window. In that case, a backward average is taken before the gap and forward average after the gap, which may result in “jumpy” results. Time-dependent reconstruction, which also accounts for the temporal changes within the time window, handled such situations more smoothly.

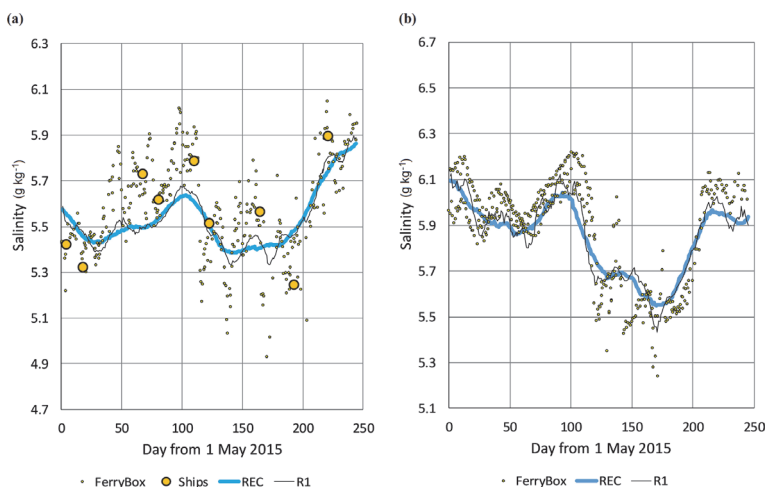
### 3.2 Data assimilation experiments

We have used a two-scale DA approach (see detailed explanation in Sect. 2.4), where observations were reconstructed on the coarse grid. Results were interpolated into the fine grid of the model and were subsequently used for relaxing the fine-scale model results towards basin-scale observational patterns. More specifically, gridded observational SST and SSS data were pre-calculated each day using the time-dependent EOF reconstruction method with a time window  $t_R = 30$  d as presented in Sect. 3.1. Reconstructed SST and SSS fields were interpolated bilinearly to the fine 0.5 nautical mile grid and used for relaxing the model results towards observational counterparts, based on Eqs. (3)–(4) with  $\Delta t = 1$  d. Two basic experiments were conducted, with relaxation time 10 d (weight of observations 0.1, experiment code DA01) and with a relaxation time of 5 d (weight 0.2, experiment code DA02). In addition, a variety of short-term trials was performed in a preparatory phase (results not graphically presented) which led to the two basic experiments. Comparison data were coded as FR for the control run without DA, and FB for observed FerryBox data.

#### 3.2.1 Example from the beginning of August

There was an interesting oceanographic situation in the beginning of August, when a moderate but extensive upwelling SST pattern at the northern coasts of the basins (Fig. 5), with some effects on SSS (Fig. 6), was combined with fast heating of the thin (6–9 m) surface layer (Fig. 7). Since the middle of July, moderate winds with speeds from 4 to 6 m s<sup>-1</sup>, which had a westerly zonal component (favouring upwelling at the northern coasts of the basins), were blowing above the Gulf of Finland. After 3 August 2015 (the maps in Figs. 5 and 6 are taken on this date), wind ceased and air temperatures increased by 10 August across the study area up to 25–27 °C in the Gulf of Finland and up to 31 °C in the southern Gulf of Riga, creating a thin layer of warm surface water. Heating of surface waters was favoured by high night-time air temperatures, higher than SST. Vertical profiles (not shown) in the Gulf of Finland revealed a deep thermocline at 40 m depth near the southern (downwelling) coast and a shallower thermocline near the northern coast; the warm surface water column near Tallinn was 2 to 3 times thicker than near Helsinki. From the end of July to 10 August, warming resulted in an





**Figure 4.** Salinity time series at locations (a) 59.8383° N, 24.8383° E (HELCOM station F3) and (b) 59.794° N, 24.822° E, during the study period. Shown by dots are the observations from FerryBox and from ship monitoring. Reconstructed time series, made using the time-dependent method, are given by solid lines: REC – basic option with 30 d interval, all observations in window were kept as they are; R1 – the same as previous but with time interval 10 d.

increase in SST (Fig. 7) near Tallinn from 16.5 to 18.5 °C and near Helsinki from 14.5 to 18 °C.

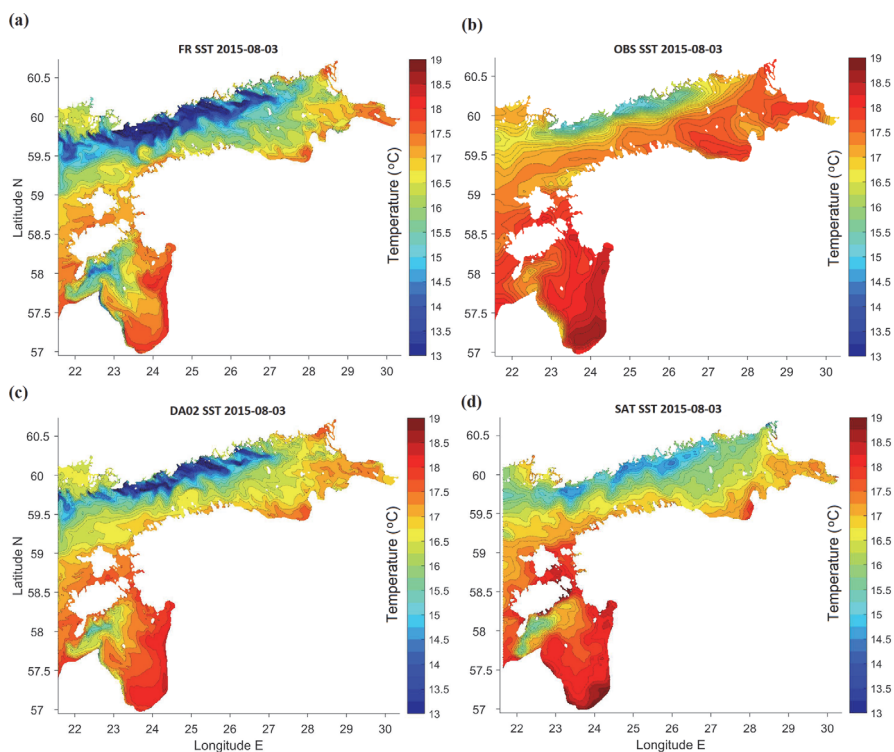
The SST maps presented in Fig. 5 include control run, reconstructed in situ observations, one experiment with DA (the other experiment yielded similar results) and satellite observations. When warm waters with SST above 17 °C dominated the study area, all the maps revealed moderate upwelling near the northern coasts of the basins. However, the minimum temperatures and the spatial extent of the colder waters were different. The warmest “cold” waters were observed on satellite images. While satellites measure SST of a thin surface layer, FerryBox and models acquire temperature over a much thicker layer. It is known that in the Gulf of Finland satellite and FerryBox can have similar SST values in cases of winds stronger than  $5 \text{ m s}^{-1}$  (Uiboupin and Laanemets, 2015); at smaller wind speeds the SST bias can be 1–3 °C in reference to FerryBox observations. Within these accuracy limitations, satellite observations presented in Fig. 5d confirm the model patterns to some extent. The control run (Fig. 5a) was characterized by SST contrasts that are too high, compared to the satellite data (Fig. 5d; for the data source see Sect. 2.2). From the earlier study by Zujev and Elken (2018), it is known that the free model without DA forecasts faster heating and cooling of shallow coastal areas and slower heat dynamics in offshore areas. Data assimilation (Fig. 5c), made using the reconstructed FerryBox data (Fig. 5b), reduced discrepancies with satellite observation. The major large-scale differences between the satellite data (Fig. 5d) and the best DA02 (Fig. 5c) can be outlined as follows: (1) the colder upwelling water extended on the

satellite image further to the east, (2) warmer waters were found on the satellite images in the southern Gulf of Riga, near the Daugava river and in the shallow areas between the Estonian islands, and (3) in the Gulf of Riga, a strip of colder waters was modelled along the western coast, while satellite observations revealed warmer waters near this coast.

There were also numerous mesoscale features evident on SST (Fig. 5) and SSS (Fig. 6) maps, like colder upwelling filaments along the northern coasts of the Gulf of Finland and the Gulf of Riga, and decaying anticyclonic warm-core eddies near the southern coast of the Gulf of Finland. The Irbe Front (Lilover et al., 1998; Raudsepp and Elken, 1999), formed by the salinity difference between the Gulf of Riga and the Baltic Proper, was found by the SSS maps in the outward position, stretching from the strait towards the open sea. This salinity structure was also repeated in the SST patterns; the satellite observations confirmed the predicted outward position during the taken snapshot. The model predicted that in the Gulf of Riga the Daugava river waters were spreading by narrow coastal strips of lower salinity in both the NE and NW directions (Fig. 6).

### 3.2.2 Time series in the areas of dense observations

Locations with dense observations allow us to validate the model and visually evaluate assimilation quality. We compared SST and SSS data of the control run (FR) and DA options DA01 and DA02 with FerryBox data (FB) at two points near Tallinn and Helsinki (Fig. 7). While SST followed the seasonal cycle, with weather-dependent deviations, then SSS



**Figure 5.** Maps (longitude E, latitude N) of SST in the study area on 3 August 2015: (a) free model run without DA, (b) in situ observations reconstructed using EOF method, (c) DA with relaxation time 5 d (weight of observations 0.2), (d) CMEMS product based on satellite observations.

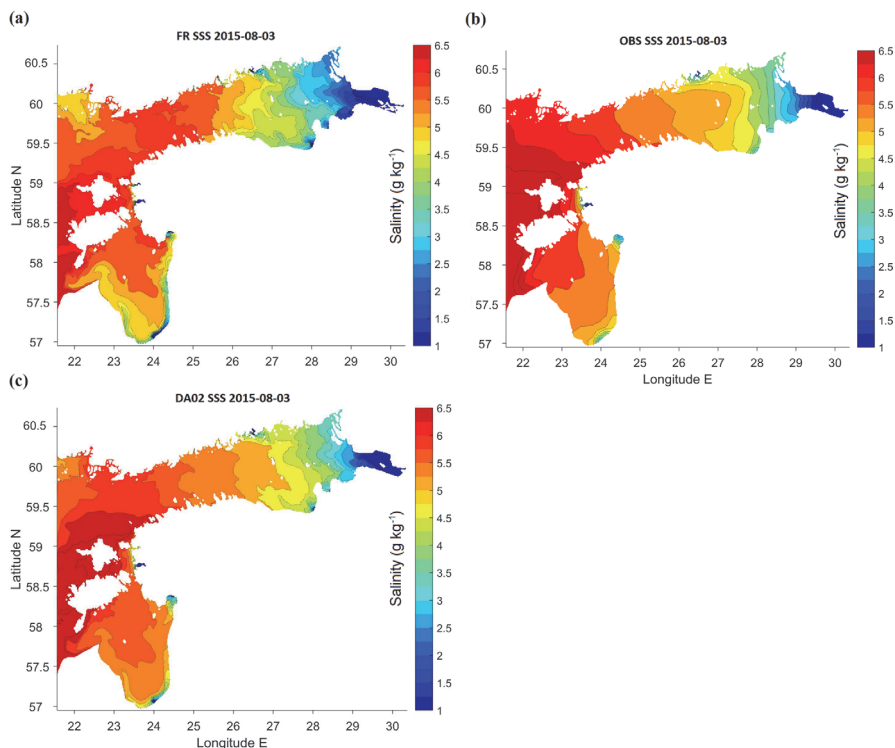
behaviour was more irregular. In the given variation scales of SST and SSS ( $16^{\circ}\text{C}$  and  $2\text{ g kg}^{-1}$  respectively), all the compared SST data sources were more similar to each other than that of SSS. Still, most of the time the assimilation curve (DA02) was closer to the FerryBox observations than the control run, for both SST and SSS.

Warm conditions in the beginning of August (Sect. 3.2.1) are clearly visible on SST time series (Fig. 7a, c). Comparing the values near Tallinn and Helsinki, the southern part of the Gulf of Finland was roughly  $2^{\circ}\text{C}$  warmer than the northern part, whereas the northern part had an unstable day-to-day pattern, possibly due to the fluctuations of the upwelling pattern. This is consistent with the spatial maps given in Fig. 5. Near the southern coast, an upwelling event occurred in September, reducing SST during a few days by nearly  $4^{\circ}\text{C}$  (Fig. 7a). A larger SST drop during the southern coast upwelling (at easterly winds), compared to the northern coast upwelling (at westerly winds of the same magnitude), is explained by the steeper topography slopes in the southern part of the Gulf of Finland (Laanemets et al., 2009). This upwelling event was properly resolved by all the data sets, with

DA02 being closest to observations. In general, a free model without DA expected warming at a lower rate during summer and was more precise in autumn, while both assimilation experiments properly corrected the SST and SSS values. However, in some cases, assimilated temperature was somewhat higher than observed and modelled SST.

Assimilation resulted in one major SSS improvement in early summer when the model predicted upwelling with salinity near Helsinki that is too high. Nevertheless, in some cases DA made minor corrections at one of the locations, ignoring observations and sticking to the control run (e.g. late July to early August near Tallinn, and the middle of October near Helsinki). When the model overshoots at both locations, DA properly corrects temperature and salinity values. This implies that DA of surface observations tends to correct the mean values better than the cross-gulf gradients, for which 3D circulation (presently not assimilated) has a significant impact.

In the salinity time series, a “freshwater event” with reduced salinity was observed in the Gulf of Finland at the end of September and beginning of October. In the daily



**Figure 6.** Maps (longitude E, latitude N) of SSS in the study area on 3 August 2015: (a) free model run without DA, (b) observations reconstructed using EOF method, (c) DA with relaxation time 5 d (weight of observations 0.2).

SSS data (Fig. 7b, d) the event was spiky, possibly due to the mesoscale features not assimilated in the present study: without DA, the eddies tend to have a random phase, and the spikes in the time series of different model options and observations do not need to be coherent. However, in the weekly averaged data (not shown) the mesoscale activity was suppressed and the fresh event appeared simultaneously in all the data within the central and western part of the Gulf of Finland.

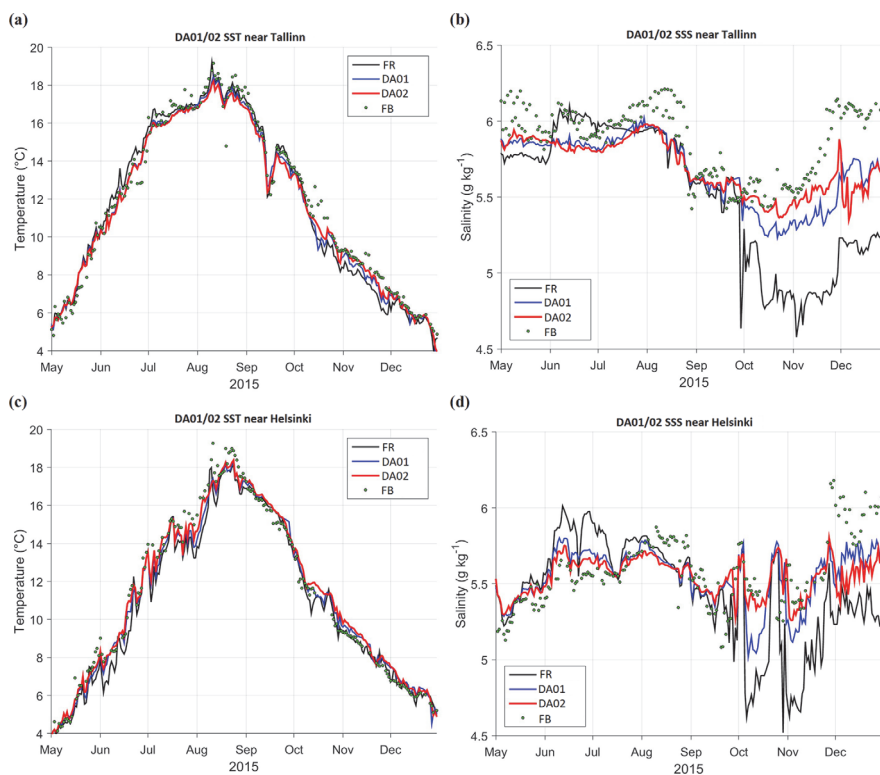
Increasing assimilation weight in Eq. (4) two times did not make assimilation results two times closer to the observations. As can be seen from Fig. 7, the results of assimilation experiments DA01 and DA02, with relaxation times of 10 and 5 d respectively, were not placed between the free run and the observations proportionally to the corresponding weights 0.1 and 0.2. They diverged as the study region experienced a temperature drop or daily trend change. Both options of assimilated SST could either coincide for a long time or go in parallel, but DA02 was systematically closer to the FerryBox observations. Salinity fluctuations had larger amplitudes in the free run without assimilation, but both DA options, with a “thumb” rule – the bigger the weight, the bigger

the change – had properly corrected them. Still, in December DA01 showed better results, being closer to the FerryBox salinity than assimilation DA02.

### 3.2.3 Spatio-temporal dynamics

We have chosen to compare assimilation with the best results (DA02) to the control run without data assimilation (FR) and track the continuous time–latitude changes of SST and SSS (Fig. 8) in two sub-basins – the Gulf of Finland and Gulf of Riga – along the coast-to-coast transects given in Fig. 2a. Using DA, temperature was corrected approximately by 1–2 °C, and salinity by less than 1 g kg<sup>-1</sup>. Major systematic change (in the Gulf of Finland this was validated as improvement; see Sect. 3.2.4 for further details) was seen near the coasts and in the spring and autumn periods, while summer temperatures underwent minor corrections. Salinity corrections had a more uniform distribution and smooth drifting pattern – DA consistently increased SSS values with time in both of the sub-basins.

Data assimilation had increased SST in the Gulf of Finland in open waters during the warming period and in late autumn



**Figure 7.** Time series of SST (a, c) and SSS (b, d) near Tallinn (a, b, 59.4833° N, 24.7667° E) and Helsinki (c, d, 59.9500° N, 24.8833° E), locations shown in Fig. 2a. FerryBox data (FB) are shown by dots, black lines represent control run (FR) without DA, red lines correspond to DA with relaxation time 5 d (weight of observations 0.2, DA02), blue lines for 10 d (weight 0.1, DA01).

all across the gulf and had decreased in the coastal areas during the warming period, whereas near the northern coast this decrease continued until September. In the Gulf of Riga, the SST increase dominated throughout the study period, but it was interrupted occasionally by basin-wide events when DA had decreased the temperature compared to the results from FR. The largest corrections of both SST and SSS were evident in the coastal waters. Salinity was increased by DA in most of the cases in the Gulf of Finland, except for May–July near Tallinn. The largest increase in SSS occurred in November and December, when control run results dropped compared to the earlier period.

Some unusual basin-wide events can be found on the difference charts in Fig. 9. For example, abrupt warming of the surface around 10 August 2015 (Sect. 3.2.1) was correctly predicted by the free run model (Fig. 7c), but it was over-smoothed by the data assimilation. A similar line in December on both charts denotes the occurrence of fronts of cold and saline water due to strong winds and storms.

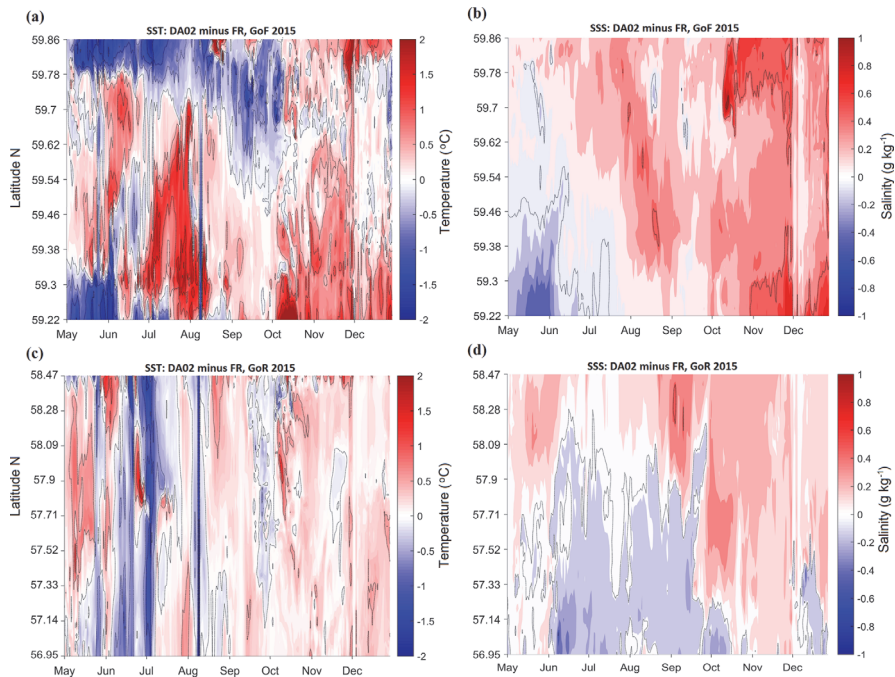
As there are not enough observations available in the Gulf of Riga for validation, we cannot definitely say whether DA improved the situation in the region and to what extent.

### 3.2.4 Evaluation of DA-based forecast performance

Ocean model performance (e.g. Stow et al., 2009; Golbeck et al., 2015; Placke et al., 2018) is usually evaluated by the differences between the observations and the model results, transferred to the times and locations of observations so that they can be directly compared. The overall mean difference (over time and space) is termed bias and the standard deviation of differences at all the observation points is denoted as RMSD (centred root-mean-square difference). The forecast skill is usually non-dimensional, with the RMSD of the studied option (in our case, DA) scaled to reference data (FR in our case) as  $\text{skill} = \text{function of } [\text{RMSD}(\text{DA}, \text{FB}) / \text{RMSD}(\text{FR}, \text{FB})]$ .

The present ocean model has a fine resolution of about 0.5 nautical miles (930 m) (Sect. 2.1); therefore for comparison with observations we used a simplified approach and took av-





**Figure 8.** Time (months of 2015) versus latitude (N) contour graph of DA anomalies of SST (**a**, **c**) and SSS (**b**, **d**) in reference to the control run (FR) without data assimilation at longitudes 23.7166° E (**a**, **b**, Gulf of Finland) and 23.5333° E (**c**, **d**, Gulf of Riga); locations shown in Fig. 2a. DA data are given for relaxation time 5 d (weight of observations 0.2).

erages of observations over the model grid cells over a daily time span (Sect. 2.2). Such a compressed fine-resolution observational data set, still having about 110 000 points for SST and SSS, originated mainly from the FerryBox (FB) lines (Fig. 2), and it covered central and western parts of the Gulf of Finland and the neighbouring part of the Baltic Proper. Areas with lower salinity in the eastern Gulf of Finland and in the Gulf of Riga had only a small number of observations.

Data from the DA experiments DA01 and DA02 were compared to the same compressed observational FB data as the data from the control run without assimilation (FR). Hernandez et al. (2015), who reviewed the problems of performance evaluations of operational ocean models, noted that most available observations are used to adjust models and reduce analysis errors. Therefore, a widespread approach is withholding part of the data set for statistical quantification of errors. In our study, the option of withholding the observations was performed: an evaluation was made of how much the DA result will change if DA is performed using 50 % of the available data (Gregg et al., 2009). The present implementation of EOF DA used about 13 000 observational averages over a coarse grid of about 15 nautical miles. The reconstruction procedure by Eqs. (1)–(2) has no direct connection to the ongoing modelling (although it includes sta-

tistical results from longer model runs), and the fields of  $\psi^0$  in Eqs. (3)–(4) are the only link where observations enter the DA process. The experiments which took every second available observation “box” into account (this resulted in a mean sampling interval along ship tracks about 20 km instead of 10 km) revealed that performing DA during the study period with a reduced data set (6.5 000 averaged observation data instead of 13 000) changed RMSD of SST by only 1 % and of SSS by 2 %, whereas the RMSD values were 0.05 °C for SST and 0.027 g kg<sup>-1</sup> for SSS. An evaluation was made over the full time span and domain using 182 000 coarse grid cells; correlation between the data sets was higher than 0.999. We have also checked reconstruction results with FerryBox data only, excluding the data from shipborne monitoring stations. Compared with the full data set, the largest (but still minor) differences with RMSD of SSS up to 0.03 g kg<sup>-1</sup> were found in the Gulf of Riga and the eastern Gulf of Finland, where FB data were missing. Consequently, for our large-scale approach DA results are robust to the reasonable variation of data amount, and we used FB data for reference in the performance evaluations.

Evaluated forecast performance metrics are presented in Table 2. Only those fine-grid points which had respective value of FerryBox observations on the same day were used

**Table 2.** Statistics of daily data in  $0.5 \times 1$  arcmin (N and E) grid cells with FerryBox (FB) observations: free model run without data assimilation (FR), data assimilation DA01 (observation weight 0.1), DA02 (weight 0.2) and FB. Bias, RMSD and correlation are taken with reference to FB.

	FR	DA01	DA02	FB
<b>SST (<math>^{\circ}\text{C}</math>)</b>				
Mean	12.03	12.15	12.25	12.48
SD	3.98	3.92	3.93	3.97
Bias	-0.45	-0.33	-0.23	0
RMSD	0.72	0.59	0.56	0
Correlation	0.98	0.99	0.99	1.00
<b>SSS (<math>\text{g kg}^{-1}</math>)</b>				
Mean	5.61	5.79	5.85	5.93
SD	0.35	0.29	0.31	0.37
Bias	-0.31	-0.14	-0.08	0
RMSD	0.35	0.24	0.23	0
Correlation	0.52	0.76	0.78	1.00

for metrics calculation. Wet points of the model without corresponding observation value were left out from the procedure.

The statistical properties presented in Table 2 reflect that DA improves the model performance significantly: RMSD of SST was reduced by 22 % and SSS by 34 %, compared to the control run. From DA01 to DA02, a slight improvement of DA performance was observed; therefore we adopted DA02 as the major result. The spatial pattern of RMSD change between the DA and FR (Fig. 9) reveals that larger reduction rates (up to 50 %), for both SST and SSS, were found in the observation-covered areas in the Gulf of Finland. Overly cold waters produced by FR near the northern coast of the Gulf of Finland were effectively corrected by DA (see also Fig. 5); therefore highest improvement percentage scores were detected in this region. Near the western open boundary, non-assimilated SST and SSS values of the larger model were advected into the area, and therefore RMSD reduction was small, or even negative for SSS.

The applied EOF DA method does not assimilate mesoscale variability. Applying the weekly average statistics like Zujev and Elken (2018) further reduced RMSD by 13 % for SST and 9 % for SSS, compared to the daily data in Table 2. Weekly statistics suppresses the mesoscale variability and reveals a better match between the DA and the observations. DA decreased the bias, especially for SSS. At the same time, the correlation of SSS between DA and observations increased considerably. We may conclude that DA made major improvements in the forecasting of SSS. Still, the forecast RMSD in reference to the observations is 62 % of the observed standard deviations, which suggests that there may be further room for improvement. Modelling of SST is already more accurate than SSS without DA: RMSD of the control

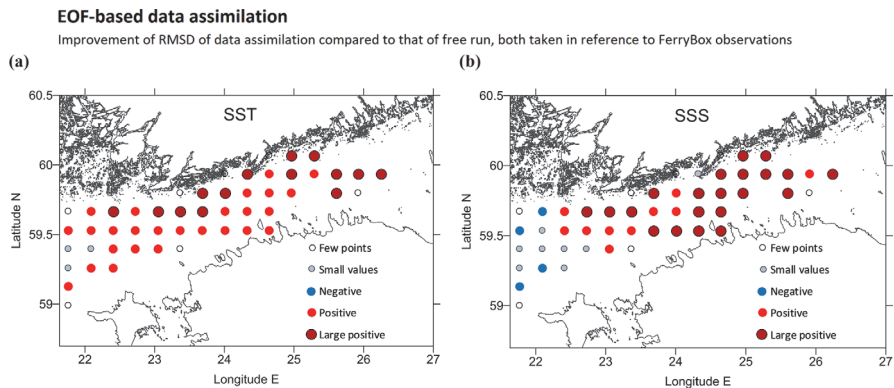
run (FR) makes 18 % of the standard deviation of observations for SST and 94 % for SSS.

#### 4 Discussion

The Baltic Sea is considered as one of the most studied marine areas in the world (e.g. Andersen et al., 2017). However, the large observational data sets are distributed unevenly. If we divide our study area into 744 eddy-averaging grid cells of  $5 \times 10$  arcmin by N and E, then during the study period 330 000 FerryBox observations covered only 18 % of the sea region. Shipborne monitoring added 8 % more coverage of the area, but with a much smaller frequency of sampling. Having in mind that the ocean models tend to deviate in the NE Baltic from the observations not only by constant bias but also for large-scale and longer-term responses, the introduction of non-local, region-wide data assimilation is of high importance.

It is interesting to consider how our statistical evaluations of model and DA performance, given in Table 2, compare with other Baltic Sea studies. For remote sensing versus in situ reference, Kozlov et al. (2014) have found RMSD  $1.31^{\circ}\text{C}$  in the Curonian Lagoon. Uiboupin and Laanemets (2015) have estimated RMSD of various satellite products to FerryBox in the Gulf of Finland from  $0.29$  to  $0.98^{\circ}\text{C}$ . Our control run gave RMSD of  $0.72^{\circ}\text{C}$ . Golbeck et al. (2015) compared SST from 13 models with satellite data and found yearly RMSD for SST of  $0.65$ – $0.87^{\circ}\text{C}$  in the Baltic Sea. They found a larger relative spread of SSS ensemble members than of SST: deviations in the Gulf of Finland between the models were up to nearly  $1 \text{ g kg}^{-1}$ , while the average SSS is only about  $4 \text{ g kg}^{-1}$ . Unfortunately, there were not enough validating observations for SSS available. Fu et al. (2011) found even larger RMSD for SST for the control run,  $1.0^{\circ}\text{C}$ , based on satellite observations. They also used DA with ensemble optimal interpolation and found that DA reduced RMSD between the forecasts and observations by 25 % for SST and 34 % for SSS. With our simpler and less computationally demanding EOF DA technique, similar RMSD reductions have been obtained (Sect. 3.2.4) compared to earlier studies.

We have developed and tested an EOF-based relaxation technique where the large-scale observed fields to be assimilated are pre-calculated independently from the ongoing model. From sparse observations, it is possible to estimate the amplitudes of only the gravest, large-scale EOF modes. The EOF DA method handles large-scale features over the sea basin(s), like change of mean SST, SSS and their gradients, including differential heating in coastal and offshore areas, major patterns from upwelling, and spreading of river discharge. The method can work well with irregular data but cannot resolve mesoscale features in the areas of dense observations, because the EOF amplitudes of higher modes get noisy, according to our experiments. Optimal interpolation,



**Figure 9.** Improvement of RMSD of DA compared to that of FR, both taken in reference to 110 000 FerryBox observations. Comparison is made for  $20 \times 20$  grid cells (about 10 nautical miles) for SST (a) and SSS (b) over the whole study period. Legend codes: few points – less than 100 observations in a box, small values – absolute percentage change less than 10 %, negative – DA RMSD growth more than 10 %, positive – DA improvement (RMSD reduction) from 10 % to 30 %, large positive – improvement more than 30 %.

successive corrections and similar methods usually assume localized covariance and/or radius of influence (e.g. Axell and Liu, 2016); they work well in resolving mesoscale features in dense sampling areas, but regions of rare observations remain unaffected by DA. For the mesoscale range, in our study area there are only satellite observations of surface variables available. They were omitted from our study, since salinity as a variable of primary interest can be presently only be determined in situ in the Baltic. It is possible to implement on top of EOF DA more traditional localized DA methods to assimilate mesoscale data when and where such data are available. Studies on using EOF DA for handling large-scale data are also ongoing in the UK Met Office by Daniel Lea (Haines, 2018).

We have tested the EOF-based DA in a centred time window of 30 d, based mainly on available FerryBox data during the study period. As shown by reconstruction experiments by Elken et al. (2019), the time-dependent method can also work with backward observations as if it occurs during operational forecasts. When more observations become available, for example from new automated buoy stations, Argo floats and gliders, the time window can be shortened. A full covariance matrix estimated from the model results is the backbone of the EOF DA method. Prior and/or complementary to implementation of the method into operational practice, detailed covariance studies using results from multiple models could be useful, as well as additional reconstruction and DA studies using more data sources over longer periods.

The EOF DA method has some practical advantages. Firstly, for assimilation of basin-scale patterns, it can be implemented on a coarse grid, and therefore it has small computational effort compared to the localized methods (like optimal interpolation etc.) that should be usually implemented on the model resolution, i.e. on the fine grid. Secondly, inter-

mediate results are in the form of maps that are easily understandable and can be checked visually or taught to be analysed by artificial intelligence. For optimizing the observational data needs, the concept of OSEs (observing system experiments), which check various data configurations for DA performance, is high on the agenda. Since the quality of DA and forecasting are primarily determined by the quality of EOF reconstruction (when extensive mesoscale observations are not available), then it would be possible to save a significant amount of computing power and perform most of the experiments using EOF basis vectors.

There are obvious possible extensions of the EOF DA method to other variables and layers: improvement of stratification modelling; extension to biogeochemical models; and DA of oxygen, nitrogen and phosphorus. Applicability depends on how well the model reproduces the studied fields and their covariance as well as how much variance is explained by the major EOF modes. There are a number of questions that may be addressed, such as the following: What is the minimal amount of observations needed to produce decent results? What areas are reconstructed with higher accuracy with given observation design, nearshore, offshore, open basins? What areas are most problematic to reconstruct, complicated coastline, straits and channels, semi-enclosed basins, regions of river influence? Are there some specific locations that can be used as a proxy for larger regions? Is it possible to measure SST/SSS just at these points in order to give enough input for successful reconstruction?

## 5 Conclusions

The present study was aimed to implement EOF-based statistical reconstruction technique into the data assimilation of the forecast model, and to study the feasibility of such assimilation method. Gridded EOF modes were determined from the 5 yr long model results. “Observational” EOF amplitudes were found each day to minimize the RMSD between the reconstructed and observed values at the observation points, using a time-dependent technique where both the amplitudes and their time rate of change were searched for the best fit. In this procedure, a time window of 30 d was selected that ensured acceptable SST and SSS reconstruction patterns by three leading EOF modes throughout the whole study period from 1 May to 31 December 2015. The study used about 330 000 FerryBox observations along four ship tracks from 1 May to 31 December 2015, and 370 observations from research vessels. Statistically gridded observations were assimilated into the model daily by the relaxation techniques, using restoring times of 5 and 10 d.

The tested EOF-based data assimilation (DA) method decreased RMSD of surface temperature (SST) and salinity (SSS) in the NE Baltic Sea by 22 % and 34 %, respectively, compared to the control run without DA. Using the observation-estimated amplitudes of the pre-calculated gravest model-based EOF modes, the method is able to follow on the regular grid the pointwise observed temporal changes of the mean state and of the major basin-scale gradients. DA with EOF reconstruction technique was found to be feasible for further implementation studies, since (1) the method that works on the large-scale patterns (mesoscale features are neglected by taking only the leading EOF modes) improves the high-resolution model performance by comparable or even better degree than in the other published studies, and (2) the method is computationally effective.

*Code availability.* The model code has been developed by the Baltic MFC partners. Presently it is frozen and not being developed anymore. The DA scripts and demonstrated model results can be requested by contacting the corresponding author. All the observational data used are freely available as described in Sect. 2.2.

*Author contributions.* MZ carried out data assimilation experiments and performed analysis of the results. JE worked on theoretical aspects and performed gridded reconstruction of observations. PL worked with the circulation model. All authors contributed to the discussion, planning and writing.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* The study was supported by the PhD programme for Mihhail Zujev and the institutional research funding. A larger BAL MFC team worked on the development of the HBM model within the EU projects MyOcean, MyOcean2 and MyOcean-FO. There is ongoing activity to develop and maintain Baltic monitoring and forecasting services within the CMEMS. This cooperation is highly acknowledged.

*Review statement.* This paper was edited by Markus Meier and reviewed by four anonymous referees.

## References

- Alenius, P., Myrberg, K., and Nekrasov, A.: The physical oceanography of the Gulf of Finland: a review, *Boreal Environ. Res.*, 3, 97–125, 1998.
- Andersen, J. H., Carstensen, J., Conley, D. J., Dromph, K., Fleming-Lehtinen, V., Gustafsson, B. G., Josefson, A. B., Norikko, A., Villnäs, A., and Murray, C.: Long-term temporal and spatial trends in eutrophication status of the Baltic Sea, *Biol. Rev.*, 92, 135–149, <https://doi.org/10.1111/brv.12221>, 2017.
- Axell, L. and Liu, Y.: Application of 3-D ensemble variational data assimilation to a Baltic Sea reanalysis 1989–2013, *Tellus A*, 68, 24220, <https://doi.org/10.3402/tellusa.v68.24220>, 2016.
- Berg, P. and Poulsen, J. W.: Implementation details for HBM, DMI Technical report No. 12–11, Copenhagen, 2012.
- Buizza, R., Brönnimann, S., Haimberger, L., Laloyaux, P., Martin, M. J., Fuentes, M., Alonso-Balmaseda, M., Becker, A., Blaschek, M., Dahlgren, P., and De Boisseson, E.: The EU-FP7 ERA-CLIM2 project contribution to advancing science and production of earth system climate reanalyses, *B. Am. Meteorol. Soc.*, 99, 1003–1014, <https://doi.org/10.1175/BAMS-D-17-0199.1>, 2018.
- Bullock Jr., O. R., Foroutan, H., Gilliam, R. C., and Herwehe, J. A.: Adding four-dimensional data assimilation by analysis nudging to the Model for Prediction Across Scales – Atmosphere (version 4.0), *Geosci. Model Dev.*, 11, 2897–2922, <https://doi.org/10.5194/gmd-11-2897-2018>, 2018.
- Carrasi, A., Bocquet, M., Bertino, L., and Evensen, G.: Data assimilation in the geosciences: An overview of methods, issues, and perspectives, *Wires Clim. Change*, 9, e535, <https://doi.org/10.1002/wcc.535>, 2018.
- Elken, J., Zujev, M., and Lagema, P.: Reconstructing sea surface temperature and salinity fields in the northeastern Baltic from observational data, based on sub-regional Empirical Orthogonal Function (EOF) patterns from models, in: 2018 IEEE/OES Baltic International Symposium (BALTIC), 12–15 June 2019, Klaipeda, Lithuania, 1–8, <https://doi.org/10.1109/BALTIC.2018.8634845>, 2018.
- Elken, J., Zujev, M., She, J., and Lagema, P.: Reconstruction of large-scale sea surface temperature and salinity fields using sub-regional EOF patterns from models, *Front. Earth Sci.*, 7, 232, <https://doi.org/10.3389/feart.2019.00232>, 2019.
- Fu, W., She, J., and Zhuang, S.: Application of an Ensemble Optimal Interpolation in a North/Baltic Sea model: Assimilating temperature and salinity profiles, *Ocean Model.*, 40, 227–245, <https://doi.org/10.1016/j.ocemod.2011.09.004>, 2011.



- Fujii, Y., Remy, E., Zuo, H., Oke, P. R., Halliwell, G. R., Gasparin, F., Benkiran, M., Loose, N., Cummings, J., Xie, J., and Xue, Y.: Observing system evaluation based on ocean data assimilation and prediction systems: on-going challenges and future vision for designing/supporting ocean observational networks, *Front. Mar. Sci.*, 6, 417, <https://doi.org/10.3389/fmars.2019.00417>, 2019.
- Golbeck, I., Li, X., Janssen, F., Brüning, T., Nielsen, J. W., Huess, V., Söderkvist, J., Büchmann, B., Siiriä, S. M., Vähä-Piikkiö, O., Hackett, B., Kristensen, N., Engedahl, H., Blockey, E., Sellar, A., Lagema, P., Ozer, J., Legrand, S., Ljungemyr, P., and Axell, L.: Uncertainty estimation for operational ocean forecast products – a multi-model ensemble for the North Sea and the Baltic Sea, *Ocean Dyn.*, 65, 1603–1631, <https://doi.org/10.1007/s10236-015-0897-8>, 2015.
- Golbeck, I., Izotova, J., Jandt, S., Janssen, F., Lagema, P., Brüning, T., Huess, V., and Hartman, A.: Quality Information Document (QUID) Baltic Sea Physical Analysis and Forecasting Product, <https://resources.marine.copernicus.eu/documents/QUID/CMEMS-BAL-QUID-003-006.pdf> (last access: 10 August 2020), 2018.
- Goodliff, M., Bruening, T., Schwichtenberg, F., Li, X., Lindenthal, A., Lorkowski, I., and Nerger, L.: Temperature assimilation into a coastal ocean-biogeochemical model: assessment of weakly and strongly coupled data assimilation, *Ocean Dyn.*, 69, 1217–1237, <https://doi.org/10.1007/s10236-019-01299-7>, 2019.
- Gregg, W. W., Friedrichs, M. A., Robinson, A. R., Rose, K. A., Schlitzer, R., Thompson, K. R., and Doney, S. C.: Skill assessment in ocean biological data assimilation, *J. Marine Syst.*, 76, 16–33, <https://doi.org/10.1016/j.jmarsys.2008.05.006>, 2009.
- Haines, K.: Ocean Reanalyses, in: *New Frontiers in Operational Oceanography*, Florida State University, 545–562, <https://doi.org/10.17125/gov2018.ch19>, 2018.
- Hernandez, F., Blockley, E., Brassington, G. B., Davidson, F., Divakaran, P., Drévilion, M., Ishizaki, S., Garcia-Sotillo, M., Hogan, P. J., Lagema, P., and Levier, B.: Recent progress in performance evaluations and near real-time assessment of operational ocean products, *J. Oper. Oceanogr.*, 8, s221–s238, <https://doi.org/10.1080/1755876X.2015.1050282>, 2015.
- Holland, W. R. and Malanotte-Rizzoli, P.: Assimilation of altimeter data into an ocean circulation model: Space versus time resolution studies, *J. Phys. Oceanogr.*, 19, 1507–1534, [https://doi.org/10.1175/1520-0485\(1989\)019<1507:AOADIA>2.0.CO;2](https://doi.org/10.1175/1520-0485(1989)019<1507:AOADIA>2.0.CO;2), 1989.
- Høyer, J. L. and She, J.: Optimal interpolation of sea surface temperature for the North Sea and Baltic Sea, *J. Mar. Syst.*, 65, 176–189, <https://doi.org/10.1016/j.jmarsys.2005.03.008>, 2007.
- Huess, V.: Product User Manual For Baltic Sea Physical Analysis and Forecasting Product, available at: <http://marine.copernicus.eu/documents/PUM/CMEMS-BAL-PUM-003-006.pdf>, last access: 2 May 2020.
- Janssen, F., Schrum, C., and Backhaus, J. O.: 1999. A climatological data set of temperature and salinity for the Baltic Sea and the North Sea, *Deutsche Hydrographische Zeitschrift*, 51, 5–245, <https://doi.org/10.1007/BF02933676>, 1999.
- Jevrejeva, S., Drabkin, V. V., Kostjukov, J., Lebedev, A. A., Lepäranta, M., Mironov, Y. U., Schmelzer, N., and Sztobryn, M.: Baltic Sea ice seasons in the twentieth century, *Clim. Res.*, 25, 217–227, <https://doi.org/10.3354/cr025217>, 2004.
- Johansson, J.: Total and regional runoff to the Baltic Sea, Baltic Sea environment fact sheet, available at: <http://www.helcom.fi/baltic-sea-trends/environment-fact-sheets/>, 2017 (last access: 8 April 2020).
- Karlson, B., Andersson, L. S., Kaitala, S., Kronsell, J., Mohlin, M., Seppälä, J., and Wranne, A. W.: A comparison of Ferrybox data vs. monitoring data from research vessels for near surface waters of the Baltic Sea and the Kattegat, *J. Marine Syst.*, 162, 98–111, <https://doi.org/10.1016/j.jmarsys.2016.05.002>, 2016.
- Köuts, T. and Omstedt, A.: Deep water exchange in the Baltic Proper, *Tellus A*, 45, 311–324, <https://doi.org/10.3402/tellusa.v45i4.14895>, 1993.
- Kozlov, I., Dailidienė, I., Korosov, A., Klemas, V., and Mingėlaitė, T.: MODIS-based sea surface temperature of the Baltic Sea Curonian Lagoon, *J. Marine Syst.*, 129, 157–165, <https://doi.org/10.1016/j.jmarsys.2012.05.011>, 2014.
- Laanemets, J., Zhurbas, V., Elken, J., and Vahtera, E.: Dependence of upwelling-mediated nutrient transport on wind forcing, bottom topography and stratification in the Gulf of Finland: model experiments, *Boreal Environ. Res.*, 14, 213–225, 2009.
- Lagema, P.: Operational forecasting in Estonian marine waters, Thesis on Natural and Exact Sciences, B128, Tallinn University of Technology, 2012.
- Le Traon, P. Y., Reppucci, A., Alvarez Fanjul, E., Aouf, L., Behrens, A., Belmonte, M., Bentamy, A., Bertino, L., Brandt, V. E., Kreiner, M., and Benkiran, M.: From observation to information and users: the Copernicus Marine Service perspective, *Front. Mar. Sci.*, 6, 234, <https://doi.org/10.3389/fmars.2019.00234>, 2019.
- Liblik, T., Laanemets, J., Raudsepp, U., Elken, J., and Suhova, I.: Estuarine circulation reversals and related rapid changes in winter near-bottom oxygen conditions in the Gulf of Finland, *Baltic Sea, Ocean Sci.*, 9, 917–930, <https://doi.org/10.5194/os-9-917-2013>, 2013.
- Lilover, M. J., Lips, U., Laaneau, J., and Liljebadh, B.: Flow regime in the Irbe Strait, *Aquat. Sci.*, 60, 253–265, <https://doi.org/10.1007/s000270050040>, 1998.
- Lips, U., Lips, I., Kikas, V., and Kuvaldina, N.: May. Ferrybox measurements: a tool to study meso-scale processes in the Gulf of Finland (Baltic Sea), 2008 IEEE/OES US/EU-Baltic International Symposium, Date 27–29 May 2008, location Tallinn, Estonia, IEEE, <https://doi.org/10.1109/BALTIC.2008.4625536>, 2008.
- Liu, Y. and Fu, W.: Assimilating high-resolution sea surface temperature data improves the ocean forecast potential in the Baltic Sea, *Ocean Sci.*, 14, 525–541, <https://doi.org/10.5194/os-14-525-2018>, 2018.
- 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, <https://doi.org/10.5194/bg-14-2113-2017>, 2017.
- Männik, A. and Merilain, M.: Verification of different precipitation forecasts during extended winter-season in Estonia, *HIRLAM Newsletter*, 52, 65–70, 2007.
- Martin, M. J., Balmaseda, M., Bertino, L., Brasseur, P., Brassington, G., Cummings, J., Fujii, Y., Lea, D. J., Lellouche, J. M., Mogensen, K., and Oke, P. R.: Status and future of data assimilation in operational oceanography, *J. Oper. Oceanogr.*, 8, s28–s48, <https://doi.org/10.1080/1755876X.2015.1022055>, 2015.

- Meier, H. E. M., Eilola, K., Almroth-Rosell, E., Schimanke, S., Kniebusch, M., Höglund, A., Pemberton, P., Liu, Y., Väli, G., and Saraiva, S.: Disentangling the impact of nutrient load and climate changes on Baltic Sea hypoxia and eutrophication since 1850, *Clim. Dyn.*, 53, 1145–1166, <https://doi.org/10.1007/s00382-018-4296-y>, 2019.
- Moore, A. M. and Reason, C. J.: The response of a global ocean general circulation model to climatological surface boundary conditions for temperature and salinity, *J. Phys. Oceanogr.*, 23, 300–328, [https://doi.org/10.1175/1520-0485\(1993\)023<0300:TROAGO>2.0.CO;2](https://doi.org/10.1175/1520-0485(1993)023<0300:TROAGO>2.0.CO;2), 1993.
- Moore, A. M., Martin, M. J., Akella, S., Arango, H., Balmaseda, M. A., Bertino, L., Ciavatta, S., Cornuelle, B., Cummings, J., Frolov, S., and Lermusiaux, P.: Synthesis of ocean observations using data assimilation for operational, real-time and reanalysis systems: A more complete picture of the state of the ocean, *Front. Mar. Sci.*, 6, <https://doi.org/10.3389/fmars.2019.00090>, 2019.
- Omstedt, A. and Axell, L. B.: Modeling the variations of salinity and temperature in the large Gulfs of the Baltic Sea, *Cont. Shelf Res.*, 23, 265–294, [https://doi.org/10.1016/S0278-4343\(02\)00207-8](https://doi.org/10.1016/S0278-4343(02)00207-8), 2003.
- Placke, M., Meier, H.E., Gräwe, U., Neumann, T., Frauen, C., and Liu, Y.: Long-term mean circulation of the Baltic Sea as represented by various ocean circulation models, *Front. Mar. Sci.*, 5, 287, <https://doi.org/10.3389/fmars.2018.00287>, 2018.
- Raudsepp, U. and Elken, J.: Application of the Bryan-Cox-Type Ocean Model to reproduce synoptic and mesoscale variability of the Irbe Strait salinity front, *Deutsche Hydrografische Zeitschrift*, 51, 477–488, <https://doi.org/10.1007/BF02764168>, 1999.
- Ravichandran, M., Behringer, D., Sivareddy, S., Girishkumar, M. S., Chacko, N., and Harikumar, R.: Evaluation of the global ocean data assimilation system at INCOIS: the tropical Indian Ocean, *Ocean Model.*, 69, 123–135, <https://doi.org/10.1016/j.ocemod.2013.05.003>, 2013.
- Rodhe, J.: The Baltic and North Seas: a process-oriented review of the physical oceanography, *The sea*, 11, 699–732, 1998.
- Savchuk, O. P.: Large-scale nutrient dynamics in the Baltic Sea, 1970–2016. *Front. Mar. Sci.*, 5, 95, <https://doi.org/10.3389/fmars.2018.00095>, 2018.
- She, J.: Assessment of Baltic Sea observations for operational oceanography, in: Proceedings of the 8th EuroGOOS International Conference (Bergen: EuroGOOS), edited by: Buch, E., Fernández, V., Eparkhina, D., Goringe, P., and Nolan, G., 7, 79–87, 2018.
- She, J., Meier, M., Darecki, M., Goringe, P., Huess, V., Kouts, T., Reissmann, J. H., and Tuomi, L.: Baltic Sea Operational Oceanography-A Stimulant for Regional Earth System Research, 3–5 October 2017, Bergen, Norway, *Front. Earth Sci.*, 8, <https://doi.org/10.3389/feart.2020.00007>, 2020.
- Stauffer, D. R. and Seaman, N. L.: Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data, *Mon. Weather Rev.*, 118, 1250–1277, [https://doi.org/10.1175/1520-0493\(1991\)119<0734:UOFDDA>2.0.CO;2](https://doi.org/10.1175/1520-0493(1991)119<0734:UOFDDA>2.0.CO;2), 1990.
- Stow, C. A., Jolliff, J., McGillicuddy Jr., D. J., Doney, S. C., Allen, J. I., Friedrichs, M. A., Rose, K. A., and Wallhead, P.: Skill assessment for coupled biological/physical models of marine systems, *J. Marine Syst.*, 76, 4–15, <https://doi.org/10.1016/j.jmarsys.2008.03.011>, 2009.
- Stramska, M. and Bialogrodzka, 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, <https://doi.org/10.1016/j.oceano.2015.04.004>, 2015.
- Tuomi, L., She, J., Lorkowski, I., Axell, L., Lagema, P., Schwichtenberg, F., and Huess, V.: Overview of CMEMS BAL MFC Service and Developments, Proceedings of the Eight EuroGOOS International Conference, 3–5 October 2017, Bergen, Norway, 261–267, ISBN 978-2-9601883-3-2, 2018.
- Uiboupin, R. and Laanemets, J.: Upwelling characteristics derived from satellite sea surface temperature data in the Gulf of Finland, Baltic Sea, *Boreal Environ. Res.*, 14, 297–304, 2009.
- Uiboupin, R. and Laanemets, J.: Upwelling parameters from bias-corrected composite satellite SST maps in the Gulf of Finland (Baltic Sea), *IEEE Geosci. Remote S.*, 12, 592–596, <https://doi.org/10.1109/LGRS.2014.2352397>, 2015.
- Yurkovskis, A., Wulff, F., Rahm, L., Andruzaitis, A., and Rodriguez-Medina, M.: A nutrient budget of the Gulf of Riga; Baltic Sea, *Estuar. Coast. Shelf S.*, 37, 113–127, <https://doi.org/10.1006/ecss.1993.1046>, 1993.
- Zujev, M. and Elken, J.: Testing marine data assimilation in the northeastern Baltic using satellite SST products from the Copernicus Marine Environment Monitoring Service, Proceedings of the Estonian Academy of Sciences, 67, 217–230, <https://doi.org/10.3176/proc.2018.3.03>, 2018.

# Curriculum Vitae

## Personal data

Name: Mihhail Zujev  
Date of birth: 11.04.1979  
Place of birth: Estonia  
Citizenship: Estonia

## Contact data

E-mail: mihhail.zujev@taltech.ee

## Education

2013–2020 Tallinn University of Technology, Earth Sciences – PhD  
2011–2013 Tallinn University of Technology, Earth Sciences – MSc  
2007–2011 Estonian Maritime Academy, Hydrography – BSc  
1986–1994 High school #19

## Language competence

English Fluent  
Estonian Fluent  
Russian Fluent

## Professional employment

2013 – Tallinn University of Technology



## Elulookirjeldus

### Isikuandmed

Nimi:	Mihhail Zujev
Sünniaeg:	11. aprill 1979
Sünnikoht:	Eesti
Kodakondsus:	Eesti

### Kontaktandmed

E-post:	mihhail.zujev@taltech.ee
---------	--------------------------

### Hariduskäik

2013–2020	Tallinna Tehnikaülikool, Maa-teadused – PhD
2011–2013	Tallinna Tehnikaülikool, Maa-teadused – MSc
2007–2011	Eesti Mereakadeemia, hüdrograafia – BSc
1986–1994	Tallinna Tõnismäe Reaalkool

### Keelteoskus

Inglise keel	Kõrgtase
Eesti keel	Kõrgtase
Vene keel	Kõrgtase

### Teenistuskäik

2013–	Tallinna Tehnikaülikool
-------	-------------------------

ISSN 2585-6901 (PDF)  
ISBN 978-9949-83-712-0 (PDF)