

STYRELSEN FÖR  
**VINTERSJÖFARTSFORSKNING**  
WINTER NAVIGATION RESEARCH BOARD

Research Report No 126

Lars Axell, Tomas Landelius

**BALTIC FORECAST IMPROVEMENTS USING REMOTE  
SENSING DATA – 2 (BALTIC FIRE-2): FINAL REPORT**

Finnish Transport and Communications Agency

Finnish Transport Infrastructure Agency

Finland

Swedish Maritime Administration

Swedish Transport Agency

Sweden

Talvimerenkulun tutkimusraportit — Winter Navigation Research Reports  
ISSN 2342-4303  
ISBN 978-952-311-874-4

## FOREWORD

In this report no 126, the Winter Navigation Research Board presents the results of research project Baltic Forecast Improvements using Remote Sensing Data – 2 (Baltic FIRE-2): Final Report. The project was a follow up for Baltic Fire-1. Building on the previous research, satellite data usage was further enhanced to improve the initial conditions form SMHI's operational ice-ocean forecast model to improve forecast quality.

The Winter Navigation Research Board warmly thanks Lars Axell and Tomas Landelius for this report.

Helsinki

August 2023

Ville Häyrynen

Finnish Transport and Communications Agency

Anders Dahl

Swedish Maritime Administration

Helena Orädd

Finnish Transport Infrastructure Agency

Stefan Eriksson

Swedish Transport Agency

# Baltic Forecast Improvements using Remote Sensing Data – 2 (Baltic FIRE-2): Final Report

Project number: W22-6 BalticFire-2

Authors: Lars Axell (Project leader) and Tomas Landelius,  
Research and Development,  
Swedish Meteorological and Hydrological Institute (SMHI),  
Sweden

Date: 2022-11-30

## 1. Introduction

The numerical ice-ocean forecasting model used at SMHI today is NEMO-Nordic 1.0. It covers the Baltic Sea and North Sea region out to the western English Channel in the west and up to a latitude line between Scotland and Norway in the northern North Sea. The model code is based on NEMO (Nucleus for European Modelling of the Ocean; Madec et al. 2019) version 3.6. The general model setup was described by Hordoir et al. (2015, 2019), and the performance of its ice model was described by Pemberton et al. (2017). Soon, however, the forecasting model at SMHI will be upgraded to NEMO-Nordic 2.0 (see Kärnä et al., 2021), which is based on NEMO version 4.0.

Today's ice ocean forecasts made by NEMO-Nordic include the most important ice variables. However, to make good forecasts requires using good initial conditions (the state that we start the forecast model from), good boundary conditions (including meteorological forcing), and a good numerical model that can accurately calculate future states (a forecast). What we are lacking today is mainly the first criterion, i.e. a good initial condition. Without proper data assimilation, where we use observations to correct the initial conditions, it doesn't really matter much if the numerical model is good or not, if we start each forecast from a sea state with large errors. As a result, today's ice forecasts are of little use. Large areas covered with sea ice may be ice free in the forecast, and vice versa. Though the thermodynamics in the model is generally good from a climatological point of view, ice extent is challenging to forecast, because small errors in ocean heat content may lead to unacceptable errors in e.g. ice extent. This project is about increasing the use of remote sensing data (satellite data) with the goal of improving the initial states of ice-ocean forecasts, but also weather forecasts due to the loose coupling between the models.

Daily ice charts are available for ice assimilation, but they are created manually by ice analysts with some degree of subjectivity. Though the ice edge is usually well defined, the ice concentration and ice thickness both have large uncertainties. In addition, snow depth is

not available from these charts. Further, in atmospheric data assimilation it is well established that direct assimilation of satellite data is preferable to assimilation of products which can result in inconsistencies.

In this project we have set out to start studying how to use satellite data (Sentinel-1 C-band synthetic aperture radar data and AMSR2 passive microwave data, both of which are rather insensitive to the state of the atmosphere) together with Radiative Transfer Models (RTMs) to compute model equivalents, i.e. what the satellite should observe given the ice-ocean model state. In this way we should be able to improve the constraint on the ice-ocean forecast model with respect to ice extent, ice concentration, ice thickness, sea surface temperature and possibly also snow depth on the ice.

## 2. Work done in the Work Packages (WPs)

The Baltic FIRE project was divided into two parts (henceforth called Baltic FIRE-1 and Baltic FIRE-2, respectively), and in this year's contract (Baltic FIRE-2) the following work was planned (list of WPs, descriptions and briefly what we have done):

### WP0: Administration and dissemination

Early this year (April 12, 2022), we had a meeting with a representative from the Swedish Ice Service to discuss interesting situations during our target period, the ice winter 2016/2017. This resulted in three approximately 5-7-day periods in February, March and April 2017, that we selected for extra study in the coming data assimilation experiments (WP3); see the Results section for details.

An in-house presentation was given at SMHI on June 2, 2022, during which the project was presented. Further, on June 20, 2022, an informal meeting took place at the Swedish Maritime Administration (SMA) in Norrköping, Sweden, during which the progress of the project was discussed together with representatives from SMA.

Future dissemination is planned for next year, after the end of this project. We think this will lead to 1-2 papers in peer-reviewed journals. Further, a presentation of the work is planned for the upcoming conference EGU 2023 in Vienna, and probably in other conferences as well.

### WP1: Improve the new observation operator for satellite data

It was foreseen that the work with the RTMs would need to be revisited during the second part of the project and that knowledge gained from Baltic FIRE-1 should form a basis for an iterative process with the goal to see to that the quality of the modeled satellite data was good enough for use in the data assimilation.

The NEMO runs in Baltic FIRE-1 often had too much ice. This meant that the NEMO data couldn't be used as is for machine learning where input is paired with output. Instead we had

to select cases where NEMO agreed with the ice chart and then model ice and open sea cases separately. In Baltic FIRE-2, a new NEMO version was used, that together with modifications described below show reduced biases in the ice cover. However, there was still a noticeable bias and cases with and without ice in both model and ice charts had to be singled out for training and tuning.

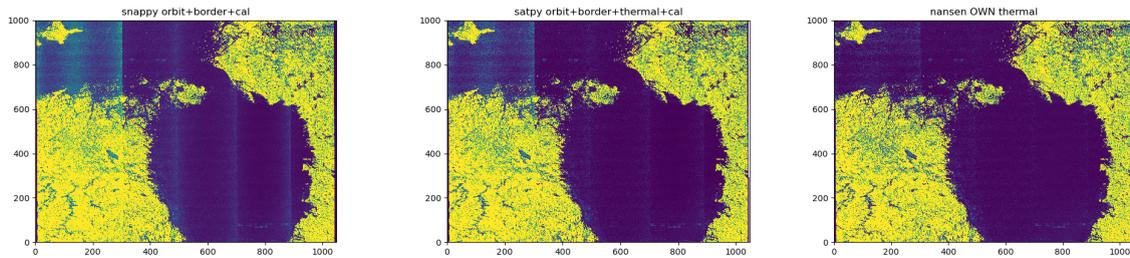
Even if the machine learning approach with KNNR from Baltic FIRE-1 showed promising results we had not dropped the idea to use physical radiative transfer models like SMRT. Nor had we settled for KNNR as the optimal machine learning approach. Ensemble methods are also used in connection to machine learning. By averaging over a group (ensemble) of predictors one often gets better predictions than with the best individual predictor. As an example of an ensemble method, you can train a group of Decision Tree regressors, each on a different random subset of the training set. Such an ensemble of Decision Trees is called a Random Forest and is one of the most powerful Machine Learning algorithms available today. One way to get spread in the ensemble is to train its members on different random subsets of the training set. When sampling is performed with replacement, this method is called bagging and when sampling is performed without replacement, it is called pasting. The Random Forest algorithm introduces extra randomness when growing trees. Instead of searching for the very best feature when splitting a node in the tree, it searches for the best feature among a random subset of features.

Boosting refers to any Ensemble method that can combine several weak (simple) learners into a strong learner. The general idea of most boosting methods is to train predictors sequentially, each trying to correct its predecessor. The main difference between random forests and gradient boosting lies in how the decision trees are created and appended. Unlike random forests, the decision trees in gradient boosting are built sequentially. Extreme Gradient Boosting (XGBoost) is an efficient open-source implementation of the gradient boosting algorithm. It has shown good results on standard machine learning datasets like the University of California at Irvine (UCI) machine learning repository and is often a component of winning solutions in machine learning competitions (Chen and Guestrin, 2016). In Baltic FIRE-2 we tried both the RandomForestRegressor (<https://scikit-learn.org>) and XGBoost (<https://xgboost.readthedocs.io>). Both these ensemble methods outperformed KNNR and XGBoost was better and faster than the Random Forest. The results using XGBoost are presented in this report.

We have had ongoing discussions with scientists from Chalmers University of Technology as well as the Technical University of Denmark who are specialists in both active and passive microwave measurements related to sea ice. The ambition was to tune the SMRT parameters to come up with a setting that could challenge the performance of the KNNR method. From these contacts we learned that the setting used for SMRT during Baltic FIRE-1 was probably not optimal. We had modeled the snow and ice covers with single layers but the experts noted that crucial scattering happens at the interfaces between layers and hence that multi-layered descriptions of the snow and ice packs should be employed. Below we address the developments made for the active (Sentinel-1 C-SAR) and passive (AMSR2) data respectively.

## Sentinel-1 C-SAR

As described in the final report of Baltic FIRE-1, the Sentinel-1 C-SAR data is occasionally affected by significant thermal noise. During Baltic FIRE-1, the noise reduction procedure provided by ESA (European Space Agency) was applied but there were still situations when the thermal noise remained at a troublesome level. In Baltic FIRE-2 we have adapted an alternative method for noise reduction developed at the Nansen Environmental and Remote Sensing Center (Park et al., 2018). This improvement can be seen in Figure 1.



**Figure 1:** Example of thermal noise reduction of Sentinel-1 C-SAR backscatter. *Left: Without thermal noise correction. Middle: Standard ESA noise reduction. Right: Alternative noise reduction using methods developed at NERSC. Note that the images are upside down with respect to reality.*

In Baltic FIRE-1 we tried to model the Sentinel-1 C-SAR backscatter from snow and ice surfaces with the SMRT model. Since the SMRT model lacks support for open sea surfaces we instead opted for a machine learning model (KNNR) for those surfaces. However, when we tried to apply the KNNR model to the snow/ice surfaces it turned out that this approach outperformed the SMRT model which showed large systematic errors and low correlation. From discussions with the radar and RTM experts in Chalmers and DTU we learned that a single layer would probably not be enough to model the scattering processes in snow and ice. Instead we opted for a three layer model with input from the NEMO, see Table 1.

*Table 1: NEMO input to the three layered snow and ice profiles used with SMRT.*

	Temperature	Salinity	Thickness
Snow layer 1	Snow temperature	Zero	Snow thickness / 3
Snow layer 2	Snow temperature	Mean ice salinity / 2	Snow thickness / 3
Snow layer 3	Surface ice temp	Mean ice salinity	Snow thickness / 3
Ice layer 1	Surface ice temp	Mean ice salinity	Ice thickness / 3
Ice layer 2	Mean ice temp	Mean ice salinity	Ice thickness / 3
Ice layer 3	Sea surface temp	Sea surface salinity	Ice thickness / 3
Sea	Sea surface temp	Sea surface salinity	—

The reason for choosing KNNR was that this model is known to perform well (decent error statistics) when there is lots of data. It also works like a test regarding whether there is enough information in the input to model the output or not. A drawback is that it is time

consuming. In Baltic FIRE-2 we turned to XGBoost as this is a well known model that has shown good performance on a vast number of problems as mentioned above.

Four XGBoost models were trained, one for each combination of surface (water or ice) and polarization (HH/HV). For the water surface models all the input regarding snow and ice were omitted. The NEMO input came from a reference run without data assimilation but with the same settings as the ones used in the later data assimilation experiments. The number of water cases outnumbers the ice cases and in order to have enough ice cases the thresholds for filtering out these were set to 95 % ice fraction in NEMO and 90 % ice cover in the ice chart. For the water cases the criteria is that there is zero ice in both NEMO and the ice chart.

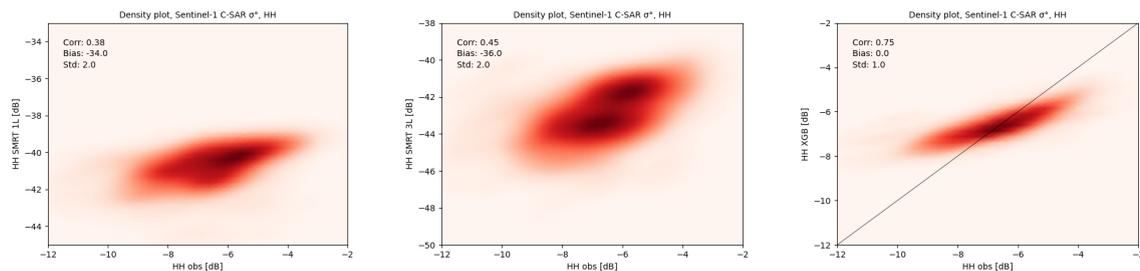
When training machine learning models one likes to have all inputs and outputs to have a mean value close to zero, a standard deviation of about one and a Gaussian like distribution. In order to obtain this to some extent we transformed the inputs and outputs as well as the observations with a power transformation, since many of our variables have long tailed distributions. The training was done using the scikit-learn and xgboost Python libraries (<https://scikit-learn.org>, <https://xgboost.readthedocs.io>) using the PowerTransformer for the transformations and XGBRegressor as the model. In order to tune the hyperparameters of the XGBRegressor we used the SklearnTuner with the BayesianOptimizationOracle from the Keras library ([https://keras.io/keras\\_tuner](https://keras.io/keras_tuner)). Cross validation was used with KFold(3) and 100 trials. The resulting model parameters were fairly similar and we opted for using the same settings for all four models: XGBRegressor(colsample\_bylevel=0.4, colsample\_bynode=0.6, colsample\_bytree=1.0, learning\_rate=0.06, max\_depth=5, n\_estimators=700, subsample=0.8). To evaluate the models we used data from March 2017. These data were not used during training. A nice thing with the XGBRegressor is that you obtain an estimate of the importance of different inputs. The order of importance and the obtained correlation (in the independent data from March 2017) is given in Table 2.

*Table 2: Importance of NEMO inputs for the XGBoost models of different surfaces (i: ice, w: water) and polarizations of Sentinel-1 C-SAR backscatter. Order of importance (decreasing order with 1 most important and entries above 0.05 marked in green) at the top of each cell together with the relative importance (sum normalized to 1) at the bottom. The values associated with the leftmost column is the correlation obtained for a given surface and polarization. Notation: sst - sea surface temperature, sss - sea surface salinity, ice th - ice thickness, ice mt - ice mean temperature, ice st - ice surface temperature, ice ms - ice mean salinity, Vb - brine volume, snw th - snow thickness, snw mt - snow mean temperature, sat az - satellite azimuth angle, sat el - satellite elevation, u gst - gust wind in u-direction, v gst - gust wind in v-direction, u - wind in u-direction, v - wind in v direction.*

	sst	sss	ice th	ice mt	ice st	ice ms	ice age	Vb	snw th	snw mt	sat az	sat el	u gst	v gst	u	v
hh i 0.67	13 0.03	7 0.05	4 0.08	11 0.03	6 0.06	9 0.04	5 0.06	14 0.03	2 0.08	8 0.05	2 0.08	1 0.30	15 0.03	16 0.03	10 0.03	12 0.03
hv i 0.38	16 0.03	7 0.06	3 0.11	15 0.03	12 0.04	5 0.07	4 0.08	13 0.04	2 0.12	8 0.06	6 0.07	1 0.12	13 0.04	11 0.05	10 0.05	9 0.05
hhw 0.97	7 0.01	8 0.01									2 0.18	1 0.59	4 0.06	5 0.04	6 0.04	3 0.06
hvw 0.76	7 0.03	8 0.02									6 0.07	3 0.21	3 0.16	1 0.22	4 0.16	5 0.14

The results from modeling the HH backscatter from Sentinel-1 C-SAR over snow and ice with different approaches are shown Figure 2. Here the criteria for filtering out pure snow/ice cases was that the NEMO ice fraction as well as the ice concentration from the ice chart

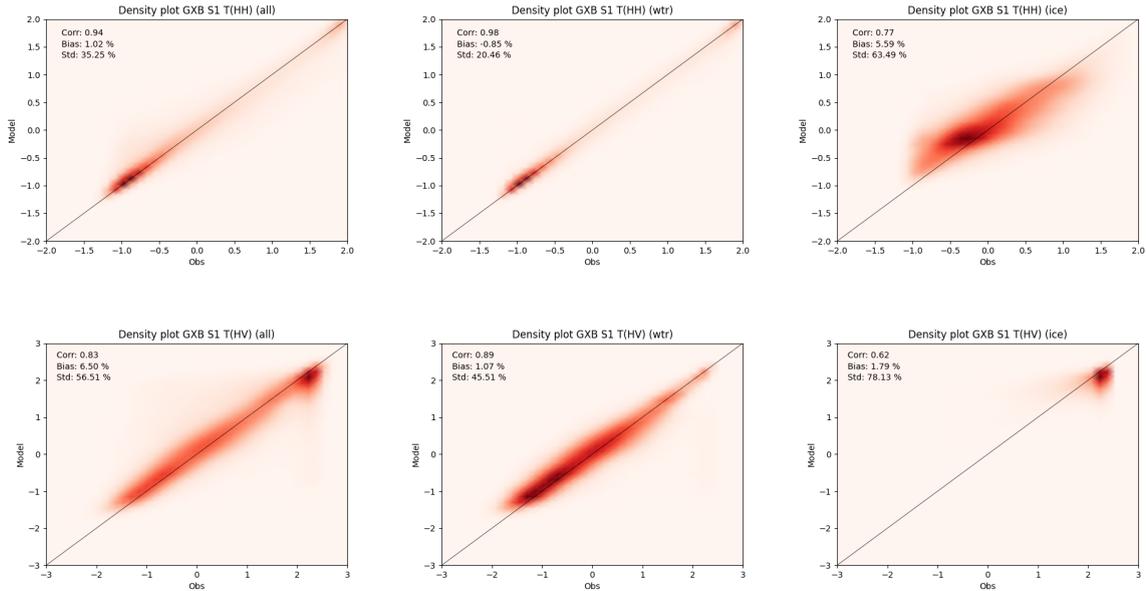
should both be above 99 %. Adding extra layers of snow and ice improves the SMRT results to some extent (e.g. correlation increase from 0.38 to 0.45) but the machine learning model based on XGBoost is clearly superior (correlation is 0.75). The experiments with the SMRT model were done at the end of the project and we have not had time to investigate the reason for the significant bias in the SMRT results. We will continue this work in the CAISA project where the experts on radar and RTMs are also involved. The tilt in the pattern of the XGBoost model is due to the transform applied to the output in order to make its distribution more Gaussian. The parameters of the transform are estimated from all cases (open sea and snow/ice). Since the open sea cases dominate and the backscatter behaves differently in these cases the estimate will fit better to the open sea cases. Still, the correlation is rather high and the bias is low also for the scarcer snow/ice cases. The results for the HV polarization with its lower signal to noise ratio are considerably worse and were left out here.



**Figure 2: Density plots for Sentinel-1 C-SAR HH backscatter from snow/ice surfaces. The observed data is on the x-axis in all panels. Left: HH modeled with SMRT using one snow and one ice layer. Middle: HH modeled with SMRT using three snow and three ice layers. Right: HH modeled with XGBoost.**

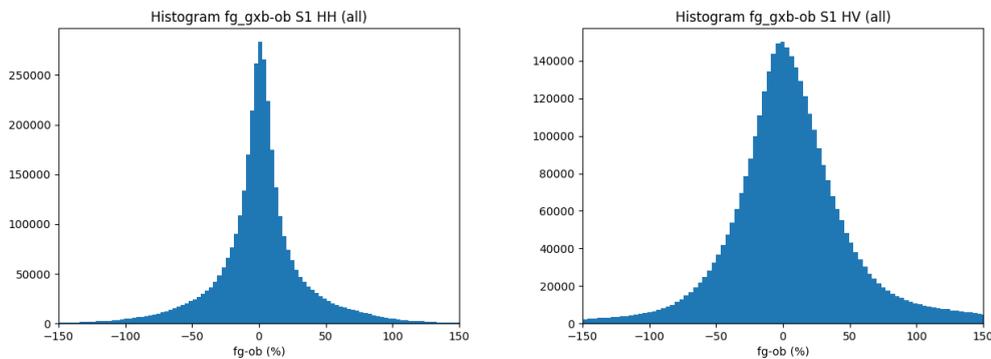
Results from modeling the HH as well as HV backscatter (here shown in terms of transformed values) from Sentinel-1 C-SAR over snow and ice with XGBoost are shown Figure 3. Note that due to more relaxed thresholds for the ice cases the rightmost panel in Figure 2 and the top right panel in Figure 3 are not identical since the latter includes more data (some water) and also is given in terms of transformed values. Also note that the densities are normalized with respect to the number of cases which means that the ice cases are in a minority in the leftmost panels of the figure. The reason why the plot for HV with all cases still contains the distinct feature of the ice cases must be that this polarization is sensitive to even small amounts of ice. Remember that the water case (middle panels) is restricted to zero ice. The separation of water and ice is more pronounced in the HV polarization where high values indicate ice whereas there is more overlap between the two cases in the HH polarization. On the other hand the quality of the model for the HV is lower. Hence assimilating a combination of the two should be beneficial.

The correlation numbers differ somewhat (a bit higher) from Table 2 since here they were computed for all values, not only for the independent data of March 2017. The error statistics are better for the water cases and overall better for HH than for HV. As can be seen from Table 2 the most important information for the water situations is that regarding satellite angles and the wind, neither of which are assimilated. However, since the ice fraction depends on knowing the contribution from both water and snow/ice surfaces the modeling of the water part is still important.



**Figure 3: Density plots for Sentinel-1 C-SAR HH and HV backscatter modeled with XGBoost. Top left: HH for all cases. Top middle: HH for open sea. Top right: HH for snow/ice. Bottom row shows the same as the top row but for HV polarization.**

The histogram for the differences between modeled and observed backscatter from the Sentinel-1 C-SAR is shown in Figure 4. Note that the error for the HH polarization is significantly lower than for the HV polarization and that the latter is more Gaussian than the former. These differences include contributions from both the model and observation errors. Still it can be seen as an indication that the observation error for the HV data in the data assimilation procedure should probably be set to a larger value than the error for the HH observations.



**Figure 4: Histograms for normalized (with respect to the standard deviation of the observed values) differences between modeled (XGBoost) and observed backscatter from Sentinel-1 C-SAR (all cases). Left: HH polarization. Right HV polarization.**

## AMSR2

In Baltic FIRE-1 we realized that the winter season 2016-2017 did not contain enough cases where the large footprints of the lower frequency AMSR2 channels were fully covered with ice or snow. As noted above we need clean cases with and without ice in order to train the machine learning observation operators and/or tune the parameters of the RTMs.

One idea was to extend the study with more winters in order to find enough data for such clean cases. However, the work on assimilation of Sentinel-1 C-SAR data took more time than expected and extending the project with more seasons turned out to be out of reach.

Work on AMSR2 data will continue in a related project called “Consistent Air-Ice-Sea Data Assimilation of Satellite Observations (CAISA)”, funded by the Swedish National Space Agency. Here the contributions from land surfaces in coastal areas will be taken into consideration. This will leave us with more cases where the AMSR2 footprint is fully covered with snow or ice.

### Test cases

Equipped with information from Baltic FIRE-1 regarding when satellite data was available we set up a meeting with people working at the operational ice service. Based on this information and their experiences from the ice season 2016-2017 we decided on three time periods for detailed case studies:

- Case A: February 9-15, 2017. This period includes the time of maximum ice extent during the selected winter season, 2017-02-12.
- Case B: March 15-20, 2017. A low pressure system gave rise to winds that caused ice drift and built up ice ridges.
- Case C: April 7-14, 2017. A brash ice barrier was located near the ice edge during the start of the period but began to crack up towards its end.

Results from the data assimilation experiments during these three periods are presented later on in this report; see section 3.

## WP2: Set up an ensemble forecast system for sea ice

Data assimilation implies correcting a numerical model by using information from observations, to help keep the model from deviating too far from reality. Since it is normally impossible to have observations available for all grid cells in the model, and for all model variables, some kind of a priori information is needed to spread the information from the observations to nearby grid cells for which we do not have observations. This information is used in the data assimilation step, which is an integral part of a forecasting system today. In some data assimilation systems, e.g. Optimal Interpolation (OI) and conventional three-dimensional variational data assimilation (3D-Var), this a priori information is determined ahead of time, e.g. by spreading the information using Gaussian functions that decay with the distance from the observation, so-called structure functions. In so-called ensemble data assimilation, an ensemble of model states is used to calculate statistics with more natural, but also more complex, structure functions.

The data assimilation system used at SMHI today with ocean forecast models is an example

of such an ensemble data assimilation system. Recently it was decided to start calling this system the Nordic Variational (NOVA) data assimilation system, as it is mainly used for the NEMO-Nordic setup and employs a so-called variational approach, as opposed to e.g. the Parallel Data Assimilation Framework (PDAF) used within Copernicus for daily forecasts in the Baltic Sea region. The form of NOVA employed at SMHI today is three-dimensional Ensemble Variational data assimilation (3D EnVar), and uses ensemble statistics from a static library of model states, created by running the NEMO-Nordic forecast model for several years.

In this project, that will not be possible unless we also calculate model equivalents of the satellite backscatter for several years, which we are not able to do now. Instead, we will employ four-dimensional Ensemble Variational (4D EnVar) data assimilation, and use an ice ensemble forecasting system to generate the required ensemble members. This has the advantage that the ensemble statistics has the potential to be more precise compared to when using a static ensemble. The drawback is that it is very expensive to run a full ensemble forecast system just to be able to perform data assimilation. However, since the success of this project heavily depends on the cross correlations between the satellite backscatter signal and the model parameters we wish to update (ice concentration, thickness etc.), we think the extra cost of running a full ensemble forecast system is justified. To save computing time, however, we decided to use only a small part of the full NEMO-Nordic model domain, that includes only the Gulf of Bothnia, with an open boundary in the south close to Åland (approximately latitude 60 degrees North).

To set up such an ensemble forecasting system, we needed to set up a script system that would make it possible to use perturbed atmospheric forcing. This was done by creating perturbed weight files that are used for interpolation from the atmospheric forcing grid. The perturbations were made in such a way that the original atmospheric forcing fields were shifted approximately 55 km north, south, east or west, thus simulating uncertainties of air temperature and wind forcing fields, and the timing of fronts etc. This gave us five different forcing fields to use (one unperturbed and four perturbed). In addition we implemented the possibility to run the model with or without the mixing effects of Langmuir circulation included (parameterization according to Axell, 2002), as well as using different numerical values of the strength of the ice. These are model parameters that have already been implemented into NEMO-4.0, to help testing possible improvements in mixed layer depths and sea surface temperature (the Langmuir Circulation) as well as changes in the ability of sea ice to drift and compress (the ice strength parameter). Using all possible combinations we were thus able to run an ensemble forecast system with 20 ensemble members (19 perturbed ensemble members and one unperturbed member). In case this is not enough, it is possible to include more uncertainties in the forcing or model parameters in the future, or introduce uncertainties in the initial conditions.

Figure 5 below shows an example of ensemble statistics of the variable ice concentration from the new ice ensemble forecasting system, for the date 2017-01-03. Panel (a) shows the unperturbed ensemble member (without data assimilation in this case), i.e. with optimal atmospheric forcing (unperturbed), without the mixing effects of Langmuir circulation and with the standard NEMO-value of the ice strength parameter. Panel (b) shows the ensemble mean, panel (c) shows the standard deviation (high values where the variability is high), panel (d) and (e) show the maximum and minimum values, respectively, and panel (f) shows

the operational ice chart for the same date.

When creating an ensemble forecast system, it is important that the spread of the ensemble encompasses future observations. Figure 5 shows that the ice extent in the analyzed ice chart in panel (f) falls somewhere between the extent in the maximum and minimum fields in panels (d) and (e), but rather close to the minimum extent in panel (e). Thus, the spread of the ensemble seems large enough, and perhaps even a bit too large.

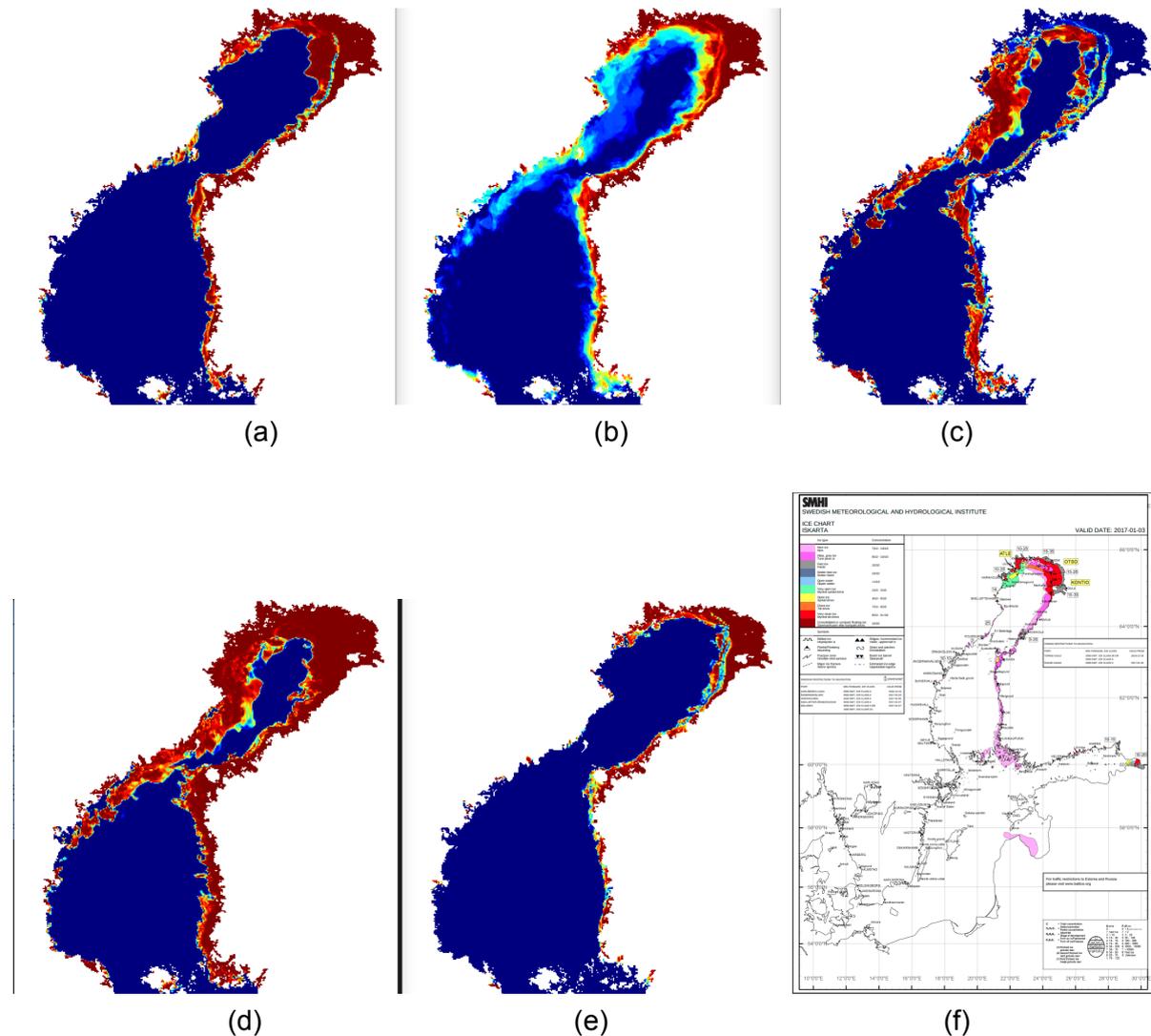


Figure 5: Examples of ice concentration fields in the ensemble ice forecasting system for the date 2017-01-03: (a) Unperturbed member, (b) ensemble mean, (c) standard deviation, maximum ice concentration, (e) minimum ice concentration, and (f) the operational ice chart (downloaded from SMHI: <https://www.smhi.se/data/oceanografi/havsis>).

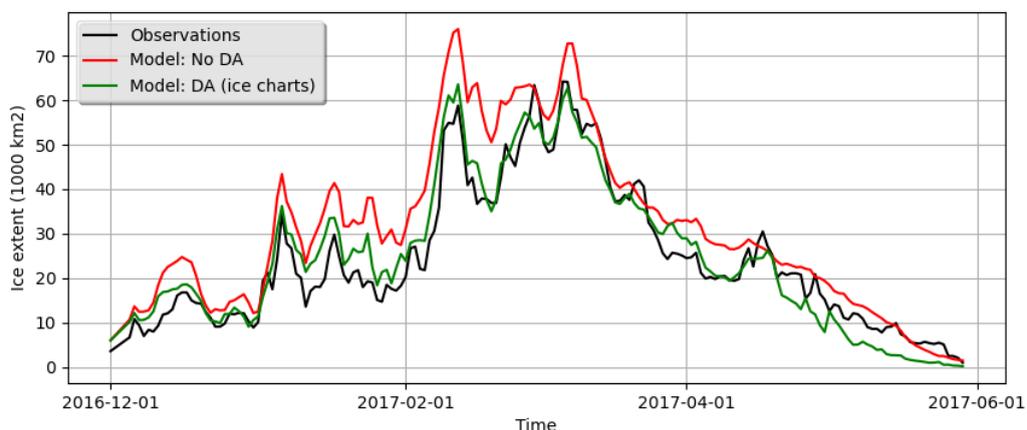
### WP3: Data assimilation experiments

Before this project began, we had already decided that we should use the data assimilation system used operationally at SMHI (NOVA), which is already used for the assimilation of Sea Surface Temperature (SST) as well as salinity and temperature profiles. Assimilation of

sea ice concentration and sea ice thickness had already been implemented into NOVA prior to this project, but now we have also added the possibility of calculating increments (corrections) for snow thickness as well. The main part, however, was to make it possible to read satellite backscatter observations and its model equivalents using the new observation operator developed in WP1.

The next step was to work on the NEMO ocean model code itself, to make it possible for NEMO to read the increments calculated by NOVA. This was done for the model code version NEMO-3.6, but we soon discovered that the data assimilation increments resulted in strange behavior in the sea ice model. Instead of trying to fix the problem in the relatively old model version NEMO-3.6, we decided to implement the ice data assimilation routines in the more recent model code NEMO-4.0, which resolved most of the problems. As a result, our local version of NEMO-4.0 can now read and handle model increments for sea ice concentration, sea ice thickness, and snow thickness. Thus, it is more akin to the forthcoming NEMO-Nordic 2.0, also based on NEMO-4.0, which will soon be operational at SMHI.

To test the new data assimilation system, we simulated the whole target period 2016-12-01 to 2017-05-31, both with data assimilation of ice concentration and ice thickness (using ice charts from the Finnish and Swedish Ice Services at FMI and SMHI, respectively), and without data assimilation (as a reference). The resulting sea ice extent for the whole target period can be seen in Figure 6 below. We can see that in the reference run without data assimilation (“Model: No DA”), the model overestimates the ice extent compared to observations (ice charts). When we assimilate ice concentration and thickness from ice charts, however (“Model: DA (ice charts)”), the results are much improved. As a result, the Mean Absolute Error (MAE) decreased from  $7.3 \times 10^3 \text{ km}^2$  to  $3.7 \times 10^3 \text{ km}^2$ , by assimilating ice chart data. The ensemble forecast system developed in WP2 was used for ensemble statistics. The conclusion so far is that the developed ensemble forecast system can be used with good results together with NOVA and NEMO-4.0.



*Figure 6. Improvements of modeled sea ice extent by assimilation of sea ice concentration and sea ice thickness from operational ice charts.*

## 3. Results

### 3.1 Timeseries of ice extent

After some tuning of parameters in the NOVA data assimilation code (e.g. assumed observation errors and maximum allowed difference between the model first guess and the observation), we simulated the whole target period with and without data assimilation, just like we did above in the case of assimilation of ice chart data. We made three assimilation runs in total (see Figure 7 below): One with assimilation of Sentinel-1 C-SAR HH polarization data (“Model (DA: S1HH)”), one with HV polarization data (“Model (DA: S1HV)”), and one with both HH and HV polarization data. However, the last experiment gave very bad results and is thus not shown in Figure 7. It is clear that assimilation of either HH or HV polarization data from Sentinel-1 improves the sea ice extent compared to the reference run without data assimilation, and that HH polarization gives best results. The calculated Mean Absolute Errors decreased from  $7.3 \times 10^3 \text{ km}^2$  to  $4.7 \times 10^3 \text{ km}^2$ , and  $5.3 \times 10^3 \text{ km}^2$  by assimilating Sentinel-1 C-SAR HH and HV polarization, respectively. In comparison, assimilation of ice charts directly gave slightly better results, with  $\text{MAE} = 3.7 \times 10^3 \text{ km}^2$ .

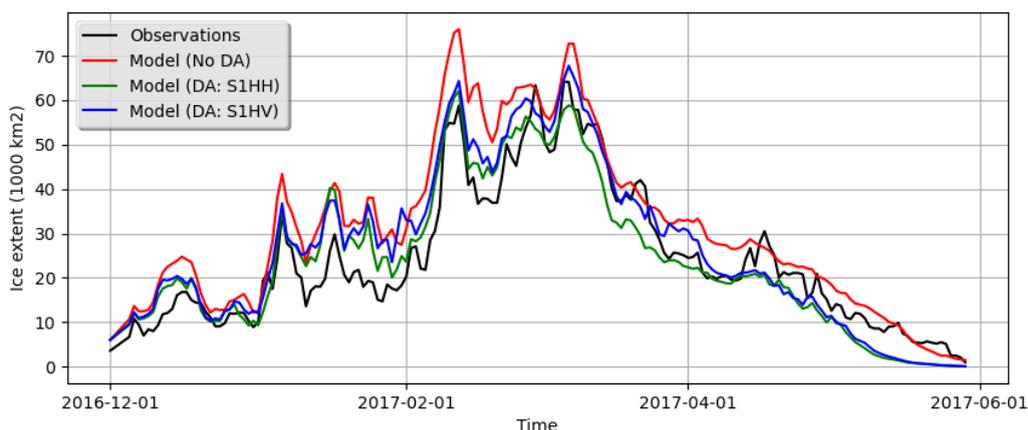


Figure 7: Improvements of modeled sea ice extent by assimilation of Sentinel-1 C-SAR data, using HH polarization and HV polarization.

### 3.2 Test case A: Maximum ice extent in February, 2017

This test case includes the period February 9-15, 2017, and includes the date of maximum ice extent during the selected winter season, 2017-02-12. Figure 8 below shows the ice chart and the observed backscatter from Sentinel-1 C-SAR with polarizations HH and HV, respectively. For this particular date, only the Bothnian Sea is covered by the satellite observations, but the ice edge in the northern Bothnian Sea can be seen clearly in both polarizations in (b) and (c) and compares well with the ice chart in (a), but the HV polarization in (b) indicates some ice south of the ice edge as well, which were also present in the both polarizations 11 hours earlier (not shown) but are not present in the ice charts.

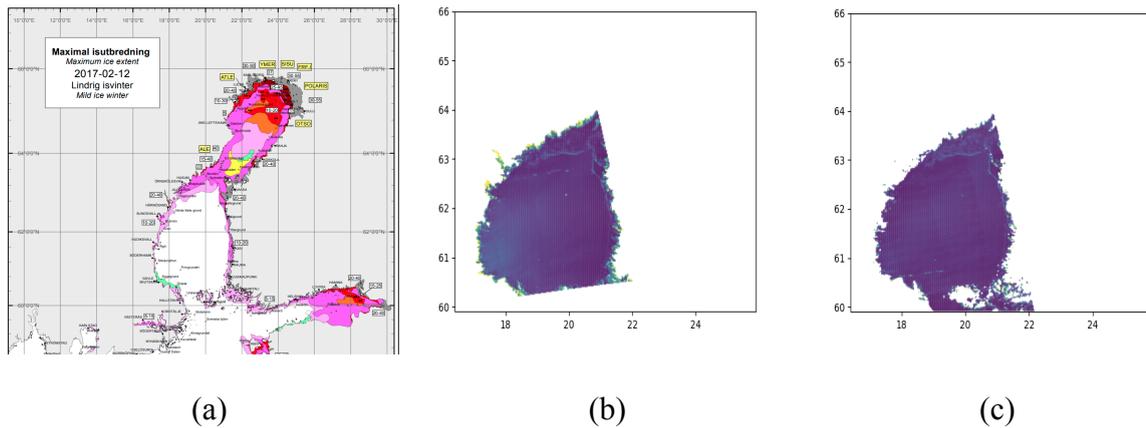


Figure 8. Ice situation according to (a) the ice chart, (b) Sentinel-1 C-SAR HH polarization, and (c) Sentinel-1 C-SAR HV polarization. Both Sentinel images were from 2017-02-11 16 UTC.

Figure 9 below shows modeled ice concentration from the same date, for the case of a free run (no data assimilation; left panel) and with data assimilation (middle and right panels) of Sentinel data. It is clear that assimilation improves the ice extent (as we saw in section 3.1 above), and that for this date the data assimilation has managed to keep the amount of ice close to the western and eastern coasts of Bothnian Sea closer to observations compared to the free run in the left panel. Also, in panel (b) we see that data assimilation of the HH polarization data gives rise to a small amount of ice south of the ice edge in the northern Bothnian Sea, not unlike what we saw in the Sentinel data in Figure 8 (b) above. However, although assimilation of Sentinel data gives a positive impact, there is still too much ice in some locations, and actually too little ice off the coast in the northwestern Bothnian Sea.

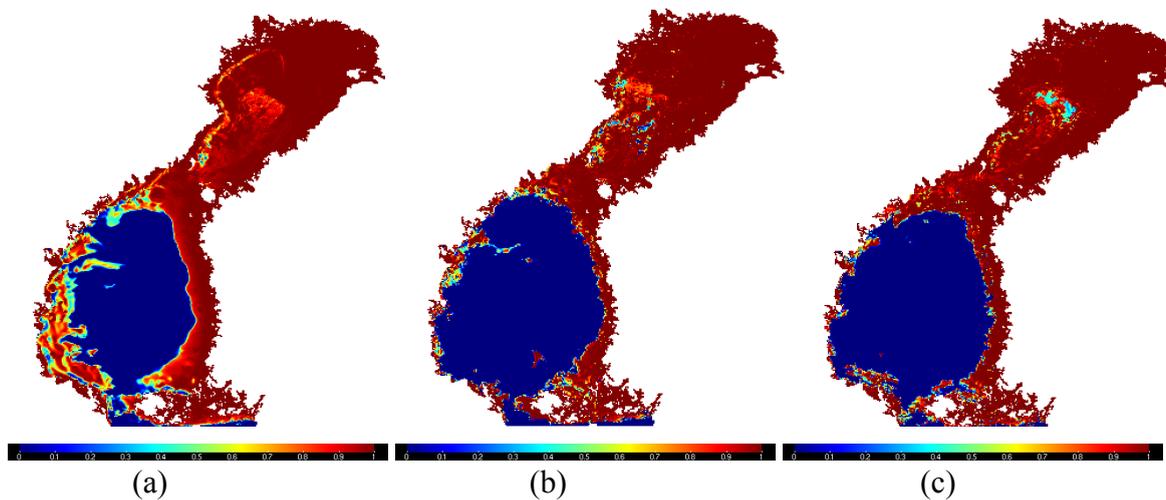


Figure 9. Ice situation according to the model (a) without data assimilation, (b) data assimilation of Sentinel-1 C-SAR HH polarization data, and (c) data assimilation of Sentinel-1 C-SAR HV polarization data.

### 3.3 Test case B: Formation of ice ridges in March, 2017

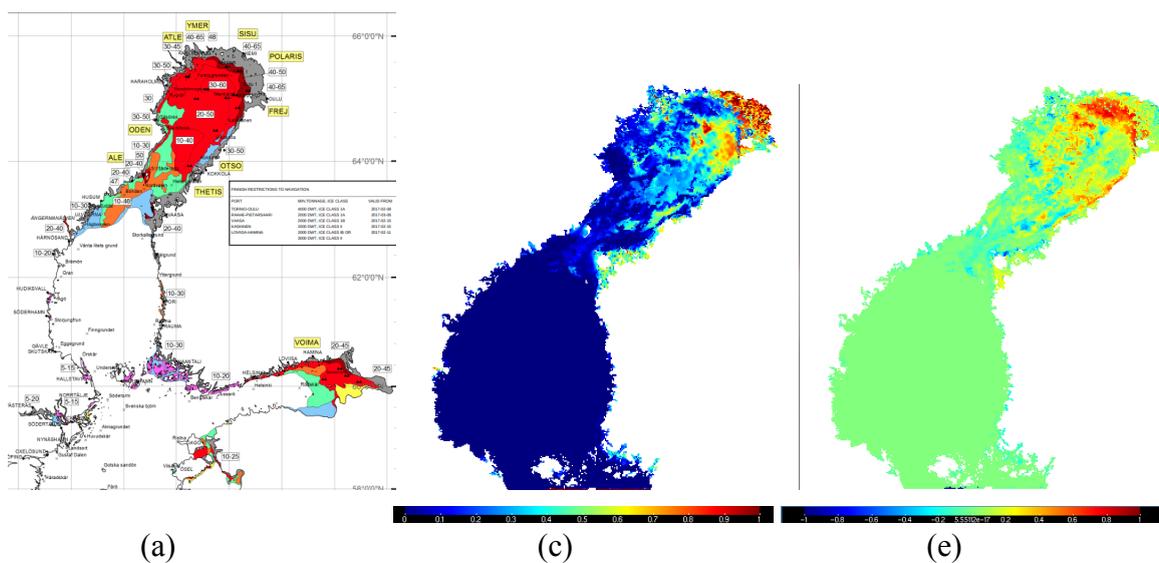
A low-pressure system generated winds that moved ice and created ridges the days before

March 17, 2017. In this test case we will see how well the model is able to simulate this process, while assimilating Sentinel-1 C-SAR HH polarization data which gave best results above.

Figure 10 (a) shows the ice chart on March 15 (before the wind event) and Figure 10 (b) shows the ice chart on March 17 (after the event). A strong southwestern wind pushes the ice toward the fast ice zone in the northeastern Bothnian Bay, while the ice sheet leaves the coast of Sweden, more or less, and leaves large areas of open water on March 17.

Panels (c) and (d) in Figure 10 show the modeled sea ice thickness before and after the wind event, respectively. We see that the model is able to simulate the ice drift rather well and the opening of the ice off the Swedish coast, but not quite as much as the ice charts suggest.

Panel (e) shows the difference in sea ice thickness (panel (d) minus panel (c)), which makes it easy to see the increase outside of the fast ice zone in northeastern Bothnian Bay. The increase is almost one meter in thickness. Panel (f) shows the difference in ice concentration instead, in which we clearly can see the regions where the ice left the Swedish coast, and also the Northern Quark region.



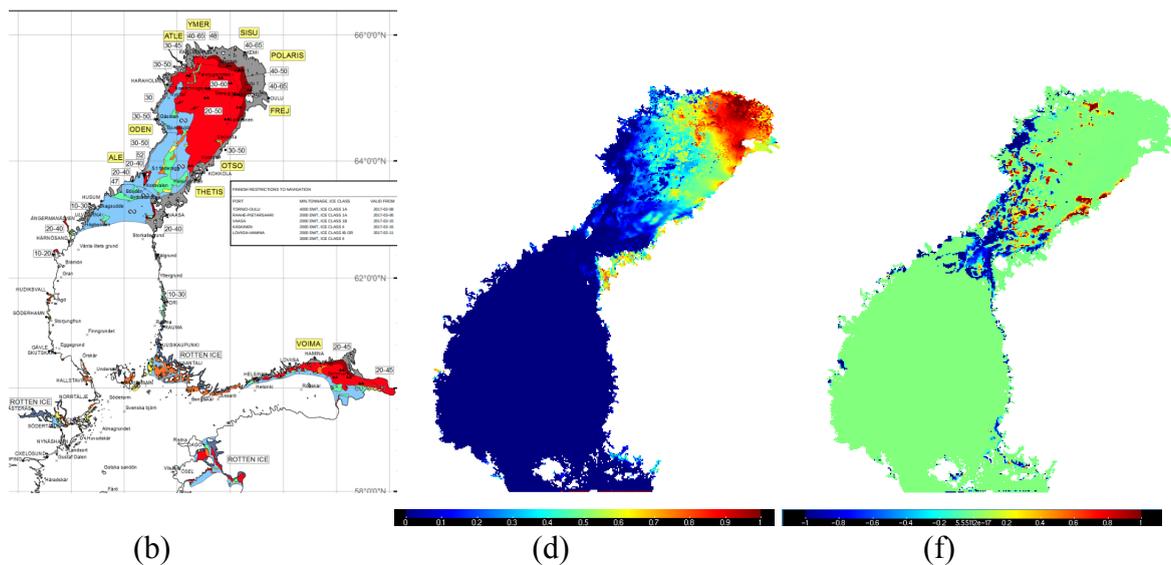


Figure 10. The ice situation before and after a low-pressure event, forcing ice to leave great parts of the Swedish coast while pushing against the fast ice zone in the northeastern Bothnian Bay, generating ice ridges. Panels (a) and (b) show ice charts on 2017-03-15 and 2017-03-17, respectively. Panels (c) and (d) show modeled ice thickness before and after the wind event, respectively, while assimilating Sentinel-1 C-SAR HH polarization data. Panels (e) and (f) show the difference in sea ice thickness and sea ice concentration, respectively.

### 3.4 Test case C: The breakup of a brash ice barrier in April

During the days around April 10, 2017, a brash ice barrier was present near the ice edge in the Bothnian Bay, but was subsequently broken up by northern winds. In this test case we will see how well the model is able to simulate this process, again while assimilating Sentinel-1 C-SAR HH polarization data.

Panels (a) and (b) in Figure 11 below show the operational ice charts for the dates 2017-04-10 and 2017-04-12, respectively. Panel (a) indicates where the brash ice barrier is located along the ice edge in the Bothnian Bay that runs almost in a straight line from southeast to northwest. Panel (b) shows the situation two days later, when the brash ice barrier has disappeared through the action of the northern winds which pushed the ice southward.

Panels (c) and (d) show the simulated sea ice thickness before and after the brash barrier has disappeared. However, none of the images show any signs of a brash ice barrier in the model, which seems to imply that the NEMO ice model is not very good at predicting these features, not even using data assimilation of satellite data.

Panels (e) and (f) in Figure 11 show the model differences in sea ice thickness and sea ice concentration, respectively. The positive values (yellow/orange) near the location of the ice edge merely indicates that the ice edge is moving southwards during these days and does not shed any more light on the brash ice barrier, simply because it does not seem to be well simulated by the model. However, the general breakup of the ice sheet seems to be well simulated, indicated by the southward motion of the ice as well as the thinning of the ice in northern Bothnian Bay.

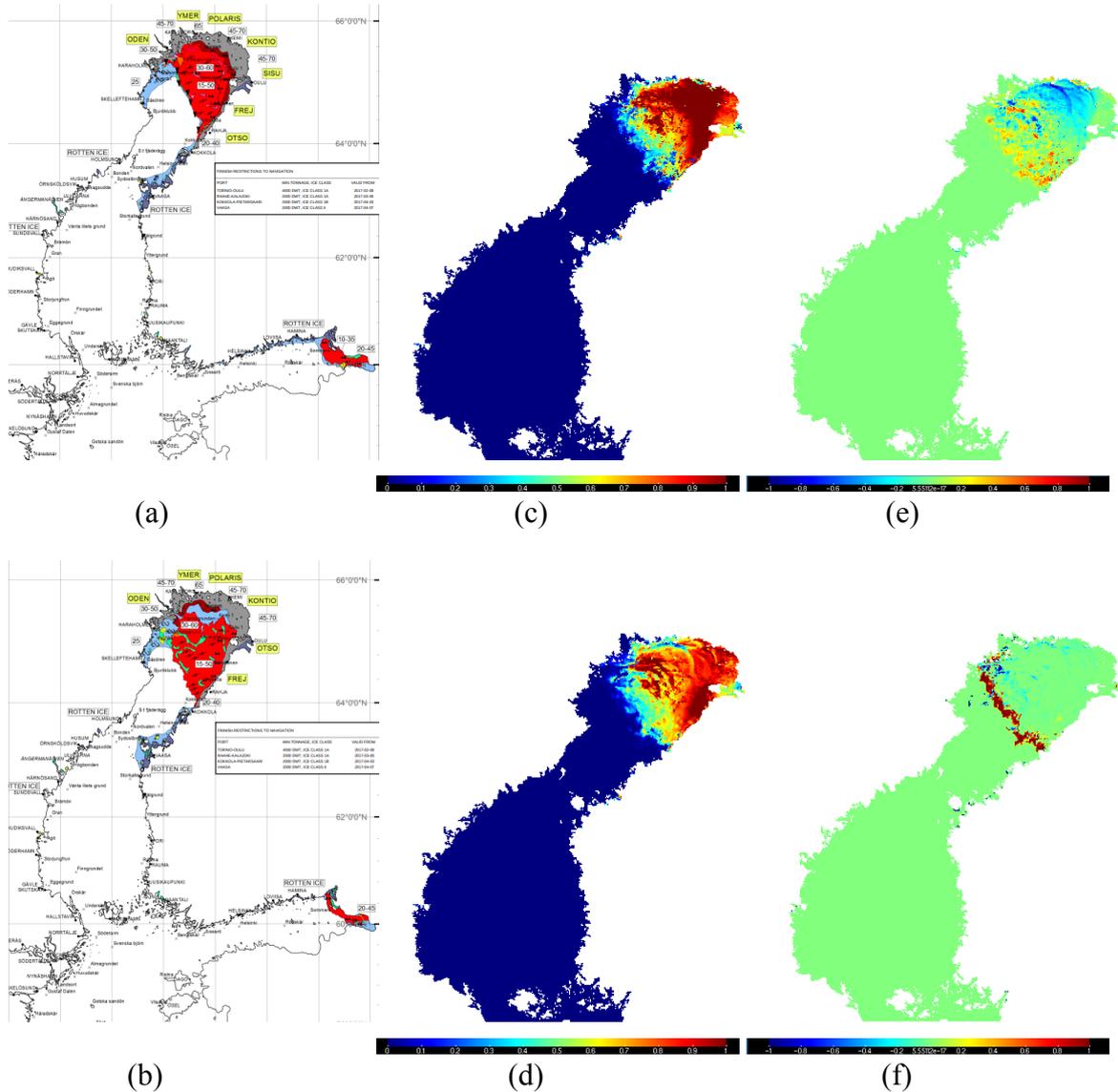


Figure 11. The ice situation before and after the disappearance of a brash ice barrier along the ice edge in the Bothnian Bay, seemingly caused by northern winds that start to break up the whole ice field in Bothnian Bay. Panels (a) and (b) show ice charts on 2017-04-10 and 2017-04-12, respectively. Panels (c) and (d) show modeled ice thickness before and after the wind event, respectively, while assimilating Sentinel-1 C-SAR HH polarization data. Panels (e) and (f) show the difference in sea ice thickness and sea ice concentration, respectively.

## 4. Discussion and Future Outlook

The quality control of the observations is crucial. Improved removal of thermal and texture noise in the Sentinel-1 data was achieved by applying methods developed at Nansen Environmental and Remote Sensing Center. Unfortunately we could not use the AMSR2 data as planned. The large footprints of the non-weather-dependent channels left us with too few cases to be considered as pure snow/ice. In order to move forward one needs to look at more ice seasons and probably also extend the modeling to include snow covered coastal regions.

In the observation operator we replaced KNNR (from Baltic FIRE-1) with XGBoost. XGBoost is well known for its general applicability to non-deep-learning problems. It outperformed KNNR both in terms of speed and quality. The observation operator based on XGBoost showed dependence on NEMO variables in case of snow/ice surfaces. Both HH and HV polarization are sensitive to ice thickness, ice age, snow thickness and the satellite azimuth and zenith angle where the latter is of greatest importance. The model for the open water does not depend on any NEMO variables but is still necessary in order to obtain a combined estimate for a general situation and in this respect reflects the ice concentration which is described by NEMO.

As an alternative to the machine learning approach, physical modeling of the backscatter from snow/ice surfaces with the SMRT model was improved by describing the snow and ice pack with three instead of one layer which was used in Baltic FIRE-1. Still, the results turned out inferior to those obtained with the empirical XGBoost model. The high bias in SMRT results could point to an error in how the model was set up and this should be investigated further.

To improve the results on the ice-ocean model side, it could prove beneficial to try to reduce the cold bias in free simulations with the model (i.e. without data assimilation), which today generates slightly too much sea ice. One simple fix could be to include the Langmuir circulation in the unperturbed ensemble member, which should increase the mixed layer depth and thus make the model cool slower during winter.

Further, regarding the ensemble forecasting system, we saw that the ensemble spread seems large enough to encompass future states, but that maybe the spread was a bit too large. Future investigations should look into this, and see if this is a general problem with our new ensemble forecasting system, and see if the reason may be the perturbations in the atmospheric forcing or the perturbations in the physical model parameters. Another possible reason may be that it is only the unperturbed ensemble member that has been updated using data assimilation, whereas the other members have been running without data assimilation. The reason for that was partly to make the system quicker to run and make experiments on, but also to make sure the spread was large enough. In future experiments we may test applying data assimilation on all ensemble members, to see if that improves the results.

Another aspect of the ensemble forecasting system is the number of ensemble members that we use. Since the data assimilation result relies heavily on the cross covariances between the Sentinel-1 back scatter and the near-surface model variables, it is of great importance that the ensemble is good enough to be used for this kind of data assimilation. In this project we have used 20 ensemble members, which is already a bit heavy to run on a computer, but it may be of importance to increase the number to 40 or 80, for example by introducing perturbations to the initial conditions.

One technical problem we ran into regarding the data assimilation code NOVA, was that it requires a lot of internal computer memory to run. The reason is that it needs to store lots of model data simultaneously (e.g. the model state for each ensemble member for 24 different time steps), and the amount partly depends on the number of observations. Since the

Sentinel data were co-located on the model grid, the number of observations could easily amount to tens of thousands which made it impossible to run the code. To solve the problem temporarily, it was decided to resample the observations on a smaller resolution, 11 km instead of 1.85 km, which was used in the results presented above. This could be the reason why the brash ice barrier did not show up in the simulations above, it was simply smeared out when we resampled the observations to lower resolution before the data assimilation step. So, in a future version of NOVA it will hopefully be possible to use the observations at higher resolution, which may improve results regarding brash ice barriers.

Finally, in theory at least, the results should improve when we assimilate both HH and HV polarization data. However, the results were much worse than when we assimilated only one polarization type at a time, which is a bit puzzling. So this is also something to look into in the future, especially in the ongoing project CAISA that we mentioned above in section 2.

To summarize, we have

- implemented the possibility to assimilate ice-related variables into NEMO-4.0, which will soon be used operationally at SMHI;
- shown that a simple 20-member ice ensemble forecasting system can be used to generate an ensemble good enough for ice data assimilation;
- shown that assimilation of ice variables from operational ice charts greatly improves the results in our ice-ocean model;
- shown that it is possible to improve ice-ocean forecasts most significantly also by assimilating Sentinel-1 C-SAR backscatter data directly, using 4D EnVar in the NOVA data assimilation system;
- found that the best observation operator was obtained using a machine-learning technique called XGBoost, rather than the physically based Radiative Transfer Model SMRT.

## References

Axell, L.B. (2002). Wind-driven internal waves and Langmuir circulations in a numerical ocean model of the southern Baltic Sea, *J. Geophys. Res.*, 107 (C11), 3204, doi: 10.1029/2001JC000922.

Chen, T. and Guestrin, C. (2016): XGBoost. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, doi:10.1145/2939672.2939785

Hordoir, R., Axell, L., Löptien, U., Dietze, H., and Kuznetsov, I. (2015). Influence of sea level rise on the dynamics of salt in-flows in the Baltic Sea, *J. Geophys. Res.*, 120, 653–6668, <https://doi.org/10.1002/2014JC010642>.

Hordoir, R., Axell, L., Höglund, A., Dieterich, C., Fransner, F., Gröger, M., Liu, Y., Pemberton, P., Schimanke, S., Andersson, H., Ljungemyr, P., Nygren, P., Falahat, S., Nord, A., Jönsson, A., Lake, I., Döös, K., Hieronymus, M., Dietze, H., Löptien, U., Kuznetsov, I., Westerlund, A., Tuomi, L., and Haapala, J. (2019). Nemo-Nordic 1.0: a NEMO-based ocean model for the Baltic and North seas – research and operational applications, *Geosci. Model Dev.*, 12, 363–386, <https://doi.org/10.5194/gmd-12-363-2019>.

Kärnä, T. et al. (2021). *Geosci. Model Dev.*, 14, 5731–5749, <https://doi.org/10.5194/gmd-14-5731-2021>.

Madec, G., Bourdallé-Badie, R., Chanut, J., Clementi, E., Coward, A., Ethé, C., Iovino, D., Lea, D., Lévy, C., Lovato, T., Martin, N., Masson, S., Mocavero, S., Rousset, C., Storkey, D., Vancoppenolle, M., Müller, S., Nurser, G., Bell, M., and Samson, G. (2019). NEMO ocean engine, Zenodo, <https://doi.org/10.5281/zenodo.3878122>.

Park, Jeong-Won; Korosov, Anton; Babiker, Mohamed; Sandven, Stein; and Won, Joong-Sun (2018): Efficient noise removal of Sentinel-1 TOPSAR cross-polarization channel, *IEEE Transactions on Geoscience and Remote Sensing*, 56(3), 1555-1565, doi:10.1109/TGRS.2017.2765248

Pemberton, P., Löptien, U., Hordoir, R., Höglund, A., Schimanke, S., Axell, L., and Haapala, J. (2017). Sea-ice evaluation of NEMO-Nordic 1.0: a NEMO-LIM3.6-based ocean–sea-ice model setup for the North Sea and Baltic Sea, *Geosci. Model Dev.*, 10, 3105–3123, <https://doi.org/10.5194/gmd-10-3105-2017>, 2017.