https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

# Barents-2.5km v2.0: An operational data-assimilative coupled ocean and sea ice ensemble prediction model for the Barents Sea and Svalbard

Johannes Röhrs[1], Yvonne Gusdal[1], Edel Rikardsen[1], Marina Duran Moro[1], Jostein Brændshøi[2], Nils Melsom Kristensen[1], Sindre Fritzner[3], Keguang Wang[1], Ann Kristin Sperrevik[1], Martina Idžanović[1], Thomas Lavergne[1], Jens Debernard[1], and Kai H. Christensen[1]

[1]Norwegian Meteorological Institute, Henrik Mohns Plass 1, 0371 Oslo, Norway
[2]Norwegian Defence Research Establishment, Instituttveien 20, 2007 Kjeller, Norway
[3]UiT The Arctic University of Norway

**Correspondence:** J. Röhrs (johannes.rohrs@met.no)

**Abstract.** An operational ocean and sea ice forecast model, *Barents-2.5*, is implemented at MET Norway for short-term forecasting at the coast off Northern Norway, the Barents Sea, and waters around Svalbard. Primary forecast parameters are the sea ice concentration (SIC), sea surface temperature (SST), and ocean currents. The model is also a substantial input for drift modeling of pollutants, ice berg, and in search-and-rescue pertinent applications in the Arctic domain. *Barents-2.5* has recently been upgraded to include an Ensemble Prediction System with 24 daily realizations of the model state. SIC, SST and *in-situ* hydrography are constrained through the Ensemble Kalman Filter (EnKF) data assimilation scheme executed in daily forecast cycles with lead time up to 66 hours. While the ocean circulation is not directly constrained by assimilation of ocean currents, the model ensemble represents the given uncertainty in the short-term current field by retaining the current state for each member throughout forecast cycles. Here we present the model setup and a validation in terms of SIC, SST and in-situ hydrography. The performance of the ensemble to represent the models uncertainty, and the performance of the EnKF to constrain the model state are discussed, in addition to the model's forecast capabilities for SIC and SST.

## 1 Introduction

Rapid changes in water mass distribution and seasonal sea ice cover are underway in the Barents Sea and the areas around Svalbard (Lind et al., 2018). Alongside, we see changes in human activities such as shipping, oil exploration, and fisheries. The ability to provide efficient emergency services in this region for search-and-rescue, accidental oil spills, ship drift, etc, is much reduced compared to mainland Norway due to the vast distances and challenging environmental conditions. Contingency models for decision support are based on short term forecast models for the weather, surface waves, ocean circulation and sea ice distribution, and increased predictability in these models translate directly into better emergency preparedness (Röhrs et al., 2023). Of particular interest are upper ocean transport processes (e.g. Strand et al., 2017, 2021) and energy exchanges across the air-ice-sea interfaces.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

In addition, Arctic Ocean ecosystems are vulnerable to change, and the question how projected long term decline in sea ice cover, and the associated "Atlantification" of the Barents Sea (Asbjørnsen et al., 2020), is impacting key species at different trophic levels are important topics for investigation (Ingvaldsen et al., 2021). Hence accurate models of upper ocean transport processes are also important for understanding connectivity of fish stocks and other physical-biological couplings (Röhrs et al., 2014; Strand et al., 2017)

A high resolution (2.5 km) numerical weather prediction (NWP) system, Arome-Arctic (MET-AA), for the Barents Sea and the areas around Svalbard was established in 2015 (Müller et al., 2017). This development paved the way for a coupled ocean and sea ice forecast model that operates on the same horizontal grid as the NWP forcing. Here we present this coupled ocean-sea ice forecasting system, the *Barents-2.5* model.

Ocean forecasting in this region is challenging for several reasons: The first challenge is that the hydrography is poorly observed and in-situ data are scarce compared to other regions. We have frequent passes of polar orbiting satellites, but since the skies seldom are clear we have relatively few high resolution observations of sea surface temperature (SST) from infrared sensors. We do have abundant medium-resolution observations of the sea ice cover from passive microwave imagery (Spreen et al., 2008), and frequent high-resolution images from synthetic aperture radars (SAR). Combined, these sensor technologies provide sub-daily coverage in the region.

A second challenge in high latitudes is that the atmospheric low pressure systems can be small but very intense (Noer et al., 2011; Furevik et al., 2015). Correctly describing the temporal and spatial development of polar lows in the NWP system is difficult, although progress is being made with improvements in physics (Batrak and Müller, 2019) and data assimilation (DA) methodology (Hallerstig et al., 2021; Mile et al., 2022). We have, however, significant knowledge gaps when it comes to rapid changes in upper ocean conditions and sea ice cover due to complex air-sea interactions including waves, and many surface layer processes are quite crudely parameterised in the models, if at all.

A third challenge is that the inertial period and the dominant tidal period (M2) are very similar at the latitudes considered here (e.g. Röhrs and Christensen, 2015). Hence it is difficult to disentangle the transient Ekman dynamics from the more predictable tidal circulation. A general lack of long observation time series of ocean currents or mean sea level imply that we cannot easily remove the tidal component of observed drift velocities to isolate effects of air-sea interactions on the upper ocean transport.

Ocean forecasting in high latitudes must encompass the uncertainties that result from the above challenges. As widely applied in NWP, Ensemble Prediction Systems (EPS) can directly quantify the uncertainty in forecasts. Ensemble prediction has not yet been widely used in regional ocean modeling, but recent works have shown that models may predict uncertainty correctly if the underlying ocean model exhibits statistical skill (Jacobs et al., 2021). Our goal has been to implement an operational EPS for the Barents Sea that assimilates the available observations in this region. The uncertainty provided by the EPS may be used directly in forecast applications, e.g. to quantify the reliability of a prediction during decision support. In addition, the ensemble spread can be used to estimate model background errors, which is needed in DA schemes such as the Ensemble Kalman Filter (EnKF) applied in this work (Evensen, 1994; Sakov and Oke, 2008).

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

This paper describes the operational setup of the *Barents-2.5* forecast model. In section 2, we provide details on the configuration and setup of the ocean and ice model, including the coupling scheme and forcing data. Section 3 describes the DA scheme and the assimilated observations. Section 4 documents how the model is operationalized, which includes details on the setup as an EPS. Model performance is assessed in section 5, focusing on validation of SST and sea ice concentration (SIC) as central state variables. Finally, in section 6 we discuss the model's capability as forecast tool and inform about our plans for future model development.

## 2 Model physics and configuration

### 2.1 Ocean circulation model

*Barents-2.5* is built on the Regional Ocean Modeling System (ROMS) version 3.7, which applies a topography following coordinate system in the vertical (Shchepetkin and McWilliams, 2005) and a curvilinear horizontal grid. ROMS solves the Boussinesq primitive equations. The modeled state variables are temperature, salinity, surface elevation, and horizontal current velocities. The setup in *Barents-2.5* includes a second order turbulence closure scheme with turbulent kinetic energy and a generic length scale as state variables (Warner et al., 2005).

The model domain and bathymetry is shown in Fig. 1. The model resolution is approximately 2.5 km, which varies slightly throughout the domain consisting of 737 x 947 grid points. The bathymetry is based on the GEBCO global data set and interpolated to the model grid. Minimum depth is set to 10m. The bathymetry is smoothed to reduce pressure gradient errors as required in ROMS. Additionally, the coastline has been modified to match numeric grid point criteria by ROMS (i.e., every water grid point must have at least two adjacent open boundaries).

#### 2.1.1 Discretisation and advection schemes

The applied vertical discretization uses a stretched, topography-following coordinate system. *Barents-2.5* has 42 layers with emphasis on the upper mixed layer using the ROMS-specific parameters shown in Tab. 1. This choice of values results in an upper layer of approx. $0.2 - 1.2$ m and maintains an increased resolution in the upper 100 m.

ROMS uses split explicit time stepping, solving slow baroclinic modes separately from fast barotropic modes. *Barents-2.5* uses 90 second outer time steps for the baroclinic mode, solving the 3D momentum equations, and 30 inner time steps (i.e., 3 seconds) for the solution of barotropic 2D momentum such as tides.

Momentum and tracers are advected using a 3rd order upwind scheme in the horizontal and a 4th order centered scheme in the vertical. Turbulent kinetic energy and length scale are advected vertically and horizontally using a 4th order centered scheme.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

### 2.1.2 Sub-scale processes

Vertical mixing is modeled using a second order scheme for turbulent kinetic energy (TKE) and a generic length scale (GLS). In this setup, which is recommended in Umlauf and Burchard (2005) and Warner et al. (2005), the GLS has non-physical units of $m^2/s^2 \cdot m^{-0.67}$. Parameters for the turbulence scheme are documented Tab. 2. It is noted that turbulence dissipation rate and turbulent length scale may be calculated from TKE and GLS according to eqs. 14 and 15 in Warner et al. (2005). The `CANUTO_A` stability function is chosen for the diffusion of momentum and tracers (Canuto et al., 2001; Warner et al., 2005).

The boundary condition for TKE at the surface is based on the model of Craig and Banner (1994) using a flux condition, wherein the energy flux at the surface is proportional to the air-side friction velocity with a factor of 100. Surface roughness is set by the wind stress using a Charnok constant of 1400. Buoyancy and shear are horizontally smoothed using the `N2S2_HORAVG` pre-compiler option. The background vertical diffusivity is set to $10^{-6}\,m^2/s$ for tracers, $10^{-5}\,m^2/s$ for momentum and $5 \cdot 10^{-6}\,m^2/s$ for TKE and GLS. Furthermore, the pre-compiler options `RI_SPLINES`, `SPLINES_VVISC` and `SPLINES_VDIFF` are used.

Harmonic horizontal diffusion of tracers is applied using a diffusivity of $2\,m^2/s$. Towards the boundaries, the diffusivity is increased up to 10-fold within a distance of 40 grid points using an arctan shaped smooth transition.

Explicit harmonic horizontal diffusion is applied using a viscosity of $50\,m^2/s$ only in the sponge zone of 30 grid points, increasing from zero to a viscosity of $100\,m^2/s$ using an arctangent shaped smooth transition. A harmonic horizontal diffusivity for TKE and GLS is set to $0.1\,m^2/s$

Tracers are mixed along surfaces of constant geopotential, while momentum is mixed along the bottom topography following coordinate surfaces. Quadratic bottom friction is applied using a drag coefficient of 0.003 where the water depth is greater than 100 m. In shallower regions, the bottom drag coefficient is increased up to 0.009 for the shallowest parts with a water depth of 10 m with linear transition as a function of water depth. The bottom drag is limited such that the current cannot reverse sign using the pre-compiler option `LIMIT_BSTRESS` in ROMS. Limiting the bottom drag in such fashion is useful to avoid numeric instabilities in very shallow waters during strong storm surges. A full list of all applied pre-compiler options for ROMS, static files such as grid files, and run-time options are provided in a code repository for *Barents-2.5* [1].

## 2.2 Sea ice model

The sea ice model used in *Barents-2.5* is the Los Alamos sea ice model (CICE), version 5.1 (Hunke et al., 2017). CICE describes both dynamic and thermodynamic processes, using a elastic-viscous-plastic (EVP) rheology (Hunke and Dukowicz, 1997). The model solves the evolution of the sea ice state, which is described by the ice thickness distribution (ITD) function (Thorndike et al., 1975; Hibler III, 1980):

$$\frac{\partial g}{\partial t} = -\nabla \cdot (g\mathbf{u}) - \frac{\partial (fg)}{\partial h} + \psi \tag{1}$$

[1] https://doi.org/10.5281/zenodo.7607191

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

where $g(\mathbf{x},h,t)dh$ is defined as the fractional area covered by ice in the thickness range $(h,h+dh)$ at a given time $t$ and location $\mathbf{x}=(x,y)$. The ice velocity $\mathbf{u}$ is calculated from the sea ice momentum equation that account for air and water drag, Coriolis force, sea surface tilt, and the divergence of internal ice stress. The evolution of internal stress is described by the EVP rheology, with the ice strength reformulated according to (Rothrock, 1975), and the advection using the incremental remapping scheme (Lipscomb and Hunke, 2004). The subgrid sea ice deformation and the redistribution of various ice categories follow (Rothrock, 1975) , with a modified expression for the participation function (Lipscomb et al., 2007).

The sea ice thermodynamic growth and melting $f$ is determined by solving the one-dimensional vertical heat balance equations for each ice thickness category and snow. The sea ice heat balance equation is solved with the mushy-layer scheme that also accounts for the evolution of sea ice salinity (Turner et al., 2013). The upper snow and ice boundary is assumed to be in balance with short- and longwave radiation and the sensible, latent and conductive heat fluxes when the surface is below freezing point. When the surface temperature reaches the melting point, it is held constant and the extra heat is used to melt the snow and ice surface. The bottom water-ice interface is assumed to be at thermodynamic balance, such that growing or melting results from the divergence between ice conductive heat flux and the under-ice oceanic heat flux.

An example of the CICE input parameter file, including all choices of numeric schemes and input parameters, is provided in a code repository for *Barents-2.5* [2].

### 2.3 Ocean-ice coupling

The ROMS-CICE coupling utilizes the Model Coupling Toolkit (MCT,Larson et al. (2005)) for intermodel exchange of state variables and fluxes, as implemented in the *metroms* framework[3] (Debernard et al., 2021)). The coupling was briefly explained in Naughten et al. (2017) and Naughten et al. (2018), and in more detail by Duarte et al. (2022). The underlying philosophy behind the coupling is that surface fluxes are calculated in the component with most information about the surface, utilizing required information from the other component. Finally, the heat, water and momentum fluxes are passed conservatively back to the other component. The same fluxes are transferred as in Duarte et al. (2022). In *metroms*, the coupling is based on the principle of 'levitated' ice, so there is no actual exchange of mass between the ocean and the ice. Freshwater and salt fluxes from the ice model are converted to a virtual salt flux before they are used in ocean model. In this 'mass less' state, the ice does not displace water, and it is only seen by the ocean as a source of surface fluxes responding to the present ocean state. The ice and ocean models are run concurrently, with exchange of information every baroclinic ocean time step (90 s, section 2.1.1). The information used by each model is therefore time-lagged compared with its own state.

### 2.4 Boundary conditions and river input

The boundary conditions for *Barents-2.5* are provided by TOPAZ version 4, a coupled ocean and sea ice DA system based on the *HYCOM* ocean model and a EnKF DA scheme. The model system is configured for the North Atlantic and Arctic Ocean with a horizontal resolution of 12-16 km (Xie et al., 2017). It provides daily averages of temperature, salinity, sea surface

[2]https://doi.org/10.5281/zenodo.7607191
[3]https://github.com/metno/metroms.git

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

elevation, and ocean current velocities for the ocean component. For the sea ice component it provides daily averages of SIC, sea ice thickness, first year ice age, snow depth, and ice velocity. The inverse barometric effect due to the local atmospheric pressure is added to the sea surface elevation boundary conditions because TOPAZ does not include the barotropic signal of atmospheric pressure.

The numeric schemes for boundary conditions used in ROMS are given in Tab. 3, these differ for the various state variables. A sponge zone with up to 10-fold increased horizontal tracer diffusivities and viscosity is implemented within 40 grid points from the boundary. Nudging of passive tracers towards the boundary fields from TOPAZ is imposed within the sponge zone. 2D momentum anomalies are radiated out of the model domain using the tangential phase speed of the barotropic signal, using ROMS' `RADIATION_2D` pre-compiler option. Details about the implementation of boundary conditions in CICE are given in (Duarte et al., 2022).

Point sources for river influx along the coast in the model domain are shown in Fig. 1. At each of the 318 river locations, daily values for temperature, salinity and mass flux are specified. Climatological values for rivers on the Svalbard archipelago are used, while river data on mainland Norway originates from daily estimates provided by the Norwegian Water Resources and Energy Directorate.

### 2.5 Atmospheric forcing

Surface forcing in *Barents-2.5* is provided by ensemble forecasts from the integrated forecast system at European Centre for Medium Weather Forecasts (ECMWF-ENS), which has a horizontal resolution of about 18km. Hourly surface fields of wind speed, air temperature, humidity, rain fall, and cloud cover are used by the bulk flux module of ROMS to calculate surface fluxes as upper boundary condition in the ocean. CICE uses air temperature, humidity, density, precipitation rate and winds to calculate surface stress, heat fluxes and snow aggregation on sea ice.

*Barents-2.5* is configured as EPS with 24 members, wherein 20 members are forced with random members from ECMWF-ENS, and four members are forced by MET-AA (Müller et al., 2017). MET-AA is nested into ECMWF's high-resolution forecasts and covers the identical domain as *Barents-2.5* with the same horizontal resolution of 2.5 km, hence providing higher resolution forcing than ECMWF-ENS.

Tidal forcing is provided as amplitudes and phases of the 10 major tidal constituents in the model domain (Tab. 4). These are obtained from the TPXO global inverse barotropic model (Egbert and Erofeeva, 2002) and imposed on velocities and free surface elevation. The tidal signal is also added to velocities and surface elevation during the processing of boundary data.

## 3 Data assimilation scheme

The model state in *Barents-2.5* is constraint by a DA scheme, the EnKF, that moderates the combination of a model background state and observations. Central in this moderation are the model error, estimated by the spread of the EPS, and the observation errors. A new analysis is obtained on a daily basis, which is aimed to reduce the errors in the model state compared to the avail-

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

able observations. The assimilated variables are SST, SIC and in-situ hydrography, i.e. salinity and temperature observations (Tab. 5).

### 3.1 Ensemble Kalman Filter

The assimilation method used is the deterministic version of the EnKF (Evensen, 1994; Burgers et al., 1998). The same DA scheme was also applied in a preceeding setup of *Barents-2.5* that assimilates sea ice variables but not the hydrography (Fritzner et al., 2019, 2020). The EnKF is a sequential ensemble based assimilation method that has been used for various geophysical applications in recent years (Houtekamer and Zhang, 2016). The Deterministic Ensemble Kalman Filter (DEnKF) (Sakov and Oke, 2008) do not require perturbation of observations to maintain ensemble spread as is required by the original formulation of the EnKF. Perturbing observations introduces additional sampling error in the analysis, which for applications with few ensemble members might be a significant contribution (Sakov and Oke, 2008; Whitaker and Hamill, 2002).

The standard analysis equation solved by the EnKF is given by (Evensen, 2003):

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{P}_b\mathbf{H}^T\left(\mathbf{H}\mathbf{P}_b\mathbf{H}^T + \mathbf{R}\right)^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}_b), \tag{2}$$

where $\mathbf{x}_a \in \mathbb{R}^{n\times N}$ is the analysis vector representing the updated variables after assimilation, $\mathbf{x}_b \in \mathbb{R}^{n\times N}$ the model background, and $\mathbf{y} \in \mathbb{R}^{m\times N}$ is the observation vector. $N$ is the number of ensemble members, $n$ the number of variables multiplied by the number of spatial grid points in our model, $m$ the total number of observations, $\mathbf{R} \in \mathbb{R}^{m\times m}$ the observation covariance matrix, and $\mathbf{H} \in \mathbb{R}^{m\times n}$ is the observation operator. The key property of the EnKF is that the background error covariance matrix $\mathbf{P_b}$ $\in \mathbb{R}^{n\times n}$, providing the model uncertainty and covariance between variables and locations, is estimated as the variance of the ensemble of background states,

$$\mathbf{P_b} = \overline{((\mathbf{x}_b - \overline{\mathbf{x}_b})(\mathbf{x}_b - \overline{\mathbf{x}_b})^T)}. \tag{3}$$

In the equation above, the overbars signify the average operator.

The implementation of the EnKF in *Barents-2.5* is accomplished using the *enkf-c* software package (Sakov, 2014). The EnKF is executed offline, i.e., in between daily forecast cycles, thereby providing updated initial conditions $\mathbf{x_a}$ to ROMS and CICE at 00 UTC each day. Observations of the previous 24 hours $\mathbf{y}$ are synchronously evaluated and compared with the model state $\mathbf{x_b}$ at time 00 UTC, hence assuming that observed variables retain validity for 24 hours.

The *enkf-c* software package allows configuration of the DA scheme in terms of a spread reduction factor $\sigma_r$ and a localisation radius $r_{loc}$ for each observation type (Tab. 5), a global moderation factor $K = 1.5$, and an inflation factor that is set to 1.05 for all model variables except for SIC using 1.1 in order to increase ensemble spread in this variable. In addition, the impact of observations have been moderated by means of scaling the observation errors.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

### 3.2 In-situ Observations

In situ observations of salinity and temperature are routinely collected from the in-situ thematic assembly centre of the Copernicus Marine Environment Monitoring Service[4]. The data set consists of observations from a variety of platforms, such as drifting buoys, profiling floats, moorings, CTD casts and thermosalinographic data from monitoring cruises and the Ships-of-Opportunity Programme. Additional in situ observations are retrieved from the Global Telecommunication System (GTS), and from MET Norway's observation database[5]. The observations are run through some basic quality control procedures prior to assimilation, and duplicate observations between the different sources are removed.

### 3.3 SST observations

Level 2P satellite SST from the infrared sensors AVHRR (Metop satellites), VIIRS (Suomi NPP and NOAA-20 satellites), and SLSTR (Sentinel-3 mission) are used for assimilation. A bias correction scheme is applied to ensure consistency between the different SST products. The SLSTR measurements are used as a reference because this instrument uses a dual view technique which gives more accurate SST measurements. The bias correction scheme is described in more detail in Iversen et al. (2022).

All in-situ and SST observations are processed using the python toolbox pyromsobs [6]. When more than one observation of a given state variable is available within the same grid cell at the same time, they are replaced with a so-called super-observation, which is a mean of the available observations.

### 3.4 SIC observations

Observations of SIC are computed from microwave brightness temperature observations from the Advanced Microwave Scanning Radiometer-2 (AMSR2) sensor on board of the Global Change Observation Mission – Water "Shizuku" (GCOM-W1) satellite. The SIC algorithm *SIRANO* is an evolution of that described in Lavergne et al. (2019) and involves dynamic tuning of the algorithm coefficients, atmospheric correction of the brightness temperatures using a radiative transfer model, and per-pixel uncertainties. Two SICs are computed independently and then combined before they are assimilated: first a low resolution, low uncertainty SIC from the 18.7 and 36.5 GHz imagery channels, then a higher resolution, higher uncertainty SIC from the 89.0 GHz imagery channels. The two SICs are then combined using a pan sharpening method to preserve the higher resolution and low noise characteristics from each. Because the nominal spatial resolution of the 89.0 GHz imagery of AMSR2 is 3x5 km, the resulting SIC fields have a spatial resolution slightly coarser than the 2.5 km grid spacing of the model.

In the operational setup, SICs are computed for each incoming satellite orbit (every 100 min) and then combined and projected to cover the geographic domain of the model in 12 hourly fields, including uncertainties. Because of the orbit of the GCOM-W1 satellite, the domain is observed in the early morning (from 00 UTC to 04 UTC) which is adequate for an assimilation at 00 UTC model time.

[4]http://marine.copernicus.eu
[5]http://frost.met.no
[6]https://github.com/metno/pyromsobs

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

## 4 Operational implementation

*Barents-2.5* is executed four times daily, spreading the 24 ensemble members into 4 sets of 6 members – as visualized in Fig. 3. While a new analysis by the EnKF is only computed at 00 UTC, each bulletin time benefits from the updated NWP forecasts. In practise, the 00 UTC model run is executed with a six hour real-time delay to allow for enough time processing incoming observations. The atmospheric forcing from MET-AA is avaliable with 4 hours delay. The 06, 12 and 18 UTC model runs are executed with a 4 hour delay as these only need to wait for the updated atmospheric forecast. The first member in each set is forced by the most recent MET-AA forecast, the remaining by ECMWF-ENS members.

### 4.1 Ensemble prediction system

At each of the four bulletin times each day, six ensemble members are executed with a 66 hour forecast range. Each member is initialized by the state of the same member from the previous day, with individual analysis increments from the the EnKF DA scheme. This amounts to 24 unique members. The forecast runs after 06, 12 and 18 UTC require a spin up run from the analysis time at 00 UTC.

A conscious choice for the design of the ensemble forecast system is to initialize the ensemble runs from the model states of previous day forecasts, instead of using perturbed states of the identical analysis which is more common in NWP. The latter approach yields estimate of model uncertainty when the initial condition is constrained by observation networks and most of the uncertainty arises internally within the forecast cycle.

The *Barents-2.5* EPS, however, retains the state of previous forecast runs in order to preserve sufficient ensemble spread. The EnKF moderates the ensemble spread. Such an approach is beneficial because the initial state of the ocean circulation is poorly observed. At the same time, the mesoscale circulation exhibit features with time scales larger than the assimilation window and forecast range. Therefore, small perturbation of the initial state would not yield ensemble spread in the forecast that covers the actual uncertainty in mesoscale circulation.

To represent the actual model uncertainties in the circulation field, we initialize the EPS forecast with largely varying initial conditions in the mesoscale circulation. At the same time, the EnKF reduces the ensemble spread for variables that are observed well, i.e., SST and SIC. The ensemble spread for these variables is evaluated in section 5.3.

### 4.2 Operationalisation

*Barents-2.5* is operational in its current setup since September 2021, tagged as version 2.0. The model is implemented at MET Norway and part of Norway's national ocean and weather forecast system. Timely triggering of the model components, as well as pre-processing steps for observation data, forcing and boundary conditions, are managed through ECMWF's scheduling software package *ecflow*[7]. Backup measures are in place to secure a reliable operation of the *Barents-2.5* setup to deal with exceptions from normal operation:

[7]https://github.com/ecmwf/ecflow, accessed 2023-01-29

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

- Computing and data storage facilities are setup in two physically independent data rooms at MET Norway to mitigate the risk of hardware failures and planned hardware maintenance. All model pre- and post-processing is setup in both rooms and there is a continuous synchronization between the data rooms. Model runs can be executed in either of the two data rooms.
- For pre-processing of forcing and boundary data, alternative data sources have been setup, e.g. older forecast cycles from the ECMWF-ENS atmospheric forcing and TOPAZ model can be accessed to produce necessary input data.
- Initial states can be obtained from the previous two days forecast cycles of Barents-2.5, in case of lost forecast cycles during temporary outages.
- The forecast runs can be executed without analysis increments by the EnKF.

These exceptions are dealt automatically through triggering in *ecflow*. While the first three type of exceptions rarely occur, a failure of the EnKF update is more common. This may occur either due to unexpected deviations during observation data processing, or due to poorly performing statistic representations in the EnKF algorithm. In such cases, which occur about twice a month, the model continues without DA increments.

Any type of deviations alert a triggering for the model developers at MET Norway, such that repeated failures are added to the exceptions that can be dealt with by the system within the *ecflow* scheduling software. Model developers are also alerted 24/7 if the forecasts from Barents-2.5 are delayed more than a few hours.

## 5 Model performance

The general performance of the *Barents-2.5* model is evaluated by validating SST and SIC as key model state variables. Both are frequently observed from satellite platforms and are of central importance as forecast parameters. Model performance is evaluated i) in terms of validation metrics close to model analysis time (0-24hours), ii) in terms of skill in the forecast range up to 66 hours, and iii) ensemble spread. The spread controls the EnKF data assimilation scheme, and is in addition used to predict forecast uncertainty.

Validation of SST is based in comparison with Sentinel-3 SLSTR swaths. Multiple swaths in the model domain may be available each day, but only cloud free daylight conditions provide SST measurements. A composite map of comparing SST observations with model values available in a 24 h period are shown in Fig. 4. As the coverage of SST observations varies, the validation of SST is not uniform in season and location, providing generally more data during spring and summer. A validation of SIC against the SIRANO passive microwave SIC (ref. Sec. 3.4) is shown in Fig. 5. For SIC, a validation of the full model domain is possible during most days.

### 5.1 Validation of model analysis

SST is validated separately for the regions defined in Fig. 6, and displayed during the course of one year in Fig. 7. Only SST observations from the Sentinel 3 satellite have been used in validations in order to avoid bias differences of various sensor

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


platforms in the validation. These same data are also included for the assimilation in the analysis time step at 00:00 each day, but the assimilation occurs always in the analysis cycle after the direct validation, as the model value is obtained from the exact satellite retrieval time in forecast run closest to the analysis time. Hence, the observations are considered as independent data for validation.

SIC in forecast ranges of 0-24, 24-48 and 48-66 hours is validated by comparison with the SIRANO passive microwave SIC product. The validation is carried out for data that is assimilated in the following forecast cycle, hence these are relatively independent from the current forecast cycle. Daily validation metrics during the course of a year are shown in Fig. 8.

As a second reference for validation of SIC, a comparison with subjective ice charts is provided in Fig. 8a. The ice charts are based on a manual production by ice analysts, who use synthetic aperture radar and visible band satellite imagery to locate the ice edge and ice classes on a daily basis from Monday to Friday. The ice charts provide a higher-resolution assessment of ice conditions that is independent from the passive microwave data, and exhibits different systematic errors.

The largest anomalies in the model errors for SIC are related to the model spinup and failures in the data assimilation scheme. During the beginning of the year in consideration, deviations occurred more frequent and during the second half of the displayed period very few failures of the assimilation scheme occurred, maintaining low model error. Likewise, the model performance deteriorates within a few days when no data assimilation is applied for subsequent cycles, as seen for two periods in March and May 2022 in Fig. 7 and 8.

### 5.2 Forecast skill

Predictive skill in the model forecast is assessed by computing model validation metrics as function of forecast lead time, e.g. the validation metrics can be obtained by comparison with a value of older forecast cycles. For SIC, we see that the model error is larger for higher lead times, indicating skill in the model analysis (see Fig. 8). In the periods with no data assimilation, the model error grows rapidly and is distinguished in various lead times. Hence we conclude that skill in ice forecast is provided by assimilation of SIC.

For SST, we provide a mapping of model error as function of model lead time as averages over periods of three months. (Fig. 9). Time averaging of the model errors for many forecast cycles is necessary in this case, because individual days rely on the availability of sparse SST observations, covering different regions from day to day as shown for one particular day in Fig. 4.

SST validation metrics are compared with a reference forecast (black line). As reference we are using the observation itself, commonly referred to as *persistence forecast*. By this definition, the error of the persistence forecast is nil at zero lead time. The error of larger lead times are given by comparison with past observations, whenever there is spatial overlap between observation times.

Skillful model predictions may be defined in two ways:

- The model has skill for lead times where the error of the model forecast is lower than the reference forecast.
- The model analysis and forecast is thought to contain skill if the model error grows with lead time.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

The model error may be defined as root mean square error (RMSE) or mean error (ME). In order to specifically evaluate the capability of the forecast model to predict SST changes, we define an additional forecast that is the sum of past observation and the trend of the model forecast (dashed graphs in Fig. 9). This forecast is also nil at zero lead time, and validation of it is only possible when there is spatial overlap between observation times.

Skill in model SST is present for forecast ranges beyond 12 hours in spring and summer time, extending beyond the forecast range of 66 h. For the autumn and winter time, we see that the model error is on par with the errors in the persistence forecasts beyond 12 h lead time. In general, combining the model trend with past observations performs better at intermediate lead times, but the valuable information content in observations vanishes for longer lead times. In this range, dynamical changes that are resolved by the model provide superior value in predicting the ocean state conditions.

Skill in SIC analysis is concluded from the fact that the short forecast range (0-24h) yields lower errors than the longer range forecasts. However, the modeled SIC field exhibit large systematic errors in terms of RMSE and ME (Fig. 8). This uncertainty is also present in the difference between the two observation types used for validation, e.g. passive microwave SIC and ice charts. Some forecast cycles show skill compared to the persistence forecast (not shown), but most cycles do not provide a better forecast than using the last observation as reference. Model skill is present in situations with rapid movements in the sea ice edge that are driven by strong winds. Re-freezing of the sea ice cover in the winter are not as skillful in the model predictions as summer melting and dynamic response to wind forcing.

### 5.3 Ensemble spread

The ensemble of model states is an integral part of the EnKF data assimilation scheme: it provides model error covariances to weigh observational impact with model state in order to compute model increments in each analysis cycle through Eq. 3. The objective of the ensemble spread is to reflect the actual uncertainty of the model, which is possible to assess by comparing observations with the ensemble.

A common means to verify the ensemble spread are rank histograms, which rank each observation value into the range of model values at the same time and location. If each ensemble member has a truly equal likelihood of occurrence in the validation, the observation ranks should be uniformly distributed across the ensemble members. The rank histogram for SST is shown in Fig 10. We see a fair ensemble spread for this variable across the center part of the rank histogram, but note a skewed offset that reflects a negative overall bias in model SST. We also acknowledge that the most extreme SSTs are not represented in the model.

The spread in SIC is evaluated using a reliability diagram, as shown in Fig. 11. The reliability diagram is a more adequate means to assess non-gaussian distributed variables, e.g. SIC, and it allows to directly compare ranges in ice classes that are commonly used in ice charting. Ice classes are defined by thresholds in SIC, i.e. very close drift ice ($0.9 < SIC < 1.0$), close drift ice ($0.7 < SIC < 0.9$), open drift ice ($0.4 < SIC < 0.7$), very open drift ice ($0.1 < SIC < 0.4$), open water ($0.01 < SIC < 0.1$) and ice free ($0 < SIC < 0.01$). The latter three classes are used as one ($0 < 0.1$) in this analysis because they are indistinguishable in the SIRANO SIC that is used for validation.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


The reliability diagram compares the forecast probability for cumulative ranges of ice classes with the occurrence in the observations for the same situations. An ideal ensemble would match the forecast probability with the validation, e.g. if 50% of ensemble members show a SIC below 0.4, then 50% of those cases should validate with SIC < 0.4. Deviations from the diagonal in Fig. 11 are due to i) imperfect ensemble spread and ii) model biases. We can also distinguish the skill for various forecast probabilities.

In Fig. 11a we assess the forecast probabilities for occurrences of open water up to a given ice class. The opposite approach (Fig. 11b) assesses forecast probabilities of high SICs down to a certain ice class minimum. Both figures indicate an offset that represents a positive SIC bias of the model, e.g. forecast probabilities for ice tend to be too low from the ensemble. In addition, the reversed S-shape in both figures indicates too low ensemble spread.

Parts of the the lack in ensemble spread owes to observation errors, however not all and there is a clear lack in variability in SIC between ensemble members. The highest forecast probabilities of open water, e.g. 80% chance for the occurrence of $0 < SIC < 0.4$ is nevertheless addressed fairly well by the model. Also low chances of occurrences for high SIC, e.g. in the range $0.4 < SIC < 1$, is predicted accurately by the model. Addressing the lack of sufficient EPS spread, forecast probabilities would benefit from empirical scaling to match the observed occurrence frequencies.

### 5.4 Analysis increments

The EnKF provides analysis increments for each ensemble member. The objective is to reduce deviations from the observation. In this section we show some characteristics of the analysis increments for an exemplary day in order to illustrate the functioning of the DA scheme. Similar model increments are calculated for the analysis in each forecast cycle, applied daily at 00 UTC.

Figure 12 shows forecast increments of SST, surface salinity and SIC on the 2022-12-15. The differences result from the observations that were available in the preceeding 24 hours. The most substantial increments in model hydrography occur where both SST observations present (Fig 2) and spread in SST is large (Fig. 12d). In-situ observations provide rather localized increments in surface fields. SIC increments are most pronounced in the marginal ice zone where the ensemble has large spread (Fig. 12d).

To confirm that the DA increments move the model state towards the observed ocean state, we present a binned scatter histogram of observed vs. model values for SST and SIC in Fig. 13. The correlation for SST is slightly improved for the analysis compared to the background, starting at already good fit in the background model state. In the case of SIC, which generally has larger errors than SST, we see more radical updates to the model state in each cycle.

Ensemble spread is an essential characteristic of the EPS, and sufficient ensemble spread is required by the EnKF to yield appropriate model increments. The spread for SST, salinity and SIC during one analysis cycle is shown in Fig. 12d-f. Spread in SST and surface salinity is present throughout the model domain, however pronounced in certain areas as for example below the sea ice cover. This indicates pronounced uncertainty in the model state, e.g. the present hydrography below the sea ice cover is not well known and the model is expected to exhibit larger errors in these areas.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

The ensemble spread is reduced in each analysis such that model state variables converge towards the observations. The spread increment is shown in Fig. 12g-i. For SST, we see a substantial reduction in spread that is associated with coverage by SST observations during that analysis cycle. All shown variables experience a spread reduction in the marginal ice zone, owing to the impact of SIC observations in this area. There is little modification to ensemble spread below the dense sea ice cover and in other places where no observations are present during the analysis cycle.

## 6 Discussion

The *Barents-2.5* ocean and ice forecast model provide short-term predictions on a daily basis with a forecast range up until 66 hours. This fully operational model suite has been built from various components and code repositories; the dynamic ocean component is based on ROMS v3.7 (Shchepetkin and McWilliams, 2005), the dynamic sea ice compenent is CICE 5.1 (Hunke et al., 2017), model coupling is achieved by MCT, DA is handled through *enkf-c* (Sakov, 2014), and operational scheduling is implemented in *ecflow*, including the maintanence of the ensemble. In addition, *Barents-2.5* involves static data such as the GEBCO bathymetry and TPXO tidal forcing, and a constantly updated stream of forcing data. These are the atmospheric forcing provided by ECWMF-ENS and MET-AA, lateral boundary conditions provided by *TOPAZ*, and river influx. Finally, observations of SIC, SST and in-situ hydrography are assimilated. All components combined constituate the *Barents-2.5* model, and it is worth to stress that model capabilities and the presented validation are a consequence of the interactions between all the components.

### 6.1 Forecast capabilities

Forecast skill and validation metrics have been presented for SST and SIC. These are central variables describing the state of ocean and sea ice conditions that are frequently observed by satellite imagery. Forecast skill and ensemble spread for surface currents are the subject of a separate investigation [8].

*Barents-2.5* shows predictive skill in temperature forecasts (Fig. 9) compared to the persistence reference by Sentinel-3 SST imagery. Forecast skill is better during spring and summer time when SST observations are more abundant and solar radiation drives rapid changes in SST. Mismatches in the modeled ice edge dominate the SST errors in the marginal ice zone, therefore all regions that contain ice have been excluded in the consideration for SST validation. SST fields are directly needed in forecast applications, e.g. to warn about icing on ships (Samuelsen, 2018) and to provide surface forcing in NWP models (e.g. Müller et al., 2017).

Skillful SST predictions in the forecast range are a consequence from the models capability to represent realistic air-sea fluxes, vertical mixing and water mass transport (Janssen, 2012). A requirement for low errors at the analysis time is the successful constraint by observations. Without a functioning DA scheme, the model state tends to drift away from temperature fields, owing to compounding of systematic model errors. Such model state drift is particularly evident for SIC, during the periods when no DA is applied (ref. Fig.8). SST validation is not comparable between daily cycles, but similar drift is expected.

[8] Idzanovic et al., Forecast uncertainty and ensemble spread in surface currents from a regional ocean model, manuscript in preparation

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

For ice concentration, we see that the model state looses skill within few days without DA. Deterioration of model skill within few days is generally expected when DA is halted in regional ocean models (e.g. Moore et al., 2011).

Validation of SST indicates larger errors during summer time (Fig. 7) but also a higher skill compared to the persistence forecast (Fig. 9). Variability in SST during the summer time is generally higher, resulting from the availability of direct sunlight to heat up the surface. Most of the model region is in high latitudes with midnight sun. The diurnal effect may dominate in early spring time, while in the summer the solar heating is largely controlled by cloud cover. Skilful SST predictions by the ocean model thus require accurate cloud cover and atmospheric radiative fluxes, highlighting the role of the used NWP models. Due to the strong coupling with the atmosphere, SST in summer time changes more rapidly than in the winter season. *Barents-2.5* is forced by hourly fields from Arome-Arctic, giving SST forecasts an advantage over the persistence forecast that is shown as reference in Fig. 9.

While errors in SIC are variable throughout the validated year (Fig. 8), we see a persistent advantage of short lead times compared to longer lead times. The model analysis is hence skillful, but the remaining systematic model errors make it difficult to use predicted ice concentration fields directly in forecasting applications. Use of ice charts based on remote sensing alongside model forecast remains necessary. Predictions of ice drift in ocean-ice forecast models have been shown to be skillful (Schweiger and Zhang, 2015), resulting from the skilfull NWP and realistic description of sea ice rheology. The freezing and melting along the ice edge, however, depends on a number of thermodynamic processes (Fichefet and Maqueda, 1997), and these are critically sensitive to water mass properties and higher resolution details in the atmospheric forcing (Kusahara et al., 2017).

### 6.2 Configuration of the ensemble and assimilation schemes

The information provided by the EPS is useful in two ways. Firstly, ensemble spread guides the DA scheme. Secondly, the ensemble allows to assess model uncertainty in the forecast range. Spread in SST is satisfactory (Fig. 10), with only slight underestimation of extremes in SST. Spread in ice concentration is too low (Fig. 11), but the model shows a capability to predict the uncertainty in ice forecasts. Difficulties in representing ice concentration spread in similar EnKF setups have previously been reported by Lisæter et al. (2003).

The initialisation of ensemble in each forecast cycle is based on an approach where the individual members retain their identity. DA increments are provided for each member. The spread is hence reduced in each cycle but the differences that resulted from previous cycles remain. The major difference between the ensemble members does therefore result from subsequent diverging history of the members. Each member is also forced by different atmosphere ensemble members, but the most important spread in mesoscale ocean circulation stems from the history of the EPS.

The full EPS and EnKF system requires time to spin up the ensemble and DA increments, due to the system dependence on its own history of ensemble spread. Initialisation of the EPS took place for during March 2021, and the EnKF was activated in September 2021. About 3 months were required for a sufficient spin-up of the full DA system.

The choice of a 24 hour analysis cycle is based on the practical need to issue daily forecasts. The assimilation cycle length is a compromise between amount of observations, the models ability to maintain spread, and need for rapidly updated forecast

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


cycles. The number of ensemble members is limited by computational resources. The EnKF DA scheme relies on a statistical description of possible model states compared to the available observations, and therefore the ensemble needs to cover the degrees of freedom for the ocean state within a localisation radius (Tab. 5), which is maintained in the current setup using 24 ensembles.

### 6.3 Integration with contingency models

The *Barents-2.5* model forecasts are routinely used in trajectory models that serve as decision support tools in emergency response situations. The OpenDrift framework (Dagestad et al., 2018) encompasses oil spill transport, ice berg drift forecasts, and leeway drift (e.g. person-in-water, large floating objects) applications. All of these models require timely computation of trajectory forecasts that are based on surface current, wave, and wind forecasts. Among these forcing variables, surface current forecasts exhibit by far the largest uncertainty (Dagestad and Röhrs, 2019). OpenDrift make use of the *Barents-2.5* EPS by issuing ensembles of trajectory forecasts that is based on the ensemble spread form *Barents-2.5* currents and respective wind forcing, allowing to assess the uncertainty of trajectory simulations in a given situation. In practise, the difference in trajectory forecasts from various ensemble members varies from case to case. The *Barents-2.5* model is useful to distinguish cases of low and high uncertainty in drift applications (de Aguiar et al., 2023).

### 6.4 Model development history and outlook

An earlier implementation of the *Barents-2.5* has been operational since March 2019, with DA scheme for sea ice concentrations using a combined optimal interpolation and nudging scheme (Fritzner et al., 2018). The most notable refinements in the herein described model version 2.0 marks the implementation as EPS with 24 members and the introduction of the EnKF DA scheme for both ice and ocean variables. Forecast skill of SST are a direct consequence of these developments.

A number of future developments are planned to further improve forecast capabilities. Asynchronous assimilation of ice concentration data, i.e. swath data, are in a development stage. SST assimilation will also benefit from asynchronous assimilation. In asynchronous DA, each satellite swath is compare to the model field at the respective retrieval time instead of applying observations at a fixed time of the day. This may lead to a reduction of the representation error in observations and hence introduce a tighter constraint on the model state.

The lack of ensemble spread in SIC is a major weakness in the current model setup. A more explicit perturbation of ice variables is planned to introduce larger spread in SIC, using either a perturbation of the initial conditions or physics perturbation in the sea ice component, e.g. by variation of empirical parameters in CICE. Constraint of mesoscale ocean current fields beyond assimilation of SST is planned by assimilation of sea level anomaly from satellite altimetry and HF radar currents.

A stronger coupling with NWP models and wave prediction models is foreseen as soon as such coupling is shown to improve forecasts. Atmosphere models require SSTs as lower boundary conditions, and these may benefit from skillful predictions of SST by *Barents-2.5*, provided that the coupling will not introduce artifacts through poorly defined feedback loops. At present, the NWP model Arome-Arctic, which provides surface forcing for *Barents-2.5*, uses static SST fields based on a daily satellite retrieval product. Coupling can either be implemented as fully coupled online, or by using SST forecasts of the previous

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

forecast cycle. A coupling with wave prediction models is foreseen, in which the wave prediction model provides surface wave dissipation as boundary condition for turbulent kinetic energy of the ocean (Breivik et al., 2015, e.g.). Such coupling has been shown to improve SST predictions in the global ocean forecast model at the ECMWF (Janssen, 2012).

*Code and data availability.* A source code for the coupled ROMS/CICE ocean and sea ice model is available through https://doi.org/10.5281/zenodo.5067164 tagged as version 0.4.1 of the METROMS repository. Specific configuration and grid files for *Barents-2.5* are archived at https://doi.org/10.5281/zenodo.7607191. The source code for the enkf-c v.2.9.9 is obtained from https://github.com/sakov/enkf-c.git, commit *7eea4d8* as of Jul 8, 2021. The source code for the ecflow scheduling software is available at https://github.com/ecmwf/ecflow.

Archived data from the operational model runs of *Barents-2.5* is disseminated on https://thredds.met.no/thredds/fou-hi/barents_eps.html. Observations of satellite SST and in-situ hydrography are retrieved from CMEMS https://doi.org/10.48670/moi-00036. Observations of sea ice concentration are available through https://thredds.met.no/thredds/osisaf/osisaf_seaiceconc.html .

*Author contributions.* Model setup and implementaion: JR, JB, ER, SF, NMK, KW, JD, KHC. Observation preparation: AKS, TL. Configuration of DA scheme: SF, AKS, JR, MDM. Model validation: YG, ER, MDM, MI, AKS. Manuscript preparation: all authors. Project management: JR, KHC

*Competing interests.* There are no competing interests associated with this work.

*Acknowledgements.* We acknowledge funding by the Research Council of Norway, 237906 (CIRFA), 302917 (SIRANO) and 300329 (Ecopulse), and the Norwegian FRAM Flagship program project SUDARCO (No. 2551323). This study has been conducted using E.U. Copernicus Marine Service Information, i.e. model data as boundary condition and observation data used in assimilation. We acknowledge the R&D contribution of ESA CCI Sea Ice and EUMETSAT OSI SAF projects to the preparation of the regional SIC product.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

## References

Asbjørnsen, H., Årthun, M., Skagseth, O., and Eldevik, T.: Mechanisms Underlying Recent Arctic Atlantification, Geophysical Research Letters, 47, e2020GL088 036, https://doi.org/10.1029/2020GL088036, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL088036, 2020.

Batrak, Y. and Müller, M.: On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice, Nature Communications, 10, 4170, https://doi.org/10.1038/s41467-019-11975-3, number: 1 Publisher: Nature Publishing Group, 2019.

Breivik, O., Mogensen, K., Bidlot, J.-R., Balmaseda, M. A., and Janssen, P. A. E. M.: Surface wave effects in the NEMO ocean model: Forced and coupled experiments, Journal of Geophysical Research: Oceans, 120, 2973–2992, https://doi.org/10.1002/2014JC010565, _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2014JC010565, 2015.

Burgers, G., Leeuwen, P. J. v., and Evensen, G.: Analysis Scheme in the Ensemble Kalman Filter, Monthly Weather Review, 126, 1719–1724, https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2, publisher: American Meteorological Society Section: Monthly Weather Review, 1998.

Canuto, V. M., Howard, A., Cheng, Y., and Dubovikov, M. S.: Ocean Turbulence. Part I: One-Point Closure Model - Momentum and Heat Vertical Diffusivities, Journal of Physical Oceanography, 31, 1413–1426, https://doi.org/10.1175/1520-0485(2001)031<1413:OTPIOP>2.0.CO;2, 2001.

Craig, P. D. and Banner, M. L.: Modeling Wave-Enhanced Turbulence in the Ocean Surface Layer, Journal of Physical Oceanography, 24, 2546–2559, https://doi.org/10.1175/1520-0485(1994)024<2546:MWETIT>2.0.CO;2, 1994.

Dagestad, K.-F. and Röhrs, J.: Prediction of ocean surface trajectories using satellite derived vs. modeled ocean currents, Remote Sensing of Environment, 223, 130–142, https://doi.org/10.1016/j.rse.2019.01.001, 2019.

Dagestad, K.-F., Röhrs, J., Breivik, O., and Ådlandsvik, B.: OpenDrift v1.0: a generic framework for trajectory modelling, Geosci. Model Dev., 11, 1405–1420, https://doi.org/10.5194/gmd-11-1405-2018, 2018.

de Aguiar, V., Röhrs, J., Johansson, M., and Eltoft, T.: Assessing Ocean Ensemble Drift Predictions by Comparison with Observed Oil Slicks, in review for Frontiers in Marine Science, 2023.

Debernard, J., Kristensen, N. M., Maartensson, S., Wang, K., Hedstrom, K., Brændshøi, J., and Szapiro, N.: metno/metroms: Version 0.4.1, https://doi.org/10.5281/zenodo.5067164, language: eng, 2021.

Duarte, P., Brændshøi, J., Shcherbin, D., Barras, P., Albretsen, J., Gusdal, Y., Szapiro, N., Martinsen, A., Samuelsen, A., Wang, K., and Debernard, J. B.: Implementation and evaluation of open boundary conditions for sea ice in a regional coupled ocean (ROMS) and sea ice (CICE) modeling system, Geoscientific Model Development, 15, 4373–4392, https://doi.org/10.5194/gmd-15-4373-2022, publisher: Copernicus GmbH, 2022.

Egbert, G. D. and Erofeeva, S. Y.: Efficient inverse modeling of barotropic ocean tides, J. Atmos. Ocean. Tech., 19, 183–204, https://doi.org/10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2, 2002.

Evensen, G.: Inverse methods and data assimilation in nonlinear ocean models, Physica D: Nonlinear Phenomena, 77, 108–129, https://doi.org/10.1016/0167-2789(94)90130-9, 1994.

Evensen, G.: The Ensemble Kalman Filter: theoretical formulation and practical implementation, Ocean Dynamics, 53, 343–367, https://doi.org/10.1007/s10236-003-0036-9, 2003.

Fichefet, T. and Maqueda, M. A. M.: Sensitivity of a global sea ice model to the treatment of ice thermodynamics and dynamics, Journal of Geophysical Research: Oceans, 102, 12 609–12 646, https://doi.org/10.1029/97JC00480, 1997.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023




Fritzner, S., Graversen, R., and Christensen, K. H.: Assessment of High-Resolution Dynamical and Machine Learning Models for Prediction of Sea Ice Concentration in a Regional Application, Journal of Geophysical Research: Oceans, 125, e2020JC016 277, https://doi.org/10.1029/2020JC016277, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020JC016277, 2020.

Fritzner, S. M., Graversen, R. G., Wang, K., and Christensen, K. H.: Comparison between a multi-variate nudging method and the ensemble Kalman filter for sea-ice data assimilation, Journal of Glaciology, 64, 387–396, https://doi.org/10.1017/jog.2018.33, publisher: Cambridge University Press, 2018.

Fritzner, S. M., Graversen, R., Christensen, K. H., Rostosky, P., and Wang, K.: Impact of assimilating sea ice concentration, sea ice thickness and snow depth in a coupled ocean-sea ice modelling system, Ocean Sci., 13, 491–509, https://doi.org/10.5194/tc-13-491-2019, accepted: 2019-10-16T11:59:18Z Publisher: European Geosciences Union, 2019.

Furevik, B. R., Schyberg, H., Noer, G., Tveter, F., and Röhrs, J.: ASAR and ASCAT in Polar Low Situations, J Atmos Ocean Tech, 32, 783–792, https://doi.org/10.1175/JTECH-D-14-00154.1, 2015.

Hallerstig, M., Magnusson, L., Kolstad, E. W., and Mayer, S.: How grid-spacing and convection representation affected the wind speed forecasts of four polar lows, Quarterly Journal of the Royal Meteorological Society, 147, 150–165, https://doi.org/10.1002/qj.3911, 2021.

Hibler III, W. D.: Modeling a Variable Thickness Sea Ice Cover, Monthly Weather Review, 108, 1943–1973, https://doi.org/10.1175/1520-0493(1980)108<1943:MAVTSI>2.0.CO;2, 1980.

Houtekamer, P. L. and Zhang, F.: Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation, Monthly Weather Review, 144, 4489–4532, https://doi.org/10.1175/MWR-D-15-0440.1, publisher: American Meteorological Society Section: Monthly Weather Review, 2016.

Hunke, E., Lipscomb, W., Jones, P., Turner, A., Jeffery, N., and Elliott, S.: CICE, The Los Alamos Sea Ice Model, Tech. Rep. CICE; 005315WKSTN00, Los Alamos National Lab. (LANL), Los Alamos, NM (United States), https://www.osti.gov/biblio/1364126, 2017.

Hunke, E. C. and Dukowicz, J. K.: An Elastic–Viscous–Plastic Model for Sea Ice Dynamics, Journal of Physical Oceanography, 27, 1849–1867, https://doi.org/10.1175/1520-0485(1997)027<1849:AEVPMF>2.0.CO;2, 1997.

Ingvaldsen, R. B., Assmann, K. M., Primicerio, R., Fossheim, M., Polyakov, I. V., and Dolgov, A. V.: Physical manifestations and ecological implications of Arctic Atlantification, Nature Reviews Earth & Environment, 2, 874–889, https://doi.org/10.1038/s43017-021-00228-x, number: 12 Publisher: Nature Publishing Group, 2021.

Iversen, S. C., Sperrevik, A. K., and Goux, O.: Improving the SST in a regional ocean model through refined SST assimilation, EGUsphere, pp. 1–24, https://doi.org/10.5194/egusphere-2022-957, publisher: Copernicus GmbH, 2022.

Jacobs, G., D'Addezio, J., Ngodock, H., and Souopgui, I.: Observation and model resolution implications to ocean prediction, Ocean Modelling, p. 101760, https://doi.org/10.1016/j.ocemod.2021.101760, 2021.

Janssen, P.: Ocean wave effects on the daily cycle in SST, J. Geophys. Res., 117, C00J32, https://doi.org/10.1029/2012JC007943, 2012.

Kusahara, K., Williams, G. D., Massom, R., Reid, P., and Hasumi, H.: Roles of wind stress and thermodynamic forcing in recent trends in Antarctic sea ice and Southern Ocean SST: An ocean-sea ice model study, Global and Planetary Change, 158, 103–118, https://doi.org/10.1016/j.gloplacha.2017.09.012, 2017.

Larson, J., Jacob, R., and Ong, E.: The model coupling toolkit: A new fortran90 toolkit for building multiphysics parallel coupled models, The International Journal of High Performance Computing Applications, 19, 277–292, https://doi.org/10.1177/1094342005056115, 2005.

Lavergne, T., Sørensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., Bell, L., Dybkjær, G., Eastwood, S., Gabarro, C., Heygster, G., Killie, M. A., Brandt Kreiner, M., Lavelle, J., Saldo, R., Sandven, S., and Pedersen, L. T.: Version 2 of the EUMETSAT OSI SAF and ESA

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023




CCI sea-ice concentration climate data records, The Cryosphere, 13, 49–78, https://doi.org/10.5194/tc-13-49-2019, publisher: Copernicus GmbH, 2019.

Lind, S., Ingvaldsen, R. B., and Furevik, T.: Arctic warming hotspot in the northern Barents Sea linked to declining sea-ice import, Nature Climate Change, 8, 634–639, https://doi.org/10.1038/s41558-018-0205-y, number: 7 Publisher: Nature Publishing Group, 2018.

Lipscomb, W. H. and Hunke, E. C.: Modeling Sea Ice Transport Using Incremental Remapping, Monthly Weather Review, 132, 1341–1354, https://doi.org/10.1175/1520-0493(2004)132<1341:MSITUI>2.0.CO;2, 2004.

Lipscomb, W. H., Hunke, E. C., Maslowski, W., and Jakacki, J.: Ridging, strength, and stability in high-resolution sea ice models, Journal of Geophysical Research: Oceans, 112, https://doi.org/10.1029/2005JC003355, 2007.

Lisæter, K. A., Rosanova, J., and Evensen, G.: Assimilation of ice concentration in a coupled ice–ocean model, using the Ensemble Kalman filter, Ocean Dynamics, 53, 368–388, https://doi.org/10.1007/s10236-003-0049-4, 2003.

Mile, M., Azad, R., and Marseille, G.-J.: Assimilation of Aeolus Rayleigh-Clear Winds Using a Footprint Operator in AROME-Arctic Mesoscale Model, Geophysical Research Letters, 49, e2021GL097 615, https://doi.org/10.1029/2021GL097615, 2022.

Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B., Foley, D., Doyle, J. D., Costa, D., and Robinson, P.: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems: Part II – Performance and application to the California Current System, Progress in Oceanography, 91, 50–73, https://doi.org/10.1016/j.pocean.2011.05.003, 2011.

Müller, M., Batrak, Y., Kristiansen, J., Køltzow, M. A. O., Noer, G., and Korosov, A.: Characteristics of a Convective-Scale Weather Forecasting System for the European Arctic, Monthly Weather Review, 145, 4771–4787, https://doi.org/10.1175/MWR-D-17-0194.1, 2017.

Naughten, K. A., Galton-Fenzi, B. K., Meissner, K. J., England, M. H., Brassington, G. B., Colberg, F., Hattermann, T., and Debernard, J. B.: Spurious sea ice formation caused by oscillatory ocean tracer advection schemes, Ocean Modelling, 116, 108–117, https://doi.org/10.1016/j.ocemod.2017.06.010, 2017.

Naughten, K. A., Meissner, K. J., Galton-Fenzi, B. K., England, M. H., Timmermann, R., Hellmer, H. H., Hattermann, T., and Debernard, J. B.: Intercomparison of Antarctic ice-shelf, ocean, and sea-ice interactions simulated by MetROMS-iceshelf and FESOM 1.4, Geoscientific Model Development, 11, 1257–1292, https://doi.org/10.5194/gmd-11-1257-2018, 2018.

Noer, G., Saetra, O., Lien, T., and Gusdal, Y.: A climatological study of polar lows in the Nordic Seas, Quarterly Journal of the Royal Meteorological Society, 137, 1762–1772, https://doi.org/10.1002/qj.846, 2011.

Rothrock, D. A.: The energetics of the plastic deformation of pack ice by ridging, Journal of Geophysical Research (1896-1977), 80, 4514–4519, https://doi.org/10.1029/JC080i033p04514, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/JC080i033p04514, 1975.

Röhrs, J. and Christensen, K. H.: Drift in the uppermost part of the ocean, Geophy. Res. Lett., 42, 1–8, https://doi.org/10.1002/2015GL066733, 2015.

Röhrs, J., Christensen, K. H., Vikebø, F. B., Sundby, S., Saetra, O., and Broström, G.: Wave-induced transport and vertical mixing of pelagic eggs and larvae, Limnol. Oceanogr., 59(4), 1213–1227, https://doi.org/10.4319/lo.2014.59.4.1213, 2014.

Röhrs, J., Sutherland, G., Jeans, G., Bedington, M., Sperrevik, A. K., Dagestad, K.-F., Gusdal, Y., Mauritzen, C., Dale, A., and LaCasce, J. H.: Surface currents in operational oceanography: Key applications, mechanisms, and methods, Journal of Operational Oceanography, 16, 60–88, https://doi.org/10.1080/1755876X.2021.1903221, 2023.

Sakov, P.: EnKF-C user guide, https://doi.org/10.48550/arXiv.1410.1233, arXiv:1410.1233 [cs], 2014.

Sakov, P. and Oke, P. R.: A deterministic formulation of the ensemble Kalman filter: an alternative to ensemble square root filters, Tellus A: Dynamic Meteorology and Oceanography, 60, 361–371, https://doi.org/10.1111/j.1600-0870.2007.00299.x, number: 2 Publisher: Stockholm University Press, 2008.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023




Samuelsen, E. M.: Ship-icing prediction methods applied in operational weather forecasting, Quarterly Journal of the Royal Meteorological Society, 144, 13–33, https://doi.org/10.1002/qj.3174, 2018.

Schweiger, A. J. and Zhang, J.: Accuracy of short-term sea ice drift forecasts using a coupled ice-ocean model, Journal of Geophysical Research: Oceans, 120, 7827–7841, https://doi.org/10.1002/2015JC011273, 2015.

Shchepetkin, A. F. and McWilliams, J. C.: The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model, Ocean Modelling, 9, 347–404, https://doi.org/10.1016/j.ocemod.2004.08.002, 2005.

Spreen, G., Kaleschke, L., and Heygster, G.: Sea ice remote sensing using AMSR-E 89 GHz channels, J. Geophys. Res, 113, C02S03, 2008.

Strand, K. O., Sundby, S., Albretsen, J., and Vikebø, F. B.: The Northeast Greenland Shelf as a Potential Habitat for the Northeast Arctic Cod, Frontiers in Marine Science, 4, https://doi.org/10.3389/fmars.2017.00304, 2017.

Strand, K. O., Huserbråten, M., Dagestad, K.-F., Mauritzen, C., Grøsvik, B. E., Nogueira, L. A., Melsom, A., and Röhrs, J.: Potential sources of marine plastic from survey beaches in the Arctic and Northeast Atlantic, Science of The Total Environment, 790, 148 009, https://doi.org/10.1016/j.scitotenv.2021.148009, 2021.

Thorndike, A. S., Rothrock, D. A., Maykut, G. A., and Colony, R.: The thickness distribution of sea ice, Journal of Geophysical Research, 80, 4501–4513, https://doi.org/10.1029/JC080i033p04501, 1975.

Turner, A. K., Hunke, E. C., and Bitz, C. M.: Two modes of sea-ice gravity drainage: A parameterization for large-scale modeling, Journal of Geophysical Research: Oceans, 118, 2279–2294, https://doi.org/10.1002/jgrc.20171, 2013.

Umlauf, L. and Burchard, H.: Second-order turbulence closure models for geophysical boundary layers. A review of recent work, Cont. Shelf Res., 25, 795–827, https://doi.org/10.1016/j.csr.2004.08.004, 2005.

Warner, J. C., Sherwood, C. R., Arango, H. G., and Signell, R. P.: Performance of four turbulence closure models implemented using a generic length scale method, Ocean Model., 8, 81–113, https://doi.org/10.1016/j.ocemod.2003.12.003, 2005.

Whitaker, J. S. and Hamill, T. M.: Ensemble Data Assimilation without Perturbed Observations, Monthly Weather Review, 130, 1913–1924, https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2, publisher: American Meteorological Society Section: Monthly Weather Review, 2002.

Xie, J., Bertino, L., Counillon, F., Lisæter, K. A., and Sakov, P.: Quality assessment of the TOPAZ4 reanalysis in the Arctic over the period 1991–2013, Ocean Sci., 13, 123–144, https://doi.org/10.5194/os-13-123-2017, 2017.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

| Transform function | Stretching function | $\theta_S$ | $\theta_B$ | $H_c$ |
|---|---|---|---|---|
| 2 | 4 | 6. | 0.3 | 100$m$ |

**Table 1.** Parameters for the setup of the topography following vertical coordinate

| P | M | N | Kmin | Pmin | CMU0* | C1 | C2 | C3M* | C3P | SIGK | SIGP |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2.0d0 | 1.0d0 | -0.67d0 | 1.0d-8 | 1.0d-8 | 0.5270d0 | 1.0d0 | 1.22d0 | 0.05d0 | 1.0d0 | 0.8d0 | 1.07d0 |

**Table 2.** Parameters for the GLS turbulence scheme used in Barents-2.5. *) parameters associated with the choice of stability function, which is set using the `Canuto_A` pre-compiler option.

| Variable | numeric scheme |
|---|---|
| free surface | Chapman exlicit |
| 2D momentum | Shchepetkin |
| 3D momentum | oblique radiation and nudging |
| salinity and temperature | oblique radiation with nudging |
| TKE and GLS | gradient |

**Table 3.** Numeric boundary solution schemes used in the ocean component for Barents-2.5. All four model boundaries use the same schemes. Nudging towards the boundary TOPAZ fields is imposed within an area of 40 grid points from the boundaries with decaying nudging coefficients.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

| constituent | K2 | S2 | M2 | N2 | K1 | P1 | O1 | MN4 | M4 | MS4 |
|---|---|---|---|---|---|---|---|---|---|---|
| period [h] | 11.96723 | 12 | 12.4206 | 12.65835 | 23.93447 | 24.06589 | 25.81934 | 6.269174 | 6.210301 | 6.103339 |

**Table 4.** Tidal constituents used for tidal forcing in Barents-2.5.

| | SIC | SST | *in-situ* temperature | *in-situ* salinity |
|---|---|---|---|---|
| $\sigma_r$ | 10 | 10 | 4 | 10 |
| $r_{loc}$ | 50 km | 50 km | 8m km | 80 km |
| platform | AMSR-E | AVHRR, VIIRS, SLSTR | profiling floats, drifting buoys, moorings, CTD casts, ships | |

**Table 5.** Observation platforms of assimilated variables and configuration parameters for observation types in EnKF-c.

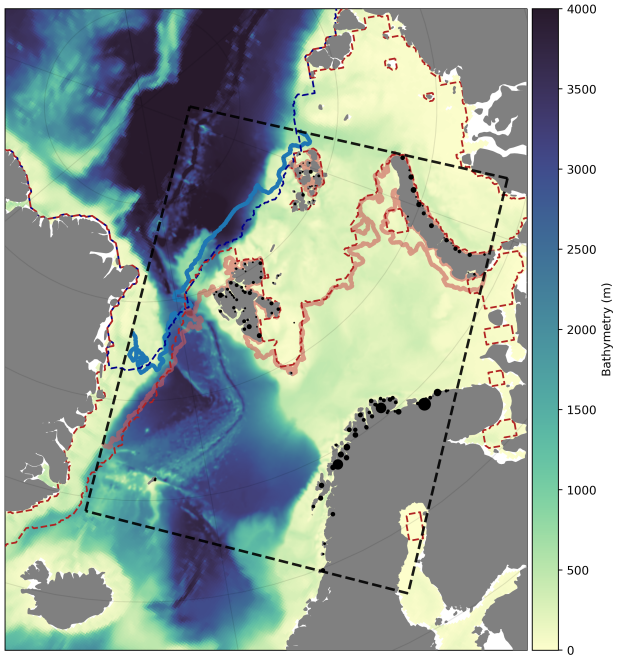


**Figure 1.** Bathymetry of Barents-2.5 plotted within TOPAZ4 model, explicitly showing the Barents-2.5 domain. The locations of river input are marked by black dots, with dot sizes representing the yearly mean of river transport. The sea ice edge, here represented by 15% sea ice concentration, is shown at the time of minimal (blue) and maximal (red) sea ice extent for each model (Barents-2.5 in solid lines, and TOPAZ4 in dashed lines).

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions
Open Access
EGU

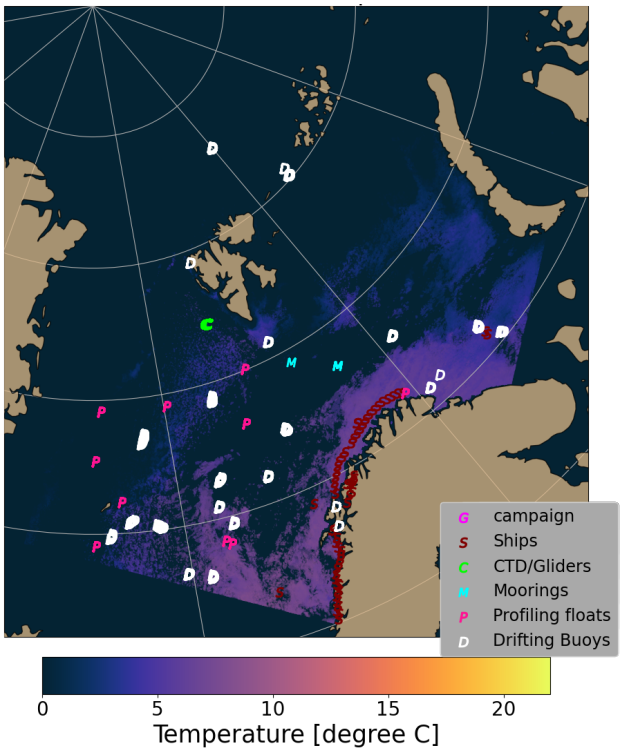


**Figure 2.** In-situ data coverage during the assimilation cycle at 2022-12-15


**Figure 3.** Work flow diagram for operational ensemble prediction system. Each horizontal bar represents a group 6 individual model runs. At each 00 UTC bulletin time, the EnKF DA scheme provides an analysis as new initial conditions for all 24 members. Every subsequent bulletin time (06,12,18 UTC) includes a short spin up run from the analysis time at 00 UTC, and a forecast run of 66 hours using updated atmospheric forcing.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions
Open Access
EGU

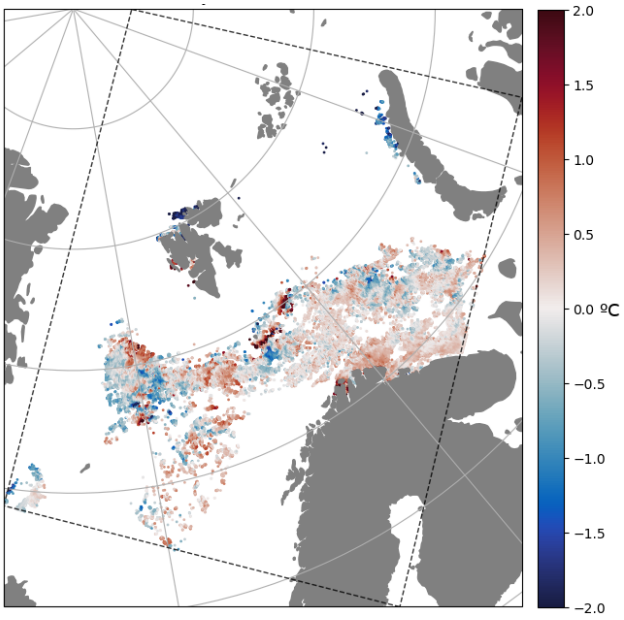


**Figure 4.** Difference between model SST and observed SST from Sentinel-3. The displayed data includes all swaths that pass the model domain during a 24 h period (2022-12-15). The respective model values for computing differences are taken from the satellite retrieval time.

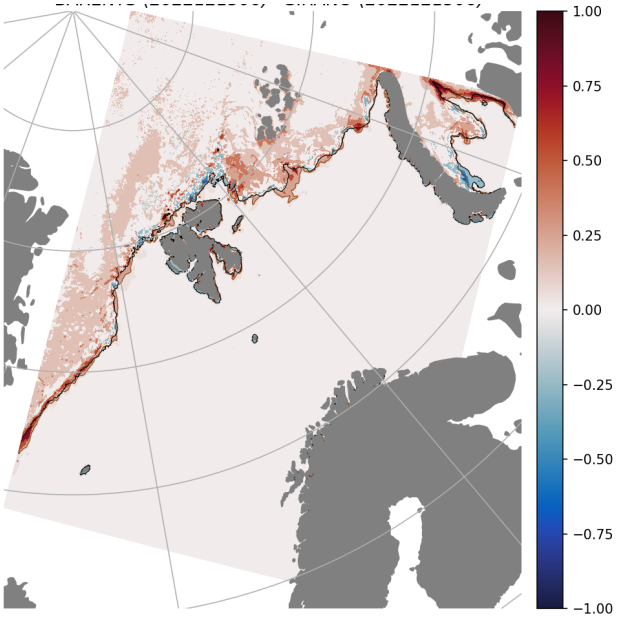


**Figure 5.** Difference between model and observed SIC at 2022-12-15.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions

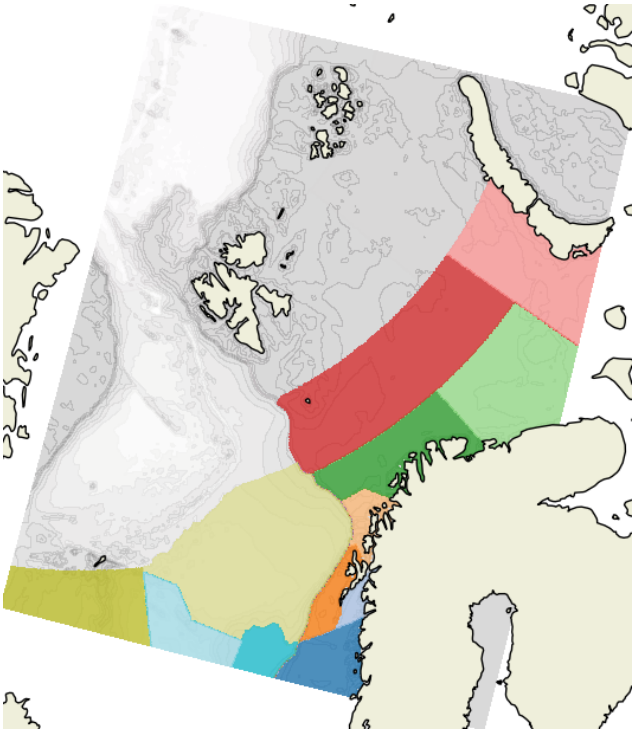

**Figure 6.** Definition of regions used in SST validation. Areas that include the marginal ice zone (without color) are excluded from the SST validation.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions


**Figure 7.** Time evolution of ME (a) and RMSE (b) for SST, separated for various regions in the model domain as defined by the colours in Fig. 6. For each available swath, Sentinel-3 SLSTR observations are compared with the model value that is closest to the satellite retrieval time in the 0-24h forecast range. Monthly averages of available data are shown.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions
Open Access
EGU

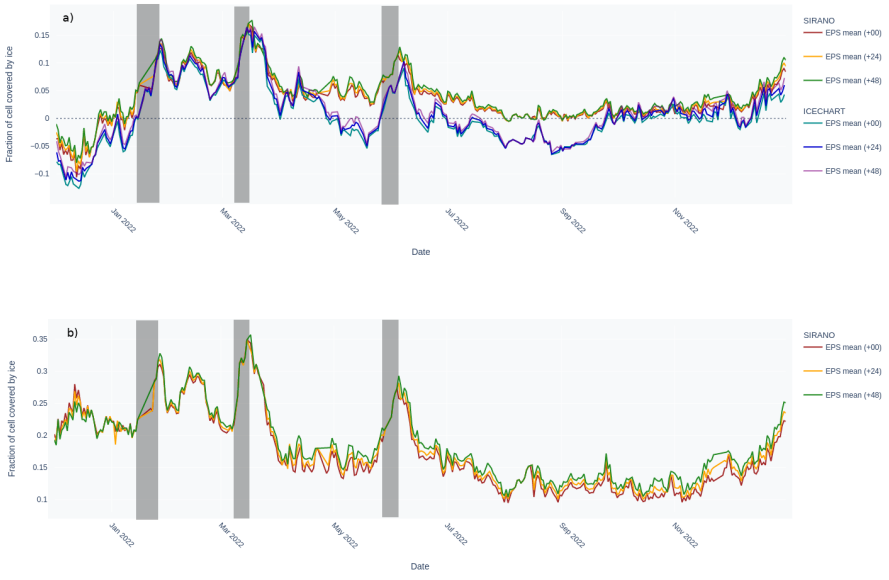


**Figure 8.** Time evolution of ME (a) and RMSE (b) for ice concentration, based on daily comparison of the model values in various forecast ranges averages over the entire model domain. Differences are shown for comparison with the passive microwave ice concentration product that is assimilated, and for a comparison with subjected ice charts that are produced by ice analysts. Shaded intervals mark model cycles when no data was assimilated for more than 3 consecutive days.

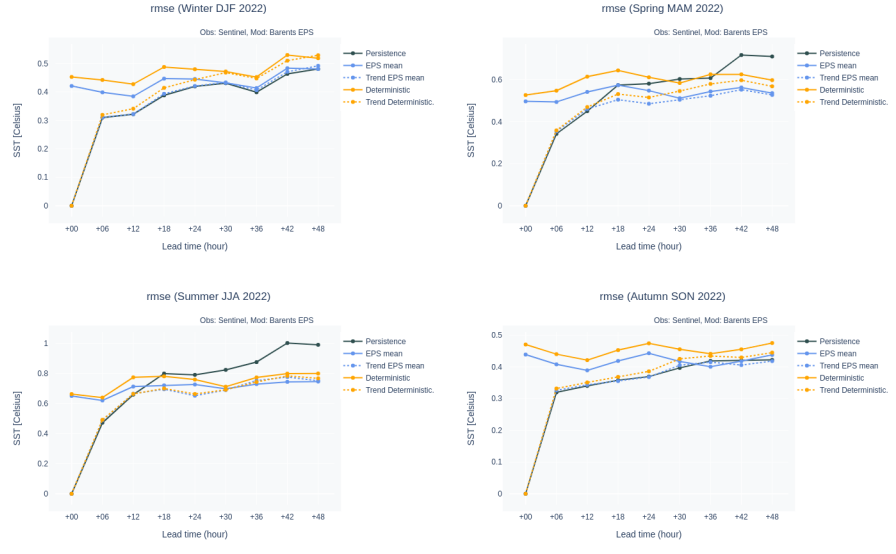


**Figure 9.** Root mean square error (RMSE) of SST as function of lead time, for each season. The black line shows the root mean square difference between overlapping scenes of SST observations, that are shifted by the respective lead time.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions
Open Access
EGU

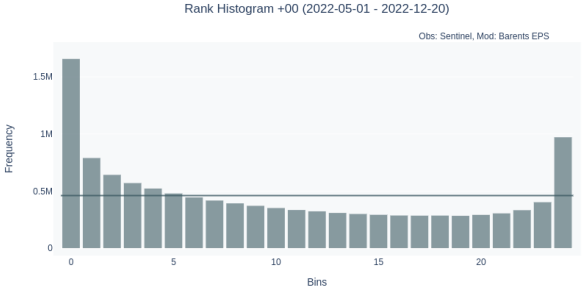


**Figure 10.** Rank histogram of SST model ensemble, evaluating the rank location of SST observations in the ensemble of model values.

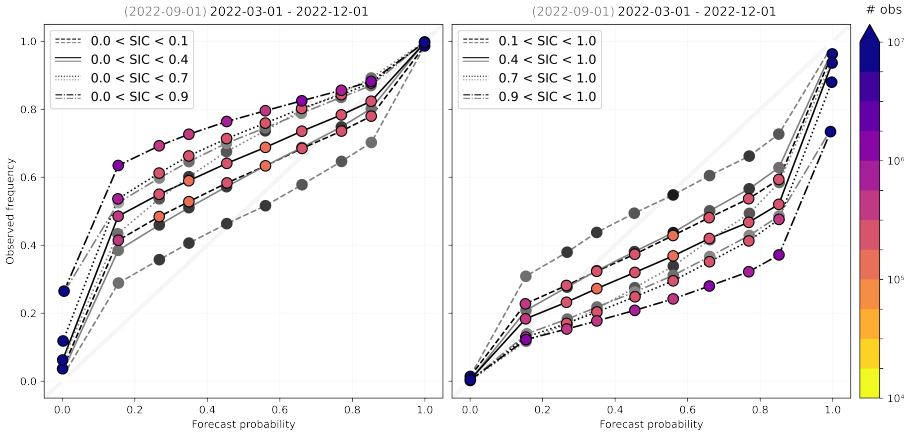


**Figure 11.** Reliability diagram for SIC probabilities derived from the EPS. Reliabilitis have been computed for eight ice concentration intervals, stating hos well the ensemble predicts the chance of low SIC occurence (left panel) and high SIC occurence (right panel) The winter season has been excluded in the shaded curves, leaving a time period where the model generally performs better.

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


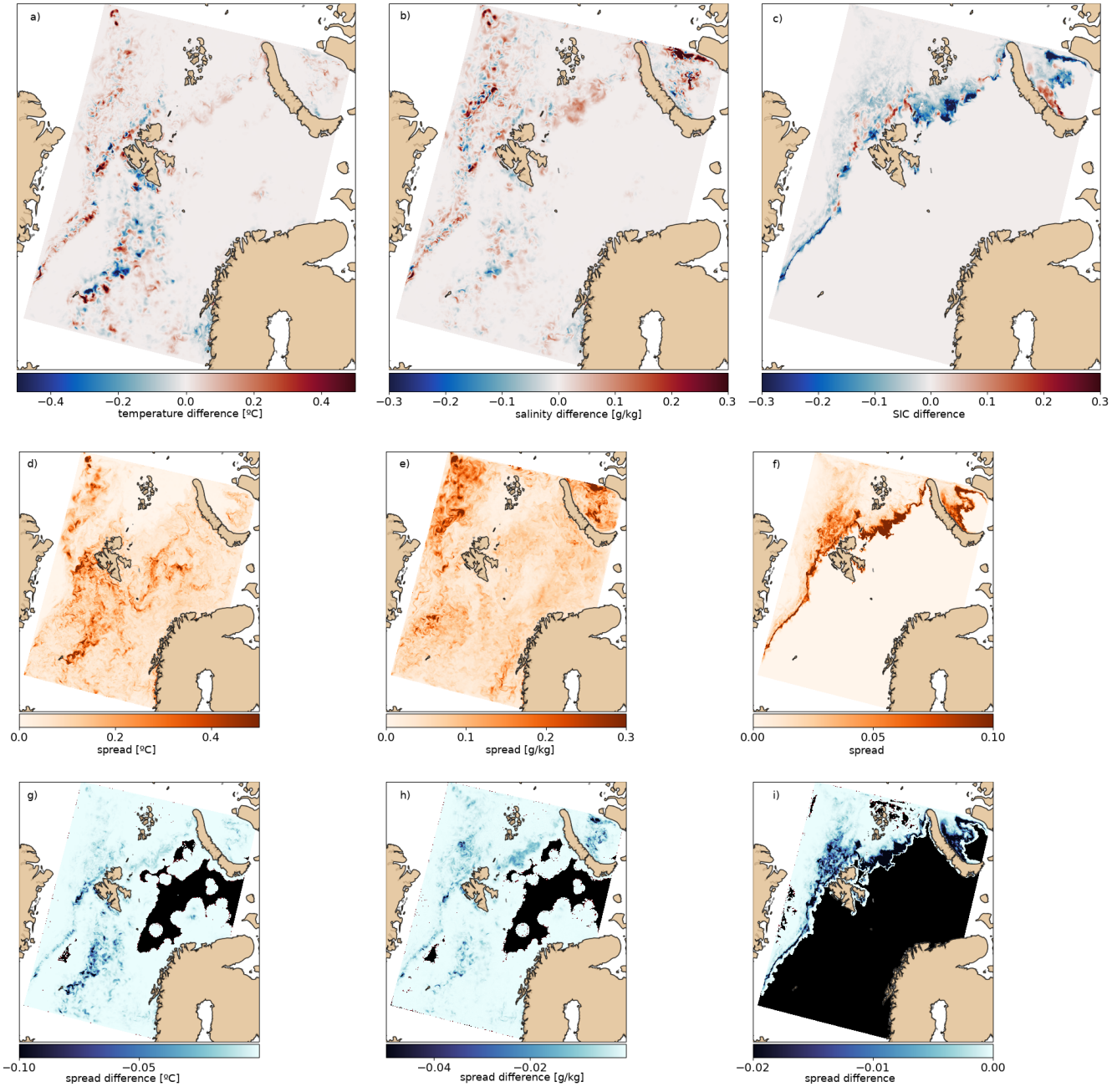


**Figure 12.** Analysis increments of SST (a), surface salinity (b) and ice concentration (c) for an assimilation cycle at 2011-12-15, as calculated by the EnKF. Ensemble spread of the model background state for SST (d), surface salinity (e) and ice concentration (f). Spread increments for SST (g), surface salinity (h), and ice concentration (i).

https://doi.org/10.5194/gmd-2023-20
Preprint. Discussion started: 6 March 2023


Geoscientific Model Development Discussions
Open Access
EGU

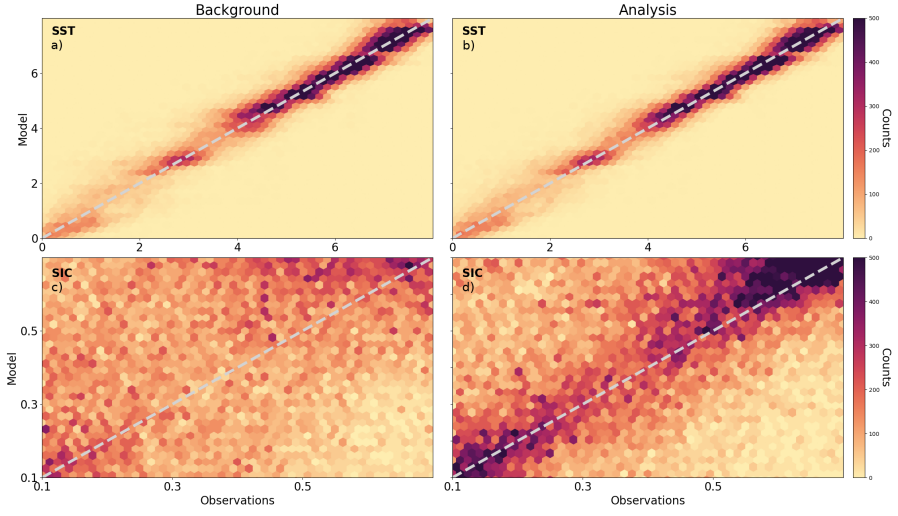


**Figure 13.** Binned scatter plot of model versus observation values for SST (a,b) and ice concentration (c,d). Panels a,c show a comparison for the background model state, and b,d for the analysis on 2022-12-15, i.e. the same DA cycle as in Fig. 12, based on the observations used in Figs. 2 and 5