Improving ~~the~~ SST in a regional ocean model through refined SST assimilation

Silje Christine Iversen[1], Ann Kristin Sperrevik[2], and Olivier Goux[3]

[1]Department of Physics and Technology, UiT The Arctic University of Norway, Tromsø, Norway
[2]Division for Ocean and Ice, Norwegian Meteorological Institute, Oslo, Norway
[3]CERFACS/CECI CNRS UMR 5318, Toulouse, France

Correspondence: Silje Christine Iversen (silje.c.iversen@uit.no)

Abstract. Infrared (IR) and passive microwave (PMW) satellite sea surface temperature (SST) retrievals are valuable to assimilate into high-resolution regional ocean forecast models. Still, there are issues related to these SSTs that need to be addressed to achieve improved ocean forecasts. Firstly, satellite SST products tend to be biased. Assimilating SSTs from different providers can thus cause the ocean model to receive inconsistent information. Secondly, while ~~the~~ PMW SSTs are valuable for constraining ~~the model in cloudy regions~~models during cloudy conditions, the spatial resolution of these retrievals is rather coarse. Assimilating PMW SSTs into high-resolution ocean models will spatially smooth the modeled SST and consequently remove finer SST structures. In this study, we implement a bias correction scheme that corrects ~~the~~ satellite SSTs before assimilation. We also introduce a special observation operator, called the supermod operator, into the Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation algorithm. This supermod operator handles the resolution mismatch between the coarse observations and the finer model. We test the bias correction scheme and the supermod operator using a setup of ROMS covering the shelf seas and shelf break off Norway. The results show that the validation statistics in the modeled SST improve if we apply the bias correction scheme. We also find improvements in the validation statistics when we assimilate PMW SSTs in conjunction with the IR SSTs. However, our supermod operator must be activated to avoid smoothing the modeled SST structures of spatial scales smaller than twice the PMW SST footprint. Both the bias correction scheme and the supermod operator are easy to apply, and the supermod operator can easily be adapted for other observation variables.

1 Introduction

Satellite-based sea surface temperature (SST) retrievals account for the majority of observations assimilated into most ocean forecast models and can be retrieved by measuring both the infrared (IR) and passive microwave (PMW) radiation emitted by the ocean surface. ~~In cloudy regions such as high northern latitudes, the cloud~~ Cloud coverage reduces the number of SST retrievals by ~~the~~ IR sensors. Consequently, SSTs from PMW radiometers serve as a complementary data set as ~~the~~ PMW radiation penetrates clouds. While the spatial resolution of IR SSTs typically is ∼1 km, which is similar to or finer than those of regional ocean forecast models (∼1–10 km), a disadvantage with ~~the~~ PMW SSTs is their relatively coarse resolution. For the instrument AMSR-2 onboard GCOM-W1, the elliptic PMW SST footprint is ∼35 × 62 km (Imaoka et al., 2010).

Indeed, mesoscale ocean models can resolve circulation features of smaller spatial scales than ~~those the~~ PMW SSTs can provide. Such differences in resolved scales are commonly referred to as representation errors within the field of data assimilation (Janjić et al., 2018). Previous studies have considered the differences in resolved scales while assimilating coarse observations such as those of sea ice concentration, satellite sea surface salinity, and scatterometer ocean surface winds (Buehner et al., 2013; Martin et al., 2019; Mile et al., 2021). ~~Furthermore, the difference between the footprint sizes of the PMW and IR SSTs has been considered~~ In these studies, the observation operator was modified such that the coarse observations could be compared with the average of the model values located within an area similar to the spatial resolution of the observations. Similar ideas for the observation operator have also been applied while producing global SST analyses where IR and PMW SSTs are merged using statistical methods (Donlon et al., 2012; Brasnett and Colan, 2016). However, ~~an attempt with regards to assimilating~~ as far as the authors are aware, such an observation operator has not been described and implemented for the assimilation of PMW SSTs into a high-resolution regional ocean forecast model~~has so far not been undertaken~~.

~~We implement a special~~ To tackle the mismatch in spatial resolution, we implement an observation operator, called the supermod operator, into the Regional Ocean Modeling System (ROMS; Haidvogel et al., 2008; Shchepetkin and McWilliams, 2009) 4-dimensional variational (4D-Var) data assimilation system (Moore et al., 2011; Gürol et al., 2014). This operator considers the resolution mismatch between the PMW SSTs and the model by comparing each PMW SST observation with the model mean over an area similar to the observation footprint. The implementation of the operator follows a similar methodology as described in Mile et al. (2021). We show that, compared to a traditional observation operator, the supermod operator prevents PMW SSTs from constraining the spatial scales of the model that the PMW SSTs do not resolve properly. Thus, the supermod operator enables us to assimilate PMW SSTs without smoothing the modeled SST structures. We test the operator using a configuration of ROMS that covers the shelf seas and shelf break off Norway. This region is subject to high cloud coverage, meaning that there is a need for assimilating PMW SSTs to better constrain the model. ~~Moreover, the risk of smoothing modeled SST structures is high in this region since the spatial scales of the SST structures typically are small~~ Also, the Rossby radius of deformation, which affects the horizontal scales of mesoscale processes, decreases with increasing latitude. Consequently, SST structures arising from mesoscale processes are typically smaller at high latitudes, which leads to a greater risk of smoothing the modeled SST in these regions when assimilating coarse PMW SSTs.

To further improve the modeled SST, we take into account that satellite SSTs from different providers tend to be biased due to differences and limitations in the SST retrieval processes and algorithms (Chan and Gao, 2005; Høyer et al., 2012; Wang et al., 2016). Specifically, we implement a bias correction scheme which is inspired by Høyer et al. (2014) and use this scheme to correct the assimilated satellite ~~SST observations~~ SSTs for biases. The scheme ensures consistency between SST observations from different providers, thus preventing the model from receiving inconsistent information during the assimilation of different ~~satellite SST~~ products. By reducing ~~the inconsistent information~~inconsistencies, the probability of artificial noise in the modeled SST is hampered.

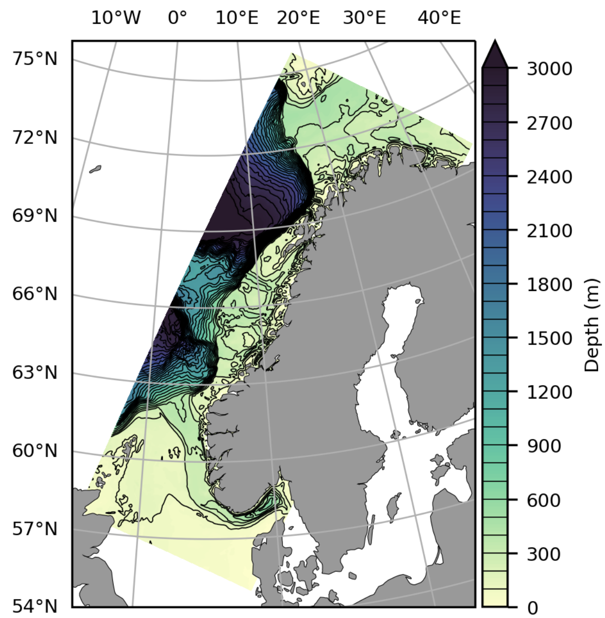

Figure 1. Bathymetry of the model region covering the shelf seas and shelf break off Norway.

2 Data and methods

2.1 Model configuration and data assimilation setup

~~We use a configuration of ROMS to model the shelf seas and shelf break off Norway (Fig. 1).~~ ROMS is a free-surface, hydrostatic, primitive equation ocean model using stretched, terrain-following vertical coordinates, which is beneficial for modeling shallow, coastal waters. The model domain used in this study covers the shelf seas and shelf break off Norway (Fig. 1). We present only a brief summary of the model configuration since a thorough description can be found in Röhrs et al. (2018).

The model domain has a horizontal resolution of ~~the model is~~ 2.4 km, and there are 42 layers in the vertical direction which are distributed so that the uppermost layer has a thickness of $\sim$0.2–1.2 m. Vertical mixing is parameterized using a second-order scheme for turbulent kinetic energy and a generic length-scale ~~(GLS)~~ based on a setup recommended by Umlauf and Burchard (2005) and Warner et al. (2005). The configuration differs slightly from that of Röhrs et al. (2018) with different choices for the *Kmin* and *Pmin* parameters, which in our case are set to 1.0×10^{-8}.

We force the model with three-hourly operational forecast data from the Integrated Forecast System of the European Centre for Medium-Range Weather Forecasts (ECMWF, 2016). Air–sea fluxes for momentum and heat are calculated via the COARE3.0 bulk flux algorithms (Fairall et al., 2003) within ROMS. Open boundary conditions consist of daily averages of all state variables from the TOPAZ4 reanalysis (Xie et al., 2017). The sea surface elevation and currents are adjusted to include the effect of local atmospheric pressure, as TOPAZ4 does not include the inverse barometer effect. Harmonic tidal forcing, consisting of eight tidal constituents from the TPXO9 global inverse barotropic model (Egbert and Erofeeva, 2002), is also imposed along the open boundaries. Riverine runoff is based on modeled river discharges from the Norwegian Water Resources and Energy Directorate (Beldring et al., 2003).

The model is initialized from the TOPAZ4 reanalysis on 1 January 2017 and run with 4D-Var assimilation of satellite SST and in situ temperature and salinity observations as described in Röhrs et al. (2018). This reanalysis is used to initialize a set of different experiments (see Sect. 4.1), all of which assimilate only satellite SST observations and are run for the period 21 April 2018–22 June 2018.

The assimilation algorithm used for the experiments is the dual formulation of the ROMS 4D-Var system (RBL4DVAR; Moore et al., 2011; Gürol et al., 2014). Generally, in 4D-Var data assimilation, the analysis is found by minimizing a cost function in the form:

$$
\begin{aligned}
J(\boldsymbol{x}) = &\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^b)^T\mathbf{B}^{-1}(\boldsymbol{x}-\boldsymbol{x}^b)+ \\
&\frac{1}{2}\left(\boldsymbol{y}-\mathcal{H}(\boldsymbol{x})\right)^T\mathbf{R}^{-1}(\boldsymbol{y}-\mathcal{H}(\boldsymbol{x})),
\end{aligned}
\tag{1}
$$

where the control vector, $\boldsymbol{x}$, holds the control variables, which include the prognostic variables (temperature, salinity, horizontal velocity, and sea surface height) at the model grid points. The background control vector is denoted $\cancel{\boldsymbol{x}^{\boldsymbol{b}}}$ $\boldsymbol{x}^b$, while the vector $\boldsymbol{y}$ ~~is the observations. In our setup, we gather the observationsusing a temporal spacing of 15 minutes~~contains the observations. The observation operator, $\mathcal{H}$, returns the model equivalents of the observations by interpolating the model values to the observation locations. ~~This interpolation is dynamic in nature since $\mathcal{H}$ includes the nonlinear model.~~ In case the observed quantity is not part of the control vector, $\mathcal{H}$ may also transform the model values to correspond to the observed quantity. In 4D-Var, the nonlinear model is a part of $\mathcal{H}$ and facilitates temporal mapping. $\mathbf{B}$ is the background error covariance matrix, while $\mathbf{R}$ is the observation error covariance matrix. $\mathbf{R}$ is ~~required~~ assumed to be diagonal in most operational data assimilation systems, including ~~ROMS, as the 4D-Var minimization algorithm involves the inverse of $\mathbf{R}$ (e.g., Gürol et al., 2014). With a diagonal $\mathbf{R}$, we assume~~ the implementation in ROMS used in this study (Gürol et al., 2014). This assumption means that the observation errors are assumed to be uncorrelated in both time and space.

The observation errors used to construct $\mathbf{R}$ may represent the measurement error, which is the error associated with the accuracy of the measuring instrument, and representation errors (Janjić et al., 2018). The latter is connected to errors arising from the pre-processing and ~~quality-control~~ quality control of the observations, the observation operator, and the mismatch between the resolutions of the observations and the model grid. ~~In~~

In ROMS, the construction of the background error covariance matrix $\mathbf{B}$ follows the approach described in Weaver and Courtier (2001) where $\mathbf{B}$ is separated into a univariate correlation matrix, a multivariate balance operator, and background error standard deviations. The latter is, in this study, ~~the standard~~ calculated from a multi-year simulation with the same model configuration. To avoid unrealistically high values of the standard deviations due to large seasonal variations, background standard deviations are calculated separately for each day of the year, using all of the time records in the multi-year run that fall within ± 14 days of the day in question. Consequently, the $\mathbf{B}$ used in this study changes slightly from one assimilation cycle to the next. We do

Table 1. IR and PMW satellite SST observations used in this study and their approximate spatial resolution at nadir.~~The factor α is used to estimate the observation errors (see Sect. 2.1).~~

Satellite	Sensor	Data creator	Method	Res. (km) ~~α~~
Metop-B	AVHRR	OSI SAF	IR	1.1 ~~2~~
N20	VIIRS	NOAA	IR	0.75 ~~2~~
GCOM-W1	AMSR-2	RSS	PMW	35 $\times$ 62 ~~0.3 (2)~~
NPP	VIIRS	NOAA	IR	0.75 ~~–~~
Sentinel-3A	SLSTR	EUMETSAT	IR	1 ~~–~~

not apply any explicit balance relations between the control variables. However, the linearized physics in the Tangent Linear and Adjoint models implicitly account for correlations between the control variables.

The standard deviation of the error attributed to an observation (σ_o) is estimated as a function of the ~~standard deviation of the background error~~ background error standard deviation at the observation location (σ_b), a factor based on the reported quality level of the satellite products (Q), and a parameter α:

$$\sigma_o^2 = \alpha \cdot Q \cdot \sigma_b^2. \tag{2}$$

The reported quality level reflects the uncertainty of the observation and is assigned to each SST measurement by the data creator (Donlon et al., 2007). Only quality level 4 (QL4, acceptable quality) and 5 (QL5, best quality) are used~~in this study~~. To ensure that more weight is given to the observations with best quality during assimilation, observations with ~~quality levels 4 and 5~~ QL4 and QL5 are assigned a Q of 1.1 and 0.9, respectively. The parameter α ~~varies for each satellite product (Table 1)~~ is set to $\alpha = 2$ for all of the satellite products to ensure that the different experiments are comparable and is chosen by validating a set of experiments with different choices of α against independent data sets. ~~Notice that for the PMW SSTs we use two different values for α: $\alpha = 0.3$ when the supermod operator (see Sect. 3) is activated, and $\alpha = 2$ when the supermod operator is not activated.~~

2.2 Observations

~~Satellite~~ We use satellite SST observations from IR and PMW radiometers~~are used in this study~~, and a summary of the products is found in Table 1. All observations are level-2 preprocessed (L2P) products downloaded from the Physical Oceanography Distributed Active Archive Center (PO.DAAC), the Satellite Application Facility on Ocean and Sea Ice (OSI SAF) FTP server, and the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) data center online ordering client.

The IR SSTs have a spatial resolution ranging from 0.75–1.1 km at nadir and are thus capable of resolving SST structures of similar scales as the structures captured by the model. When several of these observations fall within the same model grid cell during the same time interval, the observations are averaged to create superobservations. In our setup, each individual observation time is rounded to the nearest quarter hour before superobservations are calculated.

PMW radiometers can, ~~contrary to~~ unlike IR radiometers, retrieve SSTs during cloudy conditions. Challenges related to PMW SST retrievals that require affected pixels to be discarded include rain, radio-frequency inference, sun glitter, and emission from sea ice surfaces and land (Minnett et al., 2019). The latter contamination affects PMW retrievals from regions within ∼100 km from land. The spatial resolution of PMW SSTs depends on the frequency used to derive the SSTs ~~and~~ along with the antenna size and is significantly coarser than those of the IR SSTs. For AMSR-2 GCOM-W1, the primary band for PMW SST retrieval at 6.9 GHz corresponds to an elliptic footprint, elongated in the along-track direction of the satellite, with sizes of approximately 35×62 km (Imaoka et al., 2010).

All satellite SST observations are corrected by applying the single-sensor error statistics (SSES) bias, ~~as recommended~~ in line with the recommendations in Donlon et al. (2007). ~~We use quality level 4 and 5~~ As previously mentioned, we use QL4 and QL5 for the assimilated IR SST products. For the PMW SSTs, we use only ~~quality level 5~~ QL5 observations and discard the ~~quality level 4~~ QL4 observations due to unwanted artifacts in the SST field. These artifacts are likely caused by radio-frequency interference from oil rigs (Alerskans et al., 2020). Evaluation of the QL5 PMW SSTs revealed that observations of dubious quality are still present in the data set, both at the edges of areas from which lower quality observations were removed, as well as in areas corresponding to locations of oil rigs. Thus, an additional quality control of the PMW SSTs was implemented which flags observations as bad if they are within a radius of 75 km of an oil rig.

The SLSTR instrument onboard Sentinel-3A uses a dual view technique where the swath is scanned twice through two different atmospheric path lengths. This capability allows for more accurate SST measurements (Luo et al., 2020). ~~Unlike the other SST products, which provide the sub-skin temperature, SLSTR Sentinel-3A provides the skin temperature. An offset of 0.17 °C is added to these skin temperatures so that all of the SSTs used throughout this study represent the sub-skin temperature (Donlon et al., 2002; Høyer et al., 2014). Due to the higher accuracy of the SLSTR sensor onboard Sentinel-3A, we use quality level 5~~ We use QL5 SSTs measured by this sensor to correct the other satellite products for biases (see Sect. 2.3) and for validation purposes. Other SSTs used for validation are bias-corrected SSTs from VIIRS NPP and SSTs measured by drifting buoys from the CORA data set (Cabanes et al., 2013). The spatial coverage of the drifting buoys is shown in Fig. 2.

While SLSTR Sentinel-3A provides the skin temperature, the other satellite products used in this study provide the sub-skin temperature. IR sensors intrinsically measure the skin temperature. However, these skin temperatures are converted to sub-skin temperatures during the data creation process if in situ data are used to tune parameters included in the SST retrieval algorithms. To ensure that all of the SSTs used throughout this study represent the sub-skin temperature, an offset of 0.17 °C is added to the SLSTR Sentinel-3A SSTs (Donlon et al., 2002; Høyer et al., 2014). This constant offset is generally valid for wind speeds above 6 m s^{-1} but may be greater for lower wind speeds. Our results indicate, however, that using this constant offset to convert the SLSTR Sentinel-3A SSTs to sub-skin temperatures does not degrade the modeled SST resulting from assimilating bias-corrected SSTs (see Sect. 4.2).

Figure 3 shows time series of the number of observations inside the model domain during the chosen assimilation period in 2018. For the satellite observations, we count only observations meeting the set minimum quality level. Cloudiness causes

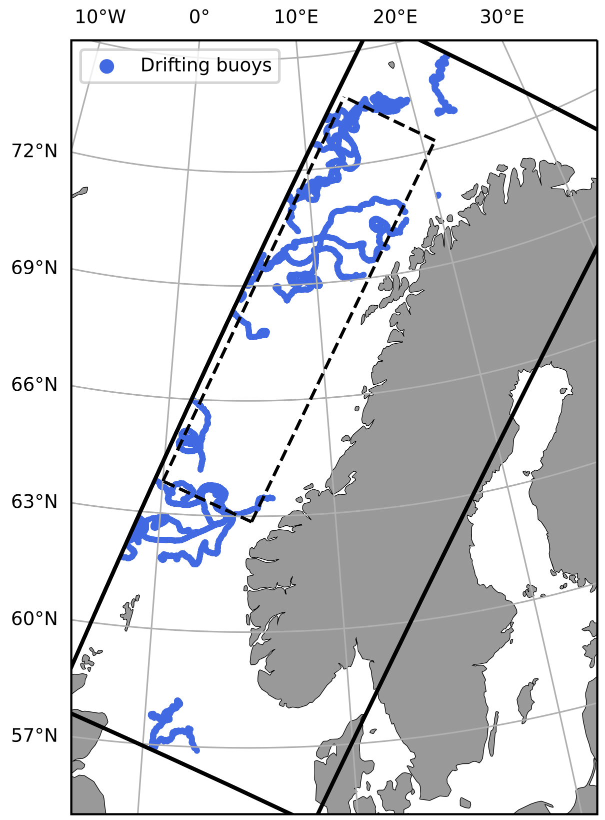

Figure 2. ~~Colored lines show the spatial~~ Spatial coverage of CORA drifting buoys during the period ~~21~~ 27 April 2018–22 June ~~2018.~~ 2018, which is the period used for validation. There are in total $\sim$~~13 000~~ 10 600 observations of this kind. ~~Solid~~ The solid black box shows the extent of the model domain, while the dashed black box shows the domain selected for SST ~~variance~~ power spectrum calculations (see Sect. 4.3).

the number of IR SST observations to vary. The largest number of observations ~~is in~~ occurs at the beginning of May, and there is a drop in the ~~observational coverage~~ number of available observations while approaching June. The number of PMW SST observations is stable during the whole period, which we expect since PMW SSTs can be measured in cloudy conditions. Moreover, the number of available PMW SSTs is on average smaller than the number of available IR SSTs since the PMW SSTs are provided on a coarser grid.

2.3 SST bias correction algorithm

High northern latitudes are recognized as challenging regions for satellite SST retrieval (Donlon et al., 2010; Wang et al., 2016; Jia and Minnett, 2020). In situ data are used to derive the satellite SST retrieval algorithms and to correct and thus improve the quality of the retrieved SST. However, the number of available in situ SST measurements at high northern latitudes is low

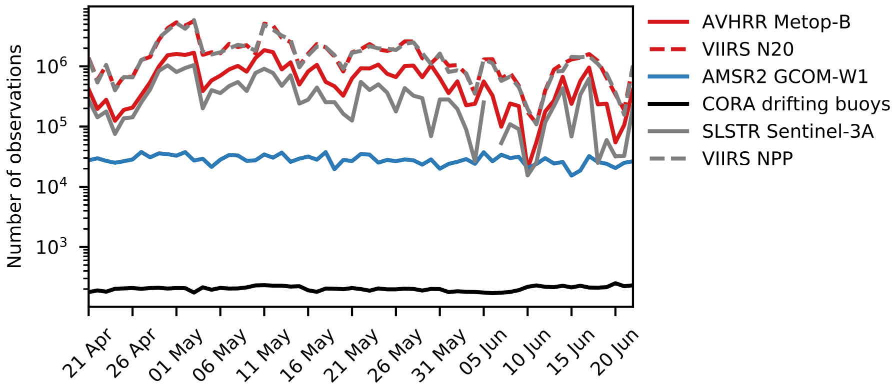

Figure 3. Time series of the number of observations available inside the model domain for each sensor and satellite pair and for the CORA drifting buoys.

compared to equatorial and mid-latitude regions, and this data scarcity can lower the quality of high-latitude IR and PMW SST retrievals (O'Carroll et al., 2019).

~~The SST retrieval~~ SST retrievals in the IR ~~is~~ are subject to additional challenges. First, the process of cloud screening, where cloudy and clear-sky conditions are separated, is critical (O'Carroll et al., 2019). Any undetected clouds will result in erroneous SST retrievals. ~~That the high northern latitudes are exposed to frequent cloud coverage challenges the identification process.~~ Additionally, the brightness temperatures derived from the measured IR ~~radiance~~ radiances are modified by gases, such as water vapor, and aerosols in the atmosphere. Atmospheric correction algorithms are thus required to retrieve accurate SSTs. At high latitudes, these corrections tend to be problematic, ~~hereby~~ thereby degrading the accuracy of the measured SSTs (Kumar et al., 2003; Merchant et al., 2006; Vincent, 2018; Jia and Minnett, 2020).

For the PMW part of the spectrum, challenges that must be corrected for (in addition to those mentioned in Sect. 2.2, which cause the measurements to be discarded) include sea surface roughness generated from strong winds and atmospheric effects such as those associated with modification of the PMW radiation by water vapor and liquid water (Wentz et al., 2000; Minnett, 2014).

SST products from various satellites and sensors are subject to regional and temporal biases due to the difficulties concerning the SST retrieval process and the differences in applied retrieval algorithms (Chan and Gao, 2005; Høyer et al., 2012; Wang et al., 2016). To this end, Høyer et al. (2014) developed and employed a bias correction scheme ~~which~~ that they used to correct SST products from different providers. Their goal was to merge these products, and the correction was performed to avoid any biases from affecting the final product. Bias correction schemes have also been implemented in other studies in order to correct the satellite SSTs being assimilated into ocean models. For example, Waters et al. (2015) employed a scheme where SSTs were corrected using a set of chosen reference observations. This bias correction was performed prior to the assimilation of the observations. While and Martin (2019), on the other hand, implemented a variational bias correction scheme where the SSTs were corrected within the data assimilation algorithm. In such variational schemes, bias correction is achieved by

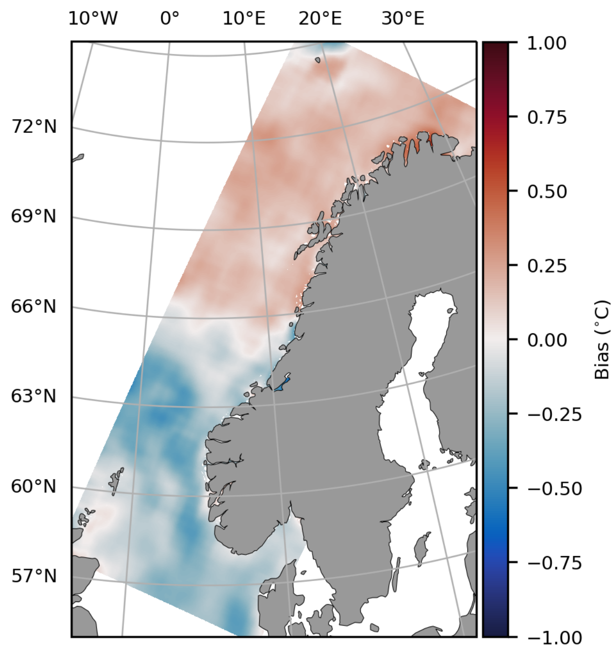

Figure 4. ~~Mean bias~~ Bias for AVHRR Metop-B on 22 May 2018. The bias is calculated using 11 daily difference fields from 17 May 2018 to 27 May 2018.

incorporating additional terms into the cost function. One key advantage of variational schemes is that they do not rely on having full coverage of reference observations at all times.

We implement a similar bias correction scheme as in Høyer et al. (2014), with some modifications, to ensure consistency between the different satellite SST products being assimilated into our model. The bias correction is an individual step during the preparation of observations for assimilation and is not part of the data assimilation process itself. Biases are calculated independently for all satellite products using SLSTR Sentinel-3A SST as reference. Figure 4 illustrates such a bias field for AVHRR Metop-B, and the method is described ~~in what follows~~below.

Daily SST fields on a grid with a resolution of ∼25 km are produced for each satellite product, including the reference product, by first resampling the available L2P SST products to this coarser grid and then performing a simple averaging of the observations falling into each grid cell. To minimize the possibility of using SST observations affected by greater diurnal warming events (Eastwood et al., 2011; Karagali et al., 2012), we exclude observations located in regions where the wind speed is less than 3 m s^{-1} during local summer daytime (defined here as May–August at 10–14 UTC). This wind speed threshold is less strict than that recommended in Donlon et al. (2002) but was chosen to ensure sufficient spatial coverage in the daily fields as the recommended threshold resulted in degraded bias estimates in cloudy regions. We do, however, find that the low threshold value of 3 m s^{-1} performs slightly worse in the few occasions when our region is affected by diurnal warming. In the future, the wind speed threshold should be reassessed and adjusted, possibly in conjunction with a somewhat coarser analysis grid.

Daily difference fields are produced by subtracting the reference daily SST field from the daily SST fields of the other satellite products. We remove all of the values where the difference exceeds ±2 °C. This data removal is performed to exclude unrealistic values which we detected along the coast and ~~nearby~~ in regions affected by cloud contamination.

By aggregating 11 daily difference fields, we produce a ~~mean~~ bias that is valid on the central day. The aggregation is performed as a simple temporal averaging as in Høyer et al. (2014). The ~~bias correction scheme will thus correct for biases persistent for approximately 11 days. Correcting for biases occurring for shorter temporal scales would require a shorter temporal averaging interval, and additional reference SST products would have to be included to ensure sufficient spatial coverage. The mean~~ biases are finally resampled to the model grid and, to avoid unwanted noise, spatially smoothed by applying a uniform filter. This filter has the size of 40 grid points (∼96 km) in both horizontal directions.

3 Supermod operator

As noted, observations with higher spatial resolution than the model are combined into superobservations. This procedure reduces the representation errors, as information on smaller spatial scales than what the model is able to represent is smoothed. For the case of PMW ~~SST~~SSTs, however, each observation represents an SST average over the area covered by its footprint. For AMSR-2 GCOM-W1, this footprint is roughly 375 times larger than the area of a grid cell in our model domain. Thus, we expect the model to contain potentially valuable information on finer spatial scales than what is present in the observations. If observations of this type were assimilated with a traditional observation operator that does not account for the observation footprints, we would expect large, spatially correlated representation errors (Liu and Rabier, 2002). Such errors are particularly problematic in ~~operational~~ data assimilation systems designed to handle uncorrelated observation errors.

An observation operator that ensures that each PMW SST observation is compared with a model mean over an equivalent area as the SST footprint is implemented in ROMS to avoid corrections of finer scales. The implementation of the operator follows a similar methodology as described in Mile et al. (2021). We will refer to this operator as the supermod operator~~hereafter~~.

Our implementation of the supermod operator builds on the assumption that the observation footprints are squares with sides of length $(1+2L)dx$, where $(1+2L)$ is the number of grid cells and dx is the horizontal resolution of the model. L can be set separately for each individual observation, which makes the operator easy to apply to new observation sources and different model configurations. As the center of the observation footprint may be located between model grid points, the supermod value is calculated as a weighted mean of the values in grid cells that fall completely or partially within the footprint area. This is done to conserve the effective resolution of the supermod values and to ensure that the model footprint is centered at the observation location. ~~Another factor~~ Other factors that could affect the effective resolution of the supermod values ~~is the existence of~~ are land points within the footprint or if the footprint extends beyond the boundaries of the model domain. The operator is designed to reject observations with ~~these~~ such footprint features. Figure 5 illustrates how the operator works.

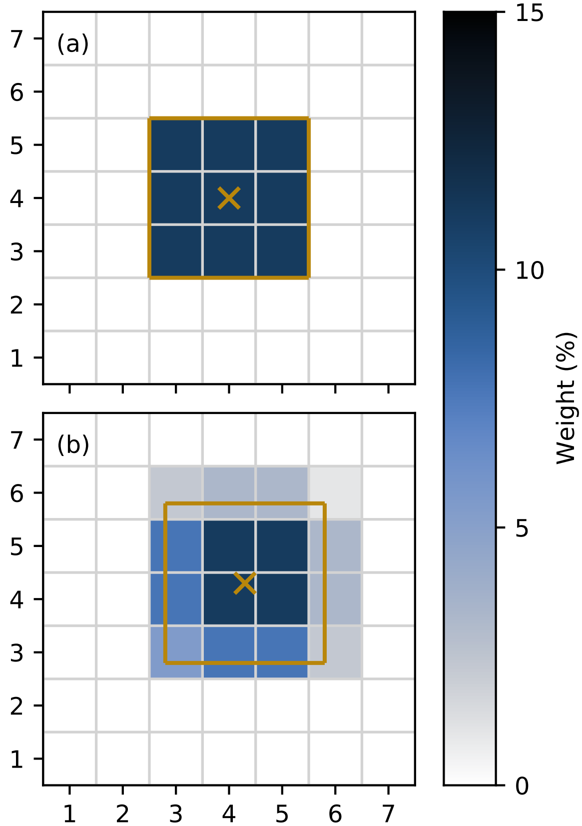

Figure 5. Schematic figure showing the supermod operator with $L = 1$ for (a) an observation that falls exactly in the center of a model grid cell; (b) an observation that is shifted towards the upper right corner of a grid cell. The orange crosses indicate the observation locations, the mesh illustrates grid cells, and the orange squares indicate the footprint area of the observations. The colormap indicates the weight (%) by which the grid cells contribute to the supermod value.

Idealized tests showed a sudden increase in the standard deviations of the innovations (observations minus background) when the footprints were increased and ~~became overlapping~~started to overlap. Thus, a thinning procedure is performed on the PMW observations prior to assimilating the observations. As the supermod operator can only represent square footprints, we have chosen L conservatively to ~~make sure~~ ensure that the square used to represent the footprint will not be smaller than the largest size of the true, elliptic footprint of the satellite retrievals. With the grid resolution of our model being 2.4 km, $L = 13$ gives us footprints with sides of 64.8 km. The PMW observations are thus thinned during a pre-processing step to ensure a distance of at least 64.8 km between the observation points.

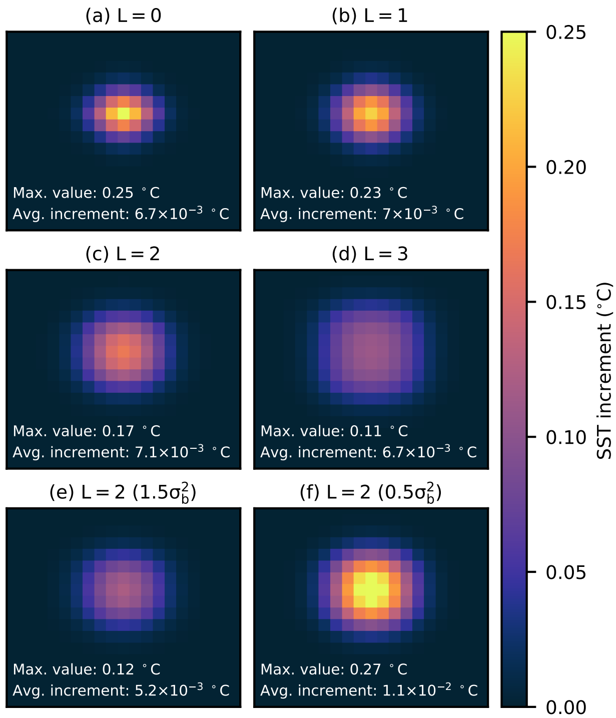

Figure 6. SST increment for (a–d) varying sizes of L and with the observation error (σ_o^2) set equal to the background error (σ_b^2); (e) $L = 2$ and $\sigma_o^2 = 1.5\sigma_b^2$; (f) $L = 2$ and $\sigma_o^2 = 0.5\sigma_b^2$.

3.1 Idealized experiment

An idealized scenario ~~with no surface forcing~~for a domain with flat bathymetry, no land, and ~~a flat bathymetry is applied in order~~ no surface forcing is applied to test the behavior of the supermod operator in ROMS. A coarse grid with a horizontal resolution of ~30 km is chosen to reduce the computational costs. The ocean is set to be initially at rest with a uniform temperature and salinity of 12 °C and 34, respectively. The standard deviations of the background errors are also uniform for every prognostic variable, and for temperature, it is set to 0.02 °C.

A single PMW SST observation, with a temperature of 12.5 °C and a standard deviation of the observation error set similar to the standard deviation of the background, is assimilated into the idealized setup for varying footprint sizes. Figure 6a–d show the resulting SST increments. Increasing the footprint size causes the increment to spread laterally such that the observation affects a greater area. The increment amplitude decreases simultaneously. Furthermore, for each footprint size in Fig. 6a–d, we calculate the average increment using all values greater than zero (see white text in Fig. 6a–d). That the average increment ~~does not change significantly while increasing the footprint illustrates~~ remains relatively unchanged as the footprint size increases, demonstrates that the inclusion of the supermod operator spreads the information from the PMW SST observation as expected.

Table 2. Summary of experiments showing the sources (sensor and satellite) of the assimilated satellite SSTs, whether the observations are corrected for biases through the bias correction scheme, whether the supermod operator is activated, and the approximate number of observations assimilated ($\sim$Nobs).

Experiment	Assimilated ~~SST~~ SSTs	Bias correction	Supermod	$\sim$Nobs
IR1	AVHRR Metop-B VIIRS N20	No	–	2.4×10^7
IR2	AVHRR Metop-B VIIRS N20	Yes	–	2.4×10^7
PMW1	AMSR-2 GCOM-W1	Yes	No	~~1.7~~ 1.6 $\times 10^6$
PMW2	AMSR-2 GCOM-W1	Yes	Yes	~~2.3~~ 2.2 $\times 10^4$
COMB1	AVHRR Metop-B VIIRS N20 AMSR-2 GCOM-W1	Yes	No	2.5×10^7
COMB2	AVHRR Metop-B VIIRS N20 AMSR-2 GCOM-W1	Yes	Yes	2.4×10^7

Figure 6e and f show the effects of increasing and decreasing the observation error by a factor of 1.5 and 0.5, respectively, when the footprint parameter is set to $L = 2$. Increasing (decreasing) the error decreases (increases) the increment amplitude, which is what we would expect. Further tests revealed that the amplitude reduction with increasing L is not as pronounced when the observation error is set much smaller than the background error. In the extreme cases of a very small observation error, the increment might increase with increasing L.

4 Results

4.1 Experimental design

Different experiments are set up to (1) test the effect of correcting ~~the~~ satellite SSTs for biases before assimilation and (2) determine the effect of assimilating PMW SST observations with and without the application of the supermod operator. Table 2 summarizes the experiments, all of which cover the period 21 April 2018–22 June 2018.

To assess the bias correction scheme, IR SSTs are assimilated with (IR2) and without (IR1) ~~applying a correction of~~ correcting for biases. Similarly, to assess the supermod operator, we assimilate PMW SSTs by first treating the SSTs as regular point observations (PMW1) and then by applying the supermod operator (PMW2). Notice that PMW2 includes the process of ~~thinning the observations~~observation thinning. As a consequence, PMW2 assimilates only ∼~~2~~ 1.4% of the observations assimilated in PMW1.

The argument for including PMW SSTs in the assimilated data set is to supplement the IR SSTs during periods of high cloud coverage. Hence, we run two additional experiments to determine the effects of combining IR and PMW SSTs. ~~In these experiments, we correct all of the observations for biases.~~ Experiment COMB1 assimilates IR and PMW SSTs without employing the supermod operator, i.e., the experiment is a combination of IR2 and PMW1. The final experiment, COMB2, is a combination of IR2 and PMW2, thus assimilating IR and PMW SSTs with the supermod operator activated.

The assimilation window is set to three days in all experiments. The assimilated observations update the background model state produced for such a three-day cycle to produce the analysis. The ocean state at the end of the analysis is then used as ~~the initial conditions for~~ initial conditions to produce the background state of the next assimilation cycle, ~~where a background is produced~~ and the process repeats. Thus, the background can be considered a forecast, and we evaluate the experiments using ~~the backgrounds~~ these background states in the following. ~~If not stated otherwise, the~~ The evaluation period excludes the two first assimilation cycles and covers 27 April 2018–22 June 2018. Note that the bias correction scheme is applied in all experiments but IR1.

4.2 Error statistics

Table 3 shows validation statistics ~~in the~~ of SST for the background ~~.~~ states. PMW SST measurements are unavailable in the coastal zone as emissions from land contaminate the retrievals in this region. The experiments assimilating only PMW SSTs thus perform quite poorly in these regions compared to the experiments that include IR SSTs. To assess the experiments, we have chosen to exclude observations within these areas from the validation, and focus instead on the regions where an impact from including PMW SST can be expected.

In comparing IR1 and IR2, we find that correcting the observations for biases reduces the RMSE and the bias of the modeled SST. We expect improved statistics when validating against satellite SSTs since the validating data set contains SLSTR Sentinel-3A SSTs, which ~~were~~ are the SSTs used to bias correct the observations assimilated in IR2. The SSTs from VIIRS NPP are also bias corrected with respect to SLSTR Sentinel-3A SSTs in the same way as the assimilated observations. However, the in situ SSTs measured by drifting buoys are completely independent of the assimilated observations. The improvements in the error statistics calculated using this independent data set reveal ~~the benefits~~ a clear benefit of applying the bias correction scheme prior to assimilation.

The RMSE is higher in PMW1 than in IR2 both when validating against ~~satellite and in situ SSTs~~ SSTs from satellites and drifting buoys. However, the bias of PMW1 is smaller than the bias of IR2 when validating against ~~in situ SSTs~~ SSTs from drifting buoys, indicating that PMW SSTs are valuable to assimilate to reduce the model's systematic error. When validating against IR SSTs, PMW1 has a larger bias than IR2. This larger bias does not reflect the ability of the PMW SST data set to adjust the model~~but was found~~. Rather, we find the bias to be a ~~direct consequence~~ result of having no PMW observations in ~~the~~ coastal zones, causing also the parts of the coastal current that extend beyond these zones to be too cold during its heating phase as summer approaches.

Table 3. SST RMSE (°C) and bias (°C). Validation metrics are calculated using SLSTR Sentinel-3A and bias-corrected VIIRS NPP ~~SST~~ SSTs as reference (left columns) and SSTs from drifting buoys as reference (right columns).~~Observations in coastal regions, where land emissions contaminate PMW SSTs, are excluded from the reference data sets.~~

	Satellites		**Drifting buoys**	
Experiment	RMSE	Bias	RMSE	Bias
Free model run	0.786	−0.514	~~0.770~~ 0.785	−~~0.606~~0.630
IR1	0.534	−0.224	~~0.406~~ 0.408	−~~0.173~~ 0.181
IR2	0.470	−0.154	~~0.377~~ 0.375	−0.143
PMW1	~~0.598~~ 0.602	−~~0.177~~ 0.187	~~0.402~~ 0.410	−~~0.088~~ 0.104
PMW2	~~0.612~~ 0.685	−~~0.237~~ 0.392	~~0.412~~ 0.505	−~~0.121~~ 0.314
COMB1	~~0.465~~ 0.464	−~~0.143~~ 0.142	~~0.368~~ 0.365	−~~0.115~~ 0.116
COMB2	~~0.468~~ 0.470	−~~0.152~~ 0.154	~~0.371~~ 0.367	−0.142

~~Applying~~ When applying the supermod operator ~~generally increases the errors, as can be seen by comparing PMW1 and PMW2~~to assimilate the PMW SSTs, both RMSEs and biases are increased in comparison with PMW1. ~~An explanation is~~ PMW2 does, however, validate better than the free model run. Compared with PMW1, we find that the increments in PMW2 are relatively weak~~compared to the increments in PMW1, which could be caused by both the observation thinning as well as how the supermod operator is designed to work (increasing the footprint results in a decrease in the increment amplitude). Weaker increments cause smaller adjustments of the background and, depending on the capability of the PMW SSTs to represent the true state, greater errors during validation~~. The mean increment of PMW1 and PMW2 during the first assimilation window~~in PMW1 and PMW2, where both experiments have similar~~ , where the experiments share the same initial conditions, demonstrate this increment difference: the mean increment in PMW1 is ∼~~0.48~~ 0.39 °C (increment standard deviation: ∼~~0.41~~ 0.49 °C), whereas PMW2 has a mean increment of ∼~~0.21~~ 0.04 °C (increment standard deviation: ∼~~0.15~~ 0.06 °C). That the increments are reduced in PMW2 is in accordance with how the supermod operator works. As demonstrated in Fig. 6, the amplitude of the increment decreases with increasing footprint size when the observation error remains unchanged.

The difference between the SST increment amplitudes in PMW1 and PMW2 is, in addition, visually illustrated in Fig. 7a and b~~, which show the increments inside a selected region in the model domain~~. Even though the number of observations in PMW2 ~~are thinned~~is significantly reduced compared to PMW1 due to thinning, the supermod operator is able to spread the impact of the observations over the footprint area ~~to create~~ yielding an increment pattern ~~which is similar to the pattern created~~ that overall resembles the pattern seen in PMW1. ~~However, the adjustments of the modeled SST are smaller in PMW2. That PMW2 is able to produce a similar increment pattern as~~ Both PMW1 ~~, only with reduced increment amplitudes, indicates that the information provided by the additional SSTs~~ and PMW2 are cooled in the North Sea in the southern part of the domain,

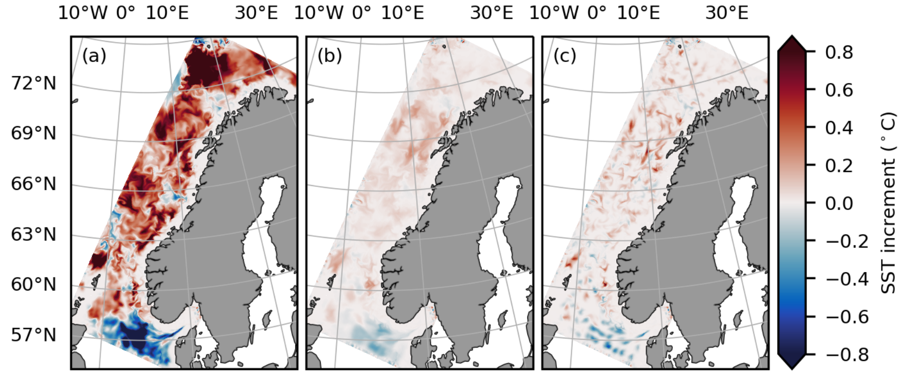

Figure 7. SST increments ~~during~~ at the ~~first~~ last time step in the first assimilation window~~shown for a selected region in the model domain~~. Increments ~~belong to~~ of experiments (a) PMW1 and (b) PMW2. (c) ~~shows the resulting~~ increments if PMW1 assimilates observations thinned in the same way as the observations used in PMW2.

while the larger part of the domain is heated. Also, a circular structure in the increments can be seen just north of 69°N in both panels. The most notable difference between the experiments is that the increments are weaker and more smooth in PMW2, which is what we would expect as the supermod operator is designed to adjust the average SST in an area corresponding to the footprint of the observations. Note that the high level of detail in the increments in PMW1 ~~is redundant. This redundancy can be a symptom of significant correlations in the PMW SSTs observation error, as demonstrated by Liu and Rabier (2002). Assimilating a thinned version~~ does not necessarily lead to a more detailed SST field in the analysis state. In many cases, detailed increments reflect that structures present in the background state have been strongly modified or even removed. Other notable differences between the experiments are the areas of strong, positive increments in PMW1 (north of 72°N and close to the coast around 63°N), which are not subject to enhanced heating in PMW2, as well as some cold patches in PMW1 that are not seen in PMW2. Furthermore, to separate the effect of the supermod operator and the observation thinning, we assimilate a thinned set of the PMW SSTs without applying the supermod operator~~is, however, not sufficient~~. We find that only thinning the observations, and not applying the supermod operator, is not sufficient in order to spread the increments over the observation footprint (Fig. 7c). The thinning itself does not handle the spatial resolution mismatch between the PMW SSTs and the model, and the supermod operator is required in order to spread the information from each observation over an area of similar size ~~as~~ to the observation footprint.

~~Both~~ When validating against satellite SSTs, the performance of COMB1 is found to be superior to that of IR2, while COMB2 is found to produce results that are identical to those of IR2. For the validation against SSTs from drifting buoys, both COMB1 and COMB2 ~~validate~~ perform better than IR2~~, which implies~~. These results imply that the PMW SST data set provides the model with information not seen by the IR SSTs and that the supermod operator is able to pass along this information ~~. To ensure that these better validations are statistically significant, we performed a Monte Carlo approach where we randomly selected half of the calculated SST differences (model minus observations) for experiments IR2, COMB1, and COMB2, followed by a calculation of the SST RMSE and bias for each experiment. This selection and the subsequent~~

~~calculations were repeated 1000 times. Our tests showed that the better validation of COMB1 and COMB2 compared to IR2 is statistically significant at the 99 % level. Furthermore, we~~ to some extent. We also find that COMB1 validates better than COMB2. This can partly be explained by the differences between the magnitudes of the increments resulting from assimilating the PMW SST observations, as discussed when comparing PMW1 and PMW2. ~~The better validation of~~ To verify that the addition of PMW SSTs to the assimilated data set indeed has a positive impact with regards to the validation against SSTs from drifting buoys, we tested for statistical significance as follows. For each of the experiments IR2, COMB1~~can also reflect that a smoothed version of a model field that resolves small-scale features generally validates better than the model field resolving these small-scale features. Small features captured by the model tend to have incorrect positions compared to the reference observations used for validation (Dagestad and Röhrs, 2019; Jacobs et al., 2021).~~ , and COMB2 every individual SST difference (model minus observation) were randomly assigned to one of 22 subsets. RMSE and bias were calculated for these unique subsets, and a Wilcoxon signed-rank test was used to test these RMSEs and biases for statistical significance. The tests showed that the differences between the RMSEs in COMB1, COMB2, and IR2 are statistically significant at the 99% confidence level. The same applies to the better bias in COMB1 compared to IR2. There is, however, no significant difference between the biases calculated from COMB2 and IR2.

~~We perform two additional validations to test if~~ To investigate whether the improved error statistics of COMB1 and COMB2 can be explained by the ~~additional~~ extra information PMW SSTs ~~bring~~ provide to the model when IR SSTs are unavailable~~. First, we validate the background during the~~ , error statistics are now calculated separately for cloudy and clear-sky conditions. The validation of SST during cloudy conditions is calculated using only reference satellite SSTs within regions where IR SSTs were unavailable during the previous analysis cycle. Similarly, only reference satellite SSTs in regions where there were clear skies in the same analyses are used for validation of clear-sky conditions. As we use the background model states for validation purposes regions that experienced heavy cloud cover and thus poor coverage of IR SSTs during the previous assimilation cycle are not necessarily poorly sampled by IR SST during the period ~~6 May 2018–16 May 2018. Then, validation is performed during the cloudy period 7 June 2018–17 June 2018.~~ used for validation. The resulting error statistics are shown in Table 4. We find that ~~the cloudy period has an RMSE of 0.443 °C and 0.428 °C in IR2~~ both COMB1 and COMB2 ~~, respectively. The bias is −0.110 °C in IR2, while it is slightly reduced to −0.098 °C in COMB2. Conversely, the clear-sky period validation shows that both the RMSEs and biases are similar in~~ perform better than IR2 during cloudy conditions, both in terms of RMSE and bias. Using the same method as before to calculate the statistical significance, we find that the lower RMSE and bias in COMB1 and COMB2 ~~, differing only by 0.001 °C. Similar results are found when comparing~~ compared to IR2 are statistically significant. There is, however, no significant improvement in COMB2 over IR2 ~~and~~ during clear-sky conditions. The improvement seen in COMB1 ~~; that is , the RMSEs and biases differ the most during the cloudy period.~~ is also more pronounced during cloudy conditions than during periods with good IR SST coverage, which indicates that PMW SSTs indeed provide valuable extra information in the absence of IR SSTs.

Figure 8 illustrates how PMW observations compensate for IR SST data deficiency. ~~SST observations~~ SSTs assimilated in IR2 during the assimilation cycle covering 8 June 2018–10 June 2018 are shown in Fig. 8a, while the SST increments ~~from the beginning~~ at the last time step of this cycle ~~at 8 June 2018~~ are shown in Fig. 8b. Figure 8e shows the increments inside the black

Table 4. SST RMSE (°C) and bias (°C) calculated using satellite SSTs as reference. Validation performed during cloudy (left column) and clear-sky conditions (right column). Number of reference observations is $\sim 1.7 \times 10^6$ (cloudy) and $\sim 3.5 \times 10^7$ (clear-sky).

	Cloudy		**Clear-sky**	
Experiment	RMSE	Bias	RMSE	Bias
IR2	0.401	−0.081	0.530	−0.188
COMB1	0.346	−0.004	0.523	−0.169
COMB2	0.393	−0.075	0.529	−0.187

box in Fig. 8b. The presence of clouds limits the IR SST coverage, particularly in the southwestern part of the domain. A cluster of IR SST observations around the Shetland Islands causes a cooling of the model. However, the general lack of observations in this region leaves behind a larger area with no adjustments to the SST field. For COMB2, the assimilated observations and the increments are shown in Fig. 8c and d, respectively. ~~We notice~~ The increments inside the black box in Fig. 8d are shown in Fig. 8f. Notice that the regions without IR SST observations in Fig. 8a are now covered by PMW SSTs. This additional information to the model causes the SST field to adjust such that a larger area is cooled.

4.3 SST ~~variance~~ power spectra

Spectral analysis is a technique that can be used to decompose the information contained in an observed or modeled field into different spatial scales. Specifically, spectra can be computed for modeled or observed SST fields to evaluate the spatial resolution of the SST structures contained within the field (e.g., Reynolds and Chelton, 2010; Brasnett and Colan, 2016; Castro et al., 2017; Pearson et al., 2019; Schubert et al., 2019; Janeković et al., 2022). The reason for this is that a spectrum holds information about the SST ~~gradients~~ variability at different spatial scales. ~~A~~ Hence, a reduction in the spectral density at ~~higher frequency scales (i.e., shorter wavelengths) thus reflects that the SST gradients~~ shorter wavelengths reflects that SST structures at these scales have ~~become weaker~~been dampened, i.e., that the field has become more smooth.

~~Usually, the~~ SST data are often decomposed into spectral space using the discrete Fourier transform (DFT), a transformation that requires the data to be periodic. Periodicity can be retained by spatially detrending the data or by windowing, i.e., multiplying the field by a function such that the interior of the domain retains its structures while the boundaries drop off and approach zero. ~~A~~ As noted by Denis et al. (2002), a disadvantage of detrending and windowing is that these methods ~~contaminate the largest scales in the spectral analysis (Denis et al., 2002). We use~~ modify the resulting spectrum by removing important information in the original data. To avoid these problems, they propose an alternative to the DFT~~to avoid these problems~~, namely the discrete cosine transform (DCT)~~, where~~. We have chosen to use this method for the spectral analysis performed in this study. When applying the DCT, the input fields are made periodic in space by mirroring the fields prior to the transformation. The reader is referred to Denis et al. (2002) for a thorough description of the DCT.

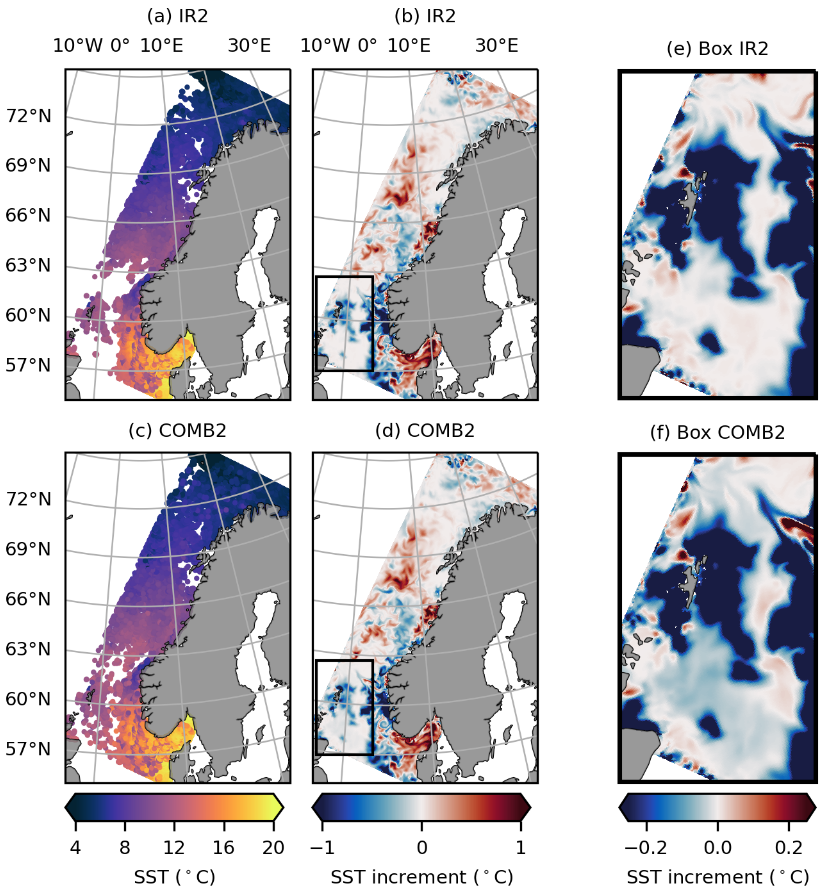

Figure 8. SST observations assimilated in (a) IR2 and (c) COMB2 during the assimilation cycle covering 8 June 2018–10 June 2018, and ~~average~~ SST increments ~~at 8 June 2018~~ in (b) IR2 and (d) COMB2 at the last time step of this assimilation cycle. A zoom in of the increments inside the black box in (b) and (d) is shown in (e) and (f), respectively. Note the change in the scale of the colorbar. ~~We suspect that the "blob" of warming between 58–59° N and 1–3° E in Fig. 8d arises from unwanted SST artifacts in the PMW SST data set. These artifacts are likely caused by radio-frequency interference from oil platforms (Alerskans et al., 2020). Careful consideration of the PMW SSTs revealed that it is not sufficient to discard all but the quality level 5 data to avoid erroneous data. Thus, we will have to implement an additional quality-control which is based on the positions of oil platforms to use this PMW SST data set in the future.~~

For each time step of the background, we apply the DCT to the two-dimensional SST field inside the dashed box shown in Fig. 2. This transformation into spectral space, as well as the subsequent calculation of the two-dimensional spectral variance array, ~~follow~~ follows the methodology described in Denis et al. (2002). The spectral ~~variances are distributed between their corresponding wavelengths using the binning process presented in Ricard et al. (2013), so we end up with a~~ variance array consists of elements, $\sigma^2(m,n)$, where m and n are the adimensional wavenumber axes (i.e., the grid cell indexes in both horizontal directions), and each of these variance elements is associated with a normalized wavenumber,

$$\kappa = \sqrt{\frac{m^2}{N_i^2} + \frac{n^2}{N_j^2}}, \tag{3}$$

where N_i and N_j are the domain's total number of grid cells in each horizontal direction. The wavenumbers are adimensional, but they can be converted into wavelengths through the following equation,

$$\lambda = \frac{2\Delta}{\kappa}, \tag{4}$$

where Δ is the grid cell spacing. All individual spectral variance elements, $\sigma^2(m,n)$, are distributed between wavelength bins to create a one-dimensional SST ~~variance spectrum. The SST variance~~ power spectrum. As described in Denis et al. (2002), the variance elements that contribute to wavelength bin κ are those variances that fall within the wavelength band κ and $\kappa + \Delta\kappa$. However, we have modified this approach according to Ricard et al. (2013) such that the variances within a wavelength band are proportionally distributed between the two bounding wavelength bins based on the variances' proximity to the bins. The one-dimensional SST power spectra calculated for each time step are subsequently averaged in time to provide a mean SST ~~variance spectrum.~~

~~Upper panel shows SST variance spectra with wavelengths ranging from 278.4 km (length of the shortest edge of the domain used to calculate spectra) to 4.8 km (double the grid spacing). Shading is applied to indicate the 95 % confidence interval on the mean spectrum and is calculated using the jackknife method. Notice that the shading is hard to detect due to the smallness of the confidence interval. Inset zoom shows the spectra at wavelengths ranging from 120 km to 30 km. The lower panel shows the ratio of the experiments' spectra to the spectrum from IR2. A gray dashed line is drawn at 60 km, which is approximately the size of the major axis of the elliptic PMW SST footprint.~~ power spectrum. Notice that while the mean SST power spectra are functions of the adimensional wavenumber, we will refer to the wavelength calculated from Eq. 4 when we present the results.

Figure 9 shows the mean SST ~~variance~~ power spectra calculated for each experiment (except IR1). For wavelengths smaller than ∼120 km, we find that the spectrum from PMW1 has ~~a significantly smaller variance~~ significantly lower power than the spectrum calculated from IR2 as well as that of the free model run (not shown). Assimilating PMW SSTs without using the supermod operator ~~has thus~~ thus has a smoothing effect on the modeled SST, indicating that the average effect of the high level of detail seen in the increments (Fig. 7) is a removal or smoothing of structures present in the background state. This smoothing effect is also present in the spectrum calculated from COMB1, where IR SSTs are added to the assimilated data set: the spectrum from COMB1 follows that of PMW1 at most wavelengths and is more or less similar to PMW1's spectrum at wavelengths in the range ∼35–120 km. This result is within expectations, as PMW SSTs represent a mean value over a large footprint of the actual SST field. With a traditional observation operator, this mean value is compared to individual model grid points, and any small-scale deviation from this mean value in the model is thus damped in the analysis.

The spectrum from PMW2 has more ~~variance~~ power than that from ~~IR2~~ PMW1 at all scales smaller than ∼~~200 kmand follows the spectrum of a~~ 120 km, which means that PMW2 resolves more SST structures at these scales. The peak in the

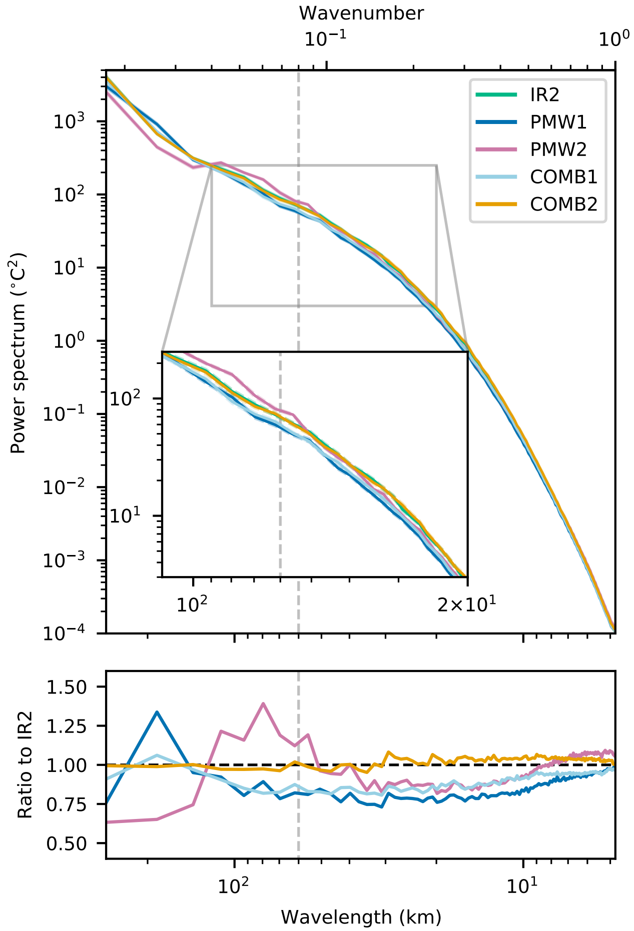

Figure 9. Upper panel shows SST power spectra with wavelengths ranging from 278.4 km (length of the shortest edge of the domain used to calculate spectra) to 4.8 km (double the grid spacing). Shading is applied to indicate the 95% confidence interval on the mean spectrum and is calculated using the jackknife method. Notice that the shading is hard to detect due to the narrowness of the confidence interval. Inset zoom shows the spectra at wavelengths ranging from 120 km to 20 km. The lower panel shows the ratio of the experiments' spectra to the spectrum from IR2. A gray dashed line is drawn at 60 km, which is approximately the size of the major axis of the elliptic PMW SST footprint.

ratio between PMW2 and IR2, which is found at ~80 km, is also present in the ratio between the free model run and IR2 (not shown). ~~Furthermore, the variance~~ In fact, we find that the spectrum of PMW2 follows the spectrum of the free model run at most spatial scales, but with less power at all spatial scales. Furthermore, we see in Fig. 9 that the power spectrum from COMB2 is more or less identical to that of IR2 at all spatial scales. Even with lower observation errors for the PMW SSTs (decreasing the α) and thus stronger increments, COMB2 stays similar to IR2 with ~1 as the ratio between COMB2's and IR2's spectrum at all spatial scales (not shown). These findings suggest that using the supermod operator prevents ~~the gradients~~

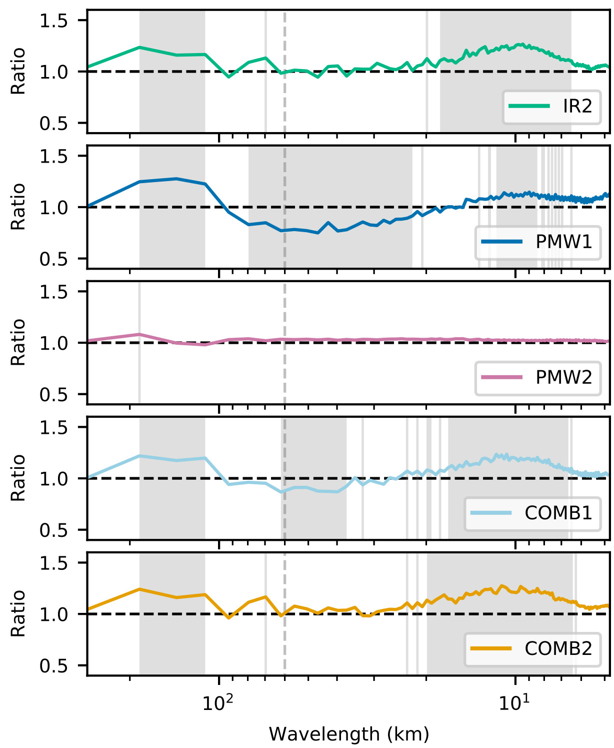

Figure 10. Ratio of the analysis spectrum to the background spectrum for each experiment. Gray shading is applied at wavelengths where there is no overlap of the 95% confidence intervals of the mean analysis spectrum and the mean background spectrum. Gray dashed line drawn at 60 km.

smoothing of small-scale structures present in the field~~from being smoothed~~: as PMW SSTs are now compared to mean model values, they do not penalize variations of the SST at spatial scales they do not resolve.

SST ~~variance spectra were~~ power spectra are also calculated for each experiment using the last day of the analyses from each analysis cycle~~(not shown). The~~. These spectra were compared ~~to~~ with the corresponding background spectra to examine if the assimilation changes the spatial scales of the background SST structures within each experiment. Figure 10 shows the ratio of the analysis spectrum to the background spectrum for each experiment. We find that the analyses in PMW1 experience a loss of SST structures with spatial scales of ∼~~10–100~~ 20–80 km and that the analyses in COMB1 have ~~less~~ fewer SST structures at scales of ∼~~50–70~~ 40–60 km. The other experiments' analyses do not experience a significant loss due to the assimilation of the observations.

5 Discussion

Our results suggest that it is beneficial to assimilate the PMW SSTs ~~should be assimilated~~ in conjunction with the IR SSTs in order to ~~lower~~ reduce the errors in ~~the~~ modeled SST. This is verified by comparing the error statistics of COMB1 and COMB2 to IR2. The comparison of error statistics during clear-sky and cloudy conditions ~~imply that the error improvements~~ suggests that the reduced errors in COMB1 and COMB2 in large part originate from the additional information given to the model ~~in cloudy regions where~~ during cloudy conditions when there is a shortage of IR SSTs. Figure 8, which shows the SST increments in such cloudy regions, demonstrates the supermod operator's ability to pass along the relevant information seen by the PMW SSTs. ~~The~~ Here, the assimilated PMW SST observations cool the model SST in a region of sparse IR SST coverage. ~~While the~~ The few available IR SSTs in that region confirm this cooling~~, they~~. However, these IR SSTs are not sufficient to correct the SST of the whole region when assimilated on their own.

While error statistics show that COMB1 validates better than COMB2, we find that COMB1 ~~smooths SST structures. Hence, the smaller SST RMSE and bias of COMB1 do not reflect that this experiment is a better quality product~~yields smoother SST fields than what is found in IR2. This indicates that neglecting to account for the mismatch in spatial scales when assimilating PMW SSTs may be disadvantageous. The smoothing of COMB1 is demonstrated by the calculated SST ~~variance~~ power spectra. We find that both PMW1 and COMB1 return smoother SST fields than IR2 at spatial scales smaller than $\sim$120 km. This upper limit is approximately twice the size of the major axis of the elliptic PMW SST footprint. Thus, the limit corresponds to the expected effective resolution of the SST structures resolved by the PMW SST observations. It is striking that the affected spatial scales in PMW1 and COMB1 correspond to scales smaller ~~than~~ or similar to this effective resolution. Furthermore, ~~that the~~ the fact that additional information from the IR SSTs in COMB1 does not prevent ~~the smoothingreveals~~ smoothing, indicates that the PMW SSTs ~~greatly impact the final reanalysis~~ have a substantial impact on the final product. However, assimilating the PMW SSTs through the supermod operator, which was performed in ~~PMW2 and~~ COMB2, does not result in a smoothing of the SST structures. This demonstrates that the operator is a good alternative approach to assimilate observations of coarser spatial resolution than the model.

~~Finally, all of the experiments cover the~~ As demonstrated in Sect. 3.1, the amplitude of the increments decreases with increasing footprint size when the supermod operator is applied if the observation errors are kept unchanged. A suitable choice for the observation error covariance is important for the overall performance of all assimilation algorithms. To assess the observation errors prescribed to the different products assimilated in this study, we have applied the diagnosis proposed by Desroziers et al. (2005). The results indicate that the chosen observation errors are too high for all assimilated products. For PMW SSTs, we see the same tendency of overestimation both with and without the supermod operator activated. By setting $\alpha = 0.8$ to reduce the observation errors, we obtain improved error statistics for PMW2 without any indication of overfitting to the observations, whereas the results for PMW1 barely change (not shown). However, for PMW1 this lower error introduces artifacts in the modeled SST fields. These artifacts are similar to those that can be seen in the PMW observations, which may be caused by the large footprint size as well as the regridding and interpolation during the processing. Due to this degradation of PMW1, and to be able to fairly assess all experiments, the observational error was constructed using the higher $\alpha = 2$ for all experiments shown in this study. Moreover, the assimilation of PMW SSTs using the supermod operator could potentially be improved by optimization of the prescribed footprint size. If a smaller footprint could be applied without smoothing the modeled SST, this could both yield larger increments and increase the number of observations available for assimilation. Furthermore, the effect of a less strict thinning of observations could be explored.

Finally, all of the experiments were during local spring, which is a period when the modeled SST undergoes great changes. Such changes make it challenging to sustain high-skilled forecasts, and we chose ~~the~~ this period due to these challenges. However, the chosen period is not heavily affected by clouds. The impact of assimilating PMW SSTs is possibly greater during winter when the oceanic regions along the Norwegian coastline experience high cloud coverage.

6 Conclusions

Correcting the satellite SSTs for biases through the implemented bias correction scheme improves the modeled SST. The bias correction scheme is easy to implement and apply since it is separate from the data assimilation process.

While assimilating IR SSTs reduces the modeled SST errors, an additional reduction is achieved if PMW SSTs are assimilated in conjunction with the IR SSTs. This error reduction is mainly caused by the information the PMW SSTs provide in cloudy regions. However, if we assimilate the PMW SSTs without considering their large footprint sizes, we end up smoothing the modeled SST structures of spatial scales smaller than twice the PMW SST footprint. By introducing the supermod operator, we have shown that the PMW SST observations can be assimilated into the ocean model without causing any spatial smoothing of the modeled SST. Our supermod operator is easy to implement and can be used to assimilate other observation variables having a coarser spatial resolution than the resolution of the model.

Code and data availability. Supermod operator code is available at: https://github.com/siljeci/ROMS_supermod. Model output from each data assimilation experiment is available at: https://thredds.met.no/thredds/projects/supermodop.html.

Author contributions. SCI: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft preparation, Writing - review & editing. AKS: Conceptualization, Data curation, Methodology, Project administration, Supervision, Validation, Writing - review & editing. OG: Formal analysis, Software, Validation, Writing - review & editing.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. This research was funded by the Centre of Integrated Remote Sensing and Forecasting for Arctic Operations (CIRFA) under the Research Council of Norway (RCN), grant number 237906. Data from the European Organisation for the Exploitation of Me-

teorological Satellites (EUMETSAT) Satellite Application Facility on Ocean and Sea Ice (OSI SAF) are accessible through https://osi-saf.eumetsat.int. Data were also provided by Group for High Resolution Sea Surface Temperature (GHRSST), the National Oceanic and Atmospheric Administration (NOAA) and Remote Sensing Systems (RSS). These data can be downloaded from he Physical Oceanography Distributed Active Archive Center (PO.DAAC). Data were also created and distributed by EUMETSAT.

References

Alerskans, E., Høyer, J. L., Gentemann, C. L., Pedersen, L. T., Nielsen-Englyst, P., and Donlon, C.: Construction of a climate data record of sea surface temperature from passive microwave measurements, Remote Sens. Environ., 236, 111 485, https://doi.org/10.1016/j.rse.2019.111485, 2020.

Beldring, S., Engeland, K., Roald, L., Sælthun, N., and Voksø, A.: Estimation of parameters in a distributed precipitation-runoff model for Norway, Hydrol. Earth Syst. Sci., 7, 304–316, https://doi.org/10.5194/hess-7-304-2003, 2003.

Brasnett, B. and Colan, D. S.: Assimilating Retrievals of Sea Surface Temperature from VIIRS and AMSR2, J. Atmos. Ocean Tech., 33, 361–375, https://doi.org/10.1175/JTECH-D-15-0093.1, 2016.

Buehner, M., Caya, A., Pogson, L., Carrieres, T., and Pestieau, P.: A New Environment Canada Regional Ice Analysis System, Atmos. Ocean, 51, 18–34, https://doi.org/10.1080/07055900.2012.747171, 2013.

Cabanes, C., Grouazel, A., von Schuckmann, K., Hamon, M., Turpin, V., Coatanoan, C., Paris, F., Guinehut, S., Boone, C., Ferry, N., de Boyer Montégut, C., Carval, T., Reverdin, G., Pouliquen, S., and Le Traon, P.-Y.: The CORA dataset: validation and diagnostics of in-situ ocean temperature and salinity measurements, Ocean Sci., 9, 1–18, https://doi.org/10.5194/os-9-1-2013, 2013.

Castro, S., Emery, W., Wick, G., and Tandy, W.: Submesoscale Sea Surface Temperature Variability from UAV and Satellite Measurements, Remote Sens., 9, 1089, https://doi.org/10.3390/rs9111089, 2017.

Chan, P.-K. and Gao, B.-C.: A Comparison of MODIS, NCEP, and TMI Sea Surface Temperature Datasets, IEEE Geosci. Remote S., 2, 270–274, https://doi.org/10.1109/LGRS.2005.846838, 2005.

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

Denis, B., Côté, J., and Laprise, R.: Spectral Decomposition of Two-Dimensional Atmospheric Fields on Limited-Area Domains Using the Discrete Cosine Transform (DCT), Mon. Weather Rev., 130, 1812–1829, https://doi.org/10.1175/1520-0493(2002)130<1812:SDOTDA>2.0.CO;2, 2002.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Q. J. Roy. Meteor. Soc., 131, 3385–3396, https://doi.org/10.1256/qj.05.108, 2005.

Donlon, C. J., Minnett, P. J., Gentemann, C., Nightingale, T. J., Barton, I. J., Ward, B., and Murray, M. J.: Toward Improved Validation of Satellite Sea Surface Skin Temperature Measurements for Climate Research, J. Climate, 15, 353–369, https://doi.org/10.1175/1520-0442(2002)015<0353:TIVOSS>2.0.CO;2, 2002.

Donlon, C. J., Robinson, I., Casey, K. S., Vazquez-Cuervo, J., Armstrong, E., Arino, O., Gentemann, C., May, D., LeBorgne, P., Piollé, J., Barton, I., Beggs, H., Poulter, D. J. S., Merchant, C. J., Bingham, A., Heinz, S., Harris, A., Wick, G., Emery, B., Minnett, P., Evans, R., Llewellyn-Jones, D., Mutlow, C., Reynolds, R. W., Kawamura, H., and Rayner, N.: The Global Ocean Data Assimilation Experiment High-resolution Sea Surface Temperature Pilot Project, B. Am. Meteorol. Soc., 88, 1197–1214, https://doi.org/10.1175/BAMS-88-8-1197, 2007.

Donlon, C. J., Casey, K., Gentemann, C., LeBorgne, P., Robinson, I., Reynolds, R. W., Merchant, C., Llewellyn-Jones, D., Minnett, P. J., Piolle, J. F., Cornillon, P., Rayner, N., Brandon, T., Vazquez, J., Armstrong, E., Beggs, H., Barton, I., Wick, G., Castro, S., Høyer, J., May, D., Arino, O., Poulter, D. J., Evans, R., Mutlow, C. T., Bingham, A. W., and Harris, A.: Successes and Challenges for the Modern Sea Surface Temperature Observing System, in: Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society, vol. 2, pp. 249–258, ESA Publication WPP-306, Venice, Italy, https://doi.org/10.5270/OceanObs09.cwp.24, 2010.

Donlon, C. J., Martin, M.and Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W.: The Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system, Remote Sens. Environ., 116, 140–158, https://doi.org/10.1016/j.rse.2010.10.017, 2012.

Eastwood, S., Le Borgne, P., Péré, S., and Poulter, D.: Diurnal variability in sea surface temperature in the Arctic, Remote Sens. Environ., 115, 2594–2602, https://doi.org/10.1016/j.rse.2011.05.015, 2011.

ECMWF: ECMWF IFS CY41r2 High-Resolution Operational Forecasts, Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory [data set], https://doi.org/10.5065/D68050ZV, accessed 11 March 2020, 2016.

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.

Fairall, C. W., Bradley, E. F., Hare, J. E., Grachev, A. A., and Edson, J. B.: Bulk parameterization of air–sea fluxes: Updates and verification for the COARE algorithm, J. Climate, 16, 571–591, https://doi.org/10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2, 2003.

Gürol, S., Weaver, A. T., Moore, A. M., Piacentini, A., Arango, H. G., and Gratton, S.: B-preconditioned minimization algorithms for variational data assimilation with the dual formulation, Q. J. Roy. Meteor. Soc., 140, 539–556, https://doi.org/10.1002/qj.2150, 2014.

Haidvogel, D. B., Arango, H., Budgell, W. P., Cornuelle, B. D., Curchitser, E., Di Lorenzo, E., Fennel, K., Geyer, W. R., Hermann, A. J., Lanerolle, L., Levin, J., McWilliams, J. C., Miller, A. J., Moore, A. M., Powell, T. M., Shchepetkin, A. F., Sherwood, C. R., Signell, R. P., Warner, J. C., and Wilkin, J.: Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the Regional Ocean Modeling System, J. Comput. Phys., 227, 3595–3624, https://doi.org/10.1016/j.jcp.2007.06.016, 2008.

Høyer, J. L., Karagali, I., Dybkjær, G., and Tonboe, R.: Multi sensor validation and error characteristics of Arctic satellite sea surface temperature observations, Remote Sens. Environ., 121, 335–346, https://doi.org/10.1016/j.rse.2012.01.013, 2012.

Høyer, J. L., Le Borgne, P., and Eastwood, S.: A bias correction method for Arctic satellite sea surface temperature observations, Remote Sens. Environ., 146, 201–213, https://doi.org/10.1016/j.rse.2013.04.020, 2014.

Imaoka, K., Kachi, M., Kasahara, M., Ito, N., Nakagawa, K., and Oki, T.: Instrument performance and calibration of AMSR-E and AMSR2, Int. Arch. Photogramm., 38, 13–18, 2010.

Jacobs, G. A., D'Addezio, J. M., Bartels, B., and Spence, P. L.: Constrained scales in ocean forecasting, Adv. Space Res., 68, 746–761, https://doi.org/10.1016/j.asr.2019.09.018, 2021.

Janeković, I., Rayson, M., Jones, N., Watson, P., and Pattiaratchi, C.: 4D-Var data assimilation using satellite sea surface temperature to improve the tidally-driven interior ocean dynamics estimates in the Indo-Australian Basin, Ocean Model., 171, 101 969, https://doi.org/10.1016/j.ocemod.2022.101969, 2022.

Janjić, T., Bormann, N., Bocquet, M., Carton, J. A., Cohn, S. E., Dance, S. L., Losa, S. N., Nichols, N. K., Potthast, R., Waller, J. A., and Weston, P.: On the representation error in data assimilation, Q. J. Roy. Meteor. Soc., 144, 1257–1278, https://doi.org/10.1002/qj.3130, 2018.

Jia, C. and Minnett, P. J.: High latitude sea surface temperatures derived from MODIS infrared measurements, Remote Sens. Environ., 251, 112 094, https://doi.org/10.1016/j.rse.2020.112094, 2020.

Karagali, I., Høyer, J., and Hasager, C.: SST diurnal variability in the North Sea and the Baltic Sea, Remote Sens. Environ., 121, 159–170, https://doi.org/10.1016/j.rse.2012.01.016, 2012.

Kumar, A., Minnett, P. J., Podestá, G., and Evans, R. H.: Error Characteristics of the Atmospheric Correction Algorithms Used in Retrieval of Sea Surface Temperatures from Infrared Satellite Measurements: Global and Regional Aspects, J. Atmos. Sci., 60, 575–585, https://doi.org/10.1175/1520-0469(2003)060<0575:ECOTAC>2.0.CO;2, 2003.

Liu, Z.-Q. and Rabier, F.: The interaction between model resolution, observation resolution and observation density in data assimilation: A one-dimensional study, Q. J. Roy. Meteor. Soc., 128, 1367–1386, https://doi.org/10.1256/003590002320373337, 2002.

Luo, B., Minnett, P. J., Szczodrak, M., Kilpatrick, K., and Izaguirre, M.: Validation of Sentinel-3A SLSTR derived Sea-Surface Skin Temperatures with those of the shipborne M-AERI, Remote Sens. Environ., 244, 111 826, https://doi.org/10.1016/j.rse.2020.111826, 2020.

Martin, M. J., King, R. R., While, J., and Aguiar, A. B.: Assimilating satellite sea-surface salinity data from SMOS, Aquarius and SMAP into a global ocean forecasting system, Q. J. Roy. Meteor. Soc., 145, 705–726, https://doi.org/10.1002/qj.3461, 2019.

Merchant, C. J., Embury, O., Le Borgne, P., and Bellec, B.: Saharan dust in nighttime thermal imagery: Detection and reduction of related biases in retrieved sea surface temperature, Remote Sens. Environ., 104, 15–30, https://doi.org/10.1016/j.rse.2006.03.007, 2006.

Mile, M., Randriamampianina, R., Marseille, G., and Stoffelen, A.: Supermodding – A special footprint operator for mesoscale data assimilation using scatterometer winds, Q. J. Roy. Meteor. Soc., 147, 1382–1402, https://doi.org/10.1002/qj.3979, 2021.

Minnett, P. J.: Sea Surface Temperature, in: Encyclopedia of Remote Sensing, edited by Njoku, E. G., pp. 754–759, Springer New York, https://doi.org/10.1007/978-0-387-36699-9_166, 2014.

Minnett, P. J., Alvera-Azcárate, A., Chin, T. M., Corlett, G. K., Gentemann, C. L., Karagali, I., Li, X., Marsouin, A., Marullo, S., Maturi, E., Santoleri, R., Saux Picart, S., Steele, M., and Vazquez-Cuervo, J.: Half a century of satellite remote sensing of sea-surface temperature, Remote Sens. of Environ., 233, 111 366, https://doi.org/10.1016/j.rse.2019.111366, 2019.

Moore, A. M., Arango, H. G., Broquet, G., Powell, B. S., Weaver, A. T., and Zavala-Garay, J.: The Regional Ocean Modeling System (ROMS) 4-dimensional variational data assimilation systems. Part I - System overview and formulation, Prog. Oceanogr., 91, 34–49, https://doi.org/10.1016/j.pocean.2011.05.004, 2011.

O'Carroll, A. G., Armstrong, E. M., Beggs, H. M., Bouali, M., Casey, K. S., Corlett, G. K., Dash, P., Donlon, C. J., Gentemann, C. L., Høyer, J. L., Ignatov, A., Kabobah, K., Kachi, M., Kurihara, Y., Karagali, I., Maturi, E., Merchant, C. J., Marullo, S., Minnett, P. J., Pennybacker, M., Ramakrishnan, B., Ramsankaran, R., Santoleri, R., Sunder, S., Saux Picart, S., Vázquez-Cuervo, J., and Wimmer, W.: Observational Needs of Sea Surface Temperature, Front. Mar. Sci., 6, 420, https://doi.org/10.3389/fmars.2019.00420, 2019.

Pearson, K., Good, S., Merchant, C. J., Prigent, C., Embury, O., and Donlon, C.: Sea Surface Temperature in Global Analyses: Gains from the Copernicus Imaging Microwave Radiometer, Remote Sens., 11, 2362, https://doi.org/10.3390/rs11202362, 2019.

Reynolds, R. W. and Chelton, D. B.: Comparisons of Daily Sea Surface Temperature Analyses for 2007–08, J. Climate, 23, 3545–3562, https://doi.org/10.1175/2010JCLI3294.1, 2010.

Ricard, D., Lac, C., Riette, S., Legrand, R., and Mary, A.: Kinetic energy spectra characteristics of two convection-permitting limited-area models AROME and Meso-NH: KE Spectra Characteristics of Two Convection-Permitting Models, Q. J. Roy. Meteor. Soc., 139, 1327–1341, https://doi.org/10.1002/qj.2025, 2013.

Röhrs, J., Sperrevik, A. K., and Christensen, K. H.: NorShelf: An ocean reanalysis and data-assimilative forecast model for the Norwegian Shelf Sea, Tech. Rep. ISSN 2387-4201 04/2018, Norwegian Meteorological Institute, Oslo, Norway, https://www.met.no/publikasjoner/met-report, 2018.

Schubert, R., Schwarzkopf, F. U., Baschek, B., and Biastoch, A.: Submesoscale Impacts on Mesoscale Agulhas Dynamics, J. Adv. Model. Earth Sy., 11, 2745–2767, https://doi.org/10.1029/2019MS001724, 2019.

Shchepetkin, A. F. and McWilliams, J. C.: Correction and commentary for "Ocean forecasting in terrain-following coordinates: Formulation and skill assessment of the regional ocean modeling system" by Haidvogel et al., J. Comp. Phys. 227, pp. 3595–3624, J. Comput. Phys., 228, 8985–9000, https://doi.org/10.1016/j.jcp.2009.09.002, 2009.

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.

Vincent, R. F.: The Effect of Arctic Dust on the Retrieval of Satellite Derived Sea and Ice Surface Temperatures, Sci. Rep., 8, 9727, https://doi.org/10.1038/s41598-018-28024-6, 2018.

Wang, H., Guan, L., and Chen, G.: Evaluation of Sea Surface Temperature From FY-3C VIRR Data in the Arctic, IEEE Geosci. Remote S., 13, 292–296, https://doi.org/10.1109/LGRS.2015.2511184, 2016.

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.

Waters, J., Lea, D. J., Martin, M. J., Mirouze, I., Weaver, A., and While, J.: Implementing a variational data assimilation system in an operational 1/4 degree global ocean model, Q. J. Roy. Meteor. Soc., 141, 333–349, https://doi.org/10.1002/qj.2388, 2015.

Weaver, A. and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, Q. J. Roy. Meteor. Soc., 127, 1815–1846, https://doi.org/10.1002/qj.49712757518, 2001.

Wentz, F. J., Gentemann, C., Smith, D., and Chelton, D.: Satellite Measurements of Sea Surface Temperature Through Clouds, Science, 288, 847–850, https://doi.org/10.1126/science.288.5467.847, 2000.

While, J. and Martin, M. J.: Variational bias correction of satellite sea-surface temperature data incorporating observations of the bias, Q. J. Roy. Meteor. Soc., 145, 2733–2754, https://doi.org/10.1002/qj.3590, 2019.

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.