Assimilation of heterogeneous measurements at different spatial scales in the Arctic ocean and Norwegian sea

Daniel Ørnes Halvorsen*, Alberto Dallolio, Morten Omholt Alver
Department of Engineering Cybernetics,
Norwegian University of Science and Technology
Trondheim, Norway
Email: *daniel.o.halvorsen@ntnu.no

***Abstract*—The assimilation of ocean temperature measurements into ocean models provides useful insights on how to design heterogeneous ocean observation systems. In systems of this kind, ocean models can be complemented with multiscale operational assets such as satellites and in-situ unmanned vehicles. In this article, the authors simulate three different ocean model domains with horizontal resolutions of $20\,\mathrm{km}$, $4\,\mathrm{km}$ and $800\,\mathrm{m}$. The assimilated data sets are a global observation product including sea surface temperature, a vertical temperature profile measurement data set from the Norwegian Sea, and sea surface temperature measurements from an unmanned surface vehicle operating in the coastal waters of Frohavet (Central Norway). The key outcomes of the study suggest that global covering data sets should be assimilated in coarse model domains when available, while the intermediate and local data sets can be assimilated if they are covering areas of specific interest, and can be omitted otherwise.**

***Index Terms*—Ocean modeling, Data assimilation, Multiscale observational systems**

I. Introduction

In a time of unprecedented focus on the ocean, monitoring and forecasting of the aquatic environment are becoming increasingly important in order to understand the impact that human activities such as fisheries, aquaculture and shipping have on the environment. Commonly, studies of this kind make use of ocean models of varying complexity and spatio-temporal resolution, as well as local or global observations of the ocean state. Ocean models are usually subject to errors, either by simplifications in the mathematical description of real processes, or by uncertainties in relevant model parameters and inputs [1]. Additionally, also ocean observations are susceptible to measurement errors. Combined, the ocean models and observations form a synergistic toolbox for a more correct and reliable estimate of the ocean state. The combination is usually done by methods for data assimilation and, in this study, the Ensemble Kalman Filter (EnKF) is used for model corrections [2].

Today, ocean studies making use of data assimilation techniques benefit from an increasingly large pool of available measurement, including both global coverage data sets from organizations such as the Copernicus Marine Environmental Monitoring Service (CMEMS), and local coverage sets obtained through sporadic field campaigns [3]. In this article, the authors investigate the effect of assimilating ocean observations of different spatial coverage into ocean model domains of varying spatial and temporal resolution. The objective is to determine how to design an ocean observing system given an area of interest and available multiscale operational assets.

The observation data used in this study is a combination of publicly available data sets from CMEMS and Met Office Hadley Centre for Climate Science and Services [4]–[7], in addition to data from an ocean observation field campaign conducted by the Norwegian University of Science and Technology (NTNU) in April, 2022, in the semi-enclosed sea named Frohavet (Central Norway). The purpose of the field campaign was to conduct an initial test of the "observational pyramid", which is an observation system consisting of a small satellite, aerial drones, Autonomous Surface Vehicles (ASVs) and Autonomous Underwater Vehicles (AUVs), together with ground proofing from biologists [8]. In other words, a tool to study an ocean phenomenon on all spatial and temporal scales at the same time. When conducting such large-scale field campaigns with multiscale operational assets making heterogeneous measurements of the same physical or biological quantity, an understanding of when and where to assimilate the different observations into the ocean model for state estimation and model correction is necessary, and is therefore the purpose of this study. Similar work on observation systems for the ocean can be found in [9]–[13].

The manuscript is organized as follows. In Section II an overview of the theory and methods is given. First, the ocean model SINMOD is presented along with the three computational domains of interest. Then, a presentation of the EnKF equations as well as considerations taken when implementing the filter is provided. Three sources of ocean measurements are available: the sea surface temperature (SST) observation product from CMEMS [4]–[6], the EN4 quality controlled ocean data product [7], and the SST measurement provided by the AutoNaut wave-propelled ASV [14]. Finally, the assimilation setup of the various computational domains, with or without data assimilation, is shown. In Section III the

results from the various simulations are presented, while a discussion of the data and of the consequences of assimilating data at different scales is provided in Section IV. Section V lists some concluding remarks from this work.

II. Theory and Method

A. SINMOD ocean model

The model in this study is a physical-chemical-biological model of the Arctic Ocean and the Norwegian Sea based on the primitive Navier-Stokes equations. The model uses z-coordinates vertically, and a regular horizontal grid based on the polar stereographic projection [15], [16]. It operates in a nested set-up such that the model domains covering the largest scales run with lower spatial and temporal resolution and provides boundary conditions for model domains covering smaller scales with higher spatial and temporal resolution. The model domains of interest in this study are shown in Figures 1a to 1c which are the GIN domain, Nor4km domain and Mids domain with $20\,\mathrm{km}$, $4\,\mathrm{km}$, $800\,\mathrm{m}$ horizontal resolution respectively. The red rectangle in Fig. 1a and 1b indicates the boundaries for Nor4km and Mids respectively.

Each instance of the model is parallelized using OpenMP [17], and a Message Passing Interface (MPI) [18] is used to facilitate communication between each ensemble member and the process which executes the EnKF calculations. This allows the EnKF and each ensemble member to run on different computing nodes with relatively low overhead.

For this specific study, the biology calculations are not activated.

B. Ensemble Kalman Filter formulation

In this study, the Ensemble Kalman Filter (EnKF) is used to assimilate temperature measurements into the SINMOD ocean model described in Section II-A. An ensemble of model simulations is run in parallel to represent the error statistics of the ocean state and, in particular, the temperature state. The method is proven to handle strong nonlinear dynamics and large state vectors which is the case for ocean models such as SINMOD [2], [19].

1) Equations: The implementation of the EnKF follows the formulation presented in [2], [20]–[22] with a minor difference in the representation of the measurement error introduced by [23].

The forecast ensemble consisting of N ensemble members, each containing n state variables, can be written as the $n \times N$ matrix

$$X^f = [x_1, \ldots, x_N], \tag{1}$$

where the column vectors are the state vectors x_i of dimension n. The ensemble mean is defined to be

$$E(X^f) = \frac{1}{N}\sum_{k=1}^{N} x_k = \frac{1}{N}(Xe_{N\times 1})e_{1\times N}, \tag{2}$$

with dimension $n \times n$. The vector e, whose entries are all ones, is introduced as it is closer in terms of notation to how the averaging is implemented numerically. The ensemble spread is defined as

$$\theta = X^f - E(X^f) = X - \frac{1}{N}(Xe_{N\times 1})e_{1\times N} \tag{3}$$

giving the ensemble covariance matrix

$$\Psi = \frac{\theta\theta^T}{N-1}. \tag{4}$$

Measurements are organized in a vector, d, with a size equal to the number of measurements, $m \times 1$. The corresponding measurement error covariance matrix, R, is of size $m \times m$. Independent random perturbations are drawn from a normal distribution with expectation 0 and variance corresponding to the values in R. These perturbations are added to N instances of the measurement vector d to form a measurement matrix D with size $m \times N$ such that

$$D = [d_1, \ldots, d_N], \quad d_j = d + v_j, \quad v_j \sim N(0, R) \tag{5}$$

where v_j are the independent random perturbations with standard deviation corresponding to the uncertainty of each measurement. The analysis step is then given by

$$X^a = X + \Psi H^T(H\Psi H^T + R)^{-1}(D - HX^f), \tag{6}$$

where H is the observation matrix with rows indicating the linear combination of state variables corresponding to each measurement.

To avoid computing Φ directly which is impractical for state vectors with millions of variables, equation (6) can be rewritten to

$$X^a = X^f + \frac{1}{N-1}\theta(H\theta)^T P^{-1}(D - HX^f), \tag{7}$$

where

$$P = \frac{1}{N-1}H\theta(H\theta)^T + R, \tag{8}$$

$$H\theta = HX^f - \frac{1}{N}((HX^f)e_{N\times 1})e_{1\times N} \tag{9}$$

2) Sequential setup: Depending on the number of available measurements in each assimilation cycle, and on how the memory is distributed to each rank on the computing nodes, the observation matrix H of size $m \times n$ may become too large to be stored in memory. For certain simulations in this study, H can be as large as $72\,\mathrm{GB}$. To handle these large matrices the available measurements are split into P batches and the assimilation cycle is run sequentially P times in a loop. Table I lists the different simulations and whether the assimilation cycle is implemented sequentially or not.

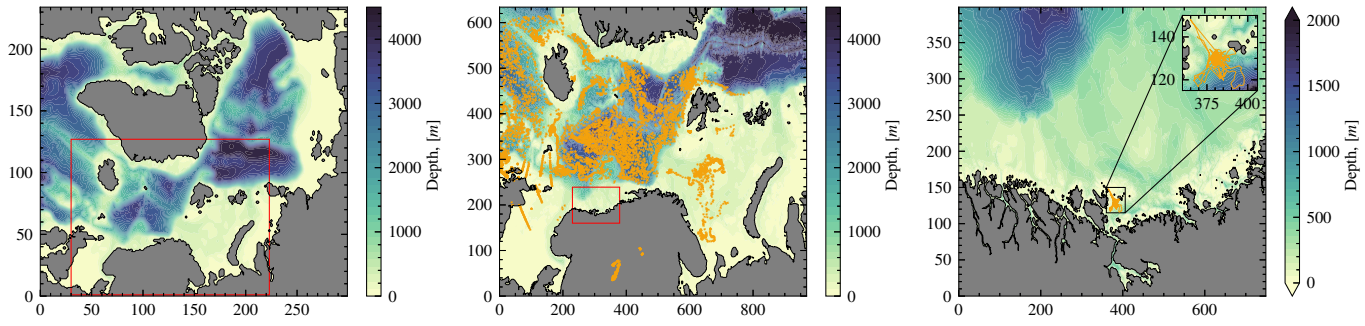


Fig. 1. Computational model domain with depth fields for GIN with a horizontal resolution of 20 km, shown in Fig. 1a, Nor4km with a horizontal resolution of 4 km, shown in Fig. 1b and Mids with a horizontal resolution of 800 m, shown in Fig. 1c. The red rectangular box in the first two domains indicates the sub-domain boundaries. The orange points in 1b indicate the location of the measurements in the EN4 data set for the time period January, 2021 to May, 2022. The orange line in Fig. 1c indicates the path covered by the AutoNaut in April, 2022.

3) Localization: The ensemble undersampling is due to finite ensemble size and causes spurious long-range cross-correlations between state variables. This in turn leads to a nonphysical and erratic analysis update of the ensemble members [19], [24], [25]. The long-range correlations can be dampened by vertical and horizontal covariance localization which is implemented as a Schur product of a quasi-Guassian, isotropic, distance-dependent localization function (the Gaspari-Cohn function [26]), and the covariance terms in the analysis step, as shown in Equations (10) and (11):

$$\rho_h \circ \rho_v \circ \theta(H\theta)^T, \tag{10}$$

$$\rho_h \circ \rho_v \circ H\theta(H\theta)^T, \tag{11}$$

where the Gaspari-Cohn matrix ρ is

$$\rho_{h,v}(d,c) = \begin{cases} -\frac{1}{4}(|d|/c)^5 + \frac{1}{2}(|d|/c)^4 + \frac{5}{8}(|d|/c)^3 \\ \qquad -\frac{5}{3}(|d|/c)^2 + 1, & 0 \le |d| \le c, \\ \frac{1}{12}(|d|/c)^5 - \frac{1}{2}(|d|/c)^4 + \frac{5}{8}(|d|/c)^3 \\ +\frac{5}{3}(|d|/c)^2 - 5(|d|/c) + 4 - \frac{2}{3}c/|d|, & c \le |d| \le 2c, \\ 0, & 2c \le |d| \end{cases} \tag{12}$$

where d is the spatial distance between two state variables, and $c = \sqrt{10/3} * L$, where L is the localisation radius. This smoothly dampens any covariance structures outside of a preset localization distance which depends on the horizontal and vertical resolution in the model of interest.

4) System noise: In this work, the system noise is modeled by adding spatially coherent perturbations to the atmospheric wind velocity field, [27], in the hourly ERA5 data set from the European Centre for Medium-Range Weather Forecasts (ECMWF). This data set covers the Earth on a 30 km grid and includes 137 vertical layers from the surface and up to 80 km [28]. The wind velocity field is interpolated to fit the horizontal resolution of the different model domains in the ocean model.

The perturbations are generated every time a new wind field is indexed by the ocean model to be used as atmospheric forcing and are computed by taking the difference of the wind velocity field at two randomly selected dates within a time span of one year and scaling the difference with a temporal autocorrelation factor, γ. The perturbation is aggregated over time but dampened with a factor μ. The zonal, τ_x, and meridional, τ_y, wind field components applied to ensemble member i at time t_k are then given by

$$\tau^i_{x,y}(t_k) = \tau^i_{x,y}(t_k) + \mu\delta\tau^i_{x,y}(t_k) \tag{13a}$$

$$\delta\tau^i_{x,y}(t_k) = f\delta\tau^i_{x,y}(t_{k-1}) + \gamma\{\tau_{x,y}[t^i_{k1}] - \tau_{x,y}[t^i_{k2}]\} \tag{13b}$$

where the perturbation amplitude is set to $\mu = 0.2$ and the temporal autocorrelation is set to $\gamma = \sqrt{1-(f^2)}$, where $f = 0.98$. t^i_{k1} and t^i_{k2} are two randomly selected dates.

5) Simplifications: The EnKF offers no guarantees of the maximum or minimum values of *a posteriori* estimates, even for directly observed states. To ensure that the analysis update is within physically accepted limits, any updates to the temperature are limited to the interval $[-1.8°C, 35°C]$, where the lowest value represents the freezing temperature of saline water.

C. Measurement data

In this work, three sources of observations are used and assimilated at different spatial and temporal scales.

1) CMEMS SST product: The observation data set covering the largest spatial scales is a sea surface temperature observation product from the Copernicus Marine Environmental Monitoring Service (CMEMS) [4]–[6]. The data set consists of daily global ocean maps of the foundation SST using both in-situ and satellite data from both infrared and microwave radiometers, provided by the Group for High-Resolution Sea

Surface Temperature (GHRSST) project. The operational setup is run by the UK's Met Office. Its horizontal grid resolution is $0.05^\circ \times 0.05^\circ$ which transforms to a $5.53\,\mathrm{km}$ horizontal resolution along the latitudinal axis, and a longitudinal horizontal resolution dependent on the latitude. The horizontal resolution at the southernmost point in Norway corresponds to $1.21\,\mathrm{km}$, whereas the northernmost point in Norway corresponds to $1.81\,\mathrm{km}$.

For this study, only every third day of observations is assimilated into the largest model domain with spatial resolution $20\,\mathrm{km}$ in the time interval January 2020 to May 2022. The uncertainty of the SST observation product was set to $\Delta T = 0.02^\circ C$.

2) EN4 quality controlled ocean data: The observation data set EN4 is assimilated into the intermediate scale model with a resolution of $4\,\mathrm{km}$. EN4 consists of data from a number of ocean profiling instruments that provide temperature and salinity measurements [7]. Its primary source is the World Ocean Database 2009 (WOD09) National Oceanic and Atmospheric Administration (NOAA) and the National Centers for Environmental Information (NCEI). Additionally, a subsection of the data set provided by the Arctic Synoptic Basin Wide Oceanography (ASBO) project is included to improve arctic coverage. Moreover, the Global Temperature and Salinity Profile Program (GTSPP) and Argo data from the Argo global data assembly centers (GDACs) are included. Together, these sources are processed and quality controlled to form the EN4 quality-controlled ocean data which are distributed on a monthly basis about two months after they are collected.

The depths at which the profiles provide measurements vary with location and time, and do not exactly match the fixed depth layers of the ocean model. To account for this, the vertical profile measurements are interpolated to the ocean model depth layers using a cubic spline interpolation. Furthermore, the deepest measurement of every profile is removed to increase the interpolation accuracy. For this study, every available profile within the Nor4km model domain, indicated by orange dots in Fig. 1b, has been assimilated at simulation time in the time interval July 2021 to May 2022, which amounted to roughly one profile observation per hour. The measurement error was set to $\Delta T = 0.002^\circ$ according to [29].

3) The AutoNaut: a wave-propelled USV: The AutoNaut[1] is a commercially available surface vehicle produced and supplied with a proprietary control and communication system [30]. The data used in this work were collected by a 5 meter long version of the AutoNaut in which payload control, navigation and communication systems are developed by NTNU with a publicly available hardware and software architecture[2] based on the DUNE unified navigation environment [31], as described in [32].

The AutoNaut USV carries an innovative propulsion system that relies on sea surface waves to produce forward thrust, making it suitable for sustained operations at sea without human assistance. The ground speed of the vehicle is mainly determined by waves and normally reaches 0.5 to 3 knots, although drifting forces generated by winds and sea currents may under certain circumstances degrade speed and impact navigation performance [33]. The USV's heading is governed by an electric stern rudder powered by an onboard battery bank. A set of three solar panels supplying up to 300 W allows continuous energy harvesting and recharging of the battery bank, extending vehicle autonomy to several weeks. Unlike common robotic marine platforms, the AutoNaut is therefore less constrained by the energy limitations that normally affect propulsion and payload usage, ensuring long-duration autonomous missions.

The uncertainties of the AutoNaut temperature measurements were set to $\Delta T = 0.005^\circ C$.

D. Assimilation tree

A series of simulations are run in a tree branching setup as shown in Fig. 2, where model domains with coarse spatial resolution generate boundary conditions for model domains with finer spatial resolution. The propagation of boundaries is indicated by arrows. Gray blocks indicate free-running models with no assimilated measurements, whereas green blocks indicate models with assimilation of measurements. An overview of the model domains, the horizontal resolution, the corresponding assimilated measurement, and the simulation time period is presented in Table I. In order to effectively describe the individual simulations the following naming convention is used. Consider the horizontal axis in Fig. 2, showing the same models with different boundary conditions and available measurements. Two levels are available for GIN, three levels for Nor4km, and four levels for Mids, where the last levels for all models have data assimilation enabled. The level associated with each simulation will be referred to as L1 to L4 and added to the name of the model, e.g., the Nor4km simulation without data assimilation of the EN4 data set, that uses boundary conditions from the GIN simulation with data assimilation of the SST observation product from CMEMS, will be named Nor4km-L2.

III. Results

A. Large scales

Two simulations are run for the GIN model covering the largest spatial scales of $20\,\mathrm{km}$ horizontal resolution. One in a free-run setup with no measurements, and the other with assimilation of the SST observation product from CMEMS (see Section II-C1). A series of plots is presented in Fig. 3, each showing the difference in SST between the free-run simulation (GIN-L1) and the data assimilated simulation (GIN-L2), at different time steps. The time steps are February 26, 2021, June 25, 2021 and March 29, 2022 in Fig. 3a, 3b and 3c respectively. The red color in the figures indicates a warmer SST in the free-run simulation. The warmer bias in the free-run simulation can be seen for the winter and spring time steps in the North Sea, Norwegian Sea and the Barents Sea.

[1] http://www.autonautusv.com

[2] http://autonaut.itk.ntnu.no/doku.php

TABLE I
SIMULATIONS WITH THE CORRESPONDING ASSIMILATION SETUP.

Model	Horizontal resolution	Measurements	Time period	Batching	Vertical localisation	Horizontal localisation
GIN-L1	20 km	-	January 2020 - May 2022	-	-	-
GIN-L2	20 km	CMEMS SST product	January 2020 - May 2022	Yes	100 m	80 km
Nor4km-L1	4 km	-	July 2021 - May 2022	-	-	-
Nor4km-L2	4 km	-	July 2021 - May 2022	-	-	-
Nor4km-L3	4 km	EN4	July 2021 - May 2022	No	50 m	8 km
Mids-L1	800 m	-	January 2022 - May 2022	-	-	-
Mids-L2	800 m	-	January 2022 - May 2022	-	-	-
Mids-L3	800 m	-	January 2022 - May 2022	-	-	-
Mids-L4	800 m	Autonaut	March 2022 - May 2022	No	20 m	1 km

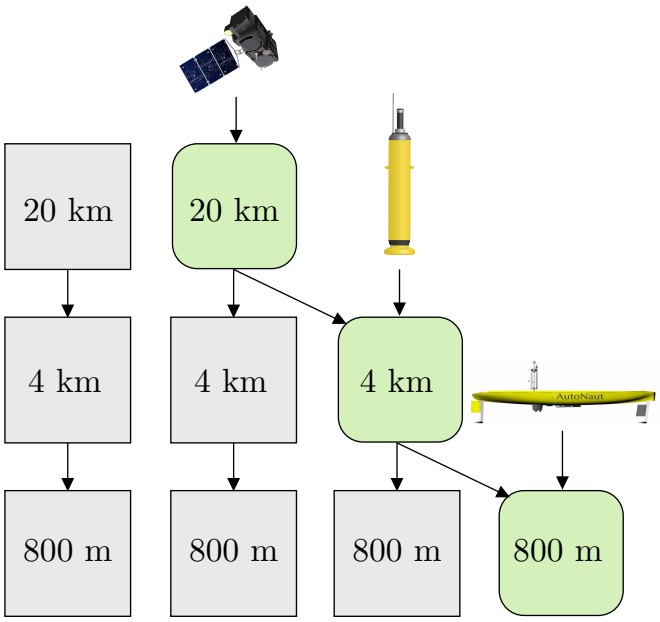


Fig. 2. Branching assimilation setup for the three model domains, indicated with their respective horizontal resolution, with three different sources of measurements: the SST observation product represented by a satellite, the EN4 data set represented by an ARGO floater, and lastly the AutoNaut. Gray boxes indicate simulations without data assimilation. Green boxes indicate simulations with data assimilation. The black arrows indicate the direction of propagation of boundary or measurement information.

Similarly, the Icelandic Sea and the Baffin Bay in Greenland are generally colder in the free-run simulation for the same time steps.

B. Intermediate scales

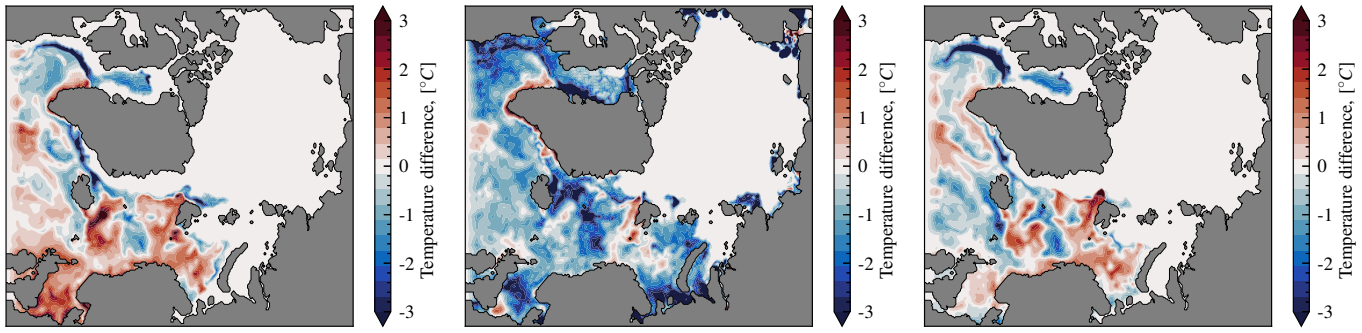
Three simulations are run for the Nor4km model covering the intermediate spatial scales of 4 km horizontal resolution: Nor4km-L1, Nor4km-L2 and Nor4km-L3, where the last simulation includes data assimilation of the EN4 data set presented in Section II-C2. Fig. 4 shows the difference in SST between the three simulation setups at the time April 12, 2022. The SST difference between Nor4km-L1 and Nor4km-L2 is presented in Fig. 4a. A clear temperature bias is seen in shallow waters, e.g., in the North Sea and between Scotland and Iceland, indicating that information from the SST observation product assimilated in GIN-L2 propagates to Nor4km-L2 via the boundary conditions. However, for parts of the Norwegian and Greenland Sea with greater depths (2000 m and below), the temperature bias is not present.

The SST difference between Nor4km-L1 and Nor4km-L3 is presented in Fig. 4b. Here, the same warm temperature bias is seen for shallower parts of the ocean. However, the deeper parts are now also altered compared to the free-run model, but with a varying warm/cold bias. This is seen more clearly in Fig. 4c showing the difference between Nor4km-L2 and Nor4km-L3. Corrections to the Nor4km-L3 simulation are taking place in the deeper regions due to the EN4 profile measurements. The corrections are insignificant outside of this deep-water area.

C. Small scales

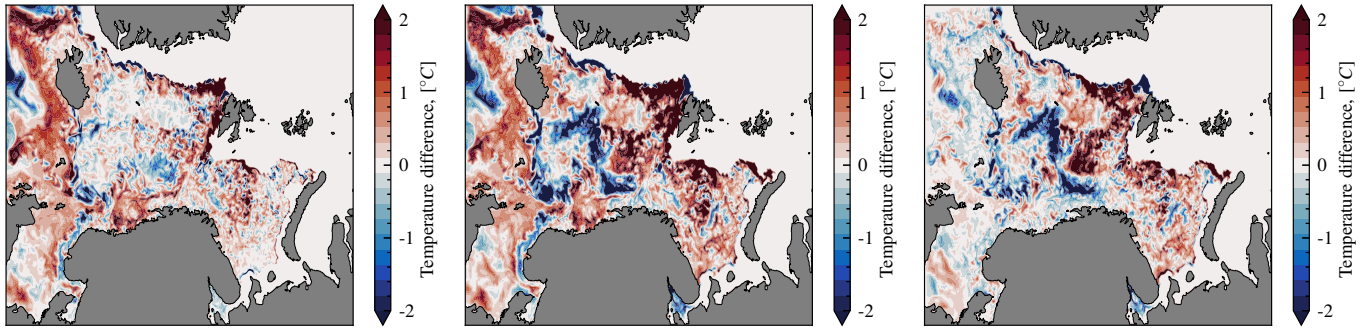
For the domain covering the smallest spatial scales of 800 m horizontal resolution, four simulations are run: Mids-L1, Mids-L2, Mids-L3 and Mids-L4, where Mids-L4 assimilates measurements from the AutoNaut (see Section II-C3). A subdomain of Mids including Frohavet is investigated. The SST difference between Mids-L1 and Mids-L2, Mids-L1 and L3 and Mids-L1 and Mids-L4 are represented in Figures 5a to 5c. Moreover, the position of the assimilated AutoNaut measurements in the Mids-L4 simulation for a one-hour time interval before and after the analysis step is indicated by the orange line in Fig. 5c. All three comparison plots show the warm bias from Mids-L1. Additionally, a significant local correction is seen in the areas surrounding the AutoNaut measurements.

Fig. 6a and 6b present time series of the surface temperature for all Mids simulations at two locations. While the former is far from the shore with deeper waters, the latter is in the AutoNaut's operational area in shallower waters. Both time series are for the simulation period January, 2022 to May, 2022. The time series for both the deep and shallow points show a similar temperature trend where the Mids-L1 and Mids-L2 simulations are overlapping in value. Not until March does Mids-L1 deviate from the other three simulations, and at that point, Mids-L1 is seen to hold higher surface temperature values primarily for the deep point far away from the coast for the remainder of the simulation time, and for the time period March, 2022 to April, 2022 for the shallow



(a) February 26, 2021 (b) June 25, 2021 (c) March 29, 2022

Fig. 3. Comparison plots at three different time steps showing the difference between the free-run GIN-L1 simulation and the GIN-L2 simulation with assimilation of the SST observation product. The time steps are indicated in the sub-captions of each figure.



(a) (b) (c)

Fig. 4. Three comparison plots of the Nor4km simulations at the time step April 12, 2022. Fig. 4a shows the difference between Nor4km-L1 and Nor4km-L2. Fig.4b shows the difference between Nor4km-L1 and Nor4km-L3, and lastly, Fig. 4c shows the difference between Nor4km-L2 and Nor4km-L3.

point. For the simulations Mids-L2, Mids-L3 and Mids-L4, the values follow the same trend, all deviating from the free-run simulation.

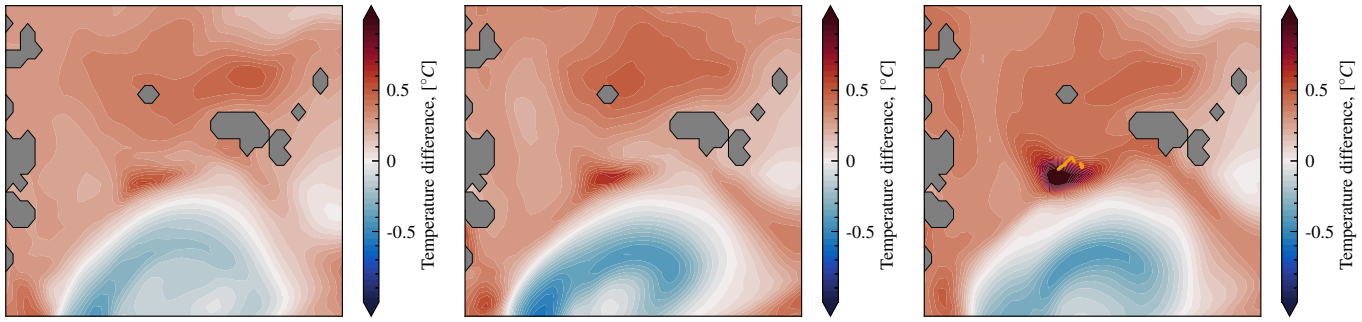
Fig. 6c and 6d show a time series of both SST values and the root mean square error (RMSE) of the simulated SST and that measured by the AutoNaut. The spatial location for each time point is varying according to the position of the AutoNaut such that the values are taken from the model grid cells closest to the ASV's location at the time. In Fig. 6c, the orange line indicates the vehicle's measurements. The small bias in the SST which was seen for the shallow point in Fig. 6b for the time period March, 2022 to early April, 2022, is also seen in Fig. 6c. However, the Mids-L4 simulation which assimilates the AutoNaut data is seen to follow the local observations more closely. Eventually, all four simulations follow the temperature trend measured by the ASV. The order of magnitude of the deviations from simulated SST and measured SST seen in Fig. 6d show the same result, that the Mids-L4 simulation is closest in value to the local measurements. Note that there is a small period for which the vehicle did not collect temperature measurements (April 7, 2022 to April 11, 2022) and in this time period the RMSE for Mids-L4 increased.

IV. Discussion

A. Technical aspects of the EnKF implementation

The EnKF framework utilized in this study is the first version of this scale to be implemented for the SINMOD ocean model and some technical considerations should be discussed to facilitate improvement for future versions.

The amount of computational resources required to run the EnKF framework varies drastically from case to case, depending on how many measurements are to be assimilated for each time step. The Nor4km-L3 and Mids-L4 which assimilates a number of measurements on the order of $\mathcal{O}(1) - \mathcal{O}(100)$, the assimilation step takes only a couple of seconds. However, for the GIN-L2 simulation, which assimilates the SST observation product for every wet surface cell which accounts to approximately 28000 measurements assimilated into 687000 state variables, the assimilation step takes close to 15 minutes. This is not much if the framework is to be utilized in an operational setup where new large-scale satellite measurements only become available every couple of hours and the simulation is only meant to forecast a couple of days in advance. However, for longer simulations spanning years, a subset of the measurements would be enough to correct the model by allowing further correction of the other state vari-



(a) (b) (c)

Fig. 5. Three comparison plots for a subdomain of the Mids simulations at the time step April 6, 2022. Fig. 5a shows the difference between Mids-L1 and Mids-L2. Fig. 5b shows the difference between Mids-L1 and Mids-L3, and Fig. 5c shows the difference between Mids-L1 and Mids-L4. The orange line in Fig. 5c indicates the position of the AutoNaut in a one-hour time interval before and after the time of the displayed temperature field.

ables through the covariance structure and localization length. On the other hand, GIN requires relatively few CPU cycles to simulate and its main computational cost is associated with the assimilation step. For Nor4km, the model domain is smaller in spatial scale, but the amount of state variables is roughly 13 times more than in the GIN model, meaning its main computational cost is associated with running the model itself and not the assimilation step. To summarize the computational footprint of this study, a total of 90'000 CPU hours were used to develop and run the 9 simulations outlined in Fig. 2.

Due to a limited computational budget, the number of ensemble members varied from 23 to 29 for each of the simulations GIN-L2, Nor4km-L3 and Mids-L4. As explained in Section II-B3, the small ensemble size may cause spurious cross-correlations across large distances, and this was observed especially in regions close to the ice edge, where the edge moved in and out of neighboring wet cells, causing a large variance in the temperature. Localization prevented the update of state variables due to long-range correlations. However, the choice of horizontal localization radius should, for the Mids-L4 simulation, be chosen larger than the one presented in Table I. This would likely cause a more distinct and lasting update for a larger region in the local area, and not only in the very close proximity of the AutoNaut. Although, increasing the radius too much would reintroduce the spurious cross-correlations. In that case, the simulations would benefit more from increasing the number of operational assets producing observations and increasing the ensemble size, as the localization radius should then be chosen not only based on the covariance structure but also on the available assets in the area. The same reasoning applies to Nor4km-L3. However, that simulation and assimilated measurement are designed to cover a larger spatial scale than the Mids-L4 simulation, and the localization radius should be increased in any case.

On the topic of system noise, as explained in Section II-B4, only the perturbations on the atmospheric forcing were applied to model the total system noise. For state variables near the top surface layers, this worked fairly well. However, for deeper layers, the ensemble spread was limited. After some initial spin-up, the deeper state variables still experienced a covariance with a magnitude surpassing the uncertainty in the available measurements, allowing for model corrections on the deeper layers as well. An alternative would be to add perturbations to other elements of the model, for instance, parameters related to vertical and horizontal mixing or direct perturbation of the state variables [27].

B. Systematic bias

There are clear deviations from the free-run simulations GIN-L1, Nor4km and Mids-L1, which are used as bench marks for most of the comparison results in Section III, and in the corresponding simulations which directly assimilates measurements or indirectly through boundary conditions. Even though the deviations in Nor4km and Mids are seen to be warmer in the free-run models, it is not necessarily true for all time steps through a full year simulation. The GIN model is analyzed for three different time steps and compared to GIN-L2, some areas, specifically the Baffin bay and regions around Greenland are seen to be colder at all times. However, the majority of the computational domain is warmer during winter times and colder in the summer. This indicates that there is no constant systematic bias which can easily be corrected for throughout the model.

Another remark is that the measurements assimilated into GIN-L1 is the foundation SST, i.e., the surface temperature free of diurnal temperature variability. It is defined to be the temperature at the first time of the day when the heat gain from the solar radiation absorption exceeds the heat loss at the sea surface, meaning that the measurement will give a cold temperature relative to other times of the day. The time of assimilation is set to midnight which doesn't necessarily exactly fit the time of the measurement, and may cause a model correction which is generally colder than it should be.

As presented in Section III-B, the deviations from free-run simulation and assimilated simulation for the Nor4km model seemingly varies with depth, which is an effect not seen in the GIN-L2 simulation. For the GIN-L2 simulation, the deep-water areas receive updates directly from the measurements,

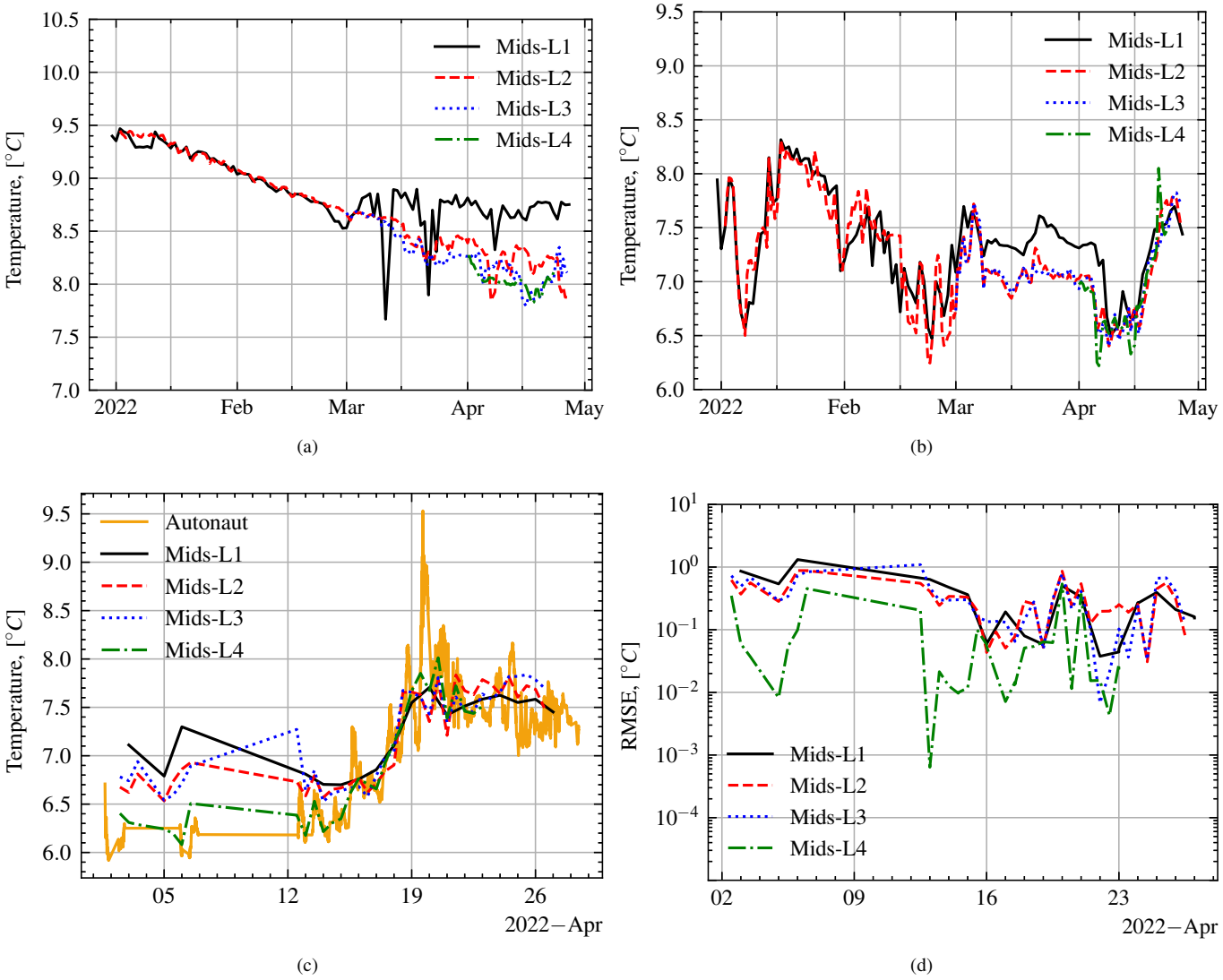


Fig. 6. Four time series of the temperature field in the Mids simulations at the varying location. Fig. 6a shows the time series of a point in relatively deep waters in the Mids domain, i.e., the point $(200, 350)$ in the coordinate system seen in Fig. 1c. Fig. 6b shows the time series of the same simulation, but for a point closer to the operational area of the AutoNaut, i.e., $(380, 138)$. Fig. 6c shows a time series where the temperature is evaluated in the cell closest to the position of the ASV in each time step. Fig. 6d shows the time series of the RMSE of the simulated temperature at the ASV's location and the AutoNaut measurements.

both in the surface layer, and in the lower layers within the vertical localisation radius. For the Nor4km-L2, the same area only receives this cooling effect through the boundary conditions, and an increased vertical mixing may neglect the effect altogether making the warm bias less persistent for Nor4km-L2 in the deeper areas than in the shallower areas compared to the GIN-L2 simulation. Once the EN4 data set is assimilated, with a high measurement density in the deeper areas, the same area which didn't experience corrections in Nor4km-L2, gets corrections in Nor4km-L3. These are seen to be of both cold and warm character in Fig. 4c and don't show a clear trend other than the corrections being larger for deeper waters than for shallow water. Note that the effect from GIN on the boundaries of Nor4km is mainly felt in the area drawn by the line from the North sea going westwards passing through the Celtic sea and Icelandic sea to Greenland. The rest of the boundaries are mainly covered by land or ice.

The warm bias close to shore during winter or spring is notable in all three model domains. Interestingly, neither the large scale corrections in GIN-L2 or the intermediate scale corrections in Nor4km-L3 are capable of providing an accurate temperature field in very local regions in the Mids model. The deviation from the free-run model seen in Mids-L2 and Mids-L3 indicates that the assimilation of the SST observation product in the coarse large scale model has an impact on even the smaller scales, and the assimilation of the EN4 data in the intermediate scales are close to negligible. Once the local asset data from the AutoNaut is included in Mids-L4,

a large correction is seen on all time steps it is operating, surpassing the correction from the largest scales, indicating that local assets in high resolved domains are necessary to build a complete picture of the ocean state when investigating local phenomena.

C. Effect of assimilating at different scales

As indicated in the previous sections, the choice of when and where to assimilate available measurements in model domains of different horizontal resolution, will affect the simulations to varying degree. The largest impact on all non-free-run simulations in this study is seen from the large scale covering SST observation product. Its effect is best seen in figure 6a where all simulations subject to the assimilated SST observation product is seen to deviate from the free-run model, while the simulations covering other measurements are all seen to be following each other, suggesting that the coarser data set contribute on all scales to make a good foundation for assimilation at smaller scales.

The effect of the intermediate scale measurements, the EN4 data set, is insignificant on a large scale compared to assimilation of the SST product. However, as presented in Section III-B it contributes with large corrections for the deep water areas, as the data set is of three-dimensional character and provides measurements not only on the surface layer, but also in the deeper layers. Nonetheless, the effects of these corrections are close to negligible on the boundary of the Mids domain. For studies of areas closer to the location of the observing instruments in the EN4 data set, the corrections would be more relevant. Note that the location of the majority of the measurements in the EN4 data set is far away from boundary for the Mids domain. If the measurements were closer they might have influenced the simulations in the smaller domain to a larger extent.

Assimilating the ASV measurements in the Mids domain would not likely yield a different result if the EN4 data set was assimilated or not in the Nor4km domain. However, removing the SST observation product from the GIN model would definitely make a difference on the Mids simulations. Recall that red colors in the comparison plots are to be interpreted as the free-run simulation being warmer than simulations with assimilation. As seen in Section III-C, the cooling effect caused by the ASV measurements for this specific time step and area would not likely be as lasting if not the boundaries to the model domain was not already corrected for the systematic bias from the coarser domains. The simulated temperature in the cells surrounding the ASV measurement would likely be warmer, and the cooling effect would rapidly be mixed and diminished into the neighbouring warm cells.

D. Recommendations

Based on the findings in this study, some general points can be made on the design of observational systems with multiscale heterogeneous measurements together with an ocean model corrected with an EnKF framework.

A specific location of interest as well as the spatial resolution should be chosen before the design of the simulations with assimilation take place. Basing the recommendations on observational systems with models and measurements similar to those used in this study, an area of interest studied by coarse models such as the GIN model, subject to either warm or cold temperature bias would benefit from large scale corrections from low resolved global covering data sets such as the SST observation product from CMEMS. The otherwise simplified GIN simulation will be corrected to provide a better foundation for further analysis in the coarse computational domain. Other measurements on this scale are not tested in this study, but are not expected to have a significant impact on the coarse simulations as both the spatial and temporal resolution is of such greater extent than simulations like, e.g., the Nor4km model.

For studies in areas of intermediate length scales such as the Nor4km simulation, the initial assimilation of the global covering data set in the boundary domain, GIN, should be present in any case. Moreover, if the studies are conducted on shallow water areas in this domain, the results of this paper don't suggest any benefit from assimilating new measurements on this scale. However, if the studies are conducted on phenomena on greater depths, assimilation of vertical profile measurements such as the EN4 data sets will impact the result greatly. A decision on whether or not to assimilate global covering data sets such as satellite measurements into this intermediate length scale domain will depend on the computational efficiency of running the model with or without assimilation. For the case of a reanalysis study, adding more measurements to the intermediate scales is likely to increase accuracy of the simulations. However, for an operational set-up the relative cost vs accuracy gain can be too high, and assimilating large scale measurements in the coarser model domains may be sufficient.

Furthermore, if the studies of interest concern very local areas such as the subdomain of Mids presented in fig 5, assimilation of measurements from local operative assets will likely only be useful if these assets operate in close proximity to the point of interest and the regions surrounding it, or if the simulation time is long enough for the effect to propagate throughout the computational domain. If the assets are operating far away from the area of interest, they are not likely to have a significant effect on the simulation in this area.

V. Conclusion

In this work the Ensemble Kalman Filter was implemented for the SINMOD ocean model and used to assimilate data from three different sources into model domains with varying spatial resolution. The results of this study suggests that ocean models as part of observational systems for the ocean should assimilate observations of global coverage into the coarsest model domains whenever available to reduce propagation of systematic bias in nested sub domains. Assimilation of measurements at greater depths or of local extent can be omitted unless the measurements cover regions in close proximity to

the investigated phenomena of interest.
An increased ensemble size as well as a more optimally chosen localization radius would enhance the presented results. However, the discussion on the effect of assimilating heterogeneous measurements at different spatial scales would likely be unchanged.
Investigation of how variables not included in the state space, e.g. biological variables, are influenced by assimilation of measurements at varying spatial scales will be left for future work.

ACKNOWLEDGEMENT

The simulations were performed on resources provided by Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway
The work was funded internally by the Department of Engineering Cybernetics, NTNU.

REFERENCES

[1] P. F. Lermusiaux, C.-S. Chiu, G. G. Gawarkiewicz, P. Abbot, A. R. Robinson, R. N. Miller, P. J. Haley, W. G. Leslie, S. J. Majumdar, A. Pang, and F. Lekien, "Quantifying uncertainties in ocean predictions," Report, 2006.
[2] G. Evensen, "The ensemble kalman filter: theoretical formulation and practical implementation," *Ocean Dynamics*, vol. 53, no. 4, pp. 343–367, 2003.
[3] D. Murphy and C. Janzen, *Advances in In-Situ Ocean Measurements*. Cham: Springer International Publishing, 2018, pp. 141–162. [Online]. Available: https://doi.org/10.1007/978-3-319-66493-4_8
[4] S. Good, E. Fiedler, C. Mao, M. J. Martin, A. Maycock, R. Reid, J. Roberts-Jones, T. Searle, J. Waters, J. While, and M. Worsfold, "The current configuration of the ostia system for operational production of foundation sea surface temperature and ice concentration analyses," *Remote Sensing*, vol. 12, no. 4, 2020.
[5] C. J. Donlon, M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, "The operational sea surface temperature and sea ice analysis (ostia) system," *Remote Sensing of Environment*, vol. 116, pp. 140–158, 2012.
[6] J. D. Stark, C. J. Donlon, M. J. Martin, and M. E. McCulloch, "Ostia : An operational, high resolution, real time, global sea surface temperature analysis system," in *OCEANS 2007 - Europe*, 2007, Conference Proceedings, pp. 1–4.
[7] S. A. Good, M. J. Martin, and N. A. Rayner, "En4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates," *Journal of Geophysical Research: Oceans*, vol. 118, no. 12, pp. 6704–6716, 2013.
[8] A. Dallolio, G. Quintana-Diaz, E. Honoré-Livermore, J. L. Garrett, R. Birkeland, and T. A. Johansen, "A satellite-usv system for persistent observation of mesoscale oceanographic phenomena," *Remote Sensing*, vol. 13, no. 16, 2021.
[9] P. McGillivary, K. Rajan, J. B. de Sousa, and F. Leroy, "Integrating autonomous underwater vessels, surface vessels and aircraft as persistent surveillance components of ocean observing studies," in *2012 IEEE/OES Autonomous Underwater Vehicles (AUV)*, 2012, Conference Proceedings, pp. 1–5.
[10] G. E. Berget, J. Eidsvik, M. O. Alver, F. Py, E. I. Grøtli, and T. A. Johansen, *Adaptive Underwater Robotic Sampling of Dispersal Dynamics in the Coastal Ocean*, ser. Springer Proceedings in Advanced Robotics. Springer, 2022, vol. 20, book section Chapter 37, pp. 610–625.
[11] T. O. Fossum, J. Eidsvik, I. Ellingsen, M. O. Alver, G. M. Fragoso, G. Johnsen, R. Mendes, M. Ludvigsen, and K. Rajan, "Information-driven robotic sampling in the coastal ocean," *Journal of Field Robotics*, vol. 35, no. 7, pp. 1101–1121, 2018.
[12] T. O. Fossum, "Adaptive sampling for marine robotics," Thesis, Norwegian University of Science and Technology, 2019. [Online]. Available: http://hdl.handle.net/11250/2608426
[13] F. A. Michelsen and M. Alver, "Optimal in situ measurement networks for ocean data," in *OCEANS 2017 - Aberdeen*, 2017, Conference Proceedings, pp. 1–9.
[14] AutoNautLtd., "Autonaut - wave propelled unmanned surface vehicle," http://www.autonautusv.com/.
[15] D. Slagstad and T. A. McClimans, "Modeling the ecosystem dynamics of the barents sea including the marginal ice zone: I. physical and chemical oceanography," *Journal of Marine Systems*, vol. 58, no. 1-2, pp. 1–18, 2005.
[16] P. Wassmann, D. Slagstad, C. W. Riser, and M. Reigstad, "Modelling the ecosystem dynamics of the barents sea including the marginal ice zone ii. carbon flux and interannual variability," *Journal of Marine Systems*, vol. 59, no. 1-2, pp. 1–24, 2006.
[17] R. Chandra, L. Dagum, D. Kohr, R. Menon, D. Maydan, and J. McDonald, *Parallel programming in OpenMP*. Morgan kaufmann, 2001.
[18] L. Clarke, I. Glendinning, and R. Hempel, "The mpi message passing interface standard," *Int. J. Supercomput. Appl.*, vol. 8, 05 1996.
[19] P. Sakov and P. R. Oke, "A deterministic formulation of the ensemble kalman filter: an alternative to ensemble square root filters," *Tellus A: Dynamic Meteorology and Oceanography*, vol. 60, no. 2, pp. 361–371, 2016.
[20] P. L. Houtekamer and H. L. Mitchell, "Data assimilation using an ensemble kalman filter technique," *Monthly Weather Review*, vol. 126, no. 3, pp. 796–811, 1998.
[21] ——, "A sequential ensemble kalman filter for atmospheric data assimilation," *Monthly Weather Review*, vol. 129, no. 1, pp. 123–137, 2001.
[22] P. L. Houtekamer and F. Zhang, "Review of the ensemble kalman filter for atmospheric data assimilation," *Monthly Weather Review*, vol. 144, no. 12, pp. 4489–4532, 2016.
[23] J. Mandel, "Efficient implementation of the ensemble kalman filter," Report, 2006.
[24] P. Sakov and L. Bertino, "Relation between two common localisation methods for the enkf," *Computational Geosciences*, vol. 15, no. 2, pp. 225–237, 2010.
[25] L. Nerger, P. Kirchgessner, and A. Bunse-Gerstner, "On the choice of an optimal localization radius in ensemble kalman filter methods," *Monthly Weather Review*, vol. 142, no. 6, pp. 2165–2175, 2014.
[26] G. Gaspari and S. E. Cohn, "Construction of correlation functions in two and three dimensions," *Quarterly Journal of the Royal Meteorological Society*, vol. 125, no. 554, pp. 723–757, 1999.
[27] C. L. Keppenne, M. M. Rienecker, J. P. Jacob, and R. Kovach, "Error covariance modeling in the gmao ocean ensemble kalman filter," *Monthly Weather Review*, vol. 136, no. 8, pp. 2964–2982, 2008.
[28] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J. Thépaut, "The era5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, 2020.
[29] R. Cowley, R. E. Killick, T. Boyer, V. Gouretski, F. Reseghetti, S. Kizu, M. D. Palmer, L. Cheng, A. Storto, M. Le Menn, S. Simoncelli, A. M. Macdonald, and C. M. Domingues, "International quality-controlled ocean database (iquod) v0.1: The temperature uncertainty specification," *Frontiers in Marine Science*, vol. 8, 2021.
[30] P. Johnston and M. Poole, "Marine surveillance capabilities of the autonaut wave-propelled unmanned surface vessel (USV)," in *OCEANS 2017 - Aberdeen*, June 2017, pp. 1–46.
[31] J. Pinto, P. S. Dias, R. Martins, J. Fortuna, E. R. B. Marques, and J. B. de Sousa, "The lsts toolchain for networked vehicle systems," *2013 MTS/IEEE OCEANS - Bergen*, p. 1–9, jun 2013. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6608148
[32] A. Dallolio, B. Agdal, A. Zolich, J. A. Alfredsen, and T. A. Johansen, "Long-endurance green energy autonomous surface vehicle control architecture," *Oceans 2019 Mts/Ieee Seattle*, 2019.
[33] A. Dallolio, H. Øveraas, J. A. Alfredsen, T. I. Fossen, and T. A. Johansen, "Design and Validation of a Course Control System for a Wave-Propelled Unmanned Surface Vehicle," *Field Robotics*, vol. Volume 2, pp. 748–773, 2022.