AGU ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems®

RESEARCH ARTICLE
10.1029/2022MS003218

Special Section:
Data assimilation for Earth system models

Bridging Gaps in the Climate Observation Network: A Physics-Based Nonlinear Dynamical Interpolation of Lagrangian Ice Floe Measurements via Data-Driven Stochastic Models

Jeffrey Covington[1], Nan Chen[1], and Monica M. Wilhelmus[2]

[1]Department of Mathematics, University of Wisconsin–Madison, Madison, WI, USA, [2]Center for Fluid Mechanics, School of Engineering, Brown University, Providence, RI, USA

Key Points:

- Developed systematic data-driven reduced-order stochastic model to significantly lower the computational cost for dynamical interpolation
- Estimated key physical parameters such as the floe thickness and recovered the unobserved ocean flow field with uncertainty quantification
- Recovered important non-Gaussian distributions of the curvature and angular displacement of the missing floes in the Beaufort Sea

Correspondence to:
N. Chen,
chennan@math.wisc.edu

Citation:
Covington, J., Chen, N., & Wilhelmus, M. M. (2022). Bridging gaps in the climate observation network: A physics-based nonlinear dynamical interpolation of Lagrangian ice floe measurements via data-driven stochastic models. *Journal of Advances in Modeling Earth Systems*, *14*, e2022MS003218. https://doi.org/10.1029/2022MS003218

Received 26 MAY 2022
Accepted 31 AUG 2022

Abstract Modeling and understanding sea ice dynamics in marginal ice zones rely on measurements of sea ice. Lagrangian observations of ice floes provide insight into the dynamics of sea ice, the ocean, and the atmosphere. However, optical satellite images are susceptible to atmospheric noise, leading to gaps in the retrieved time series of floe positions. This paper presents an efficient and statistically accurate nonlinear dynamical interpolation framework for recovering missing floe observations. It exploits a balanced physics-based and data-driven construction to address the challenges posed by the high-dimensional and nonlinear nature of the coupled atmosphere-ice-ocean system, where effective reduced-order stochastic models, nonlinear data assimilation, and simultaneous parameter estimation are systematically integrated. The new method succeeds in recovering the locations, curvatures, angular displacements, and the associated strong non-Gaussian distributions of the missing floes in the Beaufort Sea. It also accurately estimates floe thickness and recovers the unobserved underlying ocean field with an appropriate uncertainty quantification, advancing our understanding of Arctic climate.

Plain Language Summary Tracking individual ice floes is a unique measurement of areas of the Arctic where the ice cover interacts with the open ocean. Unfortunately, optical satellite images of these areas are frequently obscured by clouds, leading to missing observations of the ice floes. Traditional methods of filling in these gaps in the data set have issues. Linear interpolation, which averages between available observations to fill in missing ones, fails to recover the curvature of the floes. Dynamical interpolation methods, which take into account the physical properties of the ice floes, are very computationally expensive. This paper presents a nonlinear dynamical interpolation framework for recovering missing floe observations, which is both computationally efficient and statistically accurate. The framework incorporates a model of the atmosphere, ocean, and sea ice and systematically develops data-driven reduced-order stochastic models, which significantly accelerate the dynamical interpolation while retaining accuracy. In addition, the framework estimates key physical parameters, such the floe thickness. This new method succeeds in recovering the locations, curvatures, angular displacements, and strong non-Gaussian statistics of the missing floes in a data set of ice floes in the Beaufort Sea. These results can provide complete data sets that advance our understanding of Arctic climate.

1. Introduction

Sea ice plays a key role in the Arctic climate system (Bhatt et al., 2014; Leppäranta, 2011; Maslowski et al., 2012; Thomas, 2017; Weeks & Ackley, 1986; Weiss, 2013). It modulates important momentum, heat, and material transfer processes between the ocean and the atmosphere (Squire, 2020; Thomson et al., 2018; Timmermans et al., 2018). Given the sensitivity of the sea ice cover to global warming trends, the observation and modeling of sea ice are critical for understanding global climate, including monitoring the drastic changes in the Arctic and assessing possible future climate scenarios.

Earth system models typically characterize sea ice as a continuum with viscous-plastic rheology primarily through ice concentration, volume, and thickness (Hibler, 1979; Hunke & Dukowicz, 1997; Tremblay & Mysak, 1997). While this traditional modeling approach yields realistic results at the basin scale, at scales of $\mathcal{O}(10)$ km and smaller, sea ice exhibits brittle behavior with the motion of individual fragments deviating from a continuum description. In this case, the discrete element method (DEM) (Cundall, 1988; Cundall & Strack, 1979; Hart

et al., 1988), which characterizes the trajectories of individual ice floes, as opposed to clusters of ice, becomes the natural choice to describe sea ice dynamics. Compared to continuum models, the modeling of individual floes provides a richer representation of sea ice dynamics through local interactions with the oceanic and atmospheric components (Lindsay & Stern, 2004; Tuhkuri & Polojärvi, 2018). In addition, since the DEM models are developed under Lagrangian coordinates, there is no need for an advective transport scheme to move floes between grid cells as in continuum models, significantly reducing computational costs. The DEM models can also change the spatial resolution as the geophysical situation requires, allowing greater flexibility in the study of sea ice.

The unique advantages and wide applications of models based on the DEM highlight the need for observational sea ice data within the Lagrangian framework. Observed floe trajectories facilitate the development and calibration of DEM models and provide insight into the evolution of sea ice properties. However, despite the increase in satellite missions and the improved techniques in acquiring remote sensing observations, most existing observational products are based on Eulerian descriptions of the sea ice drift field. Exceptions include the Arctic Ocean Sea Ice Drift Reprocessed (Ardhuin & Prévost, 2020) and the Making Earth System data records for Use in Research Environments (MEaSUREs) programs (Kwok, 1999). Yet, these measurements cannot adequately resolve sea ice motion at small scales due to the spatial resolution (31.25 km for Arctic Ocean Sea Ice Drift Reprocessed product) or the sampling frequency (3-day interval for MEaSUREs) of the data. On the other hand, in situ field measurements using buoys on ice floe surfaces have provided invaluable information. However, trajectories are often sparse (Brunette et al., 2022; Gabrielski et al., 2015; Hutchings et al., 2012; Itkin et al., 2017; Lei et al., 2020), with coastal regions undersampled relative to the central Arctic.

Recently, a new Lagrangian floe tracking algorithm, called the Ice Floe Tracker (Lopez-Acosta et al., 2019; Wolfe et al., 1998), was developed and applied to optical satellite images. It creates Lagrangian sea ice measurements in the low-sampled regions between the ice pack and the open ocean, commonly known as the marginal ice zones (MIZ). This data set was the first of its kind in that it provides not only the Lagrangian trajectories but also the floe sizes and geometries together with the angular displacements of floes in the MIZ extending throughout the 21st century. Given the demonstrated link between floe rotation rates and the characteristics of the underlying small-scale ocean eddies in the Beaufort Gyre MIZ, these sea ice floe observations have proved to be essential for recovering the state of the underlying turbulent ocean field, providing a unique insight into the multi-scale nature of the ocean (Manucharyan et al., 2022).

Atmospheric noise is visible in optical images and leads to many 1- or 2-day gaps in the retrieved trajectories within the Lagrangian Ice Floe Tracker data set. See Figure 1 for an example. A commonly used approach to filling these gaps is to interpolate between the available observations through linear interpolation (Manucharyan et al., 2022). These trajectories (and their curvature) are used to characterize ocean circulation/eddy behavior. However, such an approach ignores the MIZ dynamics, leading to trajectories lacking many physical properties. Linear interpolation also fails to retrieve the curvature of the trajectories, which is essential for characterizing the turbulent ocean flow field within the meso/submeso-scale range in polar regions. Alternatively, physics-based dynamical interpolation incorporates both the available partial observations and knowledge of the coupled atmosphere-ocean-floe dynamics. While the resulting interpolated trajectories are expected to reflect reality better, traditional dynamical interpolation approaches are often extremely slow and computationally expensive due to the high dimensionality and nonlinearity of the underlying system (Branicki et al., 2018; Grooms & Majda, 2013).

This paper presents an efficient and statistically accurate nonlinear dynamical interpolation framework for recovering the missing floe observations in Lagrangian trajectories. It exploits a balanced physics-based and data-driven construction to address the challenges posed by the high-dimensional and nonlinear nature of the coupled system. This new method involves a sequential prediction-correction procedure. The error from predicting the missing values in the coupled atmosphere-ocean-floe system is mitigated by incorporating the available observations of floe positions and orientations via Bayesian inference. One crucial feature of the presented framework is that it exploits a data-driven reduced-order stochastic modeling strategy to advance the statistical forecast of the atmosphere and ocean fields, which are the underlying driving forces of the sea ice motion, but are not considered by the direct curve-fitting algorithms. Particularly, these simple stochastic models describe the time evolution of the leading spectral modes of the atmosphere and ocean fields, where effective stochastic forcing is adopted to characterize the fluctuations at the unresolved scales. Therefore, the resulting surrogate models significantly reduce the computational cost at the forecast step, which is the most time-consuming part in

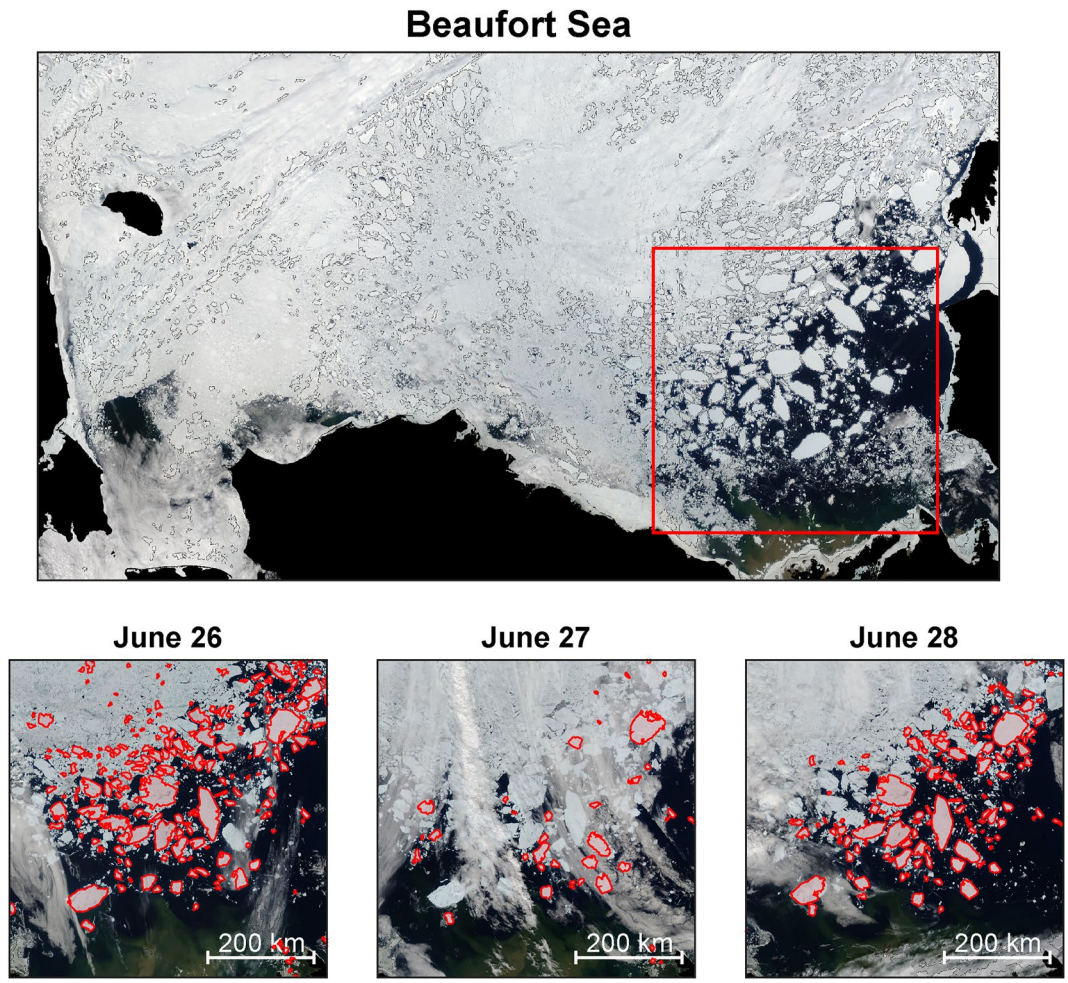

Figure 1. Sea ice floes in the Beaufort Sea marginal ice zones (MIZ). Top panel: Representative Moderate Resolution Imaging Spectroradiometer True Color image (downloaded from the NASA Worldview application) displayed in a WGS 84/ NSIDC Sea Ice Polar Stereographic North 70″ N projection. For only this figure, the image is oriented 90° from standard Polar Stereographic coordinates so that the top of the image is roughly north. The red box outlines the region of interest. Bottom panels: The observed MIZ of the Beaufort Sea (the box area of the top panel) is shown on three consecutive dates (26–28 June 2008). Identified floes are contoured with red. On 27 June 2008, the atmospheric noise acts to blur ice floe contours impeding the effective identification of most of the floes.

traditional dynamical interpolation approaches. It is worth highlighting that closed analytic formulae are available for expressing the statistics associated with these simple stochastic models, facilitating the systematic and efficient data-driven model calibration. The calibrated models succeed in accurately predicting the atmosphere and ocean states and the associated uncertainty. The latter is crucial in reaching the least biased state estimate using nonlinear dynamical interpolation, especially in the presence of strong turbulence, which is again completely missed by deterministic curve fitting methods. In addition, the framework allows for the simultaneous estimation of several critical physical parameters that cannot be directly inferred from satellite images but are essential for dynamical interpolation, such as the thickness of each floe, using only relatively short floe trajectories.

The rest of the paper is organized as follows: It starts with the development of the physics-based data-driven dynamical interpolation framework. Then the new method is tested on both a synthetic data experiment and the real data set of floe trajectories in the Beaufort Sea MIZ. The focus here is on the noninteracting floes, but the framework can be easily extended to the interacting ones. The study also includes analysis of the resulting interpolated Lagrangian ice floe trajectories and angular displacements as well as the recovery of several key physical properties of the floes and their associated statistics. The recovered floe trajectory utilizing the traditional linear interpolation approach will serve as a benchmark solution.

2. The Reduced-Order Modeling and Nonlinear Dynamical Interpolation Framework

This section presents an overview of the new modeling and nonlinear dynamical interpolation framework consisting of four key steps. The technical details of the model, the data, and the methods are included in the Appendices A–C.

This framework works with a coupled atmosphere-ice-ocean system. While this system can take the form of a coupled dynamical model, the framework also allows the atmospheric and/or oceanic components to be given as numerical data. Here, a DEM model is used to characterize the sea ice dynamics, where the individual floe shapes and sizes are drawn from a library of floe observations in the Beaufort Gyre MIZ (Lopez-Acosta & Wilhelmus, 2021). Given that the observations contain only nearly noninteracting and shape-preserving floes, the ice floe motion can be assumed to be mainly driven by oceanic and atmospheric forcing, which are calculated from surface integrals over floe shapes. The ocean component is given by a two-layer quasi-geostrophic (QG) model that generates eddies from baroclinic instabilities, the so-called Phillips model (Vallis, 2017), in which the cumulative impact of many passing floes on the turbulent eddy field is represented via a quadratic surface drag. The model has been systematically calibrated to capture the key features of the interaction between the Arctic Ocean and the ice floes. The optimization criterion used here is to match the simulated and observed scale-dependency of the floe rotation variance, where the bulk vertical shear of the background horizontal velocity, the deformation radius, and the effective ratio between the top and bottom layer depths are the tuning parameters. See Manucharyan et al. (2022) for the detailed calibration strategy. The spatial resolution is 128×128 gridpoints. Because of the high latitude, the deformation radius is small—on the same order as the model resolution. Despite potential unresolved effects, their impact is mitigated by the relatively large floe areas and statistical averaging of the dynamical interpolation. The atmospheric component is taken from a reanalysis product (ERA5) (Copernicus Climate Change Service (C35), 2017; Olauson, 2018), which provides Eulerian wind velocity fields over the observational period. As the wind field exhibits larger-scale features, a coarser spatial resolution of 11×11 gridpoints is used. Note that the focus here is in the MIZ of the Beaufort Sea (see Figure 1). Hence, a double-periodic boundary condition is adopted for simplicity. The potential model error and bias introduced from the various approximations can be mitigated at the statistical forecast stage using the stochastic corrections and Bayesian inference in the framework introduced in the rest of this section. The domain size, as shown in Figure 1, is roughly 600×600 km. Figure 2 includes a schematic illustration of the main steps of the framework.

Step 1. Development of low-cost data-driven reduced-order stochastic models

The ensemble forecast adopts a probabilistic characterization of the model state and is thus a natural way to predict complex turbulent systems (Leutbecher & Palmer, 2008; Palmer, 2019; Toth & Kalnay, 1997). However, the high dimensionality and nonlinearity of the coupled atmosphere-ice-ocean system make a single realization of the model forecast very computationally expensive, let alone the forecast of the entire ensemble. Therefore, the first step in this framework is to develop data-driven reduced-order models with the aim to significantly lower the computational cost of the forecast step.

The top section of Figure 2 outlines the development of such reduced-order models for the turbulent ocean field. Given a long simulation generated from the original two-layer QG ocean model, the spectral decomposition of the velocity field is used. Most of the energetic modes are concentrated within a circular area centered at $\boldsymbol{k} = (0, 0)$ with a relatively small radius in spectral space $|\boldsymbol{k}| \leq K$, where $\boldsymbol{k} = (k_1, k_2)$ is the spectral index. The reduced-order model is set up to only describe the temporal evolution of the dynamics of this small set of spectral modes. It is expected to retain most of the key features of the original ocean field but significantly lower the computational cost. Yet, given the nonlinearity of the original ocean model, the governing equation of each spectral mode is fully coupled with all other modes, including those omitted in the reduced-order model. To effectively characterize the temporal evolution of each spectral mode in the reduced-order model, a linear stochastic model is developed as a surrogate (Gardiner, 2009),

$$\frac{\mathrm{d}u}{\mathrm{d}t} = (-a + i\omega)u + f + \sigma\dot{W}, \tag{1}$$

where u is a complex variable for a single spectral mode, a and ω are the damping and oscillation frequencies, respectively, f is the forcing of the system, $\dot{W}$ is a complex-valued white noise, and σ is the amplitude of the noise. In Equation 1, damping and stochastic noise are adopted to parameterize the contribution of the extremely

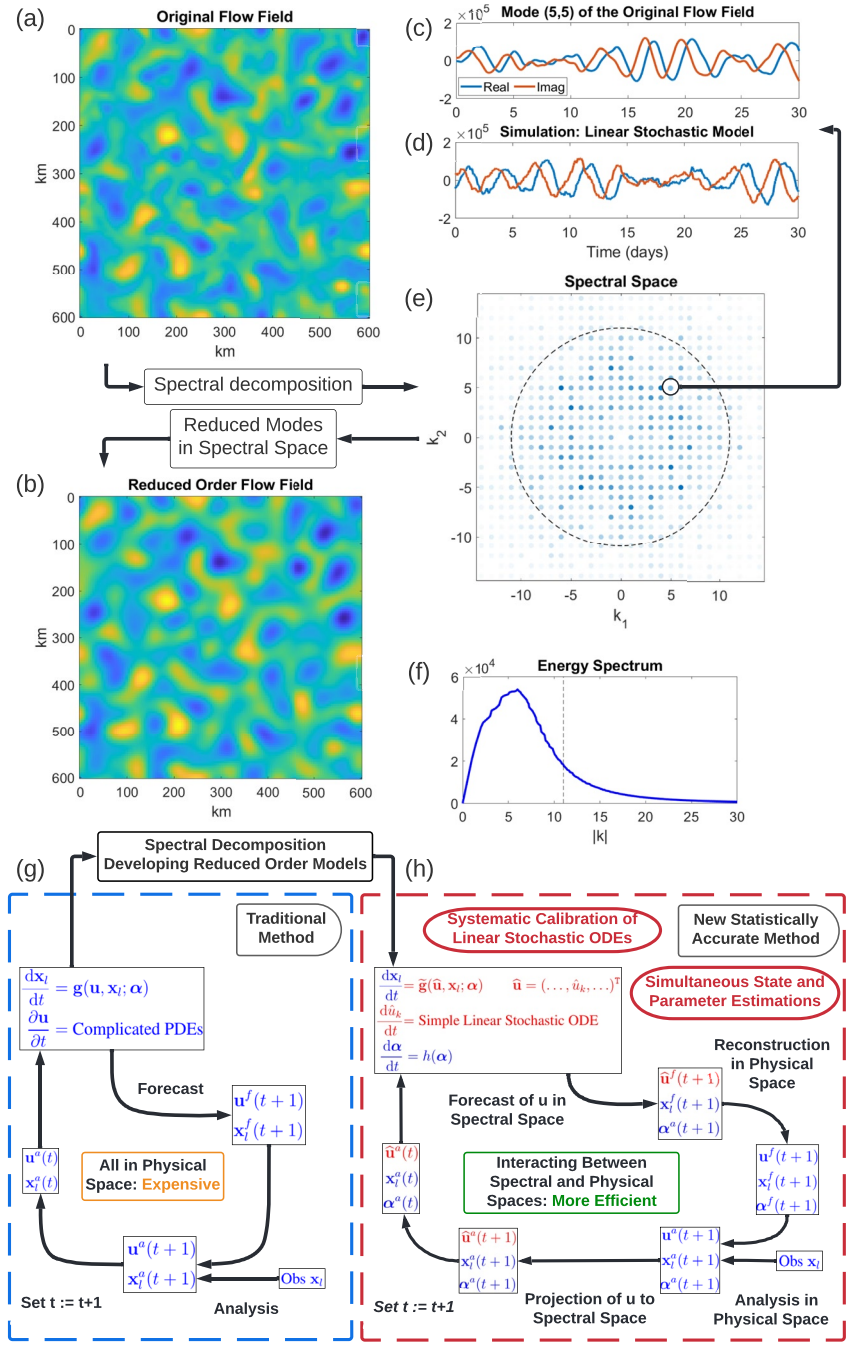

Figure 2. Schematic diagram of the new method. Panels (a)–(f): A spectral decomposition is applied to the output of a complicated ocean model. Only a small set of the most energetic spectral modes are retained. The governing equations of these energetic modes are modeled by the low-cost linear stochastic models, thereby significantly reducing the computational cost. The illustration also compares the true ocean flow field and its reconstructed state. The original field is generated from the two-layer quasi-geostrophic (QG) model, while the reconstructed one only uses modes for which $|\mathbf{k}| \leq 11$. The top right corner compares the time series of the mode $\mathbf{k} = (5, 5)$ associated with the QG model and a random realization from the calibrated linear stochastic model. Panels (g)–(h): The dynamical interpolation is performed via nonlinear data assimilation. The traditional method requires running the original system in the physical space and is extremely expensive. Here, $\boldsymbol{x}$, $\boldsymbol{u}$ and $\boldsymbol{\alpha}$ denote the ice floes, the ocean and the atmospheric state variables, and the model parameters, respectively. In contrast, the new and efficient method for dynamical interpolation alternates between physical and spectral spaces using the reduced-order stochastic models. It has the main benefit of allowing for the simultaneous estimation of state variables and key physical parameters.

complicated, nonlinear, and deterministic part of the original governing equation, which leads to a cheaper and more effective way to reproduce the statistical forecast results (Berner et al., 2017; Branicki et al., 2018; Farrell & Ioannou, 1993; Majda, 2016; Majda & Chen, 2018; Qi & Majda, 2018). Independent linear stochastic models are used to characterize the temporal evolution of each mode in spectral space, which nevertheless allows a fully correlated spatial pattern in physical space. Since the QG model generates an incompressible flow field, the spectral representation of the ocean is based on the stream function. Similarly, a pair of linear stochastic models is utilized to approximate the two-dimensional velocity components associated with each spectral mode of the atmospheric wind field.

Note that despite the independence between the linear stochastic models for different Fourier modes, a strong correlation still exists between the state variables in physical space after the spatial reconstruction via the inverse Fourier transform. The simple structure of the linear stochastic model allows for systematic model calibration and large computational savings, making an ensemble forecast feasible. It is worthwhile to highlight that the most important factor impacting the performance of the dynamical interpolation is the leading-order statistics of the ensemble forecast rather than the dynamics of individual model trajectories. In this way, because the independent stochastic models are accurate with respect to the statistics of the QG model, the forecast produces good results.

Step 2. Systematic model calibration

The linear stochastic model in Equation 1 can be calibrated systematically by taking advantage of the analytic formulae for its four fundamental statistics: the mean, the variance, and the real and imaginary parts of the decorrelation time. The values of these four statistics have a unique one-to-one correspondence with the four parameters: a, ω, f, and σ. Therefore, once the values of these statistics are computed numerically from the time series of a single spectral mode in the original QG ocean model, these values are plugged into the closed analytic formulae, determining the four parameters in the linear stochastic model associated with that spectral mode.

Rather than simulating the 30,000 modes of the two-layer QG model, setting K to be 11 in step 1 of the framework results in a reduced-order model containing only about 400 modes. This simulation still resembles the full QG system while being much more computationally inexpensive, and roughly 30 times faster, see panels (a) and (b) of Figure 2. In addition, after applying the proposed calibration procedure, a random realization of the time series from the linear stochastic model is also statistically similar to the truth by capturing the mean, variance, and decorrelation time. See panels (c) and (d) of Figure 2 for a comparison of mode (5, 5). This similarity is essential for an accurate ensemble forecast using the linear stochastic reduced-order models. Finally, the same linear stochastic models are adopted as surrogate models to describe the atmospheric wind field based on the ERA5 reanalysis data.

Step 3. Physics-based dynamical interpolation via nonlinear data assimilation

Dynamical interpolation exploits the optimal combination of the ensemble forecast from the model and the information from the partial observations via nonlinear data assimilation. The incorporation of the underlying dynamics sets dynamical interpolation apart from pure curve fitting methods. The basic dynamical interpolation scheme used here is the ensemble Kalman smoother (EnKS), in which an ensemble of model trajectories represents the estimate of the system state. Each ensemble member contains trajectories for all state variables, including the sea ice, the ocean, and the atmosphere. The EnKS provides point estimates through the ensemble mean and quantifies the uncertainty through the ensemble spread. The resulting distribution is called the posterior distribution, which contrasts with the prior distribution solely obtained from the forecast step of the model.

The traditional EnKS contains a straightforward prediction-correction loop in physical space that requires repeatedly integrating the expensive original dynamical model. In contrast, the new method here uses the linear stochastic models to approximate the ocean and atmospheric flow fields, and the prediction-correction procedure alternates between the physical and the spectral spaces. Specifically, the prediction of the ocean and atmospheric flow fields, which involves running the linear stochastic models forward, is implemented in the spectral space. On the other hand, the correction of all the state variables, which applies the Bayesian formula that optimally combines the model and observational information, is carried out in the physical space. Spectral decomposition and flow field reconstruction are adopted after each correction and prediction step, respectively, see panels (g) and (h) of Figure 2. Since only a few spectral modes are involved in the sequential prediction-correction procedure, the computational efficiency is preserved. Note that the DEM sea ice model remains highly nonlinear, which makes the entire dynamical interpolation nonlinear. To further improve the numerical stability and mitigate erro-

neous spurious long-term correlations, localization and fixed lag strategies are incorporated into the basic version of the EnKS (Anderson, 2012; Evensen, 2009).

Each set of floe observations is processed sequentially in time. The algorithm represents the model state with an ensemble of model trajectories. During the prediction correction loop, the ensemble is forecast forward in time up to the next available observation. The sea ice variables are forecast in physical space according to the sea ice model. The ocean and atmosphere variables are forecast in spectral space according to the statistically accurate stochastic forecast models. After the forecast of the ensemble, all variables are transformed back to physical space where they are compared to the observations in the analysis step. The ensemble is updated, in a Bayesian sense, according to the new observations. Finally, the ocean and atmosphere variables are transformed back to spectral space for the next iteration of the loop.

Step 4. Efficient parameter estimation of important physical quantities

The main practical challenge of using general dynamical interpolation methods to analyze the sea ice cover is the lack of access to the entire parameter space from a single remote sensing instrument. For example, the thickness of the floes determines the inertia of floe motion and is crucial to the coupled system. To overcome this challenge, an efficient parameter estimation algorithm is embedded into the dynamical interpolation framework. Here, the unobserved physical quantities are treated as the augmented state variables, which are simultaneously estimated with the actual variables of the model state. The uncertainty in the estimated parameters, due to the relatively short Lagrangian trajectories, is also quantified in the algorithm. See Appendix E.

3. Results of Interpolating the Floe Trajectories and Angular Displacements

3.1. Setups of the Two Experiments

The new dynamical interpolation framework is first applied to a synthetic data experiment and then to the real observation scenario.

The synthetic data experiment uses the two-layer QG ocean model and the reanalysis data for the atmospheric winds to force the ice floes governed by the DEM model. The ice floe shapes, sizes, positions, and orientations are initialized from a library of floes in the Beaufort Gyre MIZ (Lopez-Acosta & Wilhelmus, 2021), which is generated from optical remote sensing imagery using the Ice Floe Tracker algorithm (Lopez-Acosta et al., 2019). The thickness of each floe is randomly drawn from a background distribution (Kwok, 2018) and is assumed to be constant during the entire observational period, see panel (c) of Figure 4. Note that the stochastic approximate models are not utilized to generate the synthetic data; rather, they are only used to dynamically interpolate the missing floe observations. For the real data experiment, Lagrangian sea ice floe trajectories are obtained using the Ice Floe Tracker algorithm within the study area delineated in Figure 1 during the spring-to-summer transition of 2008. The same linear stochastic models that are calibrated for the synthetic data experiment are adopted to carry out the dynamical interpolation. See Table 1 for the summary of the models used to generate the true signal and those adopted to implement the dynamical interpolation in the two experiments.

The floe locations and angular displacements are the only observational information in the dynamical interpolation for the coupled atmosphere-ice-ocean system. These two quantities are obtained from the satellite images at a frequency of roughly every 24 hr. The observational uncertainty, which is used in the dynamical interpolation algorithm, is set to be 0.25 km and 5°, respectively.

Both experiments contain 38 floe trajectories of various lengths in time, see the bottom panel of Figure 3. Excluding the first and the last point in each floe trajectory, in total, there are 164 remaining candidate observational points for the 38 trajectories. These 164 candidates are randomly divided into four sets, where each set contains 41 data points. Then four independent dynamical interpolation simulations are carried out. In each simulation, the 41 candidate observations in the corresponding set are artificially removed as the missing observations. Note that the missing observations referenced in the real data experiment are not the actual missing ones in the satellite images obscured by clouds, but are rather the artificially removed ones. Such a setup guarantees that the true values of these missing floes are known, and therefore it allows the qualitative study of the accuracy of the dynamical interpolation. Nevertheless, this 3:1 ratio between the number of observed and missing floe observations mimics the real-world situation in the MIZ during the boreal summer. The number of ensemble members used here is 600.

Table 1
Summary of the Models Used for Both the Synthetic and the Real Data Experiments

	Atmosphere	Ocean	Sea ice
(a) Synthetic data experiment			
Truth	ERA5 reanalysis	Two-layer QG	The known DEM model
Interpolation	Calibrated LSM	Calibrated LSM	The known DEM model
(b) Real data experiment			
Truth	Not needed	Not needed	Satellite observations
Interpolation	Calibrated LSM	Calibrated LSM	The known DEM model

Note. In each experiment, the first row "truth" stands for the underlying systems that generate the true signal, while the second row "interpolation" indicates the model used for dynamical interpolation. The same calibrated linear stochastic model (LSM) is utilized for the real data as for the synthetic data experiments. In the real data case, the true signals of the atmosphere and ocean components are not needed. In the synthetic data case, the true atmosphere and ocean models are used to drive the DEM model to generate the observed floe trajectories and angular displacements.

Figure 3 displays the 38 sea ice floe trajectories in the real data experiment, which are retrieved from satellite remote sensing imagery using the Ice Floe Tracker algorithm during the spring-to-summer transition of 2008. Each floe trajectory is represented by the transition from fully transparent to opaque, with each floe assigned a specific color. The index in these floes corresponds to those in Figures 5 and 7.

3.2. Results of the Synthetic Data Experiment

Figure 4 illustrates the parameter estimation of the floe thicknesses from the dynamical interpolation. To quantify the uncertainty in the estimated thickness of each floe, the posterior distribution characterized by the ensemble members is also included via a violin plot. The results shown here are calculated using all the ensemble members from the four simulations, but different simulations lead to similar distributions for all the floes. The estimations are overall reasonably accurate, especially given such a small number of observations within a large domain. Particularly, the truth of all the 38 floes is consistently covered by the posterior distribution. In addition, for two-thirds of the floes, the true thickness value lies in a high likelihood region of the distribution within one standard deviation from the mean. Note that the error in the thickness estimation can offset the error from recovering the atmosphere and ocean flow fields, and therefore the overall interpolation results remain accurate, as will be seen below. It is worthwhile to highlight that both the background thickness distribution, from which the true thickness values are drawn, and the estimated posterior distributions, exhibit strongly fat-tailed non-Gaussian behavior, as is clear in the violin plot. These non-Gaussian features are the unique outcome of the highly nonlinear dynamics of the sea ice floes. In addition, because the ensemble is transformed from the prior distribution to the posterior during each update of the nonlinear EnKS, rather than resampled, these important non-Gaussian features can be preserved through the ensemble update (see the Appendix D.) By contrast, simply considering the ensemble mean and standard deviation would underestimate the likelihood of large floe thicknesses while simultaneously overestimating the likelihood of small thicknesses. Notice that such a non-Gaussian feature is found in all model variables but is especially illustrated by thickness estimation. These findings imply the necessity of incorporating both the nonlinear sea ice dynamics and the nonlinear data assimilation scheme into the dynamical interpolation framework.

Panel (a) in Figure 5 compares linear and dynamical interpolation for recovering the floe location and angular displacement. The ensemble mean estimate using the dynamical interpolation almost always outperforms linear interpolation in recovering the floe locations. Specifically, the absolute error using the linear interpolation is nearly three times as large as that using the dynamical interpolation. The linear interpolation also, by design, completely fails to recover the curvature and the nonlinear evolution of the floe trajectories. In contrast, the dynamical interpolation accurately captures these important physical features. In addition to the point estimate using the ensemble average, the ensemble provides the quantification of the estimated uncertainty. The

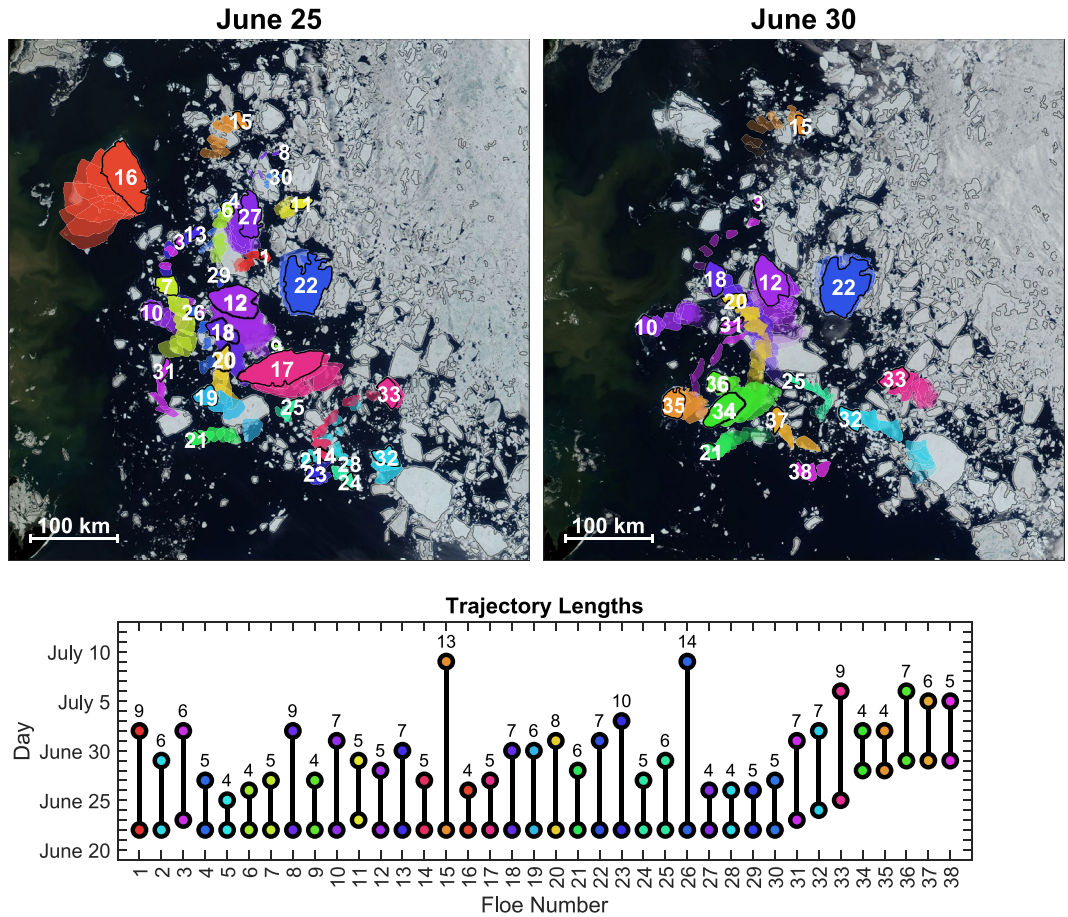

Figure 3. Sea ice floe trajectories retrieved from optical satellite remote sensing imagery. Moderate Resolution Imaging Spectroradiometer True Color images (downloaded from the NASA Worldview application) acquired on the 25th and 30th of June, 2008 are displayed in a WGS 84/NSDIC Sea Ice Polar Stereographic North 70° N Projection, on top of which retrieved ice floe trajectories are displayed in color. In both images, the evolution of floe positions is represented as a shift in opacity from transparent to opaque objects. Final floe positions are marked using black contour lines. Note that the recovered floe trajectories have different lengths and periods. Information regarding the acquisition period of each floe trajectory is shown in the bar plot underneath. Only a subset of the 38 noninteracting floes used in this study are shown for clarity.

uncertainty overall remains at a relatively low level, indicating the confidence of the posterior mean estimate. Among all the 164 recovered missing observations, roughly 80% of the true observations fall within two standard deviations around the ensemble mean estimate. This implies the accuracy and robustness of the dynamical interpolation. In addition, the recovery of the angular displacement using the dynamical interpolation is quite accurate.

Panel (a) of Figure 6 illustrates the recovered ocean field represented by the stream functions utilizing the dynamical interpolation. The result shown here is on a specific day in the middle of the entire time period. The accuracy in recovering the ocean field remains in a similar level on other days. The overall pattern correlation between the truth and the recovered ocean field is around 0.3. Nevertheless, given the fact that there are only 17 floes inside this large domain on this day, the skill of recovering the ocean field is already significant. In particular, the ocean eddies are recovered quite reasonably in the areas where the observed floes are concentrated. The pattern correlation is above 0.3 and the amplitude of the recovered eddies is similar to the truth. This is partially due to the use of the localization technique in the dynamical interpolation, where each observation affects the ability to recover the ocean in the nearby regions. The Appendix F includes a more sensitive analysis, which shows the improvement of the recovered ocean field if the density of observed floes increases.

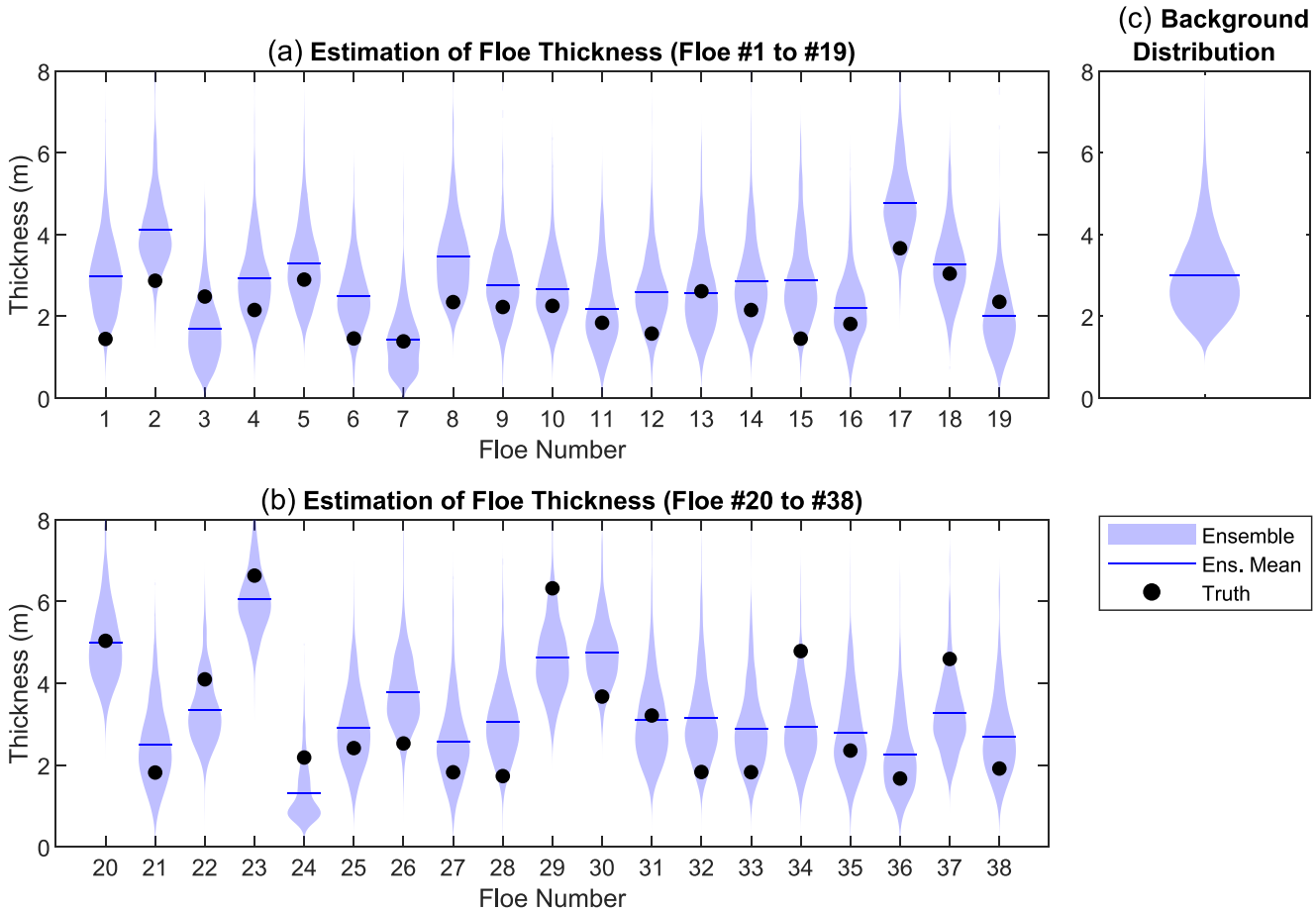

Figure 4. Parameter estimation of sea ice thickness in the synthetic data experiment. Panels (a) and (b): The black dots and solid lines indicate the truth and the ensemble mean estimate of each sea ice floe, respectively, while the shaded area in the violin plot indicates the estimated non-Gaussian PDF formed by ensembles. Panel (c): The background sea ice thickness distribution. The true value of the thickness for each sea ice floe is randomly drawn from such a distribution. It is also used as the initial distribution in the parameter estimation algorithm.

3.3. Results of the Real Data Experiment

Panel (b) of Figure 5 includes four cases of the recovered missing floes on the real data set. Similar to the conclusion from the synthetic data experiment, the dynamical interpolation being applied to the real data set also shows significant advantages over the linear interpolation in the sense that the error in the ensemble mean is overall much smaller and the uncertainty can be systematically quantified. Comparing with the analogs from the synthetic data experiment in Panel (a), the accuracy of the results in the real data test remains comparable. Figure 7 includes additional case studies of the recovered missing floe trajectories from the real data experiment. Again, the dynamical interpolation provides reasonable results in most of the cases. Panel (b) of Figure 6 displays the recovered ocean field on July 1, 2008. Although there is no true solution for the validation of the point-wise recovery skill, the overall flow amplitudes as well as the number and the size of the eddies in the recovered ocean field all look reasonable. One interesting finding is that the recovered ocean field in the north-east corner of the domain is nearly zero due to high uncertainty, which corresponds to the area beneath the large piece of the ice cover shown in Figure 3.

Figure 8 compares the physical properties of the recovered ice floes between the real observations, the dynamically interpolated data, and the direct model simulation. Since the data set consists of discrete observations, the two metrics used are the discrete curvature and the daily angular displacement. The former is calculated using the circumscribing circle of each trio of observations, while the latter is obtained by taking the difference in angle

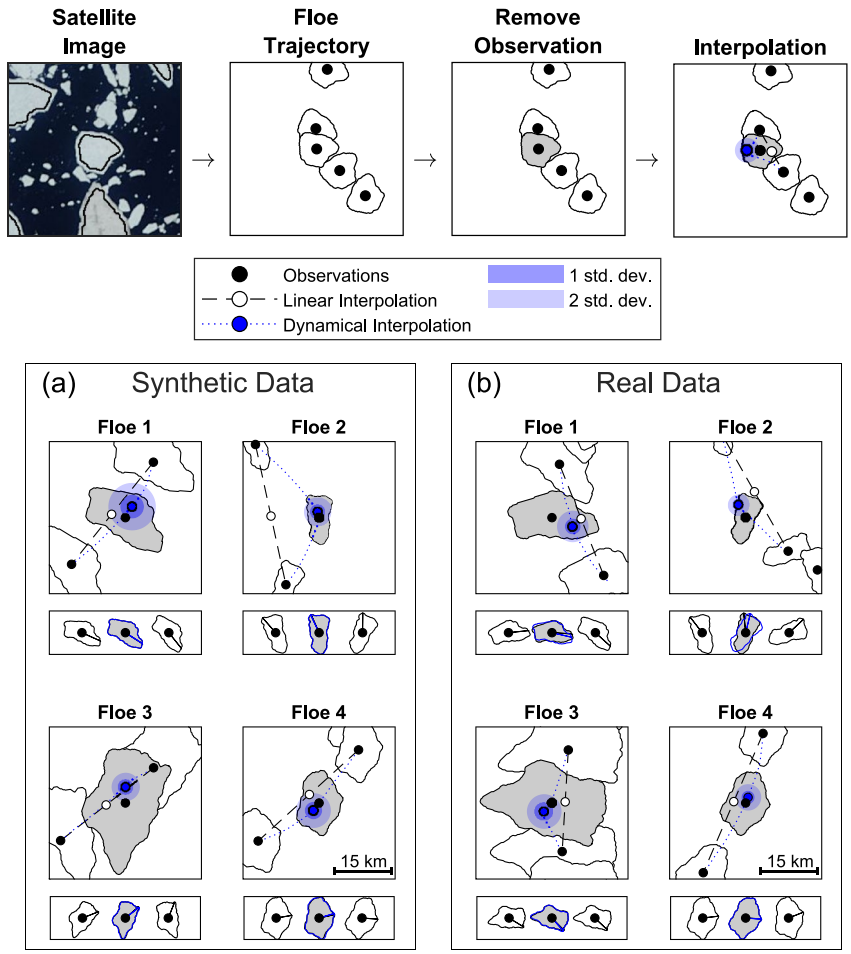

Figure 5. Comparison of recovering the missing observations using linear and dynamical interpolation schemes. The top panel illustrates the procedure of performing the interpolation experiments. A floe trajectory is first retrieved from the satellite imagery. Next, the observed floe on a specific day is artificially removed. The linear/dynamical interpolation framework is applied to recover this artificially removed observation. In the bottom part, panels (a) and (b) show the results from the synthetic and the real data experiments, respectively. In each panel, the top part shows the interpolated floe locations while the bottom part shows the interpolated angular displacement. Since the floes in the synthetic data experiment are taken from the library of sea ice floe observations, floes with the same index in the two experiments are identical (i.e., shapes and sizes are retained). In addition to the ensemble mean estimate presented by the blue marker, the uncertainty resulting from the dynamical interpolation is provided by the shaded areas. For the illustration purpose, only the two-dimensional Gaussian confidence interval is used to characterize the uncertainty in the dynamical interpolation.

between two consecutive observations. The results using the linear interpolation are omitted here, as the linear interpolation fails to provide any useful information of these two physical quantities. Panel (a) shows that the curvature of the floe trajectories from the direct model simulation is severely underestimated, which is a natural outcome of the model error. In contrast, the data resulting from the dynamical interpolation succeeds in reproducing the non-Gaussian distribution of the observed truth with a one-sided fat tail. Next, with respect to the angular displacement, as is shown in Panel (b), the real data set has a negative bias due to the influence of the Beaufort Gyre, something which is not reflected in the direct model simulation that is again due to the model error. Nevertheless, such a bias in the direct model simulation is almost fully corrected in the dynamically interpolated data with the help of partial observations. These results indicate the importance of utilizing both the observations and a suitable model in the dynamical interpolation, as the model provides at least partial access to the crucial underlying nonlinear dynamical information. In contrast, the observations can largely reduce the biases from the model forecast.

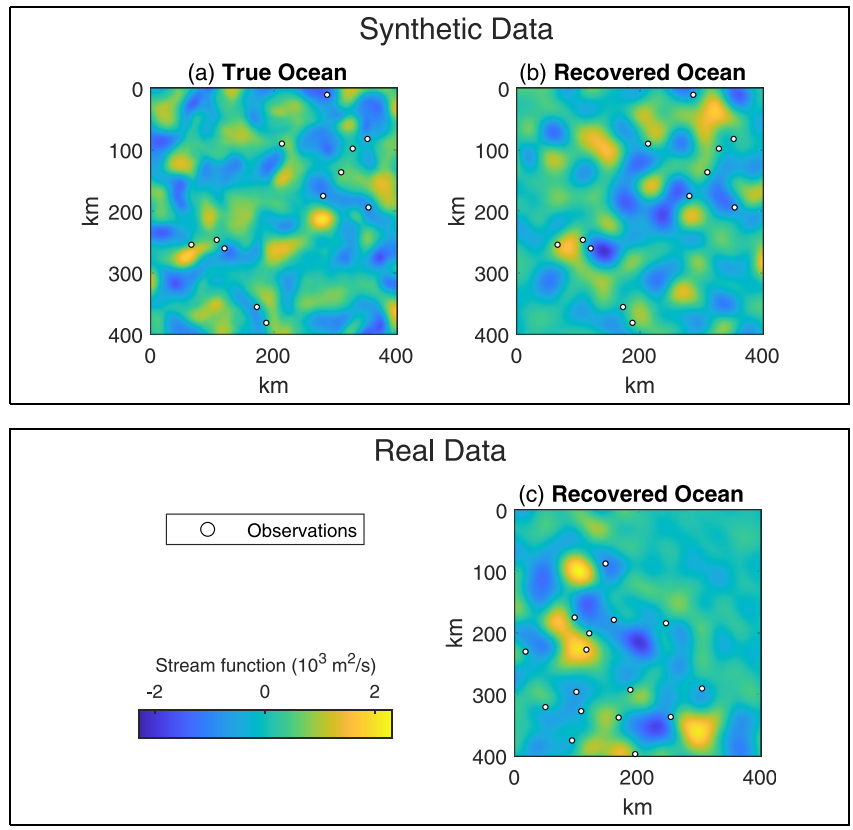

Figure 6. The recovered ocean flow field represented by the stream functions utilizing the dynamical interpolation. The top panel shows the truth and the recovered ocean field in the synthetic data experiment while the bottom panel shows the recovered ocean field of the real data. Since the primary focus is to resolve regions close to the ice edge, a 400 × 400 km domain within the original 600 × 600 km area and having the same domain center is presented. The pattern correlation between the true and recovered ocean in the 400 × 400 km subdomain is 0.32. The results shown here are on a specific day in the middle of the study period. For the real data, it is July 1. The error in recovering the ocean field remains in a similar level on other days. The white dots mark the locations of the floes. There are in total 17 floes inside the 400 × 400 km domain for both cases.

4. Conclusions and Discussion

Model error is inevitable when studying complex systems. In the dynamical interpolation framework developed here, the sea ice DEM model remains highly nonlinear while the ocean and atmospheric components are effectively approximated by linear stochastic models. Indeed, a significant error will appear if these linear stochastic models are used to study the dynamics associated with the ocean and atmospheric fields. Nevertheless, for the purpose of dynamical interpolation, the information needed from the model is merely some prior knowledge of the short-range statistical forecast of these fields, which are usually quite accurate due to the fact that these linear stochastic models are carefully calibrated.

The reduced-order models in the proposed framework are not limited to linear stochastic models. If the time series of the underlying flow fields exhibit strong non-Gaussian features, then suitable nonlinear or non-Gaussian surrogate models can be easily incorporated (Chen & Majda, 2018; Edeling & Crommelin, 2020; Gershgorin et al., 2010; Khodkar & Hassanzadeh, 2018; Majda & Qi, 2018). In particular, one such simple candidate is a family of linear models with multiplicative noise (Averina & Artemiev, 1988). Another potential alternative is the multilayer stochastic models (Kondrashov et al., 2015, 2018). On the other hand, the sea ice dynamics within the scales studied here are predominantly nonlinear. The governing equations are well understood and are crucial in the dynamical interpolation, given that the directly observed variables are sea ice floe trajectories. The strong nonlinearity in sea ice dynamics is also more deterministic and less turbulent than the atmosphere or the ocean. Therefore, linear stochastic models are not appropriate for approximating the fully nonlinear behavior of sea ice.

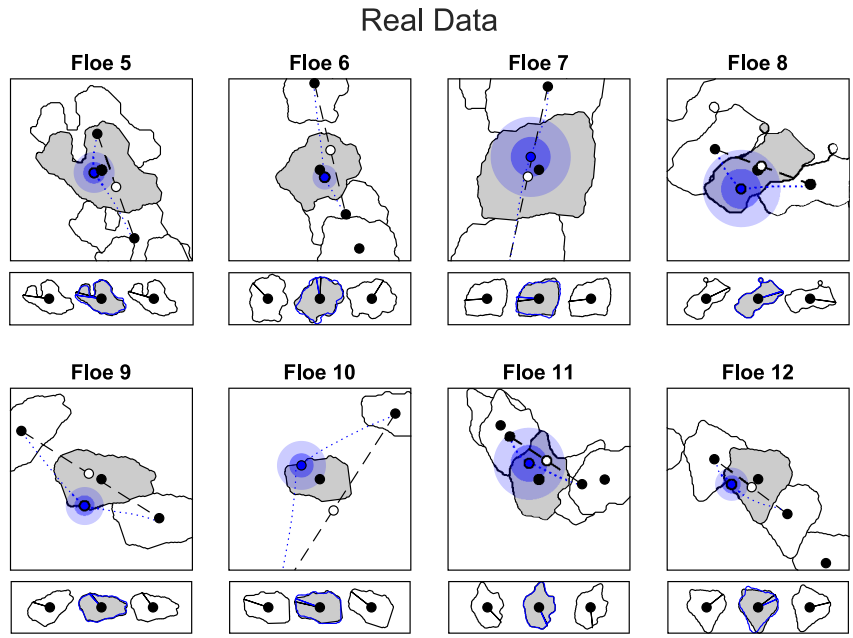

Figure 7. Additional results for the real data experiment, similar to those in Panel (b) of Figure 5.

Since the degree of freedom in characterizing the floe trajectories is much lower than the governing equations of the atmospheric and oceanic velocity fields, the nonlinear floe dynamics are explicitly incorporated into the dynamical interpolation framework.

It is also worth highlighting the importance of the prior physical model of the ocean and the prior time series of the atmosphere, which significantly facilitates the calibration of the linear stochastic models. In the absence of a suitable prior model or data for the ocean and the atmosphere, the calibration of the reduced-order stochastic surrogate models requires a more complicated iterative expectation-maximization procedure (Chen, 2020). In other words, the dynamical interpolation, the parameter estimation of the thickness, and the uncertainty quantification of the oceanic and atmospheric flow fields have to be carried out simultaneously with the floe trajectories providing the only available information. Such an iterative approach often requires an extensive observational database to ensure the accuracy of the dynamical interpolation scheme.

Another crucial point is the quantification of the uncertainty, particularly in the presence of the model error, the small number of observations, and the implications of turbulent systems. Uncertainty quantification is not available by applying direct curve fitting methods but rather a unique feature of the dynamical interpolation framework. In this study, the properties of sea ice exhibit various non-Gaussian features such as the non-symmetry in the distribution of the angular displacement, strong skewness, and fat tails with extreme events in the distributions of the curvature and thickness. These non-Gaussian features have been shown to be crucial in understanding the sea ice dynamics (Moon & Wettlaufer, 2017; Toppaladoddi & Wettlaufer, 2015). Hence, to assess uncertainty, the attributes of the entire distribution are considered.

Finally, the framework developed here has several unique implications for improving our understanding of Earth system science. First, new-generation climate models that accurately represent sea ice dynamics at the floe scale will require validation against Lagrangian observations of sea ice floes at high and moderate resolutions. The point-estimate recovery of missing observations and the associated estimates of the uncertainty mitigate some of the issues of Lagrangian optical remote sensing observations. The methodology presented here is also easily adaptable to analyze the output of other instruments. In this sense, it is expected that the continuous trajectories stemming from the nonlinear data assimilation can be used in more sophisticated deep-learning models for calibration and training. Second, the proposed framework allows for accurate parameter estimation of unobserved variables at unprecedented scales in a Lagrangian setting. For example, the retrieval of sea ice thickness is outlined here, which is a crucial variable in understanding the evolution of the sea ice cover in response to a changing climate. Lastly, data assimilation provides the key missing piece for understanding ocean transport and

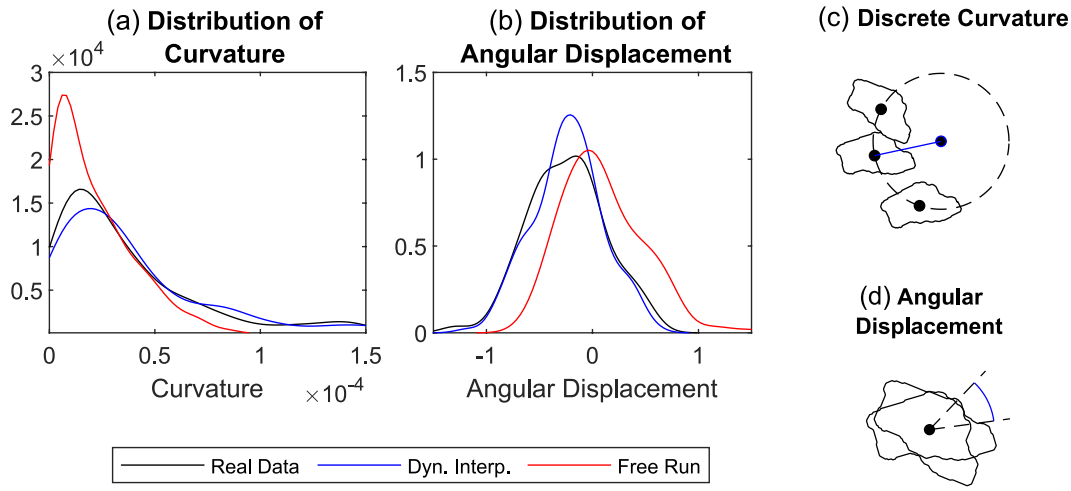

Figure 8. Comparison of the recovered properties using different interpolation methods. Panel (a): comparison of the distribution of the curvature of the recovered trajectories. Panel (b): comparison of the distribution of the angular displacement of the recovered trajectories. Panels (c)–(d): Schematic illustrations of the definitions of the discrete curvature and angular displacement used to compute the distributions in Panels (a)–(b).

mixing processes at high latitudes. Small-scale eddies have an important role in transferring energy to larger-scale structures via an inverse cascade of energy and are thus hypothesized to be the missing energy source to close the ocean energy budget. High-resolution numerical simulations have highlighted their contribution to nutrient redistribution, oxygen transport, and biogeochemical processes. However, they are hard to observe due to the lack of resolution of space-borne sensors and the sparsity of in situ instruments. Given the recently demonstrated connections between the rotation rate of sea ice floes and eddies with sub-surface expression in the western Arctic Ocean, it is anticipated that this method can be used to understand fundamental processes of ocean turbulence at small-to-moderate scales.

Appendix A: The Coupled Atmosphere-Ice-Ocean System

A1. The DEM Model

The sea ice floes are described using a DEM model. The floes can have arbitrary 2D shapes with their movements and rotations determined by the surface integrals of the ocean and wind velocities over these shapes. Both the full QG model and reanalysis data or and the approximate stochastic models can be used to drive this ice floe model. In the DEM model utilized here, the shape and thickness for each floe are assumed to be unchanging over time. Since the noninteracting floes are the primary focus of this work, the contact forces are not included in the model presented here, which greatly reduces the computational cost.

The dynamics of a single ice floe is described as follows (Chen et al., 2022; Manucharyan et al., 2022): Let $\boldsymbol{X}_{\text{ice}} = (x_{\text{ice}}, y_{\text{ice}})$ be the centroid of the floe and Ω be the angular displacement about the centroid. Also let $\boldsymbol{V}_{\text{ice}} = (u_{\text{ice}}, v_{\text{ice}})$ be the velocity of the floe and ω be the angular velocity. Then, ice floe-ocean interactions are calculated using surface integrals over the area of the floe:

$$\dot{\mathbf{X}}_{\text{ice}} = \mathbf{V}_{\text{ice}} \tag{A1}$$

$$\dot{\Omega} = \omega \tag{A2}$$

$$m\dot{\mathbf{V}}_{\text{ice}} = \iint_A \mathbf{F}_{\text{total}}\, \mathrm{d}A \tag{A3}$$

$$I\dot{\omega} = \iint_A \tau\, \mathrm{d}A \tag{A4}$$

where A is the area of the floe. $\mathbf{F}_{\text{total}}$ is the total force on the ice floe induced by the ocean, atmosphere, and other sources. τ is the resulting torque, calculated from the force.

The total force on the floe by the ocean consists of ocean drag, atmosphere forcing, Coriolis force, and the pressure gradient

$$\mathbf{F}_{\text{total}} = \mathbf{F}_{\text{ocn}} + \mathbf{F}_{\text{atm}} + \mathbf{F}_{\text{Coriolis}} + \mathbf{F}_{\text{pressure}}. \tag{A5}$$

To incorporate the ocean turning angle, define the rotation matrix $\mathbf{R}_\theta$ as

$$\mathbf{R}_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \tag{A6}$$

The force induced by the ocean drag at the point $\boldsymbol{X}_{\text{ocn}}$ within A is given by

$$\mathbf{F}_{\text{ocn}} = \rho_{\text{ocn}} C_{\text{ocn}} \|\boldsymbol{V}_{\text{ocn}} - \boldsymbol{V}_{\text{ice}}\| \mathbf{R}_\theta \left(\mathbf{V}_{\text{ocn}} - \mathbf{V}_{\text{ice}}\right) \tag{A7}$$

where θ is the fixed ocean turning angle. Similarly, the force induced by the atmosphere at the point $\boldsymbol{X}_{\text{atm}}$ within A is given by

$$\mathbf{F}_{\text{atm}} = \rho_{\text{atm}} C_{\text{atm}} \|\boldsymbol{V}_{\text{atm}} - \boldsymbol{V}_{\text{ice}}\| \left(\mathbf{V}_{\text{atm}} - \mathbf{V}_{\text{ice}}\right). \tag{A8}$$

No atmosphere turning angle was used in this study. The Coriolis force induced on the floe is constant over the area of the of the floe and is given by

$$\mathbf{F}_{\text{Coriolis}} = \rho_{\text{ice}} f_c L_{\text{ice}} \mathbf{R}_{-\pi/2} \mathbf{V}_{\text{ice}} \tag{A9}$$

where f_c is the Coriolis coefficient and L_{ice} is the thickness of the ice floe. The force induced by the pressure gradient of the ocean acting on the floe depends on the ocean velocity and so varies over the floe area, A

$$\mathbf{F}_{\text{pressure}} = \rho_{\text{ice}} f_c L_{\text{ice}} \mathbf{R}_{\pi/2} \mathbf{V}_{\text{ocn}}. \tag{A10}$$

The pressure gradient of the ocean is in geostrophic balance with the Coriolis force induced on the ocean. The torque induced on the floe at the grid point $\boldsymbol{X}_{\text{torque}}$ is given by

$$\tau = \left(\mathbf{X}_{\text{torque}} - \mathbf{X}_{\text{ice}}\right) \times \mathbf{F}_{\text{total}} = \left(x_{\text{torque}} - x_{\text{ice}}\right) F_y - \left(y_{\text{torque}} - y_{\text{ice}}\right) F_x \tag{A11}$$

where $\boldsymbol{F}_{\text{total}} = (F_x, F_y)$ are the components of the total force.

A2. The Two-Layer QG Model

The ocean model is a two-layer QG model with periodic boundary conditions on a square domain. The ocean state is characterized by the stream functions $\psi_i(x, y)$ and potential vorticities (PV) $q_i(x, y)$ of each layer $i = 1, 2$. These quantities, ψ and q, are deviations from a background mean state. The level curves of the stream function, ψ_i, correspond to streamlines of the velocity field, which guarantees an incompressible flow. The ocean velocity field for each layer can thus be calculated as

$$(u_i, v_i) = \left(-\frac{\partial \psi_i}{\partial y}, \frac{\partial \psi_i}{\partial x}\right), \quad i = 1, 2. \tag{A12}$$

The formulation of the QG equations follows the version in Arbic and Flierl (2004). The PDEs that govern the time evolution of ψ_i and q_i are as follows:

$$\frac{\partial q_1}{\partial t} + \overline{u_1}\frac{\partial q_1}{\partial x} + \frac{\partial \overline{q}_1}{\partial y}\frac{\partial \psi_1}{\partial x} + J\left(\psi_1, q_1\right) = \text{ssd} \tag{A13}$$

$$\frac{\partial q_2}{\partial t} + \overline{u_2}\frac{\partial q_2}{\partial x} + \frac{\partial \overline{q_2}}{\partial y}\frac{\partial \psi_2}{\partial x} + J\left(\psi_2, q_2\right) = -R_2 \nabla^2 \psi_2 + \text{ssd}. \tag{A14}$$

Here "ssd" represents small-scale dissipation, which are higher-order derivative terms that are ignored. J is the Jacobian

$$J(\psi, q) = \frac{\partial \psi}{\partial x}\frac{\partial q}{\partial y} - \frac{\partial \psi}{\partial y}\frac{\partial q}{\partial x}. \tag{A15}$$

The stream functions further satisfy

$$q_1 = \nabla^2\psi_1 + \frac{(\psi_2 - \psi_1)}{(1+\delta)L_d^2} \qquad q_2 = \nabla^2\psi_2 + \frac{\delta(\psi_1 - \psi_2)}{(1+\delta)L_d^2}. \tag{A16}$$

where $\delta = H_1/H_2$, H_i is the depth of each layer, and L_d is the deformation radius.

$\overline{u_1}$ and $\overline{u_2}$ are the mean ocean velocities for each layer. $\partial\overline{q}_1/\partial y$ and $\partial\overline{q}_2/\partial y$ are the mean of the PV gradients for each layer and are given explicitly by

$$\frac{\partial \overline{q_1}}{\partial y} = \frac{\overline{u_1} - \overline{u_2}}{(1+\delta)L_d^2}\frac{\partial \overline{q_2}}{\partial y} = \frac{\delta\left(\overline{u_2} - \overline{u_1}\right)}{(1+\delta)L_d^2}. \tag{A17}$$

The final parameter, R_2, is the decay rate of the barotropic mode

$$R_2 = \frac{f_0 d_{\text{Ekman}}}{2H_2} \tag{A18}$$

where f_0 is the Coriolis parameter and d_{Ekman} is the bottom boundary layer thickness. Note that in this formulation, we use a constant Coriolis force throughout the domain.

Table A1 summarizes the parameters in the DEM and two-layer QG models.

Table A1
Parameters in the Discrete Element Method, the Two-Layer QG Models and the Ensemble Kalman Smoother

Parameter	Value
Ocean density	$\rho_{\text{ocn}} = 1027$ kg/m^3
Ice density	$\rho_{\text{ice}} = 920$ kg/m^3
Air density	$\rho_{\text{atm}} = 1.2$ kg/m^3
Ocean drag coefficient	$c_{\text{ocn}} = 5.5 \times 10^{-3}$
Atmosphere drag coefficient	$c_{\text{atm}} = 1.6 \times 10^{-3}$
Coriolis coefficient	$f_c = 1.4 \times 10^{-4}$
Top layer mean ocean velocity	$\overline{u_1} = 2.58$ km/day
Bottom layer mean ocean velocity	$\overline{u_2} = 1.032$ km/day
Top layer mean potential vorticity	$\frac{\partial\overline{q_1}}{\partial y} = 0.0265$ km^{-1} day^{-1}
Bottom layer mean potential vorticity	$\frac{\partial\overline{q_2}}{\partial x} = -0.0212$ km^{-1} day^{-1}
Coriolis parameter	$f_c = 12$ days^{-1}
Coupling parameter	$R_1 = 6.9 \times 10^{-5}$ km^{-1}
Decay rate of the barotropic mode	$R_2 = 1$ day^{-1}
Deformation radius	$L_d = 5.7$ km
Ratio of upper-to lower-layer depth	$\delta = 0.8$
Turning angle of the ocean	$\theta = \pi/9$
Ensemble size	600
Localization radius	200 km
Observational noise in location	250 m
Observational noise in angular displacement	5°

A3. The Atmospheric Wind Velocity Data

The fifth generation ECMWF reanalysis data product (ERA5) (Copernicus Climate Change Service (C35), 2017; Olauson, 2018) for the global climate and weather is implemented for describing the atmospheric wind that is used to calibrate the atmospheric component of the linear stochastic models.

Appendix B: Sea Ice Floe Observations and the Processing of Satellite Images

Remote sensing measurements were retrieved from Moderate Resolution Imaging Spectroradiometer (MODIS) optical imagery (Level 1B 250 M). The data are open-access through the Earth Observing System Data and Information System (EOSDIS) Worldview platform (https://worldview.earthdata.nasa.gov). In summary, both Corrected Reflectance True and False Color images were preprocessed to reduce the imprint of atmospheric noise, allowing the segmentation of sea ice floes ranging from 4 to 75 km in the length scale as individual objects. Ice floes were then tracked in a three-stage process involving comparing geometrical parameters in successive images, finding potential matches, and selecting the best candidates based on the assessment of a similarity metric and surface area differences. The reader is referred to Lopez-Acosta et al. (2019) for a detailed description of the preprocessing, segmentation, and tracking routines.

Appendix C: Calibration of Stochastic Forecast Models

Statistically accurate stochastic models are used for the ocean and atmospheric components of the forecast model. These models can be systematically calibrated based on a 20-year simulation of the two-layer QG model in the case of the ocean component and reanalysis data set in the case of the atmospheric component. In both cases, the system state is represented in spectral space and the evolution of each spectral mode is governed by a linear stochastic model (Equation 1). Only the modes with a wave number less than a certain radius are kept: $|k| \leq 11$ in the case of the ocean and $|k| \leq 5$ in the case of the atmosphere for a total of 337 and 81 modes, respectively.

Recall the linear stochastic model in Equation 1, which is also known as the complex Ornstein-Uhlenbeck (OU) process (Gardiner, 2009). In Equation 1, a, ω, and σ are real-valued parameters with $a, \sigma > 0$, f is a complex-valued parameter, and $\dot{W}$ is a complex-valued white noise. The equilibrium distribution of this OU process is Gaussian and its mean and variance are given in terms of the model parameters:

$$\bar{u} = \frac{f}{a - i\omega} \qquad \mathrm{Var}(u) = \frac{\sigma^2}{2a}. \tag{C1}$$

The decorrelation time is defined as

$$T = \int_0^\infty \frac{\mathbb{E}\left[\left(u(t) - \bar{u}\right)\left(u(t+\tau) - \bar{u}\right)^*\right]}{\mathrm{Var}(u)}\, d\tau \tag{C2}$$

where the expectation is taken over t. The decorrelation time is also given in terms of the model parameters

$$T = \frac{1}{a - i\omega}. \tag{C3}$$

Using these equation for the equilibrium mean, variance, and decorrelation time, the four parameters of the OU process, d, ω, f, and σ, can be written explicitly in terms of these equilibrium statistics as in

$$d = \mathrm{Re}\left[\frac{1}{T}\right] \qquad \omega = -\mathrm{Im}\left[\frac{1}{T}\right] \qquad f = \frac{\bar{u}}{T} \qquad \sigma = \sqrt{2\mathrm{Var}(u)\mathrm{Re}\left[\frac{1}{T}\right]}. \tag{C4}$$

Using these formulae for the model parameters, an OU process can be fit to match a given set of equilibrium statistics. In the case of the ocean model, an independent OU process is fit to each spectral mode of the stream function using the mean, variance, and decorrelation time of a 20-year QG model simulation. For the atmosphere model, a pair of independent OU processes is fit to each spectral mode of the two-dimensional velocity field using the statistics of the velocity field from the ERA5 data set.

Appendix D: Ensemble Update

Let K denote the number of days of observations and denote the time of each day of observations by t_k for $k = 1, \ldots, K$. Denote the M-dimensional system state at time t by $\boldsymbol{\psi}(t)$ for $t_1 \leq t \leq t_K$. Then define the vector $\boldsymbol{d}_k$ of observed floe locations and orientations at time t_k by

$$\boldsymbol{d}_k = \mathcal{M}_k\left[\boldsymbol{\psi}\left(t_k\right)\right] + \boldsymbol{\epsilon}. \tag{D1}$$

$\mathcal{M}_k$ returns only the subset of system variables corresponding to the observed floe positions and orientations at time t_k. The dependence on k allows for a changing number of observed floes at each observation time. $\boldsymbol{\epsilon}$ is a small Gaussian observational noise, corresponding to the resolution of the satellite images.

The EnKS, and all smoothing algorithms, estimates

$$f\left(\boldsymbol{\psi}(t)|\boldsymbol{d}_k, \ldots, \boldsymbol{d}_0\right), \tag{D2}$$

the probability distribution of the state variable $\boldsymbol{\psi}(t)$ at a time $t_0 < t < t_k$ given the available observations $d_0, \ldots, d_k$. The estimation of the variable at any time t in the interval (t_0, t_k) contrasts smoothing with filtering, which only estimates the state at the present time t_k. This distribution at time t_k can be calculated exactly using Bayes' theorem,

$$f\left(\boldsymbol{\psi}\left(t_k\right)|\boldsymbol{d}_0, \ldots, \boldsymbol{d}_k\right) \propto f\left(\boldsymbol{\psi}(t_k)|\boldsymbol{d}_0, \ldots, \boldsymbol{d}_{k-1}\right) f\left(\boldsymbol{d}_k|\boldsymbol{\psi}(t_k)\right), \tag{D3}$$

a process called Bayesian inference. The left-hand side of equation (D3) is the posterior distribution while the right-hand side consists of two factors: the prior or forecast distribution and the observational model. The EnKS represents the distribution of the state variable with an ensemble of model trajectories. This ensemble is then iteratively forecast and updated for each set of observations, which is processed sequentially in time, using equation (D3). Note that the difference between the ensemble Kalman filter (EnKF) and EnKS is that the former only uses the information up to the current time instant while the latter uses all the available observational information, including the future one. Thus, the EnKF is more appropriate for the providing an improved initialization of the forecast while the EnKS applies to dynamical interpolation.

Let N be the size of the ensemble and denote individual ensemble members by $\boldsymbol{\psi}_k^{(n)}(t)$ for $n = 1, \ldots, N$. The superscript "(n)" distinguishes individual ensemble members from the true system state denoted by $\boldsymbol{\psi}(t)$. The subscript k denotes that the ensemble has been updated using the first k observations. To compute $\boldsymbol{\psi}_k^{(i)}(t)$ for $t > t_k$, the forecast model is used. While only the value of $\boldsymbol{\psi}_k^{(i)}\left(t_{k+1}\right)$ is required to perform the ensemble update, the ensemble at any prior time t for $t_1 \leq t \leq t_{k+1}$ can be stored in memory and updated for each new observation.

Once the ensemble has been updated for observation k, to assimilate observation $k + 1$, the $M \times N$ matrix of the ensemble members is formed

$$\boldsymbol{A}_k(t) = \left(\boldsymbol{\psi}_k^{(1)}(t) \quad \boldsymbol{\psi}_k^{(2)}(t) \quad \cdots \quad \boldsymbol{\psi}_k^{(N)}(t)\right). \tag{D4}$$

The forecast ensemble matrix, $\boldsymbol{A}_{k+1}^{\mathrm{f}} = \boldsymbol{A}_k\left(t_{k+1}\right)$, is calculated using the forecast model. Then the updated ensemble is calculated using an $N \times N$ linear transformation, $\boldsymbol{T}$, of the ensemble

$$\boldsymbol{A}_{k+1}(t) = \boldsymbol{A}_k(t)\boldsymbol{T} \tag{D5}$$

for any $t_1 \leq t \leq t_{k+1}$ where T is formed using the Kalman filter equations from the forecast ensemble, $\boldsymbol{A}_{k+1}^{\mathrm{f}}$, and the observations, $\boldsymbol{d}_{k+1}$.

The above method is modified slightly to utilize localization, which leverages the spatial structure of the system to reduce the negative effects of spurious correlations. During the update at time t_k, each variable in $\boldsymbol{\psi}$ is associated with a location in physical space. For variables of an observed floe, their assigned location

is the position of the observation. For an unobserved floe, the forecast mean position is used. For ocean and atmosphere variables, the spectral representations are transformed to physical space and the location for each grid point is used.

To update under localization, each state variable in $\boldsymbol{\psi}$ is updated individually. An ensemble corresponding to each state variable is formed from taking the values from the ensemble of model trajectories. This ensemble is then updated using only observations that are within a fixed radius of that variable's associated spatial location. The matrix $\boldsymbol{T}$ is formed using the full forecast ensemble, $\boldsymbol{A}_k^{\mathrm{f}}$, and the observation vector $\boldsymbol{d}_k$ containing only the observations within the localization radius. Then the $1 \times M$ row vector of the variable's ensemble is updated using this localized version of $\boldsymbol{T}$. This process is repeated to update all localized variables in $\boldsymbol{\psi}$.

Appendix E: Parameter Estimation

Unknown parameters, such as individual floe thicknesses, can be estimated within the proposed framework. The parameters are appended to the state vector and treated as nondynamical variables. In other words, the evolution equations of the parameters $\boldsymbol{\theta}$ are $\mathrm{d}\boldsymbol{\theta}/\mathrm{d}t = 0$. The initial values of the parameters in each ensemble member are drawn from a background distribution (displayed in panel (c) of Figure 4). Therefore, the initial ensemble includes the uncertainty of the parameters. During the ensemble forecast, the parameter values are kept constant. However, during the ensemble update, the distribution of parameter values among the ensemble members is updated as well according to the same linear transformation. In this way, the distribution of parameter values changes over the course of the algorithm even though the values do not change during the forecast. Since the parameters are nondynamical, trajectories of these parameters do not need to be considered and only the current values of the parameters need to be stored.

Appendix F: Sensitivity Analysis of the Ocean Recovery

The ocean is represented in the system state by the stream function, which encodes an incompressible flow. The synthetic data was generated using a QG incompressible ocean model. For nonlinear data assimilation, a statistically accurate stochastic forecast model is used to forecast the ensemble of model trajectories. This ensemble is then updated after each set of observations, allowing any model variables, including the ocean stream function, to be recovered at any point in time. For these experiments, the ocean is recovered for the day of July 1, which is roughly in the middle of the set of observations. Note that the ocean evolves slowly with time, and so, this is representative of the overall ocean recovery. In addition to the standard setup in the main text, the following two additional experiments of the ocean flow field recovery are included.

Figure F1 shows the recovery of the ocean stream function using synthetic data without atmosphere forcing, that is, the atmospheric component of the sea ice model is removed for both the generation of synthetic data and the forecast model. In this case, the dynamics of the floes are fully accounted for by ocean forces and the total degree of freedom of the unobserved variables is reduced. Therefore, a better recovery of the ocean state is expected due to the lack of interference from the atmosphere. Indeed, the results of the ocean recovery seem to be improved. However, the lack of atmosphere forcing, which is responsible for the large-scale floe movements, leads to floe trajectories that travel a shorter distance and hence cover a smaller portion of the region. This impacts localization during the ocean ensemble update, which removes the influence of observed floes that fall outside the localization radius of each ocean grid point. In this case, the amplitude of the recovered ocean is nearly zero in some areas, which reflects the uncertainty of the ensemble rather than an actual amplitude of individual ensemble members.

In another experiment, Figure F2 shows the ocean recovery for synthetic data that is generated with atmospheric forcing but using a 200 km by 200 km ocean (rather than a 600 km by 600 km ocean) that is extended periodically to the entire domain. This is analogous to a case in which a much higher density of observations is available. No localization is utilized here, as the typical localization radius of 200 km covers the entire ocean. In this case, the qualitative recovery of the ocean is quite good, demonstrating that a sufficiently high density of floe observations can recover ocean features well.

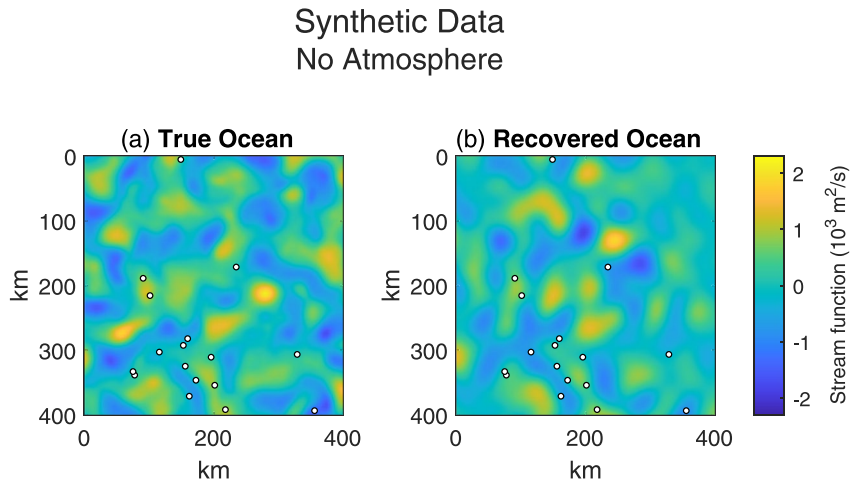

Figure F1. A comparison of the true and recovered ocean on July 1 for a synthetic data experiment with no atmosphere forcing. Panel (a) shows the stream function of the true ocean generated from a QG ocean model. Panel (b) shows the recovery of the stream function using the ensemble mean. The pattern correlation between the true and recovered ocean in the 400 × 400 km subdomain is 0.43. Each plot shows white dots indicating the position of the observed floes. While a 400 km by 400 km region is plotted, the ocean is extends another 100 km on all sides to reduce the impact of the periodicity in the ocean model.

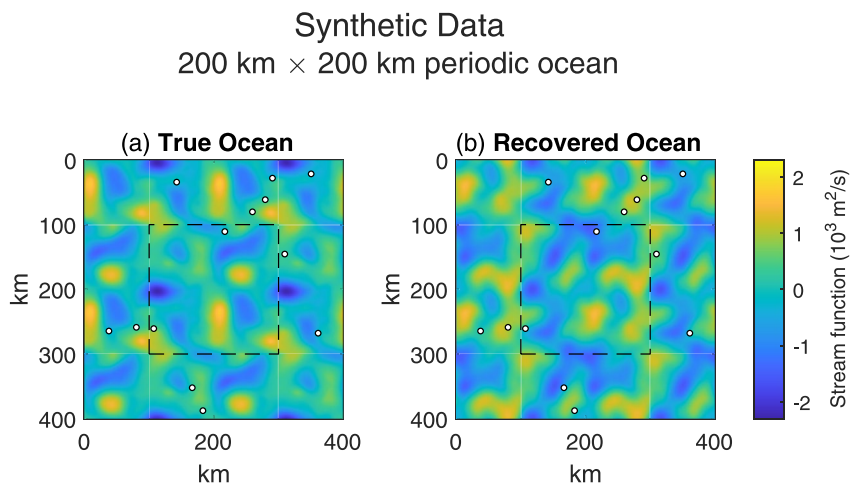

Figure F2. A comparison of the true and recovered ocean stream functions for a synthetic data experiment using a 200 km by 200 km ocean extended periodically to the whole domain. Panel (a) shows the true ocean, generated from a QG model on a 200 km by 200 km domain. The dashed black line indicates the size of the ocean, which is then periodically extended to the whole domain. Panel (b) shows the ensemble mean stream function. The pattern correlation between the true and recovered ocean is 0.60. The floe positions are indicated with white dots. Note that observations outside of the 200 km by 200 km ocean region influence the ensemble update of the ocean just as much as observations inside the region due to the periodicity of the ocean.

Appendix G: Behavior of Ensemble Members in the Dynamical Interpolation

The ensemble mean provides a complete trajectory and is the best point estimate of a floe's position available from the dynamical interpolation method. However, despite the fact that each ensemble member is based on the model forecast corrected by the partial observations, their average—the ensemble mean—is often not a physically consistent trajectory of the model. Therefore, it is important to look at each individual ensemble member and understand the additional physical properties that are not fully reflected in the ensemble mean.

Figure G1 shows an example of a continuous floe trajectory generated from the sea ice model, along with the daily observed floe positions. The trajectories between daily observations are then interpolated using different methods. Note that this floe trajectory contains a loop midway through, which is compared qualitatively with the interpolation using different methods. While the ensemble mean provides better physics than linear interpolation and contains nonlinear evolution of the trajectory, it still fails to capture this loop. On the other hand, more than half of the ensemble members recover such a loop around the correct location, something which is otherwise lost during averaging. The finding here reveals the potential role that utilizing individual ensemble members can play in identifying properties of floe trajectories that are not otherwise captured by the ensemble mean. Such a result also highlights the importance of considering the uncertainty represented by the ensemble members in addition to the mean state estimation.

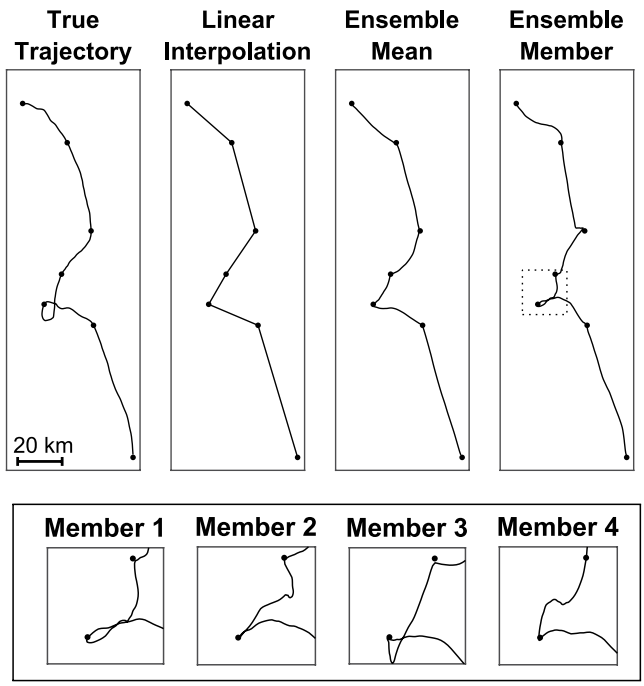

Figure G1. Comparison of different methods for the interpolation of floe trajectories. The first panel shows a floe trajectory generated from the sea ice model and sampled every time unit at the black dots. This trajectory contains a loop. The next three panels show various forms of interpolation. The first shows linear interpolation. The second shows the ensemble mean. The third shows a sampled ensemble member. The bottom panels show other sampled ensemble members.

Data Availability Statement

A library of floe observations (Lopez-Acosta & Wilhelmus, 2021) was used for the real data set experiments and to initialize the synthetic data experiments (https://doi.org/10.5281/zenodo.4796844). The ERA5 reanalysis data set (Copernicus Climate Change Service (C35), 2017) was used to calibrate the wind model and generate the synthetic data (https://doi.org/10.24381/cds.adbb2d47). The code used to process the data and create the figures was written in MATLAB. The code and output data of the experiments are available on Zenodo (https://doi.org/10.5281/zenodo.7026025).

Acknowledgments

The research of Nan Chen. was partially funded by the Office of Naval Research (ONR) Multidisciplinary University Initiative (MURI) award N00014-19-1-2421. Monica M. Wilhelmus was funded by the ONR Arctic and Global Prediction program (N00014-20-1-2753) and the ONR MURI program (N00014-19-1-2421). Jeffrey Covington was supported as research assistant under this grant and by the National Science Foundation award DMS-2023239 through the Institute for Foundations of Data Science (IFDS) at UW-Madison. The authors gratefully acknowledge Dr. Georgy Manucharyan for insightful discussions and Dr. Rosalinda Lopez-Acosta for her work on the development of the Ice Floe Tracker algorithm. The authors also gratefully acknowledge the reviewers of this manuscript for their valuable insights, feedback, and suggestions.

References

Anderson, J. L. (2012). Localization and sampling error correction in ensemble Kalman filter data assimilation. *Monthly Weather Review*, *140*(7), 2359–2371. https://doi.org/10.1175/mwr-d-11-00013.1

Arbic, B. K., & Flierl, G. R. (2004). Baroclinically unstable geostrophic turbulence in the limits of strong and weak bottom Ekman friction: Application to midocean eddies. *Journal of Physical Oceanography*, *34*(10), 2257–2273. https://doi.org/10.1175/1520-0485(2004)034<2257:bugtit>2.0.co;2

Ardhuin, F., & Prévost, C. (2020). Arctic Ocean sea ice drift reprocessed. *Copernicus Monitoring Environment Marine Service (CMEMS)*.

Averina, T., & Artemiev, S. (1988). Numerical solution of systems of stochastic differential equations. *Russian Journal of Numerical Analysis and Mathematical Modelling*, *3*(4), 267–286. https://doi.org/10.1515/rnam.1988.3.4.267

Berner, J., Achatz, U., Batte, L., Bengtsson, L., Cámara, A. d. L., Christensen, H. M., et al. (2017). Stochastic parameterization: Toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, *98*(3), 565–588. https://doi.org/10.1175/bams-d-15-00268.1

Bhatt, U. S., Walker, D. A., Walsh, J. E., Carmack, E. C., Frey, K. E., Meier, W. N., et al. (2014). Implications of Arctic sea ice decline for the Earth system. *Annual Review of Environment and Resources*, *39*(1), 57–89. https://doi.org/10.1146/annurev-environ-122012-094357

Branicki, M., Majda, A. J., & Law, K. J. (2018). Accuracy of some approximate Gaussian filters for the Navier–Stokes equation in the presence of model error. *SIAM Journal on Multiscale Modeling and Simulation*, *16*(4), 1756–1794. https://doi.org/10.1137/17m1146865

Brunette, C., Tremblay, L. B., & Newton, R. (2022). A new state-dependent parameterization for the free drift of sea ice. *The Cryosphere*, *16*(2), 533–557. https://doi.org/10.5194/tc-16-533-2022

Chen, N. (2020). Learning nonlinear turbulent dynamics from partial observations via analytically solvable conditional statistics. *Journal of Computational Physics*, *418*, 109635. https://doi.org/10.1016/j.jcp.2020.109635

Chen, N., Fu, S., & Manucharyan, G. E. (2022). An efficient and statistically accurate Lagrangian data assimilation algorithm with applications to discrete element sea ice models. *Journal of Computational Physics*, *455*, 111000. https://doi.org/10.1016/j.jcp.2022.111000

Chen, N., & Majda, A. J. (2018). Conditional Gaussian systems for multiscale nonlinear stochastic systems: Prediction, state estimation and uncertainty quantification. *Entropy*, *20*(7), 509. https://doi.org/10.3390/e20070509

Copernicus Climate Change Service (C35). (2017). ERA 5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. *Copernicus Climate Change Service Climate Data Store (CDS)*, *15*(2), 2020.

Cundall, P. A. (1988). Formulation of a three-dimensional distinct element model—Part I. A scheme to detect and represent contacts in a system composed of many polyhedral blocks. In *International journal of rock mechanics and mining sciences & geomechanics abstracts* (Vol. 25, pp. 107–116). Elsevier.

Cundall, P. A., & Strack, O. D. (1979). A discrete numerical model for granular assemblies. *Géotechnique*, *29*(1), 47–65. https://doi.org/10.1680/geot.1979.29.1.47

Edeling, W., & Crommelin, D. (2020). Reducing data-driven dynamical subgrid scale models by physical constraints. *Computers & Fluids*, *201*, 104470. https://doi.org/10.1016/j.compfluid.2020.104470

Evensen, G. (2009). *Data assimilation: The ensemble Kalman filter*. Springer Science & Business Media.

Farrell, B. F., & Ioannou, P. J. (1993). Stochastic forcing of the linearized Navier–Stokes equations. *Physics of Fluids A: Fluid Dynamics*, *5*(11), 2600–2609. https://doi.org/10.1063/1.858894

Gabrielski, A., Badin, G., & Kaleschke, L. (2015). Anomalous dispersion of sea ice in the Fram Strait region. *Journal of Geophysical Research: Oceans*, *120*(3), 1809–1824. https://doi.org/10.1002/2014JC010359

Gardiner, C. (2009). *Stochastic methods* (Vol. 4). Springer.

Gershgorin, B., Harlim, J., & Majda, A. J. (2010). Improving filtering and prediction of spatially extended turbulent systems with model errors through stochastic parameter estimation. *Journal of Computational Physics*, *229*(1), 32–57. https://doi.org/10.1016/j.jcp.2009.09.022

Grooms, I., & Majda, A. J. (2013). Efficient stochastic superparameterization for geophysical turbulence. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(12), 4464–4469. https://doi.org/10.1073/pnas.1302548110

Hart, R., Cundall, P. A., & Lemos, J. (1988). Formulation of a three-dimensional distinct element model—Part II. Mechanical calculations for motion and interaction of a system composed of many polyhedral blocks. In *International journal of rock mechanics and mining sciences & geomechanics abstracts* (Vol. 25, pp. 117–125). Elsevier.

Hibler, W. D. (1979). A dynamic thermodynamic sea ice model. *Journal of Physical Oceanography*, *9*(4), 815–846. https://doi.org/10.1175/1520-0485(1979)009<0815:adtsim>2.0.co;2

Hunke, E. C., & Dukowicz, J. K. (1997). An elastic–viscous–plastic model for sea ice dynamics. *Journal of Physical Oceanography*, *27*(9), 1849–1867. https://doi.org/10.1175/1520-0485(1997)027<1849:aevpmf>2.0.co;2

Hutchings, J. K., Heil, P., Steer, A., & Hibler, W. D. (2012). Subsynoptic scale spatial variability of sea ice deformation in the western Weddell Sea during early summer. *Journal of Geophysical Research*, *117*(C1), C01002. https://doi.org/10.1029/2011JC006961

Itkin, P., Spreen, G., Cheng, B., Doble, M., Girard-Ardhuin, F., Haapala, J., et al. (2017). Thin ice and storms: Sea ice deformation from buoy arrays deployed during N-ICE2015: Thin ice and storms. *Journal of Geophysical Research: Oceans*, *122*(6), 4661–4674. https://doi.org/10.1002/2016JC012403

Khodkar, M. A., & Hassanzadeh, P. (2018). Data-driven reduced modelling of turbulent Rayleigh–Bénard convection using DMD-enhanced fluctuation–dissipation theorem. *Journal of Fluid Mechanics*, 852. https://doi.org/10.1017/jfm.2018.586

Kondrashov, D., Chekroun, M. D., & Ghil, M. (2015). Data-driven non-Markovian closure models. *Physica D: Nonlinear Phenomena*, *297*, 33–55. https://doi.org/10.1016/j.physd.2014.12.005

Kondrashov, D., Chekroun, M. D., Yuan, X., & Ghil, M. (2018). Data-adaptive harmonic decomposition and stochastic modeling of Arctic sea ice. In *Advances in nonlinear geosciences* (pp. 179–205). Springer.

Kwok, R. (1999). Radarsat-1 data (csa). *Lagrangian Sea-Ice Kinematics*. Retrieved from https://asf.alaska.edu/data-sets/sar-data-sets/radarsat-1/
Kwok, R. (2018). Arctic sea ice thickness, volume, and multiyear ice coverage: Losses and coupled variability (1958–2018). *Environmental Research Letters*, *13*(10), 105005. https://doi.org/10.1088/1748-9326/aae3ec
Lei, R., Gui, D., Heil, P., Hutchings, J. K., & Ding, M. (2020). Comparisons of sea ice motion and deformation, and their responses to ice conditions and cyclonic activity in the western Arctic Ocean between two summers. *Cold Regions Science and Technology*, *170*(November 2018), 102925. https://doi.org/10.1016/j.coldregions.2019.102925
Leppäranta, M. (2011). *The drift of sea ice*. Springer Science & Business Media.
Leutbecher, M., & Palmer, T. N. (2008). Ensemble forecasting. *Journal of Computational Physics*, *227*(7), 3515–3539.
Lindsay, R., & Stern, H. (2004). A new Lagrangian model of Arctic sea ice. *Journal of Physical Oceanography*, *34*(1), 272–283. https://doi.org/10.1175/1520-0485(2004)034<0272:anlmoa>2.0.co;2
Lopez-Acosta, R., Schodlok, M., & Wilhelmus, M. (2019). Ice floe tracker: An algorithm to automatically retrieve Lagrangian trajectories via feature matching from moderate-resolution visual imagery. *Remote Sensing of Environment*, *234*, 111406. https://doi.org/10.1016/j.rse.2019.111406
Lopez-Acosta, R., & Wilhelmus, M. M. (2021). Library of sea ice floe remote sensing observations in the Beaufort Sea Marginal Ice Zone. *Zenodo*. https://doi.org/10.5281/zenodo.4796845
Majda, A. J. (2016). *Introduction to turbulent dynamical systems in complex systems*. Springer.
Majda, A. J., & Chen, N. (2018). Model error, information barriers, state estimation and prediction in complex multiscale systems. *Entropy*, *20*(9), 644. https://doi.org/10.3390/e20090644
Majda, A. J., & Qi, D. (2018). Strategies for reduced-order models for predicting the statistical responses and uncertainty quantification in complex turbulent dynamical systems. *SIAM Review*, *60*(3), 491–549. https://doi.org/10.1137/16m1104664
Manucharyan, G. E., Lopez-Acosta, R., & Wilhelmus, M. M. (2022). Spinning ice floes reveal intensification of mesoscale eddies in the western Arctic Ocean. *Scientific Reports*, *12*(1), 1–13. https://doi.org/10.1038/s41598-022-10712-z
Maslowski, W., Clement Kinney, J., Higgins, M., & Roberts, A. (2012). The future of Arctic sea ice. *Annual Review of Earth and Planetary Sciences*, *40*(1), 625–654. https://doi.org/10.1146/annurev-earth-042711-105345
Moon, W., & Wettlaufer, J. S. (2017). A stochastic dynamical model of Arctic sea ice. *Journal of Climate*, *30*(13), 5119–5140. https://doi.org/10.1175/jcli-d-16-0223.1
Olauson, J. (2018). ERA5: The new champion of wind power modelling? *Renewable Energy*, *126*, 322–331. https://doi.org/10.1016/j.renene.2018.03.056
Palmer, T. (2019). The ECMWF ensemble prediction system: Looking back (more than) 25 years and projecting forward 25 years. *Quarterly Journal of the Royal Meteorological Society*, *145*(S1), 12–24. https://doi.org/10.1002/qj.3383
Qi, D., & Majda, A. J. (2018). Predicting extreme events for passive scalar turbulence in two-layer baroclinic flows through reduced-order stochastic models. *Communications in Mathematical Sciences*, *16*(1), 17–51. https://doi.org/10.4310/cms.2018.v16.n1.a2
Squire, V. A. (2020). Ocean wave interactions with sea ice: A reappraisal. *Annual Review of Fluid Mechanics*, *52*(1), 37–60. https://doi.org/10.1146/annurev-fluid-010719-060301
Thomas, D. N. (2017). *Sea ice*. John Wiley & Sons.
Thomson, J., Ackley, S., Girard-Ardhuin, F., Ardhuin, F., Babanin, A., Boutin, G., et al. (2018). Overview of the Arctic sea state and boundary layer physics program. *Journal of Geophysical Research: Oceans*, *123*(12), 8674–8687. https://doi.org/10.1002/2018jc013766
Timmermans, M.-L., Toole, J., & Krishfield, R. (2018). Warming of the interior Arctic Ocean linked to sea ice losses at the basin margins. *Science Advances*, *4*(8), eaat6773. https://doi.org/10.1126/sciadv.aat6773
Toppaladoddi, S., & Wettlaufer, J. S. (2015). Theory of the sea ice thickness distribution. *Physical Review Letters*, *115*(14), 148501. https://doi.org/10.1103/physrevlett.115.148501
Toth, Z., & Kalnay, E. (1997). Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, *125*(12), 3297–3319. https://doi.org/10.1175/1520-0493(1997)125<3297:efanat>2.0.co;2
Tremblay, L. B., & Mysak, L. (1997). Modeling sea ice as a granular material, including the dilatancy effect. *Journal of Physical Oceanography*, *27*(11), 2342–2360. https://doi.org/10.1175/1520-0485(1997)027<2342:msiaag>2.0.co;2
Tuhkuri, J., & Polojärvi, A. (2018). A review of discrete element simulation of ice–structure interaction. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, *376*(2129), 20170335. https://doi.org/10.1098/rsta.2017.0335
Vallis, G. K. (2017). *Atmospheric and oceanic fluid dynamics*. Cambridge University Press.
Weeks, W. F., & Ackley, S. F. (1986). The growth, structure, and properties of sea ice. In *The geophysics of sea ice* (pp. 9–164). Springer.
Weiss, J. (2013). *Drift, deformation, and fracture of sea ice: A perspective across scales* (Vol. 83). Springer.
Wolfe, R. E., Roy, D. P., & Vermote, E. (1998). MODIS land data storage, gridding, and compositing methodology: Level 2 grid. *IEEE Transactions on Geoscience and Remote Sensing*, *36*(4), 1324–1338. https://doi.org/10.1109/36.701082